COMMUNICATIONS IN
ANALYSIS AND GEOMETRY
Volume 6, Number 2, 199-253, 1998

Gradient flows on nonpositively curved metric spaces and harmonic maps

UWE F. MAYER

The notion of gradient flows is generalized to a metric space setting without any linear structure. The metric spaces considered are a generalization of Hilbert spaces, and the properties of such metric spaces are used to set up a finite-difference scheme of variational form. The proof of the Crandall–Liggett generation theorem is adapted to show convergence. The resulting flow generates a strongly continuous semigroup of Lipschitz-continuous mappings, is Lipschitz continuous in time for positive time, and decreases the energy functional along a path of steepest descent. In case the underlying metric space is a Hilbert space, the solutions resulting from this new theory coincide with those obtained by classical methods. As an application, the harmonic map flow problem for maps from a manifold into a nonpositively curved metric space is considered, and the existence of a solution to the initial boundary value problem is established.

1. Existence theory.

1.1. Introduction.

In the past, one assumed an inner-product structure to make sense of the term *gradient*; one worked on a Hilbert space, or on the tangent space to a manifold, for example. However, it is possible to do without an inner product. The domain of the energy functionals considered herein is assumed to be a nonpositively curved metric space (L, D). Let $G : L \to \mathbf{R} \cup \{+\infty\}$ be the energy functional under consideration. One has to make sense of the equation

$$\frac{du(t)}{dt} = -\nabla G(u(t)) .$$

AMS subject classifications: 58F25, 58G11, 34G20

The time derivative is replaced by a finite difference

$$\frac{u(t+h)-u(t)}{h} = -\nabla G(u(t+h)),$$

which in the variational formulation formally translates into a penalty term:

$$u(t+h) \quad \text{minimizes} \quad u \mapsto G(u) + \frac{1}{2h}D^2(u, u(t)).$$

It will be shown that under suitable assumptions on G a unique minimizer $u(t+h)$ exists. For fixed $h > 0$ let $J_h : L \to L$ be the map which assigns the corresponding minimizer of the above time-step energy functional, formally

$$J_h = (I + h\nabla G)^{-1}.$$

Here I stands for the identity map. Applying J_h to an element u_0 of L corresponds to a discrete time step of width h along the gradient flow of G starting at u_0. Therefore $J^n_{t/n}(u_0)$ should be a reasonable approximation of the solution to the gradient flow at time t, since it corresponds to taking n steps of width t/n. Formally, one has the following equality:

$$J^n_{t/n} = \left(I + \frac{t}{n}\nabla G\right)^{-n}.$$

In the theory of semigroups of operators one considers equations of the form

$$\frac{du(t)}{dt} + A(u(t)) = 0.$$

Setting up an implicit finite-difference scheme leads to the consideration of

$$\left(I + \frac{t}{n}A\right)^{-n}. \tag{1.1}$$

The Crandall–Liggett generation theorem [2] concerns the convergence of (1.1) for a nonlinear operator A defined on a Banach space. It is possible to adapt the proof of this theorem to the current situation. Of course, as ∇G, which plays the role of A, needs not to exist, statements involving A have to replaced by equivalent statements about J_h. In the sequel it will be shown that the maps J_h are uniformly Lipschitz, and that the resolvent identity holds. These results together with a simple a priori estimate allow to show that $\{J^n_{t/n}(u_0)\}$ forms a Cauchy sequence in L. The outline of the proof follows closely the original proof of the Crandall–Liggett generation theorem in [2].

After completing the mathematical portion of this paper the author discovered that certain of the lemmas have been derived independently (and several months earlier) by J. Jost [7]. In particular, Jost proves the well-definedness and nonexpansiveness of the maps J_h, as well as the resolvent identity (Theorem 1.8, Lemma 1.12, Lemma 1.10) for the important special case of a convex energy functional. Jost's aim in [7] is to find generalized harmonic maps; he does not address the parabolic problem (gradient flows), which is studied herein.

According to M. Crandall, about 25 years ago W. Helton remarked already that the Crandall–Liggett generation theorem could be extended to metric spaces. However, W. Helton did not publish his findings, and the author of this paper does not know any further details.

The Crandall–Liggett generation theorem generalizes the Hille–Yosida generation theorem for linear operators to nonlinear operators on a Banach space. This new theory in turn is about nonlinear domain spaces. The main assumptions made will be assumptions on the convexity of both the functional G and the underlying space L. The results described herein indicate that it is natural to look at the functional itself rather than at the gradient of the functional. Although stronger assumptions like a Hilbert space setting allow one to come to stronger conclusions, they are not really necessary for a satisfactory theory. Of course, the term gradient flow has to be interpreted in a wider sense, perhaps as flow along the most rapid decrease of the given functional.

The introduction to Section 2 gives a description of the properties of the generated flow.

1.2. NPC spaces.

1.2.1. Introduction. Most of what appears in this section is included in [9].

A complete metric space (L, D) is called a nonpositively curved (NPC) space if it satisfies the following two conditions.

(a) For any two elements u, v of L there is a rectifiable curve from u to v with length $D(u, v)$. In other words, L is assumed to be a length space. Distance realizing curves are called geodesics.

(b) For any three points v, u_0, u_1 and choices of connecting geodesics γ_{v,u_0}, γ_{u_0,u_1}, $\gamma_{u_1,v}$ the following comparison principle holds. Let u_t be the point on γ_{u_0,u_1} which is a fraction t of the distance from u_0 to u_1,

$$D(u_0,u_t)=tD(u_0,u_1)\,,\qquad D(u_t,u_1)=(1-t)D(u_0,u_1). \tag{1.2}$$

The NPC hypothesis is the following inequality for $0 \le t \le 1$:

$$D^2(v,u_t)\le(1-t)D^2(v,u_0)+tD^2(v,u_1)-t(1-t)D^2(u_0,u_1). \tag{1.3}$$

Geometrically this inequality arises by considering a triangle in $\mathbf{R}^2$ with the same side lengths as the triangle in L from above. Denote the vertex corresponding to v by V, similarly the point U_t corresponds to u_t. The NPC hypothesis is that the distance $D(v,u_t)$ be bounded above by the Euclidean distance $|V-U_t|$. A computation of the Euclidean distance yields (1.3).

Examples of NPC spaces are Hilbert spaces; trees; Euclidean buildings; and complete, simply connected Riemannian manifolds with nonpositive sectional curvature. Furthermore, if X is an NPC space and if (M,g) is a Riemannian manifold, then the space $L^2(M,X)$ is also an NPC space.

NPC spaces have the following properties [9, 12].

Theorem 1.1. *Let (L,D) be an NPC space.*
(a) *For any two points u, v in L there is a unique connecting geodesic.*
(b) *Let v_0, v_1, u_0, and u_1 be four points in L and define u_t to be the point which is a fraction t of the distance from u_0 to u_1. For any $0\le t\le 1$ the following quadrilateral comparison holds:*

$$\begin{aligned} D^2(u_t,v_0)+D^2(u_{1-t},v_1) \le\ & D^2(u_0,v_0)+D^2(u_1,v_1)+2t^2D^2(u_0,u_1)\\ & +t(D^2(v_0,v_1)-D^2(u_0,u_1))\\ & -t(D(v_0,v_1)-D(u_0,u_1))^2\,. \end{aligned} \tag{1.4}$$

A consequence of (a) is that given two points u_0, u_1 there is a unique point u_t which satisfies (1.2). This fact has been used in the formulation of statement (b) of the theorem. One often writes $u_t=(1-t)u_0+tu_1$; this also enables one to introduce the concept of convexity to the NPC space setting. A function $G: L \to \mathbf{R}\cup\{+\infty\}$ is said to be convex if for all $t\in[0,1]$ one has $G(u_t)\le(1-t)G(u_0)+tG(u_1)$. The function G is said to be strictly convex if this inequality is strict for $t\in(0,1)$.

Inequality (1.3) implies that balls in L are uniformly convex. This fact has been used in [10] to generalize a well-known fact from the theory of uniformly convex Banach spaces to the NPC setting. The proof is essentially the same as in the Banach space version, see for example [5], and the idea also appears implicitly in Lemma 2 of [6].

Theorem 1.2. *Let S_k be a descending sequence of nonempty bounded convex closed sets in an NPC space L. Then the intersection $\bigcap S_k$ is nonempty.*

1.2.2. Convex functionals on NPC spaces. The aim of this section is to provide a few lemmas that will be needed in the sequel.

Lemma 1.3. *Let (L, D) be an NPC space and $G : L \rightarrow \mathbf{R} \cup \{+\infty\}$. If G is convex and lower semicontinuous then G is bounded from below on bounded subsets of L. Furthermore, G attains its infimum on nonempty bounded convex closed subsets of L. The resulting minimizer is unique if G is strictly convex.*

Proof. It is clearly enough to consider the case of a nonempty bounded convex closed set $S \subset L$ because any bounded set is contained in a closed geodesic ball of sufficiently large radius. Define

$$m = \inf_{u \in S} G(u) .$$

Assume $m < \infty$ and define for $k \in \mathrm{N}$

$$\begin{aligned} m(k) &= \begin{cases} m + 1/k & \text{for } m \neq -\infty , \\ -k & \text{for } m = -\infty , \end{cases} \\ S_k &= \{u \in S : G(u) \leq m(k)\} . \end{aligned}$$

By construction $\{S_k\}$ is a decreasing sequence of bounded sets. Each set S_k is closed because S is closed and G is lower semicontinuous. Each set S_k is convex because both S and G are convex. Theorem 1.2 is therefore applicable, and $\bigcap S_k \neq \{\}$. Let u be an element of this intersection. The lower semicontinuity of G implies $G(u) \leq m$. This is clearly impossible if $m = -\infty$. Hence m must be finite, and u is a minimizer for G restricted to S. As any strictly convex function attains its minimum at most once, this minimizer is unique if G is strictly convex. □

Corollary 1.4. *Let L be an NPC space and $G : L \rightarrow \mathbf{R} \cup \{+\infty\}$ be a convex lower semicontinuous functional. If G has a bounded minimizing sequence then G attains its infimum.*

Proof. Let $\{u_n\}_{n=1}^{\infty}$ be a minimizing sequence for G with $D(u_0, u_n) \leq C$ for $n \in \mathrm{N}$ and some $C > 0$. By the lemma G attains its minimum on the set $S = \{u \in L : D(u, u_0) \leq C\}$, and the minimum on S is the same as the infimum on all of L. □

The lemma enables one to make the following definition.

Definition 1. Let L be an NPC space and $G : L \to \mathbf{R} \cup \{+\infty\}$ be a convex lower semicontinuous functional. Denote the closed geodesic ball of radius r about $w \in L$ by $B(w,r)$. The minimal function of G centered at some point $w \in L$ is defined via

$$\begin{aligned} m(G,w) : [0,\infty) &\to \mathbf{R} \cup \{+\infty\}, \\ r &\mapsto \inf_{u \in B(w,r)} G(u). \end{aligned}$$

It is clear from the definition that $m(G,w)$ is a decreasing function, and provided $G \not\equiv \infty$, that there is a radius $R \geq 0$ such that one has $m(G,w)(r) < \infty$ for $r \geq R$. In the latter case Lemma 1.3 guarantees the existence of a point $v(r) \in L$ with $m(G,w)(r) = G(v(r))$. This point needs not to be unique if G is not strictly convex. Let r_1, r_2 be two radii for which $m(G,w)$ is finite, and for $t \in (0,1)$ let $v = tv(r_1) + (1-t)v(r_2)$ for some choice of $v(r_1)$ and $v(r_2)$. By the convexity of the distance function for NPC spaces one has $D(v,w) \leq tr_1 + (1-t)r_2$, that is $v \in B(w, tr_1 + (1-t)r_2)$. The convexity of G implies $G(v) \leq tG(v(r_1)) + (1-t)G(v(r_2)) = t\,m(G,w)(r_1) + (1-t)\,m(G,w)(r_2)$. As $m(G,w)(tr_1 + (1-t)r_2) \leq G(v)$ this shows that $m(G,w)$ is convex. The convexity implies continuity for $r > 0$. Continuity at $r = 0$ follows from the lower semicontinuity of G. This proves the following lemma.

Lemma 1.5. *Let L be an NPC space and $G : L \to \mathbf{R} \cup \{+\infty\}$ be a convex lower semicontinuous functional. The minimal function of G as defined in Definition 1 is convex and continuous.*

While the first lemma above showed that a convex lower semicontinuous functional G is bounded from below on bounded sets of L, it is still to be seen how badly G can go to negative infinity as its argument becomes unbounded. The lemma below states that the behavior is in some sense at most linear.

Lemma 1.6. *Let L be an NPC space and let $G : L \to \mathbf{R} \cup \{+\infty\}$ be a convex lower semicontinuous functional. Then there is a constant $C \geq 0$ depending only on G such that for any $w \in L$*

$$\liminf_{D(u,w)\to\infty} \frac{G(u)}{D(u,w)} \geq -C.$$

Proof. If $G \equiv \infty$ then there is nothing to prove. Assume therefore that $G \not\equiv \infty$. Pick some $w_0 \in L$ with $G(w_0) < \infty$. For any convex function

$f : [0, \infty) \to \mathbf{R}$ and $t \geq 1$ one has the following inequality

$$f(t) \geq f(0) + t(f(1) - f(0)) .$$

Apply this to the function $m(G, w_0)$. For any $u \in L$

$$\begin{aligned} G(u) &\geq m(G, w_0)(D(u, w_0)) \\ &\geq m(G, w_0)(0) + D(u, w_0)(m(G, w_0)(1) - m(G, w_0)(0)) . \end{aligned}$$

This inequality together with $D(u, w_0) \geq D(u, w) - D(w, w_0)$ yields the statement of the lemma. □

Definition 2. Assume L is an NPC space and and $G : L \to \mathbf{R} \cup \{+\infty\}$. The functional G is said to be uniformly convex if there is some $\epsilon > 0$ such that for $t \in [0, 1]$ and all $u_0, u_1 \in L$:

$$G(u_t) \leq (1 - t)G(u_0) + tG(u_1) - \epsilon t(1 - t)D^2(u_0, u_1) .$$

Remark. If there is an $\epsilon > 0$ and a $w_0 \in L$ such that $u \mapsto G(u) - \epsilon D^2(u, w_0)$ is a convex functional, then G is uniformly convex. If L is a Hilbert space or if L is an NPC space of finite diameter with curvature bounded from below then this condition is in fact equivalent to uniform convexity. In the smooth case the inequality in the definition corresponds to a positive lower bound on the second derivative.

Lemma 1.7. *Let L be an NPC space and $G : L \to \mathbf{R} \cup \{+\infty\}$, $G \not\equiv \infty$. If G is uniformly convex and lower semicontinuous then G is bounded from below; furthermore, every minimizing sequence converges to the unique minimizer of G.*

Proof. Assume $\inf G = -\infty$, then by Corollary 1.4 every minimizing sequence $\{u_k\}$ becomes unbounded. Pick some $w \in L$ with $G(w) < \infty$. For $v_k = \frac{1}{2}w + \frac{1}{2}u_k$ the definition of uniform convexity implies

$$\lim_{k \to \infty} \frac{G(v_k)}{D(w, v_k)} = -\infty ,$$

which contradicts Lemma 1.6. Hence G must have been bounded from below.

The following idea is well known. Again, let $\{u_k\}$ be a minimizing sequence. Then

$$G(\frac{1}{2}u_i + \frac{1}{2}u_j) \leq \frac{1}{2}G(u_i) + \frac{1}{2}G(u_j) - \epsilon \frac{1}{4}D^2(u_i, u_j)$$

implies

$$\lim_{i,j\to\infty} D^2(u_i,u_j)=0\,.$$

Hence every minimizing sequence is a Cauchy sequence, and this implies the lemma. □

1.3. A finite-difference scheme.

Let (L,D) be an NPC space and $G: L\to \mathbf{R}\cup\{+\infty\}$, with $G\not\equiv\infty$. The functional G will be fixed for the rest of the section. As L has no inner-product structure an equation of the form

$$\frac{du(t)}{dt}=-\nabla G(u(t))$$

is therefore to be interpreted in a wider sense. As outlined in the introduction to this part of the paper, formal discretization in time suggests a variational problem.

Definition 3. For fixed $h>0$ and $u_0\in L$ define the time-step energy functional via

$$E(u;u_0,h)=G(u)+\frac{1}{2h}D^2(u,u_0)\,, \tag{1.5}$$

and the time-step map via

$$J_h: L\to L\,,\quad E(J_h(v);v,h)=\min_{u\in L}E(u;v,h)\,. \tag{1.6}$$

It now needs to be shown that J_h is a well-defined mapping. In other words, one has to show that the implicit time-step problem has a unique solution.

In the following theorem, as always, let u_t denote the unique element of L which is a fraction t from u_0 to u_1.

Theorem 1.8 (Solvability of the time-step problem). (See also [7].) *Let (L,D) be an NPC space and $G: L\to\mathbf{R}\cup\{+\infty\}$, $G\not\equiv\infty$. Assume*
(a) *G is lower semicontinuous,*
(b) *$\exists S>0$ such that $G(u_t)\le(1-t)G(u_0)+tG(u_1)+St(1-t)D^2(u_0,u_1)$ for $t\in[0,1]$ and all $u_0,u_1\in L$.*
Then for given $v\in L$ and $0<h<\frac{1}{2S}$ there is a unique minimizer for $E(.\,;v,h)$.

Proof. Fix $h < \frac{1}{2S}$ and pick some $\epsilon \geq 0$ with $h < \frac{1}{2(S+\epsilon)}$. Define

$$E_\epsilon(u_t; v, h) = G(u_t) + \left(\frac{1}{2h} - \epsilon\right) D^2(u_t, v).$$

Take $u_0 \neq u_1 \in L$ and $t \in (0,1)$. Then by assumption (b) and the NPC hypothesis (1.3)

$$\begin{aligned} E_\epsilon(u_t; v, h) &\leq (1-t)G(u_0) + tG(u_1) + St(1-t)D^2(u_0, u_1) \\ &\quad + \Big(\frac{1}{2h} - \epsilon\Big)\Big((1-t)D^2(u_0, v) + tD^2(u_1, v) \\ &\qquad - t(1-t)D^2(u_0, u_1)\Big) \\ &< (1-t)E_\epsilon(u_0; v, h) + tE_\epsilon(u_1; v, h). \end{aligned}$$

This shows the strict convexity of the functional $E_\epsilon(.\,; v, h)$. This functional is lower semicontinuous by the assumption on G. Note that the same is true for $E(.\,; v, h)$ as this is the special case $\epsilon = 0$. Now fix some $\epsilon > 0$. Lemma 1.6 implies the existence of constants C_1 and C_2 with

$$E_\epsilon(u; v, h) \geq -C_1 - C_2 D(u, v).$$

Since $\epsilon > 0$ this shows

$$E(u; v, h) = E_\epsilon(u; v, h) + \epsilon D^2(u, v) \to \infty$$

as $D(u, v) \to \infty$. Pick a $\overline{u} \in L$ with $G(\overline{u}) < \infty$, then there is a radius R such that $E(u; v, h) > E(\overline{u}; v, h)$ for all $u \notin B(v, R)$. Here $B(v, R)$ denotes the closed geodesic ball of radius R about v. The statement of the theorem follows from an application of Lemma 1.3 with $S = B(v, R)$ and $G = E(.\,; v, h)$. ☐

Corollary 1.9. *Let L and G be as in the theorem above. Then for any $v \in L$*

$$\liminf_{D(u,v) \to \infty} \frac{G(u)}{D^2(u, v)} \geq -S.$$

Proof. The proof of the theorem showed

$$G(u) + \frac{1}{2h} D^2(u, v) \geq -C_1 - C_2 D(u, v) + \epsilon D^2(u, v)$$

for $\epsilon > 0$ sufficiently small and some constants C_1 and C_2. This implies

$$\liminf_{D(u,v) \to \infty} \frac{G(u)}{D^2(u, v)} \geq \epsilon - \frac{1}{2h}.$$

The statement of the corollary follows after one lets first $\epsilon \to 0$, and then $h \to 1/(2S)$. □

Remarks. (1) If G is convex the constant S can be chosen to be arbitrarily small and one has a unique minimizer for $E(.\,;v,h)$ for all $h > 0$.
(2) If assumption (b) is replaced by the stronger assumption of convexity for G, then the above proof still holds if one does not assume L to be NPC, but assumes only that $D^2(u_t, v)$ is a strictly convex function of t and that the conclusion of Theorem 1.2 holds. In particular the result holds for convex functionals on uniformly convex Banach spaces.
(3) Condition (b) implies that $u \mapsto G(u) + SD^2(u, w_0)$ is convex for any $w_0 \in L$. If L is a Hilbert space then these two conditions are equivalent. In the smooth case this corresponds to a lower bound for the second derivative of G.

Lemma 1.10 (Resolvent identity). (See also [7].) *Assume L and G satisfy the assumptions of Theorem 1.8. For $\mu > \lambda > 0$ and J_h as defined in (1.6) the following holds:*

$$J_\mu(v) = J_\lambda\left(\frac{\mu - \lambda}{\mu} J_\mu(v) + \frac{\lambda}{\mu} v\right) . \tag{1.7}$$

Here $\frac{\mu-\lambda}{\mu} J_\mu(v) + \frac{\lambda}{\mu} v$ denotes the element of L a fraction $\frac{\lambda}{\mu}$ along the geodesic from $J_\mu(v)$ to v.

Proof. Let $w = \frac{\mu-\lambda}{\mu} J_\mu(v) + \frac{\lambda}{\mu} v$. The following two inequalities follow directly from the minimizing properties of $J_\lambda(w)$ and $J_\mu(v)$:

$$\begin{aligned}(1.8)\quad G(J_\lambda(w)) + \frac{1}{2\lambda} D^2(J_\lambda(w), w) &\leq G(J_\mu(v)) + \frac{1}{2\lambda} D^2(J_\mu(v), w) ,\\ G(J_\mu(v)) + \frac{1}{2\mu} D^2(J_\mu(v), v) &\leq G(J_\lambda(w)) + \frac{1}{2\mu} D^2(J_\lambda(w), v) .\end{aligned}$$

Combining these two inequalities and using $\frac{\lambda}{\mu} + \frac{\mu-\lambda}{\mu} = 1$, together with $D(w, v) = \frac{\mu-\lambda}{\mu} D(J_\mu(v), v)$ and $D(w, J_\mu(v)) = \frac{\lambda}{\mu} D(J_\mu(v), v)$, yields

$$D^2(J_\lambda(w), v) \geq \frac{\mu}{\lambda} D^2(J_\lambda(w), w) + \frac{\mu}{\mu - \lambda} D^2(w, v) .$$

Applying Jensen's inequality to the triangle inequality results in the competing inequality

$$\begin{aligned} D^2(J_\lambda(w), v) &\leq (D(J_\lambda(w), w) + D(w, v))^2 \\ &\leq \frac{\mu}{\lambda} D^2(J_\lambda(w), w) + \frac{\mu}{\mu - \lambda} D^2(w, v) .\end{aligned}$$

Therefore the above inequalities are in fact all equalities. In particular equality in (1.8) implies $J_\lambda(w) = J_\mu(v)$ due to the uniqueness of the minimizer of the time-step energy. ☐

Remark. The above proof only uses the NPC property for the well-definedness of the time-step map J_h. If the existence of J_h can be guaranteed otherwise then the proof still goes through. In particular the resolvent identity holds for convex functionals on uniformly convex Banach spaces, cf. the remarks after Theorem 1.8.

Lemma 1.11. *Let $u_0 \in L$ be such that $G(u_0) < \infty$ and define recursively $u_{j+1} = J_h(u_j)$, where G is assumed to satisfy the conditions of Theorem 1.8. Fix some $w_0 \in L$ and let*

$$A = -\min\left\{0, \liminf_{D(u,w_0)\to\infty} \frac{G(u)}{D^2(u,w_0)}\right\}.$$

If $A \neq 0$ assume $jh \leq 1/(16A)$, otherwise assume $jh \leq T$ for some $T > 0$. Then

$$D^2(u_0,u_j) \leq Bjh. \tag{1.9}$$

Here B is a constant depending only on $G(u_0)$, $D(u_0,w_0)$, and A or T, respectively.

Remark. By the triangle inequality the constant A is independent of the choice of $w_0 \in L$; by Corollary 1.9 one has $A \leq S$.

Proof. As u_{l+1} minimizes $E(.\,;u_l,h)$ one has $E(u_{l+1};u_l,h) \leq E(u_l;u_l,h)$. By Definition 3 of the time-step energy functional this is equivalent to

$$\frac{1}{2h}D^2(u_{l+1},u_l) \leq G(u_l) - G(u_{l+1}). \tag{1.10}$$

The triangle inequality together with the Cauchy–Schwarz inequality implies

$$D^2(u_0,u_j) \leq \Big(\sum_{l=0}^{j-1} D(u_l,u_{l+1})\Big)^2 \leq j\sum_{l=0}^{j-1} D^2(u_l,u_{l+1}).$$

Combining these inequalities results in a telescoping sum and yields

$$D^2(u_0,u_j) \leq 2jh(G(u_0) - G(u_j)). \tag{1.11}$$

By definition of A there is a $K \geq 0$ such that $G(u) > -K - CD^2(u,w_0)$, where $C = 2A$ if $A \neq 0$, and $C = 1/(8T)$ otherwise. This together with the squared triangle inequality implies

$$D^2(u_0,u_j) \leq 2jh(G(u_0) + K + 2C(D^2(u_j,u_0) + D^2(u_0,w_0))),$$

and as by the assumptions $4Cjh \leq 1/2$, one obtains

$$D^2(u_0, u_j) \leq 4jh(G(u_0) + K + 2CD^2(u_0, w_0)),$$

and the proof is complete. $\square$

Remarks. If G is bounded from below on all of L then (1.11) shows that one can choose $B = 2(G(u_0) - \inf_{u \in L} G(u))$ and the estimate holds for all j and h. The estimate also holds for a convex functional on a uniformly convex Banach space.

Recall that for a map F from a metric space into itself the Lipschitz norm is defined as follows:

$$\|F\|_{\mathrm{Lip}} = \sup_{u \neq v} \frac{D(F(u), F(v))}{D(u, v)}.$$

Lemma 1.12. (See also [7].) *Assume L and G are as in Theorem* 1.8. *Then for given h with $0 < h < \frac{1}{4S}$ the map J_h as defined in* (1.6) *is Lipschitz continuous with*

$$\|J_h\|_{\mathrm{Lip}} \leq \frac{1}{\sqrt{1 - 4Sh}}.$$

Proof. Take two points v_0, v_1 and let $u_i = J_h(v_i)$, $i = 1, 2$, and $u_t = (1-t)u_0 + tu_1$. Using assumption (b) from Theorem 1.8 for u_t and u_{1-t} yields

$$G(u_t) + G(u_{1-t}) \leq G(u_0) + G(u_1) + 2St(1-t)D^2(u_0, u_1).$$

The quadrilateral comparison (1.4) implies

$$\begin{aligned}\frac{1}{2h}D^2(u_t, v_0) + \frac{1}{2h}D^2(u_{1-t}, v_1) \quad \leq \quad & \frac{1}{2h}D^2(u_0, v_0) + \frac{1}{2h}D^2(u_1, v_1) \\ & + \frac{t^2}{h}D^2(u_0, u_1) \\ & + \frac{t}{2h}(D^2(v_0, v_1) - D^2(u_0, u_1)) \\ & - \frac{t}{2h}(D(v_0, v_1) - D(u_0, u_1))^2.\end{aligned}$$

Adding these two inequalities together and regrouping the terms results in

$$\begin{aligned}E(u_t; v_0, h) + E(u_{1-t}; v_1, h) \quad \leq \quad & E(u_0; v_0, h) + E(u_1; v_1, h) \\ & - \frac{t}{2h}\big((1 - 4Sh)D^2(u_0, u_1) - D^2(v_0, v_1) \\ & \qquad + (D(u_0, u_1) - D(v_0, v_1))^2\big) \\ & + \frac{t^2}{h}(1 - 2Sh)D^2(u_0, u_1).\end{aligned}$$

The minimality of $E(u_i; v_i, h)$ for $i = 1, 2$ implies

$$E(u_0; v_0, h) + E(u_1; v_1, h) \leq E(u_t; v_0, h) + E(u_{1-t}; v_1, h)\,,$$

and therefore

$$\begin{aligned} 0 \;\leq\; & -\Big((1-4Sh)D^2(u_0,u_1) - D^2(v_0,v_1) + (D(u_0,u_1) - D(v_0,v_1))^2\Big) \\ & +2t(1-2Sh)D^2(u_0,u_1)\,. \end{aligned}$$

Drop the squared difference term and let $t \to 0$. This completes the proof of the lemma. □

Remarks. This lemma does not hold in a uniformly convex Banach space, even if the functional G is assumed to be convex. The reason is that the nearest-point projection onto a line can be distance increasing, even in a two-dimensional, uniformly convex, smooth Banach space. As an example, take the space $L = l^3(\mathbf{R}^2)$ and the points $v_0 = (0,0)$ and $v_1 = (5,4)$, and project them onto the line going through the origin and the point $(2,1)$. A simple computation shows that the projections u_0 and u_1 are further apart than v_0 and v_1. Let $e_1 = v_1 - u_1$ and $e_2 = (2,1)$, then $\{e1, e2\}$ constitutes a basis of L. Define a convex continuous functional G via $G(x_1e_1 + x_2e_2) = \max\{x_1, 0\}$. It is clear that $J_h(v_0) = u_0$ for any $h > 0$ because $v_0 = u_0 = (0,0)$, and that $J_h(v_1) = u_1$ if h is sufficiently large (a computation shows $h \geq 5$ suffices). Hence J_h is distance increasing.

Theorem 1.13 (Existence). *Let (L, D) be an NPC space and u_0 be a point in L. For the functional $G : L \to \mathbf{R} \cup \{+\infty\}$ assume the following properties:*
(a) *G is lower semicontinuous,*
(b) *$\exists S > 0$ such that $G(v_t) \leq (1-t)G(v_0) + tG(v_1) + St(1-t)D^2(v_0, v_1)$ for $t \in [0,1]$ and all $v_0, v_1 \in L$,*
(c) *$G(u_0) < \infty$.*
Fix any $v \in L$ and let

$$\begin{aligned} A &= -\min\left\{0, \liminf_{D(u,v)\to\infty} \frac{G(u)}{D^2(u,v)}\right\}, \\ I_A &= \begin{cases} (0,\infty) & \text{for } A = 0\,, \\ (0, \frac{1}{16A}] & \text{for } A > 0\,. \end{cases} \end{aligned}$$

Let J_h be the time-step map as defined in (1.6). Then there is a function $u : I_A \to L$ with

$$u(t) = \lim_{n\to\infty} J^n_{t/n}(u_0)\,, \tag{1.12}$$

and

$$G(u(t)) \leq G(u_0)\,. \tag{1.13}$$

The convergence is uniform for $0 < t \leq T$ *for any* $T \in I_A$. *Furthermore, the limit of* $u(t)$ *as* $t \to 0$ *exists and*

$$\lim_{t \to 0} u(t) = u_0\,. \tag{1.14}$$

Proof. This proof follows closely the original proof of the Crandall–Liggett generation theorem in [2]. For $n \geq m$ and $\mu \geq \lambda > 0$ let

$$a_{m,n} = D^2(J_\mu^m(u_0), J_\lambda^n(u_0))\,.$$

Take λ small enough so that $4S\lambda < 1$ and let $\alpha = \lambda/\mu$, $\beta = (\mu - \lambda)/\mu$, $\alpha_S = (1-4S\lambda)^{-1}\alpha$, and $\beta_S = (1-4S\lambda)^{-1}\beta$. The following chain of (in-) equalities uses the resolvent identity (Lemma 1.10), the uniform Lipschitz continuity of J_λ (Lemma 1.12), and the NPC hypothesis (1.3):

$$\begin{aligned}
a_{m,n} &= D^2(J_\lambda(\alpha J_\mu^{m-1}(u_0) + \beta J_\mu^m(u_0)), J_\lambda^n(u_0)) \\
&\leq (1-4S\lambda)^{-1} D^2(\alpha J_\mu^{m-1}(u_0) + \beta J_\mu^m(u_0), J_\lambda^{n-1}(u_0)) \\
&\leq (1-4S\lambda)^{-1}\Big(\alpha D^2(J_\mu^{m-1}(u_0), J_\lambda^{n-1}(u_0)) \\
&\qquad + \beta D^2(J_\mu^m(u_0), J_\lambda^{n-1}(u_0)) - \alpha\beta D^2(J_\mu^{m-1}(u_0), J_\mu^m(u_0))\Big) \\
&\leq \alpha_S a_{m-1,n-1} + \beta_S a_{m,n-1}\,. \qquad (1.15)
\end{aligned}$$

The following lemmas are quoted from [2].

Lemma 1.14. *For positive* α, β *with* $\alpha + \beta = 1$ *and for* $\alpha_S = (1-4S\lambda)^{-1}\alpha$, $\beta_S = (1-4S\lambda)^{-1}\beta$ *the recursion inequality* (1.15) *implies*

$$\begin{aligned}
a_{m,n} &\leq (1-4S\lambda)^{-n} \sum_{j=0}^{m-1} \alpha^j \beta^{n-j} \binom{n}{j} a_{m-j,0} \\
&\quad + \sum_{j=m}^{n} (1-4S\lambda)^{-j} \alpha^m \beta^{j-m} \binom{j-1}{m-1} a_{0,n-j}\,. \qquad (1.16)
\end{aligned}$$

Lemma 1.15. *Let* $m, n \in \mathbf{N}$, $n \geq m$, *and* α, β *positive with* $\alpha + \beta = 1$. *Then*

$$\begin{aligned}
\sum_{j=0}^{m-1} \alpha^j \beta^{n-j} \binom{n}{j} (m-j) &\leq \Big((n\alpha - m)^2 + n\alpha\beta\Big)^{1/2} \\
\sum_{j=m}^{n} \alpha^m \beta^{j-m} \binom{j-1}{m-1} (n-j) &\leq \Big(\frac{m\beta}{\alpha^2} + (\frac{m\beta}{\alpha} + m - n)^2\Big)^{1/2}.
\end{aligned}$$

Lemma 1.11 shows the existence of a constant B such that

$$\begin{aligned} a_{m-j,0} &\leq B(m-j)\mu, \\ a_{0,n-j} &\leq B(n-j)\lambda \end{aligned}$$

holds. These estimates in combination with Lemma 1.14 result in

$$\begin{aligned} a_{m,n} \leq\ & B(1-4S\lambda)^{-n}\sum_{j=0}^{m-1}\alpha^j\beta^{n-j}\binom{n}{j}(m-j)\mu \\ & +B\sum_{j=m}^{n}(1-4S\lambda)^{-j}\alpha^m\beta^{j-m}\binom{j-1}{m-1}(n-j)\lambda. \end{aligned} \tag{1.17}$$

Now apply Lemma 1.15 to this inequality,

$$\begin{aligned} a_{m,n} \leq\ & B(1-4S\lambda)^{-n}\Big\{\Big((n\alpha-m)^2+n\alpha\beta\Big)^{1/2}\mu \\ & +\Big(\frac{m\beta}{\alpha^2}+(\frac{m\beta}{\alpha}+m-n)^2\Big)^{1/2}\lambda\Big\}. \end{aligned}$$

Recall that $\alpha=\lambda/\mu$, and $\beta=(\mu-\lambda)/\mu$. Hence

$$\begin{aligned} a_{m,n} \leq\ & B(1-4S\lambda)^{-n}\Big\{\Big((n\lambda-m\mu)^2+n\lambda(\mu-\lambda)\Big)^{1/2} \\ & +\Big(m\mu(\mu-\lambda)+(m\mu-n\lambda)^2\Big)^{1/2}\Big\}. \end{aligned}$$

Note that

$$(1-x)^{-n}\leq e^{2nx} \tag{1.18}$$

for $0\leq x\leq 1/2$, see [2]. Thus for $4S\lambda\leq 1/2$ one obtains

$$\begin{aligned} a_{m,n} \leq\ & Be^{8Sn\lambda}\Big\{\Big((n\lambda-m\mu)^2+n\lambda(\mu-\lambda)\Big)^{1/2} \\ & +\Big(m\mu(\mu-\lambda)+(m\mu-n\lambda)^2\Big)^{1/2}\Big\}. \end{aligned}$$

Now set

$$\mu=t/m, \quad \lambda=t/n,$$

and recall that $a_{m,n}=D^2(J_\mu^m(u_0),J_\lambda^n(u_0))$. Then the above inequality simplifies considerably,

$$D^2(J_\mu^m(u_0),J_\lambda^n(u_0))\leq 2BTe^{8ST}\left(\frac{1}{m}-\frac{1}{n}\right)^{1/2},$$

where $n \geq m > 0$ and $0 < t \leq T$. Here $m\mu = t \leq T$ and $n\lambda = t \leq T$ so the application of Lemma 1.11 is justified by the choice of T. One also has $4S\lambda \leq 1/2$ if n is large enough so that (1.18) could be used.

This last inequality clearly implies that the sequence $\{J^n_{t/n}(u_0)\}$ is uniformly Cauchy. If one defines $u(t)$ to be the limit of this sequence then (1.12) holds. Dropping the distance term in (1.11) and setting $h = t/n$, $j = n$ results in

$$G(J^n_{t/n}(u_0)) \leq G(u_0)\,,$$

which together with the lower semicontinuity of G implies (1.13).

Lemma 1.11 also implies (again, set $h = t/n$, $j = n$)

$$D^2(u_0, J^n_{t/n}(u_0)) \leq Bt\,, \tag{1.19}$$

which in the limit as $n \to \infty$ converges to

$$D^2(u_0, u(t)) \leq Bt\,.$$

Therefore the limit of $u(t)$ as $t \to 0$ exists and (1.14) holds. The proof of Theorem 1.13 is complete. □

Remarks. **(1)** Corollary 1.9 shows $A \leq S$. If either G is bounded from below or convex then $A = 0$.

(2) The proof shows that in fact more is true. If $a(n) \in \mathbf{N}$ and $b(n) \in \mathbf{R}$ with $\lim\limits_{n\to\infty} a(n) = \infty$ and $\lim\limits_{n\to\infty} \dfrac{a(n)}{b(n)} = 1$ then

$$u(t) = \lim_{n\to\infty} J^{a(n)}_{t/b(n)}(u_0)\,.$$

In particular

$$u(t) = \lim_{h\to 0+} J^{[t/h]}_h(u_0)$$

where [.] denotes the greatest integer function; see also [2].

2. Properties of the flow.

2.1. Introduction.

The Crandall–Ligget generation theorem [2] concerns more than just the convergence of the finite-difference scheme. It also states that the limit generates a continuous semigroup of operators. This is also true in this general setting. More precisely, let the flow map $F_t(u)$ be defined as the flow at time t starting at $u \in L$. Then for fixed u the map $t \mapsto F_t(u)$ is $1/2$-Hölder continuous, and Lipschitz continuous for positive time. Also, for fixed t the map $u \mapsto F_t(u)$ is Lipschitz continuous. In case the functional under consideration is convex, this last map is nonexpanding. This is well-known for the Hilbert-space case [1, 3]. Also, if the metric space is in fact a Hilbert space, then the finite-difference scheme presented herein is equivalent to the finite-difference scheme by Crandall and Liggett for the subdifferential of the functional.

The semigroup property, of course, is also intuitively clear: A gradient flow is supposed to be a curve along the steepest descent of an energy functional. It only makes sense, that if one starts at some point and goes for some time s, then stops, starts again, and goes for another time t, then one should end up at the same position as if one would have started at the same initial point and gone for a time $s+t$. The semigroup property can be used to continue the flow for all time, because the existence-time interval is independent of the initial point of the flow.

The existence theorem from Section 1 shows that the time-step solutions $J^n_{t/n}(u)$ converge to the flow $F_t(u)$ in the metric sense. It will be shown that this also implies the convergence of the functional G. In other words, $G(J^n_{t/n}(u))$ converges to $G(F_t(u))$.

The almost-convexity assumption on the functional G allows to assign a slope to each point of the graph. This function is denoted by $|\nabla_- G|$, and is lower semicontinuous. The minus sign in the notation is a reminder that the graph may have corners, and that the absolute value of the descending slope has been chosen. As the flow is Lipschitz for positive time, it should be differentiable (almost everywhere) in some sense. In fact, it is possible to show for all times t

$$\lim_{s \to 0+} \frac{D(F_{t+s}(u), F_t(u))}{s} = |\nabla_- G|(F_t(u)) .$$

In other words, the speed of the flow is equal to the slope of the graph, which generalizes the equation $|u'(t)| = |\nabla G(u(t))|$ from the smooth case. Stationary points are exactly the places where $|\nabla_- G|$ vanishes, and once the flow reaches a stationary point, it stays there for all later times. Furthermore, the flow does proceed into the direction of the steepest descent of the functional: If the flow is not stationary one has

$$\lim_{s \to 0+} \frac{G(F_t(u)) - G(F_{t+s}(u))}{D(F_{t+s}(u), F_t(u))} = |\nabla_- G|(F_t(u)) .$$

These characterizations also allow to show that the function $G(F_t(u))$, which is Lipschitz continuous, is strictly decreasing as long as the flow is not stationary. They can also be used to show directly (without reference to the results by Crandall and Liggett in [2]) that if the underlying space is a Hilbert space, and if the flow is differentiable with respect to time, then the derivative of the flow is the element of smallest norm in the negative subdifferential of the functional, cf. [1] for the Hilbert-space case.

The behavior of the flow for large time is usually of interest if one wants to find stationary points for the functional, i.e. solve the elliptic problem. Although it is true that if $G(F_t(u))$ stays bounded from below then $|\nabla_- G|(F_t(u)) \to 0$ as $t \to \infty$, it is not true in general that the flow converges. However, if it does, then the limit is a stationary point. If one assumes a suitable compactness condition of the functional in form of a generalized Palais–Smale condition, then convergence can be asserted. Also, the mountain-pass theorem generalizes to this metric space setting.

Finally, for the important special case of a convex functional, one obtains stronger results. The flow crosses level sets perpendicularly (in a suitable sense). The only stationary points are the minimizers of the functional, and there is a minimizer if and only if the flow stays bounded. In any case, $G(F_t(u))$ converges to the infimum of G as $t \to \infty$. If the functional is additionally known to be uniformly convex, or if the flow stays bounded and the metric space is finitely compact, then the limit of the flow exists, and is a minimizer of the functional.

2.2. Continuity.

The following is a formal definition of the flow map.

Definition 4. Let I_A be as in Theorem 1.13. The flow map F is defined via

$$\begin{array}{rcl} F : \overline{I_A} \times \{u \in L : G(u) < \infty\} & \to & L \\ (t, u_0) & \mapsto & u(t) \end{array}$$

where $u(t)$ is defined in (1.12) for $t > 0$ and $u(0) = u_0$. F_t denotes the map

$$\{u \in L : G(u) < \infty\} \to L\,, \quad u \mapsto F(t,u)\,.$$

Let J_h be the time-step maps as defined in Section 1. Since one has $\left\| J^n_{t/n} \right\|_{\mathrm{Lip}} \le \|J_{t/n}\|^n_{\mathrm{Lip}} \le (1 - 4St/n)^{-n/2}$ and $\lim_{n\to\infty}(1 - 4St/n)^{-n/2} = e^{2St}$, one concludes $\|F_t\|_{\mathrm{Lip}} \le e^{2St}$ for the limit F_t of the maps $J^n_{t/n}$.

Theorem 2.1. *The maps* F_t *as defined in Definition* 4 *are Lipschitz continuous with*

$$\|F_t\|_{\mathrm{Lip}} \le e^{2St}.$$

Remark. If G is convex then this shows that the maps F_t are nonexpanding. This generalizes the result from the Hilbert-space setting [1, 3].

For the case of a convex functional G on a Hilbert space with $G(u) < \infty$ it is known that the map $t \mapsto F_t(u)$ is continuous [1, p57]. The following theorem generalizes this result to the current setting.

Theorem 2.2. *For given* $u_0 \in L$ *with* $G(u_0) < \infty$ *the map* $t \mapsto F_t(u_0)$ *is uniformly Hölder continuous with exponent* $\frac{1}{2}$ *on* $0 \le t \le T$ *for any* $T \in I_A$. *More precisely, there is a constant* B *such that for* $0 \le s \le t \le T$

$$D(F_s(u_0), F_t(u_0)) \le \sqrt{B} e^{2ST} (t-s)^{1/2}.$$

B is the constant from Lemma 1.11.

Proof. Set $\mu = t/n$ and $\lambda = s/n$ in (1.17) where $n = m$ and $0 < s < t \le T$. After simplification this results in

$$D^2(J^n_{t/n}(u_0), J^n_{s/n}(u_0)) \le B(1 - 4Ss/n)^{-n} \sum_{j=0}^{n} \binom{n}{j} \alpha^j \beta^{n-j} (n-j)\frac{t}{n},$$

where $\alpha = \lambda/\mu$ and $\beta = (\mu - \lambda)/\mu$. Note that $n\mu \le T$, $n\lambda \le T$, and $4S\lambda \le 1/2$ for $n \ge 8ST$, so that the application is justified. Now one needs the following two equalities for α, β with $\alpha + \beta = 1$, which are quoted from [2].

$$\sum_{j=0}^{n} \binom{n}{j} \alpha^j \beta^{n-j} = 1, \quad \sum_{j=0}^{n} \binom{n}{j} \alpha^j \beta^{n-j} j = \alpha n.$$

These two equalities together with $\alpha = \lambda/\mu = s/t$ and the inequality above yield

$$D^2(J^n_{t/n}(u_0), J^n_{s/n}(u_0)) \le B(1 - 4ST/n)^{-n}(t-s).$$

The case $s = 0$ follows from (1.19), if one defines J_0 to be the identity map. Taking the square root on both sides and letting $n \to \infty$ completes the proof. □

Corollary 2.3. *The flow map* F *is jointly continuous. That is, the map*

$$F : \overline{I_A} \times \{u \in L : G(u) < \infty\} \to L, \quad (t, u) \mapsto F(t, u)$$

is continuous.

Proof. This is a consequence of the uniform continuity in both variables.□

Remark. The exponent 1/2 in Theorem 2.2 is sharp in the following sense. For $\beta > \frac{1}{2}$ there are functionals G_β with corresponding flows $u_\beta(t)$ such that $t \mapsto u_\beta(t)$ is not β–Hölder continuous. For example, take L to be the closed interval $[0,1]$ and $G_\beta(u) = 1 - u^\alpha$ with $\alpha > 0$ and $\gamma = 1/(2-\alpha) < \beta$. Then $u(t) = (C_\alpha t)^\gamma$ for some constant C_α solves $\dfrac{du}{dt} = -G'_\beta(u)$, $u(0) = 0$, and is clearly not β–Hölder continuous. However, one can show that the flow is Lipschitz continuous for positive time. This result needs the semigroup property, and is therefore shown in a later section.

2.3. The semigroup property.

The aim of this section is to establish

$$F_{s+t} = F_s \circ F_t$$

for the flow map F. This is first established when the times s and t are the same.

Lemma 2.4. *For I_A as defined in Theorem 1.13 take $k \in \mathbf{N}$ with $kt \in I_A$, then for F_t as defined in Definition 4 the following holds:*

$$F_{kt} = (F_t)^k \ .$$

Proof. Only the case $k = 2$ is shown in detail. The proof for $k > 2$ proceeds analogously.

Fix $u_0 \in L$ with $G(u_0) < \infty$. The sequence $\{J^n_{t/n}(u_0)\}$ converges and is therefore bounded. Also $G(J^n_{t/n}(u_0)) \leq G(u_0)$ is true. Lemma 1.11 implies then the existence of a constant B depending only on $G(u_0)$, A, $T \in I_A$, and the bound on $D(J^n_{t/n}(u_0), w_0)$ for a fixed $w_0 \in L$, such that

$$D^2(J^j_{t/m}(J^n_{t/n}(u_0)), J^n_{t/n}(u_0)) \leq Bj\frac{t}{m}$$

for $0 < t \leq T$, $j \leq m$, and $n \in \mathbf{N}$. The proof of the existence of the flow (Theorem 1.13) shows then the uniform convergence for $n \in \mathbf{N}$ and $0 < t \leq T$ of

$$F_t(J^n_{t/n}(u_0)) = \lim_{m\to\infty} J^m_{t/m}(J^n_{t/n}(u_0)) \, .$$

This together with the continuity of F_t then implies

$$\begin{aligned} F_t(F_t(u_0)) &= \lim_{n\to\infty}\lim_{m\to\infty} J^m_{t/m}(J^n_{t/n}(u_0)) = \lim_{n\to\infty} J^n_{t/n}(J^n_{t/n}(u_0)) \\ &= \lim_{n\to\infty} J^{2n}_{2t/2n}(u_0) = F_{2t}(u_0) \, , \end{aligned}$$

which completes the proof of the lemma. □

Theorem 2.5. *The flow map F as defined in Definition* 4 *has the semigroup property, that is for $s, t \geq 0$ with $s + t \in I_A$:*

$$F_{s+t} = F_s \circ F_t .$$

Proof. The outline of this proof follows the presentation in [2]. Let k, l, r, $s \in \mathbf{N}$ and write $F(t)$ for F_t. The by the previous lemma

$$\begin{aligned} F\Big(\frac{k}{l}+\frac{r}{s}\Big) &= F\Big(\frac{ks+lr}{ls}\Big) = \Big(F\Big(\frac{1}{ls}\Big)\Big)^{ks+lr} \\ &= \Big(F\Big(\frac{1}{ls}\Big)\Big)^{ks} \circ \Big(F\Big(\frac{1}{ls}\Big)\Big)^{lr} = F\Big(\frac{k}{l}\Big) \circ F\Big(\frac{r}{s}\Big), \end{aligned}$$

hence $F(p+q) = F(p) \circ F(q)$ for p, q rational and positive. The desired statement follows now from the continuity of the flow map in the variable t. □

Corollary 2.6. *For given $u_0 \in L$ with $G(u_0) < \infty$ the map $t \mapsto G(F_t(u_0))$ is nonincreasing.*

Proof. Fix $s \geq 0$ and let $u(t)$ be the flow starting at $F_s(u_0)$. Inequality (1.13) then reads

$$G(u(t)) \leq G(F_s(u_0)) .$$

By the semigroup property $u(t) = F_t(F_s(u_0)) = F_{s+t}(u_0)$ and hence

$$G(F_{s+t}(u_0)) \leq G(F_s(u_0)) .$$

This is the desired statement, as $s, t \geq 0$ were arbitrary. □

Another consequence of the semigroup property is the existence of the flow for all time even if $A > 0$. Theorem 1.13 showed the existence of the flow up to $T_A = \frac{1}{16A}$. Using the semigroup property one sees that one can restart the flow, say, at $T_A/2$, then again at T_A, and so on. The flow is then defined inductively for all positive times. The properties of the exponential function show that the Lipschitz continuity estimate still holds.

Corollary 2.7. *In Theorem* 1.13 *assume $A > 0$ and define the flow for all positive times as outlined above. Then Theorems* 2.1 *and* 2.5 *still hold.*

Remark. This extension is tacitly understood to be used in all of the following. The notations $u(t)$ and $F_t(u_0)$ will be used synonymously.

2.4. An estimate.

The lemma in this section is essential for much of what follows. It will be used in connection with Lipschitz continuity, steepest descent, and the behavior of the flow for large times. The idea is to pull the time-step solution towards a point to get competitors for the minimization of the time-step energy.

Lemma 2.8. *Let L, G, S, and $u(t)$ be as in Theorem 1.13. Let u_1 be any point in L. Then for $t, s \geq 0$ with $s \in I_A$*

$$D^2(u(t+s), u_1) \leq e^{2Ss} D^2(u(t), u_1) - \frac{e^{2Ss}-1}{S}(G(u(t+s)) - G(u_1)).$$

Furthermore, if $G(u_1) \leq G(u(t+s))$ then

$$D^2(u(t+s), u_1) \leq e^{2Ss} D^2(u(t), u_1) - 2s(G(u(t+s)) - G(u_1)).$$

Proof. It is enough to consider the case $t = 0$ as the general statement then follows from the semigroup property. Assume $h < \frac{1}{2S}$ and let J_h denote the time-step map as in (1.6). Let u_λ be the point a fraction λ from $J_h(u_0)$ to u_1. From Definition 3 of the time-step energy, the properties of G, and the NPC hypothesis, one concludes

$$\begin{aligned}
E(J_h(u_0); u_0, h) &= G(J_h(u_0)) + \frac{1}{2h} D^2(J_h(u_0), u_0) \\
&\leq G(u_\lambda) + \frac{1}{2h} D^2(u_\lambda, u_0) \\
&\leq (1-\lambda) G(J_h(u_0)) + \lambda G(u_1) + S\lambda(1-\lambda) D^2(J_h(u_0), u_1) \\
&\quad + \frac{1}{2h}\Big((1-\lambda) D^2(J_h(u_0), u_0) + \lambda D^2(u_0, u_1) \\
&\qquad - \lambda(1-\lambda) D^2(J_h(u_0), u_1)\Big) \\
&= E(J_h(u_0); u_0, h) \\
&\quad + \lambda\Big(G(u_1) - G(J_h(u_0)) + \frac{1}{2h} D^2(u_0, u_1) \\
&\qquad - \frac{1}{2h} D^2(J_h(u_0), u_0) + (S - \frac{1}{2h}) D^2(J_h(u_0), u_1)\Big) \\
&\quad + \lambda^2 (\frac{1}{2h} - S) D^2(J_h(u_0), u_1). \qquad (2.1)
\end{aligned}$$

Subtract $E(J_h(u_0); u_0, h)$, divide by λ, and then let $\lambda \to 0$. This implies

$$D^2(J_h(u_0), u_1) \leq \frac{1}{1-2Sh}\Big(D^2(u_0, u_1) - 2h(G(J_h(u_0)) - G(u_1))\Big).$$

Now apply the above inequality n times, where $h = s/n$ and u_0 is replaced by $J^k_{s/n}(u_0)$, $k = 1, \ldots, n$. Together with $G(J^k_{s/n}(u_0)) \geq G(J^n_{s/n}(u_0))$ this implies

$$\begin{aligned} D^2(J^n_{s/n}(u_0), u_1) &\leq (1-2Ss/n)^{-n} D^2(u_0, u_1) \\ &\quad -2\frac{s}{n}\Big(\sum_{k=1}^{n}(1-2Ss/n)^{-k}\Big)(G(J^n_{s/n}(u_0)) - G(u_1)). \end{aligned}$$

An application of the sum formula for a finite geometric series yields

$$2\frac{s}{n}\sum_{k=0}^{n}(1-2Ss/n)^{-k} = \frac{(1-2Ss/n)^{-n} - 1}{S},$$

now use this in the inequality above and let $n \to \infty$. The lower semicontinuity of G implies

$$D^2(u(s), u_1) \leq e^{2Ss} D^2(u_0, u_1) - \frac{e^{2Ss} - 1}{S}(G(u(s)) - G(u_1)),$$

which of course used $u(s) = \lim_{n\to\infty} J^n_{s/n}(u_0)$. If $G(u(s)) > G(u_1)$ then one can use the estimate

$$2s \leq \frac{e^{2Ss} - 1}{S}$$

to get the second statement of the lemma. □

2.5. Lipschitz continuity.

In the section on continuity it has been shown that the flow is Hölder continuous. While this is the best possible result if one considers the flow for all time, it is not optimal if the flow is considered for positive time only. The flow becomes more regular instantaneously. This regularization property is well-known for the case of a convex functional on a Hilbert space, see [1, p57f].

Theorem 2.9. *Let L and G as in Theorem 1.13. Then the flow $t \mapsto F_t(u_0)$ is uniformly Lipschitz continuous on $[t_0, t_1]$ for any $t_1 > t_0 > 0$.*

Proof. Let $s \in I_A$. Apply Lemma 2.8 with $u_1 = u(t)$ and divide by s^2:

$$\frac{D^2(u(s+t),u(t))}{s^2} \leq 2\frac{G(u(t))-G(u(t+s))}{s}\cdot\frac{e^{2Ss}-1}{2Ss}.$$

Replacing t by $t-s$ in the application of the lemma results in

$$\frac{D^2(u(t),u(t-s))}{s^2} \leq 2\frac{G(u(t-s))-G(u(t))}{s}\cdot\frac{e^{2Ss}-1}{2Ss}.$$

The function $t \mapsto G(u(t))$ is real valued and nonincreasing, and therefore differentiable almost everywhere, see for example [13]. Because of $\lim\limits_{s\to 0}\frac{e^{2Ss}-1}{2Ss} = 1$ one obtains for almost all t

$$\limsup_{|s|\to 0}\frac{D(u(s+t),u(t))}{|s|} < \infty .$$

Let $0 < \bar{t} < t_0$ be such that the above limit is finite. Then there are constants C and $\epsilon > 0$ such that for $|s| < \epsilon$

$$D(u(\bar{t}+s),u(\bar{t})) \leq C|s| .$$

The semigroup property of the flow together with the Lipschitz continuity of $F_{t-\bar{t}}$ implies for any $t_1 \geq t \geq \bar{t}$ and $|s| < \epsilon$

$$\begin{aligned} D(u(t+s),u(t)) &= D(F_{t-\bar{t}}(u(\bar{t}+s)),F_{t-\bar{t}}(u(\bar{t}))) \\ &\leq e^{2St_1}D(u(\bar{t}+s),u(\bar{t})) \leq Ce^{2St_1}|s| . \end{aligned}$$

This shows uniform Lipschitz continuity. □

Corollary 2.10. *Let L, G, and $u(t)$ be as in the theorem above. Then the function $g : t \mapsto G(u(t))$ is continuous on $[0,\infty)$ and uniformly Lipschitz continuous on $[t_0,t_1]$ for any $t_1 > t_0 > 0$.*

Proof. Let C be the Lipschitz constant of $t \mapsto u(t)$ on $[t_0,t_1]$. Assume first $t_1 \in I_A$. Pick $t,\bar{t} \in [t_0,t_1]$ with $\bar{t} > t$, and set $u_1 = u(\bar{t})$ and $s = \bar{t}-t$ in the second statement of Lemma 2.8. This implies

$$(\bar{t}-t)(g(t)-g(\bar{t})) \leq \frac{1}{2}e^{2St_1}D^2(u(t),u(\bar{t})) .$$

Applying the Lipschitz continuity of u results in

$$g(t)-g(\bar{t}) \leq \frac{C^2e^{2St_1}}{2}(\bar{t}-t) . \tag{2.2}$$

If $t_1 \notin I_A$ one chooses a finite number of intermediate points, and, using the semigroup property, one obtains the analogous inequalities on the resulting subintervals. Combining these then yields (2.2) as well.

In either case this shows the uniform Lipschitz continuity on $[t_0, t_1]$ because g is nonincreasing.

Continuity at $t = 0$ follows from the continuity of u together with the lower semicontinuity of G, and the fact that g is nonincreasing. □

Remark. In the example from the remark in section 2.2 one has the solution $G(u(t)) = 1 - (C_\alpha t)^{\alpha/(1-2\alpha)}$ for some constant C_α. As $\alpha \to 0$ the continuity at $t = 0$ deteriorates.

2.6. Stationary points and steepest descent.

Definition 5. For $u_0 \in L$ with $G(u_0) < \infty$ define

$$|\nabla_- G|(u_0) = \max\left\{\limsup_{u\to u_0} \frac{G(u_0) - G(u)}{D(u_0, u)}, 0\right\},$$

if $G(u_0) = \infty$ set $|\nabla_- G|(u_0) = \infty$. The point u_0 is called a stationary point for the gradient flow of the functional G if $|\nabla_- G|(u_0) = 0$.

If L is a Hilbert space and ∇G exists at u_0, then $|\nabla G(u(_0))| = |\nabla_- G|(u_0)$, and u_0 is a stationary point in the language of the above definition if and only if $\nabla G(u_0) = 0$. Hence this definition generalizes the classical case. For technical reasons it is sometimes more convenient to use the following equivalent condition, which is a consequence of the assumptions on G.

Lemma 2.11. *Let G be as in Theorem* 1.13 *with $G(u_0) < \infty$. Then*

$$|\nabla_- G|(u_0) = 0 \iff \sup_{u\neq u_0} \frac{G(u_0) - G(u)}{D^2(u_0, u)} < \infty.$$

Proof. It is clear that the right-hand side implies the left-hand side. To show the other direction, proceed indirectly. In particular assume there is a $u_1 \in L$ with

$$\frac{G(u_0) - G(u_1)}{D^2(u_0, u_1)} \geq 2S.$$

Use $G(u_\lambda) \leq (1 - \lambda)G(u_0) + \lambda G(u_1) + S\lambda(1 - \lambda)D^2(u_0, u_1)$ with $u_\lambda = (1 - \lambda)u_0 + \lambda u_1$ to conclude

$$\frac{G(u_0) - G(u_\lambda)}{D(u_0, u_\lambda)} \geq S(1 + \lambda)D(u_0, u_1),$$

which implies $\limsup_{u\to u_0} \dfrac{G(u_0)-G(u)}{D(u_0,u)} \geq SD(u_0,u_1) > 0\,.$ □

Theorem 2.12. *Let L, G, and u_0 be as in Theorem 1.13. Then the flow starting at u_0 is stationary if and only if u_0 is a stationary point in the meaning of Definition 5.*

Proof. Assume first that u_0 is a stationary point. By the lemma above one has for some constant C and all $u \in L$

$$G(u) \geq G(u_0) - CD^2(u,u_0)\,.$$

This implies for the time step energy with $h < \frac{1}{2C}$ and for $u \neq u_0$

$$\begin{aligned} E(u;u_0,h) &= G(u) + \frac{1}{2h}D^2(u,u_0) \\ &\geq G(u_0) + (\frac{1}{2h} - C)D^2(u,u_0) \\ &> G(u_0) = E(u_0;u_0,h)\,. \end{aligned}$$

Hence by the definition of the time step map $J_h(u_0) = u_0$, and the flow does not move.

Assume now that u_0 is not a stationary point. Again by the lemma above, there is a $u_1 \in L$ with

$$\frac{G(u_0)-G(u_1)}{D^2(u_0,u_1)} > 2S\,.$$

By continuity $G(u(s)) - G(u_1) > 2SD^2(u_0,u_1)$ for $s > 0$ sufficiently small. Apply Lemma 2.8 with $t = 0$,

$$D^2(u(s),u_1) \leq e^{2Ss}D^2(u_0,u_1) - 2s(G(u(s)) - G(u_1))\,.$$

For small $s \geq 0$ one has $e^{2Ss} \leq 1 + 3Ss$ so that

$$D^2(u(s),u_1) \leq (1 - Ss)D^2(u_0,u_1)\,.$$

Thus the distance from $u(s)$ to u_1 is initially strictly decreasing and the flow is not stationary. □

Corollary 2.13. *Let L, G, and u_0 be as in the theorem above. If u_0 is not a stationary point then $G(u(s)) < G(u_0)$ for all $s > 0$.*

Proof. The theorem shows $D^2(u(s), u_0) > 0$ for small $s > 0$. An application of Lemma 2.8 with $u_1 = u_0$ and $t = 0$ yields $G(u(s)) - G(u_0) < 0$. □

Remark. The semigroup property then implies that $G(u(t)) < G(u(s))$ for $t > s$, provided $u(s)$ is not a stationary point. Hence the flow continues to strictly decrease G, until it possibly reaches a stationary point of G.

In the Hilbert-space case the gradient points in the direction of the steepest ascent of the functional under consideration. The gradient flow therefore moves along the steepest descent of the functional. This notion can be generalized.

Theorem 2.14. *Let L, G, and $u(t)$ be as in Theorem 1.13. Assume $u(t_0)$ is not a stationary point of G. Then*

$$\lim_{t\to t_0+} \frac{G(u(t_0)) - G(u(t))}{D(u(t), u(t_0))} = |\nabla_{-}G|(u(t_0)) . \tag{2.3}$$

The existence of the limit is part of the statement, and for $t_0 > 0$ this limit is finite.

The following lemma is used in the proof.

Lemma 2.15. *Let L, G, and $u(t)$ be as above. Assume*

$$\liminf_{t\to 0+} \frac{D(u(t_0 + t), u(t_0))}{t} = 0 .$$

Then $u(t_0 + t) = u(t_0)$ for all $t \geq 0$.

Proof. Let $s > t_0$ and use the semigroup property and the Lipschitz continuity (Theorem 2.1) to show

$$0 \leq \liminf_{t\to 0+} \frac{D(u(s + t), u(s))}{t} \leq \liminf_{t\to 0+} \frac{e^{2S(s-t_0)} D(u(t_0 + t), u(t_0))}{t} = 0 .$$

The function $d : t \mapsto D(u(t), u(t_0))$ is Lipschitz continuous for positive time and therefore differentiable almost everywhere. The triangle inequality implies

$$|D(u(s + t), u(t_0)) - D(u(s), u(t_0))| \leq D(u(s + t), u(s)) ,$$

hence $d' = 0$ almost everywhere by the argument above. The absolute continuity of d for $t > 0$ then shows that d is constant. □

Proof of the theorem. Let

$$C_1 = \liminf_{t\to t_0+} \frac{G(u(t_0)) - G(u(t))}{D(u(t), u(t_0))} .$$

If $C_1 = \infty$ then (2.3) holds trivially; hence assume $C_1 < \infty$. Set $\overline{u} = u(t_0)$ and $\overline{u}(t) = u(t_0 + t)$. This proof uses the time step maps J_h with $h = t/n$. Let $\epsilon_1 > 0$. Lemma 1.11 implies

$$D(\overline{u}, J_h^k(\overline{u})) \leq \epsilon_1 \tag{2.4}$$

for t sufficiently small and $k \leq n$. Fix any $u \in L$ with $G(u) < G(\overline{u})$. Such a u exists because $\overline{u}$ is not a stationary point of G. Pick ϵ_1 small so that

$$0 < \frac{2\epsilon_1}{D(\overline{u}, u) - \epsilon_1} < 1 .$$

Let

$$\lambda_k = \frac{D(J_h^k(\overline{u}), J_h^{k-1}(\overline{u}))}{D(u, J_h^{k-1}(\overline{u}))}$$

for $k = 1, \dots, n$, then $\lambda_k \in [0, 1)$. Let

$$u_k = (1 - \lambda_k) J_h^{k-1}(\overline{u}) + \lambda_k u ,$$

then $D(u_k, J_h^{k-1}(\overline{u})) = D(J_h^k(\overline{u}), J_h^{k-1}(\overline{u}))$, which by the minimizing property of $J_h^k(\overline{u})$ implies $G(J_h^k(\overline{u})) \leq G(u_k)$. Assumption (b) of Theorem 1.13 together with the definition of u_k implies

$$G(J_h^k(\overline{u})) \leq (1 - \lambda_k) G(J_h^{k-1}(\overline{u})) + \lambda_k G(u) + S\lambda_k (1 - \lambda_k) D^2(J_h^{k-1}(\overline{u}), u) .$$

By (2.4) one has $D(J_h^{k-1}(\overline{u}), u) \leq D(\overline{u}, u) + \epsilon_1$. Also, $1 - \lambda_k \leq 1$. Using these two estimates and the estimate for $G(J_h^k(\overline{u}))$ from above n times, one obtains

$$\begin{aligned} G(J_h^n(\overline{u})) \;\leq\; & \Big(\prod_{k=1}^n (1 - \lambda_k)\Big) G(\overline{u}) + \Big(1 - \prod_{k=1}^n (1 - \lambda_k)\Big) G(u) \\ & + S\Big(\sum_{k=1}^n \lambda_k\Big) (D(\overline{u}, u) + \epsilon_1)^2 . \end{aligned}$$

Furthermore,

$$\begin{aligned} \log\Big(\prod_{k=1}^n (1 - \lambda_k)\Big) \;&=\; \sum_{k=1}^n \log(1 - \lambda_k) \\ &\leq\; -\sum_{k=1}^n \lambda_k \\ &\leq\; -\sum_{k=1}^n \frac{D(J_h^k(\overline{u}), J_h^{k-1}(\overline{u}))}{D(\overline{u}, u) + \epsilon_1} \\ &\leq\; -\frac{D(J_h^n(\overline{u}), \overline{u})}{D(\overline{u}, u) + \epsilon_1} . \end{aligned}$$

In the last step the triangle inequality has been used. An estimate for $\sum_{k=1}^{n}\lambda_k$ in the other direction follows from an application of inequality (1.10) together with the Cauchy-Schwarz inequality and $h = t/n$,

$$\begin{aligned}\sum_{k=1}^{n}\lambda_k &\leq \sum_{k=1}^{n}\frac{\sqrt{2h(G(J_h^{k-1}(\overline{u})) - G(J_h^k(\overline{u})))}}{D(\overline{u},u)-\epsilon_1}\\ &\leq \frac{1}{D(\overline{u},u)-\epsilon_1}\cdot\Big(\sum_{k=1}^{n}2h\Big)^{1/2}\cdot\Big(\sum_{k=1}^{n}(G(J_h^{k-1}(\overline{u})) - G(J_h^k(\overline{u})))\Big)^{1/2}\\ &= \frac{\sqrt{2t(G(\overline{u}) - G(J_{t/n}^n(\overline{u})))}}{D(\overline{u},u)-\epsilon_1}.\end{aligned}$$

Combining these estimates (and using also $G(\overline{u}) - G(u) > 0$) results in

$$\begin{aligned}G(J_{t/n}^n(\overline{u})) &\leq G(u) + \exp\Big\{-\frac{D(J_{t/n}^n(\overline{u}),\overline{u})}{D(\overline{u},u)+\epsilon_1}\Big\}(G(\overline{u}) - G(u))\\ &\quad+\frac{S(D(\overline{u},u)+\epsilon_1)^2\cdot\sqrt{2t(G(\overline{u}) - G(J_{t/n}^n(\overline{u})))}}{D(\overline{u},u)-\epsilon_1}.\end{aligned}$$

Now let $n\to\infty$ and use $J_{t/n}^n(\overline{u})\to\overline{u}(t)$,

$$\begin{aligned}G(\overline{u}(t)) &\leq G(u) + \exp\Big\{-\frac{D(\overline{u}(t),\overline{u})}{D(\overline{u},u)+\epsilon_1}\Big\}(G(\overline{u}) - G(u))\\ &\quad+\frac{S(D(\overline{u},u)+\epsilon_1)^2\cdot\sqrt{2t(G(\overline{u}) - G(\overline{u}(t)))}}{D(\overline{u},u)-\epsilon_1}.\end{aligned}$$

For $\epsilon_2 > 0$ and $x \geq 0$ sufficiently small one has

$$e^{-x} \leq 1 - (1-\epsilon_2)x\,.$$

Applying this and rearranging terms results in

$$\begin{aligned}&\frac{G(\overline{u}) - G(\overline{u}(t))}{D(\overline{u}(t),\overline{u})} + \frac{S(D(\overline{u},u)+\epsilon_1)^2\cdot\sqrt{2t(G(\overline{u}) - G(\overline{u}(t)))}}{D(\overline{u}(t),\overline{u})\cdot(D(\overline{u},u)-\epsilon_1)}\\ &\qquad\geq(1-\epsilon_2)\frac{G(\overline{u}) - G(u)}{D(\overline{u},u)+\epsilon_1}\,.\end{aligned}$$

In the above the assumption that $\overline{u}$ is not a stationary point has been used to guarantee $D(\overline{u},\overline{u}(t)) \neq 0$. This assumption also allows one to use Lemma 2.15, and thus there is some $C_2 > 0$ such that for small $t > 0$

$$D(\overline{u}(t),\overline{u}) \geq C_2 t\,. \tag{2.5}$$

Hence

$$\frac{G(\overline{u})-G(\overline{u}(t))}{D(\overline{u}(t),\overline{u})}+\frac{\sqrt{2}S(D(\overline{u},u)+\epsilon_1)^2}{\sqrt{C_2}(D(\overline{u},u)-\epsilon_1)}\sqrt{\frac{G(\overline{u})-G(\overline{u}(t))}{D(\overline{u}(t),\overline{u})}}$$
$$\geq(1-\epsilon_2)\frac{G(\overline{u})-G(u)}{D(\overline{u},u)+\epsilon_1}\,.$$

Let $t\longrightarrow 0+$ and use the definition of C_1 from the beginning of the proof,

$$\liminf_{t\to 0+}\frac{G(\overline{u})-G(\overline{u}(t))}{D(\overline{u}(t),\overline{u})}+\frac{\sqrt{2C_1}S(D(\overline{u},u)+\epsilon_1)^2}{\sqrt{C_2}(D(\overline{u},u)-\epsilon_1)}\geq(1-\epsilon_2)\frac{G(\overline{u})-G(u)}{D(\overline{u},u)+\epsilon_1}\,.$$

Now let $\epsilon_1\longrightarrow 0$ and $\epsilon_2\longrightarrow 0$, and finally $u\longrightarrow\overline{u}$. This results in

$$\liminf_{t\to 0+}\frac{G(\overline{u})-G(\overline{u}(t))}{D(\overline{u}(t),\overline{u})}\geq\limsup_{u\to\overline{u}}\frac{G(\overline{u})-G(u)}{D(\overline{u},u)}\,.$$

This shows in particular the existence of the full limit of the quotient on the left-hand side, and shows equation (2.3).

Finally for $t_0>0$ the Lipschitz continuity of $t\mapsto G(u(t))$ shows the existence of a constant C_3 such that for $t-t_0>0$ small enough

$$G(u(t_0))-G(u(t))\leq C_3(t-t_0)\,,$$

so that, using (2.5) as well,

$$\lim_{t\to t_0+}\frac{G(u(t_0))-G(u(t))}{D(u(t),u(t_0))}\leq\frac{C_3}{C_2}<\infty\,.$$

The proof is complete. □

While the above theorem shows that the gradient flow moves along a path of steepest descent, it does not address the question whether there is a unique steepest direction. The following theorem answers this question in the affirmative.

Theorem 2.16. *Assume* $0<|\nabla_-G|(u(t_0))<\infty$, *where* G *and* $u(t)$ *are as in Theorem* 1.13. *Let* $u_k\longrightarrow u(t_0)$ *be any sequence of points with*

$$\lim_{k\to\infty}\frac{G(u(t_0))-G(u_k)}{D(u_k,u(t_0))}=|\nabla_-G|(u(t_0))\,.$$

Then there is a sequence $s_k\longrightarrow 0+$ *such that*

$$\lim_{k\to\infty}\frac{D(u_k,u(t_0+s_k))}{D(u_k,u(t_0))}=0\,.$$

Proof. Set $d_k = D(u_k, u(t_0))$ and by continuity of $u(t)$ pick $s_k \geq 0$ such that $D(u(t_0 + s_k), u(t_0)) = d_k$. Set $e_k = D(u_k, u(t_0 + s_k))$ and $v_k = (1/2)u_k + (1/2)u(t_0 + s_k)$. Finally, set $\beta = |\nabla_- G|(u(t_0))$. Then by the triangle inequality, by the assumptions, and by Theorem 2.14 one has

$$\begin{aligned} e_k &\leq 2d_k\,, \\ G(u_k) &= G(u(t_0)) - \beta d_k + o(d_k)\,, \\ G(u(t_0 + s_k)) &= G(u(t_0)) - \beta d_k + o(d_k)\,. \end{aligned}$$

Using the convexity property of the functional one has

$$\begin{aligned} G(v_k) &\leq \frac{1}{2}G(u_k) + \frac{1}{2}G(u(t_0 + s_k)) + \frac{S}{4}e_k^2 \\ &\leq G(u(t_0)) - \beta d_k + o(d_k) + Sd_k^2\,. \end{aligned}$$

For d_k sufficiently small this implies $G(v_k) < G(u(t_0))$. Therefore, if k is sufficiently large, one must have $v_k \neq u(t_0)$. Hence one has $0 < D^2(v_k, u(t_0)) \leq d_k^2 - e_k^2/4$ by the NPC condition for midpoints. Using this together with the estimate for $G(v_k)$ implies

$$\frac{G(u(t_0)) - G(v_k)}{D(v_k, u(t_0))} \geq \frac{\beta d_k - o(d_k) - Sd_k^2}{\sqrt{d_k^2 - e_k^2/4}} = \frac{\beta - o(d_k)/d_k - Sd_k}{\sqrt{1 - (e_k/(2d_k))^2}}\,.$$

Letting $k \longrightarrow \infty$ and using Theorem 2.14 yields

$$\beta \geq \limsup_{k \to \infty} \frac{\beta}{\sqrt{1 - (e_k/(2d_k))^2}}\,,$$

and hence $\limsup\limits_{k \to \infty} \dfrac{e_k}{d_k} = 0$. □

The two previous theorems have shown that the flow moves in the direction of steepest descent. However, they give no indication upon the speed of the flow. Now, if $u(t)$ solves $u_t = -\nabla G(u)$ in the classical sense, then certainly $|u_t| = |\nabla G(u)|$. This fact generalizes.

Theorem 2.17. *Let G and $u(t)$ be as in Theorem 1.13. Then*

$$\lim_{s \to 0+} \frac{D(u(t+s), u(t))}{s} = |\nabla_- G|(u(t))\,,$$

and the existence of the limit for all times t is part of the statement.

Proof. The semigroup property allows to assume $t = 0$. Set $u_1 = u_0$ in inequality (2.1), regroup, divide by λ, and then let $\lambda \to 0$. This implies

$$\left(\frac{1}{h} - S\right) D^2(J_h(u_0), u_0) \leq G(u_0) - G(J_h(u_0)) .$$

Iterate this formula and use the Cauchy-Schwarz inequality to get

$$\left(\frac{1}{h} - S\right) \frac{1}{n} D^2(J_h^n(u_0), u_0) \leq G(u_0) - G(J_h^n(u_0)) .$$

Now set $h = s/n$ and let $n \to \infty$, then, using the lower semicontinuity of G,

$$\frac{1}{s} D^2(u_0, u(s)) \leq G(u_0) - G(u(s)) .$$

Theorem 2.14 then implies

$$\limsup_{t \to 0+} \frac{D(u(s), u_0)}{s} \leq |\nabla_- G|(u_0) .$$

For the other direction assume without loss of generality $|\nabla_- G|(u_0) > 0$. Fix any $\epsilon > 0$ and any positive $\delta < |\nabla_- G|(u_0)$, and take some u_1 with $D(u_0, u_1) < \epsilon\delta/3S$ such that

$$\frac{G(u_0) - G(u_1)}{D(u_0, u_1)} > \delta .$$

The lower semicontinuity of G then implies

$$\frac{G(u(s)) - G(u_1)}{D(u_0, u_1)} > \delta$$

for sufficiently small $s > 0$. For small $s > 0$ one has $e^{2Ss} < 1 + 3Ss$ and hence by Lemma 2.8

$$D^2(u(s), u_1) \leq (1 + 3Ss) D^2(u_0, u_1) - 2s(G(u(s)) - G(u_1)) ,$$

so that

$$D^2(u(s), u_1) \leq (1 + 3Ss) D^2(u_0, u_1) - 2s\delta D(u(s), u_1) .$$

The estimates $D(u_0, u_1) \leq D(u_0, u(s)) + D(u(s), u_1)$ and $D(u_0, u_1) < \epsilon\delta/3S$ imply

$$D(u(s), u_1) \leq D(u_0, u_1) + \frac{D(u_0, u(s))}{D(u(s), u_1)} (D(u_0, u_1) + \epsilon\delta s) - (2 - \epsilon)\delta s .$$

Finally use $D(u_0, u(s)) \geq D(u_0, u_1) - D(u_1, u(s))$ to conclude

$$\frac{D(u_0, u(s))}{s}\left(1 + \frac{D(u_0, u_1)}{D(u(s), u_1)}\right) \geq (2-\epsilon)\delta - \epsilon\delta \frac{D(u_0, u(s))}{D(u(s), u_1)},$$

and hence

$$2 \liminf_{s \to 0+} \frac{D(u(s), u_0)}{s} \geq (2-\epsilon)\delta .$$

The desired statement follows now upon $\epsilon \to 0$ and $\delta \to |\nabla_- G|(u(t))$. □

Corollary 2.18. *Let G and $u(t)$ be as before. Then for almost all $t > 0$*

$$-\frac{dG(u(t))}{dt} = (\,|\nabla_- G|(u(t))\,)^2 .$$

Proof. Let $g(t) = G(u(t))$, and assume without loss of generality that $u(t)$ is not a stationary point of G. Then

$$\frac{g(t) - g(t+s)}{s} = \frac{G(u(t)) - G(u(t+s))}{D(u(t+s), u(t))} \cdot \frac{D(u(t+s), u(t))}{s},$$

and the corollary follows from Theorems 2.14 and 2.17, and from the Lipschitz continuity of $G(u(t))$. □

Remark. This is in fact true for all times $t > 0$ if one considers the derivative from the right; see [1] for the Hilbert-space case.

Corollary 2.19. *If $\lim_{t\to\infty} G(u(t)) > -\infty$ then $|\nabla_- G|(u(.)) \in L^2([0,\infty))$.*

Proof. Obvious. □

2.7. Consistency in the Hilbert-space case.

If the NPC space L is a Hilbert space then the solutions obtained with the method of Theorem 1.13 will coincide with those obtained by the classical Crandall–Liggett method. One way to show this is to show that the time-step map J_h does in fact correspond to a finite difference. The reader is referred to [1] for for the definition of the subdifferential $A = \partial G$ of a convex functional G defined on a Hilbert space. The functional G under consideration in Theorem 1.8 does not need to be convex, however condition (b) in that theorem implies that $G(u) + S|u - u_0|^2$ is a convex functional, where $|.|$ denotes the norm in L, and one can consider the subdifferential of the latter functional instead.

Proposition 2.20. *Let L be a Hilbert space, G as in Theorem* 1.8 *with* $G(u_0) < \infty$. *Let* $h < 1/(2S)$. *Then for the time step map J_h one has*

$$\frac{J_h(u_0) - u_0}{h} \in -\partial G(J_h(u_0))\,.$$

The idea of the proof (for a convex G) is that if the proposition is wrong, then one can find a point on the graph of G in $L \times \mathbf{R}$ that lies below the plane determined by the normal vector $(\frac{1}{h}(J_h(u_0) - u_0), 1)$ through $(J_h(u_0), G(J_h(u_0)))$, which in turn leads to a contradiction to the minimizing property of $J_h(u_0)$. For the complete proof of the proposition and for further details see [11]. This proposition shows

$$J^n_{t/n}(u_0) = (I + \frac{t}{n}\partial G)^{-n}(u_0)\,.$$

Crandall and Liggett [2] showed that if $u(t) = \lim_{n\to\infty}(I + \frac{t}{n}A)^{-n}(u_0)$ is strongly differentiable almost everywhere, and if A satisfies fairly general conditions, then $u(t)$ is a strong solution of $0 \in \frac{du}{dt} + A(u)$, $u(0) = u_0$, and conversely. This establishes the desired consistency, namely that the solution produced with the methods herein are identical with the one obtained with the classical Crandall–Liggett method.

A more direct approach to the question of consistency can be taken along the lines of the results on stationary points and steepest descent. Using Theorems 2.14 and 2.17 one can derive the following theorem directly (for the details see [11]).

Theorem 2.21. *Assume L is a Hilbert space, G and $u(t)$ are as in Theorem* 1.13, *and $u(t)$ is differentiable at t_0. Then*

$$u'(t_0) \in -\partial G(u(t_0))\,.$$

More precisely, $u'(t_0)$ is the element of smallest norm in $-\partial G(u(t_0))$.

2.8. Energy-convergence.

Theorem 1.13 shows *metric* convergence of the approximate solutions $J^n_{t/n}(u_0)$ to the gradient flow $u(t)$. The aim of this section is to show that the functional also converges along the approximate solutions, that is $G(J^n_{t/n}(u_0))$ converges to $G(u(t))$.

Lemma 2.22. *Let G be as in Theorem* 1.13. *For any $u_0, u_1 \in L$ with $G(u_0) < \infty$ one has*

$$|\nabla_- G|(u_0) \geq \frac{G(u_0) - G(u_1)}{D(u_0, u_1)} - SD(u_0, u_1)\,.$$

Proof. If $G(u_1) \geq G(u_0)$ then this is trivial. Hence assume $G(u_1) < G(u_0)$ and set $u_t = (1-t)u_0 + tu_1$. Using the convexity properties of G, and $D(u_t, u_0) = tD(u_1, u_0)$, one derives

$$\frac{G(u_0) - G(u_t)}{D(u_t, u_0)} \geq \frac{G(u_0) - G(u_1)}{D(u_0, u_1)} - S(1-t)D(u_0, u_1)\,.$$

The results follows from the definition of $|\nabla_- G|$ upon letting $t \to 0$. ☐

Lemma 2.23. *Let G be as in Theorem* 1.13. *Then for the time-step map J_h the following holds:*

$$|\nabla_- G|(J_h(u_0)) \leq |\nabla_- G|(u_0) + 2SD(u_0, J_h(u_0))\,.$$

Proof. Set $v = J_h(u_0)$ and assume $|\nabla_- G|(v) \neq 0$. Take any $w \in L$ with $D(v, w) < |\nabla_- G|(v)/2S$ and with $(G(v) - G(w))/D(v, w) > |\nabla_- G|(v)/2$. Set $u_\lambda = (1-\lambda)v + \lambda w$. The convexity properties of G together with the choice of w imply $G(u_\lambda) \leq G(v)$, hence by the minimality property of v one has $D(u_\lambda, u_0) \geq D(v, u_0)$. Let

$$u_{\lambda,t} = (1-t)u_0 + tu_\lambda\,, \qquad t = \frac{D(v, u_0)}{D(u_\lambda, u_0)}\,,$$

then $D(v, u_0) = D(u_{\lambda,t}, u_0)$. Using again the minimality property of v and the convexity properties of G one has

$$\begin{aligned} G(v) &\leq G(u_{\lambda,t}) \leq (1-t)G(u_0) + t(1-\lambda)G(v) + t\lambda G(w) \\ &\quad + t\lambda(1-\lambda)SD^2(v, w) + St(1-t)D^2(u_0, u_\lambda)\,. \end{aligned}$$

The triangle inequality implies

$$(1-t) \leq \lambda \frac{D(w, v)}{D(u_\lambda, u_0)}\,.$$

Bringing $G(v)$ to the other side, estimating $1-t$ from above, and dividing by λ implies

$$\begin{aligned} 0 &\leq \frac{D(w, v)}{D(u_\lambda, u_0)}(G(u_0) - G(v)) - t(G(v) - G(w) - SD^2(v, w)) \\ &\quad + tSD(w, v)D(u_\lambda, u_0) - t\lambda SD^2(v, w)\,. \end{aligned}$$

Now let $\lambda \to 0$, which implies $u_\lambda \to v$ and $t \to 1$, and results after regrouping in

$$\frac{G(v)-G(w)}{D(w,v)} - SD(v,w) \leq \frac{G(u_0)-G(v)}{D(u_0,v)} + SD(u_0,v)\,.$$

Finally let $w \to v$ and use the definition of $|\nabla_- G|$,

$$|\nabla_- G|(v) \leq \frac{G(u_0)-G(v)}{D(u_0,v)} + SD(u_0,v)\,.$$

By Lemma 2.22 one has

$$\frac{G(u_0)-G(v)}{D(u_0,v)} \leq |\nabla_- G|(u_0) + SD(u_0,v)\,,$$

which completes the proof. □

Theorem 2.24. *If* $|\nabla_- G|(u_0) < \infty$ *and* $\frac{k(n)T}{n} \to t$ *as* $n \to \infty$ *for some* $t \in I_A$ *as in Theorem* 1.13*, then*

$$G(J_{T/n}^{k(n)}(u_0)) \to G(u(t))\,.$$

Proof. Iterating Lemma 2.23 yields

$$|\nabla_- G|(J_{T/n}^{k(n)}(u_0)) \leq |\nabla_- G|(u_0) + 2S\sum_{l=1}^{k(n)} D(J_{T/n}^{l}(u_0), J_{T/n}^{l-1}(u_0))\,.$$

The sum can be estimated in the same way as in the computations leading to (1.11) to obtain

(2.6)
$$|\nabla_- G|(J_{T/n}^{k(n)}(u_0)) \leq |\nabla_- G|(u_0) + 2S\sqrt{\frac{2k(n)T}{n}(G(u_0) - G(J_{T/n}^{k(n)}(u_0)))}\,.$$

By the remark after the proof of Theorem 1.13 one has

$$J_{T/n}^{k(n)}(u_0) \to u(t)\,.$$

Let $n \to \infty$, then

$$\limsup_{n\to\infty} |\nabla_- G|(J_{T/n}^{k(n)}(u_0)) \leq |\nabla_- G|(u_0) + 2S\sqrt{2t(G(u_0) - G(u(t)))}\,,$$

in particular $|\nabla_- G|(J_{T/n}^{k(n)}(u_0))$ stays bounded.

Lemma 2.22 implies

$$|\nabla_{-}G|(J_{T/n}^{k(n)}(u_0)) \geq \frac{G(J_{T/n}^{k(n)}(u_0)) - G(u(t))}{D(J_{T/n}^{k(n)}(u_0), u(t))} - SD(J_{T/n}^{k(n)}(u_0), u(t)) .$$

Therefore the boundedness of the gradient terms implies

$$\limsup_{n\to\infty} G(J_{T/n}^{k(n)}(u_0)) \leq G(u(t)) ;$$

by the lower semicontinuity of G one has the reverse inequality for the lower limit. □

2.9. Behavior for large time.

As before, let G and the flow $u(t)$ be as in Theorem 1.13. Investigation of the behavior of $u(t)$ as $t \to \infty$ is the contents of this section. For this purpose it is necessary to first prove some results about the function $|\nabla_{-}G|$, as defined in Definition 5, and about the speed of the flow. It will then be possible to show that the only possible accumulation points of $u(t)$ as $t \to \infty$ are the stationary points of G.

Proposition 2.25. *The function* $|\nabla_{-}G| : L \to \mathbf{R} \cup \{+\infty\}$ *is lower semicontinuous.*

Proof. Fix some $u_0 \in L$. If $|\nabla_{-}G|(u_0) = 0$ then lower semicontinuity at u_0 is trivial. Assume now $|\nabla_{-}G|(u_0) > 0$ and $G(u_0) < \infty$; the case $|\nabla_{-}G|(u_0) = \infty$ is not excluded. Take any positive $\delta < |\nabla_{-}G|(u_0)$ and any small $\epsilon > 0$. Set $R = \min\{\epsilon, \epsilon\delta/(2S)\}$ and take a $v \in B_R(u_0)$ with

$$\frac{G(u_0) - G(v)}{D(u_0, v)} > \delta . \tag{2.7}$$

For u with $D(u, u_0) < \epsilon D(u_0, v)$ let $u_\lambda = (1-\lambda)u + \lambda v$. Then the inequality $G(u_\lambda) \leq (1-\lambda)G(u) + \lambda G(v) + S\lambda(1-\lambda)D^2(u, v)$ implies

$$\frac{G(u) - G(u_\lambda)}{D(u, u_\lambda)} \geq \frac{G(u) - G(v) - S(1-\lambda)D^2(u, v)}{D(u, v)} .$$

Letting $\lambda \to 0$ yields

$$|\nabla_{-}G|(u) \geq \frac{G(u) - G(u_0)}{D(u, v)} + \frac{G(u_0) - G(v) - SD^2(u, v)}{D(u, v)} .$$

The lower semicontinuity of G implies $G(u) \geq G(u_0) - \delta\epsilon(1-\epsilon)D(u_0, v)$ for u sufficiently close to u_0. By the previous assumptions one also has $D(u,v) \geq (1-\epsilon)D(u_0, v)$ and hence

$$\frac{G(u) - G(u_0)}{D(u,v)} \geq -\delta\epsilon\,.$$

Inequality (2.7) together with $D(u,v) \leq \delta\epsilon/S$, which follows from u, v being in the ball $B_R(u_0)$, implies

$$\frac{G(u_0) - G(v) - SD^2(u,v)}{D(u,v)} \geq \frac{\delta D(u_0,v) - \delta\epsilon D(u,v)}{D(u,v)}\,.$$

The estimate $D(u_0,v) \geq D(u,v)/(1+\epsilon) \geq (1-\epsilon)D(u,v)$ then results in

$$\frac{G(u_0) - G(v) - SD^2(u,v)}{D(u,v)} \geq (1-2\epsilon)\delta\,.$$

By combining the above estimates one obtains

$$|\nabla_{-}G|(u) \geq (1-3\epsilon)\delta$$

for u sufficiently close to u_0. Upon letting $\epsilon \to 0$ and $\delta \to |\nabla_{-}G|(u_0)$ one has

$$\liminf_{u \to u_0} |\nabla_{-}G|(u) \geq |\nabla_{-}G|(u_0)\,.$$

It remains to check the case when there are $u_k \to u_0$ with $G(u_k) < \infty$ and $G(u_0) = \infty$. Fixing any $v \in L$ with $G(v) < \infty$, then by Lemma 2.22

$$|\nabla_{-}G|(u_k) \geq \frac{G(u_k) - G(v)}{D(u_k,v)} - SD(u_k,v)\,,$$

and as $k \to \infty$ the right-hand side converges to $\infty = |\nabla_{-}G|(u_0)$ by the lower semicontinuity of G. □

One might expect that if a gradient flow does not diverge to negative infinity then it should flatten out. This is in fact true. The proof of this statement needs a few lemmas which are provided below.

Lemma 2.26. *Either* $\lim_{t\to\infty} G(u(t)) = -\infty$ *or there is a sequence* $\{t_n\}$ *with* $t_n \to \infty$ *such that* $|\nabla_{-}G|(u(t_n)) \to 0$.

Proof. Assume no such sequence exists. Then there is a $T > 0$ and an $\epsilon > 0$ such that for $t \geq T : |\nabla_- G|(u(t)) \geq \epsilon$. Corollary 2.18 then implies for almost all t

$$-\frac{dG(u(t))}{dt} \geq \epsilon^2 .$$

Integrating this Lipschitz-continuous function yields

$$G(u(T)) - G(u(t)) \geq (t-T)\epsilon^2 \to \infty ,$$

that is $\lim\limits_{t\to\infty} G(u(t)) = -\infty$. □

Lemma 2.27. *Let G and $u(t)$ be as before. Then for $s, t > 0$*

$$|\nabla_- G|(u(s+t)) - |\nabla_- G|(u(s)) \leq 2\sqrt{2}S \int_s^{s+t} |\nabla_- G|(u(\sigma))\, d\sigma .$$

Proof. By breaking $[s, s+t]$ into smaller subintervals one may assume $t \in I_A$ (see Theorem 1.13 for the definition of I_A). In equation (2.6) set $k(n) = n$, $T = t$, and replace u_0 by $u(s)$, then by the semigroup property of the flow, and by the lower semicontinuity of G and $|\nabla_- G|$, one has after passing to the limit

$$|\nabla_- G|(u(s+t)) \leq |\nabla_- G|(u(s)) + 2S\sqrt{2t(G(u(s)) - G(u(s+t)))} .$$

Rearranging terms yields

$$|\nabla_- G|(u(s+t)) - |\nabla_- G|(u(s)) \leq 2\sqrt{2}S\sqrt{\frac{G(u(s)) - G(u(s+t))}{t}} \cdot t ,$$

which can be iterated to show

$$\begin{aligned} &|\nabla_- G|(u(s+nt)) - |\nabla_- G|(u(s)) \\ &\leq 2\sqrt{2}S \sum_{k=0}^{n-1} \sqrt{\frac{G(u(s+kt)) - G(u(s+(k+1)t))}{t}} \cdot t . \end{aligned}$$

Set $\tau = nt$, then

$$\begin{aligned} &|\nabla_- G|(u(s+\tau)) - |\nabla_- G|(u(s)) \\ &\leq 2\sqrt{2}S \sum_{k=0}^{n-1} \sqrt{\frac{G(u(s+\frac{k\tau}{n})) - G(u(s+\frac{(k+1)\tau}{n}))}{\frac{\tau}{n}}} \cdot \frac{\tau}{n} . \end{aligned}$$

For fixed s, τ the function $G(u(t))$ is uniformly Lipschitz continuous on the interval $[s, s+\tau]$, which results in a uniform bound of the square root term. Define

$$\phi_n(\sigma) = \frac{G(u(s+\frac{k\tau}{n})) - G(u(s+\frac{(k+1)\tau}{n}))}{\frac{\tau}{n}}$$

for $\sigma \in (s + \dfrac{k\tau}{n}, s + \dfrac{(k+1)\tau}{n}]$. At a point σ where $G(u(t))$ is differentiable one has

$$\phi_n(\sigma) \to -\frac{dG(u(\sigma))}{d\sigma} = \Big(|\nabla_- G|(u(\sigma))\Big)^2,$$

hence almost everywhere. Finally,

$$\int_s^{s+\tau} \sqrt{\phi_n(\sigma)}\, d\sigma = \sum_{k=0}^{n-1} \sqrt{\frac{(G(u(s+\frac{k\tau}{n}))) - G(u(s+\frac{(k+1)\tau}{n}))}{\frac{\tau}{n}}} \cdot \frac{\tau}{n},$$

and an application of the Lebesgue Dominated Convergence Theorem completes the proof. □

Corollary 2.28. *The function* $|\nabla_- G|(u(t))$ *is continuous from the right.*

Proof. This is immediate from the lemma above and the fact that $|\nabla_- G|(u(t))$ is lower semicontinuous by Proposition 2.25 and locally L^2 by Corollary 2.18. □

Finally, one needs the following lemma from real analysis, which is stated without proof here.

Lemma 2.29. *Assume* $f : [a,b] \to \mathbf{R}$ *is lower semicontinuous and continuous from the right, and* $f(a) < c < f(b)$. *Then there is a point* $d \in (a,b)$ *such* $f(d) = c$ *and* $f(x) \geq c$ *for all* $x \in [d,b]$.

Theorem 2.30. *Let* G *and* $u(t)$ *be as in Theorem* 1.13, *and assume* $\lim_{t\to\infty} G(u(t)) > -\infty$. *Then* $\lim_{t\to\infty} |\nabla_- G|(u(t)) = 0$.

Proof. Assume the theorem is not true. Then there is an $\epsilon > 0$ and a sequence $t_n \to \infty$ such that $|\nabla_- G|(u(t_n)) > \epsilon$. By the boundedness assumption $G(u(t))$ has a finite limit for $t \to \infty$, hence on can take a time s_0 such that

$$16S|G(u(t)) - G(u(s))| < \epsilon^2$$

for $s, t \geq s_0$. Take a time $s_1 \geq s_0$ for which $|\nabla_- G|(u(s_1)) \leq \epsilon/2$, the existence follows from Lemma 2.26. Finally, pick some $t_N > s_1$ such that

$|\nabla_{-}G|(u(t_N)) > \epsilon$. By the lemma directly before the statement of the theorem there is a time $s_2 \in [s_1, t_N]$ such that

$$|\nabla_{-}G|(u(s_2)) = \frac{\epsilon}{2}, \quad |\nabla_{-}G|(u(t)) \geq \frac{\epsilon}{2} \text{ for } t \in [s_2, t_N].$$

Lemma 2.27 and Corollary 2.18 then imply

$$\begin{aligned} |\nabla_{-}G|(u(t_N)) - |\nabla_{-}G|(u(s_2)) &\leq 2\sqrt{2}S \int_{s_2}^{t_N} |\nabla_{-}G|(u(t))\,dt \\ &\leq \frac{8S}{\epsilon}\int_{s_2}^{t_N} \Big(|\nabla_{-}G|(u(t))\Big)^2 dt \\ &= \frac{8S}{\epsilon}\int_{s_2}^{t_N} \Big(-\frac{dG(u(t))}{dt}\Big) dt \\ &= \frac{8S}{\epsilon}\Big(G(u(s_2)) - G(u(t_N))\Big) \\ &\leq \frac{\epsilon}{2}. \end{aligned}$$

Hence $|\nabla_{-}G|(u(t_N)) \leq |\nabla_{-}G|(u(s_2)) + \epsilon/2 = \epsilon$, which is a contradiction. □

Corollary 2.31. *Assume* $u(t_n) \to \overline{u}$ *as* $t_n \to \infty$. *Then* $\overline{u}$ *is a stationary point of* G *with*

$$\lim_{t\to\infty} G(u(t)) = G(\overline{u}).$$

Proof. The lower semicontinuity of G implies $\lim_{t\to\infty} G(u(t)) \geq G(\overline{u}) > -\infty$. By the theorem one has $|\nabla_{-}G|(u(t_n)) \to 0$. The lower semicontinuity of $|\nabla_{-}G|$ then shows $|\nabla_{-}G|(\overline{u}) = 0$.

Assume now that $\lim_{t\to\infty} G(u(t)) > G(\overline{u})$ and pick some $\epsilon > 0$ such that $G(u(t)) > G(\overline{u}) + 2\epsilon$. Let n be large enough to guarantee $SD^2(u(t_n), \overline{u}) < \epsilon$ and let $u_\lambda = (1-\lambda)u(t_n) + \lambda\overline{u}$. Considering the difference quotient of G at $u(t_n)$ with u_λ results in

$$|\nabla_{-}G|(u(t_n)) \geq \frac{\epsilon}{D(u(t_n), \overline{u})}.$$

This implies $|\nabla_{-}G|(u(t_n)) \to \infty$ as $u(t_n) \to \overline{u}$, which contradicts the theorem. □

In general a functional which flattens out does not have to have a critical value. However, some functionals satisfy a generalized Palais–Smale compactness condition which allows to conclude the existence of critical points. To be precise, the functional G is said to satisfy the Palais–Smale condition

$(PS)_c$ at the level $c \in \mathbf{R}$ if every sequence $\{u_n\} \subset L$ with $G(u_n) \to c$ and $|\nabla_- G|(u_n) \to 0$ has a convergent subsequence. Theorem 2.30 can then be used to derive the theorem below. The proof is essentially the same as in the Hilbert space case.

Theorem 2.32. *Assume* $c = \lim_{t\to\infty} G(u(t)) > -\infty$ *and* G *satisfies* $(PS)_c$. *Then the set* $C = \{u \in L : G(u) = c \text{ and } |\nabla_- G|(u) = 0\}$ *is not empty, and* $\lim_{t\to\infty} \operatorname{dist}(u(t), C) = 0$. *If additionally* G *is convex then there is a* $\overline{u} \in C$ *such that* $\lim_{t\to\infty} u(t) = \overline{u}$.

The next theorem is independent of the theorems above. In [8] a very similar theorem is proved. The main difference is that there the functional is assumed to be continuous on a complete path-connected metric space, while herein one assumes the convexity condition from Theorem 1.13 and only lower semicontinuity for the functional, and NPC for the underlying metric space. The proof is an adaptation of the standard proof for the smooth case using the deformation lemma, for the details see [11].

Theorem 2.33 (Mountain-Pass Theorem). *Let* G *be as in Theorem* 1.13. *For two given points* $u_0, u_1 \in L$ *let*

$$\Gamma = \{p : [0,1] \to L : p(0) = u_0, p(1) = u_1, p \text{ is continuous}\}$$

and assume

$$\max\{G(u_0), G(u_1)\} < c := \inf_{p\in\Gamma} \sup_{s\in[0,1]} G(p(s)) .$$

If G *satisfies* $(PS)_c$ *then* c *is a stationary value of* G.

2.10. Flow for convex functionals.

In this section it is assumed that G is a convex functional. This stronger hypothesis is fulfilled in many applications, see for example the harmonic map flow.

A first result are the following two representations.

Proposition 2.34. *In the setting of Theorem* 1.13 *assume additionally that the functional* G *is convex. Then*

$$|\nabla_- G|(u(t)) = \sup_{s>0} \frac{D(u(t), u(t+s))}{s} .$$

If $|\nabla_- G|(u(t)) \neq 0$ *then also*

$$|\nabla_- G|(u(t)) = \sup_{u \neq u(t)} \frac{G(u(t)) - G(u)}{D(u, u(t))} = \sup_{s>0} \frac{G(u(t)) - G(u(t+s))}{D(u(t+s), u(t))}.$$

Proof. The convexity of G allows one to set $S = 0$ in Theorem 2.1, which implies that the flow is nonexpansive. Theorem 2.17 implies $D(u(t+s), u(t)) \leq (|\nabla_- G|(u(t)) + \epsilon)s$ provided $s > 0$ is sufficiently small (depending on the choice of $\epsilon > 0$). Thus for any (large) $s > 0$:

$$D(u(t+s), u(t)) \leq \sum_{k=1}^{n} D(u(t + \frac{ks}{n}), u(t + \frac{(k-1)s}{n}))$$

$$\leq \sum_{k=1}^{n} D(u(t + \frac{s}{n}), u(t)) \leq (|\nabla_- G|(u(t)) + \epsilon)s.$$

Upon letting $\epsilon \to 0$ one has

$$|\nabla_- G|(u(t)) \geq \sup_{s>0} \frac{D(u(t), u(t+s))}{s},$$

hence equality by Theorem 2.17.

For the other statement of the proposition it suffices to show

$$\sup_{u \neq u(t)} \frac{G(u(t)) - G(u)}{D(u, u(t))} \leq \limsup_{u \to u(t)} \frac{G(u(t)) - G(u)}{D(u(t+s), u)}.$$

Once this is established the desired equalities follow from Definition 5 and Theorem 2.14. Take any $u_1 \in G$ with $u_1 \neq u(t)$, and let $u_\lambda = (1-\lambda)u(t) + \lambda u_1$. The convexity of G together with $D(u_\lambda, u(t)) = \lambda D(u_1, u(t))$ shows

$$\frac{G(u(t)) - G(u_\lambda)}{D(u_\lambda, u(t))} \geq \frac{G(u(t)) - G(u_1)}{D(u(t), u_1)},$$

and upon letting $\lambda \to 0$

$$\limsup_{u \to u(t)} \frac{G(u(t)) - G(u)}{D(u, u(t))} \geq \frac{G(u(t)) - G(u_1)}{D(u(t), u_1)}.$$

As u_1 was arbitrary this completes the proof. □

Corollary 2.35. *A point* $u_0 \in L$ *is a stationary point of* G *if and only if* u_0 *is a minimizer of* G.

Proof. This is an immediate consequence of the proposition above and of the definition that a point u_0 is called stationary if $|\nabla_{-}G|(u_0) = 0$. □

The second result of the proposition above can also be rephrased as saying that the graph of G is always steeper in the direction of $u(t)$ than the slope of any secant to any lower level. This intuitively characterizes convexity; it can be used to prove the following result.

Theorem 2.36. *Let L, G, and $u(t)$ be as in Theorem* 1.13, *and assume G is convex. Then the function $g : t \mapsto G(u(t))$ is convex.*

Proof. One uses the following lemma from real analysis. If $g : (a, b) \to \mathbf{R}$ is Lipschitz continuous and nonincreasing, and

$$g'(t) \leq \frac{g(t+s) - g(t)}{s} \tag{2.8}$$

for all $s > 0$ and almost all $t \in (a, b)$, then g is convex. The proof of this lemma is routine and is omitted.

By Corollary 2.18 one has $-g'(t) = (|\nabla_{-}G|(u(t)))^2$ for almost all $t > 0$. If $|\nabla_{-}G|(u(t))$ vanishes then the flow is stationary from time t on, and (2.8) holds trivially. Otherwise, use Proposition 2.34 to conclude

$$-g'(t) = \sup_{s>0} \frac{G(u(t)) - G(u(t+s))}{D(u(t+s), u(t))} \cdot \sup_{s>0} \frac{D(u(t), u(t+s))}{s}.$$

Both suprema are attained as the full limit as $s \to 0$; hence it is possible to combine the two terms, and

$$-g'(t) = \sup_{s>0} \frac{G(u(t)) - G(u(t+s))}{s},$$

which is exactly (2.8). Therefore g is convex on $(0, \infty)$, but then also on $[0, \infty)$ because g is continuous at $t = 0$. □

On a Hilbert space the gradient flow is perpendicular to level sets of the functional G. If G is convex then the sublevel sets

$$\{G \leq c\} \equiv \{u \in L : G(u) \leq c\}$$

are convex sets. The gradient flow is crossing the boundary of $\{G \leq c\}$ perpendicularly. Another way of stating this fact is to say that at the time of crossing the boundary the gradient flow decreases the distance to all points in $\{G \leq c\}$. This formulation can be generalized to NPC spaces. To do so one needs the following improvement of Lemma 2.8, which is obtained by letting $S \to 0$ in the statement of that lemma.

Lemma 2.37. *Let L, G, and $u(t)$ be as in Theorem* 1.13, *and assume G is convex. Let u_1 be any point in L. Then for $t, s \geq 0$*

$$D^2(u(t+s), u_1) \leq D^2(u(t), u_1) - 2s(G(u(t+s)) - G(u_1)) .$$

Finally, recall the definition of the distance from a point $u \in L$ to a set $S \subset L$, $\operatorname{dist}(u, S) = \inf_{w \in S} D(u, w)$. In particular $\operatorname{dist}(u, S) = 0$ if $u \in S$.

Theorem 2.38. *Let L, G, and $u(t)$ be as in Theorem* 1.13, *and assume G is convex. If the sub-level set $\{G \leq c\}$ is nonempty then the function $t \mapsto \operatorname{dist}(u(t), \{G \leq c\})$ is nonincreasing, and strictly decreasing as long as $G(u(t)) > c$. In the latter case the functions $t \mapsto D(u(t), w)$ are strictly decreasing for all $w \in \{G \leq c\}$.*

Proof. Corollary 2.6 shows that the function $t \mapsto G(u(t))$ is nonincreasing. This implies that once $u(t)$ has entered a sublevel set of G it cannot leave it again. This proves the case $G(u(t)) \leq c$. Assume now that $G(u(t)) > c$, and take any $w \in \{G \leq c\}$. Lemma 2.37 with $u_1 = w$ implies

$$D^2(u(t+s), w) \leq D^2(u(t), w) - 2s(G(u(t+s)) - c) .$$

The function $t \mapsto G(u(t))$ is continuous and hence $G(u(t+s)) > c$ for $s > 0$ sufficiently small, which shows

$$D(u(t+s), w) < D(u(t), w) .$$

The level set $\{G \leq c\}$ is closed, convex, and nonempty by assumption. It has been shown in [9] that under these conditions there is a unique point $u_1(t) \in \{G \leq c\}$ which minimizes the distance from $u(t)$ to the set $\{G \leq c\}$. Setting $w = u_1(t)$ in the above argument results in

$$\begin{aligned} \operatorname{dist}(u(t+s), \{G \leq c\}) &\leq D(u(t+s), u_1(t)) \\ &< D(u(t), u_1(t)) = \operatorname{dist}(u(t), \{G \leq c\}) , \end{aligned}$$

and the proof is complete. $\square$

An example. Let $L = [0,1] \times [0,1]$ be endowed with the Euclidean distance and set

$$G(x,y) = \begin{cases} 2y - 2x & \text{for } y \geq 2x , \\ y & \text{for } y \leq 2x . \end{cases}$$

G is a continuous convex functional on L. The gradient flow starting at $(0,1)$ can easily be computed explicitly

$$u(t)=\begin{cases}(2t,1-2t) & \text{for } 0\le t\le 1/6\,,\\ (1/3,2/3-t) & \text{for } 1/6\le t\le 2/3\,,\\ (1/3,0) & \text{for } t\ge 2/3\,.\end{cases}$$

This example shows that while $u(t)$ does converge to a minimizer of G, it does not converge to the closest minimizer to its starting point, which is $(0,0)$.

In general, the flow $u(t)$ is supposed to minimize the functional G in the most effective way. Hence one would expect that $G(u(t))$ converges always to the infimum of G as $t\to\infty$. This is the statement of the next theorem.

Theorem 2.39. *Let L, G, and $u(t)$ be as in Theorem 1.13, and assume G is convex. Then*

$$\lim_{t\to\infty} G(u(t)) = \inf_{u\in L} G(u)\,.$$

Proof. Assume there is a $u_1\in L$ and an $\epsilon>0$ with $G(u(t))\ge G(u_1)+\epsilon$ for all $t\ge 0$. Then Lemma 2.37 implies for $t>0$

$$0\le D^2(u_0,u_1)-2\epsilon t\,.$$

This leads to a contradiction for t large. □

Not all gradient flows can be expected to converge as $t\to\infty$, if only for the simple reason that not every convex lower semicontinuous functional has a minimizer. A typical example would be the functional $x\mapsto e^x$ on the real line. Also, even if a functional has a unique minimizer, it is still possible that a minimizing sequence becomes unbounded. However, minimizing sequences obtained with a gradient flow are special, they minimize in some sense most efficiently.

Proposition 2.40. *Let L, G, and $u(t)$ be as in Theorem 1.13, and assume G is convex.*

- **(a)** *If G has a minimizer then the flow $u(t)$ stays bounded for all time.*
- **(b)** *If the flow has a bounded subsequence $\{u(t_n)\}$ as $t_n\to\infty$ then G has a minimizer.*
- **(c)** *If the flow has a convergent subsequence $\{u(t_n)\}$ as $t_n\to\infty$ then $\overline{u}=\lim_{t\to\infty}u(t)$ exists and is a minimizer of G.*

(d) *If* $M = \{u \in L : G(u) = \inf_{w \in L} G(w)\}$ *is nonempty and the flow has a subsequence* $\{u(t_n)\}$ *with* $\operatorname{dist}(u(t_n), M) \to 0$ *as* $t_n \to \infty$, *then* $\overline{u} = \lim_{t \to \infty} u(t)$ *exists and is a minimizer of* G.

Proof. (a) Let u_1 be a minimizer. By Lemma 2.37 one has $D(u(t), u_1) \leq D(u_0, u_1)$ for all $t \geq 0$.
(b) The sequence $u(t_n)$ is a bounded minimizing sequence by Theorem 2.39. Corollary 1.4 guarantees then the existence of a minimizer.
(c) $\overline{u} = \lim_{n \to \infty} u(t_n)$ is a minimizer of G by Theorem 2.39. By Lemma 2.37 one has $D(u(t_n + s), \overline{u}) \leq D(u(t_n), \overline{u})$ for all $s \geq 0$. This shows convergence.
(d) Let $d_n = \operatorname{dist}(M, u(t_n))$, then $d_n \to 0$ as $n \to \infty$. Pick points $u_n \in M$ such that $D(u_n, u(t_n)) = \operatorname{dist}(M, u(t_n))$, for the existence see [9]. Theorem 2.38 implies that the distance from $u(t)$ to any minimizer of G is a nonincreasing function. Hence $D(u(t), u_n) \leq d_n$ for all $t \geq t_n$, and therefore $D(u(t), u(s)) \leq 2d_n$ for all $s, t \geq t_n$. This shows convergence as L is complete. □

A bounded sequence does not necessarily have a convergent subsequence without any further compactness assumptions. To overcome this, one can either make an assumption on the functional, like the Palais–Smale condition (see Theorem 2.32), or one can make an assumption on the underlying space. The following result follows directly from Proposition 2.40.

Theorem 2.41. *Let* L *be an NPC space in which closed and bounded sets are compact. Assume* G *is lower semicontinuous, convex, and has a minimizer. Then the gradient flow for* G *converges to a minimizer of* G *as* $t \to \infty$.

One can do away with compactness assumptions if G is uniformly convex; then Lemma 1.7 together with Theorem 2.39 imply the convergence of the flow as $t \to \infty$.

Theorem 2.42. *Let* L *be an NPC space and assume* G *is lower semicontinuous and uniformly convex. Then the gradient flow for* G *converges to the unique minimizer of* G *as* $t \to \infty$.

3. Harmonic map flow of maps with an NPC target.

3.1. A short introduction to $W^{1,p}(\Omega, X)$.

Recently mathematicians have been working on generalizing the concept of a harmonic map from a manifold into another manifold, which was assumed to be embedded into some Euclidean space by the Nash embedding theorem. It has been possible to replace the target space by a nonpositively curved metric space, see the work of Jost [6, 7], and the work of Korevaar and Schoen [9, 10]. The material contained in this section follows the approach by Korevaar and Schoen. The results and definitions in this section are essentially quoted from [9] and are provided here for the convenience of the reader.

Let (M, g) be a Riemannian manifold and (X, d) be an NPC space. Let $\Omega \subset M$ be orientable, connected, and open. The case $\Omega = M$ is not excluded. Let $Q(x)$ be a Borel measurable function with separable range. $L^p(\Omega, X)$ is the set of Borel measurable functions with separable range for which the integral of the pth power of the pointwise distance to $Q(x)$ is finite. Clearly this definition depends on the choice of the function Q. In case Ω has finite volume one usually chooses Q to be a constant function; the resulting space is independent of the choice of the constant. $L^p(\Omega, X)$ is a complete metric space with the usual distance function, compare also [4]. For $u \in L^p(\Omega, X)$ one defines approximate ϵ-energy densities

$$e_\epsilon(x) = \frac{n+p}{\epsilon^n} \int_{B(x,\epsilon)} \frac{d^p(u(x), u(y))}{\epsilon^p} \, d\mu_g(y)$$

where $B(x, \epsilon)$ is the geodesic ball of radius ϵ about x. The e_ϵ are bounded continuous functions (away from $\partial\Omega$), and integration against them defines linear functionals E_ϵ on $C_c(\Omega)$, the set of continuous real valued functions with compact support in Ω. A map $u \in L^p(\Omega, X)$ has finite energy if

$$\sup\Big\{ \limsup_{\epsilon \to 0} E_\epsilon(f) \,:\, 0 \le f \le 1,\, f \in C_c(\Omega) \Big\} < \infty,$$

which by definition is equivalent to $u \in W^{1,p}(\Omega, X)$ for $p > 1$ and $u \in BV(\Omega, X)$ for $p = 1$. For such a map u it is shown that $\lim_{\epsilon \to 0} E_\epsilon(f) \equiv E(f)$ exists for each $f \in C_c(\Omega)$. For $p > 1$ the linear functional E is given by a measure which is absolutely continuous with respect to the measure $d\mu_g(x)$. This measure is denoted by $|\nabla u|_p(x) \, d\mu_g(x)$. $W^{1,1}(\Omega, X)$ is defined to be the subset of $BV(\Omega, X)$ consisting of those functions that have an absolutely continuous energy measure. That is for $u \in W^{1,p}(\Omega, X)$, $p \geq 1$,

$$E(f) = \int_\Omega |\nabla u|_p(x) f(x)\, d\mu_g(x)\,.$$

The p-energy E^u of u is defined to be the norm of the linear functional E generated by u.

For the special case $p = 2$ one defines

$$|\nabla u|^2(x) = \frac{1}{\omega_n} |\nabla u|_2(x)\,,$$

where ω_n denotes the volume of the n-dimensional unit ball. This definition is consistent with the usual way of defining $|du|^2$ for maps between Riemannian manifolds.

3.2. The Dirichlet problem.

For this section (L, D) will be a subset of $(L^2(\Omega, X), D)$ for an NPC space X and a Riemannian domain Ω. It has been remarked in [9] that $L^2(\Omega, X)$ is then an NPC space itself. The functional G is chosen to be the Dirichlet energy

$$G(u) = E^u = \frac{1}{2} \int_\Omega |\nabla u|^2(x)\, d\mu_g(x)\,.$$

The flow governed by this energy functional is known as the heat flow or the harmonic map flow. The reason is that in the classical case the Lagrange–Euler equation of this flow is exactly the heat equation, and stationary solutions are harmonic maps.

In this section the letter E will be used instead of G, and E^u will be used for $G(u)$. In [9] it has been shown that $u \longrightarrow E^u$ is a lower semicontinuous convex functional on $L^2(\Omega, X)$. The general theory is therefore applicable with $L = L^2(\Omega, X)$. Notice that $E^u < \infty$ is equivalent to $u \in W^{1,2}(\Omega, X)$. If Ω has finite volume then constant maps are minimizers for E. The following theorem is immediate.

Theorem 3.1. *For any starting point $u_0 \in W^{1,2}(\Omega, X)$ the gradient flow for the Dirichlet energy exists in the sense of Theorem* 1.13, *and $u(t) \in W^{1,2}(\Omega, X)$ for $t \geq 0$. In case Ω has finite volume the flow stays bounded for all times.*

Corollary 3.2. *Assume additionally that Ω has compact closure and assume that the target space X is sigma-compact. Then the flow $u(t)$ converges to a constant map as $t \longrightarrow \infty$.*

Proof. As the Sobolev energy of the maps $u(t)$ decreases it is uniformly bounded, and, as Ω has finite volume, the theorem states that the L^2-norm of $u(t)$ also remains bounded. Hence by the precompactness result of Korevaar and Schoen [9, Theorem 1.13] there is a convergent subsequence $\{u(t_n)\}$ with $t_n \to \infty$, convergence measured in L^2. Thus Proposition 2.40 implies the full L^2-convergence to a minimizer, which is of course a constant map in this case. □

Remark. The corollary in particular applies when Ω is a compact orientable Riemannian manifold without boundary and X is compact. In this form the corollary was suggested to the author by J. Eells.

For a given element $\phi \in W^{1,2}(\Omega, X)$ and $\partial\Omega \neq \{\}$ one has a well-defined trace map $\operatorname{tr}\phi$ provided Ω is a Lipschitz domain, cf. [9]. This allows to consider the boundary value problem for the harmonic map flow by prescribing that $u(t)$ is to have the same boundary values as ϕ.

Definition 6.

$$W^{1,2}_{\phi}(\Omega, X) = \{u \in W^{1,2}(\Omega, X) \,:\, \operatorname{tr}\phi = \operatorname{tr}u \text{ and } E^u \leq E^{\phi}\}\,.$$

The objective of this definition is to incorporate the boundary conditions into the space under consideration.

Lemma 3.3. *$(W^{1,2}_{\phi}(\Omega, X), D)$ is a nonpositively curved metric space, where D is the restriction of the metric of $L^2(\Omega, X)$.*

Proof. For any sequence $\{u_k\} \subset W^{1,2}_{\phi}(\Omega, X)$ with $u_k \to u$ one has $E^u \leq E^{\phi}$ by the lower semicontinuity of the Dirichlet energy. Furthermore, as $E^{u_k} \leq E^{\phi}$ and $u_k \to u$ in $L^2(\Omega, X)$, one has $\operatorname{tr}u_k \to \operatorname{tr}u$, see [9], and hence $u \in W^{1,2}_{\phi}(\Omega, X)$. This shows $W^{1,2}_{\phi}(\Omega, X)$ is a closed subset of $L^2(\Omega, X)$ and therefore a complete metric space. It remains to show that $W^{1,2}_{\phi}(\Omega, X)$ is a length space, which suffices as $W^{1,2}_{\phi}(\Omega, X)$ then inherits the NPC hypothesis (1.3) from $L^2(\Omega, X)$. To see that $W^{1,2}_{\phi}(\Omega, X)$ is a length space it is enough to show $W^{1,2}_{\phi}(\Omega, X)$ is a convex subset of $L^2(\Omega, X)$. This however is clear due to the convexity of the Dirichlet energy. □

It is now possible to set $L = W^{1,2}_{\phi}(\Omega, X)$ and to use Theorem 1.13. Note that the inequality $E^u \leq E^{\phi}$ in the definition of $W^{1,2}_{\phi}(\Omega, X)$ is only a technical requirement to show $W^{1,2}_{\phi}(\Omega, X)$ is closed. It has no effect on the flow as each time step decreases the energy. In the case Ω has compact closure it has been shown in [9] that any minimizing sequence converges to the unique minimizer for the Dirichlet energy in $W^{1,2}_{\phi}(\Omega, X)$. Their proof in fact shows that the Dirichlet energy is uniformly convex.

Theorem 3.4. (Solvability of the Initial Boundary Value Problem) *For any given map* $\phi \in W^{1,2}(\Omega, X)$ *the following problem admits a solution in the sense of Theorem* 1.13*:*

$$\begin{cases} u(t) \textit{ solves the harmonic map flow for } t \geq 0\,, \\ u(0) = \phi\,, \\ \operatorname{tr} u(t) = \operatorname{tr} \phi \textit{ for } t \geq 0\,. \end{cases}$$

In case Ω *has compact closure* $\overline{u} = \lim_{t \to \infty} u(t)$ *exists and is the unique harmonic function solving the Dirichlet problem with boundary data* $\operatorname{tr} \phi$.

Remark. If Ω is not a Lipschitz domain the traces need not exist. The boundary condition $\operatorname{tr} u = \operatorname{tr} \phi$ is then replaced by the condition $d(u, \phi) \in H^1_0(\Omega)$, compare [14]. Here $H^1_0(\Omega)$ stands for the Sobolev space of real valued functions that can be approximated in the (1,2)-Sobolev norm by smooth functions with compact support in Ω. The thus resulting space $W^{1,2}_{\phi}(\Omega, X)$ is still closed, because if $\{u_k\} \subset W^{1,2}_{\phi}(\Omega, X)$ with $u_k \to u$ in $L^2(\Omega, X)$ then one has a uniform bound for the norm of $d(u_k, \phi)$ in $H^1_0(\Omega)$. By weak compactness there is a $d \in H^1_0(\Omega)$ such that for a subsequence $d(u_k, \phi) \to d$ in $L^2(\Omega)$. For a subsequence one has $d(u_k, \phi) \to d$ a.e. and $d(u_k, u) \to 0$ a.e. The triangle inequality $d(u, \phi) \leq d(u, u_k) + d(u_k, \phi)$ implies in the limit $d(u, \phi) \leq d$ a.e. and as $0 \leq d(u, \phi)$ this together with the fact that $d(u, \phi) \in H^1(\Omega)$ (which follows from the lower semicontinuity of the Dirichlet energy) finally implies $d(u, \phi) \in H^1_0(\Omega)$.

3.3. Equivariant mappings.

Let (M, g) be a Riemannian manifold which is metrically complete. In case $\partial M \neq \{\}$ the boundary is assumed to be smooth and compact. Let $\Gamma = \pi_1(M)$ be the fundamental group of M and let $\widetilde{M}$ be the universal cover of M. If X is a metric space and $\rho : \Gamma \to \operatorname{Isom}(X)$ a homomorphism then ρ is called a representation of Γ. A special example is the action of Γ on $\widetilde{M}$ via deck transformations.

A map $u : \widetilde{M} \to X$ is called Γ-*equivariant* if

$$u(\gamma x) = \rho(\gamma)(u(x)) \quad \forall x \in \widetilde{M}\,, \gamma \in \Gamma\,.$$

It has been pointed out in [9] that for a Γ-equivariant map u the function $d(u(x), u(y))$ is invariant with respect to the domain action. If the map u

is locally a Sobolev map then it follows that the Sobolev energy density (compare Section 3.1) is Γ-invariant, so one may think of it as being defined on M.

Let X be an NPC space and $Q : \widetilde{M} \to X$ a Borel measurable Γ-equivariant map with separable range. The space $L^2_\rho(\widetilde{M}, X)$ is the set of Borel measurable Γ-equivariant functions from $\widetilde{M}$ into X with separable range for which

$$\int_M d^2(u(x), Q(x))\, d\mu_g(x) < \infty ,$$

endowed with the distance function

$$D(u, v) = \int_M d^2(u(x), v(x))\, d\mu_g(x) .$$

This definition makes $L^2_\rho(\widetilde{M}, X)$ into an NPC space.

As in Section 3.2 the functional G is chosen to be the Dirichlet energy, restricted to a fundamental domain, of course,

$$G(u) = \frac{1}{2} \int_M |\nabla u|^2(x)\, d\mu_g(x) ,$$

and the space L under consideration is $L^2_\rho(\widetilde{M}, X)$. The general theory is applicable as before.

Theorem 3.5. *If ϕ is a Γ-equivariant map from $\widetilde{M}$ into X with finite Dirichlet energy then the harmonic map flow starting at ϕ has a solution $u(t)$ in the sense of Theorem* 1.13. *Furthermore, $u(t)$ is Γ-equivariant for $t \geq 0$.*

3.4. Flow for the p-Sobolev energy, $p < 2$.

The setting of this section is similar to the one on the Dirichlet problem. (L, D) will be a subset of $(L^2(\Omega, X), D)$ for an NPC space X and a Riemannian domain Ω. It might seem somewhat artificial to choose $L^2(\Omega, X)$ instead of $L^p(\Omega, X)$, however this is necessary because $L^p(\Omega, X)$ is an NPC space only if $p = 2$. The functional G is chosen to be the p-Sobolev energy

$$G(u) = E^u = \int_\Omega |\nabla u|_p(x)\, d\mu_g(x) .$$

In case Ω is a Euclidean domain and $X = \mathbf{R}$ then the Euler–Lagrange equation for stationary points of this energy functional is

$$\operatorname{div}(|\nabla u|^{p-2}\nabla u) = 0\,.$$

The left hand side of this equation is usually called the p-Laplacian.

In [9] it has been shown that $u \mapsto E^u$ is a lower semicontinuous functional on $L^p(\Omega, X)$ for $p \geq 1$. If Ω has finite volume then convergence in $L^2(\Omega, X)$ implies convergence in $L^p(\Omega, X)$ for $1 \leq p \leq 2$. Hence the p-Sobolev energy functional is lower semicontinuous for $p \leq 2$ if considered as a functional on $L^2(\Omega, X)$. It now needs to be shown that the p-Sobolev energy is convex. To this end one needs another result from the theory of NPC spaces which is described below.

Let (X, d) be an NPC space and pick four points $\{q_0, r_0, r_1, q_1\}$ in X. Let $q_t = (1-t)q_0 + tq_1$ and $r_t = (1-t)r_0 + tr_1$, then using a subembedding into $\mathbf{R}^2$ (cf. [9, 12]) one can show

$$d(q_t, r_t) \leq (1-t)d(q_0, r_0) + td(q_1, r_1)\,.$$

Using the subembedding result together with the convexity of the real valued function $x \mapsto x^p$ one obtains

$$d^p(q_t, r_t) \leq (1-t)d^p(q_0, r_0) + td^p(q_1, r_1)\,.$$

Setting $q_0 = u_0(x)$, $q_1 = u_1(x)$, $r_0 = u_0(y)$, $r_1 = u_1(y)$, multiplying by $(n+p)\epsilon^{-(n+p)}f$ with $f \in C_c(\Omega)$ and $f \geq 0$, and integrating over the set $\{(x,y) \in \Omega \times \Omega : d(x,y) \leq \epsilon\}$, one obtains in the notation of Section 3.1

$$E_\epsilon^{u_t}(f) \leq (1-t)E_\epsilon^{u_0}(f) + tE_\epsilon^{u_1}(f)\,.$$

Now let $\epsilon \to 0$ and take the limsup on both sides. As $f \in C_c(\Omega)$ was arbitrary one obtains the following inequality for the energy measures

$$de^{u_t} \leq (1-t)\, de^{u_0} + t\, de^{u_1}\,.$$

This shows first of all $u_t \in W^{1,p}(\Omega, X)$ if both $u_0, u_1 \in W^{1,p}(\Omega, X)$, and $u_t \in BV(\Omega, X)$ if both $u_0, u_1 \in BV(\Omega, X)$. It also shows the desired convexity.

Lemma 3.6. *The p-Sobolev energy functional is convex on $BV(\Omega, X)$ for $p = 1$, and on $W^{1,p}(\Omega, X)$ for $1 \leq p < \infty$.*

Hence the general theory can be applied. The constant maps are clearly minimizers for E. Notice that $E^u < \infty$ is equivalent to $u \in W^{1,p}(\Omega, X)$ provided $p > 1$, and to $u \in BV(\Omega, X)$ if $p = 1$.

Theorem 3.7. *Let* $1 < p \leq 2$ *and assume* Ω *has finite volume. For any starting point* $u_0 \in W^{1,p}(\Omega, X) \cap L^2(\Omega, X)$ *the* $L^2(\Omega, X)$*-gradient flow for the p-Sobolev energy exists in the sense of Theorem* 1.13 *with* $u(t) \in W^{1,p}(\Omega, X) \cap L^2(\Omega, X)$ *for* $t \geq 0$*, and the flow stays bounded for all times. If* $p = 1$ *then* $W^{1,p}(\Omega, X)$ *needs to be replaced by* $BV(\Omega, X)$.

For a given element $\phi \in W^{1,p}(\Omega, X)$ and $\partial\Omega \neq \{\}$ one has a well-defined trace map $\operatorname{tr}\phi$ provided Ω is a Lipschitz domain, see [9]. This allows to consider the boundary value problem by prescribing that $u(t)$ is to have the same boundary values as ϕ.

Definition 7.

$$W^{1,p}_{\phi}(\Omega, X) = \{u \in W^{1,p}(\Omega, X) \,:\, \operatorname{tr}\phi = \operatorname{tr} u \text{ and } E^u \leq E^{\phi}\}\,.$$

As for the case $p = 2$ one can show that $W^{1,p}_{\phi}(\Omega, X) \cap L^2(\Omega, X)$ is an NPC space if endowed with the $L^2(\Omega, X)$ metric, $1 < p \leq 2$. In case Ω is not a Lipschitz domain one can also adapt the more general definition outlined in the section on the $p = 2$ case.

Theorem 3.8. (Solvability of the Initial Boundary Value Problem) *Let* $1 < p \leq 2$ *and assume* Ω *has finite volume. For any map* $\phi \in W^{1,p}(\Omega, X) \cap L^2(\Omega, X)$ *the following problem admits a solution in the sense of Theorem* 1.13*:*

$$\begin{cases} u(t) \textit{ is an } L^2(\Omega, X)\textit{-gradient flow for the p-Sobolev energy for } t \geq 0\,, \\ u(0) = \phi\,, \\ \operatorname{tr} u(t) = \operatorname{tr}\phi \textit{ for } t \geq 0\,. \end{cases}$$

Remark. The equivariant map flow problem also can be generalized to the case $1 \leq p < 2$. The details are omitted.

References.

[1] H. Brezis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, vol. 5 of Mathematics Studies, North-Holland, 1973.

[2] M. G. Crandall and T. M. Liggett, *Generation of semigroups of nonlinear transformations on general Banach spaces*, American J. Math., 93 (1971), pp. 265–298.

[3] M. G. CRANDALL, A. PAZY, AND L. TARTAR, *Remarks on generators on analytic semigroups*, Israel J. Math., 32 (1979), pp. 363–374.

[4] H. FEDERER, *Geometric Measure Theory*, vol. 153 of Grundlehren der mathematischen Wissenschaften, Springer-Verlag, 1969.

[5] K. GOEBEL AND S. REICH, *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*, vol. 83 of Pure and Applied Mathematics, Marcel Dekker, 1984.

[6] J. JOST, *Equilibrium maps between metric spaces*, Calc. Var. Partial Differential Equations, 2 (1994), pp. 173–204.

[7] J. JOST, *Convex functionals and generalized harmonic maps into spaces of nonpositive curvature*, Comment. Math. Helv., 70 (1995), pp. 659–673.

[8] G. KATRIEL, *Mountain pass theorems and global homeomorphism theorems*, Ann. Inst. Henri Poincaré, Analyse non linéaire, 11 (1994), pp. 189–209.

[9] N. J. KOREVAAR AND R. M. SCHOEN, *Sobolev spaces and harmonic maps for metric space targets*, Comm. Anal. Geom., 1 (1993), pp. 561–659.

[10] N. J. KOREVAAR AND R. M. SCHOEN, *Global existence theorems for harmonic maps to non-locally compact spaces*, Comm. Anal. Geom., 5 (1997), pp. 333–387.

[11] U. F. MAYER, *Gradient flows on nonpositively curved metric spaces*, Ph.D. thesis, University of Utah (1995).

[12] Y. G. RESHETNYAK, *Nonexpanding maps in a space of curvature no greater than K*, Siberian Math. J., 9 (1968), pp. 918–927.

[13] W. RUDIN, *Real and Complex Analysis*, McGraw-Hill, 1987.

[14] T. SERBINOWSKI, *Boundary regularity of energy minimizing maps*, Comm. Anal. Geom., 2 (1994), pp. 139–153.

VANDERBILT UNIVERSITY, U. S. A.

RECEIVED JANUARY 22, 1996.