NOTES ON LINEAR SEMIGROUPS AND GRADIENT FLOWS

F. MAGGI

These notes have been written in occasion of the course *Partial Differential Equations II* held by the author at the University of Texas at Austin. They are mostly based on Evans's *PDE* book, and on Brezis' *Functional Analysis* book. With respect to these classical references, the author has tried to supply additional insights and explanations meant to facilitate the student's understanding.

1. Linear semigroups

The idea explored here is the possibility of seeing the Cauchy problem for a linear PDE like the heat equation,

$$\begin{cases} u_t - \Delta u = 0 & \text{on } \Omega \times (0,\infty)\,, \\ u|_{t=0} = u_0 & \text{on } \Omega\,, \\ u|_{\partial\Omega} = 0 & \forall t \ge 0\,. \end{cases} \tag{1.1}$$

as a linear ODE on some infinite dimensional normed vector space X,

$$\begin{cases} U'(t) = AU(t)\,, & \forall t \ge 0\,, \\ U(0) = U_0\,. \end{cases} \tag{1.2}$$

In this ODE, A denotes a linear operator from X to X, and we have a continuous function $U : [0,T] \to X$ such that the limit in X of the incremental ratios $h^{-1}(U(t+h) - U(t))$ as $h \to 0$ exists for every $t \ge 0$, and is denoted by $U'(t)$.

The interpretation of (2.10) as an ODE is obtained by identifying $u(x,t)$ with $U(t) = u(\cdot,t) \in X$, where X is a space of functions of the space variable $x \in \mathbb{R}^n$. The linear operator A will of course consist in taking the spatial Laplacian of $U \in X$. We shall set $X = L^2(\Omega)$. The domain of $A = \Delta$ will then be the subspace $D(A) = H^1_0(\Omega) \cap H^2(\Omega)$ of X. So typically A will not be defined on the whole X, but just on a dense subspace, and A will be unbounded in the norm of X.

1.1. **Linear ODE in Banach spaces.**

Let us consider the general linear ODE (1.2) in a given normed vector space X, and with $A \in \mathcal{L}(X,X)$, that is A is a linear bounded operator from X to X. If X is a Banach space then we can solve (1.2) by the exponential construction.

Theorem 1.1. *If X is a Banach space and $A \in \mathcal{L}(X,X)$, then for every $t \in \mathbb{R}$ the limit*

$$\lim_{N\to\infty} \mathrm{Id} + \sum_{k=1}^{N} \frac{t^k A^k}{k!}$$

exists in $\mathcal{L}(X;X)$ and is denoted by

$$e^{t\,A}\,.$$

It commutes with A, it is equal to Id *for $t = 0$, and has the property that $U(t) = e^{t\,A}U_0$ solves* (1.2)*, that is*

$$\frac{d}{dt} e^{t\,A}U = A\, e^{t\,A}U = e^{t\,A}AU \qquad \forall U \in X\,. \tag{1.3}$$

Proof. Fix $t \in \mathbb{R}$, and denote by $S_N(t)$ the sequence of operators in the statement. For $N < M$ we have, in the norm of $\mathcal{L}(X,X)$,

$$\|S_N(t) - S_M(t)\| \le \sum_{k=N}^{M} \frac{|t|^k \, \|A\|^k}{k!} \to 0 \qquad \text{as } N, M \to \infty\,.$$

If X is a Banach space then so is $\mathcal{L}(X,X)$, thus the limit $S(t) = e^{t\,A}$ exists in $\mathcal{L}(X,X)$. Given $U_0 \in X$, $t \mapsto E_N(t)U_0$ is a smooth function, and as $N \to \infty$, $E_N(t)U_0 \to S(t)U_0$ locally uniformly in t. The derivatives of every order of $E_N(t)U_0$ also converge locally uniformly, so that $S(t)U_0$ is smooth in t, and its derivatives are easily computed. For example

$$\frac{d}{dt}E_N(t) = \sum_{k=1}^{N} \frac{t^{k-1}A^k}{(k-1)!} = AE_{N-1}(t)$$

implies that $(d/dt)S(t)U_0 = AS(t)U_0$, as claimed. □

1.2. **Semigroups arise from generators.** We single out the following properties of the map $S : [0,\infty) \to \mathcal{L}(X,X)$, $S(t) = e^{t\,A}$, constructed in Theorem 1.1 starting from $A \in \mathcal{L}(X,X)$:

(i) $S(0) = \mathrm{Id}\,$;
(ii) $S(t+s) = S(t)S(s) = S(s)S(t)$ for every $t, s \ge 0$;
(iii) $S(\cdot)U_0$ is continuous on $[0,\infty)$;
(iv) there exists $\omega \in \mathbb{R}$ such that $\|S(t)\| \le e^{\omega\,t}$ for every $t \ge 0\,$.

In general, a family of linear operators $\{S(t)\}_{t\ge0} \subset \mathcal{L}(X,X)$ is an **ω-contractive semigroup on** X is properties (i)–(iv) hold.

One may naively ask if every semigroup arises from an operator $A \in \mathcal{L}(X,X)$ by means of the exponential construction. This is morally true, but it cannot be exactly true, as something like the heat flow on $L^2(\Omega)$ has to arise from an operator like $A = \Delta$, which is defined as a linear map taking values in $L^2(\Omega)$ only if we restrict its domain to $H^2(\Omega)$, which is a proper dense subspace of $L^2(\Omega)$. But with this *caveat* in mind, namely that the generating operator A does not need to be bounded on X, and may be actually defined only on a dense subspace of X, it is essentially true that *every semigroup is generated by a linear operator defined on a dense subspace.*

Theorem 1.2. *Let $\{S(t)\}_{t\ge0}$ be an ω-contractive semigroup. Consider the linear subspace D of X and the linear operator $A : D \to X$ defined by*

$$\begin{aligned} D &= \left\{ U \in X : \text{the limit } \lim_{t\to0^+} \tfrac{S(t)U-U}{t} \text{ exists in } X \right\} \\ AU &= \lim_{t\to0^+} \frac{S(t)U-U}{t} \qquad U \in D\,, \end{aligned}$$

and set $D = D(A)$. Then

(i) *$D(A)$ is dense in X and A is closed (that is, $\{u_k\}_{k\in\mathbb{N}} \subset D$, $u_k \to u$ and $Au_k \to v$ imply $u \in D(A)$ and $v = Au$);*
(ii) *for every $U \in D(A)$ and $t > 0$, $S(t)U \in D(A)$, and $t \mapsto S(t)U$ is differentiable on $(0,\infty)$, with*

$$\frac{d}{dt}\,S(t)U = AS(t)U = S(t)AU\,. \tag{1.4}$$

Proof. Proof of (ii): We show that $S(t)U \in D(A)$ for each $U \in D(A)$ and $t > 0$. Indeed

$$\frac{S(h)S(t)U - S(t)U}{h} = S(t)\,\frac{S(h)U-U}{h} \to S(t)AU \qquad \text{as } h \to 0^+$$

as $(1/h)(S(h)U - U) \to AU$ for $h \to 0^+$, and since $S(t) \in \mathcal{L}(X, X)$. This shows that

$$S(t)U \in D(A) \text{ with } AS(t)U = S(t)AU.$$

Concerning the differentiability of $S(t)U$ in t for $t > 0$, we notice that

$$\frac{S(t+h)U - S(t)U}{h} = S(t)\,\frac{S(h)U - U}{h} \to S(t)\,AU \qquad \text{when } h \to 0^+ .$$

To compute the limit as $h \to 0^-$ set $k = -h$, then

$$\begin{aligned} \frac{S(t+h)U - S(t)U}{h} &= \frac{S(t)U - S(t-k)U}{k} = S(t-k)\frac{S(k)U - U}{k} \\ &= S(t-k)\Big(\frac{S(k)U - U}{k} - AU\Big) + S(t-k)AU . \end{aligned}$$

Considering that $\|S(t-k)\| \le e^{\omega t}$ for every $k > 0$ and that $S(t-k) \to S(t)$ in $\mathcal{L}(X, X)$ as $k \to 0$, we find that as $h \to 0^-$

$$\lim_{h\to 0^-} \frac{S(t+h)U - S(t)U}{h} = S(t)AU ,$$

that is (1.4).

Proof of (i): Let $U \in X$, and set

$$U^{(t)} = \frac{1}{t}\int_0^t S(s)U\,ds , \qquad t > 0 ,$$

so that, as $t \to 0^+$, by continuity of $\|S(s)U - U\|$ for $s \in [0, \infty)$ and by $S(0) = \mathrm{Id}$,

$$\|U^{(t)} - U\| \le \frac{1}{t}\int_0^t \|S(s)U - U\|\,ds \to 0 .$$

We claim that $U^{(t)} \in D(A)$: indeed,

$$\begin{aligned} \frac{S(h)U^{(t)} - U^{(t)}}{h} &= \frac{1}{t\,h}\Big\{\int_0^t S(h+s)U\,ds - \int_0^t S(s)U\,ds\Big\} \\ &= \frac{1}{t\,h}\Big\{\int_h^{t+h} S(s)U\,ds - \int_0^t S(s)U\,ds\Big\} \\ &= \frac{1}{t\,h}\Big\{\int_t^{t+h} S(s)U\,ds - \int_0^h S(s)U\,ds\Big\} \end{aligned}$$

so that as $h \to 0^+$,

$$\frac{S(h)U^{(t)} - U^{(t)}}{h} \to \frac{S(t)U - U}{t} .$$

This shows that $U^{(t)} \in D(A)$, thus that $D(A)$ is dense in X.

Now consider $\{u_k\}_{k\in\mathbb{N}} \subset D(A)$ with $u_k \to u$ and $Au_k \to v$ in X. Integrating (1.4) we find that

$$S(h)u_k - u_k = \int_0^h A\,S(s)\,u_k\,ds = \int_0^h S(s)\,A\,u_k\,ds .$$

Now $S(s)Au_k \to S(s)v$ uniformly as $k \to \infty$, while $S(h)u_k \to S(h)u$, so that we find

$$\frac{S(h)u - u}{h} = \frac{1}{h}\int_0^h S(s)\,v\,ds .$$

The limit as $h \to 0^+$ of the left-hand side exists, and it is equal to $S(0)v = v$, that is $u \in D(A)$ and $Au = v$. This proves that A is a closed operator. $\square$

1.3. **From generators to semigroups, general strategy.** We now discuss the converse construction to the one of Theorem 1.2, namely, given a closed operator A densely defined on X we wish to construct a semigroup having A as its generator, and thus solving the ODE (1.2). Notice that when $D(A) = X$, closedness implies boundedness, i.e. $A \in \mathcal{L}(X, X)$, and thus the exponential construction of Theorem 1.1 applies. Thus we are really concerned with the case when $D(A)$ is a proper dense subspace of X.

The idea is approximating A with operators $A_\lambda \in \mathcal{L}(X, X)$ and considering the semigroups $S_\lambda(t) = e^{t A_\lambda}$ defined from A_λ by means of the exponential construction (Theorem 1.1). Then the semigroup S generated by A will be obtained as the limit of the approximating semigroups S_λ.

The construction of the approximating operators will be as follows. Having in mind the case when $A = \Delta$ and $D(A) = H_0^1(\Omega) \cap H^2(\Omega)$, we expect that for $\mu > 0$ sufficiently small, the linear operator $\mathrm{Id} - \mu A$ will be bounded and invertible from $D(A)$ to X. Indeed there exists no $u \in H_0^1(\Omega) \setminus \{0\}$ with $u = \mu \Delta u$ and $\mu > 0$, as a reflection of the fact that $-\Delta$ is positive definite on $H_0^1(\Omega)$ (Poincaré inequality). Thus, for μ small, we shall be able to apply the exponential construction to

$$(\mathrm{Id}\, - \mu A)^{-1} \in \mathcal{L}(X, X)\,. \tag{1.5}$$

An even better idea is applying the exponential construction to

$$\frac{(\mathrm{Id}\, - \mu A)^{-1} - \mathrm{Id}}{\mu} \in \mathcal{L}(X, X)\,. \tag{1.6}$$

Indeed a formal expansion in $\mu \to 0^+$ gives

$$\frac{(\mathrm{Id}\, - \mu A)^{-1} - \mathrm{Id}}{\mu} = A + O(\mu)$$

so that we expect $(1/\mu)\,[(\mathrm{Id}\, - \mu A)^{-1} - \mathrm{Id}\,]$ to converge to A as $\mu \to 0^+$. The induced exponential semigroups will then converge to a limit semigroup $S(t)$ having A as its generator.

This strategy works and leads to the Hille-Yosida theorem. The statement uses the following terminology. If A is a linear operator defined from some subspace $D(A)$ of X with values in X, we denote by

$$\rho(A) = \Big\{\lambda \in \mathbb{R} : (\lambda\,\mathrm{Id}\, - A) \text{ is injective and surjective from } D(A) \text{ to } X\Big\}$$

the **resolvent** of A. For $\lambda \in \rho(A)$, the **resolvent operator** R_λ is defined as

$$R_\lambda = (\lambda\,\mathrm{Id}\, - A)^{-1} \in \mathcal{L}(X, X)\,. \tag{1.7}$$

Notice that (1.5) corresponds to (1.7) with $\mu = 1/\lambda$. In particular the limit as $\mu \to 0^+$ will become a limit as $\lambda \to +\infty$, and the approximating operators defined in (1.6) will take the form

$$A_\lambda = -\lambda\,\mathrm{Id}\, + \lambda^2 R_\lambda\,. \tag{1.8}$$

A formal expansion in $\lambda \to +\infty$ gives of course $A_\lambda = A + O(1/\lambda)$. The operators A_λ are called the **Hille-Yosida approximating operators** of A.

For example, let us consider the above definitions in the case of the Laplacian with homogeneous Dirichlet condition on a bounded open set Ω, that is, when $A = \Delta$, $D(A) = H_0^1(\Omega) \cap H^2(\Omega)$, and $X = L^2(\Omega)$. Let $0 < \lambda_k(\Omega) < \lambda_{k+1}(\Omega) \to \infty$ denote the sequence of the eigenvalues of $-\Delta$ on $H_0^1(\Omega)$. If $\lambda \in \mathbb{R}$ is such that $(\lambda\,\mathrm{Id}\, - A)$ is injective, then there cannot be $u \in H_0^1(\Omega) \setminus \{0\}$ such that $\Delta u = \lambda u$, that is $-\lambda \notin \{\lambda_k(\Omega) : k \geq 1\}$. Thus if

$\lambda > -\lambda_1(\Omega)$ we definitely have that $(\lambda\,\mathrm{Id} - A)^{-1}$ is injective. About surjectivity, pick any $f \in L^2(\Omega)$ and consider the minimization of

$$E(u) = \int_\Omega |\nabla u|^2 + \lambda\, u^2 - f u$$

on $H^1_0(\Omega)$. By the Poincaré inequality, for every $\varepsilon > 0$,

$$\int_\Omega |\nabla u|^2 + \lambda\, u^2 - f u \geq \int_\Omega \varepsilon |\nabla u|^2 + \big(\lambda_1(\Omega) - \varepsilon + \lambda\big) u^2 - \frac{f^2}{\varepsilon} - \varepsilon\, u^2$$

so that if ε is small enough with respect to $\lambda_1(\Omega) + \lambda > 0$, the energy $E(u)$ is coercive on $H^1_0(\Omega)$ and there exists $u \in H^1_0(\Omega)$ such that $(\lambda\,\mathrm{Id} - A)^{-1} f = u$. By the incremental ratios method we immediately see that $u \in H^2_{\mathrm{loc}}(\Omega)$, and if Ω has Lipschitz boundary, that $u \in H^2(\Omega)$. This shows that for a bounded open set Ω with Lipschitz boundary one always has

$$(-\lambda_1(\Omega), \infty) \subset \rho(\Delta)\,.$$

In particular, it definitely makes sense to consider the Hille-Yosida approximating operators $(\Delta)_\lambda = -\lambda \mathrm{Id} + \lambda^2(\lambda \mathrm{Id} - \Delta)^{-1}$ in the limit as $\lambda \to \infty$. It will turn out that for every $\lambda > -\lambda_1(\Omega)$, $(\Delta)_\lambda$ is indeed an element of $\mathcal{L}(L^2(\Omega), L^2(\Omega))$, and that its exponential map converge to the heat flow. All these facts can be proved in an abstract framework according to the Hille-Yosida theorem.

Theorem 1.3 (Hille-Yosida theorem)**.** *Let X be a Banach space and let $\omega \in \mathbb{R}$.*

Part one*: If $\{S(t)\}_{t\geq 0}$ is an ω-contractive semigroup, and A is the generator of $S(t)$, then*

$$(\omega, \infty) \subset \rho(A)\,, \qquad \|R_\lambda\| \leq \frac{1}{\lambda - \omega} \qquad \forall \lambda > \omega\,,$$

and R_λ is the **Laplace transform of** *$S(t)$,*

$$R_\lambda U = \int_0^\infty e^{-\lambda\, t} S(t)\, U\, dt\,, \qquad \forall U \in X\,.$$

Part two*: If A is a closed operator defined on a dense subspace $D(A)$ of X such that*

$$(\omega, \infty) \subset \rho(A)\,, \qquad \|R_\lambda\| \leq \frac{1}{\lambda - \omega} \qquad \forall \lambda > \omega\,, \tag{1.9}$$

then the Hille-Yosida approximating operators $A_\lambda \in \mathcal{L}(X, X)$ defined by

$$A_\lambda = -\lambda\,\mathrm{Id} + \lambda^2\, R_\lambda \qquad \lambda > \omega\,,$$

are such that: (i) for every $U \in D(A)$, $A_\lambda U \to AU$ as $\lambda \to \infty$; (ii) $\{S_\lambda(t) = e^{t\,A_\lambda}\}_{t\geq 0}$ is a semigroup with

$$\|S_\lambda\| \leq \frac{\omega\lambda}{\lambda - \omega}\,, \qquad \forall \lambda > \omega$$

(iii) the limit $S(t)$ of $S_\lambda(t)$ as $\lambda \to \infty$ exists and defines an ω-contractive semigroup whose generator is A.

Solving the heat equation To exemplify let us check how Theorem 1.3 can be applied to construct solutions to the heat equation. Set $X = L^2(\Omega)$, $A = \Delta$, $D(A) = H^1_0(\Omega) \cap H^2(\Omega)$. If $u_k \to u$ and $\Delta u_k \to v$ in $L^2(\Omega)$ for some $\{u_k\}_{k\in\mathbb{N}} \subset H^1_0(\Omega) \cap H^2(\Omega)$, then

$$\int_\Omega v\, \eta = \lim_k \int_\Omega \eta \Delta u_k = \lim_k \int_\Omega u_k \Delta \eta = \int_\Omega u \Delta \eta \qquad \forall \eta \in C^\infty_c(\Omega)\,.$$

By the L^2-regularity theory for Δ we also have that

$$\|\Delta u_k\|_{L^2(\Omega)} \geq \|\nabla^2 u_k\|_{L^2(\Omega)}$$

so that $u \in H^2(\Omega)$ and $u_k \rightharpoonup u$ in $H^2(\Omega)$ with $\Delta u = v$. Thus Δ is closed in $L^2(\Omega)$. We have already seen the argument showing that

$$(-\lambda_1(\Omega), \infty) \subset \rho(\Delta)\,.$$

Let us thus consider the map $R_\lambda = (\lambda \mathrm{Id} - \Delta)^{-1}$. For $f \in L^2(\Omega)$ set $u = R_\lambda f$, so that

$$\int_\Omega \nabla u \cdot \nabla \eta - \lambda u \eta = \int_\Omega f \eta \qquad \forall \eta \in H_0^1(\Omega)\,.$$

Testing with $\eta = u$ gives

$$\|f\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} \geq \int_\Omega f u = \int_\Omega |\nabla u|^2 + \lambda\, u^2 \geq (\lambda_1(\Omega) + \lambda) \|u\|^2_{L^2(\Omega)}$$

that is exactly

$$\|R_\lambda f\|_{L^2(\Omega)} = \|u\|_{L^2(\Omega)} \leq \frac{\|f\|_{L^2(\Omega)}}{\lambda + \lambda_1(\Omega)}\,,$$

as required. Thus the Hille-Yosida theorem provides a weak solution to the weak equation in the form of a $(-\lambda_1(\Omega))$-contractive semigroup on $L^2(\Omega)$. The solution $u(t)$ satisfies the exponential convergence-to-equilibrium estimate

$$\int_\Omega u(t)^2 \leq e^{-\lambda_1(\Omega) t} \int_\Omega u_0^2\,.$$

Solving the wave equation: Given $u_0 \in H_0^1(\Omega)$ and $v_0 \in L^2(\Omega)$ we want to solve $u_{tt} - \Delta u = 0$ in Ω with $u = 0$ on $\partial\Omega$ and $u|_{t=0} = u_0$, $(u_t)|_{t=0} = v_0$. To this end consider the space

$$X = H_0^1(\Omega) \times L^2(\Omega)$$

and denote by $U = (u, v)$ the generic element of X, endowed with the norm

$$\|U\| = \|u\|_{H_0^1(\Omega)} + \|v\|_{L^2(\Omega)}\,.$$

Next define the operator $AU = (v, \Delta u)$ with domain

$$D(A) = (H^2(\Omega) \cap H_0^1(\Omega)) \times H^1(\Omega)\,,$$

which is dense in X. Solving the ODE $U'(t) = AU(t)$ with $U(0) = (u_0, v_0)$ amounts in providing a weak solution to the wave equation.

Let $\lambda > 0$. For every $F = (f, g) \in X$, solving $\lambda U - AU = F$ means that

$$\lambda u - v = f \qquad \lambda v - \Delta u = g$$

so that $-\Delta u = g - \lambda v = g + \lambda f - \lambda^2 u$. Since $(\lambda^2 \mathrm{Id} - \Delta)$ is injective and surjective from $H_0^1(\Omega) \cap H^2(\Omega)$ to $L^2(\Omega)$, we find that there exists a unique u such that $-\Delta u = g + \lambda f - \lambda^2 u$, and then by setting $v = \lambda u - f$ we have found that for every $F \in X$ there exists a unique $U \in D(A)$ such that $\lambda U - AU = F$. Set $U = R_\lambda F$. Notice that by $\lambda v - \Delta u = g$, multiplying by v we find

$$\lambda v^2 - v \Delta u = g v$$

and then inserting $v = \lambda u - f$

$$\lambda\big(v^2 - u \Delta u\big) + f \Delta u = g v\,.$$

Integrating

$$\lambda \int_\Omega v^2 + |\nabla u|^2 = \int_\Omega g v + \int_\Omega \nabla f \cdot \nabla u \leq \big(\|g\|_{L^2(\Omega)} + \|f\|_{H_0^1(\Omega)}\big)\Big(\|v\|_{L^2(\Omega)} + \|\nabla u\|_{L^2(\Omega)}\Big)$$

that is

$$\|R_\lambda F\| = \|U\| = \|v\|_{L^2(\Omega)} + \|\nabla u\|_{L^2(\Omega)} \leq \frac{\|g\|_{L^2(\Omega)} + \|f\|_{H_0^1(\Omega)}}{\lambda} = \frac{\|F\|}{\lambda}\,.$$

Thus A generates a contractive semigroup on X, which provides a notion of weak solution for the wave equation. Notice that the estimate $\|S(t)U\| \le \|U_0\|$ now means that

$$\|\nabla u(t)\|_{L^2(\Omega)} + \|u_t(t)\|_{L^2(\Omega)} \le \|\nabla u_0\|_{L^2(\Omega)} + \|v_0\|_{L^2(\Omega)}\,,$$

where we actually know (by differentiating the total energy of the wave) that this quantity is conserved!

1.4. **Proving part one of the Hille-Yosida theorem.** It is a simple exercise to prove the following lemma.

Lemma 1.4. *Let A be a closed operator defined on a dense subset $D(A)$ of a Banach space X. Then for every $\lambda > 0$, R_λ is a closed operator, and thus belongs to $\mathcal{L}(X, X)$. Moreover, for every $\lambda, \mu \in \rho(A)$,*

$$A\,R_\lambda = R_\lambda\,A \qquad \text{on } D(A)\,, \tag{1.10}$$

$$R_\lambda\,R_\mu = R_\mu\,R_\lambda \qquad \text{on } X\,, \tag{1.11}$$

$$R_\lambda - R_\mu = (\mu-\lambda)\,R_\lambda\,R_\mu \qquad \text{on } X\,. \tag{1.12}$$

Proof. If $u_k \to u$ and $R_\lambda u_k \to v$ in X, then $\lambda\,u_k - Au_k \to v$ implies $Au_k \to \lambda u - v$. Since A is closed, $u \in D(A)$ and $Au = \lambda\,u - v$, that is $v = R_\lambda u$. Thus R_λ is closed, and by the closed graph theorem, $R_\lambda \in \mathcal{L}(X, X)$. To check (1.10) just notice that

$$\mathrm{Id} = (\lambda \mathrm{Id} - A)R_\lambda = \lambda\,R_\lambda - AR_\lambda \qquad \mathrm{Id} = R_\lambda(\lambda \mathrm{Id} - A) = \lambda\,R_\lambda - R_\lambda A$$

so that $A\,R_\lambda = R_\lambda\,A$ on $D(A)$. Next, by (1.10) we find

$$\begin{aligned} R_\lambda &= R_\lambda R_\mu(\mu \mathrm{Id} - A) = R_\lambda(\mu \mathrm{Id} - A)R_\mu \\ &= R_\lambda(\lambda \mathrm{Id} - A)R_\mu + R_\lambda(\mu-\lambda)\mathrm{Id}\; R_\mu \\ &= R_\mu + (\mu-\lambda)R_\lambda\,R_\mu\,, \end{aligned}$$

that is (1.12). As a consequence we find (1.11)

$$R_\lambda R_\mu = \frac{R_\lambda - R_\mu}{\mu - \lambda} = \frac{R_\mu - R_\lambda}{\lambda - \mu} = R_\mu R_\lambda\,.$$

□

Proof of Theorem 1.3, Part one. For $U \in X$ and $\lambda > \omega$ let us set

$$R^*_\lambda U = \int_0^\infty e^{-\lambda\,t} S(t)U\,dt\,.$$

Notice that the integral is convergent as $\|e^{-\lambda\,t}S(t)U\| \le e^{(\omega-\lambda)\,t}\|U\|$ and $\lambda > \omega$. In particular,

$$\|R^*_\lambda\| \le \frac{1}{l-\omega}\,.$$

We shall now prove that $R^*_\lambda = R_\lambda$ on X. First of all, for $U \in X$ and $\lambda > \omega$ we have

$$\begin{aligned} \frac{S(h)R^*_\lambda U - R^*_\lambda U}{h} &= \frac{1}{h}\Big\{\int_0^\infty e^{-\lambda\,t}S(t+h)U\,dt - \int_0^\infty e^{-\lambda\,t}S(t)U\,dt\Big\} \\ &= \frac{1}{h}\Big\{\int_h^\infty e^{-\lambda\,t+\lambda\,h}S(t)U\,dt - \int_0^\infty e^{-\lambda\,t}S(t)U\,dt\Big\} \\ &= -\frac{1}{h}\int_0^h e^{-\lambda\,t+\lambda\,h}S(t)U\,dt + \frac{e^{\lambda h}-1}{h}\int_0^\infty e^{-\lambda t}S(t)U\,dt \end{aligned}$$

which as $h \to 0^+$, converges to

$$AR^*_\lambda U = -S(0)U + \lambda\int_0^\infty e^{-\lambda t}S(t)U\,dt = -U + \lambda\,R^*_\lambda U\,.$$

This identity means that

$$(\lambda\,\mathrm{Id}\, - A)R^*_\lambda U = U \qquad \forall U \in X\,.$$

On the other hand $AR^*_\lambda = R^*_\lambda A$ on $D(A)$ as a consequence of $AS(t) = S(t)A$ on $D(A)$, so that we also have

$$R^*_\lambda(\lambda\,\mathrm{Id}\, - A)U = U \qquad \forall U \in D(A)\,.$$

These two identities imply that $(\lambda\,\mathrm{Id}\, - A)$ is injective and surjective from $D(A)$ to X and that $R^*_\lambda = (\lambda\,\mathrm{Id}\, - A)^{-1}$ for every $\lambda > \omega$. □

1.5. **Proving part two of the Hille-Yosida theorem.** We start proving (i). We first show that as $\lambda \to \infty$

$$\lambda\, R_\lambda\, U \to U \qquad \forall U \in X\,. \tag{1.13}$$

Indeed, let us first pick $U \in D(A)$ and set $V = \lambda\, R_\lambda\, U$ for $\lambda > \omega$ Then $\lambda\, U = \lambda V - AV$ so that (1.10) and (1.9) imply

$$\|U - \lambda R_\lambda U\| = \|U - V\| \le \frac{\|AV\|}{\lambda} = \|AR_\lambda U\| = \|R_\lambda\, AU\| \le \|R_\lambda\|\,\|AU\| \le \frac{\|AU\|}{\lambda - \omega}\,,$$

where $\|AU\| < \infty$. In the general case when $U \in X$ we just pick $U_\delta \in D(A)$ such that $\|U - U_\delta\| < \delta$ and notice that

$$\|R_\lambda U - U\| \le \|R_\lambda U_\delta - U_\delta\| + \Big(1 + \frac{1}{\lambda - \omega}\Big)\,\|U - U_\delta\|$$

so that (1.13) follows. By (1.13) and recalling that $\mathrm{Id}\, = (\lambda \mathrm{Id}\, - A)R_\lambda = \lambda\, R_\lambda - R_\lambda A$, we deduce that for every $U \in D(A)$

$$A_\lambda U = -\lambda U + \lambda^2 R_\lambda U = \lambda\, R_\lambda\, AU \to AU\,.$$

This proves assertion (i).

Since $R_\lambda \in \mathcal{L}(X,X)$ we have $A_\lambda \in \mathcal{L}(X,X)$ and thus $S_\lambda(t) = e^{t\,A_\lambda}$ defines a contraction semigroup thanks to Theorem 1.2. To estimate $\|S_\lambda\|$, let us notice without proof that if $L, M \in \mathcal{L}(X,X)$ and $LM = ML$ then

$$e^{(L+M)t} = e^{Lt}e^{Mt} = e^{Mt}e^{Lt} \qquad \text{on } X\,.$$

Hence,

$$e^{t\,A_\lambda} = e^{-\lambda\,t}\,e^{t\,\lambda^2\,R_\lambda}$$

so that

$$\begin{aligned}
\|e^{t\,A_\lambda}U_0\| &\le e^{-\lambda t}\sum_{k=1}^{\infty}\frac{t^k\,\lambda^{2k}}{k!}\,\|R_\lambda\|^k\|U_0\| \\
&\le e^{-\lambda t}\,\|U_0\|\sum_{k=1}^{\infty}\frac{t^k\,\lambda^{2k}}{(\lambda-\omega)^k\,k!} \\
&= e^{-\lambda t}\,e^{\lambda^2\,t/(\lambda-\omega)}\,\|U_0\| \le e^{\omega\lambda/(\lambda-\omega)\,t}\,\|U_0\|
\end{aligned}$$

This proves (ii).

We now prove that for every $t \ge 0$ and $U \in D(A)$, $\{S_\lambda(t)U\}_{\lambda>\omega}$ has the Cauchy property as $\lambda \to \infty$. Indeed, let $t > 0$, $U \in D(A)$, and define $\varphi : [0,t] \to X$ by setting

$$\varphi(s) = e^{(t-s)A_\lambda}e^{s\,A_\mu}U \qquad 0 \le s \le t\,.$$

We have

$$\varphi(t) - \varphi(0) = S_\lambda(t)U - S_\mu(t)U \qquad \sup_{0\le s\le t}\|\varphi(s)\| \le e^{\omega t}\,.$$

Moreover, on noticing that $A_\mu A_\lambda = A_\lambda A_\mu$ thanks to (1.11), and by using (1.3) we find

$$\varphi'(s) = e^{(t-s)A_\lambda}e^{s\,A_\mu}(A_\mu U - A_\lambda U) \qquad \forall s \in [0,t]\,,$$

where, as $\lambda \mapsto \lambda\omega/(\lambda-\omega)$ is decreasing,

$$\|e^{(t-s)A_\lambda} e^{s\,A_\mu}\| \le e^{\frac{\lambda\omega}{\lambda-\omega}(t-s)}\, e^{\frac{\mu\omega}{\mu-\omega}s} \le e^{\frac{\lambda\omega}{\lambda-\omega}t}\,, \qquad \forall \mu > \lambda\,.$$

Hence,

$$\|\varphi'(s)\| \le e^{\omega\lambda t/(\lambda-\omega)} \|A_\mu U - A_\lambda U\| \qquad \forall s \in [0,t]\,, \qquad \mu > \lambda > \omega\,.$$

Thus,

$$\|S_\lambda(t)U - S_\mu(t)U\| \le t\, e^{\omega\lambda t/(\lambda-\omega)} \|A_\mu U - A_\lambda U\| \qquad \forall U \in D(A)\,. \tag{1.14}$$

Since $A_\lambda U \to AU$ as $\lambda \to \infty$, we conclude that $\{S_\lambda(t)U\}_{\lambda>\omega}$ is a Cauchy family, and that for every $t \ge 0$ there exists a linear map $S(t) : D(A) \to X$ defined by

$$S(t)U = \lim_{\lambda\to\infty} S_\lambda(t)U\,, \qquad U \in D(A)\,.$$

Sending $\mu \to \infty$ in (1.14) while keeping λ fixed we find

$$\|S(t)U - S_\lambda(t)U\| \le t\, e^{\lambda\omega t/(\lambda-\omega)} \|AU - A_\lambda U\| \qquad \forall U \in D(A)\,. \tag{1.15}$$

By (1.15), $S_\lambda(t)U \to S(t)U$ uniformly over $t \in [0,T]$ as $\lambda \to \infty$, so that $t \mapsto S(t)U$ is continuous on $[0,\infty)$ for $U \in D(A)$. Since $\|S_\mu(t)U\| \le e^{\omega\mu\, t/(\mu-\omega)}\|U\|$ for every $U \in X$, we immediately see that $S(t) \in \mathcal{L}(X,X)$ with $\|S(t)\| \le e^{\omega t}$. From this we easily deduce that $S_\lambda(t)U \to S(t)\,U$ for every $U \in X$, and that $t \mapsto S(t)U$ is continuous on $[0,\infty)$ for every $U \in X$. The properties $S(t+s) = S(t)S(s) = S(s)S(t)$ and $S(0) = \mathrm{Id}$ are easily transferred by pointwise convergence, and so (iii) is proved up to showing that A is the generator of $\{S(t)\}_{t\ge0}$.

To this end, let B denote the generator of $\{S(t)\}_{t\ge0}$, so that B is a closed operator with domain

$$D(B) = \Big\{ U \in X : \text{the limit } BU = \lim_{t\to0^+} \tfrac{S(t)U-U}{t} \text{ exists in } X \Big\}\,.$$

as shown in Theorem 1.2, with

$$(\omega,\infty) \subset \rho(B)\,.$$

Recall that for $\lambda > \omega$

$$S_\lambda(t)U - U = \int_0^t A_\lambda S_\lambda(t)\,U\,dt = \int_0^t S_\lambda(t)\,A_\lambda U\,dt \qquad \forall U \in X\,.$$

In particular, if $U \in D(A)$, then

$$\|S_\lambda(t)\,A_\lambda U - S(t)\,AU\| \le \|S(t)\|\,\|A_\lambda U - AU\| + \|S_\lambda(t) - S(t)\|\,\|AU\|$$

so that

$$S(t)U - U = \int_0^t S(t)\,AU\,dt \qquad \forall U \in D(A)\,.$$

As a consequence if $U \in D(A)$, then

$$U \in D(B) \quad \text{with} \quad BU = \lim_{h\to0^+} \frac{S(h)U - U}{h} = AU$$

that is $B = A$ on $D(A) \subset D(B)$. Now $(\omega,\infty) \subset \rho(A) \cap \rho(B)$ and if $\lambda > \omega$

$$X = (\lambda\,\mathrm{Id} - A)(D(A)) = (\lambda\,\mathrm{Id} - B)(D(A))\,,$$

so that $(\lambda\,\mathrm{Id} - B)$ is injective and surjective from $D(A)$ to X. Since $(\lambda\,\mathrm{Id} - B)$ is also injective and surjective from $D(B)$ to X, it follows that

$$D(A) = D(B)$$

and thus A is the generator of $\{S(t)\}_{t\ge0}$.

2. Nonlinear semigroups as gradient flows

We now consider a class of nonlinear semigroups arising as "gradient flows" of convex energies. Natural examples to keep in mind are nonlinear parabolic equations of the form

$$u_t - \Delta u = f(u) \qquad f : \mathbb{R} \to \mathbb{R} \text{ decreasing}$$

and

$$u_t - \mathrm{div}\,(\nabla L(\nabla u)) = 0 \qquad L : \mathbb{R}^n \to \mathbb{R} \text{ convex}.$$

These will be the gradient flows of the convex energies

$$\int_\Omega |\nabla u|^2 - F(u) \qquad \int_\Omega L(\nabla u),$$

where, say, $F(u) = \int_0^u f(s)\,ds$. We now illustrate (i) the notion of gradient flow of a convex energy in the model case of finite dimension and (ii) in which sense a PDE on $\mathbb{R}^n$ like the heat equation can be seen as a gradient flow of the Dirichlet energy.

2.1. Gradient flows in $\mathbb{R}^n$.

Gradient flows in $\mathbb{R}^n$ are systems of ODE having the form

$$\begin{cases} X'(t) = -\nabla E(X(t)), & t \geq 0, \\ X(0) = X_0, \end{cases} \tag{2.1}$$

for a given scalar function

$$E : \mathbb{R}^n \to \mathbb{R}.$$

Here E represents the total energy of a system expressed with respect to the states variable $(x_1, ..., x_n) = X$, and these variables evolve according to an infinitesimal minimization principle, that is, the systems "sniffs around" for the nearest accessible states of lower energy by moving in the direction $-\nabla E(X(t))$. Whenever ∇E is a Lipschitz map, the Cauchy-Picard theorem implies the existence of a unique solution for every sufficiently small time. A global in time solution exists, for example, if ∇E is globally bounded, or if ∇E is just locally bounded one can prove that the solution never leaves a compact sublevel set of E. The latter is the case for convex gradient flows, see below.

The dynamics of a gradient flow can be quite rich. Every critical point X_0 of E is stationary for the flow, in the sense that $X(t) \equiv X_0$ will be the unique solution corresponding to the initial data X_0. Among critical points, local maxima will be unstable (under small perturbations of the initial data, the solution will flow away); but all local minima will be possible asymptotic states. This last remark illustrate the great importance of understanding local minimizers, and not just global minimizers, in the study of variational problems.

The following simple theorem gives a good idea of the properties of convex gradient flows.

Theorem 2.1. *Let $E : \mathbb{R}^n \to \mathbb{R}$ be a smooth convex function such that*

$$\lim_{|x| \to \infty} E(x) = +\infty. \tag{2.2}$$

Then a solution $X(t)$ of (2.1) exists for every $t \geq 0$. If E is strictly convex, then E has a unique global minimizer X_{min} on $\mathbb{R}^n$, and $X(t) \to X_{min}$ as $t \to \infty$. If in addition E is uniformly convex, that is, if there exists $\lambda > 0$ such that

$$Y \cdot \nabla^2 E(X) Y \geq \lambda |Y|^2 \qquad \forall X, Y \in \mathbb{R}^n, \tag{2.3}$$

then $X(t)$ converges to X_{min} exponentially,

$$|X(t) - X_{min}| \leq e^{-\lambda t} |X_0 - X_{min}|, \qquad \forall t \geq 0. \tag{2.4}$$

Proof. Notice that ∇E is only locally Lipschitz, so that the existence for every time does not follow immediately from the Cauchy-Picard theorem. The flow exists for every time nevertheless, because differentiating $E(X(t))$ in time we find that $E(X(t))$ is decreasing along the flow,

$$\frac{d}{dt}E(X(t)) = X'(t) \cdot \nabla E(X(t)) = -|\nabla E(X(t))|^2 \,. \tag{2.5}$$

In particular, the flow never leaves the set $\{E \leq E(X_0)\}$, which is a bounded set by the coercivity assumption (2.2). One can thus apply the local existence and uniqueness theorem indefinitely and construct a unique solution for every time.

By coercivity and strict convexity of E there exists a unique minimizer X_{min} of E on $\mathbb{R}^n$. We now prove that $X(t) \to X_{min}$ as $t \to \infty$. We first notice that (2.7) implies the **dissipation inequality**

$$\int_0^\infty |\nabla E(X(t))|^2 dt \leq E(X_0) \,. \tag{2.6}$$

Differentiating in turn $|\nabla E(X(t))|^2$, and using the second order characterization of convexity

$$Y \cdot \nabla^2 E(X) Y \geq 0 \qquad \forall X, Y \in \mathbb{R}^n \,,$$

we deduce that

$$\begin{aligned} \frac{d}{dt}|\nabla E(X(t))|^2 &= 2\,\nabla E(X(t)) \cdot \nabla^2 E(X(t))[X'(t)] && (2.7) \\ &= -2\,\nabla E(X(t)) \cdot \nabla^2 E(X(t))[\nabla E(X(t))] \leq 0 \,. && (2.8) \end{aligned}$$

Combining (2.6) and (2.7) we see that $|\nabla E(X(t))| \to 0$ as $t \to \infty$. Since $E(X(t))$ is bounded and E is coercive, for every $t_j \to \infty$ there exists a subsequence x_j of $X(t_j)$ such that $x_j \to X_\infty$ as $j \to \infty$. It must be $\nabla E(X_\infty) = 0$, and since E admits a unique global minimum and its convex, we find $X_\infty = X_{min}$. By the arbitrariness of t_j, we conclude that

$$\lim_{t \to \infty} X(t) = X_{min} \,. \tag{2.9}$$

To prove exponential convergence, we take the scalar product of

$$\nabla E(X(t)) - \nabla E(X_{min}) = \int_0^1 \nabla^2 E(sX(t) + (1-s)X_{min}) \cdot (X(t) - X_{min})\, ds$$

with $(X(t) - X_{min})$ and apply uniform convexity to find

$$\lambda\, |X(t) - X_{min}|^2 \leq \big(\nabla E(X(t)) - \nabla E(X_{min})\big) \cdot (X(t) - X_{min}) \,.$$

But thanks to (2.1)

$$\begin{aligned} \frac{d}{dt}|X(t) - X_{min}|^2 &= 2\,(X(t) - X_{min}) \cdot X'(t) \\ &= -2\,(X(t) - X_{min}) \cdot (\nabla E(X(t)) - \nabla E(X_{min})) \\ &\leq -2\,\lambda\, |X(t) - X_{min}|^2 \end{aligned}$$

so that (2.4) follows. □

2.2. **The heat equation as a gradient flow in L^2.** We now informally discuss how the following boundary value problem for the heat equation

$$\begin{cases} u_t - \Delta u = 0 & \text{on } \Omega \times (0, \infty)\,, \\ u|_{t=0} = u_0 & \text{on } \Omega\,, \\ u|_{\partial\Omega} = \psi & \forall t \geq 0\,. \end{cases} \tag{2.10}$$

(where u_0 and ψ are given initial conditions and boundary data) can be seen as a gradient flow in $L^2(\Omega)$. An important conclusion will be that in order to see (2.10) as a gradient flow we shall need to work with convex functions that are possibly infinite value and non-differentiable.

The energy functional will be the Dirichlet energy $E : L^2(\Omega) \to \mathbb{R} \cup \{+\infty\}$, defined as

$$E(u) = \begin{cases} \frac{1}{2} \int_\Omega |\nabla u|^2 & \text{if } u \in \psi + H_0^1(\Omega) \\ +\infty\,, & \text{otherwise}\,. \end{cases}$$

The differential of E at u is defined as

$$dE(u)[\varphi] = \lim_{t \to 0} \frac{E(u + t\varphi) - E(u)}{t}$$

whenever $\varphi \in L^2(\Omega)$ is such that the limit exists. This is equivalent in asking $\varphi \in H_0^1(\Omega)$, and in this case we find

$$E(u + t\varphi) = E(u) + t \int_\Omega \nabla u \cdot \nabla \varphi + t^2\, E(\varphi)\,,$$

so that $dE(u) : H_0^1(\Omega) \to \mathbb{R}$ is the linear operator defined by

$$dE(u)[\varphi] = \int_\Omega \nabla u \cdot \nabla \varphi \qquad \forall \varphi \in H_0^1(\Omega)\,.$$

Defining the gradient of E at u means representing the linear functional $dE(u)$ via a duality pairing. Considering that $dE(u)$ is defined on $H_0^1(\Omega)$, any Hilbert space H containing $H_0^1(\Omega)$ and such that

$$\sup \Big\{ dE(u)[\varphi] : \varphi \in H_0^1(\Omega)\,, \|\varphi\|_H \leq 1 \Big\} < \infty \tag{2.11}$$

is an admissible choice. Indeed, by Riesz theorem, if (2.11) holds, then there exists a unique $\nabla^H E(u) \in H$ such that

$$dE(u)[\varphi] = \langle \nabla^H E(u) | \varphi \rangle_H\,, \qquad \forall \varphi \in H\,.$$

When $u \in H^2(\Omega)$ an integration by parts gives

$$dE(u)[\varphi] = \int_\Omega \nabla u \cdot \nabla \varphi = -\int_\Omega \varphi\, \Delta u\,,$$

so that (2.11) holds with $H = L^2(\Omega)$ and

$$\nabla^{L^2} E(u) = -\Delta u \in L^2(\Omega)\,.$$

Therefore the equation $u_t = \Delta u$ can be seen as a gradient flow

$$U'(t) = -\nabla_{L^2} E(U(t))\,.$$

Let us review the argument of Theorem 2.1 in light of these definitions. The computation (2.5) takes the form,

$$\frac{d}{dt} E(u(t)) = \int_\Omega \nabla u(t) \cdot \nabla u_t(t) = -\int_\Omega (\Delta u(t))^2 = -\|\nabla_{L^2} E(u)\|^2_{L^2(\Omega)}\,,$$

so that the dissipation inequality is

$$\int_0^\infty \int_\Omega (\Delta u(t))^2 \leq \frac{1}{2} \int_\Omega |\nabla u_0|^2\,.$$

Similarly (2.7) takes the form

$$\begin{aligned} \frac{d}{dt} \|\nabla E(u(t))\|^2_{L^2(\Omega)} &= 2 \int_\Omega \Delta u(t)\, \Delta u_t(t) = 2 \int_\Omega \Delta u(t) \Delta(\Delta u(t)) \\ &= -2 \int_\Omega |\nabla(\Delta u(t))|^2 \leq 0\,. \end{aligned}$$

There is exponential convergence to the unique minimum of $E(u)$ over $L^2(\Omega)$, which is the unique harmonic function Ψ with ψ as its boundary data. The fastest way to see this is differentiating

$$\begin{aligned}\frac{d}{dt}\frac{1}{2}\int_\Omega |u(t)-\Psi|^2 &= \int_\Omega (u(t)-\Psi)\, u_t(t) = \int_\Omega (u(t)-\Psi)\,(\Delta u(t)-\Delta\Psi) \\ &= -\int_\Omega |\nabla(u(t)-\Psi)|^2 \le -\lambda_1(\Omega)\int_\Omega |u(t)-\Psi|^2\end{aligned}$$

where we have applied the Poincaré inequality to $u(t)-\Psi \in H^1_0(\Omega)$. As in the finite dimensional case, we thus have

$$\int_\Omega |u(t)-\Psi|^2 \le e^{-2\lambda_1(\Omega)t}\int_\Omega |u_0-\Psi|^2\,.$$

2.3. **Non-smooth convex functions.** The considerations from the previous section show that studying the heat equation as a gradient flow requires understanding convex functions which possibly take infinite values and are not everywhere differentiable.

Let X be a vector space. We say that $E : X \to \mathbb{R}\cup\{+\infty\}$ is convex if $E(tu+(1-T)v) \le t\,E(u)+(1-t)E(v)$ whenever $t\in[0,1]$ and $u,v\in X$. This inequality implies in particular that if $E(tu+(1-T)v)=+\infty$ for some $t\in[0,1]$, then either $E(u)$ or $E(v)$ equals $+\infty$. Thus

$$\mathrm{dom}(E) = \Big\{u\in X : E(u)<\infty\Big\}$$

is a convex subset of the vector space X. We say that E is proper if $\mathrm{dom}(E)\neq\emptyset$.

Convexity and topology. Now assume that E is a Banach space, and consider a proper convex function E on X. The natural property tying the convexity of E to the metric properties of X is the *lower semicontinuity inequality*, namely

$$E(u) \le \liminf_{k\to\infty} E(u_k) \qquad \text{whenever } u_k\to u \text{ in } X\,. \tag{2.12}$$

Notice that **morally** convex functions should automatically be lower semicontinuous, because intuitively they are upper envelopes of affine (thus, continuous) functions. However, this latter properties may fail when we allow for infinite values. For example, consider the proper convex function $E:\mathbb{R}\to\mathbb{R}\cup\{+\infty\}$ defined by

$$E(x) = \begin{cases} 0\,, & \text{for } |x|<1\,,\\ 1\,, & \text{for } |x|=1\,,\\ +\infty\,, & \text{for } |x|>1\,.\end{cases} \tag{2.13}$$

Then E is strictly larger than the supremum of its affine minorants, as

$$E(1) = 1 > 0 = \sup\{a+b : a+bx\le E(x) \quad \forall x\in\mathbb{R}\}\,.$$

Notice that, indeed, E is not lower semicontinuous at $x=\pm1$. So (2.12) is not automatically satisfied.

An important remark is that lower semicontinuity along strongly convergent sequences (2.12) implies lower semicontinuity along **weakly** convergent sequences. Indeed, pick $u_k\to u$ weakly in X, and up to extract a subsequence assume that

$$\liminf_{k\to\infty} E(u_k) = \lim_{k\to\infty} E(u_k)\,.$$

By Mazur's lemma, for suitable $0\le\lambda_k^{(j)}\le1$, $j=k,\dots,N(k)$, with $\sum_{j=k}^{N(k)}\lambda_k^{(j)}=1$ we have

$$\sum_{j=k}^{N(k)}\lambda_k^{(j)}u_j \to u \qquad \text{in } X$$

so that by (2.12) we find

$$
\begin{aligned}
E(u) &\leq \liminf_{k\to\infty} E\Big(\sum_{j=k}^{N(k)} \lambda_k^{(j)} u_j\Big) \\
&\leq \liminf_{k\to\infty} \sum_{j=k}^{N(k)} \lambda_k^{(j)} E(u_j) \leq \liminf_{k\to\infty} \max_{k\leq j\leq N(k)} E(u_j) = \lim_{k\to\infty} E(u_k)\,,
\end{aligned}
$$

as claimed.

Subdifferentials We now restrict our attention to the case when $X = H$ is a Hilbert space, and introduce the important notion of **subdifferential** of a convex proper function E on H. The *subdifferential* of E at $u \in H$ is the set

$$
\partial E(u) = \Big\{ v \in H : E(w) \geq E(u) + \langle v|w-u\rangle \ \forall w \in H \Big\}\,,
$$

and

$$
\mathrm{dom}(\partial E) = \Big\{ u \in H : \partial E(u) \neq \emptyset \Big\}\,.
$$

The subdifferential is the set of slopes v such that the affine function $w \mapsto E(u) + \langle v|w-u\rangle$ of slope v and value $E(u)$ at $w = u$ lies below E everywhere on H; the domain $\mathrm{dom}(\partial E)$ is the set of points u of H such that there exists at least one affine function passing through $(u, E(u))$ and lying below E everywhere on H. Clearly,

$$
\mathrm{dom}(\partial E) \subset \mathrm{dom}(E)\,.
$$

The inclusion may be strict: in example (2.13), $1 \in \mathrm{dom}(E)$ but $1 \notin \mathrm{dom}(\partial E)$, i.e., $\partial E(1) = \emptyset$. Important simple properties of sub-differentials are: (i) $\partial E(u)$ is a convex set for every $u \in H$; (ii) in finite dimension, E is differentiable at any u lying in the interior of $\mathrm{dom}(E)$ if and only if $\partial E(u)$ consists of a single point, the gradient of E at u; (iii) $E(u) = \min_H E$ if and only if $0 \in \partial E(u)$; (iv) if $u, v \in \mathrm{dom}(\partial E)$, then

$$
\langle v-u|v_*-u_*\rangle \geq 0 \qquad \forall v_* \in \partial E(v)\,, u_* \in \partial E(u)\,.
$$

This last property follows immediately by considering the inequalities

$$
\begin{aligned}
E(w) &\geq E(u) + \langle u_*|w-u\rangle \qquad \forall w \in H\,,\\
E(w) &\geq E(v) + \langle v_*|w-u\rangle \qquad \forall w \in H\,,
\end{aligned}
$$

testing the first one at $w = v$, the second one at $w = u$ and then adding up. Property (iv) is called **monotonicity** because in the case $H = \mathbb{R}^n$, $n = 1$, for a differentiable convex function E, it means $(E'(u) - E'(v))(u-v) \geq 0$, which of course implies $E'(u) \leq E'(v)$ whenever $u \leq v$.

2.4. **Construction of gradient flows on Hilbert space.** We now state the existence theorem for gradient flows in Hilbert spaces.

Theorem 2.2. *Let $E : H \to \mathbb{R}\cup\{+\infty\}$ be a proper, convex lower-semicontinuous function on a Hilbert space H. Assume that*

$$
\overline{\mathrm{dom}(\partial E)} = H\,.
$$

Then for every $u_0 \in H$ there exists a unique function $U(t) \in C^0([0,\infty); \mathrm{dom}(\partial E))$ with $U(0) = u_0$ and $U'(t) \in L^\infty((0,\infty); H)$, such that

$$
U'(t) \in -\partial E(U(t)) \qquad \textit{for a.e. } t > 0\,. \tag{2.14}
$$

Remark 2.1. Notice that uniqueness assertion may seem non-obvious, as the condition $U'(t) \in -\partial E(U(t))$ does not uniquely identify $U'(t)$ for any given t such that $\partial E(U(t))$ contains more than one element. However, it U and U_* are two gradient flows with initial condition u_0 in the sense of Theorem 2.2, then for a.e. $t > 0$ we obtain

$$\frac{d}{dt}\frac{\|U(t) - U_*(t)\|^2}{2} = -\langle U(t) - U_*(t) | U'(t) - U_*'(t)\rangle \leq 0$$

thanks to the monotonicity of ∂E. Thus *for every* $t > 0$

$$\|U(t) - U_*(t)\| \leq \|U(0) - U_*(0)\| = \|u_0 - u_0\| = 0\,,$$

and so $U \equiv U_*$.

The approach to Theorem 2.2 is, on a formal level, very similar to the approach to the Hille-Yosida theorem. In analogy with equations (1.5) and (1.6) we may try to define maps A_μ approximating ∂E by setting

$$A_\mu = \frac{\mathrm{Id} - (\mathrm{Id} - \mu\,\partial E)^{-1}}{\mu} = \partial E + \mathrm{O}(\mu)\,.$$

To make sense of this we will first need to show that $J_\mu = (\mathrm{Id} - \mu\,\partial E)^{-1}$ really defines a single valued map. From this point of view, the second identity sign should be interpreted as saying that $A_\mu(u)$ is $\mathrm{O}(\mu)$ distant from $\partial E(u)$. All these things turn out to be true, and we will see that A_μ are Lipschitz maps, whose classical flow converges to a solution of (2.14) as $\mu \to 0^+$.

One can develop a better geometric intuition on this construction by considering the problem in finite dimension. Consider a piece-wise affine convex energy $E : \mathbb{R} \to \mathbb{R}$. The subdifferential of E is multivalued at each point where E' jumps, but nevertheless we can visualize ∂E as a function whose graph contains a few vertical "walls". The idea is obtaining A_μ has function with Lipschitz constant of order $1/\mu$ and smaller slopes than ∂E, in other words, we approximate the "walls" from below with very steep slopes. Notice that, actually, A_μ is the subdifferential of a $C^{1,1}$ convex function E_μ (uniquely determined up to a constant). The flows $X_\mu(t)$ satisfying $X_\mu'(t) = -A_\mu(X_\mu(t))$ and $X_\mu(0) = x_0$ are compact in $\mu > 0$ and converge to a limit flow, which actually solves (uniquely!) $X'(t) \in -\partial E(X(t))$ and $X(0) = x_0$.

In the next theorem we rigorously describe the approximation procedure describe above.

Theorem 2.3. *Let E be a proper, lower-semicontinuous convex function on a Hilbert space H. Then for every $\mu > 0$ and $u \in H$ there exists a unique $J_\mu(u) \in \mathrm{dom}(\partial E)$ such that*

$$u \in J_\mu(u) - \mu\,\partial E(J_\mu(u))\,. \tag{2.15}$$

The resulting map $J_\mu = (\mathrm{Id} - \mu\partial E)^{-1} : H \to H$ is 1-Lipschitz continuous and satisfies

$$\lim_{\mu\to 0^+} J_\mu(u) = u \qquad \forall u \in \overline{\mathrm{dom}(\partial E)}\,. \tag{2.16}$$

Moreover, if we define $A_\mu : H \to H$ by setting

$$A_\mu(u) = \frac{u - J_\mu(u)}{\mu} \qquad u \in H\,,$$

then A_μ is $2/\mu$-Lipschitz continuous, A_μ is monotone on H, and

$$A_\mu(u) \in \partial E(J_\mu(u)) \qquad \forall u \in U \tag{2.17}$$

$$\|A_\mu(u)\| \leq \inf\Big\{\|z\| : z \in \partial E(J_\mu(u))\Big\}\,, \qquad \forall u \in \mathrm{dom}(\partial E)\,. \tag{2.18}$$

Proof. Step one: To show that J_μ is well-defined, given $u \in H$, consider the energy

$$F(v) = \mu\, E(v) + \frac{\|v\|^2}{2} - \langle u|v\rangle \,.$$

We claim that there exists a global minimizer v of F on H. Notice that such a minimizer will be unique (F is strictly convex as $\|v\|^2$ is so). Moreover, $\partial(\|\cdot\|^2/2)(v) = \{v\}$ and $\partial(\langle u|\cdot\rangle)(v) = \{u\}$, so that

$$0 \in \partial F(v) = \Big\{\mu\, w + v - u : w \in \partial E(v)\Big\}\,,$$

will immediately imply that $v = J_\mu(u)$ satisfies (2.15).

To prove the existence of a minimizer of F, we first notice that there exist $a \in \mathbb{R}$ and $v_0 \in H$ such that

$$E(w) \geq a + \langle v_0|w\rangle \qquad \forall w \in H\,. \tag{2.19}$$

This is a consequence of the Hahn-Banach theorem, but a direct proof is also simple. Let us assume that

$$\forall k \in \mathbb{N} \text{ there exists } w_k \in H \text{ s.t. } E(w_k) < -k - k\,\|w_k\|\,. \tag{2.20}$$

If w_k is bounded, then $w_k \rightharpoonup w_0$ and the lower-semicontinuity of E give $E(w_0) = -\infty$, a contradiction. So it must be $\|w_k\| \to \infty$, and then setting $\hat{w}_k = w_k/\|w_k\| \rightharpoonup w_0$ and using convexity we get that, for every $u \in H$,

$$E\Big(\Big(1 - \frac{1}{\|w_k\|}\Big)u + \frac{w_k}{\|w_k\|}\Big) \leq \Big(1 - \frac{1}{\|w_k\|}\Big)\, E(u) + \frac{E(w_k)}{\|w_k\|}$$

that is, by lower-semicontinuity and by (2.20)

$$E(w_0) \leq E(u) + \lim_{k\to\infty} \frac{-k - k\|w_k\|}{\|w_k\|} = -\infty\,,$$

again a contradiction. Thus (2.20) cannot hold, and there exists $C > 0$ such that $E(w) \geq -C - C\|w\|$ for every $w \in H$. Taking $-v_0 \in \partial(C + C\|\cdot\|)$, we see that (2.19) holds with $a = -C$.

Given (2.19) the existence of a minimizer for F on H follows by the direct method. Consider a minimizing sequence v_j,

$$\lim_{j\to\infty} F(v_j) = \inf_H F\,. \tag{2.21}$$

By (2.19) we have

$$+\infty > \inf_H F \geq \frac{\|v_j\|^2}{2} - \|u\|\,\|v_j\| - \mu\Big(|a| + \|v_0\|\,\|v_j\|\Big)$$

which immediately implies the boundedness of v_j, and thus $v_j \rightharpoonup v$ for some $v \in H$. By lower-semicontinuity of F and by (2.21) we find

$$\inf_H F \leq F(v) \leq \liminf_{j\to\infty} F(v_j) = \inf_H F$$

so that v is indeed a minimizer of F over H.

Step two: We prove the Lipschitz estimates for J_μ and A_μ, and the monotonicity of A_μ. If $u_k \in H$ then there exists $w_k \in \partial E(J_\mu(u_k))$ such that $u_k = J_\mu(u_k) - \mu\, w_k$; hence

$$\begin{aligned}\|u_1 - u_2\|^2 &= \|J_\mu(u_1) - J_\mu(u_2)\|^2 + \mu^2\,\|w_1 - w_2\|^2 + 2\,\mu\,\langle w_1 - w_2|J_\mu(u_1) - J_\mu(u_2)\rangle \\ &\geq \|J_\mu(u_1) - J_\mu(u_2)\|^2 + \mu^2\,\|w_1 - w_2\|^2\end{aligned}$$

thanks to the monotonicity of ∂E. In particular, $\mathrm{Lip}(J_\mu) \le 1$, and as a consequence $\mathrm{Lip}(A_\mu) \le 2/\mu$. The monotonicity of A_μ also follows from $\mathrm{Lip}(J_\mu) \le 1$: indeed,

$$\begin{aligned}
\langle A_\mu(u_1) - A_\mu(u_2)|u_1 - u_2\rangle &= \frac{1}{\mu}\langle u_1 - J_\mu(u_1) - (u_2 - J_\mu(u_2))|u_1 - u_2\rangle \\
&= \frac{\|u_1 - u_2\|^2}{\mu} + \frac{1}{\mu}\langle J_\mu(u_2) - J_\mu(u_1)|u_1 - u_2\rangle \\
&\ge \frac{\|u_1 - u_2\|}{\mu}\Big\{\|u_1 - u_2\| - \|J_\mu(u_2) - J_\mu(u_1)\|\Big\} \ge 0\,,
\end{aligned}$$

where $\mathrm{Lip}(J_\mu) \le 1$ was used in the last inequality.

Step three: We prove (2.17) and (2.18). By definition of $J_\mu(u)$, if $u \in H$ then there exists $z \in \partial E(J_\mu(u))$ such that $u = J_\mu(u) - \mu\, z$. In particular, for every $w \in H$

$$E(w) \ge E(J_\mu(u)) + \langle z|u - J_\mu(u)\rangle = E(J_\mu(u)) + \langle A_\mu(u)\,|u - J_\mu(u)\rangle$$

that is, $A_\mu(u) \in \partial E(J_\mu(u))$. We can exploit this last property together with the monotonicity of ∂E to find that if $u \in \mathrm{dom}(\partial E)$ and thus there exists $z \in \partial E(u)$, then

$$\begin{aligned}
0 &\le \langle u - J_\mu(u)|z - A_\mu(u)\rangle = \langle A_\mu(u)|z - A_\mu(u)\rangle = \langle A_\mu(u)|z\rangle - \|A_\mu(u)\|^2 \\
&\le \|A_\mu(u)\|\,\|z\| - \|A_\mu(u)\|^2
\end{aligned}$$

that is the **minimal slope property of** A_μ stated in (2.18). Finally,

$$\|J_\mu(u) - u\| = \mu\,\|A_\mu(u)\| \le \mu\,\inf\Big\{\|z\| : z \in \partial E(u)\Big\}\,.$$

This implies $J_\mu(u) \to u$ as $\mu \to 0^+$ since we are assuming $u \in \mathrm{dom}(\partial E)$, and thus

$$\inf\Big\{\|z\| : z \in \partial E(u)\Big\} < \infty\,.$$

The convergence of $J_\mu(u) \to u$ for $u \in \overline{\mathrm{dom}(\partial E)}$ follows easily by density. □

Proof of Theorem 2.2. Since A_μ is a Lipschit map we can apply the Cauchy-Picard theorem and define $U_\mu(t) \in C^1([0,\infty);H)$ in such a way that $U'_\mu(t) = -A_\mu(U_\mu(t))$ and $U_\mu(0) = u_0$. We aim to show that $U_\mu(t)$ has a limit as $\mu \to 0^+$ and that this limit $U(t)$ defines the desired solution to the gradient flow defined by E.

Step one: We obtain a basic estimate on U'_μ, namely

$$\sup_{t\ge 0}\|U'_\mu(t)\| \le \|A_\mu(U_\mu(0))\| = \|A_\mu(u_0)\|\,. \tag{2.22}$$

To this end we first compare $U_\mu(t)$ with the flow $V'_\mu(t) = -A_\mu(V_\mu(t))$ corresponding to another initial condition, $V_\mu(0) = v_0$. In this way

$$\begin{aligned}
\frac{d}{dt}\frac{\|U_\mu(t) - V_\mu(t)\|^2}{2} &= \langle U'_\mu(t) - V'_\mu(t)|U_\mu(t) - V_\mu(t)\rangle \\
&= -\langle A_\mu(U_\mu(t)) - A_\mu(V_\mu(t))|U_\mu(t) - V_\mu(t)\rangle \le 0\,,
\end{aligned}$$

that is

$$\|U_\mu(t) - V_\mu(t)\| \le \|u_0 - v_0\| \qquad \forall t \ge 0\,. \tag{2.23}$$

We fix $h > 0$, and apply this property with $v_0 = U_\mu(h)$, that is, we compare the flow $U_\mu(t)$ with that lagged flow $U_\mu(t+h) = V_\mu(t)$: we find

$$\|U_\mu(t) - U_\mu(t+h)\| \le \|u_0 - v_0\| = \|U_\mu(0) - U_\mu(h)\| \le \int_0^h \|U'_\mu(t)\|\,dt = \int_0^h \|A_\mu(U_\mu(t))\|\,dt\,.$$

Dividing by h and letting $h \to 0$ we find (2.24).

Step two: Assume now that $u_0 \in \mathrm{dom}(\partial E)$, then by (2.18) the right hand side of (2.24) is bounded uniformly in μ, that is

$$\sup_{\mu>0}\sup_{t\geq 0}\|U_\mu'(t)\| \leq \inf\Big\{\|z\| : z \in \partial E(u_0)\Big\}, \tag{2.24}$$

which implies the weak sequential compactness of $\{U_\mu'(t)\}$ in $L^\infty((0,\infty);H)$. In order to exploit this estimate we now compare the flows $U_\mu(t)$ and U_ν corresponding to $\mu, \nu > 0$. To this end we compute that

$$\begin{aligned}\frac{d}{dt}\frac{\|U_\mu(t)-U_\nu(t)\|^2}{2} &= \langle U_\mu'(t)-U_\nu'(t)|U_\mu(t)-U_\nu(t)\rangle \\ &= -\langle A_\mu(U_\mu(t))-A_\nu(U_\nu(t))|U_\mu(t)-U_\nu(t)\rangle\,.\end{aligned}$$

Considering that $U_\mu = \mu\, A_\mu(U_\mu) + J_\mu(U_\mu)$, and dropping the t-dependency for the sake of clarity, we have

$$\begin{aligned}\frac{d}{dt}\frac{\|U_\mu-U_\nu\|^2}{2} &= -\langle A_\mu(U_\mu)-A_\nu(U_\nu)|\mu\, A_\mu(U_\mu)-\nu\, A_\nu(U_\nu)\rangle \\ &\quad -\langle A_\mu(U_\mu)-A_\nu(U_\nu)|J_\mu(U_\mu)-J_\nu(U_\nu)\rangle \\ &\leq -\langle A_\mu(U_\mu)-A_\nu(U_\nu)|\mu\, A_\mu(U_\mu)-\nu\, A_\nu(U_\nu)\rangle\end{aligned}$$

as $A_\mu(U_\mu) \in \partial E(J_\mu(U_\mu))$. Hence

$$\begin{aligned}&\langle A_\mu(U_\mu)-A_\nu(U_\nu)|\mu\, A_\mu(U_\mu)-\nu\, A_\nu(U_\nu)\rangle \\ &= \mu\|A_\mu(U_\mu)\|^2+\nu\|A_\nu(U_\nu)\|^2-(\mu+\nu)\langle A_\mu(U_\mu)|A_\nu(U_\nu)\rangle \\ &\geq \mu\|A_\mu(U_\mu)\|^2+\nu\|A_\nu(U_\nu)\|^2-\mu\Big(\|A_\mu(U_\mu)\|^2+\frac{\|A_\nu(U_\nu)\|^2}{4}\Big) \\ &-\nu\Big(\|A_\nu(U_\nu)\|^2+\frac{\|A_\mu(U_\mu)\|^2}{4}\Big)\,,\end{aligned}$$

where we have used $ab \leq a^2 + b^2/4$. In conclusion,

$$\frac{d}{dt}\frac{\|U_\mu-U_\nu\|^2}{2} \leq -\frac{\mu}{4}\,\|A_\nu(U_\nu)\| - \frac{\nu}{4}\,\|A_\mu(U_\mu)\|$$

which, thanks to (2.24), implies

$$\frac{d}{dt}\,\frac{\|U_\mu-U_\nu\|^2}{2} \leq \frac{\mu+\nu}{4}\,\inf\Big\{\|z\| : z \in \partial E(u_0)\Big\}, \tag{2.25}$$

and hence

$$\frac{\|U_\mu-U_\nu\|^2}{2} \leq \frac{\mu+\nu}{4}\,\inf\Big\{\|z\| : z \in \partial E(u_0)\Big\}\,t, \qquad \forall t \geq 0\,, \mu,\nu > 0\,. \tag{2.26}$$

Now (2.26) implies that, for each $t \geq 0$, $\{U_\mu(t)\}_{\mu>0}$ is a Cauchy family in H, so that $U_{\mu'}(t) \to U(t)$ in H for some $U(t) \in H$ and along a subsequence $\mu' \to 0^+$. By (2.26), the convergence is locally uniform on $t \in [0,\infty)$, so that $U(t) \in C^0([0,\infty);H)$. By (2.24), U admits a weak derivative $U'(t) \in L^\infty((0,\infty);H)$ with

$$\int_a^b \langle U_{\mu'(t)}'|w\rangle dt \to \int_a^b \langle U'(t)|w\rangle dt \qquad \text{as } \mu' \to 0^+ \tag{2.27}$$

whenever $w \in H$ and $(a,b) \subset\subset (0,\infty)$, and

$$\sup_{t\geq 0}\|U'(t)\| \leq \inf\Big\{\|z\| : z \in \partial E(u_0)\Big\}. \tag{2.28}$$

Recalling that $U_\mu' = -A_\mu(U_\mu) \in \partial E(J_\mu(U_\mu))$ we see that

$$E(w) \geq E(J_\mu(U_\mu)) + \langle -U_\mu'|w - J_\mu(U_\mu)\rangle$$

thus that

$$(b-a)E(w) \geq \int_a^b E(J_{\mu'}(U_{\mu'})(t)) + \langle -U'_{\mu'}(t)|w - J_{\mu'}(U_{\mu'}(t))\rangle \, dt$$

Now, $J_\mu(U_\mu(t)) - U_\mu(t) \to 0$ uniformly on t as

$$\|J_\mu(U_\mu(t)) - U_\mu(t)\| = \mu \, \|A_\mu(U_\mu(t))\| \leq \mu \, \min \Big\{ \|z\| : z \in \partial E(u_0) \Big\},$$

hence $J_{\mu'}(U_{\mu'})(t) \to U(t)$ as $\mu' \to 0$ uniformly on (a,b). Therefore by (2.27) and by lower-semicontinuity of E we obtain

$$(b-a)E(w) \geq \int_a^b E(U(t)) + \langle -U'(t)|w - U(t))\rangle \, dt \, .$$

Dividing by $b-a$ and letting $b \to a$ we conclude that, for a.e. $a > 0$,

$$E(w) \geq E(U(a)) + \langle -U'(a)|w - U(a))\rangle \qquad \forall w \in H$$

i.e. $U(a) \in -\partial E(U(a))$ for a.e. $a > 0$, that is (2.14). We are left to prove that $U(t) \in \mathrm{dom}(\partial E(U(t)))$ for every $t > 0$, and not just for a.e. $t > 0$. It suffices to pick any $t_j \to t$ such that $U'(t_j) \in -\partial E(U(t_j))$ and notice that by (2.28), up to extract a subsequence, $U'(t_j) \rightharpoonup v$ weakly in H. By lower-semicontinuity of E, the inequalities

$$E(w) \geq E(U(t_j)) + \langle -U'(t_j)|w - U(t_j)\rangle \qquad \forall w \in H \, ,$$

imply that $U(t) \in \mathrm{dom}(\partial E)$, with $-v \in \partial E(U(t))$.

Final remarks: Notice that the uniqueness property of the convex gradient flow implies that actually $U_\mu(t) \to U(t)$ as $\mu \to 0^+$, and not just as $\mu' \to 0$ for a subsequence μ'. Also, the assumption $u_0 \in \mathrm{dom}(\partial E)$ is easily dropped, for, pick any $u_0 \in H$ and use the density assumption of $\mathrm{dom}(\partial E)$ into H to approximate u_0 with $u_k \in \mathrm{dom}(\partial E)$. Denote by $U^k(t)$ the corresponding flows and notice that by differentiating $\|U^k(t) - U^j(t)\|^2$ and by monotonicity of E we get

$$\|U^k(t) - U^j(t)\| \leq \|u_k - u_j\| \, .$$

Hence $\{U^k(t)\}_{k \in \mathbb{N}}$ is a Cauchy sequence and it easily seen, by repeating the argument used above, that the limit flow is indeed a gradient flow in the sense of Theorem 2.2 with $U(0) = u_0$. $\square$

2.5. **An example.** We finally discuss an example of a non-linear flow, precisely, we consider a non-linear parabolic PDE of the form

$$\begin{cases} u_t = \mathrm{div}\,(\nabla f(\nabla u)) & \text{on } \Omega,\ t > 0 \\ u(0) = u_0 \, , & \text{on } \Omega \\ u = \psi \, , & \text{on } \partial\Omega,\ t > 0. \end{cases} \tag{2.29}$$

which corresponds, formally, to the gradient flow of the energy

$$E(u) = \int_\Omega f(\nabla u) \qquad u : \Omega \to \mathbb{R}$$

defined by an integrand $f : \mathbb{R}^n \to \mathbb{R}$. Indeed,

$$E(u + t\,\varphi) = E(u) + t \int_\Omega \nabla f(\nabla u) \cdot \nabla \varphi + o(t)$$

so that, formally,

$$dE(u)[\varphi] = \int_\Omega \nabla f(\nabla u) \cdot \nabla \varphi \, .$$

If $\varphi = 0$ on $\partial\Omega$, then an integration by parts gives

$$dE(u)[\varphi] = -\int_\Omega \varphi \mathrm{div}\,(\nabla f(\nabla u))$$

so that, if $\mathrm{div}\,(\nabla f(\nabla u)) \in L^2(\Omega)$, then the L^2-gradient of E at u is given by

$$\nabla^{L^2} E(u) = -\mathrm{div}\,(\nabla f(\nabla u))\,.$$

In this sense the nonlinear parabolic PDE (2.29) can be seen as the gradient flow of the energy E. By exploiting the H^2-regularity theory for elliptic equations in divergence form it is not hard to show that (2.29) can indeed be solved in the framework described in this section. See Section 9.6.3 in Evans's PDE book.

(Francesco Maggi)
Department of Mathematics, University of Texas at Austin,
2515 Speedway STOP C1200, 78712, Austin, Texas, USA
E-mail address: maggi@math.utexas.edu