

# Relevance of the Calculus of Variations to Evolution Equations

## 1. PARABOLIC

Recall that the model parabolic equation is the heat equation. Without giving a precise definition, parabolic equations in general are equations that “behave like” the heat equation.

Many parabolic PDEs can be understood as “gradient flows”. To illustrate what this means, we first discuss the simpler case of ODEs:

1.1. **ODEs.** Suppose  $V : \mathbb{R}^n \rightarrow \mathbb{R}$  is a smooth function, and consider the equation

$$(1) \quad \frac{dx}{dt} = -\nabla V(x(t)), \quad x(0) = x_0.$$

This is a system of  $n$  coupled ODEs. An equation of this sort is sometimes said to be a “gradient flow” equation.

The equation (1) states that trajectories  $x(t)$  of solutions “flow downhill” (ie in the direction  $-\frac{\nabla V}{|\nabla V|}$  with a velocity proportional to  $|\nabla V|$ ).

Solutions enjoy a number of estimates. If  $x(t)$  is a solution, then

$$\frac{d}{dt}V(x(t)) = \nabla V(x(t)) \cdot \frac{dx}{dt} = -|\nabla V(x(t))|^2 = -\left|\frac{dx}{dt}\right|^2.$$

Thus

$$(2) \quad V(x(T)) + \int_0^T \left|\frac{dx}{dt}\right|^2 dt = V(x_0)$$

for every  $T > 0$ . If for example  $V$  is bounded below, then it follows that

$$(3) \quad \int_0^T \left|\frac{dx}{dt}\right|^2 dt \leq C$$

for all  $T > 0$ , where the constant  $C$  depends on  $V$  and on the initial data  $x_0$ .

One can use the gradient flow structure of (1) to construct solutions via “variational methods”, ie, by a procedure that involves minimization problems. Consider the following procedure for solving (1):

- fix a small parameter  $h > 0$ . We will write  $x_i^h$  to denote an approximate solution (which we are about to construct) at time  $t = ih$ .

We define  $x_0^h = x_0$ .

For  $i \geq 1$ , we iteratively define  $x_i^h$  to be the minimizer of the function

$$(4) \quad x \mapsto G^h(x; x_{i-1}^h) := \frac{1}{2h}|x - x_{i-1}^h|^2 + V(x).$$

- Then we define an approximate solution  $x^h : [0, \infty] \rightarrow \mathbb{R}^n$  by specifying that

$$x^h(t) = x_i^h + (t - ih)\frac{x_{i+1}^h - x_i^h}{h} \quad \text{for } t \in [ih, (i+1)h].$$

Thus  $x^h(ih) = x_i^h$  for every integer  $i$ , and  $x^h$  is affine on every interval  $[ih, (i+1)h]$ .

- One would then hope to prove that  $x^h(\cdot)$  converges in some sense to a limit  $x(\cdot)$ , and that this limit solves (1). (See Proposition 1 below.)

**Proposition 1.** *Assume that  $V : \mathbb{R}^n \rightarrow \mathbb{R}$  is smooth and nonnegative, with  $D^2v \|_{L^\infty(\mathbb{R}^n)} < \infty$ , and that  $x_0 \in \mathbb{R}^n$  is given.*

*For  $h > 0$ , define  $x^h : \mathbb{R} \rightarrow \mathbb{R}^n$  via the procedure described above,*

*Then there exists a subsequence  $h_k \rightarrow 0$  and function  $x : \mathbb{R} \rightarrow \mathbb{R}^n$  such that  $x^h(\cdot) \rightarrow x(\cdot)$  locally uniformly on  $\mathbb{R}$ , and in addition  $x$  solves (1).*

*sketch of proof.* The proof mimics estimates that we have already seen are valid for smooth solutions — the point is that these estimates remain approximately true for our approximate solutions.

1. Note that the minimization problem (4) solved by  $x_i^h$  implies that

$$\frac{x_i^h - x_{i-1}^h}{h} + \nabla V(x_i^h) = 0.$$

This can be seen to be a discrete form of the equation (1). Deduce that

$$\left| \frac{1}{h} [V(x_i^h) - V(x_{i-1}^h)] + \left| \frac{x_i^h - x_{i-1}^h}{h} \right|^2 \right| \leq C \frac{|x_i^h - x_{i-1}^h|^2}{h}$$

conclude from this that there exists a constant  $C$ , depending only  $V(x_0)$  such that

$$(5) \quad \int_0^T \left| \frac{d}{dt} x^h \right|^2 dt \leq C$$

for all  $h \in (0, 1]$  and all  $T > 0$ .

2. It follows from (5) and of Rellich's Compactness Theorem that there exists a subsequence  $h_k \rightarrow 0$  and a function  $x : [0, \infty) \rightarrow \mathbb{R}^n$  such that  $x^h(\cdot) \rightarrow x(\cdot)$  uniformly on  $[0, T]$ , for every  $T > 0$ , and such that  $\frac{d}{dt} x^{h_k} \rightharpoonup \frac{d}{dt} x$  weakly in  $L^2$ . Verify that  $x(\cdot)$  solves (1).

(A relatively easy way to do this is first to show that

$$x(b) - x(a) = - \int_a^b \nabla V(x(t)) dt$$

for every  $0 \leq a < b < \infty$ , and to deduce from here that  $t \mapsto x(t)$  is  $C^1$ , say, and satisfies (1).)

□

**1.2. PDEs.** As stated above, many parabolic equations can be cast in the same “gradient flow” framework.

One subtlety that arises here is that we have to understand what we mean by “gradient”. (This subtlety also arises in the ODE setting if we work on manifolds rather than on Euclidean space.)

If  $H$  is a Hilbert space, and  $I : H \rightarrow \mathbb{R}$  is a functional, we define the gradient  $\nabla I$  at a point  $u \in H$  by requiring that

$$(\nabla I(u), v) = \frac{d}{dt} I[u + tv] \Big|_{t=0}$$

(whenever this definition makes sense). Here  $(\cdot, \cdot)$  denotes the inner product in  $H$ .

We will sometimes write  $\nabla_H$  to explicitly indicate the role that the Hilbert space  $H$  plays in the definition of  $\nabla I$ .

In the following discussion, we consider the functional

$$I[u] := \int_U \frac{1}{2} |Du|^2 + F(u) \, dx$$

where  $F$  is a smooth, nonnegative function, and  $I[\cdot]$  satisfies all the assumptions we need to guarantee existence of minimizers etc. We examine the associated gradient flow equations with respect to 2 different Hilbert space structures. All of our computations are formal. However, these formal arguments can often be converted into proofs, like the one sketched above in the case of ODEs, that solutions can be constructed via approximation procedures that use minimization problems to construct time-discretized gradient flows.

**Example 1a.**

Let  $H = L^2(U)$ . Then formally,

$$\frac{d}{dt} I[u + tv]|_{t=0} = \int_U Du \cdot Dv + F'(u)v = \int_U (-\Delta u + F'(u))v = (-\Delta u + F'(u), v)_{L^2}$$

We are assuming that  $u = v$  on  $\partial U$ . Thus (write  $f = F'$ ) we conclude formally that

$$\nabla_{L^2} I[u] = -\Delta u + f(u).$$

Thus we expect that a possible gradient flow equation for the functional  $I[\cdot]$ , with respect to the  $L^2$  Hilbert space structure, is given by

$$(6) \quad u_t = \Delta u - f(u) \quad \text{in } U.$$

This suggests that we can construct solutions of the above equation (say with the same boundary conditions  $u = 0$  on  $\partial U$  that we imposed above) by exactly the same kind of approximation procedure used in the ODE case above, except that here we define  $u_i^h$  to minimize the functional

$$(7) \quad u \mapsto G[u; u_{i-1}^h] := \frac{1}{2h} \|u - u_{i-1}^h\|_{L^2}^2 + I[u]$$

in  $H_0^1(U)$ . (Thus, for the minimization problem, the boundary condition  $u = 0$  that we used on our formal calculation above makes sense.)

It is easy to see that a minimizer  $u_i^h$  satisfies the equation

$$(8) \quad \frac{1}{h}(u_i^h - u_{i-1}^h) = \Delta u_i^h - f(u_i^h) \quad \text{in } U$$

in the weak sense. This can be seen as a time-discretized version of (6), and one can prove as in the ODE examples above that approximate solutions constructed by linearly interpolating between the  $u_i^h$ s converge to a weak solution of (6).

**Example 1b.**

In this example, we again take  $H = L^2(U)$ , but with an inner product that is different from (although topologically equivalent to) the standard inner product structure which we used above. We will write  $\tilde{L}^2$  to indicate this inner product. To this end, fix a function  $w \in L^\infty(U)$  such that  $a^2 \leq w(x) \leq b^2$  a.e., for some positive constants  $a < b$ , and define

$$(u, v)_{\tilde{L}^2} := \int w(x)u(x)v(x) \, dx.$$

Clearly

$$a\|u\|_{L^2} \leq \|u\|_{\tilde{L}^2} \leq b\|u\|_{L^2}$$

so topologically, the associated Hilbert space is exactly  $L^2(U)$ . But the inner product is clearly non-standard.

Formally, it is easy to check that

$$\nabla_{\tilde{L}^2} I[u] = -\frac{1}{w} \operatorname{div}(w Du) + f(u)$$

where “div” denotes the divergence. So the gradient flow equation should be

$$u_t = \frac{1}{w} \operatorname{div}(w Du) - f(u).$$

If we proceed as in (7), but with the  $L^2$  norm replaced by the  $\tilde{L}^2$  norm, then the Euler-Lagrange equation satisfied by  $u_t^h$  is indeed a discrete form of the above PDE, analogous to (8).

**Example 2.**

Here we take  $H = H_0^1(U)$ , with the inner product  $(u, v)_{H^1} := \int_U Du \cdot Dv$ .

Then  $\nabla_{H^1} I[u] = w$  if and only if

$$(w, v)_{H^1} := \int Dw \cdot Dv = \frac{d}{dt} I[u + tv]|_{t=0} = \int_U Du \cdot Dv + F'(u)v = \int_U (-\Delta u + F'(u))v$$

for all  $v \in H_0^1(U)$ . This is equivalent to  $w$  solving the

$$-\Delta w = -\Delta u + f(u) \quad \text{in } U, w = 0 \text{ on } \partial U.$$

In other words,

$$\nabla_{H^1} I[u] = (-\Delta)^{-1}(-\Delta u + f(u))$$

Thus the gradient flow equation becomes

$$u_t = (-\Delta)^{-1}(\Delta u - f(u)).$$

In particular, if  $f = 0$  then  $\nabla_{H^1} I[u] = u$  for  $u \in H_0^1$ , and the gradient flow equation becomes  $u_t = -u$ .

As above, one can construct solutions of this equation via discretizing time and solving a variational problem to construct approximate solutions, although in the  $f = 0$  case of course this is not necessary.....

**Example 3.**

It is an interesting fact that the fourth-order PDE

$$u_t = -\Delta(\Delta u - f(u))$$

can be realized as a gradient flow for the functional  $I[\cdot]$  above, with respect to (a specific choice of) the  $H^{-1}$  norm. We omit the details.....

**Example 4.**

Many examples involve gradient flows with respect to what are known as the “Wasserstein distance”; this is a topic of active research. (We omit even the definition.....)

## 2. WAVE AND SCHRÖDINGER

Many equations of interest mathematical physics arise as the Euler-Lagrange equations for certain functionals. For example, the semilinear wave equation

$$(9) \quad u_{tt} - \Delta u + f(u) = 0$$

is formally the Euler-lagrange equation for the functional

$$(10) \quad I[u] := \int_U \frac{1}{2}(-u_t^2 + |Du|^2) + F(u) \, dt \, dx$$

where  $f = F'$  and  $U$  is an open subset of  $\mathbb{R}_t \times \mathbb{R}_x^n$ . Similarly, the Schrödinger equation

$$iu_t - \Delta u + f(u) = 0$$

(for a complex-valued function  $u$ ) is formally the Euler-lagrange equation for the functional

$$I[u] := \int_U \frac{1}{2}(\operatorname{Im}(u_t \bar{u}) + |Du|^2) + F(u) \, dt \, dx$$

Indeed, it is arguably a general physical principle that *every basic PDE of mathematical physics is the Euler-Lagrange equation for an associated functional*, which is usually referred to as an “action functional”.

Note that the above functionals lack any sort of coercivity or convexity properties, and so it is generally impossible to study existence and uniqueness of solutions of the above equations by variational methods.

However, knowing that an equation has a variational structure, and knowing the associated functional, is often useful. One reason for this is what is known as “Noether’s Principle”, which establishes a correspondence between symmetries of an action functional and conservation laws for the associated Euler-Lagrange equation.

**Example.** Here is a relatively simple but important illustration of how Noether’s principle works in practice. Consider the Lagrangian for the action functional (10)

$$L : (\mathbb{R} \times \mathbb{R}^n) \times \mathbb{R} \times (\mathbb{R}_t \times \mathbb{R}_x^n) : (q, p, z, t, x) \mapsto L(q, p, z, t, x) = \frac{1}{2}(-q^2 + |p|^2) + F(z)$$

(where  $q$  is the dummy variable standing in for  $u_t$ ). Note that the Lagrangian does not explicitly depend on the  $t$  variable (or on any of the  $x$  variables for that matter.) When this holds, Noether’s principle guarantees that if we multiply the equation (9) by  $u_t$ , then the result expression can be rewritten as a divergence. Let’s do it:

$$\begin{aligned} u_t(u_{tt} - \Delta u + f(u)) &:= \left( \frac{1}{2}u_t^2 + F(u) \right)_t - u_t \Delta u \\ &= \left( \frac{1}{2}u_t^2 + F(u) \right)_t - \sum_i (u_t u_{x_i})_{x_i} + u_{tx_i} u_{x_i} \\ &= \left( \frac{1}{2}u_t^2 + F(u) \right)_t - \sum_i (u_t u_{x_i})_{x_i} + \left( \frac{1}{2}|Du|^2 \right)_t \\ (11) \quad &= \frac{\partial}{\partial t} e(u) - \sum_i (u_t u_{x_i})_{x_i} \end{aligned}$$

for

$$e(u) := \frac{1}{2}(u_t^2 + |Du|^2) + F(u).$$

(Note that the write-hand side of (11) can be written as  $\operatorname{div}_{t,x}(e(u), u_t Du)$  where  $\operatorname{div}_{t,x}$  is the divergence in the  $t$  and  $x$  variables.)

It follows from the above that if  $u$  is a solution of (9), then  $\frac{\partial}{\partial t}e(u) = \sum_i (u_t u_{x_i})_{x_i}$ . It follows from this that if for example  $u$  is a smooth, compactly supported solution of (9) on  $\mathbb{R}^n$ , then

$$\frac{d}{dt} \int e(u)(x, t) dx = 0.$$

This is often interpreted as “conservation of energy”.

Similarly, the fact that the Lagrangian does not depend on any  $x_i$  implies that for each  $i$ ,  $u_{x_i}(u_{tt} - \Delta u + f(u))$  can be written as a divergence.

The next lemma gives one way of making Noether’s principle into a precise statement.