Variational Modelling

Lecture notes

Mark A. Peletier

September 19, 2011

Contents

0.1 Introduction

These lecture notes provide the material, some background, and some exercises for the course 'Variational Modelling', that I gave in the first half of 2011. The ultimate aim is to explain how and why Wasserstein gradient flows are natural and powerful concepts in the modelling of a variety of systems, and to provide a guide for their usage. On the way to this topic, we shall also discuss more basic ideas on the modelling of mechanical systems using concepts of force, work, power, energy, and entropy.

I try to use layout to indicate different types of text.

Grey backgrounds indicate illustrative examples.

Smaller type indicates sections that can be skipped on first reading.

These notes are very much in-progress; red type indicates remarks to myself. Don't accept anything on face value; that's actually a good general principle in science, but in

this case it's even more true since these notes are in a continuous state of flux, and probably contain (many) errors.

0.2 Acknowledgements

Thanks to Alex Cox, David Bourne, Manh Hong Duong, Joep Evers, and Michiel Renger for comments on earlier drafts of these notes.

Chapter 1

Modelling of real-world systems

In this chapter we introduce some basic ideas on the modelling of real-world systems, while focussing on mechanical systems with state spaces in $\mathbb{R}^n$. Later in these notes we will consider state spaces that are less simple than $\mathbb{R}^n$, such as infinite-dimensional spaces (Chapter 2) and spaces of measures (Chapter 3).

1.1 Generalities

Modelling is an art. It involves choices that can not be rationally justified. Personal preferences are important, and 'taste' plays a major role.

But some choices are more rational than others. We begin by formulating a few choices that, at least in my opinion, are firmly based in observation.

1.1.1 Kinematics and state

The first and most basic choice is that of the degrees of freedom of the system. Take a pendulum as an example. If a pendulum is assumed to be rigid, then the position of the pendulum can be described by the angle φ with the vertical. If, on the other hand, a bead can glide up and down the pendulum, then the system possesses an additional degree of freedom, and an additional variable is necessary to describe it. The term *kinematics* refers to the complete description of the freedom of movement of the system.

The representation of the degrees of freedom may also depend on the purpose of the modelling. The two values $\varphi = 0$ and $\varphi = 2\pi$ correspond to the same physical position of the pendulum, and it is natural to consider φ as an element of $[0, 2\pi)$ with periodicity (i.e. the quotient set $\mathbb{R}/2\pi\mathbb{Z}$). However, if we wish to keep track of how many times the pendulum has gone round one way or the other, then it is more natural to consider φ as an element of $\mathbb{R}$, allowing us to see that $\varphi = 0$ and $\varphi = 6\pi$ differ by three full turns.

The customary definition of *state* is a (preferably minimal) set of variables that is sufficient to determine the future evolution of the system. The state necessarily contains at least all the degrees of freedom of the system, but often also their time derivatives. Taking

the rigid pendulum again as an example, the angle alone is not sufficient to determine the evolution, since two initial states with the same initial angle but with different initial velocities give rise to different subsequent evolutions; therefore the state of the rigid pendulum has to contain at least the angle and its rate of change (and it turns out that these two *are* sufficient to characterize the future evolution).

In this chapter we shall denote a state as x, which will be an element of the state space $\mathbb{R}^n$ for some n.

1.1.2 Force, Power, and Work

Mechanical systems react to *forces*. The correct way to characterize a force is by defining how much work per unit of time it performs on the system, when the system undergoes an infinitesimal movement, characterized by a velocity v. (Often this velocity v is called a 'virtual motion', and this approach is known as the method of Virtual Power. Why this is the correct way we will discuss in Section 1.2.5). An example will clarify what I mean by this.

A useful example is always the movement of a point of mass m in $\mathbb{R}^n$, characterized by its position $x \in \mathbb{R}^n$. Let us make this example dimensional; mass is measured in kg, position in m, and time in s, so that velocity has dimensions $[dx/dt] = \mathrm{m/s}$ and acceleration $\mathrm{m/s^2}$.

Starting with the usual concept of a force as a vector f with dimensions $\mathrm{N} = \mathrm{kg\,m/s^2}$, we define its power, or work per unit of time, on the mass moving at velocity v as $f \cdot v$ (with dimensions $\mathrm{Nm/s} = \mathrm{W}$, Watt).

In mechanical systems in $\mathbb{R}^n$, power is typically of the form $f \cdot v$, but we will generalize below to other systems.

Note that $f \cdot v$ only 'sees' the component of f in the direction of v. This implies that we can disregard any component of f that we know beforehand to be orthogonal to v. A common situation of this type is when the position x is constrained to lie on some surface in $\mathbb{R}^n$, such as in the case of a ball rolling along a surface. If the ball is held on the surface by a force that is normal to the surface, then whatever the evolution, v will be tangential to the surface, and therefore the power of this normal force will be zero.

We need to distinguish between *internal* and *external* sources of power, i.e. internal and external forces. Internal forces arise from properties of the system itself, for instance the stresses in the bent pole of a pole jumper, or the tension in a clothes line. External forces arise from the interaction between the system and its surroundings, such as the pull of the earth's gravitational field, or the pressure of water on the hull of a ship.

In three-dimensional continuum mechanics an additional requirement helps to distinguish between internal and external sources of power: the internal sources of power should be zero when the system as a whole undergoes a rigid motion, i.e. a combination of uniform translation and rotation (see e.g. [Sal01, BMM06]). If a source of power satisfies that condition, it is internal; if not, external. The reasonability of this condition can be best

recognized by imagining a set of point particles that interact; the interparticle forces should be invariant under translation and rotation of the whole set of particles. By contrast, the forces exerted by external objects on the particles will vary with a rigid motion.

A special external force is the *inertial force*, which arises from the inertial property of mass. It is external because it *does* vary under rigid-motion transformations, as any child knows from experience with roundabouts. Specifying the inertial force is most easily done in a *Galilean* or *inertial* frame, where for instance the expression for a particle of mass m undergoing acceleration $a = \ddot{x}$ is $-ma = -m\ddot{x}$.

One definition of an inertial frame of reference is 'a frame of reference such that Newton's second law is satisfied', which is somewhat circular. Another definition is 'a frame of reference that travels with constant orientation and constant speed with respect to another inertial frame', which is still recursive but slightly more descriptive. Yet another description that one sometimes encounters is 'any frame in which the inertial power has the simplest expression'.

Exercise 1.1.1. Show that the first two definitions above are equivalent.

Writing down expressions for the power is a modelling task, and its difficulty can range from trivial to impossible.

1.1.3 No Net Power

A central postulate states

> *The total power of the internal forces equals the total power of the external forces.*

In other words, the sum of all forces has to vanish at all time, when summed with appropriate signs. Certain forces play a special role, since they act as a source of closure; the inertial force, for instance, is often available to provide whatever missing force is necessary for this equal-total-power postulate to hold. In such a situation one can also say that the inertial force equals the sum of the other forces. In other situations, like that of a gradient flow, it is not the inertial force of acceleration that closes the system, but a viscous or frictional force. We will see this below.

In the case of the particle with mass m undergoing a force f, there is no internal force, and two sources of external power: the inertial force, $m\ddot{x}$, with power $-m\ddot{x}\cdot v$, and the external force f, with power $f \cdot v$. The equal-total-power postulate then becomes

$$\forall v : \qquad -m\ddot{x} \cdot v + f \cdot v = 0,$$

in which we recoginize Newton's well-known second law,

$$m\ddot{x} = f.$$

Work is the integral of power along a path. If the system runs through a path $x(t)$, $t \in [0, T]$, while it experiences a force $f(t)$, then the work done by that force is

$$\int_0^T f(t) \cdot \dot{x}(t)\, dt.$$

Whenever f is constant, this reduces to $f \cdot (x(T) - x(0))$.

1.1.4 Conservative forces and energies

Many forces are *conservative*, meaning that they are a function of the state (i.e. $f = f(x)$), and can be written as the derivative of some other function (called a *potential* or an *energy*). The unit of energy is the Joule $\mathrm{J} = \mathrm{Nm} = \mathrm{kg\,m^2/s^2}$.

Mechanics provides many examples of such energies.

1. A *linear spring* has a linear relation between the force $f \in \mathbb{R}$ and the extension $x \in \mathbb{R}$, i.e. $f(x) = -kx$. Here the origin $x = 0$ is the *neutral position* of the spring, where $f = 0$; the sign follows from the convention that I use here that the force f is the force *exerted by* the spring. A 'reasonable' spring has $k > 0$.

   The force $f(x) = -kx$ is the (negative) derivative of the function $E(x) := \frac{k}{2}x^2$, which makes E an energy of the linear spring. Note that a *linear* spring has a *quadratic* energy. Note also that the positivity of k is related to the convexity (or concaveness, if $k < 0$) of E. This is one of many instances in which convexity properties of potentials relate to 'reasonability' properties of the underlying system. Other examples are the positivity of the thermal conductivity and of the viscosity, which follow from the entropy inequality (see e.g. [Mül85, Sec. 1.3.2.5]).

   Since f is measured in N and x in m, the dimension of the spring constant k is $\mathrm{N/m} = \mathrm{kg/s^2}$.

2. A *nonlinear* spring has a nonlinear force-displacement relationship $f(x)$. The function $E(x) = -\int_0^x f(s)\, ds$ is a corresponding energy.

3. An *electric charge* of charge q (with dimensions Coulomb, $\mathrm{C} = \mathrm{As}$) at the origin $0 \in \mathbb{R}^3$ gives rise to an electric field given by

   $$\frac{q}{4\pi\varepsilon_0}\frac{x}{|x|^3} \qquad \text{(with dimensions N/C)},$$

   where $\varepsilon_0 = 8.9 \cdot 10^{-12}\,\mathrm{C^2/Jm}$ is the *electric constant*. This field is the negative gradient $-\nabla V$ of the *electrostatic potential*

   $$V(x) := \frac{q}{4\pi\varepsilon_0}\frac{1}{|x|} \qquad \text{(with dimensions J/C)}.$$

The force experienced by a second charge, at x and with charge $\tilde{q}$ (which may be of the same sign as q, or of different sign) is given by

$$\frac{q\tilde{q}}{4\pi\varepsilon_0}\frac{x}{|x|^3} = -\tilde{q}\nabla V(x) \qquad \text{(with dimensions N)}.$$

Therefore the electrostatic force exerted by a charge q at 0 on a particle with charge $\tilde{q}$ at x is conservative, and generated by the potential $\tilde{q}V$,

$$\tilde{q}V(x) := \frac{q\tilde{q}}{4\pi\varepsilon_0}\frac{1}{|x|} \qquad \text{(with dimensions J)}.$$

Note that this potential $\tilde{q}V$ is symmetric in the roles of the two charges: if we interchange the roles of 0 and x, and of q and $\tilde{q}$, then the value of V does not change. Such a potential is called an *interaction potential*, and this particular example is called *Coulomb interaction*.

Note that an energy is only defined up to an additive constant. The following characterization of conservative fields $f : \mathbb{R}^n \to \mathbb{R}^n$ is well known:

Theorem 1.1.1. *Let $f \in C^1(\mathbb{R}^n; \mathbb{R}^n)$. The following are equivalent:*

1. *There exists $V : \mathbb{R}^n \to \mathbb{R}$ such that $f = -\nabla V$;*
2. *For any curve $x : [0,1] \to \mathbb{R}^n$, the integral $\int_0^1 f(x(t)) \cdot \dot{x}(t)\, dt$ only depends on $x(0)$ and $x(1)$ (but not on the connecting path);*
3. *For any closed curve $x : [0,1] \to \mathbb{R}^n$, $\int_0^1 f(x(t)) \cdot \dot{x}(t)\, dt = 0$.*

Exercise 1.1.2. Find an example that shows that the equivalence above does not hold if $\mathbb{R}^n$ is replaced by a set with a hole in it. Which implications do remain true?

Exercise 1.1.3. Show that the expression $\int_0^1 f(x(t)) \cdot \dot{x}(t)\, dt$ is invariant under time rescaling, i.e. if $t = \phi(s)$ with $\phi : [a,b] \to [0,1]$ being C^1 and bijective, then setting $y(s) := x(\phi(s))$ we have

$$\int_0^1 f(x(t)) \cdot \partial_t x(t)\, dt = \int_a^b f(y(s)) \cdot \partial_s y(s)\, ds.$$

Note how Theorem 1.1.1 connects the property that a force is conservative (condition 1) with properties of the power of that force (conditions 2 and 3).

Remark 1.1.2. Theorem 1.1.1 provides an interpretation of the energy V corresponding to a force field f:

The energy is the capacity of the force to do work.

This interpretation is true in the following sense. Assuming $f = -\nabla V$, we have for any two states $x, y \in \mathbb{R}^n$ and a connecting path $\hat{x} : [0,1] \to \mathbb{R}^n$ with $\hat{x}(0) = x$ and $\hat{x}(1) = y$ that

$$\int_0^1 f(\hat{x}(t)) \cdot \dot{\hat{x}}(t)\, dt = -V(\hat{x}(1)) + V(\hat{x}(0)) = V(x) - V(y).$$

Therefore the force f does work $V(x)-V(y)$ when the system moves from x to y—regardless of which path is taken, and also regardless of the speed along the path (see Exercise 1.1.3).

This also implies that $V(x) - V(y)$ is the *difference* in capacity to do work in going from x to any other point $z \in \mathbb{R}^n$, relative to the same capacity in going from y to z. The argument is as follows. Take two curves, $\hat{x}$ connecting x to z and $\hat{y}$ connecting y to z; the work done by f along these two curves is $V(x) - V(z)$ and $V(y) - V(z)$. The difference between these two amounts is exactly $V(x) - V(y)$. □

1.1.5 Dissipation

The mechanical interpretation of *dissipation* is the conversion of mechanical energy into heat.

Two important examples of dissipation are *friction* and *viscous dissipation*. Both depend on the velocity of the system.

- Typically a *frictional force* is parallel to the velocity but with opposite sign: $f = -\tilde{f}(|\dot{x}|)\dot{x}$, with $\tilde{f} \geq 0$. Note that such a force satifies $f \cdot \dot{x} = -\tilde{f}(|\dot{x}|)|\dot{x}|^2$, and therefore does not fulfill condition 3 of Theorem 1.1.1.

  A common model of friction is *Coulomb friction*, in which the frictional force f is has constant non-zero magnitude as soon as the velocity is non-zero. Its mathematical description in one-dimension is equivalent to the Heaviside graph,

$$f(\dot{x}) \in \begin{cases} \{-f_0\} & \text{if } \dot{x} < 0, \\ [-f_0, f_0] & \text{if } \dot{x} = 0, \\ \{f_0\} & \text{if } \dot{x} > 0, \end{cases} \qquad f_0 > 0 \text{ with dimension N.}$$

- *Viscous dissipation* is the dissipation of mechanical energy in viscous fluids. All fluids (liquids and gases) are viscous to some extent, although depending on the situation the viscosity might be dominant or not. Syrup is a good example from daily life of a fluid driven by viscous forces. At the other end of the scale, water

at room temperature and pressure is hardly viscous; this is nicely illustrated by the way milk tends to splash far and wide when it is spilt, which illustrates domination of inertial forces over viscous ones.

The source of viscous dissipation is friction between layers of fluid that slide over each other. The simplest models of viscous forces prescribe a linear relationship between the force and the velocity difference or velocity gradient. Such a fluid is called *Newtonian*. In one dimension this relationship is therefore

$$f = -\eta\, \dot{x},$$

where the parameter η is called the *viscosity*, with dimensions $\mathrm{Ns/m} = \mathrm{kg/s}$. In higher dimensions for instance *Stokes' law* gives the relationship between the velocity v of a sphere of radius $R\,[\mathrm{m}]$ and the force f required to pull it through a viscous fluid of dynamic viscosity $\eta\,[\mathrm{Ns/m^2}]$ as

$$f = 6\pi\eta R\, v. \tag{1.1}$$

Generalizations include power-law relationships and viscous forces that depend also on past values of the velocity.

The *power* of a viscous force is again the expression $f \cdot \dot{x}$; since viscous dissipation converts mechanical energy into heat, the rate of such dissipation is equal to minus the power of the viscous force. A linear (Newtonian) viscous force

$$f = -G\dot{x}, \qquad \text{with } G \in \mathbb{R}^{n\times n},$$

therefore gives rise to a dissipation

$$\text{viscous dissipation } = -f \cdot \dot{x} = \dot{x} \cdot G\dot{x}. \tag{1.2}$$

This expression depends quadratically on $\dot{x}$, and this property suggests defining an *inner product* on $\mathbb{R}^n$ by

$$(v, w)_G := v \cdot Gw \qquad \text{for } v, w \in \mathbb{R}^n. \tag{1.3}$$

This is indeed an inner product in the usual sense (symmetric, bilinear, and positive definite) whenever G is a symmetric and positive definite matrix.

Note that if the force does not depend linearly on the velocity, such as in the case of Coulomb friction, then the resulting power and dissipation are not quadratic, and there is no associated inner product.

Below we will mostly focus on the case of linear forces and quadratic dissipation, and use the word *dissipation* both for the quadratic object (1.2) and the induced inner product (1.3).

1.1.6 Example: a mass, a spring, and a dashpot

We can now formulate a classical example: the damped mass-spring system. In the simplest case, the setup is that of a mass that can move in one dimension, and which is connected to a spring and a *dashpot*, a friction element much like a shock absorber. The freedom of movement of the system is only the displacement of the mass $x \in \mathbb{R}$; the extension of the spring is also x.

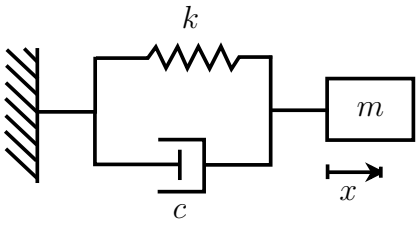

Figure 1.1: A damped mass-spring system

Assuming a linear spring with force $-kx$ and linear friction with force $-c\dot{x}$, there is no internal force, while the total external force is $-m\ddot{x} - kx - c\dot{x}$. The resulting equation is the well-known

$$m\ddot{x} + c\dot{x} + kx = 0.$$

Whether a force is internal or external also depends on the choice one makes. In the example of the mass-spring system above, if one considers the spring to be part of the surroundings, then it exerts an external force $-kx$ on the mass (i.e. in the opposite direction with respect to the displacement x); on the other hand, if one considers the spring part of the system, then it exerts an internal force kx on the wall it is connected to. Both lead to the same formula for the force balance.

If this seems confusing (and prone to mistake), I agree completely. The good news is that once one moves from this force/power-based modelling to variational modelling (which we do below), both the classification in internal and external power and the corresponding signs become completely automatic.

1.2 Duality

The reader familiar with the concept of duality, usually applied in infinite-dimensional spaces, might have noticed that the discussion above could be improved by formulating in terms of duality. We do that now.

1.2.1 Dual pairings

We recall some duality concepts. Two real normed spaces V and W are said to be *in duality* with each other if there exists a continuous bilinear form $\langle \cdot, \cdot \rangle : V \times W \to \mathbb{R}$, i.e. a mapping $\langle \cdot, \cdot \rangle : V \times W \to \mathbb{R}$ satisfying the linearity properties

- for fixed w, $\langle v, w \rangle$ is linear in v,
- for fixed v, $\langle v, w \rangle$ is linear in w,

and the estimate

$$|\langle v, w\rangle| \leq C\|v\|_V\|w\|_W.$$

Given a real normed space V, the canonical example of a dual space W is the *(topological) dual*, the space V' defined by

$$V' := \{\, \xi : V \to \mathbb{R} \ \text{ linear and continuous } \}.$$

The space V' is a linear space with respect to pointwise scalar multiplication and addition, and has the canonical norm

$$\|f\|_{V'} := \sup_{0\neq v\in V} \frac{\langle \xi, v\rangle}{\|v\|_V}.$$

Here we write $\langle \xi, v\rangle = {}_{V'}\langle \xi, v\rangle_V$ for the value of $\xi \in V'$ at $v \in V$, and this indeed defines a dual pairing in the definition above. One can think of the space V' as the *largest* space W that can be in duality with V.

1.2.2 Riesz' Lemma

In a *Hilbert space*, i.e. a complete vector space with an inner product $(\cdot,\cdot)_H$ and norm $\|z\|_H^2 := (z, z)_H$, dual objects can be mapped one-to-one to vectors in the space itself:

Lemma 1.2.1 (Riesz' Lemma)**.** *Let H be a Hilbert space. For each $\xi \in H'$ there exists a unique $w \in H$ such that*

$$ {}_{H'}\langle \xi, v\rangle_H = (v, w)_H \qquad \text{for all } v \in H.$$

In addition, $\|w\|_H = \|\xi\|_{H'}$.

See e.g. [Bre11, Th. 5.5]. Note that the vector w necessarily depends on the choice of inner product: a different inner product results in a different vector w, for the same dual object ξ.

1.2.3 The Fréchet Derivative

Suppose we have given a functional $F : V \to \mathbb{R}$. The principal way to define differentiation of F is via the *Fréchet derivative*. A functional $F : V \to \mathbb{R}$ is *Fréchet differentiable* at $v \in V$ if there exists $\xi \in V'$ such that

$$F(v + w) - F(v) - \langle \xi, w\rangle = o(\|w\|_V) \qquad \text{as } w \to 0. \tag{1.4}$$

From this definition it follows immediately that for each $w \in V$,

$$\lim_{t\to 0} \frac{F(v + tw) - F(v)}{t} = \langle \xi, w\rangle.$$

The dual object ξ is called the *Fréchet derivative* at v and is written $F'(v)$. The definition (1.4) gives a clear interpretation of this dual object, as follows: for given small perturbation w of the current position v, $\langle \xi, w\rangle$ gives the (approximate) value of $F(v+w)-F(v)$, i.e. the approximate perturbation of F.

1.2.4 Gradients

Given a Fréchet differentiable functional F on a Hilbert space H, the *gradient* $\operatorname{grad} F$ at $v \in H$ is the Riesz representation of the Fréchet derivative $F'(v)$, which is defined by the property that

$$(\operatorname{grad} F(v), w)_H = {}_{H'}\langle F'(v), w\rangle_H \qquad \text{for all } w \in H.$$

In $\mathbb{R}^n$ we usually write ∇F for $\operatorname{grad} F$.

1. In $\mathbb{R}^n$ with the standard inner product, $\operatorname{grad} F(v)$ is the vector $(\partial_1 F, \dots, \partial_n F)$ of partial derivatives.
2. In $\mathbb{R}^n$ with an inner product given by a symmetric positive definite matrix $G \in \mathbb{R}^{n\times n}$, the gradient $\operatorname{grad}_G F$ is the vector $(f_1, \dots, f_n)$ with coefficients

$$f_i = \sum_{j=1}^{n} (G^{-1})_{ij} \partial_j F(v),$$

   i.e. $\operatorname{grad}_G F = G^{-1} \operatorname{grad} F$, where we write $\operatorname{grad} F$ for the gradient with respect to the standard inner product.

Note that while the Fréchet derivative is independent of the choice of inner product, the gradient clearly depends on the inner product.

Stating that the Fréchet derivative is independent of the choice of inner product is actually only true within the class of inner products giving rise to equivalent norms. If one chooses an inner product which generates a stronger norm, then the condition that $\xi \in H'$ in (1.4) is also strengthened, implying a smaller class of Fréchet differentiable functionals.

1.2.5 Displacements, forces, and duality

We now switch to modelling again. Suppose that V is a space of displacements, for instance measured in meters. If $\mathcal{E} : V \to \mathbb{R}$ is an energy functional, then its Fréchet derivative is a *dual* object: according to the interpretation above it maps an (infinitesimal) change in displacement into an (infinitesimal) change in energy.

This suggests that a conservative force, which derives from an energy, should be viewed as a dual object; and by extension *all* forces should be thought of as dual objects. If we go back through the discussion of Section 1.1, we see that indeed the *power* is intrinsically a dual object: it maps a velocity (which is a change of the displacement per unit of time) to a change in energy per unit of time. The power can be viewed as the duality pairing corresponding to a given force.

One might ask, why do we then usually think of forces as vectors? The answer is that (a) vectors are more easily visualized, and (b) in $\mathbb{R}^n$, which is a Hilbert space, Riesz' Lemma above allows us to uniquely identify with each force (as a dual object) a corresponding 'force vector'. However, the resulting vector does depend on the choice of inner product, which is closely related to the choice of variables.

For forces in $\mathbb{R}^n$, such as arise for instance when studying rigid bodies, identifying dual objects with vectors works well, since it is usally not too restrictive to stick to a single frame of reference and a single inner product. For deformable bodies things already become more difficult: the existence of three different stress tensors in continuum elasticity, for instance (the Cauchy stress tensor and the first and second Piola-Kirchhoff stress tensors [GS08, Ch. 7]) arises from a need to describe some properties in the reference frame and others in the deformed frame. Each of the three tensors however corresponds to the *same* power duality pairing: their differences only arise from differences in expressing this object through different sets of local coordinates.

These considerations suggest that the power duality pairing should be considered as the principal object; other descriptions, such as vectorial forces or stress tensors, are representations of this object in a given set of coordinates.

1.3 Modelling gradient flows in $\mathbb{R}^n$

We now specialize to the case of conservative forces and where inertia is of vanishing importance. In this case the system is described by the potential $\mathcal{E}$ of the conservative forces and the power $(\cdot,\cdot)$ of the dissipative forces, and these are the *only* sources of power.

In this context modelling proceeds by choosing the degrees of freedom (see the discussion of *state* in Section 1.1.1) and then postulating expressions for the total energy $\mathcal{E}$ and the total dissipation $(\cdot,\cdot)$ of the system.

Let us illustrate this on the example of Section 1.1.6. Since we assume that there is no inertia, the system reduces to a spring and a dashpot. Again assuming a linear spring with force $-kx$, the energy of that spring is $\mathcal{E}(x) := kx^2/2$; the linear dashpot (with force $-c\dot{x}$) leads to a dissipation 'inner product' $(\cdot,\cdot)_c$ on $\mathbb{R}$ given by

$$(v,w)_c := cvw \qquad \text{for } v,w \in \mathbb{R}.$$

The power couple arising from the force $\mathcal{E}'$ is ${}_{(\mathbb{R})'}\langle \mathcal{E}'(x), y\rangle_{\mathbb{R}} = \mathcal{E}'(x)y = kxy$. 'No net power' then reduces to

$$0 = (\dot{x}, v)_c + \mathcal{E}'(x)v = (c\dot{x} + kx)v \qquad \forall v \in \mathbb{R},$$

or

$$\dot{x} = -\frac{k}{c}x.$$

Chapter 2

Gradient flows in Hilbert spaces

The discussion of the previous chapter tiptoed around questions of rigour. In this chapter we will explore some of the issues involved in developing a rigorous theory in a Hilbert space H. The main challenge lies in the fact that in all interesting cases the state space H is infinite-dimensional (typical examples are L^2 and H^{-1}) and $\mathcal{E} : H \to \mathbb{R}$ is *not* smooth—in fact, as we will see below, typical energy functions are not even defined on the whole of H.

Throughout this chapter H is a real Hilbert space with inner product $(\cdot, \cdot)$ and norm $\|\cdot\|$, or $(\cdot, \cdot)_H$ and $\|\cdot\|_H$ when necessary to avoid confusion. This space represents the state space, consisting of all admissible states, or in other words, it describes the *kinematics* of the system. In addition, we shall write $\operatorname{grad} \mathcal{E}$ instead of $\nabla \mathcal{E}$, to allow us to distinguish later between the 'function-space gradient' grad and the 'physical-space gradient' ∇. Both are gradients, in the same mathematical sense, but we will reserve ∇ for the operator that maps a function on $\mathbb{R}^d$ to a vector field on $\mathbb{R}^d$. In addition, by writing grad_H we will specify the space H (or more precisely the inner product $(\cdot, \cdot)_H$) with which the gradient is defined.

2.1 Mathematical properties of Hilbert-space gradient flows: the smooth case

The first main result states some simple properties that are simple to prove.

Theorem 2.1.1. *Assume that $\mathcal{E} \in C^1(H)$ and that the mapping $H \ni z \mapsto \mathcal{E}'(z) \in H'$ is locally Lipschitz continuous. Then*

1. *For each $z_0 \in H$ there exists a local solution $z \in C^1((-\delta, \delta); H)$ of the gradient-flow equation*
$$\dot{z} = -\operatorname{grad} \mathcal{E}(z) \qquad \text{with} \quad z(0) = z_0. \tag{2.1}$$
2. *If $\mathcal{E}$ is bounded from below, then this solution exists for all time $t > 0$.*

Proof. Part 1 follows from remarking that since $\|\operatorname{grad}\mathcal{E}(z)\|_H = \|\mathcal{E}'(z)\|_{H'}$ (see Lemma 1.2.1), $z \mapsto \operatorname{grad}\mathcal{E}(z)$ is also locally Lipschitz. The result then follows from standard ODE theory (e.g. [Kre89, Sec. 5.3]).

For Part 2 we estimate

$$\|z(t) - z_0\|^2 = \left\|\int_0^t \dot{z}\right\|^2 \le t\int_0^t \|\dot{z}\|^2 = -t\int_0^t (\dot{z}, \operatorname{grad}\mathcal{E}(z)) = t\big(\mathcal{E}(z_0) - \mathcal{E}(z(t))\big).$$

Since $\mathcal{E}(z(t))$ is bounded from below by a constant independent of $z(t)$, $\|z(t)\|^2$ can only increase linearly in time; therefore the solution can be continued for all positive time (see e.g. [CL72, Sec. 1.3]). □

2.2 Equivalent definitions

Theorem 2.2.1. *Assume that $\mathcal{E} \in C^1(H)$ and let $z \in C^1([0,T];H)$ be given. Then the following four statements are equivalent:*

1. $\dot{z} = -\operatorname{grad}\mathcal{E}(z)$ *on* $[0,T]$;
2. ***(Rayleigh principle I)*** *For each* $t \in [0,T]$,
$$\frac{1}{2}\|\dot{z}(t)\|^2 + (\operatorname{grad}\mathcal{E}(z(t)), \dot{z}(t)) \le \frac{1}{2}\|v\|^2 + (\operatorname{grad}\mathcal{E}(z(t)), v) \qquad \text{for all } v \in H;$$
3. ***(Rayleigh principle II)*** *For each* $t \in [0,T]$,
$$\frac{d}{dt}\mathcal{E}(z(t)) \le -\frac{1}{2}\|\dot{z}(t)\|^2 - \frac{1}{2}\|\operatorname{grad}\mathcal{E}(z(t))\|^2;$$
4. ***(Rayleigh principle III)***
$$\mathcal{E}(z(T)) - \mathcal{E}(z(0)) + \int_0^T \Big[\frac{1}{2}\|\dot{z}(t)\|^2 + \frac{1}{2}\|\operatorname{grad}\mathcal{E}(z(t))\|^2\Big]\,dt \le 0.$$

Proof. $1 \iff 2$: The expression

$$v \mapsto \frac{1}{2}\|v\|^2 + (\operatorname{grad}\mathcal{E}(z(t)), v) \tag{2.2}$$

is strictly convex and coercive, and therefore has a unique minimizer, which solves the Euler equation

$$0 = (v,w) + (\operatorname{grad}\mathcal{E}(z(t)), w) \qquad \text{for all } w \in H,$$

which is equivalent to $v = -\operatorname{grad}\mathcal{E}(z(t))$. Therefore $\dot{z}$ is the unique minimizer of (2.2) if and only if $\dot{z}(t) = -\operatorname{grad}\mathcal{E}(z(t))$.

$1 \Longrightarrow 3$: By the chain rule we have for any z the estimates

$$\begin{aligned}\frac{d}{dt}\mathcal{E}(z(t)) &= (\dot z(t), \operatorname{grad}\mathcal{E}(z(t))) \\ &\geq -\|\dot z(t)\|\,\|\operatorname{grad}\mathcal{E}(z(t))\| && \text{with equality iff } \dot z \text{ and } \operatorname{grad}\mathcal{E} \text{ are parallel,} \\ &\geq -\frac{1}{2}\|\dot z(t)\|^2 - \frac{1}{2}\|\operatorname{grad}\mathcal{E}(z(t))\|^2 && \text{with equality iff } \|\dot z\| = \|\operatorname{grad}\mathcal{E}\|.\end{aligned}$$

If z satisfies part 1 of the theorem, then the inequalities above are equalities, and part 3 is proved.

$3 \Longrightarrow 4$: This follows immediately by integration.

$4 \Longrightarrow 1$: If part 4 is satisfied, then the inequalities above are equalities for almost all t, and since $\dot z$ and $\operatorname{grad}\mathcal{E}$ are both continuous, these equalities hold for all t. Therefore z satisfies part 1. □

The simple proof of these equivalent definitions belies their tremendous importance. For instance, the 'Rayleigh Principle II' is the basis for a very influential generalization of the gradient-flow concept to metric spaces, started by DeGiorgi (see [DGMT80]) and which culminated in [AGS05]. We shall see this below. The integrated version ('III') turns out to be very useful in passing to the limit in gradient flows, as demonstrated by Sandier and Serfaty [SS04] and Stefanelli [Ste08] (see also [Ser09, AMP+11]). The Rayleigh Principle I has also been used for this purpose, especially in finite-dimensional and 'nearly finite-dimensional' situations [NO01, NO10, HN11].

2.3 Convex functions

The sections above assumed that $\mathcal{E}$ was regular. In most of the interesting cases this assumption is too strong, and we need to replace the condition of regularity by something else. It turns out that *convexity* is the natural replacement. Classical references for this section are Rockafellar [Roc72] for the finite-dimensional theory, and Ekeland-Temam [ET87] for infinite dimensions.

In the theory of convex functions it is useful to work with the set of extended reals $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$, with the obvious rules of calculation; the special case $\infty - \infty$ is undefined, while $0 \cdot (\pm\infty)$ is defined to be equal to zero. We also use the conventions $\sup \emptyset = -\infty$ and $\inf \emptyset = \infty$.

If the function $F : H \to \overline{\mathbb{R}}$ never takes the value[1] $-\infty$, then it is called *convex* provided

$$\text{for all } x, y \in H,\ \lambda \in (0,1): \qquad F(\lambda x + (1-\lambda)y) \leq \lambda F(x) + (1-\lambda)F(y). \tag{2.3}$$

The function is *strictly* convex if this inequality is strict for all $0 < \lambda < 1$. A convex function F is said to be *proper* when $F(x) > -\infty$ for all x and $F(x) < \infty$ for at least one x. The *domain* $\operatorname{dom} F$ is the set $\{x : F(x) < \infty\}$.

[1] This condition is necessary to avoid $\infty - \infty$ on the right-hand side of the inequality (2.3). For the general case convexity can be defined as convexity of the epigraph of F.

At each $x \in H$ we define the *subdifferential* $\partial F(x) \subset H'$,

$$\partial F(x) = \{\, \xi \in H' : F(y) \geq F(x) + \langle \xi, y - x \rangle \text{ for all } y \in H \,\}.$$

Lemma 2.3.1. *If F is proper and $F(x) = \infty$, then $\partial F(x) = \emptyset$. If F is differentiable at x with (Fréchet) derivative $F'(x)$, then $\partial F(x) = \{F'(x)\}$.*

Exercise 2.3.1. Prove this lemma.

Below follow a number of examples of proper convex functions. H is a real Hilbert space.

1. $F : H \to \mathbb{R}$, $F(x) := \|x\|^2$.
2. If $K \subset H$ is convex, then $F(x) := 0$ if $x \in K$, and $F(x) = +\infty$ if $x \notin K$ defines a convex function, called the *indicator function* and often denoted $\mathbb{I}_K$.
3. $F : \mathbb{R} \to \overline{\mathbb{R}}$, $F(x) = -\log x$ if $x > 0$, and $F(x) = +\infty$ if $x \leq 0$.
4. $F : \mathbb{R} \to \overline{\mathbb{R}}$, $F(x) = x \log x$ if $x > 0$, and $F(x) = +\infty$ if $x \leq 0$ (compare the difference with the previous one).
5. $F : \mathbb{R} \to \mathbb{R}$, $F(x) = e^x$.
6. If $f : \mathbb{R} \to \overline{\mathbb{R}}$ is convex, and $H = L^2(\Omega)$ for some $\Omega \subset \mathbb{R}^n$, then the function $F(u) := \int_\Omega f(u(x))\, dx$ is convex.
7. Extending the previous case, if $A : H \to H$ is a bounded linear operator, then the function $F(u) := \int_\Omega f((Au)(x))\, dx$ is convex.
8. Even further extending the example, we can also take the linear operator A to be *unbounded*. An example of this is the Dirichlet integral $\int |\nabla u|^2$ in L^2, which is the integral of the composition of $y \mapsto |y|^2$ and the unbounded operator $u \mapsto \nabla u$.

As an example of the calculation of the subdifferential, let us take the final example above, $F(u) = (1/2) \int_\Omega |\nabla u|^2$ in $H = L^2(\Omega)$. If $\xi \in \partial F(u)$, then $u \in \operatorname{dom} \partial F \subset \operatorname{dom} F$ (see part 1 of Theorem 2.3.2 below), so that $u \in \operatorname{dom} F = H^1(\Omega)$. Take $\varphi \in C^1(\overline{\Omega})$, and remark that

$$F(u + t\varphi) - F(u) = t \int \nabla u \nabla \varphi + \frac{t^2}{2} \int |\nabla \varphi|^2.$$

Therefore necessarily $\langle \xi, \varphi \rangle = \int \nabla u \nabla \varphi$; since this linear form is bounded on L^2, $\Delta u \in L^2(\Omega)$ and $\partial u / \partial n = 0$ in $L^2(\partial\Omega)$ (which is well-defined since $\Delta u \in L^2(\Omega)$). Then $\langle \xi, \varphi \rangle = -\int \varphi \Delta u$ for all $\varphi \in L^2(\Omega)$. Therefore

$$\partial F(u) := \begin{cases} \{\varphi \mapsto -\int \varphi \Delta u\} & \text{if } \Delta u \in L^2(\Omega) \text{ and } \dfrac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega, \\ \emptyset & \text{otherwise.} \end{cases}$$

Exercise 2.3.2. Calculate the subdifferentials for the other convex functions. (Hint for number 2: first consider K an interval in $\mathbb{R}$.)

The theorem below collects a number of useful properties of convex functions.

Theorem 2.3.2. *Let $F : H \to \mathbb{R}$ be convex and proper.*

1. *At each $x \in H$ the subdifferential $\partial F(x)$ exists (but may be empty); it is a weakly closed convex subset of H'. The domain* $\operatorname{dom} \partial F$ *is the set $\{x \in H : \partial F(x) \neq \emptyset\}$, and* $\operatorname{dom} \partial F \subset \operatorname{dom} F$.

2. *Convex functions have* monotone *subdifferentials: if $x, y \in \operatorname{dom} \partial F$, and $\xi \in \partial F(x)$, $\eta \in \partial F(y)$, then*
$$\langle \xi - \eta, x - y \rangle \geq 0. \tag{2.4}$$

Proof. The first part is a good exercise. For the monotonicity, note that the function $g(\lambda) := F(\lambda x + (1-\lambda)y) - \lambda F(x) - (1-\lambda)F(y)$ is convex on $[0,1]$ and vanishes at $\lambda = 0, 1$; therefore $g(\lambda) \leq 0$ for $\lambda \in [0,1]$. Since

$$F(\lambda x + (1-\lambda)y) \geq F(x) + \langle \xi, (1-\lambda)(y-x) \rangle,$$

we find

$$0 \geq g(\lambda) \geq (1-\lambda)\Big[F(x) - F(y) + \langle \xi, y - x \rangle\Big],$$

with equality at $\lambda = 1$, which implies

$$F(x) - F(y) + \langle \xi, y - x \rangle \leq 0.$$

A similar argument at $\lambda = 0$ gives

$$F(y) - F(x) + \langle \eta, x - y \rangle \leq 0,$$

and combining these two inequalities yields (2.4). □

2.4 Gradient flows of convex functionals

Brezis [Bre73] has laid the foundation of the theory of gradient flows of convex functionals in Hilbert spaces. The central existence and uniqueness theorem is the following.

Theorem 2.4.1 ([Bre73, Th. 3.1])**.** *Let H be a real Hilbert space, and let $\mathcal{E} : H \to \overline{\mathbb{R}}$ be convex and proper. Fix $u_0 \in H$. Then there exists a unique function $u \in C([0,\infty); H)$ such that*

1. *$u(0) = u_0$;*

2. *For almost every $t > 0$, $\dot{u}(t)$ exists in the classical sense;*

3. *There exists a function* $\xi : [0, \infty) \to H'$ *such that* $\xi(t) \in \partial\mathcal{E}(u(t))$ *for every* t*, and*

$$(\dot{u}(t), w)_H = -\langle \xi(t), w \rangle \qquad \text{for a.e. } t > 0 \text{ and for all } w \in H; \tag{2.5}$$

4. *At every* $t \geq 0$*,* $\xi(t)$ *is the element of minimal norm in* $\partial\mathcal{E}(u(t))$*;*

5. *At every* $t \geq 0$ *the right derivative* $\dot{u}^+(t)$ *exists, and this right derivative satisfies* (2.5) *at every* $t \geq 0$*.*

This theorem makes a number of interesting statements. For instance, the fact that $\xi(t) \in \partial\mathcal{E}(u(t))$ for all time t implies that $u(t) \in \operatorname{dom} \partial\mathcal{E}$, which can be interpreted as a regularity statement (indeed, see the example below of the diffusion equation).

Generalizations of Theorem 2.4.1 exist to functions that are not convex. For instance Brezis also considers the case where $\mathcal{E} = \mathcal{E}_0 + \mathcal{E}_1$, where $\mathcal{E}_0$ is convex and $\mathcal{E}_1 \in C^{1,1}(H)$, the class of functions with Lipschitz derivatives [Bre73, Th. 3.17]. Later authors generalize this to functions $\mathcal{E}$ that are just C^1 [DGMT80, MST89] (with additional conditions) or not even C^1 [RS06]. An important other direction of generalization is that where the Hilbert space H is replaced by a metric space, and we describe this in more detail in Chapter 3.

Property 4 of Theorem 2.4.1 implies that $\xi(t)$ is unique, since it is the unique projection of the origin onto the weakly closed convex set $\partial\mathcal{E}(u(t))$. This remark motivates the

Definition 2.4.2. *Whenever* $\partial\mathcal{E}(u)$ *is non-empty, the* gradient $\operatorname{grad}_H \mathcal{E}(u)$ *is defined to be the unique element of* H *satisfying*

$$(\operatorname{grad}_H \mathcal{E}(u), z)_H = \langle \xi, z \rangle \qquad \text{for all } z \in H,$$

where $\xi \in \partial\mathcal{E}(u)$ *is the unique subderivative of minimal norm:*

$$\|\xi\|_{H'} = \inf_{\tilde{\xi} \in \partial\mathcal{E}(u)} \|\tilde{\xi}\|_{H'}.$$

With this definition, equation (2.5) combined with statement 4 reduce to

$$\dot{u} = -\operatorname{grad}_H \mathcal{E}(u).$$

2.5 Examples

Throughout this section we use $H = L^2(\Omega)$, where $\Omega \subset \mathbb{R}^d$.

2.5.1 Zero-order functionals

Smooth convex functionals

Take $f : \mathbb{R} \to \overline{\mathbb{R}}$ convex and differentiable, and define $\mathcal{E}(u) := \int_\Omega f(u)$. (Exercise: what is the domain of $\mathcal{E}$?) The subdifferential of $\mathcal{E}$ is

$$\partial\mathcal{E}(u) := \begin{cases} \{\varphi \mapsto \int_\Omega \varphi f'(u)\} & \text{if } f'(u) \in H \\ \emptyset & \text{otherwise.} \end{cases}$$

Since $\partial\mathcal{E}$ is a singleton whenever it is non-empty, the equation (2.5) reduces to (with some abuse of notation)

$$(\dot{u}, w) = -\langle \partial\mathcal{E}(u), w\rangle \qquad \text{for all } w \in H, \tag{2.6}$$

which can also be written as

$$f'(u(t,\cdot)) \in L^2(\Omega) \quad \text{for all } t > 0, \text{ and} \tag{2.7a}$$

$$\int_\Omega \dot{u}(t,x)w(x)\,dx = -\int_\Omega w(x)f'(u(x))\,dx \qquad \text{for all } w \in L^2(\Omega),\ t > 0, \tag{2.7b}$$

which again implies

$$\dot{u} = -f'(u) \qquad \text{in } \Omega.$$

Note how (2.7a), which is a reformulation of $\operatorname{dom}\partial\mathcal{E}(u) \neq \emptyset$, may imply instantaneous regularization. For instance, if $f(u) = u^6$, then $f'(u) = 6u^5$, and $f'(u) \in L^2$ implies $u \in L^{10}$. Even if $u_0 = u(0,\cdot) \in L^2$, $u(t,\cdot)$ will be in L^{10} for all $t > 0$.

Functionals with corners

We now consider the example of $f(u) = |u|$. The subdifferential of f is

$$\partial f(u) := \begin{cases} \{+1\} & \text{if } u > 0, \\ [-1,1] & \text{if } u = 0, \\ \{-1\} & \text{if } u < 0, \end{cases}$$

which implies that the subdifferential of $\mathcal{E}(u) := \int f(u)$ is the set of all mappings

$$v \mapsto \int_\Omega \lambda(x)v(x)\,dx,$$

such that $\lambda \in L^2(\Omega)$ satisfies $\lambda(x) \in \partial f(u(x))$, i.e.

$$\lambda(x) = \pm 1 \text{ if } u(x) \gtrless 0 \quad \text{and} \quad \lambda(x) \in [-1,1] \text{ for every } x \text{ with } u(x) = 0. \tag{2.8}$$

Exercise 2.5.1. Prove this characterization of $\partial\mathcal{E}$.

Then (2.6) can be written as follows:

Find $u, \lambda : [0,\infty) \to L^2(\Omega)$, where λ satisfies (2.8), such that

$$\dot{u}(t,x) = -\lambda(t,x) \qquad \text{for a.e } t > 0,\ x \in \Omega.$$

2.5.2 The diffusion equation

Now we let $\Omega \subset \mathbb{R}^d$ be bounded, and we take $\mathcal{E}$ to be the Dirichlet integral, $\mathcal{E}(u) = \frac{1}{2}\int_\Omega |\nabla u|^2$. The subdifferential was already calculated on page 16, and as above it is a singleton when non-empty. The equation (2.6) then reduces to

$$\Delta u(t,\cdot) \in L^2(\Omega) \quad \text{and} \quad \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega \quad \text{for all } t > 0, \text{ and} \tag{2.9a}$$

$$\int_\Omega \dot{u} w = \int w \Delta u \qquad \text{for all } w \in L^2(\Omega),\ t > 0, \tag{2.9b}$$

which implies

$$\dot{u} = \Delta u, \quad \text{with boundary condition } \frac{\partial u}{\partial n} = 0. \tag{2.10}$$

Again the condition (2.9a) that follows from $\operatorname{dom} \partial\mathcal{E}(u) \neq \emptyset$ gives an instant regularization, which now is of the form $u(t,\cdot) \in H^2(\Omega)$ for all $t > 0$. (Of course we know that the diffusion equation even has a much stronger regularizing effect: $u(t,\cdot) \in C^\infty$ for all $t > 0$, as proved, e.g., in [Eva02, Sec. 2.3]. However, that result uses various properties of the Laplacian, such as linearity and translation-invariance, that Theorem 2.4.1 does not use).

Note how the boundary condition $\partial u/\partial n = 0$ arises: the non-emptiness of the subdifferential $\partial\mathcal{E}(u)$ requires the boundedness of the linear form $v \mapsto \int \nabla u \nabla v$ *in the space* $L^2(\Omega)$, and it is this condition that forces the normal derivative of u to vanish. Compare this to the case of 'natural' boundary conditions in stationary problems, which arise when minimizing (for instance) the functional

$$u \mapsto \frac{1}{2}\int |\nabla u|^2 + \int fu$$

over $L^2(\Omega)$ (where we add an additional term to prevent the problem from being trivial). Here it is the *vanishing* of the derivative $v \mapsto \int \nabla u \nabla v + \int fv$ (which of course implies boundedness in L^2) that leads to the boundary condition $\partial u/\partial n = 0$. This is an example of many aspects in which gradient flows are close to minimization problems.

Exercise 2.5.2. Work out the differential equation generated by the $L^2(\Omega)$-gradient flow of the following energies:

1. $\mathcal{E}(u) := \begin{cases} \frac{1}{2}\int_\Omega |\nabla u|^2 & \text{if } u \in H^1(\Omega) \text{ and } u = 0 \text{ on } \partial\Omega, \\ +\infty & \text{otherwise;} \end{cases}$
2. $\mathcal{E}(u) := \frac{1}{2}\int_\Omega |\nabla u|^2 + \int_\Omega f(u)$, where $f : \mathbb{R} \to \mathbb{R}$ is convex and differentiable;
3. $\mathcal{E}(u) := \frac{1}{2}\int_\Omega (\Delta u)^2$;
4. $\mathcal{E}(u) := \begin{cases} \frac{1}{2}\int_\Omega (\Delta u)^2 & \text{if } u \in H^2(\Omega) \text{ and } u = 0 \text{ on } \partial\Omega, \\ +\infty & \text{otherwise;} \end{cases}$

5. $\mathcal{E}(u) := \begin{cases} \frac{1}{2}\int_\Omega (\Delta u)^2 & \text{if } u \in H^2(\Omega) \text{ and } u = \partial u/\partial n = 0 \text{ on } \partial\Omega, \\ +\infty & \text{otherwise;} \end{cases}$

6. $\mathcal{E}(u) := \begin{cases} \frac{1}{2}\int_\Omega (\Delta u)^2 & \text{if } u \in H^2(\Omega) \text{ and } \partial u/\partial n = 0 \text{ on } \partial\Omega, \\ +\infty & \text{otherwise;} \end{cases}$

7. $\mathcal{E}(u) := \dfrac{1}{2} \displaystyle\iint_{\mathbb{R}^d \times \mathbb{R}^d} u(x)u(y)\kappa(x-y)\,dxdy$, where $\Omega = \mathbb{R}^d$ and $\kappa \in L^1(\mathbb{R}^d)$ with $\kappa \geq 0$.

Exercise 2.5.3. Repeat the previous exercise with the weighted space $L^2_c(\Omega)$, which is generated by the inner product

$$(u, v)_c := \int_\Omega c(x)\, u(x)v(x)\, dx,$$

where $c : \Omega \to \mathbb{R}$ is measurable and bounded from above and away from zero.

Chapter 3

Gradient flows in the space of measures

The recent interest in gradient flows is fueled by the discovery of gradient flows in Wasserstein metrics on the space of probability measures. One way to understand the relevance of these metrics is as follows. Let Z be a convex subset of $\mathbb{R}^d$. Define an embedding e of Z into $\mathcal{P}(Z)$, the space of probability measures on Z, by setting

$$e(z) := \delta_z, \tag{3.1}$$

where δ_z is the Dirac delta measure at z. Then the Wasserstein distance d_p of order p, which we define below, has the property that it *exactly reproduces the geometry of* Z, as embodied by the distance between two points:

$$d_p(\delta_{z_1}, \delta_{z_2}) = |z_1 - z_2| \qquad \text{for all } z_1, z_2 \in Z.$$

This property is the first indication that Wasserstein metrics are natural distance concepts on the space of probability measures. We will discuss many others below.

3.1 Measures

Let Z be a convex subset of $\mathbb{R}^d$. A 'measure' will be a Borel measure, i.e. a countably additive, $\overline{\mathbb{R}}$-valued set function defined on the Borel σ-field $\mathcal{B}(Z)$ on Z. $\mathcal{M}(Z)$ is the space of such measures, which therefore need neither be finite nor signed. The *total variation norm* $\|\mu\|_{TV}$ of a measure μ is

$$\|\mu\|_{TV} := \sup \sum_i |\mu(A_i)|, \tag{3.2}$$

where the supremum is taken over all finite collections of disjoint sets $A_i \subset Z$. The space $\mathcal{P}(Z)$ of *probability measures* is the space of nonnegative measures μ such that $\mu(Z) = 1$.

Given the embedding $e : Z \to \mathcal{P}(Z)$ described above, it is natural to ask for types of convergence that preserve the convergence in Z, i.e. such that $z_n \to z$ in Z implies $\delta_{z_n} \to \delta_z$. *Strong* convergence, that is convergence in the norm $\|\cdot\|_{TV}$, does not qualify, since $\|\delta_z - \delta_y\|_{TV} = 2$ whenever $z \neq y$. Various kinds of weak convergence do, and in these notes we will mostly use *narrow* convergence. Narrow convergence is convergence in duality with $C_b(Z)$, the space of continuous and bounded functions on Z, and is often denoted by $\Rightarrow$:

$$\mu_n \Rightarrow \mu \qquad \Longleftrightarrow \qquad \int \varphi \, d\mu_n \to \int \varphi \, d\mu \quad \text{for all } \varphi \in C_b(Z).$$

$\mathcal{P}(Z)$ is closed under narrow convergence, i.e. if $\mu_n \in \mathcal{P}(Z)$ and $\mu \in \mathcal{M}(Z)$, and $\mu_n \Rightarrow \mu$, then $\mu \in \mathcal{P}(Z)$.

Exercise 3.1.1. Prove this last statement.

The narrow topology is separable [AGS05, Remark 6.1.2] and is induced by a metric [AGS05, Remark 6.1.1], and therefore sequential compactness is the same as regular compactness. Note that if $z_n \to z$ in Z, then $\delta_{z_n} \Rightarrow \delta_z$:

$$\int \varphi \, d\mu_n - \int \varphi \, d\mu = \varphi(z_n) - \varphi(z) \to 0.$$

A useful characterization of narrowly precompact sets of probability measures is

Theorem 3.1.1 (Prokhorov's theorem [Bil99, Section 1.5])**.** *A set $F \subset \mathcal{P}(Z)$ is narrowly precompact if and only if it is* tight*, i.e.*

$$\forall \varepsilon > 0 \ \exists K \Subset Z : \ \sup_{\mu \in F} \mu(Z \setminus K) < \varepsilon.$$

3.2 Multi-particle systems

We now extend the embedding e in (3.1) to multi-particle systems. Given N particles, each with a state z_i in Z, we define the N-vector $\boldsymbol{z} = (z_1, \ldots, z_N) \in Z^N$. The corresponding embedding $e_N : Z^N \to \mathcal{P}(Z)$ is

$$e_N(\boldsymbol{z}) := \frac{1}{N} \sum_{i=1}^{N} \delta_{z_i}.$$

Given two measures of this form, $\mu = N^{-1} \sum \delta_{x_i}$ and $\nu = N^{-1} \sum \delta_{y_i}$, a natural distance between these two measures is the one inherited from Z,

$$d_1(\mu, \nu) := \inf \left\{ \frac{1}{N} \sum_{i=1}^{N} |x_i - y_{\sigma(i)}| : \ \sigma \text{ a permutation} \right\}. \tag{3.3}$$

Note that the infimum over permutations is necessary to prevent renumbering of the x_i and y_i from resulting in a different distance.

In fact the right-hand side of (3.3) is the *canonical* metric on the space Z^N/S_N of N-vectors modulo permutation, since it is the minimal distance between two elements of the two equivalence classes—one equivalence class being permutations of $\boldsymbol{z}$, the other permutations of $\boldsymbol{y}$. Defining $d_1(\mu,\nu)$ as in (3.3) simply translates that canonical metric into the language of probability measures. Or to put it differently,

> *The distance d_1 exactly captures* ***multi-particle Euclidean geometry modulo permutations****.*

Yet another way to say the same thing is that under the mapping e_N the N particles become *indistinguishable* from each other. This indistinguishability will manifest itself in various forms below.

3.3 The Wasserstein distances

We now generalize the definition of d_1 in (3.3) to the case of general measures and to general $p \in [1,\infty)$. First note that for any permutation σ, by defining $\gamma := N^{-1}\sum_{i=1}^N \delta_{x_i,y_{\sigma(i)}}$ the expression in braces in (3.3) can be written as

$$\iint\limits_{Z\times Z} |x-y|\,\gamma(dxdy),$$

while this γ has *marginals* μ and ν, i.e.

$$\gamma(A\times Z)=\mu(A) \text{ and } \gamma(Z\times A)=\nu(A) \quad \text{for all } A\subset\mathcal{B}(Z). \tag{3.4}$$

This formulation only depends on μ and ν being probability measures, giving rise to a natural generalization, that of the *Wasserstein distances*. These will be defined on the set of probability measures with finite moment of order $p\in[1,\infty)$,

$$\mathcal{P}_p(Z) := \Big\{\ \mu\in\mathcal{P}(Z) : \int_Z |x|^p\,\mu(dx)<\infty\ \Big\}.$$

Definition 3.3.1. *The* Wasserstein distance of order $p\in[1,\infty)$, $d_p:\mathcal{P}_p(Z)\times\mathcal{P}_p(Z)\to\mathbb{R}$, *is defined by*

$$d_p(\rho_0,\rho_1)^p = \inf_{\gamma\in\Gamma(\rho_0,\rho_1)} \iint\limits_{Z\times Z} |x-y|^p\,\gamma(dxdy), \tag{3.5}$$

where Γ *is the set of* couplings *of* ρ_0 *and* ρ_1*,*

$$\Gamma(\rho_0,\rho_1) := \Big\{\ \gamma\in\mathcal{P}(Z\times Z) :\ \gamma \textit{ has } \text{marginals } \rho_0 \textit{ and } \rho_1\Big\}.$$

Exercise 3.3.1. Show that if $\rho_{0,1} \in \mathcal{P}_p(Z)$, then

- $\Gamma(\rho_0, \rho_1)$ is non-empty;
- $\Gamma(\rho_0, \rho_1) \subset \mathcal{P}_p(Z \times Z)$;
- $d_p(\rho_0, \rho_1) < \infty$;
- The minimization problem (3.5) has a solution.

Let us verify that Definition 3.3.1 reduces to (3.3) when $p = 1$ and $\rho_{0,1} \in e_N(Z^N)$. Writing

$$\rho_0 = \frac{1}{N}\sum_{i=1}^{N} \delta_{x_i} \quad \text{and} \quad \rho_1 = \frac{1}{N}\sum_{i=1}^{N} \delta_{y_i},$$

each $\gamma \in \Gamma(\rho_0, \rho_1)$ has support contained in the set of points of the form (x_i, y_j); in addition, each x_i is to be paired with exactly one y_j, which makes the support of γ equivalent to a permutation σ, i.e.

$$\gamma = \frac{1}{N}\sum_{i=1}^{N} \delta_{(x_i, y_{\sigma(i)})}.$$

Then the definition (3.3.1) for $p = 1$ reduces to (3.3).

Theorem 3.3.2 ([AGS05, Remark 8.1.5-7])**.** *The Wasserstein distance d_p is a complete separable metric on $\mathcal{P}_p(Z)$. If $\rho_n, \rho \in \mathcal{P}_p(Z)$, then $d_p(\rho_n, \rho) \to 0$ if and only if*

1. *ρ_n converges narrowly to ρ, and*
2. $\displaystyle\int_Z |x|^p \, \rho_n(dx) \to \int_Z |x|^p \rho(dx)$.

3.4 Tangents—part 1

In the previous section we discussed how d_p reproduces N-particle Euclidean distance in the language of probability measures (more precisely, convex combinations of delta functions). Euclidean geometry also includes inner products, and a natural question is whether some kind of inner product structure can be recognized in the geometry given by d_p. In this section we explore this.

In Riemannian geometry, inner products are associated with tangent vectors, and therefore with derivatives of smooth curves. Assume that $t \mapsto \boldsymbol{z}(t)$ is a smooth curve in Z^N; what does this mean for $t \mapsto e_N(\boldsymbol{z}(t))$? The following lemma gives an answer in a special case, where the velocity v_i of the particle at z_i can be written as $v(z_i)$. Note that this is a natural assumption: by indistinguishability it is reasonable to let v_i depend not on i but only on z_i.

Lemma 3.4.1. *Let $Z \subset \mathbb{R}^d$, let $v : Z \to \mathbb{R}^d$ be a Borel measurable vector field, and let $\dot{z}_i(t) = v(z_i(t))$ for $i = 1, \dots N$ and $t \in (-a, a)$. Set $\rho_N(t) := e_N(\boldsymbol{z}(t))$. Then*

$$\partial_t \rho_N + \operatorname{div}(\rho_N v) = 0 \qquad \text{in } \mathcal{D}'((-a,a) \times \mathbb{R}^d). \tag{3.6}$$

Proof. The equality (3.6) means that for each $\varphi \in C_c^\infty((-a,a) \times \mathbb{R}^d))$,

$$\int_{-a}^{a} \int_{\mathbb{R}^d} \left[\partial_t \varphi + v \cdot \nabla \varphi\right] d\rho_N(t)\, dt = 0. \tag{3.7}$$

This is easily seen to be true, since for each i,

$$\begin{aligned} 0 &= \int_{-a}^{a} \frac{d}{dt} \varphi(t, z_i(t))\, dt = \int_{-a}^{a} \left[\partial_t \varphi(t, z_i(t)) + v(z_i(t)) \cdot \nabla \varphi(t, z_i(t))\right] dt \\ &= \int_{-a}^{a} \int_{\mathbb{R}^d} \left[\partial_t \varphi + v \cdot \nabla \varphi\right] d\delta_{z_i(t)}\, dt, \end{aligned}$$

after which (3.7) follows by summing over i. □

This result suggests that there might be a relationship between the derivative along a curve of measures and a corresponding vector field (called v above) on Z. This is indeed the case, but before we can explore this in more detail we need to introduce some metric-space machinery.

3.5 Absolutely continuous curves in metric spaces

Let (M, d) be a metric space, and let $f : [0, T] \to M$ be given. Metric spaces need have no differentiable structure, and therefore the 'derivative' of f need not make sense. On the other hand, for each fixed $x \in M$ the mapping $t \mapsto d(x, f(t))$ is a real-valued function, and one can wonder under what conditions this function has some minimal differentiability properties. The answer lies in the concept of *absolute continuity*.

Definition 3.5.1. *f is* absolutely continuous *(a.c.) if*

$$\begin{aligned} \limsup_{h \downarrow 0} \Bigg\{ \sum_{i=1}^{n} d(f(s_i), f(t_i)) : \\ n \in \mathbb{N},\ 0 \le s_1 < t_1 \le s_2 < t_2 \le \cdots \le s_n < t_n \le T,\ \sum_{i=1}^{n} (t_i - s_i) < h \Bigg\} = 0. \end{aligned}$$

We write $f \in AC([0, T]; (M, d))$.

Lemma 3.5.2 ([AGS05, Th. 2.1.2])**.** *If $f \in AC([0, T]; (M, d))$, then at almost every $t \in [0, T]$ the* metric derivative *exists:*

$$\left|\dot{f}\right|(t) := \lim_{h \to 0} \frac{d(f(t+h), f(t))}{|h|}.$$

Also,

1. $t \mapsto \left|\dot{f}\right|(t) \in L^1(0,T)$;

2. $d(f(s), f(t)) \leq \int_s^t \left|\dot{f}\right|(\tau)\, d\tau$ *for all* $0 \leq s < t \leq T$.

Conversely, if $w \in L^1(0,T)$ *exists such that* $d(f(s), f(t)) \leq \int_s^t w(\tau)\, d\tau$, *then* f *is absolutely continuous and* $w \geq \left|\dot{f}\right|$.

Lemma 3.5.3 ([AGS05, Remark 2.1.3])**.** *If* $(M,d) = (\mathbb{R}, |\cdot|)$, *then* f *is absolutely continuous if and only if the distributional derivative* $\partial_t f$ *is an element of* $L^1(0,T;\mathbb{R})$. *Then* $f(t) - f(s) = \int_s^t \partial_t f$, f' *(the classical derivative) exists at almost every* t, *and* $f' = \partial_t f$.

In other words, $AC(\mathbb{R}) = W^{1,1}(\mathbb{R})$. This lemma shows how the set of absolutely continuous functions f is the largest set that satisfies the two properties

1. f has a pointwise classical derivative f' almost everywhere;

2. f is the integral of f'.

Note that the first point does not imply the second, as the example of the Heaviside function shows. Even when adding the condition of continuity there remain functions that are continuous, a.e. differentiable, but not recoverable by integration; the 'devil's staircase' [GO03, p. 96] is an example.

3.6 Tangents—part 2

In Section 3.4 we established a connection between the evolution of a sum of delta functions and a vector field v defined on the support points. This connection generalizes to general elements of $\mathcal{P}_2(\mathbb{R}^d)$, as the following theorem shows.

We adopt the notation ρ_t for the time-slice of the measure ρ at time t, and similarly for other time-dependent quantities. We define $W_0^{1,2}(\rho_t)$ as the completion of $C_c^\infty(\mathbb{R}^d)$ with respect to the norm $\|u\|^2 := \int |\nabla u|^2\, d\rho_t$.

Theorem 3.6.1 ([AGS05, Th. 9.3.2])**.** *Let* $\rho : (a,b) \to \mathcal{P}_2(\mathbb{R}^d)$ *be absolutely continuous with respect to* d_2, *and let* $|\dot{\rho}| \in L^1(a,b)$ *be its metric derivative, as defined above. Then there exists a Borel vector field* $v : (t,x) \mapsto v_t(x)$ *such that* $v_t \in L^2(\rho_t)$ *and* $\|v_t\|_{L^2(\rho_t)} \leq |\dot{\rho}|(t)$ *for a.e.* t, *and*

$$\partial_t \rho + \operatorname{div} \rho v = 0 \qquad \text{in } \mathcal{D}'((a,b) \times \mathbb{R}^d). \tag{3.8}$$

Moreover, for a.e. t, $v_t \in \overline{\{\nabla p : p \in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\rho_t)}$.

Conversely, if $\rho : (a,b) \to \mathcal{P}_2(\mathbb{R}^d)$ *is narrowly continuous and satisfies* (3.8) *for some Borel vector field* v *with* $t \mapsto \|v_t\|_{L^2(\rho_t)} \in L^1(a,b)$, *then* $t \mapsto \rho_t$ *is absolutely continuous, and* $|\dot{\rho}|(t) \leq \|v_t\|_{L^2(\rho_t)}$.

This theorem establishes a one-to-one relationship, in the case of absolutely continuous curves, between distributional derivatives $\partial_t \rho$ and vector fields $v_t \in L^2(\rho_t)$. We write

$$s \in \mathrm{Tan}_\rho \quad \Longleftrightarrow \quad s + \operatorname{div} \rho v = 0 \quad \text{for some } v \in L^2(\rho).$$

Since we often will be concerned with that velocity field v that lies in $\overline{\{\nabla p\}}^{L^2(\rho)}$, we also define

$$\mathrm{Tan}_\rho^v := \overline{\{\nabla p : p \in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\rho)}.$$

Exercise 3.6.1. Show that the space Tan_ρ^v is orthogonal in $L^2(\rho)$ to the set of functions w satisfying $\operatorname{div} \rho w = 0$. Derive the characterization

$$v \in \mathrm{Tan}_\rho^v \quad \Longleftrightarrow \quad \|v\|_{L^2(\rho)} = \inf\Big\{ \|v + w\|_{L^2(\rho)} : \operatorname{div} \rho w = 0 \Big\}.$$

Show that for given $s = -\operatorname{div} \rho v$ the associated vector field v in Tan_ρ^v is unique.

In these notes we define Tan_ρ as distributions of the form $\operatorname{div} \rho v$; they are exactly those distributions that can be obtained as $\partial_t \rho$ for some absolutely continuous curve ρ. This definition mirrors the concept of a tangent in Riemannian geometry. By contrast, in [AGS05] the authors prefer to think of Tan_ρ^v as the tangent space, rather than Tan_ρ (and they use the notation Tan_ρ for what we call Tan_ρ^v). The two points of view are mathematically equivalent, and usually no confusion results.

3.7 The Benamou-Brenier formula

In a Riemannian manifold $\mathcal{M}$, the distance $d(x_0, x_1)$ between two points $x_{0,1} \in \mathcal{M}$ can be written in terms of the local metric tensor g as either

$$d(x_0, x_1) = \inf \left\{ \int_0^1 \sqrt{g_{x(t)}(\dot{x}(t), \dot{x}(t))}\, dt : x \in C^1([0,1]; \mathcal{M}),\ x(0) = x_0,\ x(1) = x_1 \right\}$$

or

$$d(x_0, x_1)^2 = \inf \left\{ \int_0^1 g_{x(t)}(\dot{x}(t), \dot{x}(t))\, dt : x \in C^1([0,1]; \mathcal{M}),\ x(0) = x_0,\ x(1) = x_1 \right\}. \quad (3.9)$$

Exercise 3.7.1. Prove that the two are equivalent.

In $\mathcal{P}_2(\mathbb{R}^d)$ a similar formula holds [BB00]:

$$d_2(\rho^0, \rho^1)^2 = \inf \left\{ \int_0^1 \|v_t\|^2_{L^2(\rho_t)}\, dt :\ \partial_t \rho + \operatorname{div} \rho v = 0,\ \rho_{0,1} = \rho^{0,1} \right\}. \quad (3.10)$$

3.8 Tangents—part 3

The previous sections suggest a defintion of *inner product* on the tangent space Tan_ρ:

$$(s_1, s_2)_\rho := \int_{\mathbb{R}^d} v_1 \cdot v_2 \, d\rho, \tag{3.11}$$

where $v_i \in \mathrm{Tan}^v_\rho$ is associated with the tangent s_i, i.e. $s_i + \operatorname{div} \rho v_i = 0$ (and this association is unique—see Exercise 3.6.1). With this definition the Benamou-Brenier formula (3.10) becomes identical to (3.9), and it also coincides with the inner product in $(\mathbb{R}^d)^N$ of two vectors $\boldsymbol{v}^1$ and $\boldsymbol{v}^2$, scaled with $1/N$,

$$(\boldsymbol{v}^1, \boldsymbol{v}^2)_{Z^N} = \frac{1}{N} \sum_{i=1}^N v_i^1 \cdot v_i^2.$$

Note that the choice of tangent representative is important in (3.11): if we replace v_i by a different $\tilde{v}_i$, still satisfying $s_i + \operatorname{div} \rho \tilde{v}_i = 0$, then the value of the inner product would change.

In which sense can we consider this an inner product? One sense is the following. In a Euclidean space X, if we fix $x, \hat{x} \in X$, and a 'tangent' vector $s \in X$, then

$$\frac{d}{dt} \|\hat{x} - (x + ts)\|_X \Big|_{t=0} = \left(\frac{\hat{x} - x}{\|\hat{x} - x\|}, s \right),$$

as can be directly verified by expanding the norm. This formula can even be used to define the inner product. In Riemannian geometry, we have the similar statement,

$$\frac{d}{dt} d(\hat{x}, x^t) \Big|_{t=0} = g_x(\hat{s}, s),$$

where now s is a unit-length ($g_x(s,s) = 1$) tangent vector at x and x^t is a unit-speed geodesic with $x^0 = x$, $\partial_t x^t|_{t=0} = s$; $\hat{s}$ is the unit-length tangent at x of the geodesic connecting x and $\hat{x}$ (provided this geodesic connection is unique and smoothly varying with x). And indeed, this property also holds for the Wasserstein distance d_2 and the inner product (3.11) [AGS05, Th. 9.4.7].

3.9 Differentiation of energies defined on the Wasserstein space

Given a smooth functional F *defined on* $L^2(\mathbb{R}^d)$ (rather than $\mathcal{P}_2(\mathbb{R}^d)$), and given a point $u \in L^2(\mathbb{R}^d)$ and a tangent $v \in L^2(\mathbb{R}^d)$, the natural definition of the derivative of F at u in the direction v is given by the Fréchet derivative, in the sense that (see Section 1.2.3)

$$\lim_{t \to 0} \frac{F(u + tv) - F(u)}{t} \;=\; {}_{L^2(\mathbb{R}^d)'}\langle F'(u), v \rangle_{L^2(\mathbb{R}^d)},$$

or using the concept of the L^2-gradient of F (Section 1.2.4),

$$\lim_{t\to 0}\frac{F(u+tv)-F(u)}{t} \;=\; (\mathrm{grad}_{L^2}F(u),v)_{L^2(\mathbb{R}^d)}.$$

The sets $L^2(\mathbb{R}^d)$ and $\mathcal{P}_2(\mathbb{R}^d)$ share the common subset of nonnegative functions in $C_c^\infty(\mathbb{R}^d)$, and as it turns out, this subset gives the key to calculating derivatives in $\mathcal{P}_2(\mathbb{R}^d)$, at least formally. Given a functional F now on $\mathcal{P}_2(\mathbb{R}^d)$, a point $\rho\in C_c^\infty(\mathbb{R}^d)$, and a tangent $s\in \mathrm{Tan}_\rho\cap C_c^\infty(\mathbb{R}^d)$, we write

$$\lim_{t\to 0}\frac{F(\rho+ts)-F(\rho)}{t} \;=\; (\mathrm{grad}_{L^2}F(\rho),s)_{L^2(\mathbb{R}^d)}.$$

This is of course still formal, since $\rho+ts$ need not be non-negative for any $t\neq 0$. However, it will serve us perfectly well for the formal discussion of these notes.

The appropriate *rigorous* concept of derivative in the case of $\mathcal{P}_2(\mathbb{R}^d)$ is more involved. To start with, as in the case of infinite-dimensional Hilbert spaces, most interesting functionals are not smooth and defined only in a small subset. As in the case of infinite-dimensional Hilbert spaces, the solution to this problem lies in appropriate concepts of convexity (now convexity along geodesics of the Wasserstein space $\mathcal{P}_2(\mathbb{R}^d)$) and subdifferentiability. See [AGS05, Ch. 10] for a full treatment.

3.10 Modelling the dissipation

The Wasserstein inner product $(s_1,s_2)_\rho$ can be interpreted as an embedding of the $\mathbb{R}^d$-inner product $v_1\cdot v_2$ in $\mathcal{P}_2(\mathbb{R}^d)$. We now discuss how this concept can be made dimensionful, and we also discuss how the choice of inner product on $\mathbb{R}^d$ propagates in the definition of dissipation.

In the case of a dimensionful (dissipation) inner product $(v_1,v_2)_G$ on $\mathbb{R}^d$ with dimension W, given by a positive definite symmetric matrix $G\in\mathbb{R}^{d\times d}$ (see Section 1.1.5), the natural corresponding dissipation inner product is

$$(s_1,s_2)_{\rho,G}=\int (v_1,v_2)_G\,d\rho=\int v_1\cdot Gv_2\,d\rho, \tag{3.12}$$

also with dimension W. Here $v_{1,2}$ are related to $s_{1,2}$ by a similar argument as in Exercise 3.6.1,

$$\|v_i\|_{\rho,G}^2=\inf\Big\{\|v_i+w\|_{\rho,G}^2:\mathrm{div}\,\rho w=0\Big\}.$$

Exercise 3.10.1. Show that the set of minimizers in this minimization problem is

$$\mathrm{Tan}^v_{\rho,G}:=\overline{\{G^{-1}\nabla p:p\in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\rho,G)}.$$

Note that $G^{-1}\nabla p$ is the gradient of p in the G-inner product (see Section 1.2.4).

The inner product $(v_1,v_2)_G$, and similarly the induced inner product (3.12) inherit their modelling interpretation from the corresponding interpretation of $(v_1,v_2)_G$ in $\mathbb{R}^d$, which follows from the assumption that a particle moving with velocity v requires a force f with $f=Gv$.

3.11 Gradient flows in Wasserstein space

Given this definition of inner product it is natural to generalize the expression (2.6) to the Wasserstein space.

Definition 3.11.1 (Gradients and gradient flows in $\mathcal{P}_2$, dimensionless form)**.** *Let* $\mathcal{E} : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ *be given. The* Wasserstein gradient *of* $\mathcal{E}$ *at* ρ*, if it exists, is the distribution on* $\mathbb{R}^d$

$$\operatorname{grad}_W \mathcal{E}(\rho) := -\operatorname{div} \rho \nabla \operatorname{grad}_{L^2} \mathcal{E}(\rho), \tag{3.13}$$

which satisfies

$$(\operatorname{grad}_W \mathcal{E}(\rho), s)_\rho = \langle \mathcal{E}'(\rho), s \rangle \qquad \text{for all } s \in \operatorname{Tan}_\rho . \tag{3.14}$$

The Wasserstein gradient flow *of* $\mathcal{E}$ *is given by*

$$\partial_t \rho_t = -\operatorname{grad}_W \mathcal{E}(\rho_t). \tag{3.15}$$

The notation $\operatorname{grad}_{L^2} \mathcal{E}$ stands for the gradient in the metric of $L^2(\mathbb{R}^d)$, as discussed in Definition 2.4.2, and the ordering of the operators is $\operatorname{div}(\rho(\nabla((\operatorname{grad}_{L^2} \mathcal{E})(\rho))))$.

As for most of these notes, Definition 3.15 is formal. In order for (3.13) to make sense, we would need $\rho \in \operatorname{dom} \mathcal{E}'$, $\operatorname{grad}_{L^2} \mathcal{E}(\rho) \in L^2$, $\nabla \operatorname{grad}_{L^2} \mathcal{E}(\rho)$ should be a function, and $\rho \nabla \operatorname{grad}_{L^2} \mathcal{E}(\rho)$ should be locally integrable; if these conditions are met then the right-hand side of (3.13) defines a distribution. Clearly, there will be many cases in which this is asking too much. On the other hand, regularizing effects often make these properties reasonable; see also the discussion in Sections 2.4 and 2.5.

Exercise 3.11.1. Show that (3.13) and (3.14) are, at least formally, equivalent.

The question in which sense (3.15) is a gradient flow occupies most of [AGS05]. One can make this gradient-flow concept rigorous in a variety of ways, depending on the degree of regularity of the energy $\mathcal{E}$. In order of increasing quality of the description, (a) as limits of time-discrete approximations ('generalized minimizing movements'), (b) as 'curves of maximal slope', a concept that codifies the idea that 'gradient flows reduce $\mathcal{E}$ as fast as possible', (c) in a pointwise (in time) differential formulation reminiscent of '$\dot{z} = -\nabla \mathcal{E}$', and (d) by an 'evolution variational inequality' that also allows for error estimates. For functionals with appropriate convexity properties the four concepts coincide.

Definition 3.11.2 (Gradient flows in $\mathcal{P}_2$, dimensionful version)**.** *For any given dimensionful Wasserstein inner product* $((\cdot,\cdot))_\rho$*, such as* $(\cdot,\cdot)_{\rho,G}$ *in* (3.12)*, the gradient flow is the solution of*

$$((\partial_t \rho_t, s))_{\rho_t} = -\langle \mathcal{E}'(\rho_t), s \rangle \qquad \text{for all } s \in \operatorname{Tan}_{\rho_t} .$$

In the case when $((\cdot,\cdot))_\rho = (\cdot,\cdot)_{\rho,G}$ and $G = cI$ this reduces to the differential equation

$$c\partial_t \rho_t = -\operatorname{grad}_W \mathcal{E}(\rho_t). \tag{3.16}$$

Exercise 3.11.2. Work out the corresponding expression for the case of a general symmetric positive definite matrix G.

3.12 Chemical potential

When ρ is interpreted as a bunch of particles $\rho_N = N^{-1}\sum_i \delta_{x_i}$—or a limit of such objects as $N \to \infty$, i.e. a probability measure—then the L^2-gradient of an energy $\mathcal{E}$ has a specific interpretation as the *chemical potential*. The term is a misnomer: the word 'chemical' suggests a relationship with processes that make and break chemical bonds, but the actual meaning has nothing to do with that. Instead, the interpretation is as follows: removing a particle at x and replacing it at y amounts to a perturbation of ρ_N of the form $N^{-1}(\delta_y - \delta_x)$. For large N, this signed measure has small total variation (see (3.2)), and it might be reasonable to approximate the difference by the derivative:

$$\begin{aligned}\mathcal{E}\Big(\rho_N + \frac{1}{N}(\delta_y - \delta_x)\Big) - \mathcal{E}(\rho_N) &\approx \Big\langle \mathcal{E}'(\rho_N), \frac{1}{N}(\delta_y - \delta_x)\Big\rangle = \Big(\operatorname{grad}_{L^2}\mathcal{E}(\rho_N), \frac{1}{N}(\delta_y - \delta_x)\Big)_{L^2} \\ &= \frac{1}{N}\big(\operatorname{grad}_{L^2}\mathcal{E}(\rho_N)(y) - \operatorname{grad}_{L^2}\mathcal{E}(\rho_N)(x)\big).\end{aligned}$$

with some obvious abuse of notation, and assuming that $\operatorname{grad}_{L^2}\mathcal{E}$ is actually a continuous function which therefore allows for pointwise evaluation. This little calculation shows that $\operatorname{grad}_{L^2}\mathcal{E}(\rho_N)(y) - \operatorname{grad}_{L^2}\mathcal{E}(\rho_N)(x)$ quantifies the change in energy when a particle is taken from x to y, and therefore $\operatorname{grad}_{L^2}\mathcal{E}$ is considered the 'potential' of a particle with respect to $\mathcal{E}$.

My guess of where the word *chemical* comes from is as follows. Chemical reactions consume and produce different types of particles. In the case of a simple reaction $A \leftrightharpoons B$, the reaction event of an A particle converting into a B particle might be represented as a particle moving from position x (where x would, amongst other things, identify the particle as an A particle) to a position y corresponding to a B particle. The corresponding change in energy $\mathcal{E}$ is what we calculated above.

In the chemical, physical, and statistical mechanics literature, the chemical potential is often indicated with the letter μ. Because of the abundance of measures in these notes, we will always write $\operatorname{grad}_{L^2}\mathcal{E}$ in full; this also helps to reduce the frequent confusion caused by the term.

Writing (3.15) as

$$\partial_t \rho = \operatorname{div} \rho \nabla \operatorname{grad}_{L^2}\mathcal{E}(\rho), \tag{3.17}$$

we see that the Wasserstein gradient flow of an energy $\mathcal{E}$ amounts to a convection equation with the gradient of the chemical potential as the convection velocity. This form is often found postulated in the physical literature. We now see that from a mathematical point of view, this amounts to postulating a Wasserstein gradient flow.

The interpretation of the gradient of the chemical potential as the convective velocity field is actually slightly off the mark. In fact the gradient of the chemical potential is a force field, not a velocity: it is a linear form on vector fields, not a vector field itself, and it also has the wrong dimensions. A constant (and a corresponding modelling argument) is required to convert one into the other. An example of such a modelling argument could be that the particles move through a viscous fluid (see Section 1.1.5) with the conversion constant given by Stokes' law (1.1); see Section 4.3.

3.13 Work and energetics along a curve in $\mathcal{P}_2(\mathbb{R}^d)$

Since the power exerted by a force $f \in \mathbb{R}^d$ on a particle moving with velocity $v \in \mathbb{R}^d$ is $f \cdot v$, the power of a force field $f : \mathbb{R}^d \to \mathbb{R}^d$ on a set of particles $\rho_t = N^{-1} \sum_i \delta_{x_i(t)}$ moving with velocities $v_i(t) = v(x_i(t))$ is

$$\frac{1}{N} \sum_{i=1}^{N} f(x_i(t)) \cdot v(x_i(t)) = \int_{\mathbb{R}^d} f \cdot v \, d\rho_t. \tag{3.18}$$

Note that this is the same expression as the Wasserstein inner product (3.11), although of the two components f and v only the latter can be interpreted as a representation of a tangent, in this case of the tangent $\partial_t \rho$. In fact, one should think of (3.18) as a linear form on $\partial_t \rho$; more precisely, we can define the corresponding linear form

$$\ell : \mathrm{Tan}_\rho \to \mathbb{R}, \qquad \langle \ell, s \rangle = \int_{\mathbb{R}^d} f \cdot v \, d\rho, \quad \text{where } s + \operatorname{div} \rho v = 0 \text{ and } v \in \mathrm{Tan}^v_\rho \,.$$

Recalling the discussion on duality in Section 1.2.5, we may recognize that we could (and maybe should) even replace the vectorial inner product $f(x) \cdot v(x)$ at the point $x \in \mathbb{R}^d$ by a duality $\langle \xi(x), v(x) \rangle$, where for each $x \in \mathbb{R}^d$, $\xi(x)$ is an element of the dual space $(\mathbb{R}^d)'$. This illustrates a relevant point: the Wasserstein *inner product* is built upon the inner product of the underlying space $\mathbb{R}^d$, while a *linear form* on the Wasserstein tangent space Tan_ρ is independent of the underlying inner product.

Given this interpretation of power, it is not surprising to see that any energy $\mathcal{E}$ is the integral along its path of the work done by the gradient $\nabla \operatorname{grad}_{L^2} \mathcal{E}$ of the chemical potential $\operatorname{grad}_{L^2} \mathcal{E}$. Indeed, if $t \mapsto \rho_t$ is a curve and $\partial_t \rho + \operatorname{div} \rho \nabla p = 0$, then

$$\begin{aligned}
\mathcal{E}(\rho_t) - \mathcal{E}(\rho_s) \qquad &= \qquad \int_s^t \langle \mathcal{E}'(\rho_\tau), \partial_\tau \rho_\tau \rangle \, d\tau \\
&\overset{(3.14)}{=} \qquad \int_s^t (\operatorname{grad}_W \mathcal{E}(\rho_\tau), \partial_\tau \rho_\tau)_{\rho_\tau} \, d\tau \\
&\overset{(3.11),(3.13)}{=} \qquad \int_s^t \int_{\mathbb{R}^d} \nabla \operatorname{grad}_{L^2} \mathcal{E}(\rho_\tau) \cdot \nabla p_\tau \, d\rho_\tau d\tau.
\end{aligned}$$

3.14 Comments on Chapter 3

Optimal transport theory. The Wasserstein distances introduced in this chapter are part of a much larger body of theory, that of *optimal transport*. The prototypical question in this field is to minimize, for given $\rho_0, \rho_1 \in \mathcal{P}(Z)$, a *cost functional*

$$\int c(x, y) \gamma(dxdy)$$

over all couplings $\gamma \in \Gamma(\rho_0, \rho_1)$. The interpretation of $c(x, y)$ is that of the cost of transferring a unit of mass from the point x to the point y; the coupling $\gamma(dxdy)$ is known

as a *transport plan* and specifies the choice of how much mass to transport from x to y. This problem was introduced, in a slightly different form, by Monge in 1781 [Mon81], and reformulated in the form above by Kantorovich in 1942 [Kan42].

The field of optimal transport (OT) has blossemed into a theory with extensive connections to many other areas of mathematics. One example of this is obviously the class of Wasserstein gradient flows, as we discuss in these notes. Other examples are the use of OT theory to characterize the geometry of the underlying metric space [vRS05]. Villani [Vil03] provides an accessible introduction; also see [AGS05, Vil09].

Non-convex subsets Z. Throughout the chapter we have worked in subsets $Z \subset \mathbb{R}^d$ that are convex. If Z is convex, then the shortest curve connecting two points is a straight line. If Z is not convex, then there exist pairs of points for which the shortest curve is no longer straight. In that case the metric on Z no longer coincides with the metric on $\mathbb{R}^d$, and therefore in Definition 3.3.1 the distance $|x - y|$ should be replaced by the metric d on Z.

However, I'm guessing that the Benamou-Brenier formula (3.9) and the local inner product (3.11) remain the same. Since the Wasserstein gradient-flow properties can be defined using only the inner product (see (3.13–3.15)), the resulting evolution should remain the same. If anyone knows any references on this, I'd be interested.

Chapter 4

Entropy

4.1 The diffusion equation is the Wasserstein gradient flow of the entropy

At the heart of both the theory and the applications of the Wasserstein gradient flows lies the example of the diffusion equation, which is the Wasserstein gradient flow of the *entropy*

$$\mathsf{Ent} : \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}, \qquad \mathsf{Ent}(\rho) = \begin{cases} \displaystyle\int_{\mathbb{R}^d} f \log f & \text{if } \rho = f\mathcal{L}^d, \\ +\infty & \text{otherwise.} \end{cases}$$

As above we often encounter expressions of the form $\log f$ with $\rho = f\mathcal{L}^d$, and we choose to use the same letter ρ for both the measure and its Lebesgue density. When we need to indicate the integration variable, $\rho(x)$ indicates the density and $\rho(dx)$ the measure (so that $\rho(x)\,dx = \rho(dx)$).

The fact that the diffusion equation is the Wasserstein gradient flow of this functional Ent is simple to check: the chemical potential $\operatorname{grad}_{L^2} \mathsf{Ent}(\rho)$ is equal to $\log \rho + 1$, and the equation (3.17) becomes

$$\partial_t \rho = \operatorname{div} \rho \nabla (\log \rho + 1) = \Delta \rho. \tag{4.1}$$

Note that the sign of Ent is the opposite of what is common in physics. In this definition the entropy is *convex* and *decreases* in time.

4.2 Entropy

Because of the tremendous relevance of entropy for Wasserstein gradient flows we spend some time discussing its properties. One argument that leads to the introduction of the entropy Ent is a counting argument, that we explain for the case of a finite state space I and corresponding entropy

$$\mathsf{Ent}(\rho) = \sum_{i \in I} \rho_i \log \rho_i, \qquad \text{for } \rho_i \geq 0,\ \sum_{i \in I} \rho_i = 1.$$

Consider N particles on the lattice described by I, i.e. consider a mapping $x : \{1, \dots, N\} \to I$. Define the *empirical measure* $\rho \in \mathcal{P}(I)$ by

$$k_i := \#\big\{j \in \{1, \dots, N\} : x(j) = i\big\}, \qquad \rho_i = \frac{k_i}{N}. \tag{4.2}$$

In going from x to ρ there is loss of information; in the same way as discussed in Section 3.2, multiple mappings x produce the same empirical measure ρ. The degree of degeneracy, the number of unique mappings x that correspond to a given ρ, is $N! \left(\prod_{i\in I} k_i!\right)^{-1}$. We now use Stirling's formula in the form

$$\log n! = n \log n - n + o(n) \qquad \text{as } n \to \infty,$$

to estimate

$$\begin{aligned}
\log N! \left(\prod_{i\in I} k_i!\right)^{-1} &= \log N! - \sum_{i\in I} \log k_i! \\
&= N \log N - N - \sum_{i\in I} (k_i \log k_i - k_i) + o(N) \\
&= -N \sum_{i\in I} \rho_i \log \rho_i + o(N) \qquad \text{as } N \to \infty.
\end{aligned}$$

One interpretation of the entropy therefore is as

$$\mathsf{Ent}(\rho) = -\lim_{N\to\infty} \frac{1}{N} \log \#\text{realizations of } \rho.$$

This shows that if the number of microscopic realizations x of the macroscopic object ρ is large, then $\mathsf{Ent}(\rho)$ is small, and vice versa.

If we allocate particles at random with the same probability for each microstate x, then the probability of obtaining each microsate is $|I|^{-N}$, and the probability of a macrostate ρ satisfies

$$\log \mathrm{Prob}(\rho) = \log |I|^{-N} N! \left(\prod_{i\in I} k_i!\right)^{-1} = -N \log |I| - N\, \mathsf{Ent}(\rho) + o(N) \qquad \text{as } N \to \infty.$$

Therefore we find for two macrostates ρ_1 and ρ_2 that

$$\mathsf{Ent}(\rho_1) - \mathsf{Ent}(\rho_2) = -\lim_{N\to\infty} \frac{1}{N} \log \frac{\mathrm{Prob}(\rho_1)}{\mathrm{Prob}(\rho_2)}. \tag{4.3}$$

In these calculations we have glossed over the fact that the probability of a given $\rho \in \mathcal{P}(I)$ is only non-zero if it is of the form (4.2); if ρ happens to be of the right form for N, then it won't be of the right form for $N + 1$. This implies that the statement (4.3) can not be true as it stands.

However, (4.3) can be made rigorous in the context of large-deviation theory (e.g. [dH00, OV05]). The problem above is avoided by a smart use of open and closed sets, and by proving for instance for all open sets $U \subset \mathcal{P}(I)$ that

$$-\lim_{N\to\infty} \frac{1}{N} \log \mathrm{Prob}\big(\{k_i/N\}_{i\in I} \in U\big) = \inf_{\rho\in U} \mathsf{Ent}(\rho) + \text{constant}.$$

4.3 An example: a convection-diffusion equation

We now describe an example of a modelling process. We consider charged particles that move through a viscous fluid under the effects of a fixed external electric field and thermal agitation (diffusion). More precisely, the ingredients are as follows:

- A surrounding viscous fluid of dynamic viscosity η (in Pa s) and temperature T (in K);
- An external electric field $E = -\nabla V$ (E in N/C, V in J/C); each particle has a charge q (in C); the external field E and potential V depend on x, but $E(x)$ and $V(x)$ are independent of the positions of the particles (this amounts to a small-concentration assumption);
- Diffusion as the result of thermal agitation by the surrounding fluid.

We now assemble a model.

1. The *kinematics* are given by the state variable $\rho \in \mathcal{P}_2(\mathbb{R}^d)$, which represents the concetration of particles, i.e.

$$\rho(A) = \frac{\text{number of particles in } A}{\text{total number of particles}}.$$

   The dimension of ρ therefore is m^{-d}. Note how this is different from the custom in continuum mechanics, where ρ is a *mass* density instead of a *number* density, with dimensions $[\rho] = \mathrm{kg\,m}^{-d}$.

Exercise 4.3.1. Repeat the derivation for the case where ρ is a mass density. Which additional assumptions are necessary?

2. The *energy* describes the forces that drive the system. Since an individual particle at x has an electrostatic energy $qV(x)$ (in $\mathrm{C} \cdot \mathrm{J/C} = \mathrm{J}$) , the total electrostatic energy of the system described by ρ is

$$\mathcal{E}_{\mathrm{e}}(\rho) = q \int_{\mathbb{R}^d} V \, d\rho.$$

   In addition the thermal agitation by the surrounding fluid also leads to an energetic contribution, which turns out to be

$$\mathcal{E}_{\mathrm{t}}(\rho) = kT \, \mathsf{Ent}(\rho),$$

   where $k = 1.4 \cdot 10^{-23}\mathrm{J/K}$ is the *Boltzmann constant*. The total energy (with dimensions J) is

$$\mathcal{E} = \mathcal{E}_{\mathrm{e}} + \mathcal{E}_{\mathrm{t}}.$$

   Note that

$$\mathrm{grad}_{L^2} \mathcal{E}(\rho) = kT(\log \rho + 1) + qV.$$

3. For the modelling of the dissipation we make an additional assumption that the particles are spherical. By Stokes' law (1.1) the velocity v and force f are then related by $f = 6\pi\eta R v =: cv$, which leads to a dissipation inner product (see Section 3.10)

$$(s_1, s_2)_{\rho,c} := c \int_{\mathbb{R}^d} v_1 \cdot v_2 \, d\rho \qquad \text{with} \quad s_i + \operatorname{div} \rho v_i = 0, \ i = 1, 2.$$

   The dimension is W.

Putting the pieces together and using expression (3.16) we find the equation

$$c\partial_t \rho = \operatorname{div} \rho \nabla \operatorname{grad}_{L^2} \mathcal{E} = kT\Delta\rho + q \operatorname{div} \rho \nabla V. \tag{4.4}$$

4.4 The Einstein relation

The derivation above differs from a more traditional derivation, which might run as follows:

1. Conservation of particles implies that $\partial_t \rho + \operatorname{div} F = 0$ for some flux function F;
2. The diffusive contribution to the flux is given by $F_{\mathrm{d}} = -D\nabla\rho$ *(Fick's law);*
3. The convection by the electrostatic field contributes a flux $F_{\mathrm{e}} = -(q/c)\rho\nabla V$.

Putting these together we find

$$\partial_t \rho = -\operatorname{div}(F_{\mathrm{d}} + F_{\mathrm{e}}) = D\Delta\rho + \frac{q}{c} \operatorname{div} \rho \nabla V. \tag{4.5}$$

Comparing this result with (4.4) provides an additional piece of information: the constant D in Fick's law should be related to the temperature T and the friction coefficient c by

$$D = \frac{kT}{c}. \tag{4.6}$$

This identity is known as the *Einstein relation* [Ein05]. It can be derived independently by taking the classical derivation above and adding the assumption that

$$\exp -\frac{qV(x)}{kT} \, \mathcal{L}^d \quad \text{is a stationary solution of (4.5).} \tag{4.7}$$

The relation (4.6) then follows immediately. Of course, this 'derivation' simply shifts the issue to the question why one requires (4.7).

When the particles are assumed to be spherical, and Stokes' law (1.1) is invoked to calculate c, the relation (4.6) becomes the *Stokes–Einstein relation* [Ein05]

$$D = \frac{kT}{6\pi\eta R}.$$

These two Einstein relations are examples of a more general class of principles, *fluctuation-dissipation theorems*. As above, these stipulate relations between frictional, dissipative effects on one hand, and fluctuating excitational effects on the other.

4.5 Relative entropy

The potential V in the example above is only defined up to a constant. If the function $x \mapsto \exp(-qV(x)/kT)$ is integrable, then we can add a constant to V such that $\mu := \exp(-qV/kT)\,\mathcal{L}^d \in \mathcal{P}(\mathbb{R}^d)$. In that case we can write the energy of the example as

$$\mathcal{E}(\rho) = kT\,\mathsf{Ent}(\rho) + q\int V\,d\rho = kT\,\mathsf{Ent}(\rho) - \int \log\frac{\mathrm{d}\mu}{\mathrm{d}\rho}\,d\rho.$$

Apart from the dimensional constants this expression is known as the *relative entropy* of ρ with respect to μ:

Definition 4.5.1. *Given $\rho, \mu \in \mathcal{P}(\mathbb{R}^d)$, the* relative entropy *of ρ with respect to μ is*

$$H(\rho|\mu) := \begin{cases} \displaystyle\int_{\mathbb{R}^d} f\log f\,d\mu & \text{if } \rho \ll \mu,\ \rho = f\mu, \\ +\infty & \text{otherwise.} \end{cases}$$

It has some useful properties:

Lemma 4.5.2. *Let $\rho, \mu \in \mathcal{P}(\mathbb{R}^d)$.*

1. *$H(\rho|\mu) \geq 0$, and $H(\rho|\mu) = 0$ if and only if $\rho = \mu$;*
2. *For fixed μ, $\rho \mapsto H(\rho|\mu)$ is strictly convex.*

Proof. The strict convexity of H follows from the same property of $z \mapsto \eta(z) := z\log z$. The positivity of H follows from remarking that $\eta(z) \geq z - 1$, so that $H(\rho|\mu) \geq \int [f-1]\,d\mu = 0$. The characterization of equality follows from the corresponding characterization of equality in this inequality, or from the strict convexity. □

Note that if $\Omega \subset \mathbb{R}^d$ has finite Lebesgue measure, then the 'usual' entropy Ent_Ω on Ω can be written as

$$\mathsf{Ent}_\Omega(\rho) := \int_\Omega \rho\log\rho = H\big(\rho\big|\mathcal{L}^d|_\Omega\big) + \log\mathcal{L}^d(\Omega).$$

This illustrates that Ent, the entropy on $\mathbb{R}^d$, can be viewed as the limit of 'relative entropy with respect to the uniform distribution' on an increasing sequence of subsets of $\mathbb{R}^d$.

Chapter 5

Other energies

5.1 Electrostatic and gravitational interaction

Interaction between particles, for instance by electrostatic or gravitational forces, gives rise to energy terms. In this section we calculate these. We restrict ourselves to three space dimensions; other dimensions require separate treatment.

5.1.1 Electrostatic interaction

A charge density $q\rho$ (with dimensions $\mathrm{C/m^3}$) creates an electrostatic potential φ (called *Poisson potential*, with dimensions J/C) satisfying *Poisson's equation*,

$$-\varepsilon_0 \Delta\varphi = q\rho \quad \text{in } \mathbb{R}^3, \qquad \varphi(x) \to 0 \text{ as } |x| \to \infty. \tag{5.1}$$

Here ε_0 is the *electric constant* (historically also called vacuum permittivity or permittivity of empty space) with value $8.854187817... \cdot 10^{-12} \mathrm{A^2\,s^4\ kg^{-1}\,m^{-3}}$. Note that equation (5.1) has a natural generalization to the case of a signed measure ρ (with dimensions such that $[\rho(A)] = \mathrm{C}$ for any Borel set A). We will comment on this generalization below.

Equation (5.1) is a reduction of the full Maxwell's equations in the case of stationary electric charges and stationary magnetic field; but we will want to apply this to moving charges. This is reasonable provided that the velocities of the particles are small; I'm guessing that they should be small with respect to the speed of light, but I have no reference (or argument) for this. All suggestions are welcome.

When $\rho = q\delta_{x_0}$, corresponding to a point charge q at x_0, the correspdonding electrostatic field $\varphi(x)$ is equal to (see below)

$$\varphi(x) = \frac{q}{4\pi\varepsilon_0} \frac{1}{|x - x_0|}.$$

The force exerted by such an electrostatic potential φ on a charge q' at x is given by $-q'\nabla\varphi(x)$. Therefore the electrostatic force exerted by two (stationary) charged particles on to each other, with charges $q_{1,2}$ (in C) at positions $x_{1,2}$, is given by *Coulomb's law*,

$$\frac{q_1 q_2}{4\pi\varepsilon_0} \frac{x_1 - x_2}{|x_1 - x_2|^3}.$$

This force is parallel to the connecting line $x_1 - x_2$, and repulsive if $q_1 q_2 > 0$.

Given a system of N particles at disjoint positions x_i with charges q_i, the total energy of this system is equal to

$$\mathcal{E} = \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \frac{q_i q_j}{4\pi\varepsilon_0} \frac{1}{|x_i - x_j|}.$$

This can be recognized by differentiating with respect to x_k, while fixing the other positions:

$$\partial_{x_k} \mathcal{E} = \frac{1}{2} \sum_{\substack{i=1 \\ i \neq k}}^{N} \frac{q_i q_k}{4\pi\varepsilon_0} \frac{x_i - x_k}{|x_i - x_k|^3} - \frac{1}{2} \sum_{\substack{j=1 \\ j \neq k}}^{N} \frac{q_k q_j}{4\pi\varepsilon_0} \frac{x_k - x_j}{|x_k - x_j|^3} = \sum_{\substack{i=1 \\ i \neq k}}^{N} \frac{q_i q_k}{4\pi\varepsilon_0} \frac{x_i - x_k}{|x_i - x_k|^3}.$$

This final expression is the joint force exerted by all the other charges on the charge k.

Still in the context of this system of N particles, we now choose all charges $q_i = q$ to be the same, so that the interaction energy between two particles is given by

$$V(x) = \frac{q^2}{4\pi\varepsilon_0} \frac{1}{|x|}. \tag{5.2}$$

We now define the number density (not *charge* density but *number* density, to connect with the earlier chapters)

$$\rho_N(x) = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i}.$$

Then

$$\frac{1}{2N^2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \frac{q_i q_j}{4\pi\varepsilon_0} \frac{1}{|x_i - x_j|} \approx \frac{1}{2} \iint V(x-y) \rho_N(dx) \rho_N(dy). \tag{5.3}$$

The approximation "$\approx$" is necessary since the right-hand side is not well-defined: it contains terms of the form

$$\frac{1}{2N^2} \iint V(x-y) \delta_{x_k}(dx) \delta_{x_k}(dy) = \frac{1}{N^2} V(0) = \infty.$$

However if ρ_N is replaced by a ρ that is sufficiently regular, then the integral becomes well-defined. This discussion motivates the definition of the electrostatic interaction energy by

$$\mathcal{E}_{\text{elec}}(\rho) := \frac{1}{2} \iint V(x-y) \rho(dx) \rho(dy), \qquad \rho \in \mathcal{P}(\mathbb{R}^3).$$

One might try to remedy the definition problems in the integral in (5.3) by defining $V(0)$ to be zero. With this choice the integral becomes well-defined, and since this V is lower-semicontinuous on $\mathbb{R}^3$ the right-hand side seems like a reasonable interaction energy. However, with this definition of V, two particles can reduce their interaction energy to zero by *being on top of one another*—a property that does not seem

to make sense. However, I don't know of an example in the electrostatic case in which this actually creates problems, and I'd be interested in any suggestions.

On the other hand, when we turn to gravitational interaction below, in which the same potential arises with the opposite sign, then the corresponding lack of lower-semicontinuity causes the value of the potential at zero to be replaced by the lower-semicontinuous limit, $-\infty$. Therefore setting $V(0)$ to zero has no effect in this case.

Lemma 5.1.1. *If $\rho \in \mathcal{P}(\mathbb{R}^3)$ with $\mathcal{E}_{\text{elec}}(\rho) < \infty$, then*

$$\mathcal{E}_{\text{elec}}(\rho) \overset{(1)}{=} \frac{1}{2}\int \psi \, d\rho \overset{(2)}{=} \frac{\varepsilon_0}{2q^2}\int |\nabla\psi|^2 \, dx.$$

Here $\psi := V * \rho$, where V is given in (5.2)*; the function $\varphi := \psi/q$ satisfies* (5.1)*.*

Proof. The definition of ψ as $V * \rho$ implies both identity (1) and the fact that $\varphi = \psi/q$ solves (5.1). Identity (2) follows formally from the following calculation:

$$\frac{1}{2}\int \psi \, d\rho \overset{(5.1)}{=} -\frac{\varepsilon_0}{2q^2}\int \psi\Delta\psi \, dx = \frac{\varepsilon_0}{2q^2}\int |\nabla\psi|^2 \, dx.$$

To make this rigorous, we note that $\psi \in L^1(\rho)$ and $\nabla\psi \in L^2(\mathbb{R}^3)$; we then approximate ψ by $\psi_n \in C_c^\infty$ in $L^1(\rho)$ such that $\nabla\psi_n \to \nabla\psi$ in $L^2(\mathbb{R}^3)$. Then

$$\int |\nabla\psi|^2 = \lim_{n\to\infty}\int \nabla\psi\nabla\psi_n = -\lim_{n\to\infty}\int \psi_n\Delta\psi = \lim_{n\to\infty}\frac{q^2}{\varepsilon_0}\int \psi_n \, d\rho = \frac{q^2}{\varepsilon_0}\int \psi \, d\rho.$$

□

5.1.2 Gravitational interaction

For gravitational interaction the picture is very similar. The only relevant difference is a sign: where particles of the same electric charge repel each other, the graviational interaction between any two particles is always attractive.

Newton's law of gravitation states that the attractive force between two masses m_1 and m_2 at distance r has the magnitude Gm_1m_2/r^2, where G is the gravitational constant with value $6.67300 \times 10^{-11}\text{m}^3\,\text{kg}^{-1}\,\text{s}^{-2}$. Therefore the interaction potential between two particles of mass m is $V(r) = -Gm^2/r$, and the gravitational energy of a mass distribution ρ (such that for any set A, $\rho(A)$ has dimension kg) is

$$\mathcal{E}_{\text{grav}}(\rho) := \frac{1}{2}\iint V(x-y)\,\rho(dx)\rho(dy) = -\frac{G}{2}\iint \frac{\rho(dx)\rho(dy)}{|x-y|}.$$

Equivalently, if $\rho(A)$ is dimensionless, and $\rho(\mathbb{R}^3)$ is normalized to be 1, then for a system of total mass M we have

$$\mathcal{E}_{\text{grav}}(\rho) := -\frac{GM^2}{2}\iint \frac{\rho(dx)\rho(dy)}{|x-y|}.$$

Similarly, for given mass distribution ρ (with dimensions kg) the *gravitational potential* is

$$\phi(x) := G \int \frac{\rho(dy)}{|x-y|}, \qquad \text{with dimension J/kg}$$

which solves the equation

$$-\Delta\phi = -4\pi G\rho \quad \text{in } \mathbb{R}^3, \qquad \text{with} \quad \phi(x) \to 0 \text{ as } |x| \to \infty.$$

Remark 5.1.2. Adding $\mathcal{E}_{\text{grav}}$ to an energy $\mathcal{E}$ can make the resulting energy unbounded from below. This phenomenon can lead to blow-up, known as *gravitational collapse* [Wol92, BCK$^+$99]. Strongly related mathematical equations arise in the modelling of aggregation in chemotaxis, such as observed in the slime mould *Dichtyostelium Discoideum* [Hor03]. □

5.1.3 Formulation as H^{-1}-norms

Let us write $\kappa(x) := 1/(4\pi|x|)$ for the fundamental solution of the Laplacian $-\Delta$ in $\mathbb{R}^3$. Then one definition of the norm of the space $H^{-1}(\mathbb{R}^3)$ is

$$\|u\|^2_{H^{-1}(\mathbb{R}^3)} := \iint\limits_{\mathbb{R}^3\times\mathbb{R}^3} \kappa(x-y)u(x)u(y)\,dxdy,$$

or if u is replaced by a measure ρ,

$$\|\rho\|^2_{H^{-1}(\mathbb{R}^3)} := \iint\limits_{\mathbb{R}^3\times\mathbb{R}^3} \kappa(x-y)\,\rho(dx)\rho(dy).$$

Then clearly we can write, for a dimensionless non-negative measure ρ,

$$\mathcal{E}_{\text{elec}}(\rho) = \frac{q^2}{2\varepsilon_0}\|\rho\|^2_{H^{-1}(\mathbb{R}^3)} \qquad \text{and} \qquad \mathcal{E}_{\text{grav}}(\rho) = -2\pi GM^2\|\rho\|^2_{H^{-1}(\mathbb{R}^3)}.$$

$H^{-1}(\mathbb{R}^3)$ is a Hilbert space with inner product

$$(u,v)_{H^{-1}} := \iint\limits_{\mathbb{R}^3\times\mathbb{R}^3} \kappa(x-y)u(x)v(y)\,dxdy,$$

which can also be written as

$$(u,v)_{H^{-1}} = (u, \kappa * v)_{L^2(\mathbb{R}^3)}$$

for any u, v for which this expression makes sense.

Since electrostatic and gravatational energies are essentially H^{-1}-norms, let us consider the energies

$$\mathcal{E}(\rho) := \mathsf{Ent}(\rho) \pm \frac{1}{2}\|\rho\|^2_{H^{-1}}, \tag{5.4}$$

where the constants have been scaled away, the plus sign corresponds to electrostatic interaction (repulsion) and the minus sign to gravitational attraction. The L^2-gradient of the H^{-1}-norm is easily calculated by remarking that

$$\|u+\varepsilon v\|_{H^{-1}}^2 = (u+\varepsilon v, \kappa * (u+\varepsilon v))_{L^2} = \|u\|_{H^{-1}}^2 + \varepsilon(v, \kappa * u)_{L^2} + \varepsilon(u, \kappa * v)_{L^2} + \varepsilon^2\|v\|_{H^{-1}}^2.$$

Since $(u, \kappa * v)_{L^2} = (v, \kappa * u)_{L^2}$ it follows that

$$\operatorname{grad}_{L^2(\mathbb{R}^3)} \frac{1}{2}\|u\|_{H^{-1}(\mathbb{R}^3)}^2 = \kappa * u.$$

Therefore the Wasserstein gradient flow of $\mathcal{E}$ in (5.4) is

$$\partial_t \rho = \operatorname{div} \rho \nabla \operatorname{grad}_{L^2} \mathcal{E} = \operatorname{div} \rho \nabla [\log \rho \pm \kappa * \rho].$$

Writing φ for $\kappa * \rho$ his system is often formulated as

$$\begin{aligned} \partial_t \rho &= \Delta \rho \pm \operatorname{div} \rho \nabla \varphi, \\ 0 &= \Delta \varphi + \rho. \end{aligned}$$

Appendix A

Physical dimensions

We follow the SI (Système International) system of units. The units in bold below are the *primary* units; the others are derived units, defined in terms of the primary ones.

physical quantity	*unit*	*symbol*	*definition*
length	**meter**	**m**	
time	**second**	**s**	
mass	**kilogram**	**kg**	
current	**Ampère**	**A**	
force	Newton	N	$\mathrm{N} = \mathrm{kg\,m/s^2}$
energy	Joule	J	$\mathrm{J} = \mathrm{N\,m}$
power	Watt	W	$\mathrm{W} = \mathrm{J/s}$
electric charge	Coulomb	C	$\mathrm{C} = \mathrm{As}$
pressure	Pascal	Pa	$\mathrm{Pa} = \mathrm{N/m^2}$
material properties			
dynamic viscosity	$\mathrm{kg/m\,s} = \mathrm{Pa\,s}$	μ, η	
kinematic viscosity	$\mathrm{m^2/s}$	ν	$\nu = \mu/\rho$ (ρ is mass density)

Bibliography

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

[AMP+11] S. Arnrich, A. Mielke, M. A. Peletier, G. Savaré, and M. Veneroni. Passing to the Limit in a Wasserstein Gradient Flow: From Diffusion to Reaction. *Arxiv preprint arXiv:1102.1202*, 2011. Accepted for publication in Calculus of Variations and Partial Differential Equations.

[BB00] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84:375–393, 2000.

[BCK+99] M. P. Brenner, P. Constantin, L. P. Kadanoff, A. Schenkel, and S. C. Venkataramani. Diffusion, attraction and collapse. *Nonlinearity*, 12:1071, 1999.

[Bil99] P. Billingsley. *Convergence of probability measures*. Wiley-Interscience, 1999.

[BMM06] C. Banfi, A. Marzocchi, and A. Musesti. On the principle of virtual powers in continuum mechanics. *Ricerche di matematica*, 55(2):139–150, 2006.

[Bre73] H. Brezis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. North Holland, 1973.

[Bre11] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, New York, 2011.

[CL72] E. A. Coddington and N. Levinson. *Theory of Ordinary Differential Equations*. Tata McGraw-Hill, 1972.

[DGMT80] E. De Giorgi, A. Marino, and M. Tosques. Problems of evolution in metric spaces and maximal decreasing curve. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)*, 68(3):180–187, 1980.

[dH00] F. den Hollander. *Large Deviations*. American Mathematical Society, Providence, RI, 2000.

[Ein05] A. Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, 17(4):548–560, 1905.

[ET87] I. Ekeland and R. Temam. *Convex analysis and variational problems*. Society for Industrial Mathematics, 1987.

[Eva02] L.C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, first edition, 2002.

[GO03] B. R. Gelbaum and J. M. H. Olmsted. *Counterexamples in analysis*. Dover, 2003.

[GS08] O. Gonzalez and A. M Stuart. *A First Course in Continuum Mechanics*. Cambridge University Press, Cambridge, 2008.

[HN11] M. Herrmann and B. Niethammer. Kramers' formula for chemical reactions in the context of Wasserstein gradient flows. *Communications in Mathematical Sciences*, 9(2):623–635, 2011.

[Hor03] D. Horstmann. From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. *Jahresber. Deutsch. Math.-Verein*, 105(3):103–165, 2003.

[Kan42] L. V. Kantorovich. On the transfer of masses. *Dokl. Akad. Nauk. SSSR*, 37(7-8):227–229, 1942.

[Kre89] E. Kreyszig. *Introductory Functional Analysis with Applications*, volume 130. John Wiley & Sons New York, 1989.

[Mon81] G. Monge. Sur la théorie des déblais et des remblais. In *Histoire de l'Academie Royale des Sciences de Paris*, pages 666–704. 1781.

[MST89] A. Marino, C. Saccon, and M. Tosques. Curves of maximal slope and parabolic variational inequalities on nonconvex constraints. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 16(2):281–330, 1989.

[Mül85] I. Müller. *Thermodynamics*. Pitman, 1985.

[NO01] B. Niethammer and F. Otto. Ostwald ripening: The screening length revisited. *Calculus of Variations and Partial Differential Equations*, 13(1):33–68, 2001.

[NO10] B. Niethammer and Y. Oshita. A rigorous derivation of mean-field models for diblock copolymer melts. *Calculus of Variations and Partial Differential Equations*, 39:273–305, 2010.

[OV05] E. Olivieri and M. E. Vares. *Large deviations and metastability*. Cambridge Univ Press, 2005.

[Roc72] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1972.

[RS06] R. Rossi and G. Savaré. Gradient flows of non convex functionals in Hilbert spaces and applications. *ESAIM: Control, Optimization and Calculus of Variations*, 12:564–614, 2006.

[Sal01] J. Salençon. *Handbook of Continuum Mechanics: General Concepts, Thermoelasticity*. Springer Verlag, 2001.

[Ser09] S. Serfaty. Gamma-convergence of gradient flows on Hilbert and metric spaces and applications. *http://www.math.nyu.edu/faculty/serfaty/gcverice2.pdf*, 2009.

[SS04] E. Sandier and S. Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Communications on Pure and Applied Mathematics*, 57(12):1627–1672, 2004.

[Ste08] U. Stefanelli. The Brezis–Ekeland principle for doubly nonlinear equations. *SIAM Journal on Control and Optimization*, 47:1615, 2008.

[Vil03] C. Villani. *Topics in Optimal Transportation*. American Mathematical Society, 2003.

[Vil09] C. Villani. *Optimal transport: Old and new*. Springer Verlag, 2009.

[vRS05] M.K. von Renesse and K.T. Sturm. Transport inequalities, gradient estimates, entropy and ricci curvature. *Communications on pure and applied mathematics*, 58(7):923–940, 2005.

[Wol92] G. Wolansky. On steady distributions of self-attracting clusters under friction and fluctuations. *Archive for Rational Mechanics and Analysis*, 119(4):355–391, 1992.