Optimal 1-rectifiable transports.

Guy Bouchitté, M. El Hajari and P. Seppecher
Université du Sud- Toulon-Var, France

Motivation: understand how optimal transport with economy of scales leads to optimal channel networks

- Stationary formulation (Monge and Kantorovich):

$$\inf\left\{\mathcal{F}(\lambda)\ ,\ -\operatorname{div}\lambda = \mu^+ - \mu^-\right\}$$

(λ transport flux measure)

- Dynamic formulation (Brenier and Benamou):

$$\inf\left\{\int_0^T \mathcal{F}'(V\rho,\rho)\,dt\ ,\ \frac{\partial\rho}{\partial t} + \operatorname{div}(V\rho) = 0\ ,\ \rho(0) = \mu^+\ ,\ \rho(T) = \mu^-\right\}$$

($\rho(t)$ mass density at time t , V the speed)

Goals

- Two main mathematical issues:

1) Functionals $\mathcal{F}$ for which optimal measures λ are one dimensional ?

2) Functionals $\mathcal{F}'$ for which optimal $\rho(t)$ are discrete measures ?

- A model case hase been widely studied

$$\mathcal{F}(\lambda) = \int_S \theta^\alpha \, dH^1 \quad \text{if} \quad \lambda = \theta \, \tau_S \, H^1 \llcorner S \qquad (+\infty \text{ otherwise}) \tag{1}$$

where $0 \leq \alpha < 1$.

Remark: $\alpha = 1$ gives Monge-Kantorovich problem (the one-rectifiability constraint dissappears after relaxation)

- J.R Banavar and All.: Universality classes of optimal channel networks, Science, 1996
- Irrigation problems: J.M. Morel, V.Caselles, M. Bernot (probability on curves)
- Q. Xia, B. Hardt: W^α Monge distance (via completion)
- G.Buttazzo, F. Santanbroggio, E.Stepanov, GMT point of view

Observation in the 2D-case

Connection with Munford-Shah image segmentation problem holds for:
$d = 2$ (static case) or for $d = 1$ (dynamic case)

Let:

- $\Omega \subset \mathbb{R}^2$ bounded with smooth boundary Γ
- μ^+, μ^- densities on Γ
- $u_0 : \Gamma \to \mathbb{R}$ is a primitive of $f := \mu^+ - \mu^-$

Then for $\lambda \in \mathcal{M}(\mathbb{R}^2; \mathbb{R}^2)$ supported in $\overline{\Omega}$

$$(i) \quad -\operatorname{div} \lambda = f \iff \exists\, u \in \mathrm{BV}(\Omega) \;:\; u = u_0 \text{ on } \Gamma \,,\; \lambda = (-\partial_2 u, \partial_1 u) \,.$$

$$(ii) \quad \lambda = \theta H^1 \llcorner S \iff \nabla u = 0 \text{ a.e. in } \Omega \setminus S \,,\; [u] = \theta \text{ on } S.$$

Thus $\inf\{\mathcal{F}_\alpha(\lambda) \;\; -\operatorname{div} \lambda = f \,,\; \operatorname{spt}\lambda \subset \overline{\Omega}\}$ is equivalent to:

$$\inf \left\{ \int_{S_u} [u]^\alpha \, dH^1 \;\; u \in \mathrm{SBV}(\Omega) \,,\; u = u_0 \text{ on } \Gamma, \; \nabla u = 0 \text{ a.e.} \right\}$$

Remark: Truncation of u (piecewise constant) $\Leftrightarrow$ Removing loops in λ.

Mass transport, economy of scales and speed

• **Monge transport** G. Monge was motivated by transporting earth from an area to an other one, "the price of the transport of a single molecule being"

(i) "proportional to its weight and" (ii) "to the distance that one makes it covering"

hence the price of the total transport is proportional to the sum of the products of the molecules each multiplied by the distance covered.

Remark: Assumption (i) says that many molecules can be transported in a single "convoy", the cost of which is proportional to the number of transported molecules. Then molecules follow a straight line (with a constant speed).

• **Economy of scales** In contrast the marginal cost of the transport decreases when the transported mass increases. We will assume that "the price of the transport of one molecule" is a concave function $g(m)$ of "its weight" m (typically $g(m) = m^\alpha$ with $0 < \alpha < 1$). This changes drastically the structure of optimal solutions: it is economic to group the transported masses as long as possible : each "molecule" will not follow a straigth line any more and the optimal strategy has to be described in a time-space setting.

Mass transport, economy of scales and speed

• **Speed of molecules** In general, transport at high speed is much more expensive. We admit that “the price of the transport of one molecule” is a convex function f of the velocity V , typically:

$$f(V) = A + BV + C\,V^p \quad \text{with} \quad p > 1\,, \quad A\,, B \text{ and } C \geq 0.$$

- $f(0) = A > 0$ means that “parking” has a cost (which is not absurd from the economical point of view).

- $B = f'(0) > 0$ in contrast favors stationary masses.

- $\alpha = 1$ and $A = B = 0\,,\ C = 1$ yields time formulation for p Wasserstein (see Benamou-Brenier)

- $\alpha < 1\,,\ p = 1$ and $A = B = 0\,,\ C = 1$ yields time formulation for for the irrigation problem studied by Xia and Morel.

Example 1

Consider two masses $M > 0$ and $m > 0$ located at time 0 at the same point x_0 to be transported respectively to points x_1 and x_2 at time $T = 1$. In other words $\rho(0) = \mu^+ = (M+m)\delta_{x_0}$ and $\rho(1) = \mu^- = M\delta_{x_1} + m\delta_{x_2}$.

We look only for two phases optimal dynamics:

a) For $[0, t_c]$ the two masses are transported together from x_0 toward a point x_c following a kinematic law $y_0(t)$.
b) For $t \in [t_c, 1]$, they are transported separately from x_c toward x_1 and x_2 following respectively the kinematic laws $y_1(t)$ and $y_2(t)$.

The cost of such a transport is:

$$\int_0^{t_c} f(\|\dot{y}_0(t)\|)g(M+m)\,dt + \int_{t_c}^1 f(\|\dot{y}_1(t)\|)g(M)\,dt + + \int_{t_c}^1 f(\|\dot{y}_2(t)\|)g(m)\,dt \quad (2)$$

to be minimized with respect to the junction time and place $(t_c, x_c) \in \mathbb{R} \times \mathbb{R}^d$ and the kinematic laws $y_0(t)$, $y_1(t)$ and $y_2(t)$ which are subjected to the constraints

$$y_0(0) = x_0,\ \ y_0(t_c) = y_1(t_c) = y_2(t_c) = x_c,\ \ y_1(1) = x_1,\ \ y_2(1) = x_2. \quad (3)$$

By the convexity of f, the velocities of the different convoys are constant.

Example 1

Thus by (2)(3), we have to minimize

$$F = t_c\, g(M+m)\, f\left(\frac{\|x_c - x_0\|}{t_c}\right) + (1-t_c)\, g(M)\, f\left(\frac{\|x_c - x_1\|}{1-t_c}\right) + (1-t_c)\, g(m)\, f(\left(\frac{\|x_c - x_2\|}{1-t_c}\right)$$

with respect to (t_c, x_c)

We draw a time-space representation of the optimal transport for

$$d = 1\ ,\ M = 1\ ,\ m = 0.5\ ,\ x_0 = 0\ , x_1 = 1\ ,\ x_2 = 0.5$$

showing the effects of the parameters α, A, B, p

(g and f are defined by $g(m) = m^{\alpha}$ and $f(V) = A + BV + V^p$)

Numerics for example 1

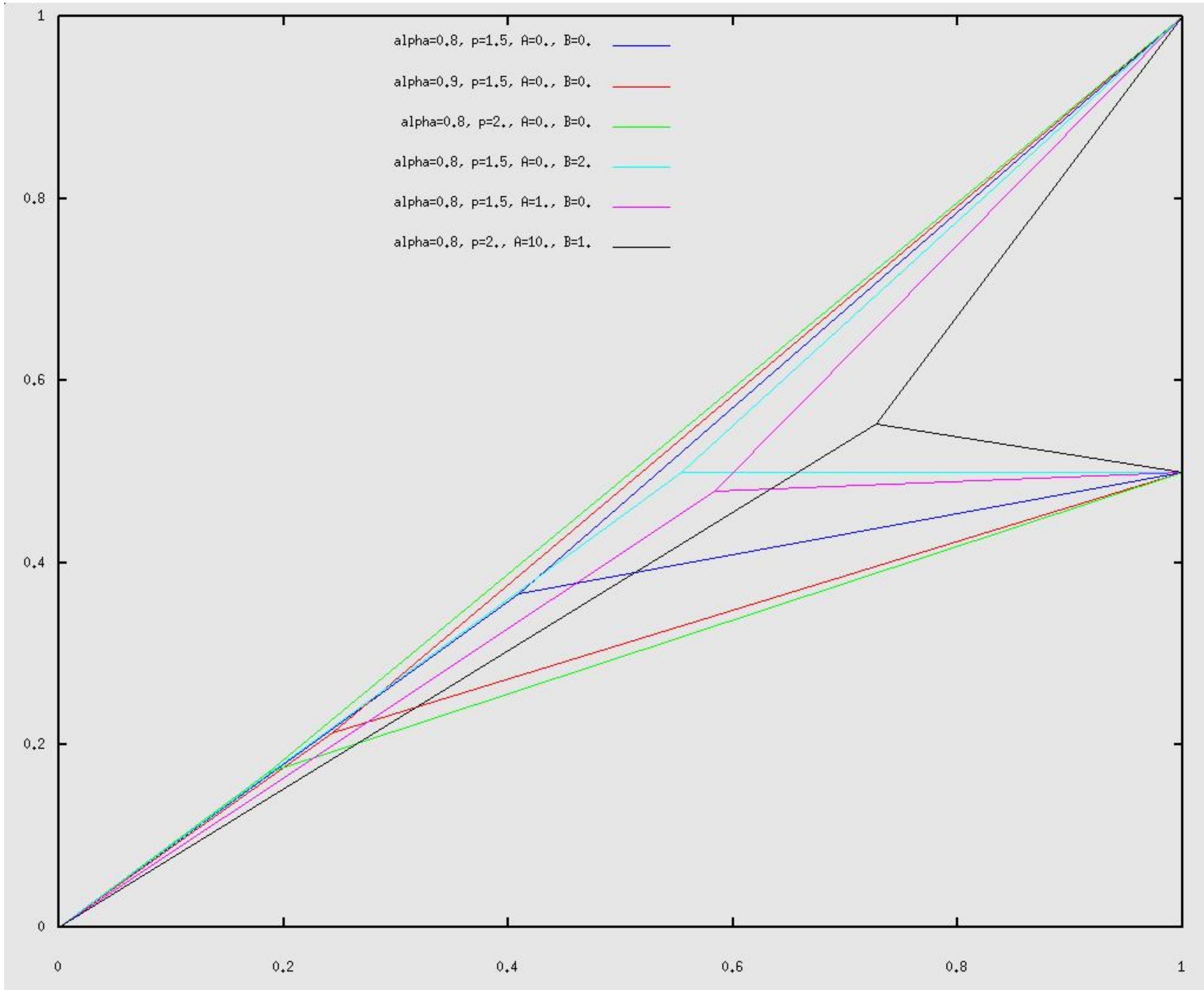

Figure 1: Optimal transport depends on p, α, A and B.

Example 2: Transport of a line density to a Dirac

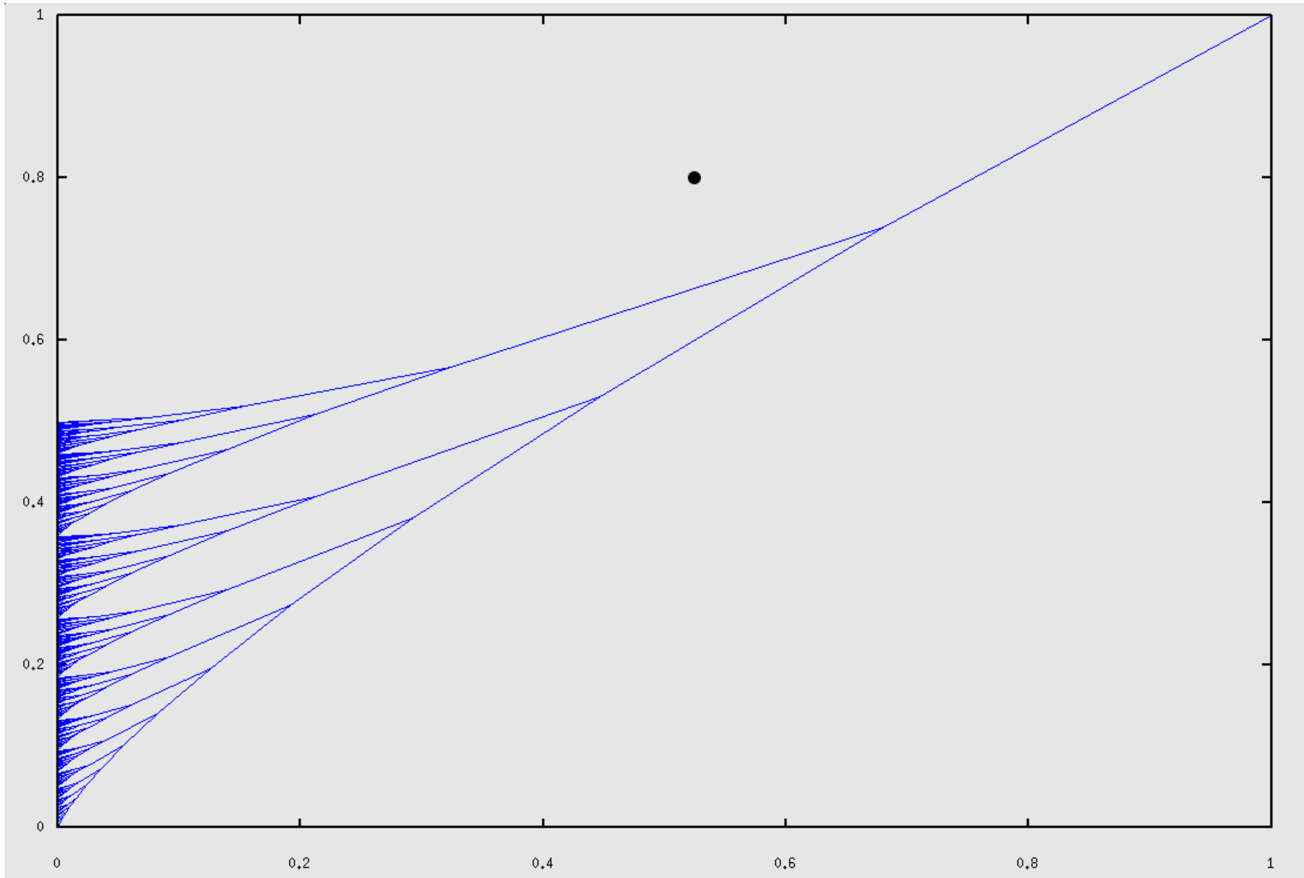

Figure 2: Auto-similar construction of optimal transport $p = 2, \alpha = 0.9$

Test for α closed to 1

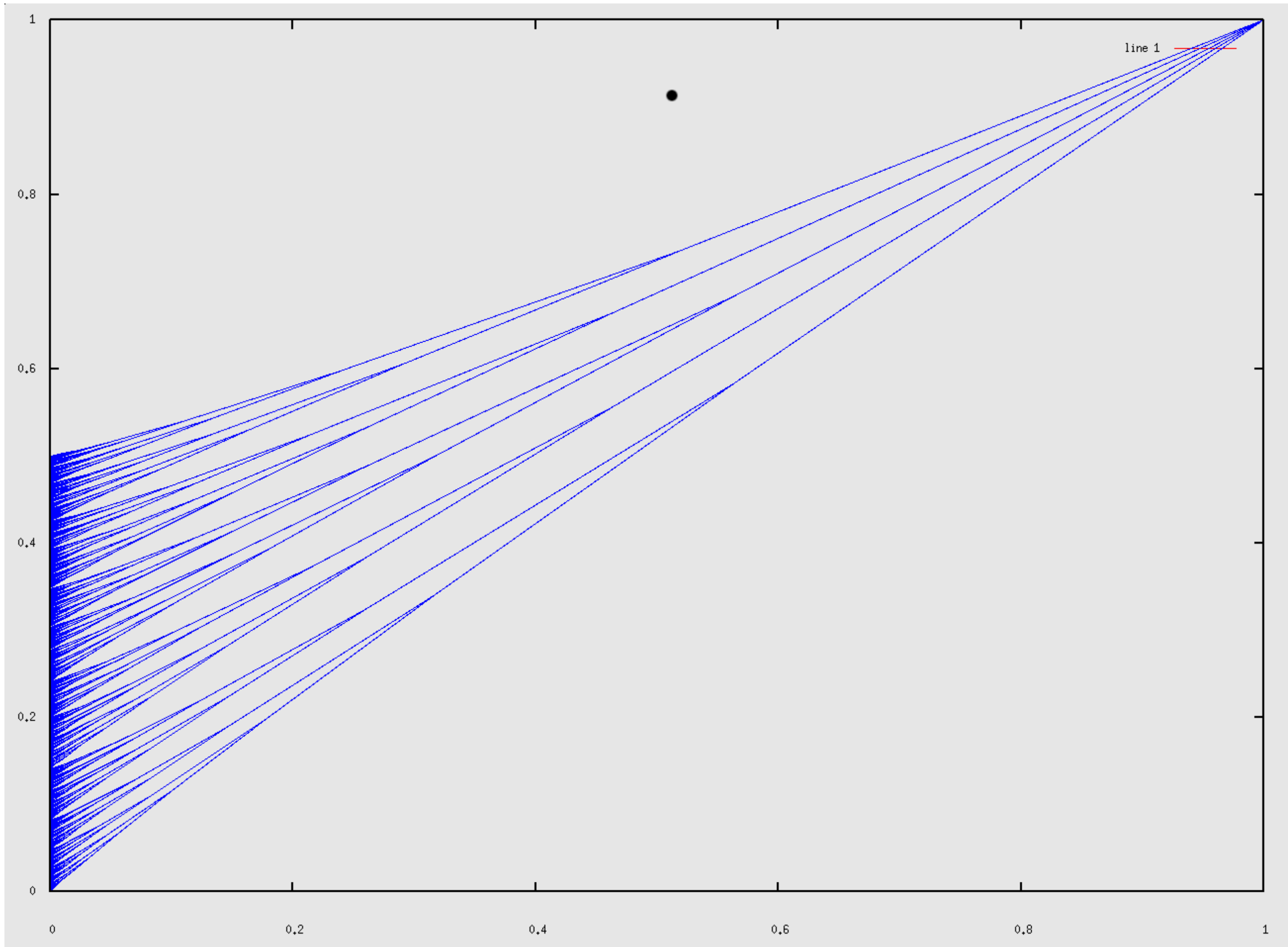

Figure 3: Transport of a line density to a Dirac, $p = 6, \alpha = 0.9$

PLAN

$$\mathcal{F}(\lambda) \ := \int_S g(\theta)\ h(\tau_S)\ dH^1$$

where:

- S is a 1-rectifiable subset of $\mathbb{R}^d$, τ_S a unit tangent vector
- $g : \mathbb{R}^+ \to [0, +\infty]$ is **concave**, monotone increasing and $\boxed{g'(0+) = +\infty}$.
- $h : \mathbb{R}^d \to [0, +\infty]$ is convex l.s.c., 1-homogeneous

Note that no lower semicontinuity result is known (except in dimension 2 via Munford-Shah functional). We will mix different techniques

1- Removing loops

2- Probability on curves and Smirnov decomposition of transport measures.

3- Intensity function and 1-rectifiability Theorem

4- Tightness results and lower semicontinuity of $\mathcal{F}$.

5- Application to dynamic formulations

6- Optimality conditions and approximation

1- Transport and loops

• **Transport measures:** A transport on $\mathbb{R}^d$ is a vector measure $\lambda \in \mathcal{M}(\mathbb{R}^d, \mathbb{R}^d)$ such that $\operatorname{div}\lambda = \mu^+ - \mu^- \in \mathcal{M}(\mathbb{R}^d)$ (thus $\int \mu^+ = \int \mu^-$).

It is called **1-rectifiable** if of the kind $\lambda = \theta\, \tau_S\; H^1 \llcorner S$ (for a suitable 1-rectifiable subset S)

The weak convergence of transports is defined by (Flat norm convergence)

$$\lambda_n \rightharpoonup \lambda \quad \Longleftrightarrow \quad \lambda_n \overset{*}{\rightharpoonup} \lambda\ ,\ \operatorname{div}\lambda_n \overset{*}{\rightharpoonup} \operatorname{div}\lambda\ .$$

• **Sub-transport:** λ' is a sub-transport of λ if there exists suitable Borel functions: $\xi, \alpha, \beta : \mathbb{R}^d \to [0,1]$ such that

$$\lambda' = \xi\,\lambda \quad , \quad \operatorname{div}(\lambda') = \alpha\,\mu^+ - \beta\,\mu^-$$

• **Loops:** It is a sub-transport sucht that $\operatorname{div}\lambda' = 0$ (i.e. $\alpha = \beta = 0$)

OBSERVATION: As g monotone $\searrow$, we have

$$\text{For all } \xi : \mathbb{R}^d \to [0,1], \qquad \mathcal{F}(\xi\lambda) \quad \leq \quad \mathcal{F}(\lambda)\ . \tag{4}$$

1.2 Sub- transports and loops

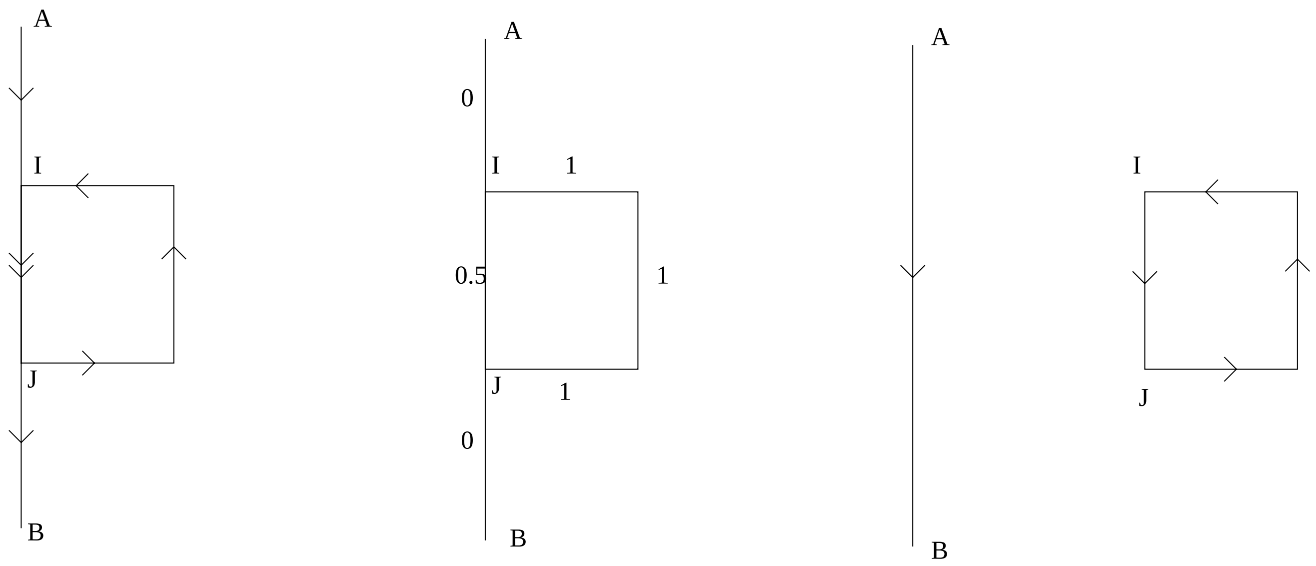

Figure 4: Removing loops: λ, ξ and decomposition of λ

1.3 Removing loops

If loops are allowed, no hope for coerciveness

$$(*) \qquad \mathcal{F}(\lambda_n) \leq C \ , \ \operatorname{div} \lambda_n = \mu^+ - \mu^- \ \Rightarrow \sup_n \int |\lambda_n| < +\infty \ .$$

We will therefore **remove loops** by using :

LEMMA 1: Given a transport measure λ, there exists a (non unique) Borel function $\xi : \mathbb{R}^d \to [0,1]$ such that: $\xi\,\lambda$ is loop free and $\operatorname{div}(1-\xi)\lambda = 0$

Proof: We consider a minimizer for problem:

$$\inf \left\{ \int \xi\,|\lambda| \ : \ \xi \in L^\infty_{|\lambda|}(\mathbb{R}^d;[0,1]) \ , \ \operatorname{div}(1-\xi)\lambda = 0 \right\} \ .$$

The property (*) will be reached for loop-free 1-rectifiable transports thanks to:

LEMMA 2: Assume that $\lambda = \theta\,\tau_S\,H^1 \llcorner S$ is loop-free and satisfies $\operatorname{div} \lambda = \mu^+ - \mu^-$ where $\int \mu^+ = \int \mu^- = M$. Then: $\boxed{0 \leq \theta \leq M}$

and by using the concavity of g which implies: $g(\theta) \geq \theta\,\dfrac{g(M)}{M}$.

2- Space of curves and probabilities

We assume that transport takes place in a **convex compact subset** Ω.

- X_Ω will denote the space of equivalent classes of Lipschitz **oriented** curves $\gamma : [0,1] \mapsto \Omega$

where $\gamma \sim \tilde{\gamma}$ means that $\tilde{\gamma} = \gamma \circ \theta$ for a strictly increasing bijection of $[0,1]$.

- X_Ω is a complete separable metric space (see Buttazzo) with

$$d(\gamma_1, \gamma_2) := \inf_\theta \sup_{t \in [0,1]} |\gamma_1(t) - \gamma_2(\theta(t))| \ ,$$

and convergence in (X_Ω, d) implies Hausdorff convergence of the images.

- X_Ω is not locally compact but for every $l > 0$ the following subset is compact

$$K_l := \{\gamma \ : \ L(\gamma) \leq l\} \quad , \quad L(\gamma) := \int_0^1 |\dot{\gamma}(t)| \, dt \ .$$

- To every $\gamma \in X_\Omega$, we can associate a line transport measure (supported in Ω)

$$< \lambda_\gamma, \phi > := \int_0^1 \phi(\gamma(t)) \cdot \dot{\gamma}(t) \, dt \quad , \quad \operatorname{div} \lambda_\gamma = \delta_{\gamma(0)} - \delta_{\gamma(1)}.$$

Remark: The map $\gamma \in X_\Omega \mapsto \lambda_\gamma \in \mathcal{M}(\Omega; \mathbb{R}^d)$ is not continuous (no control on $\dot{\gamma}$).

2.2 Representation of transports through measures on X_Ω

To every finite positive Borel measures p on X_Ω, we may associate the weighted transport denoted $\lambda(p)$ (or $\int \lambda_\gamma \, p(d\gamma)$) defined by

$$< \lambda(p), \phi > := \int_{X_\Omega} \left(\int_0^1 \phi(\gamma(t)) \cdot \dot{\gamma}(t) \, dt \right) p(d\gamma). \tag{5}$$

$$\mathrm{div}(\lambda(p)) \ = \ e_0^\sharp(p) - e_1^\sharp(p) \ ,$$

where for $i \in \{0, 1\}$, $e_i^\sharp(p)$ denotes the image of p by the continuous map $e_i : \gamma \mapsto \gamma(i)$.

Notion of complete decomposition: we say that the decomposition (5) is complete (or simply that p is complete) if:

(i) $$\mathrm{spt}\,(p) \ \subset \ \{\gamma \in X_\Omega \ : \ \gamma \text{ is simple}\}$$

(ii) $$\int_{\mathbb{R}^d} |\lambda(p)| \ = \ \int_{X_\Omega} |\lambda_\gamma| \, p(d\gamma) \ = \ \int_{X_\Omega} L(\gamma) \, p(d\gamma)$$

(iii) $$\int_{\mathbb{R}^d} |\,\mathrm{div}\,\lambda(p)| \ = \ \int_{X_\Omega} |\,\mathrm{div}\,\lambda_\gamma| \, p(d\gamma) \ = \ 2 \int_{X_\Omega} p(d\gamma)$$

2.3 Example of complete decompositions

Remark The following localized inequalities (always true) become equalities

$$|\lambda(p)| \ \le \ \int_{X_\Omega} |\lambda_\gamma|\, p(d\gamma) \quad , \quad |\operatorname{div}(\lambda(p))| \ \le \ e_0^\sharp(p) + e_1^\sharp(p).$$

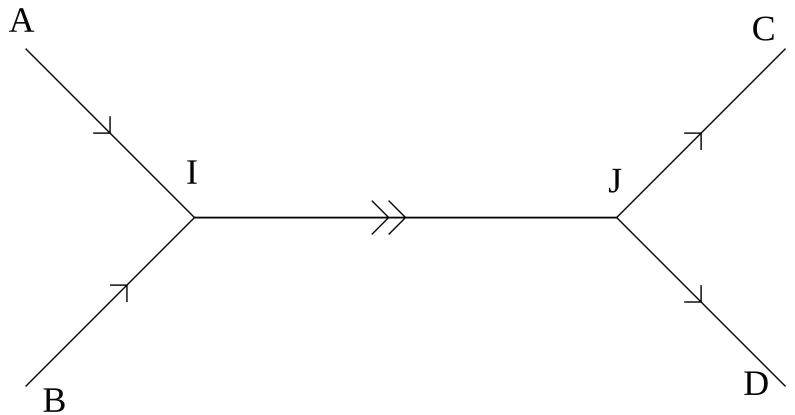

Figure 5: A transport to be decomposed

Figure 6: Two complete decompositions

2.4 Reformulation of optimal transport problem

It is tempting to reformulate the optimal transport problem as

$$\inf \left\{ \mathcal{F}(\lambda(p)) \ : \ p \in \mathcal{M}_+(X_\Omega) \ , \ e_0^\sharp(p) = \mu^+, \ e_1^\sharp(p) = \mu^- \right\}.$$

Several mathematical questions arise:

• Are all transports of the form $\lambda(p)$ and with p complete ??

Answer: YES if λ is loop-free (consequence of Smirnov Thm)

• Tightness: does $\mathcal{F}(\lambda(p_n)) \leq C$ imply that p_n is tight ?

Answer: Yes but only if p_n is complete

• Does $p_n \rightharpoonup p$ imply that $\lambda(p_n) \rightharpoonup \lambda(p)$?

Answer: No in general but OK if $L(\gamma)p_n(d\gamma)$ is tight !

• Alternative expression for $\mathcal{G}(p) := \mathcal{F}(\lambda(p))$

Needs to check that $\mathcal{G}(p) < +\infty \ \Rightarrow \ \lambda(p)$ is 1- rectifiable

• Lower semicontinity of $\mathcal{G}$?

OK if $g(t)/t$ is monotone decreasing

2.5 Smirnov decomposition

THM. Let μ^+, μ^- be two positive measures (with $\int \mu^+ = \int \mu^-$).
Let $\lambda \in \mathcal{M}(\Omega; \mathbb{R}^d)$ be a transport such that $\operatorname{div} \lambda = \mu^+ - \mu^-$.

If λ is loop-free , then it admits a complete decomposition $\lambda = \lambda(p)$ for a suitable $p \in \mathcal{M}_+(X_\Omega)$.

Futhermore, we have $\mu^+ = e_0^\sharp(p)$, $\mu^- = e_1^\sharp(p)$ so that

$$\int \mu^+ = \int \mu^- = \int_{X_\Omega} p(d\gamma)$$

Proof By SMIRNOV , any transport measure λ can be decomposed in the form:
$\lambda = \lambda(p) + \lambda^0$, where:
i) $\operatorname{div} \lambda^0 = 0$ (λ^0 accounts the loops)
ii) $\int |\lambda| = \int |\lambda(p)| + \int |\lambda^0| = \int L(\gamma)\, p(d\gamma) + \int |\lambda^0|$
iii) $\int |\operatorname{div} \lambda| = \int |\operatorname{div}(\lambda(p))| = 2p(X_\Omega)$
The first equality in ii) and the strict convexity of Euclidean norm implies that $\lambda(p) = \xi\lambda$ for $\xi \in [0, 1]$. As λ is loop free $\xi = 1$ and $\lambda^0 = 0$.

3. Intensity functions and 1-rectifiability Theorem

Given $p \in \mathcal{M}_+(X_\Omega)$, we define $\theta_p(x)$ (simple intensity) and $i_p(x)$ (total intensity) the Borel functions:

$$\theta_p(x) \ := \ p(\{x \in \gamma([0,1])\} \ = \ \int_{X_\Omega} 1_{\{x \in \mathrm{Im}(\gamma)\}} \, p(d\gamma) \ ,$$

$$i_p(x) \ := \ \int_{X_\Omega} \sharp(\{t \in [0,1] \ : \ \gamma(t) = x\}) \, p(d\gamma) \ .$$

- $i_p(x) \geq \theta_p(x)$ but $\theta_p(x) = 0 \Rightarrow i_p(x) = 0$!
- $\theta_p(x) \leq p(X_\Omega)$ and by Smirnov, we deduce Lemma 2.

Lemma 3 The function $(x,p) \in \Omega \times \mathcal{M}_+(X_\Omega) \mapsto \ \theta_p(x)$ **is upper semicontinuous.**

We introduce also the **intensity measure**

$$\mu_p(B) \ := \ \int_{X_\Omega} |\lambda_\gamma|(B) \, p(d\gamma) \ = \ \int_{X_\Omega} \left(\int_{\gamma^{-1}(B)} |\dot{\gamma}(t)| \, dt \right) p(d\gamma).$$

Then $\int \mu_p = \int L(\gamma) \, p(d\gamma)$ $(\geq \ \int |\lambda(p)|)$ with equality if p is complete).

3.2 The functional $\mathcal{G}(p)$

We set $\boxed{\beta(t) := \frac{g(t}{t}}$ and then

$$\mathcal{G}(p) \; := \int_{X_\Omega} \left(\int_0^1 \beta(\theta_p(x))\, h(\dot\gamma(t))\, dt \right) \, p(d\gamma) \; .$$

Justification If p is supported on a simple curve γ_0 i.e. $p = \theta\, \delta_{\gamma_0}$, then denoting $S = \gamma_0([0,1])$, we obtain $\theta_p = i_p = \theta$ on S and

$$\lambda(p) = \theta\, \tau_S\, H^1 \llcorner S \quad , \quad \mathcal{G}(p) = \theta \int_0^1 \beta(\theta)\, h(\dot\gamma_0(t)) \; = \; \int_S g(\theta)\, h(\tau_S)\, dH^1 \; = \; \mathcal{F}(\lambda(p)) \, .$$

LEMMA 4. Assume that g is concave $\nearrow$ and that h is convex l.s.c. postively one-homogeneous. Then the functional $p \in \mathcal{M}_+(X_\Omega) \mapsto \mathcal{G}(p)$ is lower semicontinous (for the weak convergence of measures)

3.3 Rectifiability result

LEMMA 5 (1-rectifiability) Let $p \in \mathcal{M}_+(X_\Omega)$ such that: $\mu_p(\Omega) < +\infty$ and $i_p > 0 \quad \mu_p$ a.e. . Then there exists a 1-rectifiable subset S such that:

$$\mu_p = i_p(x) \, H^1 \llcorner S \, .$$

COROLLARY Assume that $\beta(t) \nearrow +\infty$ as $t \searrow 0$ (with h coercive). Then

$\mathcal{G}(p) < +\infty \ \Rightarrow \ \lambda(p) = \theta \, \tau_S \, H^1 \llcorner S$, with S rectifiable , $0 \leq \theta(x) \leq i_p(x)$ on S.

Proof: Observe that $x \in \gamma([0,1]) \ \Rightarrow \ \liminf_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_{\gamma^{-1}(B(x,\varepsilon))} |\dot{\gamma}(t)| \, dt \ \geq \ 1$.

By Fatou $\liminf_{\varepsilon \to 0} \frac{\mu_p(B(x,\varepsilon))}{2\varepsilon} \ \geq \ \theta_p(x)$. Let $S_k := \{\theta_p > 1/k\}$.
Then $H^1(S_k) < +\infty$ and $\cup S_k \ = \ \{\theta_p > 0\} \ = \ S \cup S'$ with S one-rectifiable.
μ_p has no contribution on the purely unrectifiable part S'. Appying Fubini with $H^1 \llcorner S(dx) \otimes p(d\gamma)$ leads to $\mu_p = i_p \, H^1 \llcorner S$.
Eventually as $|\lambda(p)| \leq \mu_p$, $\lambda(p) = \xi \, \mu_p$ where ξ parallel to τ_S by the divergence condition.

3.4 Relation between $\mathcal{F}$ and G

THM Let $p \in \mathcal{M}_+(X_\Omega)$ a finite measure. Assume that $\beta(t) = g(t)/t$ is nonincreasing with $\beta'(0+) = +\infty$. Then

(i) $$\mathcal{F}(\lambda(p)) \leq \mathcal{G}(p) \quad \text{for every } p$$

(ii) $$\mathcal{F}(\lambda(p)) = \mathcal{G}(p) \quad \text{if } p \text{ is complete}$$

Hint: $\mathcal{G}(p) < +\infty$ implies that $\theta_p > 0$ on the support of μ_p. Then we apply the rectifiability result to $\lambda(p)$.

COROLLARY As every transport can be substituted with another one loop-free with lower energy (and also with complete decomposition):

$$\inf \{\mathcal{F}(\lambda) \ : \ \operatorname{div} \lambda = \mu^+ - \mu^-\} = \inf \left\{\mathcal{G}(p) \ : \ e_0^\sharp(p) = \mu^+ \ , \ e_1^\sharp(p) = \mu^-\right\} \ ,$$

$$\operatorname{Argmin} \mathcal{F} = \{\lambda(p) \ : \ p \in \operatorname{Argmin} \mathcal{G}\}$$

STILL NEEDS TO PROVE CONVERGENCE OF MINIMIZING SEQUENCES (p_n) for $\mathcal{G}$ in $\mathcal{M}_+(X_\Omega)$!

4. Tightness results and lower semicontinuity of $\mathcal{F}$

We already know that loop-free minimizing sequence of transports are uniformly bounded in variation hence weakly compact in $\mathcal{M}(\Omega; \mathbb{R}^d)$.

As regards minimizing sequences (p_n) in $\mathcal{M}_+(X_\Omega)$, we need to check the **Prokhorov's tightness criterium**.

- **Complete minimizing (p_n) are precompact**

Assume that $\int \mu^+ = \int \mu^- = 1$. If p_n is complete, $\theta_{p_n} \leq 1$ (by Lemma 2) and $g(\theta_{p_n}) \geq g(1)\,\theta$ (by the concavity of g). Thus

$$\mathcal{G}(p_n) \;\geq\; g(1) \int_{X_\Omega} L(\gamma)\, p_n(d\gamma) \;=\; g(1) \int_{X_\Omega} \mu_{p_n}.$$

For every l, the compact set $K_l = \{L(\gamma) > l\}$ satisfies $p_n(K_l) \leq \; C\, l$.

$\Rightarrow$ EXISTENCE OF SOLUTIONS

$\mathrm{Argmin}\ \mathcal{G}$ is non empty (by the l.s.c. of $\mathcal{G}$)

$\mathrm{Argmin}\ \mathcal{F}$ is non empty (by the equivalence between $\mathcal{F}$ and $\mathcal{G}$)

Remark If g is stricly $\searrow$, then optimal transports are loop-free.

4.2 Reinforced compactness

- **Why need reinforced compactness ?**

It is natural to ask wether or not

$$p_n \rightharpoonup p \quad \Rightarrow \quad \lambda(p_n) \rightharpoonup \lambda(p_n)$$

The answer is:

NO if we know merely that $\sup_n \int L(\gamma)\, p_n(d\gamma) < +\infty$

YES if the sequence $\{L(\gamma)\, p_n\}$ is tight.

Hint: The map $\Lambda: \ \gamma \in X_\Omega \mapsto \lambda_\gamma \in \mathcal{M}(\Omega; \mathbb{R}^d)$ is not continuous.

Ex.: Let $p_n := \frac{1}{n+1}\, \delta_{\gamma_n}$, $\gamma_n(t) := (1 + n^{-1})\, \exp(2i\pi nt)$. Then $p_n \rightharpoonup 0$ whereas $\lambda(p_n) \rightharpoonup \lambda_{\gamma_1}$ (factor $(1 + n^{-1})$ in order to have simple curves)

However Λ **is continuous on all compact** K_l and the contribution from $X_\Omega \setminus K_l$ is controlled by

$$\int_{X_\Omega \setminus K_l} \|\lambda_\gamma\|\, p_n(d\gamma) \quad \left(= \int_{X_\Omega \setminus K_l} L(\gamma)\, p_n(d\gamma) \right)$$

4.3 Sub-transport estimate

- **More on sub-transports**

Let $p \in \mathcal{M}_+(X_\Omega)$ and E a Borel subset. Then $\lambda(p \llcorner E)$ is a sub-transport of $\lambda(p)$.

Thus as g is **non decreasing:** $\mathcal{F}(\lambda(p \llcorner E) \leq \mathcal{F}(\lambda)$. Besides if p is complete, so is $p \llcorner E$ and $\theta_{p \llcorner E} \leq p(E)$ (by Lemma 2). Now exploiting $\beta \searrow$

Lemma 6: For every Borel subset $E \subset X_\Omega$ and every complete $p \in \mathcal{M}_+(X_\Omega)$, there holds:

$$\mathcal{G}(p) \ \geq \ G(p_E) \ \geq \ h_{\min} \, \beta(p(E)) \int_E L(\gamma) \, p(d\gamma).$$

Applying to $E = K_l$, we get

$(**)$ $\sup_n G(p_n) < +\infty \ \Rightarrow \ \{L(\gamma) \, p_n\}$ tight $\Rightarrow \ \lambda(p_n) \rightharpoonup \lambda(p)$ whenever $p_n \rightharpoonup p$

4.4 Lower semicontinuity of $\mathcal{F}$

- **A weak statement for the l.s.c of $\mathcal{F}$**

We keep the previous assumptions on g, h.

THM Let (λ_n) be a sequence of **loop-free** transport measures such that:
$\lambda_n \rightharpoonup \lambda \quad , \quad \operatorname{div} \lambda_n \rightharpoonup \operatorname{div} \lambda$.
Then $\liminf_n \mathcal{F}(\lambda_n) \geq \mathcal{F}(\lambda)$.

Proof There exists a sequence of complete measures p_n such that $\lambda_n = \lambda(p_n)$. As $\operatorname{div} \lambda_n$ is upperbounded in variation so is p_n and we may assume $p_n \to p$. Then, as we know that G is l.s.c.:

$$\liminf \mathcal{F}(\lambda_n) = \liminf \mathcal{G}(p_n) \geq \mathcal{G}(p) \geq \mathcal{F}(\lambda(p)).$$

By the implication (**): $\lambda(p)$ coincides with λ (at least when $\mathcal{F}(\lambda_n)$ is bounded)

4.5 A case excluding loops

Let $\Sigma := \{z \in S^{d-1} \ : \ h(z) < +\infty\}$ and assume

$$(***) \qquad \exists z^* \in S^{d-1} \quad : \quad z^* \cdot z < 0 \quad \forall z \in \Sigma$$

Then $\mathcal{F}(\lambda) < +\infty$ implies that λ has no loop

Proof A transport $\lambda = \theta\, \tau_S\, H^1 \llcorner S$ with $\mathcal{F}(\lambda) < +\infty$ satisfies $\tau_S \in \Sigma$ a.e. on S.
Thus $\int_S \theta\, \tau_S \cdot z^* dH^1 < 0$: incompatible with $\operatorname{div} \lambda = 0$ which forces $\int \lambda = 0$.

CONCLUSION: Under (***), $\mathcal{F}$ is lower semicontinuous

Model example:

- $z \in \mathbb{R}^d$ substituted with $(z, t) \in \mathbb{R}^d \times \mathbb{R}_+$ (space and time).
- $f : \mathbb{R}^d \to \mathbb{R}_+$ is convex l.s.c. (function of the speed)

$$h(z,t) = \begin{cases} t\, f(\frac{z}{t}) & \text{if } t > 0 \\ f^\infty(z) & \text{if } t = 0 \\ +\infty & \text{if } t < 0 \end{cases}$$

Then (***) holds for $p > 1$. If $p = 1$, loops can occur in slices of time $\{t^*\} \times \mathbb{R}^d$ (meaning that the speed in the loop is infinite !)

5. Application to dynamic formulations

- The transport of discrete measures is described at time $t \in [0,T]$ by:

$$\rho_t = \sum_i c_i(t)\, \delta_{x_i(t)} \qquad \text{(fonction from } [0,T] \text{ to } \mathcal{M}_+(\Omega))$$

$$\sigma_t = \sum_i c_i(t)\, \dot{x}_i(t)\, \delta_{x_i(t)} \qquad \text{(“momentum ”function from } [0,T] \text{ to } \mathcal{M}_+(\Omega;\mathbb{R}^d))$$

- We associate the measure $\lambda = (\sigma, \rho) \in \mathcal{M}(\Omega \times \mathbb{R}; \mathbb{R}^{d+1})$ by setting:

$$< \rho, \varphi(x,t) >= \int_0^T (\int_\Omega \varphi(x,t)\, \rho_t(dx))\, dt\, ,\ < \sigma, \phi(x,t) >= \int_0^T (\int_\Omega \phi(x,t) \cdot \sigma_t(dx))\, dt.$$

- As far as finite speeds are considered: $\sigma \ll \rho$ and $V(x,t)$ denotes the Radon-Nikodym density. The kinematic contraints (including the one at the junctions) and boundary conditions $\rho(0) = \mu^+$, $\rho(T) = \mu^-$ reduce to:

$$(dyn) \qquad \frac{\partial \rho}{\partial t} + \operatorname{div}_x(\rho V) = \mu^+ \otimes \delta_0 - \mu^- \otimes \delta_T \quad \text{as distributions on } \mathbb{R}^{d+1}$$

Remark In fact under (dyn) and condition $\int_{\mathbb{R}^{d+1}} |V|\, \rho(dxdt) < +\infty$, ρ is of the form $\int_0^T \rho_t(dx) \otimes dt$ where the map $t \mapsto \rho_t \in \mathcal{M}(\mathbb{R}^d)$ is continuous (weak topology) (see Ambrosio).

5.2 Parametrized curves in space-time

Let us argue on a situation with finitely many masses and junctions points (the number of points may change in the process !)

- Each trajectory $(x_i(t), t)\ :\ t \in [0, T|$ can be seen as an oriented curve $\gamma_i : s \in [0,1] \mapsto R^{d+1}$, with a Lipschitz $\nearrow$ parametrization $t = t(s)$ such that $t(0) = t_i^-, t(1) = t_i^+$.

- The measure $\lambda = (\sigma, \rho)$ is supported on the one dimensional subset $S = \cup_i S_i\ ,\ S_i = \mathrm{Im}(\gamma_i)$ and has the form $\lambda = \theta \tau H^1_{|S}$ where $\tau = (\tau_x, \tau_t)$ is a unit tangent vector to S with $\tau_t > 0$.

The ratio $\dfrac{|\tau_x|}{\tau_t}$ represents then the velocity V already introduced.

- A simple change of variables on the graph S_i shows that

$$\int_{t_i^-}^{t_i^+} g(c_i(t))\, f(\dot{x_i})\, dt = \int_{S_i} f(\frac{\tau_x}{\tau_t}) g(c_i(t)) \tau_t H^1(dx)$$

so that summing over i, the total cost coincides with $\mathcal{F}(\lambda) = \int_S h(\tau) g(\theta)\, dH^1$, where $\theta(x,t)$ is roughly $\sum\{c_i(t)\ :\ x_i(t) = x\}$ and h is the function defined on $\mathbb{R}^{d+1}$ by: $h(\tau) := f(\dfrac{\tau_x}{\tau_t})\tau_t$ if $\tau_t > 0$, $h(\tau) = +\infty$ otherwise.

5.3 Reformulation of the problem

We end up with the variational problem:

$$\inf \left\{ \int_0^T \mathcal{F}'(V\rho, \rho)\, dt\ ,\ \frac{\partial \rho}{\partial t} + \mathrm{div}(V\rho) = 0\ ,\ \rho(0) = \mu^+\ ,\ \rho(T) = \mu^- \right\}$$

where:

$$\mathcal{F}'(V\rho, \rho) \;=\; \int_S f\left(\frac{\tau_x}{\tau_t}\right) g(\theta(x,t))\tau_t\, dH^1 \;=\; \int_S g(\theta)\, h(\tau_S)\, dH^1.$$

EXISTENCE follows from the previous section

Remark The fact that $\sigma \ll \rho$ is forced if f has superlinear growth. In this case we know also that loops are also ruled out.

The case $f(z) = |z|$ is more delicate !!

6. Optimality conditions and approxiamation

The curve representation allows a very easy derivation of first order optimality conditions:

Let $\lambda = \lambda(p)$ an optimal transport. Recall that p is complete ($\theta_p = i_p$) and satisfies $e_1^\sharp(p) = \mu^+$, $e_0^\sharp(p) = \mu^-$.

We consider a deformation map: $\Psi_\varepsilon(x) = x + \varepsilon V(x)$ where $V(x)$ is a smooth vector field such that: $V = 0$ on $\operatorname{spt}\mu^+ \cup \operatorname{spt}\mu^-$.

Then set $p_\varepsilon = \Psi_\varepsilon^\sharp(p)$. Notice that for all γ, $\theta_{p_\varepsilon}(\Psi_\varepsilon \circ \gamma) = \theta_p(\gamma)$, there holds:

$$
\begin{aligned}
0 \;=\; \lim_{\varepsilon\to 0} \frac{\mathcal{G}(p_\varepsilon) - G(p)}{\varepsilon} \;&=\; \int_{X_\Omega} \int_0^1 \left(\beta(\theta_p)(\gamma) \, \frac{h(\dot\gamma + \varepsilon \nabla V(x)\dot\gamma) - h(\dot\gamma)}{\varepsilon} \right) dt\, p(d\gamma) \\
&=\; \int_{X_\Omega} \int_0^1 \left(\beta(\theta_p)(\gamma(t))\, \dot\gamma(t) \otimes \nabla h(\gamma) \nabla V(\gamma) \right) dt\, p(d\gamma) \\
&=\; \int_S g(\theta(x)) \, \frac{\tau \otimes \nabla h(\tau)}{h(\tau)} \cdot \nabla V(x)\, dH^1
\end{aligned}
$$

$$
\Longrightarrow \qquad \boxed{\operatorname{div}\left(g(\theta)\, \frac{\tau \otimes \nabla h(\tau)}{h(\tau)}\, H^1 \llcorner S \right) \;=\; 0 \,.}
$$

6.2 Approximation by viscosity

We condider Stokes flow with small viscosity (ε on a bounded domain $\Omega \subset \mathbb{R}^3$ and add a non linear potential $V : \mathbb{R}^3 \to \mathbb{R}^+$

$$\inf \left\{ \int_\Omega (V(u) + \varepsilon\, |\nabla u|^2) \ : \ \operatorname{div} u = \mu^+ - \mu_- \right\}$$

Choose $V(u) \ := \ |u|^p$ where $0 \le p < 1$ (so that it is better to concentrate the flow)

Remark: $p = 0$ is equivalent to find the optimal shape of the set $\{u \neq 0\}$

CONJECTURE: As $\varepsilon \to 0$, the minimizers u_ε (in $H^1(\Omega; \mathbb{R}^3)$) concentrate on 1-dimensional subsets: $u_\varepsilon \rightharpoonup \lambda$ (in the sense of measures) where λ solves

$$\inf \left\{ \int_S \theta^\alpha \, dH^1 \ : \ \lambda = \theta\, \tau_S H^1 \llcorner S \ : \ \operatorname{div} \lambda = \mu^+ - \mu_- \right\}$$

The scaling law is deduced from the 2D profile equation on a disk D_r:

$$-\Delta\varphi + |\varphi|^{p-2}\varphi = cte \ , \ \ \varphi \in H^1_0(D_r)$$

and its asymptotic as $r \to 0$ (see GB, Dubs, Seppecher, CRAS and M3AS (1997))

We obtain

$$\boxed{\alpha \ = \ \frac{2}{3-p}}$$