# Optimal Transport : Regularity and applications

Thomas Gallouët

**N° d'ordre : 797**

**N° attribué par la bibliothèque : 2012ENSL0797**

# THÈSE

en vue d'obtenir le grade de

**Docteur de l'École Normale Supérieure de Lyon -Université de Lyon-**

**Spécialité : Mathématiques**

**Unité de Mathématiques Pures & Appliquées - UMR 5669 - UMPA**

**École Doctorale Informatique & Mathématiques**

présentée et soutenue publiquement le 10 décembre 2012 par

M. Thomas Gallouët

---

**Titre :**

# Transport optimal : régularité et applications

---

Directeur de thèse :

M. Cédric VILLANI

Après avis de :

M. José Antonio CARRILLO DE LA PLATA

M. Young-Heon KIM

Devant la commission d'examen formée de :

| | |
|---|---|
| M. Vincent CALVEZ | *membre* |
| M. José Antonio CARRILLO DE LA PLATA, | *membre/rapporteur* |
| Mme. Lucilla CORRIAS | *membre* |
| M. Ludovic RIFFORD, | *membre* |
| Mme. Sylvia SERFATY, | *membre* |
| M. Cédric VILLANI, | *membre/directeur* |
| M. Hatem ZAAG, | *membre* |

*The piano has been drinking, not me*

Tom Waits

*Tadoum*

Gwenaël

# Remerciements

Tout d'abord je voudrais remercier chaleureusement Cédric Villani d'avoir encadré cette thèse. Chaque cours, chaque conférence, chaque discussion avec Cédric est un plaisir et une source incroyable de motivation. Mais de façon aussi importante sa patience et son soutien ont été determinants pour moi.

Je remercie mes rapporteurs José Antonio Carrillo et Young-Heon Kim dont les remarques pertinentes ont grandement amélioré ce manuscrit.

Je suis très honoré de la présence dans mon Jury de Vincent Calvez, José Antonio Carrillo, Lucilla Corrias, Ludovic Rifford, Sylvia Serfaty et Hatem Zaag. Leurs travaux m'ont beaucoup inspiré et je les remercie d'avoir accepté de faire partie de ce jury.

Une thèse ce n'est pas seulement un travail, c'est une période de vie indissociable des rencontres, des interactions qu'elles soient scientifiques ou non. Pour cela je voudrais remercier tout particulièrement Neil Trudinger de m'avoir accueilli en Australie. Ses travaux ont occupé une grande partie de ma thèse et cela va probablement continuer. La disponibilité des chercheurs et le style de vie Australien font que ces séjours à Canberra resteront gravés dans ma mémoire. Je remercie le MSI pour son accueil et notamment Jiakun Liu. Merci à Andrea, Colin, Carlo, Pierre pour m'avoir aidé à calculer le tenseur MTW, appris à faire les vraies pasta al mare di Napoli, pasta carbonara, pour les weekends, les soirées et fin de soirées. It was a very good experience, mate. Merci à Lionel pour sa gentillesse et les sorties d'escalade simplement exceptionnelles. Enfin merci à tous les joueurs de casual soccer, ces moments ont été une bouffée d'air. Ces séjours en Australie n'auraient pas été possibles sans l'aide précieuse de Philippe Delanoë, un grand merci à lui.

Un autre lieu très important dans ma thèse est le Laboratoire Dieudonné à Nice. Je remercie infiniment Ludovic Rifford pour ses invitations, sa disponibilité, les nombreuses discussions et relectures. C'est là-bas que s'est écrite la majorité des résultats de la première partie de cette thèse. Sur ce point merci également à Alessio Figalli pour ses idées toujours justes et son aide efficace. L'ambiance amicale régnant dans l'équipe Géométrie et Analyse fait que c'est toujours un plaisir de venir à Nice. Merci aussi à Dominique, Bernard, Marine, Romain pour leur accueil nocturne et dominical.

Un paragraphe pour lui tout seul, merci beaucoup à Vincent Calvez pour le plaisir de travailler avec lui, son enthousiasme communicatif, le temps passé à répondre à mes mails, relire ma thèse et tellement plus encore.

Merci aux nordistes Claire Chainais-Hillairet et Antoine Gloria pour leur patience.

Merci au laboratoire de L'UMPA pour toutes les années passées ici, les conditions de travail y sont absolument exceptionnelles. Ceci n'est pas seulement dû au budget conséquent mais surtout à ses membres. En particulier merci beaucoup à Magalie et Virginia pour leur aide, je n'ai jamais trouvé un secrétariat aussi efficace et chaleureux. À Gérard pour avoir réglé tous mes problèmes informatiques, à Claude Danthony pour son attention. Merci à ceux qui ont partagé mon bûro, Pierre et Maxime pour m'avoir appris de vraies maths et plus récemment Michele pour y avoir introduit un livre de finance. Un jour je vous ferai aimer le tenseur de Ma-Trudinger-Wang. L'ambiance au labo a toujours été studieuse et détendue. Merci à Rémi, Séverine, Maxime Z., Paul L., Virginie, Hélène, Marco, Ana, Romain, Frederic, Fred, Paul V., Alessandro, Pierre-Adelin, Sébastien, Léa, Ludovic,... pour les tds, les discussions, les bières au Ninkasi et les pauses en passerelle. Une mention particulière pour la relecture nocturne mais néanmoins efficace de Michele et Marielle.

Il n'est pas nécessaire d'habiter à Lyon pour boire une bière, discuter et faire du ski merci à Chen, Nathalie, Bastien, Julien, Lionel, Séverine, Veronika, Matteo, Nicky, Cédric, Thérèse, Virginie, Julien, Stacha, Soazig pour les bons moments passés un peu partout en France.

Durant cette thèse je suis passé par des moments plus ou moins agréables. Merci au docteur (the other kind of doctor) Rhul et à tout le personnel du centre Rocheplane : aide-soignantes, infirmières, hôteliers, kinés pour avoir rendu mon séjour chez eux le plus agréable qu'il puisse être. Scientifiquement ce n'est pas ma période la plus productive mais si cette thèse a pu voir le jour c'est aussi grâce à eux. *Des girandoles prolongent, dans les vergers et les allées voisins du Méandre, les verts et les rouges du couchant.*

Merci à Pascale, Émilie, Capucine, Mathilde, Ysabel, Anatole, Raphaèle, Thierry, et plus généralement à toute ma famille et ma belle-famille pour leur soutien, leur aide que ce soit pour la thèse, le pot ou dans la vie de tous les jours. À mes grand-mères pour m'avoir hébergé, nourri, blanchi pendant des périodes de rédactions. Vous m'avez promis de tout lire.

Merci à tous ceux avec qui je me suis encordé un jour : Aurélien, Edern, Raf, Mathieu, Pierre, Loic, Clément et Gilles.

Et par-dessus tout merci à Cécile et Gwenaël grâce à qui je me lève (tôt) chaque matin avec le sourire.

# Table des matières

# Présentation de la thèse

## 0.1 Le problème de Monge

Le transport optimal est un sujet ancien, abordé en premier lieu par Monge en 1781 [90]. Le problème pour Monge était de minimiser le coût de transport d'un déblai vers un remblai. Une formulation actuelle de ce problème est la suivante :

**Définition 0.1.1** (Problème de Monge). *Soit $M$ une variété riemannienne, $c : M \times M \to \mathbb{R}$ une fonction de coût et $\mu, \nu$ deux mesures de probabilités à densité par rapport au volume. Un transport $\mathbf{t}$ de $\mu$ vers $\nu$ est une application de $M$ dans $M$, $\mu$ mesurable, telle que $\mathbf{t}_{\#}\mu = \nu$ (la mesure image de $\mu$ par $\mathbf{t}$ est égale à $\nu$). Cela signifie que pour tout ensemble $B$ mesurable dans l'espace d'arrivée on a $\nu(B) = \mu\left(\mathbf{t}^{-1}(B)\right)$. Un transport optimal $T$ est une solution $\mu$ mesurable, quand elle existe, du problème de minimisation :*

$$T = \underset{\mathbf{t}_{\#}\mu=\nu}{\operatorname{argmin}} \left( \int_M c\left(x, \mathbf{t}(x)\right) d\mu(x) \right), \tag{0.1.1}$$

*c'est-à-dire le transport de coût minimum.*

La mesure $\mu$ correspond au remblai qu'il faut déplacer et la mesure $\nu$ au déblai, c'est-à-dire au lieu de stockage.

Dans son livre, Monge propose une méthode de construction du transport optimal sans prouver que cette méthode est effectivement réalisable. Cette question n'est pas pertinente à l'époque, car on pense alors que tout problème physique possède une unique solution.Il touche cependant du doigt le concept de $c$-convexité, qui s'avérera des années plus tard un outil indispensable pour les théorèmes d'existence. Longtemps après Monge, dans les années 1940-1960, Kantorovich propose une interprétation du problème de Monge par dualité [67, 68] :

**Définition 0.1.2** (Problème de Kantorovich). *Soit $M$ une variété riemannienne, $c : M \times M \to \mathbb{R}$ une fonction de coût et $\mu, \nu$ deux mesures à densité par rapport au volume. On cherche le couplage optimal $\Pi \in \mathbb{P}(M \times M)$ solution du problème de minimisation suivant :*

$$\Pi = \underset{\pi,\ \pi_1=\mu, \pi_2=\nu}{\operatorname{argmin}} \left( \int_M c\left(x, y\right) d\pi(x, y) \right), \tag{0.1.2}$$

*Par définition $\pi_1 = \mu$ signifie que la première marginale de $\pi$ est égale à $\mu$, c'est-à-dire que pour tout ensemble $B$ mesurable on a $\pi_1\left(B \times M\right) = \mu\left(B\right)$. La définition de $\pi_2$ est semblable pour la seconde marginale de $\pi$.*

Cette définition est un peu plus générale, c'est une relaxation du problème de Monge, car on peut par exemple partager la masse « localisée »en un point. Le problème de Kantorovich est linéaire, l'existence d'une telle solution se ramène à un problème de compacité faible sur l'espace de départ. Kantorovich montre que le problème de minimisation défini par 0.1.2 est équivalent au problème de maximisation suivant :

**Définition 0.1.3** (Problème de Kantorovich dual). *Soit $M$ variété riemannienne, $c : M \times M \to \mathbb{R}$ une fonction de coût et $\mu, \nu$ deux mesures à densité par rapport au volume. On cherche deux fonctions $\phi$, $\psi$ définies sur $M$ à valeurs dans $\mathbb{R}$ telles que :*

$$(\phi, \psi) = \underset{(\phi,\psi)\in E_c}{\operatorname{argmax}} \left( \int_M \phi(x)\mu(x) + \int_M \psi(y)\nu(y) \right),$$

*où*

$$E_c = \{(\phi, \psi) \mid \quad \forall (x, y) \in M \times M \quad \psi(y) - \phi(x) \leq c(x, y)\}.$$

Lorsque le maximum est atteint dans la définition 0.1.3, on a :

$$\phi(x) = \max_y (\psi(y) - c(x, y)).$$

On récupère alors $\psi$ par la formule :

$$\psi(y) = \min_x (\phi(x) + c(x, y)) =: \phi^c$$

On dit alors que $\phi$ est une fonction $c$-convexe et que $\psi$ est sa $c$-conjuguée. Cette définition dans le cas euclidien donne la convexité usuelle lorsque le coût est égal à la norme au carré, ou de façon équivalente égal à $-x \cdot y$ ($|x - y|^2 = |x|^2 + |y|^2 - x \cdot y$). C'est la formulation par dualité de Frenchel-Moreau [16]. Une bonne analogie est la suivante : une fonction $x \mapsto \varphi(x)$ convexe est une fonction sous laquelle on peut coller en tout point une droite, c'est à dire que pour tout $x_0$ il existe un $y$ tel que pour tout $x$ :

$$\varphi(x) \geq -x \cdot y + x_0 \cdot y + \varphi(x_0).$$

De la même manière un coût $c$ quelconque donne naissance à une famille de fonctions support (parametrée par $y$), appelées $c$-supports :

$$x \mapsto -c(x, y) + cste.$$

Une fonction $\phi$ $c$-convexe est alors une fonction telle que en tout point on peut coller sous la fonction un de ces $c$-supports. Explicitement, pour tout $x_0$, il existe un $y$ tel que pour tout $x$ :

$$\phi(x) \geq -c(x, y) + c(x_0, y) + \phi(x_0).$$

Le lien entre la mesure $\Pi$ et la fonction $\phi$ du problème dual est que le support de $\Pi$ est concentré sur l'ensemble des $c$-supports de $\phi$.

Le transport optimal n'a ensuite plus été étudié jusqu'à ce que Brenier s'y réintéresse dans les années 1980, dans un cadre assez inattendu de dynamique des gaz [14]. Il démontre alors l'existence et l'unicité d'un transport optimal pour un coût quadratique dans $\mathbb{R}^n$. McCann apporte ensuite la solution au problème de Monge dans le cas d'une variété riemannienne, avec le coût donné par la distance géodésique quadratique (coût associé au lagragngien$L = \frac{1}{2}|v|^2$) [88]. Dès lors le transport optimal connaît un développement très rapide dans de nombreux domaines, souvent de façon inattendue. Dans cette thèse nous nous intéressons à deux directions différentes, traitées dans deux parties distinctes. La première se concentre sur la régularité du transport optimal et sur les conditions géométriques qui l'accompagnent. La seconde est l'application de la théorie du transport optimal pour l'étude de certaines équations aux dérivées partielles.

## 0.2 Régularité du transport optimal, conséquences géométriques

La question de la régularité du transport optimal est liée à la régularité des solutions d'équations complètement non-linéaires (fully non-linear), typiquement les équations de Monge-Ampère :

*Definition* (Équation de Monge-Ampère). Une fonction $\varphi \in C^2(\mathbb{R}^n, \mathbb{R})$ est solution classique d'une équation de Monge-Ampère si, pour un $h \in C^0(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$ donné, on a :

$$\det\left[\nabla^2\varphi(x)\right] = h\left(x, \nabla\varphi(x)\right) \quad x \in \mathbb{R}^n. \tag{0.2.1}$$

Par exemple, plaçons nous dans le cas euclidien et donnons nous deux mesures à densité $C^\infty$ strictement positives $\mu(x) = f(x)dx$ et $\nu(y) = g(y)dy$. Le transport optimal $T$ tel que $T_\#\mu = \nu$ est alors donné par un potentiel $\varphi$ (c'est-à-dire que $T = \nabla\varphi$) convexe qui vérifie l'équation de Monge-Ampère avec $h(x) = \dfrac{f(x)}{g\left(T(x)\right)}$. On parle de l'équation de Monge-Ampère avec le second type de condition aux bords : $T$ envoi le support de $\mu$ sur celui de $\nu$. Le calcul qui suit permet de comprendre formellement d'où vient ce lien.

Par définition $T_\#\mu = \nu$, donc pour tout ensemble $B$ mesurable dans l'espace d'arrivée on a :

$$\int_{T^{-1}(B)} f(x)dx = \int_B g(y)dy$$

On suppose que $T$ est assez régulier pour faire le changement de variable $y = T(x)$. Il vient :

$$\int_{T^{-1}(B)} f(x)dx = \int_{T^{-1}(B)} g(T(x)) \det\left[\nabla T(x)\right] dy = \int_{T^{-1}(B)} g(T(x)) \det\left[\nabla^2\varphi(x)\right] dy$$

Il apparaît donc qu'une condition suffisante pour l'existence de $T$ est bien que pour tout $x$ de l'ensemble de départ,

$$\det\left[\nabla^2\varphi(x)\right] = \frac{f(x)}{g\left(T(x)\right)}.$$

Dans le cas plus général d'un coût $c(x,y)$, on obtient l'équation dite de Monge-Ampère généralisé :

$$\det\left[\nabla^2\varphi(x) + \nabla_{xx}c\left(x, T(x)\right)\right] = \frac{f(x)}{g\left(T(x)\right)} \nabla_{x,y}c\left(x, T(x)\right).$$

Les premières preuves de la régularité du transport optimal dans le cas d'un coût quadratique ont donc été obtenues par des spécialistes de la régularité des équations aux dérivées partielles complètement non-linéaires (*fully non-linear*), comme Delanoë, Caffarelli et Urbas [33, 18, 19, 20, 22, 107, 106]. La principale difficulté pour ce cas de transport est en effet une condition aux bords d'un genre nouveau. Le passage à un coût plus général est resté longtemps une gageure. Ce verrou a sauté après l'introduction par Ma, Trudinger et Wang d'un tenseur dont la positivité contrôle l'existence de solutions régulières au problème de Monge-Ampère généralisé [85, 105].

Le tenseur de Ma-Trudinger-Wang (MTW), assez mystérieux lors de son introduction, a été depuis très étudié. Il apparaît que sa positivité contient des informations non triviales sur la géométrie de la variété riemannienne sur laquelle on se place. Il impose par exemple la positivité des courbures sectionnelles. Ce tenseur est aussi appelé « tenseur de courbure croisé ». Une raison est qu'il est défini de manière non locale sur une variété $M$, alors que sa formulation devient locale si on se place sur le bi-produit $M \times M$. Comme par ailleurs l'information sur la courbure de $M$ est incluse dans **MTW**, il est naturel d'interpréter ce tenseur comme une courbure généralisée sur $M \times M$ [75].

La première partie de cette thèse, dédiée à la régularité du tenseur MTW, comporte trois chapitres : le premier rappelle différentes applications du tenseur MTW et s'intéresse à sa positivité pour certains coûts lagrangiens sur une variété riemannienne. Les deux autres chapitres explorent la régularité des lieux injectifs et focaux d'une variété riemannienne, en particulier lorsque MTW est positif.

### 0.2.a Interprétation du tenseur MTW

Le tenseur de Ma-Trudinger-Wang est d'ordre 4, il prend en entrée deux points d'une variété riemannienne $M$ et un vecteur tangent à la variété en chacun de ces points. Soit $x \in M$, $y \in M \setminus \mathrm{cut}(M)$ et $(\xi, \eta) \in T_xM \times T_xM$. On introduit $v \in \mathrm{I}(x)$ la vitesse telle que $\exp_x(v) = y$. Dans le cas du coût géodésique quadratique, $\frac{1}{2}d^2$, le tenseur de Ma–Trudinger–Wang en $(x, y)$ (ou $(x, v)$) évalué sur $(\xi, \eta)$ est défini par :

$$\mathfrak{S}_{(x,v)}(\xi, \eta) = -\frac{3}{2} \left.\frac{d^2}{ds^2}\right|_{s=0} \left.\frac{d^2}{dt^2}\right|_{t=0} \frac{d^2}{2}\Big(\exp_x(t\xi), \exp_x(v + s\eta)\Big). \tag{0.2.2}$$

Historiquement ce tenseur a été introduit par Ma, Trudinger et Wang pour des vitesses $\xi$ et $\eta$ orthogonales [85]. Kim et McCann ont supprimé cette restriction et introduit ainsi la terme de tenseur de courbure croisée [75]. L'introduction de la vitesse $v$ dans la définition permet de tout voir depuis le point $x$. Cette vitesse est bien définie de manière unique par $y$ et de plus $\left.\frac{d}{ds}\right|_{s=0} \exp_x(v + s\eta)$ est bien un élément de $T_yM$. Les définitions relatives au contexte riemannien sont données dans l'annexe A.

Dans le cas d'un coût $c$ vérifiant les hypothèses de régularités (1.1.a), on a une définition similaire détaillée au chapitre 1. Le tenseur MTW en $(x, y)$ évalué sur $(\xi, \eta)$ est alors noté $\mathfrak{S}_c(x, y)(\xi, \eta)$ ou MTW ou tenseur de courbure croisé.

**Définition 0.2.1.** *Pour tous réels $K$ et $C$, on dit que* $\mathbf{MTW}(K, C)$ *est vrai si pour tout* $(x, y) \in$

$M \times M \setminus c\text{-cut}(M)$ *et* $(\xi, \eta) \in T_xM \times T_xM$*, on a :*

$$\mathfrak{S}_c(x,y)(\xi,\eta) \geq K|\xi|^2|\eta|^2 - C|\langle \xi, \eta\rangle||\xi||\eta|.$$

*Dans le cas où l'on se restreint à des vecteurs tangents orthogonaux* $\langle \xi, \eta \rangle = 0$*, on dit que* $\mathbf{MTW}(K)$ *est vrai si :*

$$\mathfrak{S}_c(x,y)(\xi,\eta) \geq K|\xi|^2|\eta|^2.$$

*Cette définition a été introduite par Ma, Trudinger et Wang.*

Les premières interprétations de ce tenseur ont été proposées par Loeper dans le cas du coût donné par la distance géodésique quadratique. Il montra tout d'abord que si l'on se restreint à $x = y$, c'est-à-dire à $v = 0$, la positivité du tenseur MTW implique la positivité de toutes les courbures sectionnelles [82], Kim prolongea la réflexion en donnant des exemples de varíés riemmaniennes de courbures sectionnelles positives partout et pourtant ne vÃľrifiant pas $\mathbf{MTW}(0)$ [72].

**Théorème 0.1** (Loeper)**.** *Soit M une variété riemannienne vérifiant* $\mathbf{MTW}(0)$ *(respectivement* $\mathbf{MTW}(K)$*) pour le coût* $\frac{1}{2}d^2$*, alors les courbures sectionnelles sont positives (respectivement* $\geq K$*).*

Dans le chapitre 1, on démontre cette propriété par une méthode un peu différente de celle utilisée par Loeper. Une seconde découverte de Loeper, approfondie par Kim Figalli et McCann, est que ce tenseur conditionne la quasiconvexité, et convexité de l'ensemble des fonctions $c$-convexes [82, 44]. Précisément on a le théorème suivant :

**Théorème 0.2** (Figalli-Kim-McCann/Loper)**.** *Soit c un coût vérifiant les conditions* (1.1.a)*, alors* $\mathbf{MTW}(0,0)$ *(resp.* $\mathbf{MTW}(0)$*) est équivalent à la convexité (resp. quasiconvexité) de l'ensemble des fonctions c-convexes.*

Ce théorème n'est pas très surprenant si on pense à la régularité des solutions de l'équation de Monge-Ampère. En effet, dans le cas euclidien avec un coût quadratique, le membre de gauche de l'équation (0.2.1) peut être interprété, au sens d'Alexandrov, comme la mesure du sous-différentiel de la fonction $\varphi$. Dès lors qu'il y a plus de deux vecteurs dans ce sous-différentiel, la convexité implique que la mesure de Hausdorff $\mathrm{H}^1$ est non nulle ce qui est impossible, comme montré par exemple par Figalli et Loeper dans le cas de la dimension 2 [45]. Dans le cas d'une fonction de coût quelconque, on cherche à mesurer le $c$-sous différentiel. C'est donc bien la convexité de ce $c$-sous différentiel qui va être cruciale. En toute généralité la convexité n'est pas a priori nécessaire, il suffit que deux éléments dans le $c$-sous différentiel permettent d'obtenir une famille de mesure de Hausdorff non nulle.

Toujours pour faire plus ample connaissance avec le tenseur de Ma-Trudinger-Wang, le chapitre 1 contient le calcul de la valeur de ce tenseur dans le cas d'un coût donné par un coût lagrangien de type Tonelli et plus particulièrement de la forme énergie cinétique plus énergie potentielle. Ces calculs ont également été réalisés de manière indépendante par Lee et McCann [76].

### 0.2.b Conséquence géométrique

Dans les chapitres 2 et 3, on s'intéresse à une conséquence géométrique inattendue du transport optimal. Pour point de départ on peut mentionner les théorèmes de régularité obtenus par Figalli, Rifford et Villani dans le cas d'une variété riemannienne pour le coût géodésique quadratique [48]. Les auteurs introduisent la définition de « Transport Continue Property »qui s'intéresse au cas où le transport optimal est continu dès lors que les mesures de départ et d'arrivée sont bien à densité continue. Le théorème le plus marquant est le suivant :

**Théorème 0.3.** *Soit $M$ une variété riemannienne lisse, connexe et compacte telle que :*

- **MTW** $> 0$ *pour le coût géodésique quadratique,*
- *tous les domaines d'injectivité de $M$ sont strictement convexes.*

*Alors $M$ vérifie* (**TCP**) *pour le coût géodésique quadratique, c'est à dire que pour tout couple de mesures $\mu$ et $\nu$ à densités par rapport au volume, continues et strictement positives, le transport optimal de $\mu$ sur $\nu$ pour le coût géodésique quadratique est continu.*

La question sous-jacente naturelle est de savoir si la positivité du tenseur de Ma-Trudinger-Wang entraîne la convexité des domaines d'injectivité. Une première réponse fut apporté par Loeper et Villani [83] dans le cas où le tenseur est strictement positif et la variété $M$ non focale. Dans le chapitre 3, ce résultat est étendu au cas où le tenseur est seulement positif. Dans le cas de la dimension 2 on supprime également la condition de non focalité de la variété. Les théorèmes nouveaux sont les suivants :

**Théorème 0.4.** *Soit $(M, g)$ une variété riemannienne non focale satisfaisant* **MTW***(0). Alors tous les domaines d'injectivité de $M$ sont convexes.*

**Théorème 0.5.** *Soit $(M, g)$ une variété riemannienne compacte de dimension* 2*, analytique, satisfaisant* **MTW***(0). Alors tous les domaines d'injectivité de $M$ sont convexes.*

Les techniques de démonstration du second théorème se placent en fait dans un cadre plus général et donnent l'espoir d'obtenir un résultat sans restriction ni de régularité, ni de dimension.

L'obtention de ce théorème a nécessité une réécriture et une légère extension du théorème suivant obtenu par Li et Nirenberg :

**Théorème 0.6** (Les domaines d'injectivité (*c*ut-loci) tangents sont Lipschitz (i))**.** *Soit $M$ une variété riemannienne compacte. Il existe $\kappa > 0$ tel que les domaines d'injectivité soient $\kappa$-Lipschitz.*

On a besoin dans notre cas du caractère Lipschitz également pour des perturbations du point sur la variété et pas seulement de la vitesse. Plus précisément on obtient dans le chapitre 2 le théorème suivant :

**Théorème 0.7** (Les domaines d'injectivité (*c*ut-loci) tangents sont Lipschitz (ii))**.**

1. *Il existe $\kappa > 0$ tel que $\{I(x)\,|x \in M\}$ est $\kappa$-Lipschitz.*

2. *Si $M$ est non focal alors il existe $\kappa > 0$ tel que $\{(x, I(x)) \subset TM \,|x \in M\}$ est $\kappa$-Lipschitz.*

3. *Si $M$ est de dimension 2, alors il existe $\kappa > 0$ tel que $\{(x, I(x)) \subset TM \,|x \in M\}$ est $\kappa$-Lipschitz.*

4. *Pour tout $(x, v) \in TM$, $s \in \mathbb{R}$ et $w \in T_{\exp_x(sv)}$, on a $|t_c(\exp_x(sv), w) - t_c(x, v)| \leq d^2\left((\exp_x(sv), w), (x, v)\right)$.*

Dans le théorème précédent, $t_c$ est une fonction représentant le bord du domaine d'injectivité $I$. Elle est définie sur $UM$ :

$$t_c(x, v) := \sup\Big\{t \geq 0 \,|\, tv \in \mathrm{I}(x)\Big\} \tag{0.2.3}$$

$$= \max\Big\{t \geq 0 \,|\, d^2(x, \exp_x(tv)) = |t|_x^2\Big\}. \tag{0.2.4}$$

La démonstration du théorème 0.7, développée au chapitre 2, se fait par analogie à celle proposée par Castelpietra et Rifford [30]. L'idée est de séparer le bord des domaines d'injectivité en trois parties dissociées, la première contenant les vitesses étant à la fois dans le lieu focal et dans le bord des domaines d'injectivité, la deuxième contenant les vitesses uniformément éloignées de la première partie et la troisième contenant l'ensemble des autres vitesses. On décrit le bord à l'aide du théorème des fonctions implicites et on en déduit son caractère Lipschitz.

Les théorèmes 0.4 et 0.5 sont démontrés au chapitre 3.

## 0.3 Le transport optimal pour une vision lagrangienne de Keller-Segel

La seconde partie de cette thèse est dédiée à l'étude du comportement d'une équation 1D possédant les mêmes propriétés que l'équation de Keller-Segel.

### 0.3.a Modèle classique de Keller-Segel

Le modèle de Keller-Segel consiste en l'étude d'une population de cellules ayant la particularité de s'attirer entre elles par l'émission d'un signal chimique. On suit donc l'évolution de deux quantités, la densité de cellules (notée $\rho$) d'une part et la concentration du chémoattractant (notée $c$) d'autre part. Dans sa forme simplifiée, le modèle de Keller-Segel (ou Patlak-Keller-Segel) s'écrit [65] :

$$\partial_t \rho\,(t, x) - \Delta \rho\,(t, x) + \chi \nabla \cdot (\rho\,(t, x)\, \nabla c\,(t, x)) = 0 \qquad (t, x) \in \mathbb{R}_+ \times \mathbb{R}^n \tag{0.3.1a}$$

$$- \Delta c\,(t, x) = \rho\,(t, x) \qquad (t, x) \in \mathbb{R}_+ \times \mathbb{R}^n \tag{0.3.1b}$$

Les paramètres du modèle sont le coefficient de sensibilité des cellules $\chi$ et la masse totale $M$ qui est conservée. La particularité de ce modèle est donc la compétition entre un terme de diffusion qui régularise et étale la solution et un terme de contraction qui, au contraire, crée des singularités.

On sait résoudre l'équation (0.3.1b) dans tout l'espace. En dimension 2, par exemple, le noyau du Laplacien donne une interaction logarithmique :

$$c(t,x) = -\frac{1}{2\pi}\int_{\mathbb{R}^2} \log|x-y|\rho(t,y)\ dy\,.$$

La particularité du modèle de Keller-Segel en dimension 2 est son principe de dichotomie bien comprise par Blanchet, Dolbeault et Perthame [12] :

**Théorème 0.8** (Blanchet, Dolbeault et Perthame). *On suppose que* $\rho_0\big(|\log\rho_0| + (1+|x|^2)\big) \in L^1(\mathbb{R}^2)$.
*Si* $\chi M < 8\pi$*, alors les solutions sont globales en temps, le terme de diffusion domine.*
*Si* $\chi M > 8\pi$*, les solutions explosent en temps fini, le terme d'interaction domine.*
*Dans le cas sous-critique* $\chi M < 8\pi$*, les solutions convergent vers un profil autosimilaire.*
*Dans le cas sur-critique* $\chi M > 8\pi$*, on voit apparaître la formation d'une masse de Dirac en temps fini.*

L'outil principal pour ce théorème est la dissipation de l'énergie (ou énergie libre) définie par :

$$\mathcal{F}[\rho(t)] = \int_{\mathbb{R}^2} \rho(y)\log\rho(y)\ dy + \frac{\chi}{4\pi}\iint_{\mathbb{R}^2\times\mathbb{R}^2} \log|x-y|\rho(x)\rho(y)\ dxdy\,.$$

De nombreux chercheurs ont participé à l'étude de ce modèle, on peut citer par exemple Nanjundiah, Jäger and Luckhaus, Nagai, Biler, Herrero and Velázquez, Gajewski and Zacharias, Horstmann, Senba and Suzuki.... On trouvera plus de détails sur ces contributions dans le chapitre d'introduction de la partie 2.

### 0.3.b Le modèle de Keller-Segel avec interaction logarithmique

En dimension 2, c'est l'interaction logarithmique qui donne la structure de dichotomie. Pour la conserver en toute dimension $d$, on définit le modèle de Keller-Segel avec interaction logarithmique par :

$$\partial_t\rho = \Delta\rho + \nabla\cdot(\rho\nabla W * \rho) \quad \text{avec} \quad W(z) = \frac{\chi}{d\pi}\log|z|\,, \tag{0.3.2}$$

ainsi que l'énergie associée :

$$\mathcal{F}[\rho(t)] = \int_{\mathbb{R}^d} \rho(y)\log\rho(y)\ dy + \frac{\chi}{4\pi}\iint_{\mathbb{R}^d\times\mathbb{R}^2} \log|x-y|\rho(x)\rho(y)\ dxdy\,. \tag{0.3.3}$$

Calvez, Perthame et Sharifi Tabar ont montré que ce modèle suit bien le principe de dichotomie et de dissipation de l'énergie [25] avec cette fois comme paramètre critique $\chi M = 4d\pi$. Ce paramètre peut être deviné soit par homogénéité sur l'énergie, soit en calculant formellement le second moment :

$$\frac{d}{dt}\int |x|^2\rho\,(t,x)\,dx = M\left(2d - \frac{M\chi}{2\pi}\right). \tag{0.3.4}$$

Une autre propriété du modèle de Keller-Segel avec interaction logarithmique est qu'il s'interprète comme un flot gradient pour l'énergie $\mathcal{F}$ dans l'espace de Wasserstein, comme montré par Blanchet Calvez et Carrillo [8]. L'espace de Wasserstein contient l'ensemble des mesures absolument continues avec second moment fini :

$$W_2^{ac}(\mathbb{R}^n) = \left\{ \mu \in P_{ac}(\mathbb{R}^n) \text{ tel que } \int |x|^2 \mu(x) dx < +\infty \right\}.$$

On munit cet espace de la métrique de Wasserstein $d_W(\mu, \nu)$, donnée par l'application de transport de Brenier pour le coût quadratique $c(x, y) = |y - x|^2$ :

$$\begin{aligned} d_W^2(\mu, \nu) &= \inf_{\mathbf{t}} \left\{ \int |\mathbf{t}(x) - x|^2 d\mu(x) ; \quad \mathbf{t}_\# \mu = \nu \right\} \\ &= \int |T(x) - x|^2 d\mu(x). \end{aligned}$$

La notion de flot gradient associée à l'espace de Wasserstein fait l'objet du chapitre 5, tandis que le chapitre 6 est consacré à l'étude de l'énergie $\mathcal{F}$ dans le cas sous-critique. Le transport optimal envoyant une solution à un temps $t_0$ sur celle à un temps $t$ peut être interprété comme l'ensemble des caractéristiques de l'équation (0.3.2). Il mène donc à une interprétation lagrangienne de cette équation, c'est-à-dire en suivant les particules.

#### Problème de quantification de la masse

Une autre question intéressante concernant l'équation de Keller-Segel avec interaction logarithmique est la quantification de l'intensité de la masse de Dirac créée dans le cas sur-critique. C'est le problème de quantification de la masse [104].

**Problème 0.3.1** (Quantification de la masse)**.** *Soit $\chi$ fixé. Dans le régime sur-critique $M > \frac{4d\pi}{\chi}$, le premier point singulier contient exactement la masse critique : $\frac{4d\pi}{\chi}$.*

Au chapitre 7 de cette thèse, nous contribuons à la compréhension du problème de quantification de la masse à travers un exemple particulaire en dimension 1, défini en suivant l'intuition lagrangienne du flot gradient dans l'espace de Wasserstein.

### 0.3.c Un schéma particulaire en dimension 1

L'interprétation flot gradient est particulièrement intéressante en dimension 1, car on a alors une formule explicite permettant de ramener le système d'équation (0.3.1) à un vrai flot gradient dans l'espace $L^2(0,1)$. Cette transformation est donnée par la pseudo-inverse de la fonction de répartition de la solution au problème (0.3.1). Explicitement, pour $\mu \in W_2^{ac}(\mathbb{R})$, on définit sa masse cumulée ou fonction de répartition $M_\mu : \mathbb{R} \to [0,1]$ par :

$$M_\mu(x) = \mu(]-\infty, x[). \tag{0.3.5}$$

Sa pseudo-inverse $X_\mu : [0,1] \to \mathbb{R}$ est alors donnée par :

$$X_\mu(m) = \inf\{x \in \mathbb{R} \text{ tel que } M_\mu(x) > m\}. \tag{0.3.6}$$

De là on peut définir un schéma particulaire dont l'idée directrice est de faire porter une masse de même valeur à chaque particule. On discrétise donc l'énergie $\mathcal{F}$ en N points avec un pas de masse constant. On obtient une équation flot gradient dans $\mathbb{R}^N$ qui sera la base de notre étude. L'énergie est donnée par :

$$\mathbb{E}\left(X\right) = -\sum_{i=1}^{N-1} \log\left(X_{i+1} - X_i\right) + \chi h_N \sum_{1 \leq i \neq j \leq N} \log|X_i - X_j|, \tag{0.3.7}$$

et le flot gradient associé s'écrit :

$$\begin{cases} \dot{X}\left(t\right) = -\nabla \mathbb{E}\left(X\left(t\right)\right) & t \in \mathbb{R} \\ X\left(0\right) = X_0 & X^0 \in \mathbb{R}^N. \end{cases} \tag{0.3.8}$$

Ce schéma a été étudié en particulier par Devys.Il présente toujours une compétition entre attraction et diffusion, mais cette fois le problème de masse critique est déplacé en un problème de nombre de particules critique. En effet, on observe la collision de plusieurs particules dans le cas sur-critique. La question est alors de savoir si la collision se produit avec le nombre de particules minimum charriant la masse critique. Ce nombre est appelé le paramètre $k$ critique.

**Définition 0.3.2** (Paramètre $k$ critique)**.**

$$\chi_N^k = \frac{N+1}{k}.$$

Le dernier chapitre de cette thèse est donc consacré au problème discret de la quantification de la masse :

**Problème 0.3.3** (Quantification de la masse, cas discret)**.** *Soit $\chi$ fixé, si $\chi_N^k < \chi < \chi_N^{k-1}$ alors le premier point singulier contient, génériquement, exactement $k$ particules.*

Dans le cas discret, il est impossible d'envisager un résultat autre que générique car le système ne peut pas briser les symétries (voir figures 0.3.c et 0.3.c). Après une analyse détaillée du cas à trois particules, on s'intéresse au cas à $N$ particules. Les résultats obtenus sont de deux natures. D'une part on identifie des solutions dont on peut garantir l'explosion avec seulement le nombre de particules critiques. D'autre part on détaille le profil d'explosion de ces solutions. Les principaux outils utilisés sont les déviations standards à $k$ particules et le potentiel extérieur d'interaction.

**Définition 0.3.4** (Déviation standard et potentiel extérieur d'interaction)**.** *On se donne $\mathcal{I}$ un ensemble connexe d'indices, appelé ensemble intérieur. Typiquement, $\mathcal{I} = [l, l+p]$. On définit également $\mathcal{O} = [1, N] \backslash \mathcal{I}$ l'ensemble extérieur. La déviation standard du vecteur $(X_\mathcal{I}) = \{X_l, ...X_{l+p}\}$ est donnée par :*

$$\Pi_\mathcal{I}^2 = \sum_{i \in \mathcal{I}} \left(X_i - \overline{X}_\mathcal{I}\right)^2, \quad \textit{où} \quad \overline{X}_\mathcal{I} = \frac{1}{|\mathcal{I}|} \sum_{i \in \mathcal{I}} X_i\,. \tag{0.3.9}$$

*Le potentiel extérieur d'interaction $H_{\mathcal{I}\mathcal{O},2}$ contrôle la force avec laquelle les particules extérieures*

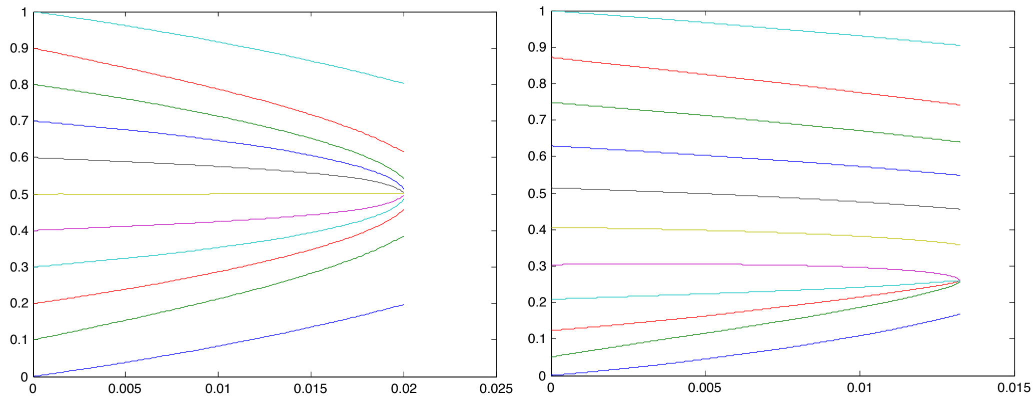



FIGURE 0.3.1 – Ici 4 particules suffisent pour exploser. Dans le premier cas, la condition initiale symétrique empêche le modèle de faire un choix parmi les cinq particules centrales, on a donc une explosion avec cinq particules. Ce phénomène est instable mais le schéma numérique utilisé conserve les symétries et permet donc de l'observer. Dans le second cas, la condition initiale n'est pas symétrique et c'est bien quatre particules seulement qui participent à l'explosion.

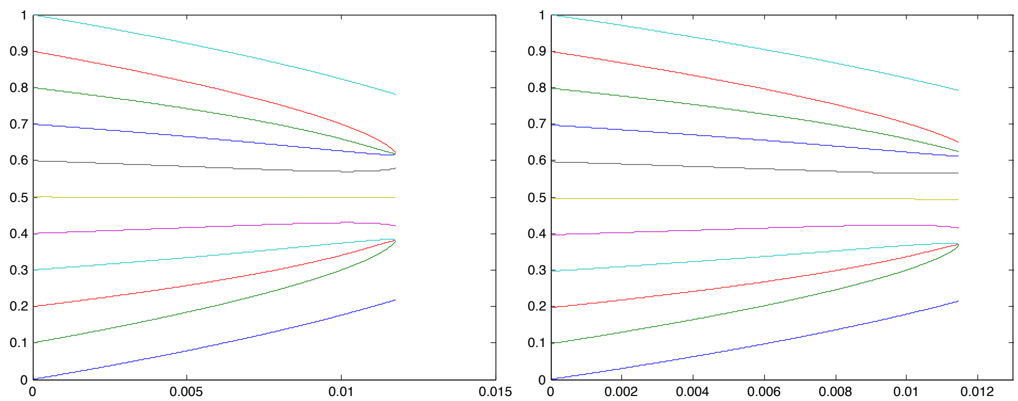



FIGURE 0.3.2 – Ici 3 particules suffisent pour exploser. Dans le premier cas, la condition initiale symétrique donne deux explosions simultanées et symétriques. Dans le second cas, la condition initiale n'est pas symétrique et l'une des deux explosions se produit juste avant l'autre.

*peuvent influer sur l'ensemble intérieur :*

$$H_{\mathcal{IO},2} = \sum_{j\in\mathcal{O}} \sum_{i\in\mathcal{I}} \frac{1}{(X_j - X_i)^2} .$$

*De façon générale, on utilisera également $H_{\mathcal{IO},m}$, qui correspond à la norme $m$ :*

$$H_{\mathcal{IO},m} = \sum_{j\in\mathcal{O}} \sum_{i\in\mathcal{I}} \frac{1}{(X_j - X_i)^m} .$$

### Stabilité des bassins d'attraction

On commence par exhiber des bassins d'attraction pour l'explosion. Lorsqu'une solution croise l'un de ces bassins, son comportement est contrôlé jusqu'au temps d'explosion. En particulier, on montre le théorème suivant :

**Théorème 0.9.** *Soit $\chi$ fixé tel que $\chi_N^k < \chi < \chi_N^{k-1}$ et $X$ une solution de l'équation* (0.3.8)*. S'il existe $t \in [0, T]$ tel que $X(t) \in D_{N,\chi}^{\varepsilon, \frac{C_N}{\varepsilon}}$, alors l'ensemble d'explosion contient seulement $k$ points.*

L'espace $D_{N,\chi}^{\varepsilon, \frac{C_N}{\varepsilon}}$ contrôle les tailles respectives de la deviation standard à $k$ particules et du potentiel d'interaction extérieur. Sa définition est la suivante :

$$D_{N,\chi}^{\varepsilon, \frac{c}{\varepsilon}} = \left\{ X \in \mathbb{R}^N \text{ tel que } \exists l > 0 \text{ avec } \Pi_{\mathcal{I}}^2 \leq \varepsilon \text{ et } H_{\mathcal{IO},2} < \frac{c}{\varepsilon} \right\}.$$

Ici $\mathcal{I} = [l, l+k-1]$, $\mathcal{O} = [1, N] \setminus [l, l+k-1]$ et $C_N$ dépend uniquement de $\chi$, $k$ et $N$.

### Rigidité de l'explosion

Lorsque la solution explose, la question naturelle qui se pose est celle du profil d'explosion. Pour répondre à ce problème, on rééchelonne la solution en ajoutant un potentiel à l'énergie et en redéfinissant le temps. On obtient ainsi une solution définie sur $\mathbb{R}^+$ qui fait apparaître le profil d'explosion. Le théorème principal obtenu au chapitre 7 peut être résumé ainsi :

**Théorème 0.10.** *On considère l'énergie rééchelonnée associée à un sous-ensemble de $k$ particules :*

$$\begin{aligned} \mathbb{E}_k(Y) = &- \sum_{i\in\mathcal{I}\setminus\{l+k-1\}} \log(Y_{i+1} - Y_i) \\ &+ \chi h_N \sum_{(i,j)\in\mathcal{I}\times\mathcal{I}\setminus\{i\}} \log|Y_i - Y_j| - \frac{\alpha}{2} \sum_{i\in\mathcal{I}} |Y_i|^2 . \end{aligned}$$

*Soit $\chi$ fixé tel que $\chi_N^k \leq \chi \leq \chi_N^{k-1}$ et $Y$ le rééchelonné d'une solution de* (0.3.8)*. On a alors :*

1. *$\dot{Y}(t) \to 0$ quand $t \to \infty$.*
2. *$\mathbb{E}_k(Y(t))$ converge vers une limite notée $E_\infty$, quand $t \to \infty$.*
3. *$(\nabla\mathbb{E}_k)(Y(t)) \to 0$ quand $t \to \infty$.*

Dit autrement, la solution converge vers un point critique de la fonctionnelle à $k$ particules à une sous-suite près. Nos $k$ particules évoluent donc comme si elles étaient solutions de l'équation flot gradient associée à l'énergie à $k$ particules. Ainsi, les profils d'explosions sont contraints par les points critiques de la fonctionnelle à $k$ particules (rigidité de l'explosion).

## 0.4 Perspectives

Les résultats obtenus au cours de ce travail de thèse ouvrent la voie à de nombreuses perspectives.

A la suite du chapitre 3, nous espérons pouvoir montrer que les domaines d'injectivité de $(M, g)$ sont convexes sans restriction ni de dimension, ni de focalité, ni de régularité sur les variétés plus que $C^4$. Pour cela, il reste toutefois des difficultés importantes à surmonter. Une première étape pour mener à bien ce travail serait une étude plus poussé du tenseur de Ma-Trudinger-Wang en toutes dimensions, à l'image des travaux de Figalli, Rifford et Villani en dimension 2 [47]. Enfin, une autre interrogation reste le comportement du transport optimal lors de la présence d'un point purement focal dans le lieu d'injectivité. Une meilleure compréhension du tenseur pourrait également permettre des avancées sur ce sujet.

En ce qui concerne la seconde partie, un prolongement direct des résultats obtenus serait de parvenir à quantifier la taille des bassins d'attraction et de montrer qu'ils permettent d'attraper toutes les solutions. Un autre développement naturel serait de s'intéresser au passage à la limite quand le nombre de particules tend vers l'infini. On pourrait également chercher à transmettre les résultats obtenus ici au cas continu en dimension 1, la difficulté venant des singularités créées aux interfaces des ensembles $\mathcal{I}$ et $\mathcal{O}$. Enfin, en suivant le travail de Devys [37], on pourrait définir un modèle particulaire qui dépasse l'explosion et s'intéresser à ce que devient la ensuite solution. La masse de Dirac grossit-elle ? Quelle peut être l'intensité des nouvelles masses de Dirac créées ? Les résultats numériques de Devys donnent déjà des indications intéressantes pour répondre à ces questions.

# Part I

# Ma-Trudinger-Wang tensor and regularity of the tangent cut loci

**Abstract**

This part is composed of three chapters. The first one can be seen as a general introduction of the Ma-Trudinger-Wang, or cross curvature, tensor and its various link with the regularity of optimal transport. We also find here the computation of the Ma-Trudinger-Wang tensor for some particular Lagrangian action. The second chapter is devoted to the proof of some regularity results on the tangent focal loci and the tangent cut loci of a smooth compact Riemannian manifold. In particular, we prove that the tangent cut loci are Lipschitz continuous. This theorem was first proved by Li-Nirenberg improving one due to Itoh-Tanaka, ([64, 77]). Then Castelpietra and Rifford used another approach to simplify the demonstration [30]. Here we follow the strategy of the latest. The idea is to split any tangent cut locus into three different categories. For each of them we obtain the Lipschitz continuity as a consequence of the implicit function theorem. The new result here is the Lipschitz continuity with respect to the position variable in some directions.

In the third chapter, we use the previous result to show that given a smooth, non-focal or two dimensional, compact Riemannian manifold, if the Ma-Trudinger-Wang (MTW) tensor is non-negative then all the tangent injectivity domains are convex. These new results extend and contain a previous one due to Loeper and Villani [83]: in the non-focal case, if the Ma-Trudinger-Wang tensor is positive then all the tangent injectivity domains are uniformly convex. The key tools are the extended MTW tensor introduced by Figalli and Rifford and the mother computation introduced by Kim and McCann [74] . This tensor is used in a bootstrap argument: for any line with endpoints in a tangent cut locus, we prove that the cut locus cannot be too far from this line, then we use the tensor to improve this default until we get convexity. This result needs a global argument at each step; therefore we obtain the convexity for all the tangent cut locus together. This work is a collaboration with Alessio Figalli and Ludovic Rifford and is a contribution to the understanding of the link between the regularity of optimal transport and the geometry on a Riemannian manifold. It enforces the idea that the Ma-Trudinger-Wang tensor contains all the ingredients, including the geometrical one, for the regularity of optimal transport.

# Chapter 1

# Regularity of optimal transport: the Ma-Trudinger-Wang tensor

## 1.1 Definitions and notation

### 1.1.a Monge problem

#### Definition

The optimal transport was introduced by Monge in 1781 [90], the idea is to transport some mass from one place to another with minimal cost. A way to set the problem is the following.

**Definition 1.1.1** (Monge Problem). *Let $M$ be a Riemannian manifold, $c : M \times M \to \mathbb{R}$ a cost function and $\mu, \nu$ two absolutely continuous probability measures. A transport $\mathbf{t}$, from $\mu$ to $\nu$ is a mapping from $M$ to $M$, $\mu$ measurable, such that $\mathbf{t}_{\#}\mu = \nu$. It means that for any measurable set $B$ in the target space we have $\nu(B) = \mu\left(\mathbf{t}^{-1}(B)\right)$. An optimal transport map $T$ is a $\mu$ measurable map of the following minimization problem:*

$$T = \underset{\mathbf{t}_{\#}\mu=\nu}{\operatorname{argmin}} \left( \int_M c\left(x, \mathbf{t}(x)\right) d\mu(x) \right). \tag{1.1.1}$$

The basic definitions and notation concerning a Riemannian manifold can be found in the appendix A (see also [53, 98]).

#### Cost assumption and regularity

The transport optimal starts with the notion of cost: how expensive is it to go from one point to another. Thus for the optimal transport to be well defined, we need assumptions on the cost function. They are enclosed in the following definition. For details we refer to [110].

*Assumptions.* . Let $c : M \times M \to \mathbb{R}$ be a cost function. The assumptions **(Super)**, **(Twist)**, **(Lip)** and **(SC)** will be used in the sequel:

1. **(Super)** if $c(x, y)$ is everywhere superdifferentiable as a function of $x$, for all y.

2. **(Twist)** if on its domain of definition, $\nabla_x c(x,\cdot)$ is injective. In this case we denote $-\left(\nabla_x c(x,\cdot)\right)^{-1}$ by $c\text{-}\exp_x$.

3. **(Lip)** $c(x,y)$ is locally Lipschitz as a function of $x$, uniformly in y.

4. **(SC)** $c(x,y)$ is locally semiconcave as a function of $x$, uniformly in y.

An important class of cost examples are those coming from a Lagrangian action:

**Definition 1.1.2** (Lagrangian action)**.** *Let $L : R \times TM \to R$ be a Tonelli Lagrangian. The cost $c$ associated to $L$ is defined by:*

$$c(x,y) = \min_{\gamma \in A_x^y} \left\{ \int_0^1 L\left(\gamma(t), \frac{d}{dt}\gamma(t), t\right) dt \right\}, \tag{1.1.2}$$

*where $A_x^y$ is the set of all absolutely continuous paths $\gamma$ defined on $[0,1]$ satisfying $\gamma(0) = 0$ and $\gamma(1) = 1$.*

The Tonelli assumption (convex in the fibre and superlinear) guarantees that $c$ satisfies the conditions 1.1.a. The most basic example is given for $L(x,v,t) = \frac{1}{2}|v|_x^2$; it leads to the quadratic geodesic cost $c(x,y) = \frac{1}{2}d^2(x,y)$, where $d$ is the geodesic distance. The notation in the twist condition are defined by analogy with this case. More informations on Tonelli Lagrangian action can be found in the recent book of Mazzuccheli [86].

### $c$-convexity

Another crucial definition is the concept of $c$-convexity. On the standard Euclidean space, $\mathbb{R}^n$, the basic tool needed to define a notion of derivation is the line. A function is differentiable in a point when it coincides at order 0 and 1 with a line. This line also gives a natural class of functions, the convex one. A function $\psi$ is convex if for any point $x_0$, one can touch $\psi$ at $x_0$, from below, with a line. At a point $x_0$, the contact line will take the form $\psi(x_0) + (x - x_0)\cdot p$. In this case we say that $p$ is an element of the subgradient of $\psi$ at $x_0$:

$$p \in \nabla^- \psi(x_0).$$

The set $\{x \mapsto (x - x_0)\cdot p$ , for $p \in \mathbb{R}^n$, $x_0 \in \mathbb{R}^n\}$ constitutes the support set where we can pick up a line to touch our function from below. A function $x \mapsto (x - x_0)\cdot p$ is a support function if for all $x \in \mathbb{R}^n$:

$$\psi(x) \geq \psi(x_0) + (x - x_0)\cdot p.$$

Roughly speaking, a function is convex if there exists a support line for any $x_0 \in \mathbb{R}^n$. If now we work in a different world and the shorter path to go from one point to another is no longer the line but a curved path, then we have other supports. The support set at $x_0$ becomes the set of all the $c$-support functions $x \mapsto D(x_0,y)(x) = c(x_0,y) - c(x,y)$ for $y \in Y$. We can thus define the $c$-convexity by analogy with convexity.

**Definition 1.1.3.** *Let $c : X \times Y \to \mathbb{R} \cup \infty$, a function $\psi : X \to \mathbb{R} \cup \infty$ is c-convex if $\forall x_0 \in X$, there exists $y_0 \in Y$ such that*

$$\psi(x) \geq \psi(x_0) + c(x_0, y_0) - c(x, y) = \psi(x_0) + D(x_0, y_0), \quad \forall x \in X.$$

*In this case we define the c-subdifferential $(\partial_c \psi(x_0))$ of $\psi$ at $x_0$ as the set of all the admissible $y_0$ and the c-subgradient $\left(\nabla_c^- \psi(x_0)\right)$ of $\psi$ at $x_0$ as the set of all the speed $p$ such that $c\text{-}\exp_{x_0}(p) \in \partial_c \psi(x_0)$.*

An equivalent way to define it is with the $c$-transform:

**Definition 1.1.4.** *Let $c : X \times Y \to \mathbb{R} \cup \infty$, and $\psi : X \to \mathbb{R} \cup \infty$. We define $\psi^c : Y \longrightarrow \mathbb{R} \cup \infty$ as*

$$\psi^c(y) = \inf_{x \in X} \left\{ \psi(x) + c(x, y) \right\}.$$

*Similarly for $\phi : Y \longrightarrow \mathbb{R} \cup \infty$ we define $\phi_c : X \longrightarrow \mathbb{R} \cup \infty$ as*

$$\phi_c(x) = \sup_{y \in Y} \left\{ \phi(y) - c(x, y) \right\}.$$

*Then $\psi$ is c-convex if $\psi = (\psi^c)_c$.*

Moreover if $c$ satisfies **(Super)** then $\psi$ is subdifferentiable at any $x_0 \in X$ such that $\partial_c \psi(x_0)$ is not empty. In particular for any $y \in \partial_c \psi(x_0)$ there exists $p \in T^* X$ such that

$$\nabla_x c(x_0, y) = p,$$

then by definition $-p \in \nabla^- \psi(x_0)$. One can remark that $\nabla_c^- \psi(x_0) \subset \nabla^- \psi(x_0)$ but the converse is not true in general.
If in addition the condition **(Twist)** and **(Lip)** are satisfied, then $\psi$ is locally Lipschitz and differentiable almost everywhere. If the condition **(Twist)** and **(SC)** are satisfied, then $\psi$ is locally semiconcave and differentiable almost everywhere in the interior of its domain. In any case when $\psi$ is differentiable at $x_0$ we have by definition $c\text{-}\exp_{x_0}(\nabla \psi(x_0)) = \partial_c \psi(x_0)$. In this case the notation $D(x_0, p)$ stands for $D\left(x_0, c\text{-}\exp_{x_0}(p)\right)$.

### 1.1.b Notation

We list here all the notation and remarks that are used in this chapter. The Riemannian manifold $M$ and the cost $c$ satisfying 1.1.a are fixed, $x, y$ are points in $M$, $p \in T_x M$ and $\psi$ is a $c$-convex function.

- $c\text{-}\exp_x(p) = (\nabla_x c(x, \cdot))^{-1}(-p)$.
- $c$-I$(x)$ is the domain of definition of $c\text{-}\exp_x$ .
- $c$-cut$(x)$ is the domain where $\nabla_x c(x, \cdot)$ is not well defined.
- $c$-NF$(x)$ is the domain where the differential of $q \mapsto c\text{-}\exp_x(q)$ is invertible.

- $\partial_c\psi(x)$ is the $c$-differential of $\psi$ at $x$.
- $\nabla_c^-\psi(x)$ is the $c$-subgradient of $\psi$ at $x$. We have $\partial_c\psi(x) = c\text{-}\exp_x\left(\nabla_c^-\psi(x)\right)$.
- $\nabla^-\psi(x)$ is the subgradient of $\psi$ at $x$. We have $\nabla_c^-\psi(x) \subset \nabla^-\psi(x)$. The converse in not true.
- The $c$-support $D(x,p) = D\left(x_0, c\text{-}\exp_x(p)\right) = D\left(x,y\right) = c(x,y) - c(\cdot, v)$.
- We define a $c$-segment with respect to $x$ by $c\text{-}[p,q]_x = c\text{-}\exp_x([p,q])$.
- A set $S$ in $M$ is $c$-convex with respect to $x$ if for any $y_0, y_1 \in S$ the $c$-segment $c\text{-}\left[\nabla_x c(x,y_0), \nabla_x c(x,y_1)\right]_x$ is in $S$. For example if $\nabla_c^-\psi(x) = \nabla^-\psi(x)$ then $\partial_c\psi(x)$ is $c$-convex with respect to $x$.
- A set $S$ in $M$ is $c$-convex with respect to another set $S'$ if for any $z \in S'$, $S$ is $c$-convex with respect to $z$.
- When $c$ is the quadratic geodesic cost we find the usual notion: $c\text{-}\exp_x = \exp_x$, $c\text{-cut}(x) = \text{cut}(x)$ and $c\text{-I}(x) = \text{I}(x)$ is the injectivity domain.
- When $c$ is given by a smooth, time independent, Tonelli Lagrangian $L$, $c\text{-}\exp_x(p)$ is solution at time 1 of the path of least action for $L$ starting at $x$ with initial velocity $p$. Moreover $\nabla_v L(x,v) = -\nabla_x c(x,y)$ .
- For convenience the manifold M is supposed to be compact. Many things can be done without this hypothesis [110] and when it comes to the regularity question for example for the quadratic geodesic cost this hypothesis is not restrictive.
- M can be a bounded domain of $\mathbb{R}^n$ or a Riemannian manifold. In this case we ask that the boundary has a Hausdorff measure at most equal to $n-1$.
- By convention the points $x$ and $y$ always lie in the interior of $M$.
- Without adding any difficulty we can suppose that the source space $M$ and the target space $M$ are in fact different.

The notion of $c$-segment, $c$-subdifferential was introduced by Ma Trudinger and Wang [85].

## 1.2 Solution of the Monge problem

### 1.2.a Existence theorem

The first solution to the Monge problem was given by Brenier [14] in the case of the quadratic cost in $\mathbb{R}^n$.

**Theorem 1.2.1** (Brenier). *Let $c : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $x, y \mapsto c(x,y) = \dfrac{|x-y|^2}{2}$ be the cost function, let $\mu \in \mathbb{P}(\mathbb{R}^n)$ absolutely continuous and $\nu \in \mathbb{P}(\mathbb{R}^n)$ then there exists a unique transport map $T$*

*solving the Monge problem. Moreover there exists a convex potential function $\varphi$, (or a $\frac{x^2}{2}$−convex function $\phi$) such that*

$$T(x) = \nabla\varphi(x) = x + \left(\nabla\varphi(x) - \frac{x^2}{2}\right) = \nabla\left(\frac{x^2}{2} + \phi(x)\right).$$

*The function $\varphi$ is unique up to a constant.*

The most general version for the solution to the Monge problem may be found in [110]. Here we consider only the particular case where $M$ is compact.

**Theorem 1.2.2** (Solution of the Monge problem)**.** *Let $M$ be a compact Riemmanian manifold, $c : M \times M \to \mathbb{R}$ be a continuous cost function satisfying* **(Super)**, **(Twist)** *and (***(Lip)** *or* **(Sc)***). Let $\mu \in \mathbb{P}(M)$ absolutely continuous and $\nu \in \mathbb{P}(M)$ then there exists a unique optimal map $T$ solving the Monge problem. Moreover there exists a $c$-convex function $\psi$ such that $\mu$ almost everywhere:*

$$\nabla\psi(x) + \nabla_x c\,(x, T(x)) = 0 \quad \textit{or equivalently} \quad T(x) = c\text{-}\exp_x\left(\nabla\psi\,(x)\right). \tag{1.2.1}$$

*In particular when the cost is the quadratic geodesic one: $c = \frac{1}{2}d^2$ the $c\text{-}\exp$ is exactly the exponential function on $M$, thus*

$$T(x) = \exp_x\left(\nabla\psi\,(x)\right).$$

The particular case of quadratic geodesic cost is due to McCann [87]. The characterization (1.2.1) states that $\psi$ is $c$-convex, differentiable and $T(x) \in \partial_c\psi(x)$, which means that $T$ is concentrated on the $c$-subdifferential of $\psi$.

### 1.2.b The Monge-Ampère equation

To avoid a long discussion we consider here a cost $c$ which is at least $C^4$ on $M \times M$, and we transport a measure with smooth density with respect to the volume. To fix the notation, we take $\mu = f dx$ and $\nu = g dy$. We also suppose that $T$ is $C^2$.
Therefore we see that $T$ gives a change of variable formula. By definition $T_{\#}\mu = \nu$. Thus for any measurable set $B$ in the target space we have

$$\int_{T^{-1}B} f(x)dx = \int_B g(y)dy.$$

We perform the change of variable $y = T(x)$, that leads to

$$\int_{T^{-1}B} f(x)dx = \int_{T^{-1}B} g(T(x))\left|\det\left[\nabla T(x)\right]\right| dy.$$

We deduce

$$\left|\det\left[\nabla T(x)\right]\right| = \frac{f(x)}{g\left(T(x)\right)}. \tag{1.2.2}$$

The characterization (1.2.1) in Theorem (1.2.2) leads to a fully non linear elliptic equation on $\psi$. Indeed we can differentiate, in smooth charts, (1.2.1) with respect to $x$ and get

$$\nabla^2\psi(x) + \nabla_{xx}c(x, T(x)) = -\nabla_{x,y}c\left(x, T(x)\right).\nabla T(x),$$

taking the absolute value of the determinant and recalling (1.2.2) we obtain a partial differential equation on $\psi$

$$\det\left(\nabla^2\psi(x) + \nabla_{xx}c\left(x, T(x)\right)\right) = \frac{f(x)}{g(T(x))}\left|\det \nabla_{x,y}c\left(x, T(x)\right)\right|. \tag{1.2.3}$$

According to (1.2.1) this equation is a Monge-Ampère type equation on $\psi$. It is a fully non linear elliptic equation. For this equation the boundary condition is given by

$$T\left(\mathrm{supp}(\mu)\right) = \mathrm{supp}\left(\nu\right).$$

This condition is called the second boundary condition. In the Euclidean case, with the cost given by the square of the norm we obtain

$$\det\left(\nabla^2\psi(x)\right) = \frac{f(x)}{g(T(x))}. \tag{1.2.4}$$

The Monge-Ampère equation is useful to answer to the question of the regularity of the potential function $\psi$. The typical question is: given two smooth densities (say $C^k$) $f$ and $g$, is the optimal transport smooth ? In the Euclidean case, several results were already known when the link was done with optimal transport. Therefore the first regularity results go to Delanoë, Caffarelli and Urbas [33, 21, 18, 19, 20, 22, 107, 106]. In the general case the question is quite difficult. We start by giving some obstruction to the regularity of the optimal transport and then we explain the breakthrough due to Ma, Trudinger and Wang [85].

## 1.3 Regularity of optimal transport map: obstructions

In this section we consider a Riemannian manifold M of dimension $n$.

### Caffarelli's counterexample, convexity of target's support

Caffarelli found a first obstruction to the optimal transport: the support of $\nu$ has to be convex. Indeed let $f$ be the normalized indicator function of the unit ball in $\mathbb{R}^2$, and $g_\varepsilon$ the normalized indicator function of two half-unit balls linked with a small shuttle. The transport $f$ to $g_\varepsilon$ for the quadratic cost ($|x-y|^2$) gives a map $T_\varepsilon$. Ninety-nine percent of the upper East side of the unit ball has to be sent into the upper east hemisphere of the target set, and the same goes for the upper West side. Moreover by continuity, the image by $T$ of a horizontal segment joining the upper East side to the upper West side must be connected. Thus we can take a $x$ into the upper East side such that $x - T_\varepsilon(x)$ is mainly oriented to the south ($T_\varepsilon(x)$ is in the shuttle). According to the ninety-nine percent properties, and passing to the limit when $\varepsilon$ goes to zero, we can find a $y$ below $x$ such that $\langle x-y, T(x)-T(y)\rangle < 0$, which is impossible since $T$ is a gradient of a convex function. We deduce that $T_\varepsilon$ cannot be continuous for $\varepsilon$ small.

#### Ma-Trudinger-Wang's counterexample, c-convexity of target's support

As Ma, Trudinger and Wang showed in [85] a similar obstruction appears when the cost is not the quadratic one ($|x-y|^2$). They proved that the the target support has to be $c$-convex. The counter example constructed in this paper is the inspiration of the one given in theorem 1.5.5.

#### Loeper's counterexample

A second obstruction was explained by Loeper. He proved that the geometry of the manifold can also be an issue for the regularity of the optimal transport map. For example if we consider the quadratic geodesic cost then the manifold has to have non-negative sectional curvature. The idea is that in a negative sectional curvature world the Pythagoras inequality is in the wrong side, the hypothenus is not the quickest path to go from one summit to the other in a rectangle-triangle. Thus if there exists one point with negative sectional curvature, Loeper managed to fix this point by symmetry. Then choosing carefully $\mu$ and $\nu$ he makes it cheaper to move this point and get a contradiction.

#### The $\partial_c\psi(x)$ have to be connected

Another important fact to obtain the regularity of the optimal transport is that all the $\partial_c\psi(x)$ have to be connected. For clarity, the proof of this result is postponed at the end of the next section, it is the object of theorem 1.5.5.
For more details and pictures on these three examples we refer to [110].

## 1.4 The Ma-Trudinger-Wang tensor

### 1.4.a Definition

A positive result on optimal transport for various costs was first obtained by Ma, Trudinger and Wang. They came up in 2002 with a tensor that seems to be the good tool to solve the regularity issue [85, 105].

**Definition 1.4.1.** *Let $c : M \times M \to R$ satisfying 1.1.a. For any $(x, y) \in M \times M \setminus c\text{-cut}(M)$, and $(\xi, \eta) \in T_xM \times T_yM$ such that*

$$\langle \nabla^2_{(x,y)} c(x,y).\xi, \eta\rangle_y = \langle \xi, \nabla^2_{(x,y)} c(x,y).\eta\rangle_y = 0\,, \tag{1.4.1}$$

*we define $\mathfrak{S}_c$ such that*

$$\mathfrak{S}_c(x,y)(\xi,\eta) = \frac{3}{2}\Sigma_{ijklrs}(c_{ij,r}c^{r,s}c_{s,kl} - c_{ij,kl})\xi^i\xi^j\eta^k\eta^l \tag{1.4.2}$$

This expression is a bit complicated, and the regularity will be based on its positivity. Fortunately we can reformulate the definition.

**Proposition 1.4.2.** *We take the same hypothesis as in definition 1.4.1, we also define* $\overline{\eta} = -\nabla^2_{(x,y)}c(x,y).\eta$ *and* $v = -\nabla_x c\,(x,y)$ *then* (1.4.2) *is equivalent to*

$$\mathfrak{S}_c(x,y)(\xi,\eta) = -\frac{3}{2}\,\frac{d^2}{dt^2}\frac{d^2}{ds^2}\bigg|_{t=0}\bigg|_{s=0}\left[c\left(\exp_x(t\xi), c\text{-}\exp_x\left(v+s\overline{\eta}\right)\right)\right]. \tag{1.4.3}$$

Moreover we can remark in the following proposition that the MTW tensor is symmetric.

**Proposition 1.4.3.**
*We take the same hypothesis as 1.4.1, we also define* $\overline{\eta} = -\nabla^2_{(x,y)}c(x,y).\eta$, $y = c\text{-}\exp_x(v)$ *and* $\widetilde{c}(x,y) = c(y,x)$, $\widetilde{\xi} = -M(x,y)\xi$, $\widetilde{\eta} = (-\overline{\eta})^t M^{-1}(x,y)$. *Then* (1.4.2) *is equivalent to*

$$\mathfrak{S}_{\widetilde{c}}(x,y)(\widetilde{\xi},\widetilde{\eta}) = -\frac{3}{2}\,\frac{d^2}{ds^2}\frac{d^2}{dt^2}\bigg|_{s=0}\bigg|_{t=0}(\widetilde{c}(\widetilde{c}-\exp_x(\widetilde{v}+t\widetilde{\xi}),\exp_y(s\widetilde{\eta}))). \tag{1.4.4}$$

$$= \frac{d^2}{dt^2}\frac{d^2}{ds^2}\bigg|_{s=0}\bigg|_{t=0} c(\widetilde{c}-\exp_x(\widetilde{v}+t\widetilde{\xi}), c\text{-}\exp_x(v+s\overline{\eta})) \tag{1.4.5}$$

$$= \mathfrak{S}_c(x,y)(\xi,\eta). \tag{1.4.6}$$

For historical reason definitions are given for $\langle \xi, \overline{\eta}\rangle_x = 0$, but this hypothesis is not needed for $\mathfrak{S}_c$, also named MTW, to be well defined. Therefore we extend these definitions without the orthogonal hypothesis. The tensor without the orthogonality condition was called the cross curvature by Kim and McCann [75]. This definition makes sense since, as we will see MTW contains some informations about the curvature of the manifold and can be see as a curvature operator on the bi-product $M \times M$, however by admiration I will call $\mathfrak{S}_c$ the Ma-Trudinger-Wang tensor.

*Proof of proposition 1.4.2 and 1.4.3* . We first observe that $M = \nabla_{x,y}c$ is an operator from $T_xM \times T_yM$ into $\mathbb{R}$: $\xi,\eta \mapsto \xi^t M\eta$. Then we compute the MTW tensor starting from (1.4.3):

$$-\frac{2}{3}\mathfrak{S}_c(x,y)(\xi,\eta) = \frac{d^2}{ds^2}\bigg|_{s=0}\frac{d}{dt}\bigg|_{t=0}\left[d_{p=t\xi}\exp_x(\xi)^t\nabla_x c\left(\exp_x(t\xi), c\text{-}\exp_x\left(v+s\overline{\eta}\right)\right)\right]$$
$$= \frac{d^2}{ds^2}|_{s=0}[\xi^t,\xi^t\nabla_{xx}c(x,c\text{-}\exp_x(v+s\overline{\eta}))]$$

where we used the fact that $\dfrac{d^2}{dt^2}\exp_x(t\xi) = 0$. Thus

$$-\frac{2}{3}\mathfrak{S}_c(x,y)(\xi,\eta)$$
$$= \frac{d}{ds}\bigg|_{s=0}\left[\xi^t,\xi^t\nabla_{xx,y}c(x,c\text{-}\exp_x(v+s\overline{\eta}))(-\overline{\eta})^t M^{-1}\left(x,c\text{-}\exp_x(v+s\overline{\eta})\right)\right]$$
$$= \left[\xi^t,\xi^t\nabla_{xx,yy}c(x,y)(-\overline{\eta})^t M^{-1}(x,y)(-\overline{\eta})^t M^{-1}(x,y)\right]$$
$$+ \left[\xi^t,\xi^t\nabla_{xx,y}c(x,y) - M^{-1}(x,y)\nabla_{,y}M(x,y)(-\overline{\eta})^t M^{-1}(x,y)(-\overline{\eta})^t M^{-1}(x,y)\right]$$

We run the symmetric computation on $-\frac{2}{3}\mathfrak{S}_{\widetilde{c}}(x,y)(\widetilde{\xi},\widetilde{\eta})$ to get

$$\begin{aligned}
-\frac{2}{3}&\mathfrak{S}_{\widetilde{c}}(x,y)(\widetilde{\xi},\widetilde{\eta}) \\
&= \Big[M^{-1}(x,y)(-\widetilde{\xi})^t M^{-1}(x,y)(-\widetilde{\xi})^t \nabla_{xx,yy}c(x,y)\widetilde{\eta},\widetilde{\eta}\Big] \\
&- \Big[M^{-1}(x,y)(-\widetilde{\xi})^t M^{-1}(x,y)(-\widetilde{\xi})^t \nabla_x M(x,y) M^{-1}(x,y)\nabla_{x,yy}c(x,y)\widetilde{\eta},\widetilde{\eta}\Big]
\end{aligned}$$

As $M^{-1}(x,y)-\widetilde{\xi}=\xi$, $(-\overline{\eta})^t M^{-1}(x,y)=\widetilde{\eta}$, $\nabla_x M=\nabla_{xx,y}$ and $\nabla_{,y}M=\nabla_{x,yy}c(x,y)$, we get that both computation are equal to

$$\big[\xi^t,\xi^t\nabla_{xx,yy}c(x,y)\widetilde{\eta}\widetilde{\eta}\big]-\xi^t\xi^t\nabla_{xx,y}c(x,y)M^{-1}(x,y)\nabla_{x,yy}c(x,y)\widetilde{\eta}\widetilde{\eta}.$$

Finally, developing in index we do get the equivalence with (1.4.2). □

These equivalent formulations have the advantage to see everything from $x$ and make the tensor local. Following this remark Figalli and Rifford extend the definition of MTW. First they noticed that in definition (1.4.3), one can see $\mathfrak{S}_c(x,y)(\xi,\eta)$ as $\mathfrak{S}_c(x,v)(\xi,\overline{\eta})$ defined for all $v\in c\text{-I}(x)$ and $(\xi,\overline{\eta})\in T_xM\times T_xM$. Secondly they observed that one can push $v$ in the definition up to the $c$-tangent focal locus ($c$-NF), as long as the differential of $r\mapsto c\text{-}\exp_x(r)$ is invertible. Therefore they came up with the following definition of the extended Ma-Trudinger-Wang tensor.

**Definition 1.4.4** (Extented MTW tensor)**.** *Let $x\in M$, $v\in c\text{-NF}(x)$, and $(\xi,\eta)\in T_xM\times T_xM$. Since $y:=c\text{-}\exp_x v$ is not c-conjugate to $x$, by the Inverse Function Theorem there is an open neighbourhood $\mathcal{V}$ of $(x,v)$ in $TM$ and an open neighbourhood $\mathcal{W}$ of $(x,y)$ in $M\times M$, such that*

$$\begin{aligned}
\Psi_{(x,v)}:\mathcal{V}\subset TM &\longrightarrow \mathcal{W}\subset M\times M \\
(x',v') &\longmapsto \big(x',c\text{-}\exp_{x'}(v')\big)
\end{aligned}$$

*is a smooth diffeomorphism from $\mathcal{V}$ to $\mathcal{W}$.*
*Then we may define the extended cost $\widehat{c}_{(x,v)}:\mathcal{W}\to\mathbb{R}$ by*

$$\widehat{c}_{(x,v)}(x',y'):=\frac{1}{2}\big|\Psi^{-1}_{(x,v)}(x',y')\big|^2_{x'}\qquad\forall\,(x',y')\in\mathcal{W}. \tag{1.4.7}$$

*If $v\in c\text{-I}(x)$ then for $y'$ close to $c\text{-}\exp_x v$ and $x'$ close to $x$ we have $\widehat{c}_{(x,v)}(x',y')=c(x',y')$. For every $x\in M$, $v\in c\text{-NF}(x)$ and $(\xi,\eta)\in T_xM\times T_xM$, the extended Ma-Trudinger-Wang tensor at $(x,v)$ is defined by the formula*

$$\overline{\mathfrak{S}}_{(x,v)}(\xi,\eta)=-\frac{3}{2}\left.\frac{d^2}{ds^2}\right|_{s=0}\left.\frac{d^2}{dt^2}\right|_{t=0}\widehat{c}_{(x,v)}\Big(\exp_x(t\xi),\,c\text{-}\exp_x(v+s\eta)\Big).$$

Be careful that this time $\eta\in T_xM$ is the previous $\overline{\eta}$. For geographic reason we will work with the second extension, The key tool for regularity will be to see when this tensor is positive, non-negative or bounded from below. It leads to the following definitions:

**Definition 1.4.5.** *Let $M$ be a smooth Riemannian manifold, and $c$ a cost satisfying 1.1.a. We* *say $M$ satisfies* $\mathbf{MTW}(K,C)$ *if for all* $(x,y) \in M \times M \setminus c\text{-cut}(M)$ *and* $(\xi,\eta) \in T_xM \times T_xM$*,*

$$\mathfrak{S}_c(x,y)(\xi,\eta) \geq K|\xi|^2|\eta|^2 - C|\langle \xi,\eta\rangle||\xi||\eta|.$$

*Similarly we say that $M$ satisfies* $\overline{\mathbf{MTW}}(K,C)$ *if for all* $(x,v) \in M \times c\text{-NF}(x)$ *and* $\xi,\eta \in T_xM \times T_xM$*,*

$$\overline{\mathfrak{S}}_c(x,v)(\xi,\eta) \geq K|\xi|^2|\eta|^2 - C|\langle \xi,\eta\rangle||\xi||\eta|.$$

*If we restrict to vectors such that* $\langle \xi.\eta\rangle = 0$*, we say that $M$ satisfies* $\mathbf{MTW}(K)$ *if*

$$\mathfrak{S}_c(x,y)(\xi,\eta) \geq K|\xi|^2|\eta|^2.$$

*Similarly we say $M$ satisfies* $\overline{\mathbf{MTW}}(K)$ *if*

$$\overline{\mathfrak{S}}_c(x,v)(\xi,\eta) \geq K|\xi|^2|\eta|^2.$$

*Finally* $\mathbf{MTW}$ *will stand for* $\mathbf{MTW}(0)$ *and* $\mathbf{MTW}_s$ *for* $\mathbf{MTW}(0)$ *in a strict form.*

Note that in previous definitions we consider $\xi,\eta$ in $T_xM$. Definitions would have been almost unchanged taking $\eta \in T_yM$, one just need to be careful with the norm $\nabla^2_{(x,y)}c(x,y)$ that can make crash the constant.

**Remark 1.4.6.** *The condition* $\mathbf{MTW}(0)$*,* $\mathbf{MTW}_s(0)$*,* $\mathbf{MTW}(0,0)$*,* $\mathbf{MTW}_s(0,0)$ *are also known under the name A3w, A3s, B3w, B3s. The second type of notation are usually used when the authors use the cross curvature name instead of MTW. This is the case in the paper of Kim and McCann [75]. The notation A3w and A3s are the original ones introduce by Ma, Trudinger and Wang [85].*

## Examples

Finding example satisfying $\mathbf{MTW}(0)$, $\mathbf{MTW}(0,0)$,..., is not an easy task. Ma, Trudinger and Wang proved in particular that the condition $\mathbf{MTW}_s(0)$ holds for the cost $c(x,y) = \sqrt{1|x-y|^2}$, $c(x,y) = \sqrt{1+|x-y|^2}$ and more generally $c(x,y) = \left(\sqrt{\varepsilon + |x-y|^2}\right)^{\frac{p}{2}}$. They also remark that $\mathfrak{S}_c = 0$ for the quadratic cost on $\mathbb{R}^n$ [85]. The first Riemmanian example satisfying $\mathbf{MTW}_s(0)$ was the sphere $S^n$ with the quadratic geodesic cost [82]. Then Figalli Rifford and Villani proved that a small $C^4$ perturbation of the metric still satisfies $\mathbf{MTW}_s(0)$ [49]. In dimension 2 Delanoë and Ge on one side and Figalli and Rifford on the other, independently give a formula for the MTW tensor it leads again to perturbation results for the MTW tensor [35, 34, 47]. However in all these cases the tensor satisfies $\mathbf{MTW}_s(0)$, the first example that allows flat sectional curvature is the product of round sphere, this result is due to Kim and McCann [75]. In details Kim and McCann showed that a product of manifold satisfying $\mathbf{MTW}(0,0)$ is $\mathbf{MTW}(0,0)$, but there is no reason for a product of $\mathbf{MTW}(0)$ to be $\mathbf{MTW}(0)$. They also show in the same paper that the tensor MTW is stable under Riemannian submersion quotient, this open to a class of new examples as $\mathbb{C}P^n$. This shows regularity of optimal maps on multiple products of round spheres, giving the first regularity results of optimal maps on curved (non flat) manifolds allowing zero curvature [42]. Another surprising result is brought by Figalli Rifford and Villani they prove that an ellipsoid of revolution does not satisfies **MTW** when it is to flat [47]. To complete this work Bonnard Caillau and Rifford proved that for such an ellipsoid the injectivity domains are not convex [13].

### 1.4.b Jacobi fields I

Before going to geometric consequences of the positivity of the MTW tensor, we first give a fundamental proposition that expresses $MTW$ with the Jacobi fields of $M$ in the case of the quadratic geodesic cost. We start with some notation and refer to appendix A for definitions related to the Riemannian structure.
For $(x, y_s) \in M \times M \setminus \mathrm{cut}(M)$ we define $\gamma_s$:

$$\gamma_s : \quad [0,1] \to M$$
$$\tau \mapsto \exp_x(\tau v_s),$$

the geodesic path starting in $x$ at time 0 and finishing in $y_s$ at time 1. We then define $J_s(\tau, \xi)$ as the Jacobi field along $\gamma_s$ in direction $\xi$, that is:

$$J_s(\tau, \xi) = \frac{d}{dt}|_{t=0}\gamma_s(t, \tau, \xi)$$

where $\gamma_s(t, \tau, \xi)$ is the geodesic path starting from $\exp_x(t\xi)$ at time $\tau = 0$ and reaching $y_s$ at time $\tau = 1$.
We can now state the following proposition:

**Proposition 1.4.7.** *Let* $c = \frac{1}{2}d^2$, $(x,y) \in M \times M \setminus cut(M)$ *with* $y = \exp_x v$, *and* $\xi, \eta \in T_xM \times T_xM$. *Then*

$$\mathfrak{S}_{x,y}(\xi, \eta) = \frac{3}{2}\frac{d^2}{ds^2}\bigg|_{s=0} \langle \dot{J}_s(0, \xi).\xi\rangle_x,$$

*where* $y_s = \exp_x(v + s\eta)$.

It says that $MTW$ is the second derivative of the initial acceleration for the Jacobi field in direction $\xi$ finishing at $y_s$.

*Proof.* By definition

$$\mathfrak{S}_{x,y}(\xi, \eta) = -\frac{3}{2}\left.\frac{d^2}{ds^2}\right|_{s=0}\left.\frac{d^2}{dt^2}\right|_{t=0} c(\exp_x(t\xi), \exp_x(v + s\eta)).$$

Still by definition

$$c(\exp_x(t\xi), \exp_x(v + s\eta)) = \frac{1}{2}\int_0^1 |\partial_\tau \gamma_s(t, \tau, \xi)|^2 d\tau.$$

Therefore

$$\left.\frac{d^2}{dt^2}\right|_{t=0} \frac{1}{2}\int_0^1 |\partial_\tau \gamma_s(t, \tau, \xi)|^2 d\tau = \frac{d}{dt}\int_0^1 \langle \frac{D}{Dt}\partial_\tau \gamma_s(t, \tau, \xi), \partial_\tau \gamma_s(t, \tau, \xi)\rangle d\tau \tag{1.4.8}$$

Then we can exchange the first derivative in $t$ and $s$, and since the path $\tau \mapsto \gamma_s(t,\tau,\xi)$ is geodesic in $\tau$, we have $\partial\tau^2\gamma = 0$. Moreover, as $\gamma_s(t,1,\xi) = y_s$, we find that $\frac{d}{dt}\gamma_s(t,1,\xi) = 0$, and since $\gamma_s(t,0,\xi)$ is a geodesic path with respect to $t$ with initial velocity $\xi$, it leads to $\frac{d}{dt}\gamma_s(t,0,\xi) = \xi$. Plugging these informations in (1.4.8) we get

$$\begin{aligned}\frac{d^2}{dt^2}\bigg|_{t=0}\frac{1}{2}\int_0^1 |\partial_\tau\gamma_s(t,\tau,\xi)|^2 d\tau &= \frac{d}{dt}\int_0^1 \frac{d}{d\tau}\langle\partial_t\gamma_s(t,\tau,\xi), \partial_\tau\gamma_s(t,\tau,\xi)\rangle d\tau \\ &= \frac{d}{dt} - \langle\partial_t\gamma_s(t,0,\xi), \partial_\tau|_{\tau=0}\gamma_s(t,\tau,\xi)\rangle \\ &= -\langle\partial_t\gamma_s(0,0,\xi), \partial_\tau|_{\tau=0}\partial_t|_{t=0}\gamma_s(t,\tau,\xi)\rangle)\end{aligned} \tag{1.4.9}$$

□

The approach by Lee and McCann in [76] is similar, but as we are working with the extended tensor, we go a little bit further and give the extended proposition.

**Proposition 1.4.8.** *For $c = \frac{1}{2}d^2$, $(x,v) \in M \times \mathrm{NF}(x)$, and $\xi,\eta \in T_xM \times T_xM$, then*

$$\overline{\mathfrak{S}}_{x,y}(\xi,\eta) = \frac{3}{2}\frac{d^2}{ds^2}|_{s=0}\langle \dot{J}_s(0,\xi).\xi\rangle_x,$$

*where $y_s = \exp_x(v+s\eta)$*

*Proof.* The proof is exactly the same as the one done above for proposition 1.4.7. □

For convenience we rewrite this proposition in the spirit of the Rifford and Figalli approach [46]. The reformulation corresponds to a time reparametrization of the Jacobi field equation. We consider the Jacobi field equation along $\tau \mapsto \gamma_s(\tau\frac{v_s}{|v_s|})$ the geodesic path starting from $x$ at time 0 and reaching $y_s$ at time $\tau_s = |v_s|$. As recalled in Appendix A the Jacobi field equation for $I$ a Jacobi field is

$$\ddot{I}(\tau) + R(\tau)I(\tau) = 0, \tag{1.4.10}$$

where R is given by the Riemmanian tensor. Note that R depends on $s$. We define the two fundamental solutions $I_0$, $I_1$ of (1.4.10) fixing the initial conditions:

$$\begin{aligned} I_0(0) = 0 \quad \dot{I}_0(0) = I_n \\ I_1(0) = I_n \quad \dot{I}_1(0) = 0. \end{aligned} \tag{1.4.11}$$

In general the dot will refer to the derivation with respect to the parameter $\tau$, the $'$ for the derivation with respect to the parameter $s$. Thus we reformulate proposition 1.4.8.

**Proposition 1.4.9.** *For $c = \frac{1}{2}d^2$, $(x,v) \in M \times \mathrm{NF}(x)$, and $\xi,\eta \in T_xM \times T_xM$, then*

$$\overline{\mathfrak{S}}_{x,v}(\xi,\eta) = -\frac{3}{2}\frac{d^2}{ds^2}\bigg|_{s=0}\langle \tau_s I_0^{-1}(\tau_s)I_1(\tau_s)\xi.\xi\rangle_x,$$

*where $v + s\eta = v_s$ and $\tau_s = |v_s|$.*

One has to be careful performing the derivation, as $I_0$ and $I_1$ depends on $s$.

*Proof.* We just have to check that $J_s(0,\xi) = -\tau_s I_0^{-1}(\tau_s)I_1(\tau_s)\xi$. To see $J_s(.,\xi)$ as a Jacobi field along $\gamma_s(.\frac{v_s}{|v_s|})$, we have to reparametrize $J_s$. Let $I_s(\tau, xi) = J_s(\frac{\tau}{\tau_s}, \xi)$. $I_s$ is a Jacobi field along $\gamma_s(.\frac{v_s}{|v_s|})$, thus we can express it in term of $I_0$, $I_1$:

$$I_s(\tau) = I_1(\tau)I_s(0) + I_0(\tau)\dot{I}_s(0). \tag{1.4.12}$$

Moreover $I_s(0) = J_s0 = \xi$ and $I_s(\tau_s) = J_s(1) = 0$, thus (1.4.12) at time $\tau_s$ gives

$$\dot{I}_s(0) = -I_0^{-1}(\tau_s)I_1(\tau_s)\xi.$$

Since $\dot{J}_s(\frac{\tau}{\tau_s}, \xi) = \tau_s \dot{I}_s(\tau, \xi)$ the proposition is proved. □

Now we can state some positive results involving $MTW$.

## 1.5 Some $\mathbf{MTW}(K, C)$ consequences

### 1.5.a Sectional curvature

Loeper was the first to obtain geometric information on Riemannian manifold with the MTW tensor [82].

**Theorem 1.5.1** (Loeper)**.**
*Let $M$ be a Riemannian manifold satisfying* **MTW** *(resp.* $\mathbf{MTW}(K)$*) for the cost* $\frac{1}{2}d^2$*, then all the sectional curvature are nonnegative ( resp* $\geq K$ *).*

*Proof.* For the proof we only have to consider the case $x = y$ i.e. $v = 0$. Thus $y_s = \exp_x(s\eta)$ for $\eta, \xi \in U_xM$, and then in (1.4.9) $I_0$ and $I_1$ are independent of $s$ and $\tau_s = s$. Therefore

$$\mathfrak{S}_{x,0}(\xi, \eta) = \frac{3}{2}\frac{d^2}{ds^2}\bigg|_{s=0} \langle sI_0^{-1}(s)I_1(s)\xi, \xi\rangle_x.$$

Using the Jacobi equation (1.4.11) we find the following limited development:

$$\begin{aligned}
I_1(s) &= Id - \frac{s^2}{2}R(0) + s^2\varepsilon(s)\\
I_0(s) &= sId - \frac{s^3}{6}R(0) + s^3\varepsilon(s)\\
sI_0^{-1}(s) &= Id + \frac{s^2}{6}R(0) + s^2\varepsilon(s)\\
sI_0^{-1}(s)I_1(s) &= (Id + \frac{s^2}{6}R(0) + s^2\varepsilon(s))(Id - \frac{s^2}{2}R(0) + s^2\varepsilon(s))\\
&= Id - \frac{s^2}{3}R(0)s^2 + s^2\varepsilon(s).
\end{aligned}$$

Two times the $s^2$ coefficient is the second derivative in $s$ therefore

$$\mathfrak{S}_{x,x}(\xi,\eta) = \langle R(0)\xi,\xi\rangle_x = \sigma_x(\xi,\eta).$$

If $\mathbf{MTW}(K)$ holds true then $\mathfrak{S}_{x,x}(\xi,\eta) \geq K$ and $\sigma_x(\xi,\eta) \geq K$. □

This first result is important as it shows that the MTW tensor contains some information about the geometry of the manifold, In particular $\mathbf{MTW}(K)$ implies that $M$ is compact.
However, as Kim shown the tensor MTW contains more information than the sectional curvature, in particular he gave examples of manifold with positive sectional curvature everywhere for which the condition **MTW** is not satisfied [72].

### 1.5.b Mother computation, and convexity result

#### Analysis

The convexity of subdifferentials, for convex functions, is the key for regularity in $\mathbb{R}^n$ with the euclidean cost [22]. As introduced by Ma Trudinger and Wang [85], for a general cost satisfying 1.1.a the subdifferential becomes the $c$-subdifferential, a segment $(1-s)p_0 + sp_1$ becomes the $c$-segment $y_s = c\text{-}\exp_x((1-s)p_0 + sp_1)$. Thus a natural question to tackle the regularity question arises: taking a $c$-convex function $\psi$, are the $c$-subdifferential $c$-convex ? It means that the $c$-segment joining two points in the $c$-subdifferential of one point stays in the $c$-subdifferential of this point. That is exactly asking the convexity of the $c$-subgradient of $\psi$. The first non trivial example of a $c$-convex function is the maximum of two $c$-convex supports. Let us explore this example.
Let $\psi_{\overline{x},y_0,y_1} = \max(D(\overline{x},p_0), D(\overline{x},p_1))$, with by definition $y_0 = c\text{-}\exp_{\overline{x}}(p_0)$ and $y_1 = c\text{-}\exp_{\overline{x}}(p_1)$. We want to know if $y_s = c\text{-}\exp_{\overline{x}}(p_s)$ with for all $s \in [0,1]$, $p_s = (1-s)p_0 + sp_1$ is in the $c$-subdifferential of $\psi_{\overline{x},y_0,y_1}$ at $\overline{x}$. To prove this we need to see that for all $x$ and for all $s \in [0,1]$ the $c$-support candidate $D(\overline{x},p_s)(x)$ is below $\psi_{\overline{x},y_0,y_1}(x) - \psi_{\overline{x},y_0,y_1}(\overline{x})$. In our case it means to ask: do we have for all $x$ and $s \in [0,1]$

$$D(\overline{x},p_s)(x) \leq \max(D(\overline{x},p_0), D(\overline{x},p_1))\,? \tag{1.5.1}$$

This question leads to the mother computation: we fix $x$ and prove that the function $h(s) = D(\overline{x},p_s)(x)$, for $s \in [0,1]$ is maximal at its endpoints. It gives (1.5.1) for all $s \in [0,1]$, this result, (1.5.1), is due to Loeper [82]. The mother computation is even more general as described in proposition 1.5.2. The proof is due to Kim and McCann [74, 73] and is the one given below. Loeper and Villani slightly extend this proposition in order to give a quantified and robust formulation [83].

#### Mother computation

Let $\overline{x}, x$ be points in $M$. For $p_0, p_1 \in c\text{-}\mathrm{I}(\overline{x})$, we set $p_s = (1-s)p_0 + sp_1$ and $y_s = c\text{-}\exp_{\overline{x}}(p_s)$, $s \in [0,1]$. We consider the $c$-convex $D(\overline{x},p_0), D(\overline{x},p_1)$ and $D(\overline{x},p_s)$. Moreover we suppose that

$\overline{x}, x \in M \setminus c\text{-cut}(y_s)$, it defines for $s \in [0,1]$

$$\overline{q}_s = -\nabla_y c(\overline{x}, y_s), \quad q_s = -\nabla_y c(x, y_s).$$

We finally denote $\overline{q} = \overline{q}_0$ and $q = q_0$. From $y_s = c\text{-}\exp_{\overline{x}}(p_s)$ we deduce

$$\dot{y}_s = (\nabla_{x,y} c)^{-1}(\overline{x}, y_s)(p_0 - p_1) = \eta_s \in T_{y_s} M.$$

and

$$\ddot{y}_s = \left\langle -(\nabla_{x,y} c)^{-1}(\overline{x}, y_s) \nabla_{x,yy} c(\overline{x}, y_s) \eta_s, \eta_s \right\rangle_x .$$

Let $h(s) = D(\overline{x}, p_s)(x)$ then

$$\dot{h}(s) = \frac{d}{ds} c\left(\overline{x}, c\text{-}\exp_{\overline{x}}\left(p_0 + s(p_1 - p_0)\right)\right) - c\left(x, c\text{-}\exp_{\overline{x}}\left(p_0 + s(p_1 - p_0)\right)\right),$$

in term of $c$ we can write it as

$$\dot{h}(s) = \langle \nabla_{,y} c(\overline{x}, y_s) - \nabla_{,y} c(x, y_s), \dot{y}_s \rangle = \langle -\overline{q}_s + q_s, \eta \rangle. \tag{1.5.2}$$

For the second derivative we have

$$\ddot{h}(s) = \frac{d^2}{ds^2} c(\overline{x}, y_s) - c(x, y_s) = \langle \nabla^2_{,yy} c(\overline{x}, y_s).\eta_s - \nabla^2_{,yy} c(x, y_s).\eta_s, \eta_s \rangle + \langle q_s - \overline{q}_s, \ddot{y}_s \rangle.$$

We define

$$\Phi(t) = \frac{d^2}{d\tau^2} c\left(c\text{-}\exp_{y_s}(\overline{q}_s + t(q_s - \overline{q}_s)), \exp_{y_s}(\tau \eta_s)\right)$$

and observe that

$$\langle \nabla^2_{,yy} c(\overline{x}, y_s).\eta_s - \nabla^2_{,yy} c(x, y_s).\eta_s, \eta_s \rangle = \Phi(0) - \phi(1)$$

and

$$\Phi'(0) = (\nabla_{x,y} c)^{-1}(\overline{x}, y_s)(\overline{q}_s - q_s) \nabla_{x,yy} c(\overline{x}, y_s) \eta_s \eta_s = \langle q_s - \overline{q}_s, \ddot{y}_s \rangle.$$

Finally

$$\ddot{h}(s) = -[\Phi(1) - \Phi(0) - \Phi'(0)]$$

By Taylor formula we get

$$\ddot{h}(s) = -\frac{2}{3}\frac{3}{2} \int_0^1 \Phi''(t)(1-t) dt$$

and finally by symmetry of the formula (see proposition 1.4.3)

$$\ddot{h}(s) = \frac{2}{3} \int_0^1 \mathfrak{S}_{\left(c\text{-}\exp_{y_s}((1-t)q_s + t\overline{q}_s), y_s\right)} \left(\nabla_{x,y} c\right)^{-1} \left(\overline{x}_t, y_s\right) \left(\overline{q}_s - q_s, p_1 - p_0\right) (1-t) dt \tag{1.5.3}$$

One may be concerned by the existence of the path $y_s$, and the $c$-segment $c\text{-}[(\overline{q}_s, q_s]_{y_s}$, since we need to avoid $c\text{-cut}(\overline{x})$, $c\text{-cut}(x)$ and $c\text{-cut}(y_s)$. The perturbation lemma proved by Figalli and Villani [50] allows us to perturb in $C^2$ topology these paths to be sure that they are well defined everywhere but a finite number of time. We summarize the mother computation in the following proposition.

**Proposition 1.5.2.** *Let* $\overline{x}, x \in M \times M \setminus c\text{-cut}(\overline{x})$*,* $p_0, p_1 \in c\text{-I}(\overline{x})$*. For any* $s \in [0,1]$ *we set* $p_s = (1-s)p_0 + sp_1$*,* $y_s = c\text{-}\exp_{\overline{x}}(p_s)$*,* $\overline{q}_s = -\nabla_y c(\overline{x}, y_s)$*,* $q_s = -\nabla_y c(x, y_s)$ *and finally*

$$h(s) = c\left(\overline{x}, c\text{-}\exp_{\overline{x}}(p_s)\right) - c\left(x, c\text{-}\exp_{\overline{x}}(p_s)\right).$$

*Then* $h$ *is* $C^2$*,*

$$\dot{h}(s) = \langle \nabla_{,y} c(\overline{x}, y_s) - \nabla_{,y} c(x, y_s), \dot{y}_s \rangle = \langle -\overline{q}_s + q_s, \eta \rangle. \tag{1.5.4}$$

*and*

$$\ddot{h}(s) = \frac{2}{3} \int_0^1 \mathfrak{S}_{\left(c\text{-}\exp_{y_s}((1-t)q_s + t\overline{q}_s), y_s\right)} \left(\nabla_{x,y} c\right)^{-1} \left(\overline{x}_t, y_s\right) \left(\overline{q}_s - q_s, p_1 - p_0\right)(1-t)dt.$$

This proposition is the starting point of any results concerning the MTW tensor, we give some of them right now.

### 1.5.c $c$-convexity of $c$-differential

We are now able to answer our primary question on the $c$-convexity of $c$-differential. The next theorem is due to Loeper [82].

**Theorem 1.5.3.** *Let* $M$ *be a compact Riemannian manifold, let* $c$ *be a cost satisfying 1.1.a then* **MTW** *holds true if and only if for all* $\psi$ *c-convex and for all* $\overline{x} \in M$*,* $\partial_c \psi(\overline{x})$ *is c-convex. In other words for any* $\overline{x} \in M$*,* $\nabla^- \psi(\overline{x}) = \nabla_c^- \psi(\overline{x})$ *if and only if* **MTW** *holds true.*

The set $\nabla^- \psi(\overline{x})$ is always convex, thus $\nabla^- \psi(\overline{x}) \subset \nabla_c^- \psi(\overline{x})$ implies that $\partial_c \psi(\overline{x})$ is $c$-convex.

*Proof.* **MTW implies $c$-convexity.** Let $\psi$ be a $c$-convex function, and $p_0, p_1 \in \nabla^-(\overline{x})$, thus $D(\overline{x}, p_0) + \psi(\overline{x})$ and $D(\overline{x}, p_1) + \psi(\overline{x})$ are $c$-support functions at $\overline{x}$. We need to show that $D(\overline{x}, p_s)$ is also a $c$-support function at $\overline{x}$ for all $s$ in $[0,1]$. We need to show that $D(\overline{x}, p_s)$ is below $\psi(\cdot) - \psi(\overline{x})$.
Using the previous computation, and notation in proposition 1.5.2, it suffices to prove that $h$ achieves its maximum at its endpoints. Could this be false ? Let $\overline{s}$ be a maximum point of $h$, if $s$ is not a endpoint then by (1.5.4) $\dot{h}(\overline{s}) = 0$, it is exactly the orthogonal condition in definition 1.4.2, therefore (1.5.3) and **MTW** gives

$$\ddot{h}(\overline{s}) \leq 0. \tag{1.5.5}$$

Let us suppose an instant that **MTW**$s$ holds true then (1.5.5) is changed with the stronger inequality

$$\ddot{h}(\overline{s}) > 0.$$

We immediately obtain a contradiction with the fact that $h$ is maximal at $\overline{s}$ and consequently $h$ achieve its maximum at its endpoints:

$$D(\overline{x}, p_s)(x) < \max(D(\overline{x}, p_0)(x), D(\overline{x}, p_1)(x)) \tag{1.5.6}$$

thus for all $x$

$$D(\overline{x}, p_s)(x) + \psi(\overline{x}) \leq \psi(x).$$

It gives $p_s \in -\nabla^-(\overline{x})$ and proves the $c$-convexity of $\partial_c\psi(\overline{x})$.

We now come back to the hypothesis **MTW**. There exists $C > 0$ such that $\mathbf{MTW}(0, -C)$ holds true [83]. Moreover (1.5.3) gives

$$\ddot{h}(t) \geq -\overline{C}|\dot{h}(t)|. \tag{1.5.7}$$

To conclude we take $\delta > 0, k > 1$ and define on $[0, 1]$ the auxiliary function

$$g(t) = h(t) - \delta(t - \frac{1}{2})^k.$$

As before let us suppose that the maximum of $g$ is achieved in $\bar{t}$ we get

$$|\dot{h}(\bar{t})| = k|\delta|(\bar{t} - \frac{1}{2})^{k-1}$$

and

$$\ddot{h}(\bar{t}) + k(k-1)\delta(\bar{t} - \frac{1}{2})^{k-2} \leq 0$$

Plugging these two equations in 1.5.7 we get

$$(k-1)\delta \leq \overline{C}|\delta|(\bar{t} - \frac{1}{2}).$$

This is impossible for $k$ fixed large enough. Then we make $\delta$ going to 0 to obtain the contradiction thus $h$ achieves its maximum at its endpoints and we conclude as in the previous case.

***c*-convex implies MTW.**

To prove the converse implication we suppose that there exists $\overline{x}, y \in M \times M \setminus c\text{-cut}(x)$, ($y = c\text{-}\exp_{\overline{x}} v$) and $\eta, \xi$ two tangent vectors at $\overline{x}$, such that $\mathfrak{S}_{(\overline{x},y)}(\xi, \eta) < 0$. Then $\mathfrak{S}$ is also negative for a neighbourhood $W$ of $\overline{x}, y, \xi, \eta$. We consider the $c$-segment $y_s = c\text{-}\exp_{\overline{x}}(v + s\eta)$ for $s \in [-\varepsilon, \varepsilon]$. As usual we call $h(s) = D(\overline{x}, p_s)(x)$, by the mother computation 1.5.2 we find that $\dot{h}(0) = \langle q - \overline{q}, \nabla_{\overline{x},y}c^{-1}\eta\rangle = 0$ as long as we take $q$ such that $q - \overline{q}$ and $\xi$ are parallel.

We also get

$$\ddot{h}(0) = \frac{2}{3}\int_0^1 \mathfrak{S}_{(\overline{c\text{-}\exp}_y((1-t)q + t\overline{q}), y)}(\nabla_{x,y}c)^{-1}(\overline{x}_t, y)(\xi, \eta)(1-t)dt.$$

For $\varepsilon, t$ small enough everything lies in $W$ so $\ddot{h}(0) < 0$ and $\dot{h}(0) = 0$ thus 0 is strict local maximum for $h$. It means

$$D(\overline{x}, y) > \max(D(\overline{x}, y_{-\varepsilon}), D(\overline{x}, y_\varepsilon)).$$

Therefore let us define $\psi = \psi_{\overline{x}, y_{-\varepsilon}, y_\varepsilon} = \max(D(\overline{x}, y_{-\varepsilon}), D(\overline{x}, y_\varepsilon))$ then $y_{-\varepsilon}, y_\varepsilon \in \partial_c\psi(\overline{x})$ but the $c$-segment $c$-$[y_{-\varepsilon}, y_\varepsilon] \notin \partial_c\psi(\overline{x})$: the $c$-differential of $\psi$ at $\overline{x}$ is not $c$-convex. □

This result is the starting point to construct a counter example for the optimal transport when **MTW** does not hold.

### 1.5.d Convexity of $c$-convex set

A very similar theorem deals with the convexity of the $c$-convex functions. It comes from a very interesting paper, with economics application, of Figalli Kim and McCann [44], we also mention that Sei independently found the same result [99] . The question is what happens if we look for the convexity of the function $h$ instead of just the maximal-endpoints property also known as the quasi-convexity. It leads us to the following theorem:

**Theorem 1.5.4.** *Let $M$ be a compact Riemannian manifold, let $c$ a cost satisfying 1.1.a then* $\mathbf{MTW}(0,0)$ *is equivalent to the convexity of the set of c-convex functions.*

*Proof.* **$\mathbf{MTW}(0,0)$ implies the convexity of the set of $c$-convex functions.** Let $f_0, f_1$ be two $c$-convex functions, with c-transform $g_0, g_1$. We define $f_t = (1-t)f_0 + tf_1$. If one equals to infinity then it is ok. Otherwise let $x_0 \in M$ and let $y_i \in \partial_c f_i(x_0)$ for $i = 0, 1$ then

$$f_t(x) \geq f_t(x_0) + (1-t)D(x_0, p_0) + tD(x_0, p_1).$$

If

$$(1-t)D(x_0, p_0) + tD(x_0, p_1) \geq D(x_0, p_t), \tag{1.5.8}$$

then

$$y_t = (\nabla_x c(x))^{-1}(-p_t) \in \partial_c f_t(x_0)$$

and $f_t$ is $c$-convex.

Thus we need to prove (1.5.8) which is stronger than (1.5.6). Using notation and computation of proposition 1.5.2 we get the equation (1.5.3). Since we suppose $\mathbf{MTW}(0,0)$ we have that $\ddot{h}$ is positive for any $t \in [0,1]$ and therefore the convexity of the function $D(x_0, p)$ with respect to the $p$ variable is exactly (1.5.8).

We can notice that $\mathbf{MTW}s(0,0)$ implies the strict convexity. We can also, as we will see later in chapter 3, obtain uniform convexity estimations.

**The convexity of the set of $c$-convex functions implies $\mathbf{MTW}(0,0)$.** This demonstration is similar to the one done in the proof of theorem 1.5.3, but this time we get that $\ddot{h}(s) < 0,\ \forall\ s \in [-\varepsilon, \varepsilon]$ thus $h$ is concave on this segment. As a consequence the function $\theta = (1-t)D(x_0, p_{-\varepsilon}) + tD(x_0, p_\varepsilon)$ can not be $c$-convex. Indeed, at $\overline{x}$ $\nabla\theta = (1-t)p_{-\varepsilon} + tp_{-\varepsilon} = p_t$, any $c$-support at $\overline{x}$ must check that $\nabla_x c(\overline{x}, y) = -p_t$ it leaves only one candidate: $D(\overline{x}, p_t)$ but the concavity of $h$ gives that for any $x$ close to $\overline{x}$ in the $\xi$ direction $\theta - \theta(\overline{x})$ is below $D(\overline{x}, p_t)$. □

### 1.5.e Nothing without MTW

At the end, we have to convince the reader that MTW is the good tool for regularity. Following an idea of Loeper [82] we prove that without **MTW** we may not construct smooth optimal map even for some nice density distributions.

**Theorem 1.5.5.** *Let $M$ be a compact Riemannian manifold and let $c$ be a cost satisfying 1.1.a. If there exists $\overline{x}, y \in M \times M \setminus c\text{-cut}(x)$ and $\xi, \eta \in T_xM \times \mathbb{T}_xM$ such that $\mathfrak{S}_{(\overline{x},y)}(\xi,\eta) < 0$ then* *there exist* $f,g$ *two* $C^\infty$ *densities on* $M$ *such that the optimal transport from* $f dvol$ *to* $g dvol$ *is not continuous.*

*Proof.* By 1.5.3 we already know that there exist $y_0$ and $y_1$ such that the $c$-differential at $\overline{x}$ of the $c$-convex function $\psi = \psi_{\overline{x},y_0,y_1} = \max(D(\overline{x},y_0), D(\overline{x},y_1))$ is not $c$-convex. The next step is to construct two measures such that $\psi$ is the Kantorovich potential for the associated Monge problem. For the source measure, $\mu$, we take any smooth positive probability measure, for the target one, $\nu$, we partition M into two sets

$$X_0 = \{x \in M, y_0 \in \partial_c\psi(x)\}$$

and

$$X_1 = X \setminus X_0,$$

note that $\mu$ almost everywhere: $\forall\ x \in X_1$ we have $y_0 \in \partial_c\psi(x)$. Then we define $a_0 = \mu(X_0)$, $a_1 = \mu(X_1)$ and finally $\nu = a_0 y_0 + a_1 y_1$. The transport map $T(X_i) = y_i$ is the optimal transport from $\mu$ to $\nu$ since the associated transport plan is supported on the $c$-convex set $\partial_c\psi$. Thus $\psi$ is the optimal Kantorovich potential, unique as we can suppose $\psi(\overline{x}) = 0$. Therefore

$$\int \psi\mu - \int \psi^c\nu = \int c(x, T(x))\mu.$$

Let $g_\varepsilon$ be a sequence of smooth densities weakly converging to $\nu$. The optimal transport from $f$ to $g_\varepsilon$ gives a unique Kantorovich potential $\psi_\varepsilon$ such that $\psi_\varepsilon(\overline{x}) = 0$ and a unique transport map $T_\varepsilon$. Moreover we have

$$\int \psi_\varepsilon\mu - \int \psi_\varepsilon^c\nu_\varepsilon = \int c(x, T_\varepsilon(x))\mu. \tag{1.5.9}$$

We can extract a subsequence of $(\psi_\varepsilon, T_\varepsilon)$ uniformly convergent to $\psi_0$ and simply convergent to $T_0$. Then we pass to the limit in (1.5.9) to get:

$$\int \psi_0\mu - \int \psi_0^c\nu = \int c(x, T_0(x))\mu. \tag{1.5.10}$$

By uniqueness $\psi_0 = \psi$.

It remains to show that all $\psi_\varepsilon$, for $\varepsilon$ small enough, cannot be differentiable. Could this be false, by lemma 1.5.6 given at the end of the proof, for any $p \in \nabla^-\psi(\overline{x})$ we can find sequences $x_k \to x$, $p_k \in T_{x_k}M \to p$, such that $p_k \in \nabla^-\psi_k(x_k)$. If $\psi_k$ is differentiable let $y_k = \exp_{x_k} p_k \in \partial_c\psi_k(x_k)$, by compactness of $M$ we can extract a converging subsequence of $y_k$, we note $y$ the limit. The continuity of the exponential map on $TM$ gives $\exp_{\overline{x}} p = y$, and $y \in \partial_c\psi(\overline{x})$ indeed

$$\psi_k(x) - \psi_k(x_k) \geq c(x_k, y_k) - c(x, y_k)$$

so

$$\psi(x) - \psi(\overline{x}) \geq c(\overline{x}, y) - c(x, y).$$

Finally we shown that $\nabla^-\psi(\overline{x}) \subset \nabla_c^-\psi(\overline{x})$. Since $\nabla^-\psi(\overline{x})$ is convex it means that $\partial_c\psi(\overline{x})$ is

$c$-convex, this is our contradiction. □

To summarize a $c$-convex function not $c$-regular gives a non smooth transport problem and nearby the transport has to be non smooth. To finish we have to prove the following lemma.

**Lemma 1.5.6.** *Let $\psi$ a c-convex function, $p \in \nabla^- \psi \overline{x}$, and $\psi_k$ some c-convex function converging uniformly on $M$ to $\psi$. Then there exist $x_k \to x, p_k \in T_{x_k}M$ such that in any chart $p_k \to p$.*

This lemma says that a vector in a subgradient cannot appear from nowhere.

*Proof.* As we ask for a local result we work in a map around $x$, by adding a second order polynomial we can suppose that all $\psi_k$, $\psi$ are convex, $\psi(\overline{x}) = 0$ and $p = 0$. To make $\overline{x}$ a strict local minimum we define $\phi(x) = \psi(x) + |x - \overline{x}|^2$, and $\phi_k(x) = \psi_k(x) + |x - \overline{x}|^2$, of course $(\phi_k)$ is uniformly convergent to $\phi$. The point $\overline{x}$ is a strict local minimum point for $x$. Then, if we note $x_k$ a minimum point for $\phi_k$ we have $x_k \to \overline{x} = 0$, $0 \in \nabla^- \phi_k(x_k)$, so that there is $p_k \in \nabla^- \psi_k(x_k)$ such that $0 = p_k + 2x_k$. This $p_k$ suits. □

### 1.5.f Transport Continuity Property

We saw that the non-negativity of the Ma-Trudinger-Wang tensor is mandatory to obtain some regularity on the Kantorovich potential $\psi$. Ma, Trudinger, Wang and Liu on one side and Figalli, Kim and McCann on the other proved that the non-negativity is enough to define a regularity theory for the Monge-Ampère equation (5.3.4) in $\mathbb{R}^n$ [41, 43, 80, 81, 85, 105, 110]. We can, for example, give the following theorem.

**Theorem 1.5.7** (Optimal transport regularity theory)**.** *Let $X$ and $Y$ be the closures of bounded open sets in $\mathbb{R}^n$, and let $c : X \times Y \to R$ be a smooth cost function satisfying* **(Twist)** *such that $\nabla^2_{x,y}$ is non singular. Moreover we suppose that* **MTW** *holds true in the interior of $X \times Y$. Let $\Omega \Subset X$ and $\Lambda \Subset Y$ be $C^2$-smooth connected open sets and let $f \in L^1(\overline{\Omega})$, $g \in L^1(\overline{\Lambda})$ be positive probability densities. Let $\Psi$ be the Kantorovich potential associated to the optimal transport from $\mu = f dx$ to $\nu = g dy$, for the cost c. If:*

- *$\Lambda$ is uniformly c-convex with respect to $\Omega$ and $\Omega$ is uniformly $\widetilde{c}$-convex with respect to $\Lambda$;*
- *and $f$ is bounded away from zero and infinity on $\Omega$, $g$ is bounded away from zero and infinity on $\overline{\lambda}$;*

*then there exist $\overline{\alpha} > 0$ such that $\psi \in C^{1,\overline{\alpha}}(\Omega)$.*
*Moreover if:*

- *for $k \in \mathbb{N}$ and $\alpha \in (0,1)$, $f \in C^{k,\alpha}\left(\overline{\Omega}\right)$ and $g \in C^{k,\alpha}\left(\overline{\Lambda}\right)$;*
- *$\Lambda$ and $\Omega$ are of class $C^{k+2,\alpha}$;*

*then $\psi \in C^{k+2,\alpha}\left(\overline{\Omega}\right)$.*

Many contributions have been done to reach this result, the starting point is the work of Delanoë and Urbas [33, 106, 107] followed, in the case $k \geq 1$, by Ma, Trudinger, Wang and Trudinger, Wang [85, 105]. Then Loeper [82] obtained the Hölder continuity of the optimal transport map under the condition $MTW_s(0)$ with the exponent $\overline{\alpha} = \dfrac{1 - n/p}{4n - 2 + 1 - n/p}$, where $p \in ]n, +\infty]$ is such that $f \in L^p$. We also mention the paper of Loeper and Villani for the Riemannian version of this lemma [83]. Still under the condition $MTW_s(0)$, the sharp exponent ( $\overline{\alpha} = \dfrac{\beta(n+1)}{2n^2 + \beta(n-1)}$ with $\beta = 1 - \dfrac{n+1}{2p}$ ) for the Hölder continuity was given by Liu [80], his result also allows $p$ to be a little bit larger: $p \in ](n+1)/2, +\infty[$. The interior $C^{2,\alpha}$ regularity was obtained by Trudinger Wang and Liu [81]. Finally the Hölder continuity of optimal maps under $\mathbf{MTW}(0)$ is the most recent result, it has been obtained by Figalli Kim and McCann [41, 43].
The extension of this theorem in the Riemannian case needs two steps. First, one has to prove that the optimal transport map is continuous, or equivalently that the potential $\psi$ is $C^1$. Then taking smooth charts we want to apply the optimal transport regularity theory to obtain more regularity on $T$. The second step is not obvious since the cost has no reason to be smooth in charts when we get close to a focal point. To avoid this problem there are two strategies: either we prove that the optimal transport map $T$ stays far from the focal set, or we control the tensor near these points. The first strategy is the one used by Delanoë and Loeper in the case of the sphere for the quadratic geodesic case [36] and again by Loeper and Villani for a non focal Riemannian manifold [83]. This strategy is also used by Figalli Kim and McCann to prove the regularity of optimal maps for the quadratic geodesic distance on the product of round spheres [41], this result is the first one dealing with flat sectional curvature and non trivial cut locus.
On the other side, the first step leads to the definition of the transport continuity property (**TCP**) [48].

**Definition 1.5.8** (TCP)**.** *Let $M$ be a Riemannian manifold and $c$ a cost. We say that $M$ satisfies* (**TCP**) *if for any pair of probability measure on $M$ $(\mu, \nu)$ associated to smooth, positive densities with respect to the volume, the optimal transport map $T$ sending $\mu$ onto $\nu$ is continuous.*

Figalli, Rifford and Villani give necessary and sufficient conditions for a Riemannian manifold to satisfies (**TCP**) when the cost is given by the quadratic geodesic distance [48].

**Theorem 1.5.9** (Necessary conditions)**.** *Let $M$ be a smooth compact connected Riemannian manifold satisfying* (**TCP**) *for the quadratic geodesic cost. Then*

- *The condition* $\mathbf{MTW}(0)$ *holds true.*
- *All the injectivity domains of $M$ are convex.*

**Theorem 1.5.10** (Sufficient conditions)**.** *Let $M$ be a smooth compact connected Riemannian manifold such that*

- *The condition* $\mathbf{MTW}_s(0)$ *holds true,*
- *All the injectivity domains of $M$ are strictly convex.*

*Then $M$ satisfies* (**TCP**) *for the quadratic geodesic cost.*

The proofs of both theorems are based on the mother computation. We can notice that there is a gap between the necessary and the sufficient conditions. To fill it, one needs to relax the necessary conditions. Another way to simplify the theorem 1.5.10 is the contribution of the chapter 3. Indeed, we prove that in many cases, the condition $\mathbf{MTW}(0)$ (reps. $\mathbf{MTW}_s(0)$) implies the convexity (reps. strict convexity) of the injectivity domains.
Another way to attack the regularity issue in Riemannian space is trough small $C^4$ variations of the cost of known example like $\mathrm{S}^n$ with the quadratic geodesic cost. In this direction we can mention the work of Delanoë and Ge, Figalli and Rifford for a perturbation of $\mathrm{S}^2$ and Figalli, Rifford and Villani for perturbation of $\mathrm{S}^n$ [35, 34, 46, 49]. We finish this chapter with the computation of the Ma-Trudinger-Wang tensor in some particular cases.

## 1.6 Some particular Lagrangian costs

### 1.6.a Jacobi fields II

Let $M$ be a compact Riemannian manifold. We consider a cost $c$ given by a time independent Tonelli Lagrangian $L$ :

$$\begin{aligned} TM &\to \mathbb{R} \\ (x,v) &\mapsto L(x,v). \end{aligned}$$

Let $s,t \in [0,1]$, $(x,y_s) \in M \times M \subset c\text{-cut}(M)$ and $\xi, \eta \in T_xM$. A path of least action action for $L$ going from $\exp_x(t\xi)$ at time 0 to $c\text{-}\exp_x(v+s\eta) = y_s$ at time 1 is noted:

$$\begin{aligned} \gamma(\cdot,t,s) : \quad & [0,1] \to M \\ & u \mapsto \gamma(u,t,s). \end{aligned}$$

In particular we have $\gamma(0,t,s) = \exp_x(t\xi)$ and $\gamma(1,t,s) = y_s$. We recall the Euler-Lagrange equation that satisfies a path of least action for $L$.

**Lemma 1.6.1.** *Let $s,t$ be fixed. the function $u \mapsto \gamma(u,t,s)$ satisfies*

$$\frac{d}{du} L_v\left(\gamma(u,t,s), \frac{d}{du}\gamma(u,t,s)\right) = L_x\left(\gamma(u,t,s), \frac{d}{du}\gamma(u,t,s)\right).$$

We then define the $c$-Jacobi field $J_s : [0,1] \times T_xM$ above $\gamma(\cdot,0,s)$, it is a way to understand the variation of a path of least action, $\gamma(\cdot,0,s)$, in direction $\xi$. It is defined by

$$\begin{cases} J_s(u,\xi) = J(u,\xi,s) = \left.\dfrac{d}{dt}\right|_{t=0} \gamma(u,t,s), \qquad u \in [0,1] \\ J_s(0,\xi) = \xi, \\ J_s(1,\xi) = 0. \end{cases}$$

Here $s$ is fixed and is just a parameter useful for later. This is exactly the definition of Jacobi fields when $L = \frac{1}{2}|v|_x^2$. We can set a proposition in the spirit of proposition 1.4.7 to make the link with the MTW tensor.

**Proposition 1.6.2.** *Let $L$ such that $L_v = v$. Let $(x, y) \in M \times M \setminus c\text{-cut}(M)$ with $y = c\text{-}\exp_x v$, and $(\xi, \eta) \in T_xM \times T_xM$ then*

$$\mathfrak{S}_{x,y}(\xi, \eta) = \frac{3}{2}\frac{d^2}{ds^2}\bigg|_{s=0} \langle \dot{J}_s(0, \xi), \xi\rangle_x,$$

*where $y_s = \exp_x(v + s\eta)$.*

*Proof.* For the sake of simplicity we note $\gamma$ for $\gamma(u, t, s)$, $L_x$ the differential with respect to $x$ at $\left(\gamma(u,t,s), \frac{d}{du}\gamma(u,t,s)\right)$ and $L_v$ the one with respect to $v$ at $\left(\gamma(u,t,s), \frac{d}{du}\gamma(u,t,s)\right)$. Finally $\frac{D}{D\tau}$ stands for the covariant derivative along the path indexed by $\tau$, see Appendix A for more details. We compute $\frac{d}{dt}\int_0^1 L\left(\gamma, \frac{d}{du}\gamma\right) du$ using lemma 1.6.1:

$$\begin{aligned}
\frac{d}{dt}\int_0^1 L\left(\gamma, \frac{d}{du}\gamma\right) du &= \int_0^1 \frac{d}{dt}L\left(\gamma, \frac{d}{du}\gamma\right) du \\
&= \int_0^1 \left\langle L_x, \frac{d}{dt}\gamma\right\rangle_x + \left\langle L_v, \frac{D}{Dt}\frac{d}{du}\gamma\right\rangle_x du \\
&= \int_0^1 \left\langle L_x, \frac{d}{dt}\gamma\right\rangle_x + \frac{d}{du}\left\langle L_v, \frac{d}{dt}\gamma\right\rangle_x - \left\langle \frac{d}{du}L_v, \frac{d}{dt}\gamma\right\rangle_x du \\
&= \int_0^1 \frac{d}{du}\left\langle L_v, \frac{d}{dt}\gamma\right\rangle_x du \\
&= \left[\left\langle L_v, \frac{d}{dt}\gamma\right\rangle_x\right]_0^1 \\
&= -\left\langle L_v\left(\gamma(u,t,s), \frac{d}{du}\gamma(u,t,s)\right), \frac{d}{dt}\gamma(0,t,s)\right\rangle_x.
\end{aligned}$$

This is the formula of first variation. Since the path $t \mapsto \gamma(0, t, s)$ is geodesic we have

$$\frac{D^2}{Dt^2}\gamma(0, t, s) = 0,$$

therefore

$$\begin{aligned}
\frac{d^2}{dt^2}\int_0^1 L\left(\gamma, \frac{d}{du}\gamma\right) du\bigg|_{t=0} &= -\left\langle L_v, \frac{d^2}{dt^2}\gamma\right\rangle_x\bigg|_{u=0,t=0} - \left\langle \frac{d}{dt}L_v, \frac{d}{dt}\gamma\right\rangle_x\bigg|_{u=0,t=0} \\
&= -\left\langle \frac{d}{dt}L_v\bigg|_{u=t=0}, J(0,\xi,s)\right\rangle_x = -\left\langle \frac{d}{dt}L_v\bigg|_{u=t=0}, \xi\right\rangle_x \\
&= -\left\langle \frac{d}{du}\bigg|_{u=0} J(u,\xi,s), \xi\right\rangle_x = -\left\langle \dot{J}(0,\xi,s), \xi\right\rangle_x.
\end{aligned}$$

According to definition 1.4.3 the proposition is proved. □

In the case of $L_v$ just being invertible one has to be careful with this computation.

### 1.6.b The sectional curvature again

We reprove here the theorem 1.5.1 with a more flexible method. The goal consists in making clear the computation before going to more general Lagrangian.

*Second proof of theorem 1.5.1.* Following the notation above we consider the Jacobi field $J$ defined by

$$\begin{cases} J_s(u,\xi) = J(u,\xi,s) = \left.\dfrac{d}{dt}\right|_{t=0} \gamma(u,t,s), & u \in [0,1] \\ J_s(0,\xi) = \xi, & \\ J_s(1,\xi) = 0. & \end{cases}$$

By definition $J_s$ satisfies the Jacobi field equation above $\gamma_s = \gamma(\cdot, 0, s)$, see Appendix A for more details, that is

$$\ddot{J}_s(u) + R_s(u) J_s(u) = 0,$$

where $R$ is given by the Riemannian tensor. We develop the solution $J_s$ around $s = 0$ forgetting $\xi$. The prime stands for the $s$ derivative and the dot for the $u$ derivative, we omit to write the terms of order greater than 2. We obtain

$$J(u,s) = J(u,0) + sJ'(u,0) + \frac{s^2}{2} J''(u,0),$$
$$\dot{J}(u,s) = \dot{J}(u,0) + s\dot{J}'(u,0) + \frac{s^2}{2} \dot{J}''(u,0).$$

The initial conditions $J(0,s) = \xi$, $J(1,s) = 0$ gives

$$J'(0,0) = 0 = J''(0,0),$$
$$J'(1,0) = 0 = J''(1,0).$$

We also develop $R(u,s) = R_s(u)$:

$$R(u,s) = R(u,0) + sR'(u,0) + \frac{s^2}{2} R''(u,0).$$

Since $R(u,s) = s^2 R(1,1) = s^2 R(\eta,\eta)$ we get $R(u,0) = R'(u,0) = 0$, $R''(u,0) = 2R(1,1)$. Plugging this into the Jacobi field equation, and identifying the coefficient in front of $1, s, s^2$ we find

$$\begin{aligned}\ddot{J}(u,0) &= -R(u,0)J(u,0) = 0,\\ \ddot{J}'(u,0) &= -R'(u,0)J(u,0) - R(u,0)J'(u,0) = 0,\\ \ddot{J}''(u,0) &= -R''(u,0)J(u,0) - R(u,0)J''(u,0) - 2R'(u,0)J'(u,0)\\ &= -2R(1,1)J(u,0).\end{aligned}$$

We deduce

$$J(u,0) = J(0,0) + u(J(1,0) - J(0,0)) = \xi - u\xi$$

and

$$\ddot{J}''(u,0) = -2R(1,1)(\xi - u\xi).$$

Thus integrating from 0 to $u$ leads to

$$\dot{J}''(u,0) - \dot{J}''(0,0) = -2R(1,1)(u\xi - \frac{u^2}{2}\xi).$$

In particular with $u = 1$ we get

$$J''(1,0) - J''(0,0) - \dot{J}''(0,0) - \dot{J}''(0,0) = -2R(1,1)(\frac{1}{2}\xi - \frac{1}{6}\xi) = 2R(1,1)\frac{1}{3}\xi,$$

therefore

$$\dot{J}''(0,0) = -\frac{2}{3}R(1,1)\xi.$$

Plugging this into (1.6.2) we obtain the lemma. □

The advantage of this method is to avoid the renormalisation which can be a difficulty when dealing with a general Lagrangian.

### 1.6.c With $L = \frac{1}{2}|v|_x^2 + V(x)$

This case is treated by Lee and McCann in [76]. The approach is a bit different Lee and McCann compute directly the tensor, here I use a limited development of the Jacobi Field.Anyway both method gives the same results. In this case subsection we compute the value of $\dot{J}''(0,\xi,0)$ for a Lagrangian cost given by $L = \frac{1}{2}|v|_x^2 + V(x)$. The $c$-Jacobi field $J_s$ is defined by $J_s(u) = J(u,\xi,s) = \frac{d}{dt}\Big|_{t=0}\gamma(u,t,s)$, with $J(0,\xi,s) = \xi$ and $J(1,\xi,s) = 0$. By definition $J_s$ satisfies the $c$-Jacobi field equation:

$$\ddot{J}_s(u) + R(u,s)J_s(u) + \nabla^2 V(u,s)J_s(u) = 0$$

, with

$$\begin{aligned}J_s(0) &= \xi,\\ J_s(1) &= 0.\end{aligned}$$

As we done before we develop $J_s, R$ and $\nabla^2 V$ and omit to write down the term of order more than 2. Thus

$$J(u,s) = J(u,0) + sJ'(u,0) + \frac{s^2}{2}J''(u,0)$$
$$\dot{J}(u,s) = \dot{J}(u,0) + s\dot{J}'(u,0) + \frac{s^2}{2}\dot{J}''(u,0)$$

The initial conditions $J(0,s) = \xi$, $J(1,s) = 0$ implies

$$J'(0,0) = 0 = J''(0,0)$$
$$J'(1,0) = 0 = J''(1,0).$$

We denote $\nabla^2 V(u,s) = HessV(u,s) = M(u,s)$ then

$$M(u,s) = M(u,0) + sM'(u,0) + \frac{s^2}{2}M''(u,0),$$

where

$$M'(u,0) = \left\langle \nabla^2(V)', \gamma'(u,0) \right\rangle$$
$$M''(u,0) =< Hess(V)', \gamma''(u,0) > + < (Hess(V)'')\gamma'(u,0), \gamma'(u,0) > .$$

We do the same development for $R(u,s)$:

$$R(u,s) = R(u,0) + sR'(u,0) + \frac{s^2}{2}R''(u,0),$$

where

$$R(u,0) = R(\dot{\gamma}(u,0), e_i)\dot{\gamma}(u,0)$$
$$R'(u,0) = R(\dot{\gamma}'(u,0), e_i)\dot{\gamma}(u,0) + R(\dot{\gamma}(u,0), e_i)\dot{\gamma}'(u,0)$$
$$R''(u,0) = R(\dot{\gamma}''(u,0), e_i)\dot{\gamma}(u,0) + R(\dot{\gamma}(u,0), e_i)\dot{\gamma}''(u,0) + 2R(\dot{\gamma}'(u,0), e_i)\dot{\gamma}'(u,0).$$

Since $\dot{\gamma}(u,0) = 0$ we have $R(u,0) = R'(u,0) = 0$ and

$$R''(u,0) = 2R(\dot{\gamma}'(u,0), e_i)\dot{\gamma}'(u,0). \tag{1.6.1}$$

Therefore we first need to compute in smooth charts

$$\gamma(u,s) = \gamma(u,0) + s\gamma'(u,0) + \frac{s^2}{2}\gamma''(u,0).$$

By hypothesis $\gamma(u,0)=x$, $\gamma(0,s)=x$ and $\dot\gamma(0,s)=s\eta$ thus

$$\begin{aligned}\gamma'(0,s)&=\gamma''(0,s)=0\\ \dot\gamma'(0,s)&=\eta\end{aligned}$$

the Euler-Lagrange equation

$$\ddot\gamma(u,s)=-\nabla V(\gamma(u,s)),$$

implies

$$\begin{aligned}\ddot\gamma(u,0)&=-\nabla V(\gamma(u,0))=0\\ \ddot\gamma'(u,0)&=-HessV(x)\gamma'(u,0)=-M(x)\gamma'(u,0)\\ \ddot\gamma''(u,0)&=-HessV(x)\gamma''(u,0)-HessV'(x)\gamma'(u,0)\gamma'(u,0)\\ \ddot\gamma''(u,0)&=-M(x)\gamma''(u,0)-M'(u)\gamma'(u,0).\end{aligned}$$

We note $Z(t)=\begin{pmatrix}Z_1(t) & Z_2(t)\\ Z_3(t) & Z_4(t)\end{pmatrix}$ the solution in $\mathbb{R}^{2n}$ of the equation

$$\dot Z(t)=\begin{pmatrix}0 & Id\\ -M(x) & 0\end{pmatrix}Z(t).$$

Therefore we obtain

$$\begin{aligned}\gamma(u,0)&=x\\ \gamma'(u,0)&=Z_2(u)\eta\\ \gamma''(u,0)&=\Pi_1(-Z(u)\int_0^u Z^{-1}(v)\begin{pmatrix}0\\ M'(u)Z_2(v)\eta\end{pmatrix})du\\ \gamma''(u,0)&=:L_0^B(-M_V'(u)Z_2^2(v)\eta\eta)\end{aligned}$$

where $\Pi_1$ is the projection along the first $n$ coordinates, and $L_0^B$ the operator combining the integration and $\Pi_1$. Coming back to (1.6.1) we obtain

$$R''(u,0)=2\dot Z_2(u)R(\eta,e_i)\eta$$

Plugging everything into the $c$-Jacobi field equation and identifying the term of order $1,s,s^2$ we get

$$\begin{aligned}
\ddot{J}(u,0) &= -M(x)J(u,0)\\
\ddot{J}'(u,0) &= -M'(u,0)J(u,0) - M(u,0)J'(u,0) = 0\\
\ddot{J}''(u,0) &= -2\dot{Z}_2(u)R(1,1)J(u,0)\\
&\quad - M''(u,0)J(u,0) - M(u,0)J''(u,0) - 2M'(u,0)J'(u,0).
\end{aligned}$$

From the first equation we deduce

$$\begin{aligned}
J(u,0) &= Z_1(u)J(0,0) + Z_2(u)Z_2^{-1}(1)Z_1(1)J(0,0)\\
&= Z_1(u)\xi - Z_2(u)Z_2^{-1}(1)Z_1(1)\xi.
\end{aligned}$$

With the equation from order $s$ equation we find

$$J'(u,0) = Z_2(u)\dot{J}'(0,0) - \Pi_1(Z(u)\int_0^u Z^{-1}(u))\begin{pmatrix} 0 \\ M'(x)[Z_1(u)\xi - Z_2(u)Z_2^{-1}(1)Z_1(1)\xi] \end{pmatrix})du,$$

with

$$\dot{J}'(0,0) = Z_2^{-1}(1)\Pi_1(Z(u)\int_0^1 Z^{-1}(u))\begin{pmatrix} 0 \\ M'(x)[Z_1(u)\xi - Z_2(u)Z_2^{-1}(1)Z_1(1)\xi] \end{pmatrix})du.$$

Finally the third equation leads to

$$\begin{aligned}
&J''(u,0) =\\
&Z_2(u)\dot{J}''(0,0) - \Pi_1(Z(u)\int_0^u Z^{-1}(u)\begin{pmatrix} 0 \\ 2\dot{Z}_2(u)R(u,1)[Z_1(u)\xi - Z_2(u)Z_2^{-1}(1)Z_1(1)\xi] \end{pmatrix} du\\
&- \Pi_1(Z(u)\int_0^u Z^{-1}(u)\begin{pmatrix} 0 \\ M''(u)[Z_1(u)\xi - Z_2(u)Z_2^{-1}(1)Z_1(1)\xi] \end{pmatrix} du\\
&- \Pi_1(Z(u)\int_0^u Z^{-1}(u)\begin{pmatrix} 0 \\ 2M'(u)J'(u,0) \end{pmatrix} du.
\end{aligned}$$

and we deduce

$$\begin{aligned}
\dot{J}''(0,0) = Z_2^{-1}(1)[\Pi_1(Z(1)\int_0^1 Z^{-1}(u)\begin{pmatrix} 0 \\ 2\dot{Z}_2(u)R(u,1)[Z_1(u)\xi - Z_2(u)Z_2^{-1}(1)Z_1(1)\xi] \end{pmatrix} du\\
+ \Pi_1(Z(1)\int_0^1 Z^{-1}(u)\begin{pmatrix} 0 \\ M''(u)[Z_1(u)\xi - Z_2(u)Z_2^{-1}(1)Z_1(1)\xi] \end{pmatrix} du\\
+ \Pi_1(Z(1)\int_0^1 Z^{-1}(u)\begin{pmatrix} 0 \\ 2M'(u)J'(u,0) \end{pmatrix} du.
\end{aligned} \tag{1.6.2}$$

### 1.6.d Examples

**If $V = 0$**

When $V = 0$ we have

$$M(u) = 0 \qquad M'(u) = 0 \qquad M''(u) = 0 \tag{1.6.3}$$

$$Z_1(u) = I_n \quad Z_2(u) = uI_n. \tag{1.6.4}$$

Thus we find again

$$\begin{aligned} J(u,0) &= \xi - u\xi \\ J'(u,0) &= 0 \\ \dot{Z}_2(u) &= Z_1(u) \\ \dot{J}''(0,0) &= -\frac{2}{3}R(\eta,\xi)\eta \end{aligned}$$

**When $V$ is maximal:** $\nabla^2 V(x) = 0$

in this case we get

$$M(u) = 0 \tag{1.6.5}$$

$$R_1(u) = R_4(u) = I_n \quad R_2(u) = uI_n \quad R_3(u) = 0 \tag{1.6.6}$$

$$\gamma(u,0) = x \quad \gamma'(u,0) = u\eta \quad \gamma''(u,0) = -\frac{u^4}{12}\nabla^2 V'\eta\eta. \tag{1.6.7}$$

Therefore

$$\begin{aligned} J(u,0) &= \xi - u\xi \\ J'(u,0) &= [\frac{u}{12} - \frac{u^3}{6} + \frac{u^4}{12}]\nabla^2 V'\eta\xi \\ \dot{J}''(0,0) &= \frac{2}{3}R(\eta,\xi)\eta + \frac{1}{30}\nabla^2 V''\eta\eta\xi \\ &\quad - \frac{1}{1260}\nabla^2 V'(\nabla^2 V'(\eta,\eta),\xi) + \frac{17}{2520}\nabla^2 V'(\eta,\nabla^2 V'(\eta),\xi)), \end{aligned}$$

and finally for a point $x$ such that $V$ is maximal at $x$ we find

$$\begin{aligned} \mathfrak{S}_{x,x}(\xi,\eta) = \sigma(\xi,\eta) + \frac{1}{20}\nabla^2 V''(\eta,\eta,\xi,\xi) - \frac{1}{840}\nabla^2 V'(\nabla^2 V'(\eta,\eta),\xi)\xi \\ + \frac{17}{1680}\nabla^2 V'(\eta,\nabla^2 V'(\eta),\xi))\xi. \end{aligned} \tag{1.6.8}$$

# Chapter 2

# The tangent cut loci are Lipschitz continuous

## 2.1 Position of the problem

Let $(M, g)$ be a smooth compact Riemannian manifold of dimension $n \geq 2$. The injectivity domain at some point $x \in M$ is defined as

$$\mathrm{I}(x) = \Big\{ v \in T_x M \,|\, \exists\, t > 1 \text{ s.t. } d(x, \exp_x(tv)) = |tv|_x \Big\}, \tag{2.1.1}$$

or equivalently

$$\mathrm{I}(x) = \Big\{ v \in T_x M \,|\, \exists\, t > 1 \text{ s.t. } d^2(x, \exp_x(tv)) = |tv|_x^2 \Big\}, \tag{2.1.2}$$

where $\exp_x$ denotes the exponential mapping at $x$, $d$ the geodesic distance on $M \times M$, $|v|_x = \sqrt{g_x(v,v)} = \sqrt{\langle v, v\rangle_x}$ and $d^2$ the squared geodesic distance or quadratic geodesic distance (associated to the Lagrangian $|v|_x^2$). The injectivity domain $I(x)$ is an open star-shaped subset of $T_x M$; is boundary $\mathrm{TCL}(x)$, which is called the tangent cut locus at $x$, can be described thanks to a function $t_c$ defined on $UM \subset TM$: for any $(x, v) \in UM$ *i.e.* $(x, v) \in TM$ and $||v|| = 1$, we define $t_c$ by

$$\begin{aligned} t_c(x, v) &:= \sup \Big\{ t \geq 0 \,|\, tv \in \mathrm{I}(x) \Big\} & (2.1.3)\\ &= \max \Big\{ t \geq 0 \,|\, d^2(x, \exp_x(tv)) = |t|_x^2 \Big\}. & (2.1.4) \end{aligned}$$

Then, for every $x \in M$, there holds

$$\mathrm{I}(x) = \Big\{ tv \,|\, 0 \leq t < t_c(x, v),\ v \in U_x M \Big\} \quad \text{and} \quad \mathrm{TCL}(x) = \Big\{ t_c(x, v)v \,|\, v \in U_x M \Big\}. \tag{2.1.5}$$

We immediately see that $t_c$ is bounded from below by the injectivity radius of $M$ and bounded from above by the diameter of $M$.

We now define the nonfocal domain at some $x \in M$ as

$$\mathrm{NF}(x) = \Big\{ v \in T_xM \,|\, d_{tv} \exp_x \text{ is not singular for any } t \in [0,1] \Big\}. \tag{2.1.6}$$

It is an open star-shaped subset of $T_xM$ whose boundary $\mathrm{TFL}(x)$ is called the tangent focal domain at $x$ and can be described by the function $t_f$ defined on $UM \subset TM$ by

$$t_f(x,v) \;:=\; \sup\Big\{ s \geq 0 \,|\, sv \in \mathrm{NF}(x) \Big\}. \tag{2.1.7}$$

Then, for every $x \in M$, there holds

$$\mathrm{NF}(x) = \Big\{ tv \,|\, 0 \leq t < t_f(x,v),\; v \in U_xM \Big\} \quad \text{and} \quad \mathrm{TFL}(x) = \Big\{ t_f(x,v)v \,|\, v \in U_xM \Big\}. \tag{2.1.8}$$

Similarly, we define the boundary function : $t_b$: $UM \to \mathbb{R}^+$ for any subset of $TM$ with starshaped fibres. We then define the notion of $\kappa$–Lipschitz continuity for such a function.

**Proposition 2.1.1** ($\kappa$–Lipschitz continuity). *Let $O \subset TM$ be such that for any $x \in M$, the fibre $O_x$ is starshaped. The set $O$ is $\kappa$ Lipschitz continuous if for any $\overline{x}, \overline{v} \in UM$, there exists a $\kappa$–Lipschitz continuous function $\tau$ defined on a neighbourhood (in $UM$) of $(\overline{x}, \overline{v})$ such that $t_b(x,v) \leq \tau(x,v)$ and $t_b(\overline{x}, \overline{v}) = \tau(\overline{x}, \overline{v})$, where $t_b$ is the boundary function for $O$.*

This proposition means that if the boundary of $O$ is locally below a $\kappa$-Lipschitz continuous function then it is $\kappa$-Lipschitz continuous.

*Proof.* The proof is a straightforward contradiction argument. □

Our aim in this chapter is to prove the following theorem:

**Theorem 2.1.2** (Lipschitz continuity of the tangent cut loci).

1. *There exists $\kappa > 0$ such that for each $x \in M$ the set $I(x)$ is $\kappa$-Lipschitz continuous.*
2. *If $M$ is non-focal then there exists $\kappa > 0$ such that $\{(x,p) \,|x \in M,\, p \in I(x)\}$ is $\kappa$-Lipschitz continuous.*
3. *If $M$ has dimension 2 then there exists $\kappa > 0$ such that $\{(x,p) \,|x \in M,\, p \in I(x)\}$ is $\kappa$-Lipschitz continuous.*

To this purpose we first prove the two following theorems:

**Theorem 2.1.3** (Lipschitz continuity of the tangent focal loci). *There exist a $\kappa$ such that $\{(x,p) \,|x \in M,\, p \in \mathrm{NF}(x)\}$ is $\kappa$-Lipschitz continuous.*

**Theorem 2.1.4** (Semiconcavity of the tangent focal loci).
*The set $\{(x,p) \,|x \in M,\, p \in \mathrm{NF}(x)\}$ is semiconcave.*

The definition of semiconcavity is similar as the definition 2.1.1, where we ask $\tau$ to be semiconcave instead of Lipschitz continuous.

**Remark 2.1.5.** *The first item of theorem 2.1.2 is a result due to Li-Nirenberg, Itoh-Tanaka and Castelpietra-Rifford [64, 77, 30]. The second and third item of theorem 2.1.2 are new. It is the Lipschitz continuity with respect to the manifold in some directions. The proof is based on the idea given by Rifford and Castelpietra in [30].*

## 2.2 Preliminary results

Before we start the proof of theorem 2.1.3 and 2.1.2, we need to give some definitions and tools. We recall that $M$ is a compact Riemannian manifold.

### First variation formula

We prove here one of the most basic, however important formulae. We give it with the notation in a general setting; typically we ask $L$ to be a Tonelli Lagrangian. Let $L$ be a $C^2$ Lagrangian defined on $M$; for any absolute continuous path $\gamma$: $[0,1] \to M$, we define the action of $L$ along $\gamma$ by

$$A(\gamma) = \int_0^1 L(\gamma(t), \dot{\gamma}(t))dt,$$

and the cost $c_L$ as

$$c_L(x,y) = \inf\Big\{A(\gamma)|\gamma \in AC_x^y\Big\}.$$

Here $AC_x^y$ denotes the set of absolutely continuous paths $\gamma$: $[0,1] \to M$ with $\gamma(0) = x$ and $\gamma(1) = y$.

**Lemma 2.2.1.** *Let $(s,t) \mapsto \gamma(s,t)$ be a function in $C^1([0,1]^2, M)$ such that $\gamma(0,t) = \gamma_0(t)$ is a geodesic path; then, denoting $\frac{d}{ds}|_{s=0}\gamma(.,t) = h(t)$ and $\frac{d}{dt}|_{t=t}\gamma(0,.) = \dot{\gamma}(t)$ we have:*

$$\frac{d}{ds}|_{s=0}A(\gamma(s,.)) = (\nabla_v L)(\gamma(1), \dot{\gamma}(0)) \cdot h(1) - (\nabla_v L)(\gamma(0), \dot{\gamma}(0)) \cdot h(0) \tag{2.2.1}$$

*Moreover we suppose that for all $s \in [0,1]$, $\gamma(s,0) = \gamma_0(0)$, then there exists $K_f > 0$ such that*

$$|A(\gamma(s,.)) - A(\gamma(0,.))| \leq s|(\nabla_v L)(\gamma_0(1), \dot{\gamma}_0(0)) \cdot h(1)| + \frac{s^2}{2}K_f. \tag{2.2.2}$$

*The constant $K_f$ depends only on the $C^2$ norm of $L$ and on a compact subset $K \in TM$ such that $(\gamma(s,t), \dot{\gamma}(s,t))$, $(\frac{D}{Ds}\gamma(s,t), \frac{D}{Ds}\dot{\gamma}(s,t))$ and $(\frac{D^2}{Ds^2}\gamma(s,t), \frac{D^2}{Ds^2}\dot{\gamma}(s,t))$ are in $K$.*

Hereafter the notation $\frac{D}{Du}$ stands for the covariant derivatives along the path $\gamma_0(u)$.

*Proof.* We first remark that

$$\frac{d}{ds}A(\gamma(s,.)) = \int_0^1 \frac{d}{ds}L(\gamma(s,t),\dot\gamma(s,t))dt$$
$$= \int_0^1 (\nabla_v L)(\gamma(s,t),\dot\gamma(s,t)) \cdot \frac{D}{Ds}\frac{d}{dt}\gamma(s,t) + (\nabla_x L)(\gamma(s,t),\dot\gamma(s,t)) \cdot \frac{d}{ds}\gamma(s,t)dt.$$

Since $\frac{D}{Ds}\frac{d}{dt}\gamma(s,t) = \frac{D}{Dt}\frac{d}{ds}\gamma(s,t)$, we integrate by parts the first term and get:

$$\frac{d}{ds}A(\gamma(s,.)) = [(\nabla_v L)(\gamma(s,t),\dot\gamma(s,t)) \cdot \frac{d}{ds}\gamma(s,t)]_{t=0}^1 \tag{2.2.3}$$
$$+\int_0^1 [(\nabla_x L)(\gamma(s,t),\dot\gamma(s,t))\frac{d}{ds}\gamma(s,t) - \frac{d}{dt}(\nabla_v L)(\gamma(s,t),\dot\gamma(s,t))] \cdot \frac{d}{ds}\gamma(s,t)dt.$$

Since $\gamma(0,t)$ is a geodesic path, it satisfies the Euler equation. Therefore taking $s=0$ in (2.2.3) we obtain (2.2.1).

In order to prove (2.2.2), let us denote $A(\gamma(s,.)) = f(s)$. Since $\gamma$ is $C^2$ with respect to $s$, and $L$ is smooth, a Taylor formula gives $|f(s)-f(0)| \le s|f'(0)| + \frac{s^2}{2}\sup_{s\in[0,1]}|f''(s)|$. Moreover the equality $\frac{d}{ds}\gamma(s,0) = h(0) = 0$ and (2.2.1) implies that $f'(0) = (\nabla_v L)(\gamma(1),\dot\gamma(0)) \cdot h(1)$. Let us then compute $f''$. One has

$$f''(s) = \int_0^1 \frac{d^2}{ds^2}L(\gamma(s,t),\dot\gamma(s,t))dt$$

and then

$$f''(s) = \int_0^1 (\nabla^2_{vv}L)(\gamma(s,t),\dot\gamma(s,t)) \cdot (\frac{D}{Ds}\frac{d}{dt}\gamma(s,t), \frac{D}{Ds}\frac{d}{dt}\gamma(s,t))$$
$$+ (\nabla^2_{xx}L)(\gamma(s,t),\dot\gamma(s,t)) \cdot (\frac{d}{ds}\gamma(s,t), \frac{d}{ds}\gamma(s,t))$$
$$+ 2(\nabla^2_{xv}L)(\gamma(s,t),\dot\gamma(s,t)) \cdot (\frac{D}{Ds}\frac{d}{dt}\gamma(s,t), \frac{D}{Ds}\gamma(s,t))dt$$
$$+ \int_0^1 (\nabla_v L)(\gamma(s,t),\dot\gamma(s,t)) \cdot \frac{D^2}{Ds^2}\frac{d}{dt}\gamma(s,t)$$
$$+ (\nabla_x L)(\gamma(s,t),\dot\gamma(s,t)) \cdot \frac{D}{Ds}\frac{d}{ds}\gamma(s,t)dt.$$

The inequality (2.2.2) follows. □

## Continuity

**Lemma 2.2.2** (Continuity of $t_c$ and $t_f$)**.** *Let $M$ be a compact Riemannian manifold, the functions $t_c$ and $t_f$ are continuous on $TM$.*

**Lemma 2.2.3.** *For $M$ a compact Riemannian manifold we have for all $x \in M$:*

$$\mathrm{I}(x) \subset \mathrm{NF}(x).$$

We refer to [53, 98] [110] [30] for a proof of these statements. In the first reference the proof is performed in the Lagrangian world using the index (second variation formula). In the second and third references the point of view is respectively Hamiltonian and Lagrangian, but both use the same idea. Before the cut locus they take advantage of the regularity of the distance function to construct a Lipschitz continuous inverse to the exponential map. This prevents the singularity of the differential of the exponential map.
Let us now identify the purely focal points in TFCL = TFL ∩ TCL. For any $v \in U_xM$ we define $\delta(v)$ by:

$$\delta(v) = \Big\{ \max |v - w|, \text{ for all } w \in U_xM, \text{ such that } \exp_x(t_c(x,v)v) = \exp_x(t_c(x,w)w) \Big\}.$$

The function $\delta(\cdot)$ is equal to zero if and only if $t_c(x,v)v \in \mathrm{TCL}(x) \cap \mathrm{TFL}(x)$. Moreover in this case $s \to \exp_x(st_c(x,v)v)$ is the unique (energy) minimizing curve. This set is the purely focal one. The purely locus set is when $t_c < t_f$.
The next section is devoted to the proof of Theorem 2.1.3.

## 2.3 Proof of Theorem 2.1.3: Lipschitz continuity of the tangent focal loci

We split this section in three parts; in a first part we give some basics on Jacobi fields and focalisation; in a second part we give a heuristic proof of theorem 2.1.3, working for example in the case of a sphere. The third part is a rigorous proof using the Hamiltonian structure hidden in the Jacobi field equation. This proof is based on the one given in the paper of Castelpietra and Rifford [30], the main difference is that we adopt here a Lagrangian point of view whereas Castelpietra and Rifford used an Hamiltonian point of view.

### 2.3.a Focalization and Jacobi fields

For the proof of Theorem 2.1.3, we need to consider the Jacobi fields; we refer to chapter one and [53, 98] for more details. Let $(x, v) \in TM$, we consider the geodesic path $\gamma_0 : t \in \mathbb{R}^+ \mapsto \exp_x(tv)$. We choose an orthonormal basis of $T_xM$: $B = (v, e_2, ..., e_i, ..., e_n)$ and define by parallel transport an orthonormal basis of $T_{\exp_x(tv)}M$: $B(t) = (e_1(t), e_2(t), ..., e_i(t), ..., e_n(t))$. We identify $T_{\exp_x(tv)}M$ with $\mathbb{R}^n$ thanks to the basis $B(t)$. By definition the Jacobi field equation along $\gamma_0$ is given by:

$$\begin{aligned} &\ddot{J}(t) + R(t)J(t) = 0, \quad t \in \mathbb{R}^+, \\ &J(0) = h, \quad h \in T_xM, \\ &\dot{J}(0) = q, \quad p \in T_xM, \end{aligned} \tag{2.3.1}$$

where $R$ is symmetric, given by the the Riemannian tensor: in the basis $B(t)$, $R_{ij} = \langle R(e_i, e_j)e_i, e_j\rangle$. It describes how a small perturbation of the geodesic path evolves along the geodesic path. Since a focal point is related to the size of the neighbourhood one can visit by perturbing the geodesic path, we understand that both notions are linked. The Jacobi field equation (2.3.1) is a linear equation of order two, we therefore define, $J_0^1 : t \mapsto M_n(\mathbb{R})$ as the solution of the following matrix Jacobi field equation,

$$\begin{aligned}
&\ddot{J}(t) + R(t)J(t) = 0, \quad t \in \mathbb{R}^+,\\
&J(0) = I_n,\\
&\dot{J}(0) = 0.
\end{aligned}$$

We similarly define $J_1^0$ as the solution of:

$$\begin{aligned}
&\ddot{J}(t) + R(t)J(t) = 0, \quad t \in \mathbb{R}^+,\\
&J(0) = 0,\\
&\dot{J}(0) = I_n.
\end{aligned}$$

Any solution $J$ of the Jacobi field equation (2.3.1) can be written for any $t \in \mathbb{R}^+$

$$J(t) = J_0^1(t)J(0) + J_1^0(t)\dot{J}(0). \tag{2.3.2}$$

Let us now exhibit two very particular families of Jacobi fields. For any $h \in T_xM$ we define the path

$$\gamma_\alpha(s,t) = \exp_{\exp_x(sh)}(tv), \qquad (s,t) \in [0,1] \times \mathbb{R}^+, \tag{2.3.3}$$

$$\gamma_\beta(s,t) = \exp_x(t(v+sh)), \qquad (s,t) \in [0,1] \times \mathbb{R}^+. \tag{2.3.4}$$

It leads to the following families of Jacobi fields

$$J_\alpha(t) = \left.\frac{d}{ds}\right|_{s=0} \gamma_\alpha(s,t) = (d_{x=x}\exp_{\cdot}(tv)) \cdot (h), \tag{2.3.5}$$

$$J_\beta(t) = \left.\frac{d}{ds}\right|_{s=0} \gamma_\beta(s,t) = (d_{p=tv}\exp_x) \cdot (th). \tag{2.3.6}$$

The Jacobi field $J_\beta$ is nothing but $J_1^0(\cdot)h$; indeed $J_\beta(0) = 0$ and $\dot{J}_\beta(0) = h$. The Jacobi field $J_\alpha$ is exactly $J_0^1(\cdot)h$: $J_\alpha(0) = h$ and $\dot{J}_\alpha(0) = 0$. The link with focalization is enclosed in the following lemma.

**Lemma 2.3.1.** *Let $(x,v) \in U_xM$ then*

$$t_f(x,v) = \inf\left\{t \in \mathbb{R}^+, \quad \exists q \in U_xM \text{ with } J_1^0(t)q = 0.\right\} \tag{2.3.7}$$

*The direction $q$ is called a focal direction at $(x,v)$.*

*Proof.* The proof is a direct consequence of (2.3.6): for any $t > 0$, $J_1^0(t)h = (d_{p=tv}\exp_x)\cdot(th)$. □

### 2.3.b Heuristic proof

We use here Lemma 2.3.1 to give a heuristic proof of Theorem 2.1.3, assuming that $J_0^1\left(t_f(x,v)\right))$ is invertible.

**Lemma 2.3.2.** *The matrix $S(t) = \left(J_0^1\right)^{-1}(t)J_1^0(t)$ is well defined and symmetric on a neighbourhood of $(x,v,t_f(x,v))$.*

Remark that $J_0^1(t)$ and $J_1^0(t)$ depend smoothly on $(x,v)$.

*Proof.* The matrix is well defined since the set of invertible matrices is open. Moreover for any $t>0$ the matrix $S(t)$ can be seen as the operator $q \in T_xM \mapsto -J_q(0)$, where $J_q$ is the unique Jacobi field such that $\dot{J}_q(0) = q$ and $J_q(t) = 0$. To check the symmetry of $S(t)$, we remark that for any $q, q' \in U_xM$:

$$\langle S(t)q, q'\rangle - \langle q, S(t)q'\rangle = -\left\langle J_q(0), \dot{J}_{q'}(0)\right\rangle + \left\langle \dot{J}_q(0), J_{q'}(0)\right\rangle.$$

Let $f(\tau) = -\left\langle J_q(\tau), \dot{J}_{q'}(\tau)\right\rangle + \left\langle \dot{J}_q(\tau), J_{q'}(\tau)\right\rangle$ we have

$$\begin{aligned}\frac{d}{d\tau}f &= -\left\langle J_q(\tau), \ddot{J}_{q'}(\tau)\right\rangle + \left\langle \ddot{J}_q(\tau), J_{q'}(\tau)\right\rangle\\ &= -\left\langle J_q(\tau), -R(\tau)J_{q'}(\tau)\right\rangle + \left\langle -R(\tau)J_q(\tau), J_{q'}(\tau)\right\rangle\\ &= 0.\end{aligned}$$

Thus $f(0) = f(t) = -\left\langle 0, \dot{J}_{q'}(t)\right\rangle + \left\langle \dot{J}_q(t), 0\right\rangle = 0$. It proves the lemma. □

The idea to prove Theorem 2.1.3 is to apply the implicit function theorem to $q^tS(t)q$. To this purpose we need to show in particular that $q^t\dot{S}(t_f)q = q^t\frac{d}{dt}S(t_f)q \neq 0$.

**Lemma 2.3.3.** *Let $(x,v) \in UM$ and let $q$ be a focal direction as defined in Lemma 2.3.1 ($q \in \operatorname{Ker}\left(J_1^0(t_f(x,v))\right)$) then*

1. *the quantity $q^t\,\dot{S}\left(t_f(x,v)\right)\,q$ is not equal to zero;*
2. *for any $0 < t < t_f(x,v)$ the matrix $K(t) = \left(J_1^0\right)^{-1}(t)J_0^1(t)$ is symmetric decreasing (all eigenvalues are decreasing);*
3. *there exists a neighbourhood of $(x,v,t_f(x,v)$ in $UM\times\mathbb{R}^+$: $O_{x,v}$, such that for any $(x',v',t) \in 0_{x,v}$ if $q^t\,S\,(t)\;q = 0$ then $t \geq t_f(x',v')$.*

*Proof.* A simple computation gives that

$$\begin{aligned}q^t\,\dot{S}\left(t_f(x,v)\right)\,q = &-q^t\left(J_0^1\right)^{-1}\dot{J}_0^1\left(J_0^1\right)^{-1}\left(t_f(x,v)\right)J_1^0\left(t_f(x,v)\right)q\\ &- q^t\left(J_0^1\right)^{-1}\left(t_f(x,v)\right)\dot{J}_1^0\left(t_f(x,v)\right)q.\end{aligned}$$

Thanks to the identity

$$\left(J_0^1\right)^t\,\dot{J}_1^0 - \left(\dot{J}_0^1\right)^t J_1^0 = I_n, \tag{2.3.8}$$

we find that

$$q^t \dot{S}\left(t_f(x,v)\right) q = -\left(\left(J_0^1\right)^{-1} q^t\right)^t \left(J_0^1\right)^{-1} q = -|\left(J_0^1\right)^{-1} q|_y^2.$$

Therefore, by compactness, we find $\delta$ such that $q^t \dot{S}\left(t_f(x,v)\right) q \leq -\delta$.

To prove the identity (2.3.8) one can check that the derivative along the field is equal to zero and that the value at time $t = 0$ is equal to $I_n$.

For $t \leq t_f(x,v)$, the matrix $K(t)$ is well defined; a similar computation as the one done before for $S$ proves that $K$ is symmetric, and for any $h \in T_xM$, $h^t\, \dot{K}(t)\, h < 0$.

For the third item we proceed by contradiction. We suppose that $t < t_f(x',v')$ and $q^t\, S\,(t)\; q = 0$. Then For $t < t_f(x,v)$, large enough such that $S(t)$ exists, we have that $K(t)$ is invertible and $K^{-1}(t) = S(t)$. therefore for any $q \in \operatorname{Ker}\left(J_1^0(t_f(x,v))\right)$ we have

$$q^t\, S\,(t)\; q = q^t\, K^{-1}(t)K(t)K^{-1}(t)\, q = h_t^t\, K(t)\, h_t,$$

where $h_t = K^{-1}(t)q$ or equivalently $q = K(t)h_t$. Without lost of generality we suppose that $K(t)$ is diagonal and we denote its eigenvalues by $(\lambda_i)_{i\in[1,n]}$. Then passing to the limit when $t$ goes to $t_f x, v$ on one side, we find $h(t_f(x,v)) = 0$, and the other side we find for any $i \in [1,n]$:

$$q_i = \lim_{t \to t_f(x,v)} \lambda_i(t) h_i(t).$$

It implies that $\lim_{t \to t_f(x,v)} |\lambda_i(t)| = +\infty$. Since $t \mapsto \lambda_i(t)$ is decreasing, we get that

$$\lim_{t \to t_f(x,v)} \lambda_i(t) = -\infty.$$

□

To conclude the proof we construct the function $\tau$ needed in definition 2.1.1 thanks to the implicit function theorem. Let $(\overline{x},\overline{v}) \in UM$ and $q \in U_xM$ be the focal direction associated then the function

$$\begin{aligned} \Psi: \quad & UM \times \mathbb{R}^+ \to \mathbb{R} \\ & (x,v,t) \mapsto q^t S\,(t)\, q \end{aligned}$$

is well defined on a neighbourhood of $(\overline{x},\overline{v},t_f(\overline{x},\overline{v}))$. Moreover $\Psi(\overline{x},\overline{v},t_f(\overline{x},\overline{v})) = 0$ and

$$|\partial_t \Psi(\overline{x},\overline{v},t_f(\overline{x},\overline{v}))| = \left| q^t \dot{S}\left(t_f(x,v)\right) q \right| \geq \delta,$$

therefore we can apply the implicit function theorem to get a function $\tau$ defined on a neighbourhood $O_{\overline{x},\overline{v}}$ of $(\overline{x},\overline{v})$ such that $\Psi(x,v,\tau(x,v)) = 0$. By Lemma 2.3.3, we find that $t_f(x,v) \leq \tau(x,v)$, there only remains to check that $\tau$ is Lipschitz continuous. By compactness, there exists $K > 0$

such that

$$\begin{aligned}|d_{\overline{x},\overline{v}}\tau| &= \left|\frac{1}{\partial_t\Psi(\overline{x},\overline{v},t_f(\overline{x},\overline{v}))}d_{\overline{x},\overline{v}}\Psi\right| \\ &\leq \frac{K}{\delta}.\end{aligned}$$

This concludes the heuristic proof.

**Remark 2.3.4.** *The main restriction of this proof is that in general $J_0^1\left(t_f(x,v)\right)$ has no reason to be invertible. Moreover, as we shall see in the rigorous proof, the main ingredients come from the hidden symplectic structure.*
*This proof works in the case of the round metric on the unit sphere, as well as small $C^4$ perturbations of this metric.*

### 2.3.c Rigorous proof

To understand this proof well, it can be interesting for the reader to compare each step with the heuristic proof. We start with some remarks on the symplectic structure coming with a Riemannian manifold.

**Definition 2.3.5** (The symplectic form)**.** *Let $M$ be a Riemannian manifold of dimension $n$, for any $x \in M$ we define the symplectic form $\sigma$:*

$$\begin{aligned}\sigma:\quad &(T_xM\times T_xM)^2 \to \mathbb{R},\\ &(h,q),(h',q') \mapsto \langle h,q'\rangle - \langle h',q\rangle = (h,q)^t\,\mathbb{J}\,(h',q').\end{aligned}$$

*The matrix* $\mathbb{J} = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$.

**Definition 2.3.6** (Lagrangian subspace.)**.** *A subspace $L \in T_xM\times T_xM$ is said to be Lagrangian if* $\dim(L) = n$ *and* $\sigma|_{L\times L}$ *is equal to* $0$.

For example the vertical subspace $\{0\}\times T_xM \in T_xM\times T_xM$ and the horizontal subspace $T_xM\times\{0\} \in T_xM\times T_xM$ are Lagrangian. The matrix $J_1^0$ and $J_0^1$ are the fundamental solution of the Jacobi field equation (2.3.1) on those subspaces.

**Lemma 2.3.7.** *Let $L$ be a Lagrangian subspace and $E, F$ be two vectorial spaces of dimension $n$ such that $E \overset{\perp}{\oplus} F = T_xM\times T_xM$, moreover suppose that $L\cap E\times\{0\} = \{0\}$ then there exist a symmetric matrix $S$ such that*

$$L = \left\{(Sq,q)_{E,F}\,,\quad q\in F\right\}.$$

*We say that $L$ is a graph above $F$.*

*Proof.* The matrix $S$ exists since $L$ has dimension $n$ and no direction in $E$. To see that $S$ is symmetric we look at the symplectic form on two vector of $L$: let $q, q' \in F$ by definition

$$\begin{aligned} 0 &= \sigma\left((Sq, q), (Sq', q')\right) \\ &= \langle Sq, q' \rangle - \langle Sq', q \rangle \\ &= \langle Sq, q' \rangle - \langle q, Sq' \rangle . \end{aligned}$$

□

**Remark 2.3.8.** *That is exactly the method we used to prove lemma 2.3.2.*

An important link between the symplectic form and the Jacobi field is that the symplectic form is preserved along the flow of the Jacobi field equation.

**Lemma 2.3.9.** *Let $J_1$ and $J_2$ be two solution of the Jacobi field equation* (2.3.1) *then for any $t > 0$*

$$\sigma\left(\left(J_1(t), \dot{J}_1(t)\right), \left(J_2(t), \dot{J}_2(t)\right)\right) = \sigma\left(\left(J_1(0), \dot{J}_1(0)\right), \left(J_2(0), \dot{J}_2(0)\right)\right).$$

*Equivalently defining $M(t) = \begin{pmatrix} J_1(t) & \dot{J}_1(t) \\ J_2(t) & \dot{J}_2(t) \end{pmatrix}$, we have $M^t(t)\mathbb{J}M(t) = \mathbb{J}$, we say that $M(t)$ is symplectic.*

*Proof.* The proof is a simple computation, let

$$f(t) = \sigma\left(\left(J_1(t), \dot{J}_1(t)\right), \left(J_2(t), \dot{J}_2(t)\right)\right).$$

Since $R$ is symmetric we have

$$\begin{aligned} \dot{f}(t) &= \left\langle J_1(t), \ddot{J}_2(t) \right\rangle - \left\langle J_2(t), \ddot{J}_1(t) \right\rangle \\ &= \langle J_1(t), -R(t)J_2(t) \rangle - \langle J_2(t), -R(t)J_1(t) \rangle \\ &= 0 \end{aligned}$$

□

**Remark 2.3.10.** *The equality $M^t(t)\mathbb{J}M(t) = \mathbb{J}$ implies the identity* (2.3.8)*, it explains why the proof of this identity is exactly the one we have just done.*

We now define a particular Lagrangian subspace in order to find a new formulation for $t_f$.

**Definition 2.3.11.** *Let $(x, v) \in UM$ we define $L_{t,v}$ by*

$$L_{t,v} = \left\{(h, q) \in T_xM \times T_xM | J_0^1(t)h + J_1^0(t)q = 0\right\}.$$

*Equivalently with $M(t) = \begin{pmatrix} J_1(t) & \dot{J}_1(t) \\ J_2(t) & \dot{J}_2(t) \end{pmatrix}$ and*

$$V_{t,v} = \{0\} \times T_{exp_x(tv)M} \in T_{exp_x(tv)M} \times T_{exp_x(tv)M},$$

*the vertical subspace at* $\exp_x(tv)$ *we have*

$$L_{t,v} = M^{-1}(t)V_{t,v}.$$

*The space* $L_{t,v}$ *is the set of initial conditions such that at time* $t$ *the Jacobi field is equal to* $0$.

**Proposition 2.3.12.** *The space* $L_{t,v}$ *is a Lagrangian subspace of* $T_xM \times T_xM$.

*Proof.* Since $M^t(t)\mathbb{J}M(t) = \mathbb{J}$ the matrix $M(t)$ is invertible therefore $L_{t,v}$ is a vectorial subspace of dimension $n$. To see that it is Lagrangian we used that $\sigma$ is preserved along the flow. Let $(h,q)$ and $(h',q')$ in $L_{t,v}$, we denote by $J_{h,q}$ the solution of the Jacobi field equation (2.3.1) with $J_{h,q}(0) = h$, $\dot{J}_{h,q}(0) = q$, then for any $u > 0$

$$\begin{aligned}\sigma\left((h,q),(h',q')\right) &= \sigma\Big((J_{h,q}(t),\dot{J}_{h,q}(t)),(J_{h',q'}(t),\dot{J}_{h',q'}(t))\Big)\\ &= \sigma\Big((0,\dot{J}_{h,q}(t)),(0,\dot{J}_{h',q'}(t))\Big) = 0.\end{aligned}$$

□

We thus can give a new formulation of lemma 2.3.1.

**Lemma 2.3.13.** *Let* $(x,v) \in U_xM$ *then*

$$t_f(x,v) = \inf\left\{t \in \mathbb{R}^+ | L_{t,v} \cap V_{0,v} \neq \{0\}\right\}. \tag{2.3.9}$$

*The set* $L_{t,v} \cap V_{0,v}$ *is called the focal set at* $(x,v)$.

*Proof.* Let $q \in U_xM$, $q \neq 0$ such that $(0,q) \in L_{t,v} \cap V_{0,v}$ then $J_{0,q}(t) = J_1^0(t)q = 0$ therefore lemma 2.3.1 concludes the proof. □

**Remark 2.3.14.** *In the heuristic proof the hypothesis "*$J_0^1(t_f(x,v))$ *invertible" exactly says that* $L_{t_f(x,v),v}$ *is a graph above* $V_{0,v}$, *given by the application*

$$S(t_f(x,v)) = \left(J_0^1(t_f(x,v))\right)^{-1} J_1^0(t_f(x,v)).$$

*In the general case* $J_0^1(t_f(x,v))$ *has no reason to be invertible, to adapt the proof we need to find another way to write* $L_{t_f(x,v),v}$ *as a graph.*

We recall that we identify $T_{exp_x(tv)}M$ with $\mathbb{R}^n$ through the basis

$$B(t) = (e_1(t), ..., e_i(t), ..., e_n(t)).$$

According to lemma 2.3.7 the obstruction to see $L_{t_f(x,v),v}$ as a graph above $V_{0,v}$ comes from the intersection of $L_{t_f(x,v),v}$ with the horizontal space. By definition we have

$$L_{t_f(x,v),v} \cap H_{0,v} = \mathrm{Ker} J_0^1(t_f(x,v)).$$

We identify, for any $u \geq 0$, $H_{u,v}$ with Vect $\left(e'_1(u), ..., e'_i(u), ..., e'_n(u)\right)$ and $V_{u,v}$ with Vect $(f_1(u), ..., f_i(u), ..., f_n(u))$, where $e'_i(u) = e_i(u) \times \{0\} \in T_{\exp_x(uv)} \times T_{\exp_x(uv)}$ and $f_i(u) = \{0\} \times e_i(u) \in T_{\exp_x(uv)} \times T_{\exp_x(uv)}$. With these notation, without lost of generality, we can suppose we have $l > 1$ such that $\mathrm{Ker} J_0^1(t_f(x,v)) = $ Vect $\left(e'_l, ...e'_n\right)$. Therefore we change, for any $i \geq l$, $e'_i(u)$ by $f_i(u)$ and $f_i(u)$ by $-e'_i(u)$ to get two new orthonormal spaces of dimension $n$:

$$E(u) = \text{Vect}\left(e'_1(u), ..., e'_{l-1}(u), f_l(u), ..., f_n(u)\right)$$
$$F(u) = \text{Vect}\left(f_1(u), ..., f_{l-1}(u), -e'_l(u), ..., -e'_n(u)\right).$$

**Remark 2.3.15.** *The change of coordinates is symplectic, that is $p^t \mathbb{J} P = J$, where $P$ is the change of basis matrix. Therefore for any $(z,w), (z',w') \in E \times F$ we have*

$$\sigma\left((z,w),(z',w')\right) = \langle z, w' \rangle - \langle z', w \rangle.$$

By construction for any $u \geq 0$ we have

1. $E(u) \overset{\perp}{\oplus} F(u) = T_{\exp_x(uv)} \times T_{\exp_x(uv)}$
2. $L_{t_f(x,v),v} \cap E(0) = \{0\}$.

Since $L_{u,v'}$ is smooth with respect to $(x', v', u)$, there exist a neighbourhood of $(x, v, t_f(x,v))$: $O_{x,v,t_f(x,v)} \subset TM \times \mathbb{R}^+$ such that for any $\left(x', v', t\right) \in O_{x,v,t_f(x,v)}$ we have

$$L_{t,v'} \cap E(0) = \{0\}. \tag{2.3.10}$$

Moreover lemma 2.3.7 implies that there exist a smooth function

$$\begin{aligned} S: \quad & O_{x,v,t_f(x,v)} \to S_n(\mathbb{R}) \\ & \left(x', v', t\right) \mapsto S(t), \end{aligned}$$

such that for any $w \in F(0)$, $S(t)w \in E(0)$ and

$$L_{t,v'} = \left\{(S(t)w, w)_{E(0)\times F(0)} \text{ with } w \in F(0)\right\}.$$

**Remark 2.3.16.** *The matrix $S(t)$ depends on $(x', v', t)$, The subspaces $E(u)$ and $F(u)$ also depend on $(x', v', t)$, but the indices $l$ used to define $E(u)$ and $F(u)$ for any $(x', v', t) \in O_{x,v,t_f(x,v)}$ only depends on $x, v, t_f(x,v)$.*

The following lemma is the key tool to apply later the theorem of implicit function.

**Lemma 2.3.17.** *let $(x, v) \in TM$,*

1. *Let $q \in U_xM$ such that $(\{0\}, q) \in L_{t_f(x,v),v} \cap V_{0,v}$ then $q \in F(0)$ and $q^t\, S(t_f(x,v))q = 0$.*
2. *There exists $\delta > 0$ such that for any $(x, v) \in TM$, $||\dot{S}(t_f(v))|| \geq \delta$.*
3. *Moreover for any $(x', v', t) \in O_{x,v,t_f(x,v)}$ if $q^t\, S(t)q = 0$ then $t_f(v') \leq t$.*

A priori $q$ is defined only in $TxM$ but we define it in any $Tx'M$ thanks to the identification with the coordinates. The dot always stands for the derivative along the Jacobi Field ( $\frac{d}{dt}$).

*Proof.* Let $q \in L_{t_f(x,v),v} \cap V_{0,v}$. Since $L_{t_f(x,v),v} \cap H_{0,v} = \text{Vect}\left(e'_l, ...e'_n\right)$, using the symplectic form $\sigma$, we find that for any $i \in [l,n], q_l = 0$. It gives that $q \in F(0)$. Moreover $S(t_f(x,v))q \in V_{0,v}$ thus for any $i \in [1, l-1]$, $(S(t_f(x,v))q)_i = 0$, consequently $q^t\, S(t_f(x,v))q = 0$.

To compute the derivative in the $t$ parameter we again use the symplectic form. Let $(0,z) \in V_{t,v}$ for any $t$ such that $(x,v,t) \in O_{x,v,t_f(x,v)}$ there exists $\phi(t) = (h_t, q_t) = (S(t)w_t, w_t)_{E(0)\times F(0)} \in L_{t,v}$ such that $M(t)\phi(t) = (0,z)$. On one side

$$\begin{aligned}
\sigma\left(\phi(t), \dot\phi(t)\right) &= \sigma\left((S(t)w_t, w_t)_{E(0)\times F(0)}, (\dot S(t)w_t + S(t)\dot w_t, \dot w_t)_{E(0)\times F(0)}\right) \\
&= \sigma\left((S(t)w_t, w_t), (S(t)\dot w_t, \dot w_t)\right) + \sigma\left((S(t)w_t, w_t), (\dot S(t)w_t, 0)\right) \\
&= \left\langle \dot S(t)w_t, w_t \right\rangle.
\end{aligned}$$

on the other side we have

$$\sigma\left(\phi(t), \dot\phi(t)\right) = \sigma\left(M(t)\phi(t), M(t)\dot\phi(t)\right)$$

Since $M(t)\phi(t) = (0,z)$ we have $M(t)\dot\phi(t) = -\dot M(t)\phi(t)$, moreover

$$\dot M(t) = \begin{pmatrix} 0 & I_n \\ -R(t) & 0 \end{pmatrix}.$$

Thus

$$\begin{aligned}
\sigma\left(\phi(t), \dot\phi(t)\right) &= -\sigma\left(M(t)\phi(t), \dot M(t)\phi(t)\right) \\
&= -\sigma\left(M(t)\phi(t), \begin{pmatrix} 0 & I_n \\ -R(t) & 0 \end{pmatrix} M(t)\phi(t)\right) \\
&= -\sigma\left((0,z), (-z,0)\right) = -|z|^2.
\end{aligned}$$

Finally we proved that

$$\left\langle \dot S(t)w_t, w_t \right\rangle = -|z|^2.$$

By compactness we deduce $\delta > 0$ such that $||\dot S(t_f(x,v))|| \geq \delta$.
For the third item we reason by contradiction: we take $(x',v',t') \in O_{x,v,t_f(x,v)}$ and suppose that $q^t\, S(t')q = 0$ and $t' < t_f(x',v')$. By definition $q \in V_{0,v'} \cap F(0)$ thus for any $i \in [l,n], q_l = 0$. Since $t' < t_f(x',v')$ the space $L_{t'(x,v),v}$ is a graph on the horizontal space. Precisely, according to (2.3.2), for any $t \in ]0, t_f(x',v')[$

$$L_{t_f(x,v),v} = \left\{(h, \left(J_1^0(t_f(x,v))\right)^{-1} J_0^1(t)h) | h \in H_{0,v}\right\}.$$

We denote $\left(J_1^0(t_f(x,v))\right)^{-1} J_0^1(t) = K(t)$. The exact same computation done above proves that for any $h \in H_{0,v}$:

$$\left\langle \dot{K}(t)h, h \right\rangle < 0.$$

Since $t\left(J_1^0(t_f(x,v))\right)^{-1}$ converges to $I_n$ when $t$ goes to zero, we deduce that for $t$ small enough $K$ is symmetric positive definite.
For any $h \in H_{0,v}$ and $q' \in V_{0,v}$ we denote $h = (h_1, h_2)$, where $h_1 \in H_{0,v} \cap E(0)$, $h_2 \in H_{0,v} \cap F(0)$ and $q' = (q'_1, q'_2)$, $q'_1 \in V_{0,v} \cap E(0)$, $q'_2 \in V_{0,v} \cap F(0)$. With this notation we have $\left((h,q') = (h_1, q'_2), (q'_1, h_2)\right)_{E\times F}$ and we define for $i \in [1,4]$ the bloc matrix $S_i(t)$, $K_i(t)$ such that

$$(q'_1, q'_2) = (K_1(t)h_1 + K_2(t)h_2, K_3(t)h_1 + K_4(t)h_2),$$

and

$$(h_1, q'_2) = \left(S_1(t)q'_1 + S_2(t)h_2, S_3(t)q'_1 + S_4(t)h_2\right).$$

Since by hypothesis $Lt', v'$ is a graph on $H_{0,v}$ and $F(0)$ we deduce that $S_1(t') = K_1^{-1}(t')$ and in particular we see that $K_1(t')$ is invertible. In the focal direction $q \in F(0) \cap V_{0,v'}$ we have $q = (q_1, 0)_{F(0)}$

$$0 = q^t\, S(t')\, q = q_1^t\, S_1(t')\, q_1 = h_1^t\, K_1(t')\, h_1,$$

where $h_1(t) = K_1^{-1}(t)q_1$. To get a contradiction we just have to remark that, for any $A > 0$, taking $O_{x,v,t_f(x,v)}$ smaller if we need, for any $(x', v', t) \in O_{x,v,t_f(x,v)}$, with $t \le t_f(x', v')$ then

$$h_1^t(t)K_1(t)h_1(t) \le -A h_1^t(t)h_1(t).$$

In the direction $(x, v)$ for any $t \le t_f(x, v)$ we have

$$\left((S_1\,(t)\,q_1, S_3\,(t)\,q_1), (q_1, 0)\right)_{E\times F} = \left((h_1\,(t)\,, 0), (K_1\,(t)\,h_1\,(t)\,, K_3\,(t)\,h_1\,(t))\right) \in L_{t,v}.$$

By definition of $q$ we have $S_1(t)q_1 = h_1(t) \to 0$ when $t \to t_f(x,v)$ and $K_1(t)h_1(t) = q_1$. Without lost of generality we can suppose that $K_1(t)$ is diagonal. Therefore any eigenvalue $\lambda_i(t)$ corresponding to a $q_i \neq 0$ goes to $-\infty$ ( it cannot goes to $+\infty$ since we proved that $t \mapsto K(t)$ decreases ). The eigenvalues are continuous with respect to $(x', v', t)$ therefore shrieking $O_{x,v,t_f(x,v)}$ if we need, we have $h_1^t(t)K_1(t)h_1(t) \le -A h_1^t h_1$. □

**Remark 2.3.18.** *The last proof just says that when the Lagrangian space $L_{t,v}$ has a vertical component it cannot be, before the focalization time, in the same time a graph above the horizontal space and $F$.*

To conclude the proof of theorem 2.1.3 we apply the implicit function theorem in order to find the function $\tau$ needed in definition 2.1.1. We define $\Psi$ by: Let $(\overline{x}, \overline{v}) \in UM$ and $q \in U_xM$ be the focal direction associated. Then the function

$$\begin{aligned} \Psi : \quad & UM \times \mathbb{R}^+ \to \mathbb{R} \\ & (x,v,t) \mapsto q^t S\,(t)\, q \end{aligned}$$

is well defined on a neighbourhood of $(\overline{x}, \overline{v}, t_f(\overline{x}, \overline{v}))$. Moreover $\Psi(\overline{x}, \overline{v}, t_f(\overline{x}, \overline{v})) = 0$ and by lemma 2.3.17 we have:

$$|\partial_t \Psi(\overline{x}, \overline{v}, t_f(\overline{x}, \overline{v}))| = \left| q^t \dot{S}\left(t_f(x, v)\right) q \right| \geq \delta.$$

Therefore we can apply the implicit function theorem to get a function $\tau$ defined on a neighbourhood $O_{\overline{x},\overline{v}}$ of $(\overline{x}, \overline{v})$ such that $\Psi(x, v, \tau(x, v)) = 0$. By lemma 2.3.17 we find that $t_f(x, v) \leq \tau(x, v)$. There only remains to check that $\tau$ is Lipschitz continuous. By compactness there exist $K > 0$ such that

$$\begin{aligned} |d_{\overline{x},\overline{v}} \tau| &= \left| \frac{1}{\partial_t \Psi(\overline{x}, \overline{v}, t_f(\overline{x}, \overline{v}))} d_{\overline{x},\overline{v}} \Psi \right| \\ &\leq \frac{K}{\delta}. \end{aligned}$$

It concludes proof of theorem 2.1.2.

**Remark 2.3.19.** *This method also proves theorem 2.1.4, indeed we easily see that the second differential of $\tau$ at $(x, v)$ is bounded.*

## 2.4 Proof of theorem 2.1.2: Lipschitz continuity of the tangent cut loci

Let $x \in M$, $e_v \in U_x M$, $v = t_c(e_v) e_v$. We want to find a function $\tau$ needed in Theorem 2.1.1. The proof follows the one given by Castelpietra and Rifford in [30], it consists to find such a function thanks to the implicit function theorem. The construction on the function $\tau$ will depends on $x, v$ and $\delta(v)$.

### 2.4.a At the intersection with the tangent focal locus

If $v \in \mathrm{TFL}(x) \cap \mathrm{TCL}(x)$ then $t_c(x, e_v) = t_f(x, e_v)$ and for any $(y, e_w) \in U_x M$ $t_c(y, e_w) \leq t_f(y, e_w)$. By theorem 2.1.3 $t_f$ is $\kappa$ Lipschitz continuous so $\tau = t_f$ works.

This shows that in the non focal case, the function $t_c$ is Lipschitz continuous on $UM$.

### 2.4.b Far from the tangent focal locus

If $v \notin \mathrm{TFL}(x) \cap \mathrm{TCL}(x)$ then $\delta(v) > 0$. Let $\overline{v} \in \bar{\mathrm{I}}(x)$ such that $|v - \overline{v}| = \delta(v)$ and $\exp_x v = \exp_x \overline{v} = y$. let $K \subset TM$ be a compact neighbourhood of the geodesic path $t \in [0, 1] \mapsto \exp_x(t\overline{v})$ and $0 < \varepsilon < t_{inj}$ such that $B(y, \varepsilon) \subset K(y)$. For any $\eta \in T_y S$ with $z = \exp_y \eta \in B(y, \varepsilon)$, we construct a path $s, t \in [0, \varepsilon] \times [0, 1] \mapsto \gamma(s, t)$ satisfying the following conditions for any $(s, t) \in [0, \varepsilon] \times [0, 1]$:

1. $\gamma(0, t) = \gamma(t) = \exp_x(t\overline{v})$.

2. $\gamma(s, 1) = \exp_y(s\eta) = z_s$.

3. $\gamma(s,0)=x$.

4. $\gamma(.,.)\in C^1([0,1]^2,M)$.

5. $(\gamma(s,t),\dot\gamma(s,t))\in K$ .

Working in smooth charts this construction is easy to realize. Note that $s\leq\varepsilon\leq t_{inj}$ implies $s\mapsto \exp_y(s\eta)$ is a minimizing geodesic path, therefore $d^2(y,z_s)=s^2$ and $z_s\in B(y,\varepsilon)$. However $t\mapsto\gamma(s,t)$ and $s\mapsto\gamma(s,t)$ are not necessarily geodesic paths away from $s=0$ and $t=1$. Anyway the first variation formula 2.2.2 of 2.2.3 applied to $\gamma$ gives $K$ such that

$$\begin{aligned} d^2(x,z_s)&\leq A(\gamma(s,t))\leq A(\gamma(0,t))+s\left\langle d_{\overline{v}}\exp_x\overline{v},\eta\right\rangle+\frac{s^2}{2}K \\ &\leq d^2(x,y)+s\left\langle d_{\overline{v}}\exp_x\overline{v},\eta\right\rangle+\frac{K}{2}d^2(y,z_s). \end{aligned} \tag{2.4.1}$$

We can similarly add a perturbation of $x$. Hence we define $u: B(x,\varepsilon)\times B(y,\varepsilon)\to\mathbb{R}^+$ by

$$\begin{aligned} u(x',z)=d^2(x,y)+\langle d_{\overline{v}}\exp_x\overline{v},(\exp_y)^{-1}(z)\rangle-\langle\overline{v},(\exp_x)^{-1}(x')\rangle \\ +K\left(d^2(x,x')+d^2(y,z)\right). \end{aligned} \tag{2.4.2}$$

Note that if we compare to the right hand side of 2.4.1 we have changed $\frac{1}{2}K$ to $K$; this modification shows that $d^2(x',z)=u(x',z)$ if and only if $z=y$ and $x'=x$ otherwise $d^2(x',z)<u(x',z)$. Moreover $u$ is $C^1$ and

$$(d_{x'=x,z=y}u)\cdot(\zeta,\eta)=-\langle\overline{v},\zeta\rangle+\langle d_{\overline{v}}\exp_x\overline{v},\eta\rangle.$$

**Remark 2.4.1.** *We proved here the semiconcavity of the cost function $c(x,.)$ with $d_{\overline{v}}\exp_x\overline{v}$ as a supergradient at $y$.*

By continuity of $\exp_x$ there exits $\varepsilon>0$ such that for any $(x',w)\in B\left((x,v),\varepsilon\right)\subset TM$, $\exp_{x'}(w)=z\in B(y,\varepsilon)$. Let $\gamma(x',w,\theta)=\exp_{x'}(\theta w)$ we define $\Phi:\ B\left((x,v),\varepsilon\right)\to\mathbb{R}$ by

$$w\mapsto u(\exp_x'(w))-A(\gamma(x',w,\theta)).$$

According to the first variation formula (2.2.1) of (2.2.3), $\Phi$ is $C^1$ on $B\left((x,v),\varepsilon\right)$ and the differential at $x,v$ in the direction $\zeta,\xi$ (*i.e.* $x'=\exp_x(r\zeta)$, $w=v+s\xi$) is given by :

$$\begin{aligned} (d_{x,v}\Phi)(\zeta,\xi)&=\langle d_{p=\overline{v}}\exp_x\overline{v},d_{p=v}\exp_x\xi\rangle_y-\langle d_{p=v}\exp_x v,d_{p=v}\exp_x\xi\rangle_y+\langle v-\overline{v},\zeta\rangle \\ &=\langle q-\overline{q},\eta\rangle_y+\langle v-\overline{v},\zeta\rangle, \end{aligned} \tag{2.4.3}$$

where $d_{\overline{v}}\exp_x\overline{v}=-\overline{q}$, $d_v\exp_x v=-q$ and $d_v\exp_x\xi=\eta$.

The set $O_{x,v} = \left\{(x', v', t) \in UM \times \mathbb{R}^+\right\}$ such that $(x', tv') \in B\left((x, v), \varepsilon\right)$ is an open subset of $UM \times \mathbb{R}^+$, moreover $(x, e_v, t_c(e_v)) \in O$. We define $\Psi$ by

$$\begin{aligned} \Psi : \quad & O_{x,v} \to \mathbb{R} \\ & \Psi(t, \zeta) \mapsto \Phi(t\zeta). \end{aligned}$$

By definition $\Psi(x, e_v, t_c(e_v)) = u(x, y) - A(\gamma(x, v, \theta)) = 0$ and for $(x', v', t) \neq (x, e_v, t_c(e_v))$ if $\Psi(x, v', t) = 0$ then (2.4.1) implies

$$d^2(x', \exp_{x'}(tv')) < A(\gamma(x', v', t)).$$

Hence $t > t_c(x', v')$. Furthermore we compute

$$\frac{\partial}{\partial t}\Psi(x, e_v, t_c(e_v)) = d_{p=v}\Phi(x, e_v) = \langle q - \overline{q}, -\frac{1}{t_c(e_v)} q\rangle_y.$$

Since the geodesic flow is Lipschitz continuous, there exists $A > 0$ such that

$$\frac{1}{A} \leq |q - \overline{q}|_y \leq A|v - \overline{v}|_x.$$

Since $|q|_y^2 = |\overline{q}|_y^2$, and $t_c$ is bounded by say $C$ uniformly on $TM$ we have

$$\frac{1}{t_c(e_v)}|\langle q - \overline{q}, q\rangle_y| = \frac{1}{t_c(e_v)}|q - \overline{q}|^2 \geq \frac{1}{2AC}\delta(v)^2 > 0. \tag{2.4.4}$$

Therefore

$$\left|\frac{\partial}{\partial t}\Psi(x, e_v, t_c(e_v))\right| \geq \frac{1}{2C'}\delta(v)^2 > 0. \tag{2.4.5}$$

Consequently we can apply the implicit function theorem to $\Psi(x', v', t) = 0$ at $(x, e_v, t_c(e_v))$; we have a neighbourhood of $(x, e_v)$: $O_{x,v} \subset UM$ and a function $\tau \in C^1(O_{x,v}, \mathbb{R}^+)$ such that

$$\forall (x', v') \in O_{x,v}\, t_c(x', v') \leq \tau(x', v') \quad t_c(x, e_v) = \tau(x, e_v). \tag{2.4.6}$$

The implicit function theorem also gives the differential of $\tau$:

$$\begin{aligned} d_{x'=x, v'=e_v}\tau(\zeta, \xi) &= -\frac{1}{d_{p=v}\Phi(x, e_v)} d_{x'=x, p=v}\Phi(\zeta, \xi) \\ &= \frac{t_c(e_v)}{\langle q - \overline{q}, q\rangle_y}\left[\langle q - \overline{q}, \eta\rangle_y + \langle v - \overline{v}, \zeta\rangle_x\right] \\ &\leq \frac{C''\left(\|\eta|_y + |\zeta|_x\right)}{\delta(v)}. \end{aligned} \tag{2.4.7}$$

We fix $\delta > 0$ and distinguish two cases.

**Case 1:** $\delta(v) \geq \delta$

In this case (2.4.7) becomes:

$$|d_{x'=x,v'=e_v}\tau(\zeta,\xi)| \leq \frac{C}{\delta}\left(|\zeta|+|\xi|\right).$$

Therefore the function $\tau$ is $\kappa$ Lipschitz–continuous, near $(x,e_v)$, for any $\kappa \leq \frac{C}{2\delta}$. In this case we are done. In particular it proves the non-focal case of Theorem 2.1.2.

**Case 2:** $\delta(v) \leq \delta$

In this case $v$ is near a purely focal point. We need to be slightly more precise regarding the estimation of $|d_{x'=x,v'=e_v}\tau(\zeta,\xi)|$. First of all we can rewrite (2.4.4) as

$$\left|\frac{\partial}{\partial t}\Psi(x,e_v,t_c(e_v))\right| \geq \frac{1}{2C'}|v-\overline{v}|^2. \tag{2.4.8}$$

The estimation of the derivative along $x, e_v$ is a bit more tricky. Since the symplectic form is preserved along the Jacobi field we have for any $t>0$:

$$\begin{aligned}\sigma\left((0,v-\overline{v}),(\zeta,\xi)\right)& \\ =\sigma\Big((J_1(t)(v-\overline{v}),\dot{J}_1(t)(v-\overline{v})),&(J_0(t)\zeta+J_1(t)\xi,\dot{J}_0(t)\zeta+\dot{J}_1(t)\xi)\Big)\end{aligned} \tag{2.4.9}$$

thus

$$\begin{aligned}-\left\langle v-\overline{v},\zeta\right\rangle_x - \Big\langle J_0(t)\zeta+J_1(t)\xi,\dot{J}_1(t)(v-\overline{v})\Big\rangle_y &= \\ \Big\langle J_1(t)(v-\overline{v}),\dot{J}_0(t)\zeta+\dot{J}_1(t)\xi\Big\rangle_y&.\end{aligned} \tag{2.4.10}$$

A Taylor formula together with the fact that $\exp_x(v)=\exp_x(\overline{v})$ gives that there exists $C\in\mathbb{R}_+$ such that

$$|d_{p=v}\exp_x\left(v-\overline{v}\right)|_y = |J_1(t_c(e_v))(v-\overline{v})|_y \leq A|v-\overline{v}|^2. \tag{2.4.11}$$

Thus the right hand side of (2.4.10) is smaller then $A|v-\overline{v}|^2$. Thanks to (2.4.7), we can show the Lipschitz continuity separately on each variable; we conclude by examining three different cases. The first case is a perturbation along the variable $v$. The second and third cases deal with a perturbation along the variable $x$.

- If we only consider a perturbation along the speed ($\zeta=0$) then (2.4.10) and (2.4.11) give

$$\left|\Big\langle\eta,\dot{J}_1(t)(v-\overline{v})\Big\rangle_y\right| \leq A|v-\overline{v}|^2. \tag{2.4.12}$$

Moreover a Taylor formula on $q-\overline{q}=d_{p=v}\exp_x(v)-d_{p=\overline{v}}\exp_x(\overline{v})$ gives, for $\delta(v)$ small enough,

$$\dot{J}_1(t)(v-\overline{v}) = q-\overline{q}+o(|v-\overline{v}|^2). \tag{2.4.13}$$

We deduce that there exist $C > 0$ and $\delta > 0$ such that for any $x \in M$ and $v \in \mathrm{I}(x)$ with $\delta(v) \leq \delta$ we have

$$\left|d_{x'=x,p=v}\Psi(0,\xi)\right| = \left|\langle \eta, q - \overline{q}\rangle_y\right| \leq C|v-\overline{v}|^2.$$

Together with (2.4.8), we obtain

$$d_{x'=x,v'=e_v}\tau(0,\xi) \leq \frac{2C'C|v-\overline{v}|^2}{|v-\overline{v}|^2} \leq C.$$

It proves the Lipschitz continuity in the $v$ variable. We recall that the constant $C$ can grow after each inequality but is uniform on $TM$. We now want to look for the Lipschitz continuity in the $x$ variable.

• If the perturbation $\zeta$ is collinear to $v$ ($\zeta = \pm v$) then

$$|d_{x'=x,p=v}\Psi(\zeta,0)| = |\langle v, v-\overline{v}\rangle_x| = |v-\overline{v}|_x^2.$$

Together with (2.4.8) we obtain that

$$d_{x'=x,v'=e_v}\tau(\zeta,0) \leq C.$$

This is exactly the Lipschitz continuity at $(x,v)$ in the $x$ variable along the geodesic direction given by $v$.

• If the perturbation $\zeta$ is in $\mathrm{Ker}J_0\left(t_c(e_v)\right)$ then Equation (2.4.10) becomes

$$-\langle v-\overline{v}, \zeta\rangle_x = \left\langle J_1(t)(v-\overline{v}), \dot{J}_0\left(t_c(e_v)\right)\zeta\right\rangle_y, \tag{2.4.14}$$

together with the estimation (2.4.13) we obtain $C > 0$ such that:

$$d_{x'=x,v'=e_v}\tau(\zeta,0) \leq C.$$

Therefore the function $t_c$ is Lipschitz continuous along these directions.

In dimension two, for any $(x,v) \in M$ we can take a basis with one direction along $e_v$ and the other one in $\mathrm{Ker}J_0\left(t_c(e_v)\right)$. We deduce that $t_c$ is Lipschitz continuous on $UM$. It concludes the proof of Theorem 2.1.2.

**Remark 2.4.2.** *The Lipschitz continuity in the $x$ variable in the geodesic direction has its own importance, in particular it allows us to show Lemma 3.2.3 in the next section.*

**Remark 2.4.3.** *We do not know if in any dimension the function $t_c$ is Lipschitz continuous on $UM$. However, for any $n$–dimensional Riemannian manifold, such that*

$$\dim\left[\mathrm{Ker}J_0\left(t_c(e_v)\right)\right] = n-1,$$

*we proved that $t_c$ is Lipschitz continuous on $UM$. It is for example the case of $S^n$. More generally we proved the Lipschitz continuity for any perturbation in* $[\mathrm{Ker}J_0\left(t_c(e_v)\right)]$.

# Chapter 3

# MTW condition vs. convexity of injectivity domains

## 3.1 Introduction

Let $(M, g)$ be a smooth compact Riemannian manifold of dimension $n \geq 2$. The injectivity domain at some $x \in M$ is defined as

$$\mathrm{I}(x) = \Big\{ v \in T_xM \,|\, \exists\, t > 1 \text{ s.t. } d(x, \exp_x(tv)) = |tv|_x \Big\},$$

where $\exp_x$ denotes the exponential mapping at $x$, $d$ the geodesic distance on $M \times M$ and $|v|_x = \sqrt{g_x(v, v)} = \sqrt{\langle v, v\rangle_x}$. It is an open star-shaped subset of $T_xM$. Thanks to the Itoh-Tanaka Theorem [30, 64, 77] its boundary $\mathrm{TCL}(x)$, which is called tangent cut locus at $x$, is Lipschitz. Its image by the exponential mapping is called the cut locus of $x$,

$$\mathrm{cut}(x) = \exp_x\big(\mathrm{TCL}(x)\big).$$

The geodesic distance from $x$, that is the function $y \mapsto d(x, y)$ is smooth outside $\mathrm{cut}(x)$. Indeed, the distance $d$ is smooth outside the set

$$\mathrm{cut}(M) = \Big\{ (x, y) \in M \times M \,|\, y \in \mathrm{cut}(x) \Big\}.$$

For every $x \in M$, $v \in \mathrm{I}(x)$, and $(\xi, \eta) \in T_xM \times T_xM$, the Ma–Trudinger–Wang tensor (or MTW tensor for short) at $(x, v)$ evaluated on $(\xi, \eta)$ is defined by the formula

$$\mathfrak{S}_{(x,v)}(\xi, \eta) = -\frac{3}{2} \left.\frac{d^2}{ds^2}\right|_{s=0} \left.\frac{d^2}{dt^2}\right|_{t=0} \frac{d^2}{2}\Big( \exp_x(t\xi),\, \exp_x(v + s\eta) \Big). \tag{3.1.1}$$

(The MTW tensor was introduced for the first time in [85] in a slightly different way, see also [110].) Since $v \in I(x)$, $\exp_x(v) \notin \mathrm{cut}(x)$, hence the pair $(\exp_x(t\xi), \exp_x(v + s\eta))$ does not belong to $\mathrm{cut}(M)$ provided $s, t$ are small enough and the right-hand side in (3.1.1) is well-defined. As noticed by Loeper in [82], if $\xi,\eta$ are two unit orthogonal vectors in $T_xM$, then

$$\mathfrak{S}_{(x,0)}(\xi,\eta) = \sigma_x(P)$$

is the sectional curvature of $M$ at $x$ along the plane $P$ generated by $\xi$ and $\eta$.

**Definition 3.1.1.** *We say that* $(M,g)$ *satisfies* **(MTW)** *if the following property is satisfied:*

$$\forall\,(x,v) \in TM \text{ with } v \in \mathrm{I}(x),\ \forall\,(\xi,\eta) \in T_xM \times T_xM, \qquad \Big[\langle \xi,\eta\rangle_x = 0 \Longrightarrow \quad \mathfrak{S}_{(x,v)}(\xi,\eta) \geq 0\Big].$$

First, by Loeper's remark [82, 110], if $(M,g)$ satisfies **(MTW)** then it must have nonnegative sectional curvature. In [83], Loeper and Villani proved that if $(M,g)$ is nonfocal and satisfies a stronger form of the **(MTW)** condition, then all its injectivity domain must be uniformly convex. Following [83], the aim of the present paper is to study the effects of the **(MTW)** condition on convexity properties of injectivity domains. Before to present our results, let us briefly recall the link between **(MTW)** and the regularity of optimal transports with quadratic Riemannian costs.

Let $\mu,\nu$ be two probability measures on $M$ and let $c : M \times M \to \mathbb{R}$ be the quadratic cost defined by

$$c(x,y) = \frac{d(x,y)^2}{2} \qquad \forall\,(x,y) \in M \times M.$$

The Monge problem with measures $\mu,\nu$ and cost $c$ consists in finding a measurable map $T : M \to M$ which minimizes the cost functional

$$\int_M c(x,T(x))\,d\mu(x)$$

under the constraint $T_{\#}\mu = \nu$ ($\nu$ is the image measure of $\mu$ by $T$). If $\mu$ is absolutely continuous, then according to McCann [88], this minimizing problem has a solution $T$, unique up to modification on a $\mu$-negligible set. A first question is whether the optimal transport map can be expected to be continuous. To this purpose, we introduce the following definition.

**Definition 3.1.2.** *We say that* $(M,g)$ *satisfies the* transport continuity property *(abbreviated* (**TCP**)*) if, whenever $\mu$ and $\nu$ are absolutely continuous measures with respect to the volume measure, with densities bounded away from zero and infinity, the optimal transport map $T$ with measures $\mu,\nu$ and cost $c$ is continuous, up to modification on a set of zero volume.*

The following results relating (**TCP**) condition with **(MTW)** and convexity properties of injectivity domains were obtained in [48].

**Theorem 3.1.3.** *Assume that* $(M,g)$ *satisfies the* (**TCP**) *condition. Then* $(M,g)$ *satisfies* **(MTW)** *and all its injectivity domains are convex.*

**Theorem 3.1.4.** *Assume that $M$ has dimension 2. Then the* (**TCP**) *condition holds if and only if* $(M,g)$ *satisfies* **(MTW)** *and all its injectivity domains are convex.*

Let us now state our results. The nonfocal domain at some $x \in M$ is defined as

$$\mathrm{NF}(x) = \Big\{ v \in T_x M \,|\, d_{tv} \exp_x \text{ is not singular for any } t \in [0,1] \Big\}.$$

It is an open star-shaped subset of $T_xM$ whose boundary $\mathrm{TFL}(x)$ is called the tangent focal domain at $x$. The set $\overline{\mathrm{NF}}(x) = \mathrm{NF}(x) \cup \mathrm{TFL}(x)$ can be shown to be semiconvex (see [30]), and the following inclusion always holds:

$$\mathrm{I}(x) \subset \mathrm{NF}(x) \qquad \forall\, x \in M,$$

see for instance [53, Corollary 3.77] or [110, Problem 8.8].

**Definition 3.1.5.** *We say that $(M, g)$ is nonfocal provided*

$$\mathrm{TCL}(x) \subset \mathrm{NF}(x) \qquad \forall\, x \in M.$$

In [83], Loeper and Villani proved that if $(M, g)$ is nonfocal and satisfies a strict form of the **(MTW)** condition (i.e. $\mathfrak{S}_{(x,v)}(\xi, \eta) \geq c|\xi|^2|\eta|^2$ with $c > 0$ in Definition 3.1.1), then all its injectivity domain are uniformly convex. Our first result removes the strictness assumption in the Loeper-Villani theorem.

**Theorem 3.1.6.** *Let $(M, g)$ be a nonfocal Riemannian manifold satisfying* **(MTW)**. *Then all injectivity domains of $M$ are convex.*

Our second result removes the non focal assumption for some particular cases.

**Theorem 3.1.7.** *Let $(M, g)$ be a compact, analytic two-dimensional Riemannian manifold satisfying* **(MTW)**. *Then all injectivity domains of $M$ are convex.*

## 3.2 Preliminary results

Let $M$ be a Riemannian manifold,we denote by $UM \subset TM$ the unit tangent bundle. The distance function to the cut locus at some $x \in M$, $t_{cut} : UM \to (0, \infty)$, is defined by

$$\begin{aligned} t_{cut}(x, v) &:= \sup\Big\{ t \geq 0 \,|\, tv \in \mathrm{I}(x) \Big\} \\ &= \max\Big\{ t \geq 0 \,|\, d(x, \exp_x(tv)) = |t|_x \Big\}. \end{aligned}$$

Then, for every $x \in M$, there holds

$$\mathrm{I}(x) = \Big\{ tv \,|\, 0 \leq t < t_{cut}(x, v),\ v \in U_xM \Big\} \qquad \text{and} \qquad \mathrm{TCL}(x) = \Big\{ t_{cut}(x, v)v \,|\, v \in U_xM \Big\}.$$

For every $x \in M$, we denote by $\rho_x$ the radial distance on $T_xM$, that is

$$\rho_x(v, w) = \begin{cases} |v|_x + |w|_x & \text{if} \quad g_x(v, w) \neq |v|_x|w|_x \\ |v - w|_x & \text{if} \quad g_x(v, w) = |v|_x|w|_x. \end{cases}$$

Then the radial distance to $\mathrm{I}(x)$ satisfies for any $v \in T_xM$,

$$\begin{aligned}\rho_x\big(v, \mathrm{I}(x)\big) &:= \inf\Big\{\rho_x(v,w) \,|\, w \in \mathrm{I}(x)\Big\} \\ &= \begin{cases} \left| v - t_{cut}\left(x, \frac{v}{|v|_x}\right)\frac{v}{|v|_x}\right|_x & \text{if } v \notin \mathrm{I}(x), \\ 0 & \text{otherwise.}\end{cases}\end{aligned}$$

To describe $\mathrm{TCL}(x)$ we define a function $\delta$.

**Definition 3.2.1.** *For any $v \in \mathrm{TCL}(x)$ we define $\delta(v) = \max|v-w|$, $w \in \mathrm{TCL}(x)$ such that $exp_x v = exp_x w$. Moreover for any set $V(x) \subset T_xM$ starshaped, we define $\delta(V(x)) = \min(\delta(v), v \in V(x) \cap \mathrm{TCL}(x))$. Finally we define for $V \subset TM$ :*

$$\delta(V) = \min(\delta(V(x)), x \in M).$$

This property allows us to split $\mathrm{TFL} \cap \mathrm{TCL}$ into two sets: the purely focal set where $\delta = 0$ and the other part $V$ where $\forall (x,v) \in V(x), \delta(v) > 0$. A compact nonfocal Riemannian manifold $M$ satisfies $\delta(M) > 0$. A manifold $M$ satisfying $\delta(M) > 0$ is a bit more general than the nonfocal manifold, since we only avoid purely focal speed.

**Lemma 3.2.2.** *Assume that $V \subset TM$ is stable for the exponential map and satisfy $\delta(V) > 0$. Then, there exist $\delta, K > 0$ such, that for every $v \in V(x)$,*

$$|v|_x^2 - d\big(x, \exp_x(v)\big)^2 \le \delta \quad \Longrightarrow \quad \rho_x\big(v, \mathrm{I}(x)\big) \le K\left(|v|_x^2 - d\big(x,\exp_x(v)\big)^2\right).$$

*In particular assume that $(M,g)$ is nonfocal. Then, there exist $\delta, K > 0$ such, that for every $v \in T_xM$,*

$$|v|_x^2 - d\big(x, \exp_x(v)\big)^2 \le \delta \quad \Longrightarrow \quad \rho_x\big(v, \mathrm{I}(x)\big) \le K\left(|v|_x^2 - d\big(x,\exp_x(v)\big)^2\right).$$

*Proof of Lemma 3.2.2.* First, for every $(x,v) \in V(x)$ we set

$$\psi_x(v) := d_v \exp_x(v) \in V(\exp_x(v)),$$

so that if $\gamma : [0,1] \to M$ is a constant-speed minimizing geodesic path going from $x$ to $y$, with initial velocity $v_0$ and final velocity $v_1$, the map $\psi_x$ is defined by $v_0 \mapsto v_1$. As $\delta(V) > 0$ there exists $\Delta > 0$ such that, for every $v \in \mathrm{TCL}(x)$, there is a geodesic path starting at $x$ with initial velocity $w$ (with $|w|_x = |v|_x$), and finishing at $y = \exp_v(x)$ with final velocity $\psi_x(w)$, satisfying

$$|v|_x^2 - \langle \psi_x(v), \psi_x(w)\rangle_y > \Delta, \tag{3.2.1}$$

see for instance [83, Proposition C.5(a)]. Let $v \in \mathrm{TCL}(x) \cap V(x)$ and $y := \exp_x(v)$ be fixed. As before, consider a minimizing geodesic path from $x$ to $y$ with initial velocity $w$ satisfying (3.2.1). Since $d^2(x,\cdot)$ is locally semiconcave on $M$, $2\psi_x(w)$ is a supergradient for $d^2(x,\cdot)$ at $y$, and the distance from $x$ to its cut locus is uniformly bounded from below (see [110, Definition 10.5 and

Proposition 10.15]), it is easy to show the existence of a smooth function $h : M \to \mathbb{R}$, whose $C^2$ norm does not depend on $x$ and $v$, and such that

$$\begin{cases} d(x,y)^2 = h(y) = |v|_x^2, \\ \nabla h(y) = 2\psi_x(w) \\ d(x,z)^2 \leq h(z), \ \forall\, z \in M, \end{cases}$$

see for instance [83, Proposition C.6]. This gives

$$|(1+\varepsilon)v|_x^2 - d\big(x, \exp_x((1+\varepsilon)v)\big)^2 \geq (1+\varepsilon)^2|v|_x^2 - h\big(\exp_x((1+\varepsilon)v)\big) \qquad \forall\, \varepsilon.$$

Hence, if $C_0$ denotes a uniform bound for the $C^2$ norm of $h$ independent of $x$ and $v$, we get

$$|(1+\varepsilon)v|_x^2 - d\big(x, \exp_x((1+\varepsilon)v)\big)^2 \geq 2\varepsilon\left(|v|_x^2 - \langle \psi_x(v), \psi_x(w)\rangle\right) - C_0\varepsilon^2.$$

Then, using (3.2.1), we deduce that

$$|(1+\varepsilon)v|_x^2 - d\big(x, \exp_x((1+\varepsilon)v)\big)^2 \geq \varepsilon\Delta \qquad \forall\, \varepsilon \in (-\varepsilon_0, \varepsilon_0),$$

where $\varepsilon_0 := \Delta/C_0$. Since

$$\rho_x\big((1+\varepsilon)v, \mathrm{I}(x)\big) = |(1+\varepsilon)v - v|_x = \varepsilon|v|_x,$$

we finally obtain

$$\rho_x\big((1+\varepsilon)v, \mathrm{I}(x)\big) \leq \frac{|v|_x}{\Delta}\left(|(1+\varepsilon)v|_x^2 - d\big(x, \exp_x((1+\varepsilon)v)\big)^2\right) \qquad \forall\, \varepsilon \in (-\varepsilon_0, \varepsilon_0).$$

To conclude the proof it suffices to observe that, by a simple compactness argument, one can easily check that there exists $\delta > 0$ such that any $w \in T_xM \setminus \mathrm{I}(x)$, with $|w|_x^2 - d\big(x, \exp_x(w)\big)^2 \leq \delta$, has the form $(1+\varepsilon)v$ for some $v \in \mathrm{TCL}(x) \cap V(x)$ and $\varepsilon \in [0, \varepsilon_0)$. □

**Lemma 3.2.3.** *There exist $\delta, K > 0$ such that, for every $(x,v) \in TM$ with $\rho_x(v, \mathrm{I}(x)) \leq \delta$,*

$$K^{-1}\rho_x\big(v, \mathrm{I}(x)\big) \leq \rho_y\big(w, \mathrm{I}(y)\big) \leq K\rho_x\big(v, \mathrm{I}(x)\big),$$

*and for every $v \in I(x)$ with $\rho_x(v, \mathrm{TFL}(x)) \geq \delta$,*

$$K^{-1}\rho_x\big(v, \mathrm{TFL}(x)\big) \leq \rho_y\big(w, \mathrm{TFL}(x)\big) \leq K\rho_x\big(v, \mathrm{TFL}(x)\big),$$

*where $y = \exp_x(v)$ and $w = -d_v \exp_x(v) = -\psi_x(v)$, in particular $x = \exp_y(w)$ .*

*Proof of Lemma 3.2.3.* First, we observe that both inequalities trivially hold whenever $v$ belongs to $\mathrm{I}(x)$. Indeed, $\rho_x(v, \mathrm{I}(x)) = 0$ is equivalent to $\rho_y(w, \mathrm{I}(y)) = 0$, so all terms vanish.
Let $(x,v) \in TM$ be fixed. We set $e_v = \dfrac{v}{|v|_x}$ and

$$y = \exp_x(v), \quad w = -\psi_x(v), \quad \overline{w} := t_{cut}\,(y, e_w)\, e_w,$$

and in addition

$$\overline{v} := t_{cut}\,(x, e_v)\, e_v, \quad z := \exp_x(\overline{v}), \quad w' := -\psi_x(\overline{v}).$$

Note that since $\overline{v}$ belongs to $\mathrm{TCL}(x)$, the velocity $w'$ necessarily belongs to $\mathrm{TCL}(z)$, so it satisfies

$$w' = t_{cut}\,(z, e_{w'})\, e_{w'}.$$

Moreover, there holds

$$\rho_x\big(v, \mathrm{I}(x)\big) = |v - \overline{v}|_x \quad \text{and} \quad \rho_y\big(w, \mathrm{I}(y)\big) = |w - \overline{w}|_y\,.$$

Equip $TM$ with any distance $d_{TM}$ which in charts is locally bi-Lipschitz equivalent to the Euclidean distance on $\mathbb{R}^n \times \mathbb{R}^n$. We may assume that $|v|_x$ is bounded. Since the geodesic flow is Lipschitz on compact subsets of $TM$ , there holds

$$d_{TM}\,\big((y, w), (z, w')\big) \le K' \big|v - \overline{v}\big|_x,$$

for some uniform constant $K'$. In fact, if $v$ is close to $\mathrm{I}(x)$ then $\overline{v}$ is close to $v$, and so also $y$ and $z$ are close to each other, so the above inequality follows from our assumption on $d_{TM}$. Then, assuming that $\rho_x(v, \mathrm{I}(x)) \le \delta$ for $\delta > 0$ small enough and taking a local chart in a neighbourhood of $y$ if necessary, we may assume that $y, z, w, \overline{w}, w'$ are in $\mathbb{R}^n$. Moreover, up to a bi-Lipschitz transformation which may affect the estimates only up to a uniform multiplicative constant, we may assume for simplicity that $d_{TM}$ coincides with the Euclidean distance on $\mathbb{R}^n \times \mathbb{R}^n$. Since $y$ is perturbed along the geodesic flow, Remark 2.4.2 gives

$$\begin{aligned}
|w - \overline{w}|_y &= |w|_y - t_c(y, e_w) = |v|_x - |\overline{v}|_x + |\overline{v}|_x - t_c(y, e_w) \\
&= |v|_x - |\overline{v}|_x + \big|w'\big|_z - t_c(y, e_w) \\
&= |v - \overline{v}|_x + t_c(z, e_{w'}) - t_c(y, e_w) \\
&\le |v - \overline{v}|_x + KK'\, |v - \overline{v}|_x\,.
\end{aligned}$$

□

## 3.3 Proof of Theorem 3.1.6: convexity in the non focal case

Let $(M, g)$ be a smooth compact Riemannian manifold of dimension $n \ge 2$ which is nonfocal and satisfies **(MTW)**, and let $\delta, K > 0$ be two constants such that all properties of Lemmas 3.2.2-3.2.3 are satisfied. For every $\mu > 0$, we set

$$\mathrm{I}^\mu(x) = \{v \in T_xM \mid \rho_x(v, \mathrm{I}(x)) \le \mu\}\,.$$

Since $M$ is assumed to be nonfocal, there is $\overline{\mu} > 0$ such that $\mathrm{I}^{\overline{\mu}}(x)$ does not intersect $\mathrm{TFL}(x)$ for any $x \in M$.

**Lemma 3.3.1.** *Taking $K > 0$ larger if necessary, for every $x \in M$ and any $v_0, v_1 \in \mathrm{I}(x)$ there* *holds*

$$v_t := (1-t)v_0 + tv_1 \in \mathrm{I}^{K|v_1-v_0|_x}(x)$$

*and*

$$\bar{q}_t := -d_{v_t}\exp_x(v_t) \in \mathrm{I}^{K|v_1-v_0|_x}\big(y_t := \exp_x(v_t)\big).$$

*Proof of Lemma 3.3.1.* Since the functions $v \in U_xM \mapsto t_{cut}(x,v)$ are uniformly Lipschitz, there is $K > 0$ such that

$$\rho_x\big(v_t, \mathrm{I}(x)\big) \le K|v_1 - v_0|_x \qquad \forall\, v_0, v_1 \in \mathrm{I}(x),\ \forall\, x \in M.$$

The definition of $\mathrm{I}^{K|v_1-v_0|_x}(x)$ together with Lemma 3.2.3 yield both inclusions. □

Our proof requires the use of the extended MTW tensor which was initially introduced by the first and third author in [46]. To define this extension, we let $x \in M$, $v \in \mathrm{NF}(x)$, and $(\xi,\eta) \in T_xM \times T_xM$. Since $y := \exp_x v$ is not conjugate to $x$, by the inverse function theorem, there exist an open neighbourhood $\mathcal{V}$ of $(x,v)$ in $TM$, and an open neighbourhood $\mathcal{W}$ of $(x,y)$ in $M \times M$, such that

$$\begin{array}{rccc} \Psi_{(x,v)} : \mathcal{V} \subset TM & \longrightarrow & \mathcal{W} \subset M \times M \\ (x',v') & \longmapsto & \big(x', \exp_{x'}(v')\big) \end{array}$$

is a smooth diffeomorphism from $\mathcal{V}$ to $\mathcal{W}$. Then we may define $\widehat{c}_{(x,v)} : \mathcal{W} \to \mathbb{R}$ by

$$\widehat{c}_{(x,v)}(x',y') := \frac{1}{2}\big|\Psi^{-1}_{(x,v)}(x',y')\big|^2_{x'}, \qquad \forall\,(x',y') \in \mathcal{W}. \tag{3.3.1}$$

If $v \in \mathrm{I}(x)$ then for $y'$ close to $\exp_x v$ and $x'$ close to $x$ we have $\widehat{c}_{(x,v)}(x',y') = c(x',y') := d(x',y')^2/2$. For every $x \in M$, $v \in \mathrm{NF}(x)$ and $(\xi,\eta) \in T_xM \times T_xM$, the extended Ma–Trudinger–Wang tensor at $(x,v)$ is defined by the formula

$$\overline{\mathfrak{S}}_{(x,v)}(\xi,\eta) = -\frac{3}{2}\,\frac{d^2}{ds^2}\Big|_{s=0}\,\frac{d^2}{dt^2}\Big|_{t=0}\,\widehat{c}_{(x,v)}\Big(\exp_x(t\xi),\, \exp_x(v+s\eta)\Big).$$

The following lemma may be seen as an "extended" version of [83, Lemma 2.3].

**Lemma 3.3.2.** *There exist constants $C, D > 0$ such that, for any $(x,v) \in TM$ with $v \in \mathrm{I}^{\overline{\mu}}(x)$,*

$$\overline{\mathfrak{S}}_{(x,v)}(\xi,\eta) \ge -C\,|\langle \xi,\eta\rangle_x|\,|\xi|_x|\eta|_x - D\rho_x(v,\mathrm{I}(x))|\xi|_x^2|\eta|_x^2 \qquad \forall\,\xi,\eta \in T_xM.$$

We also give a local version of this theorem when $M$ is not nonfocal.

**Lemma 3.3.3.** *Let $V \subset TM$ and $\mu > 0$ such that $\rho(V \cap I, TFL) > \mu$. Then there exist constants $C, D > 0$ such that, for any $(x,v) \in TM$ with $v \in V(x) \cap \mathrm{I}^{\mu}(x)$,*

$$\overline{\mathfrak{S}}_{(x,v)}(\xi,\eta) \ge -C\,|\langle \xi,\eta\rangle_x|\,|\xi|_x|\eta|_x - D\rho_x(v,\mathrm{I}(x))|\xi|_x^2|\eta|_x^2 \qquad \forall\,\xi,\eta \in T_xM.$$

*Proof of Lemma 3.3.2.* The tensors $\mathfrak{S}$ and $\overline{\mathfrak{S}}$ coincide on the sets of $(x,v) \in TM$ such that

$v \in \mathrm{I}(x)$, hence

$$\forall\,(x,v) \in TM \text{ with } v \in \mathrm{I}(x),\ \forall\,(\xi,\eta) \in T_xM \times T_xM,$$
$$\Big[\langle \xi,\eta\rangle_x = 0 \Longrightarrow \quad \overline{\mathfrak{S}}_{(x,v)}(\xi,\eta) \geq 0\Big].$$

Let $\mathrm{I}^{\overline{\mu}}(M)$ be the compact subset of $TM$ defined by

$$\mathrm{I}^{\overline{\mu}}(M) := \cup_{x\in M} \left(\{x\} \times \mathrm{I}^{\overline{\mu}}(x)\right).$$

The mapping

$$(x,v) \in \mathrm{I}^{\overline{\mu}}(M) \longmapsto \big(x, \exp_x(v)\big)$$

is a smooth local diffeomorphism at any $(x,v) \in \mathrm{I}^{\overline{\mu}}(M)$ and the set of $(x,v,\xi,\eta)$ with $(x,v) \in \mathrm{I}^{\overline{\mu}}(M)$ and $\xi,\eta \in U_xM$ such that $\langle \xi,\eta\rangle_x = 0$ is compact. Then there is $D > 0$ such that

$$\overline{\mathfrak{S}}_{(x,v)}(\xi,\eta) \geq -D\rho_x(v,\mathrm{I}(x)),$$

for every $x,v,\xi,\eta$ with $(x,v) \in \mathrm{I}^{\overline{\mu}}(M)$ and $\xi,\eta \in U_xM$ such that $\langle \xi,\eta\rangle_x = 0$. By homogeneity we infer that

$$\overline{\mathfrak{S}}_{(x,v)}(\xi,\eta) \geq -D\rho_x(v,\mathrm{I}(x))|\xi|_x^2|\eta|_x^2,$$

for every $x,v,\xi,\eta$ with $(x,v) \in \mathrm{I}^{\overline{\mu}}(M)$ and $\xi,\eta \in T_xM$ such that $\langle \xi,\eta\rangle_x = 0$. We conclude as in the proof of [83, Lemma 2.3]. □

The proof of 3.3.3 is exactly the same mutatis mutandis.

The following lemma will play a crucial role.

**Lemma 3.3.4.** *Let $h : [0,1] \to [0,\infty)$ be a semiconvex function such that $h(0) = h(1) = 0$ and let $c \geq 0$ be a fixed constant. Assume that there are $t_1 < \ldots, t_N \in (0,1)$ such that $h$ is not differentiable at $t_i$ for $i = 1,\ldots,N$, is $C^2$ on $(0,1) \setminus \{t_1,\ldots,t_N\}$, and satisfies*

$$\ddot{h}(t) \geq -|\dot{h}(t)| - c \qquad \forall\, t \in [0,1] \setminus \big\{t_1,\ldots,t_N\big\}. \tag{3.3.2}$$

*Then*

$$h(t) \leq c\,t(1-t) \qquad \forall\, t \in [0,1]. \tag{3.3.3}$$

*Moreover, if in addition there is $\varepsilon \geq 0$ such that*

$$c \leq \|h\|_\infty + \varepsilon, \tag{3.3.4}$$

*then*

$$\|h\|_\infty \leq \varepsilon/3. \tag{3.3.5}$$

*Proof of Lemma 3.3.4.* Let $a > 0$ and $f : [0,1] \to \mathbb{R}$ be the semiconvex function defined by

$$f(t) = h(t) - at(1-t) \qquad \forall\, t \in [0,1].$$

Let $\bar{t}$ be a maximum point for $f$. Since $f$ is semiconvex, it has to be differentiable at $\bar{t}$, so $\bar{t} \neq t_i$ for $i = 1, \ldots, N$. If $\bar{t} \in (0,1)$, then there holds $\dot{f}(\bar{t}) = 0$ and $\ddot{f}(\bar{t}) \leq 0$. Thus, using (3.3.2) we get

$$|\dot{h}(\bar{t})| = a|2\bar{t} - 1| \leq a,$$

$$0 \geq \ddot{f}(\bar{t}) = \ddot{h}(\bar{t}) + 2a \geq -|\dot{h}(\bar{t})| - c + 2a \geq a - c.$$

This yields a contradiction as soon as $a > c$, which implies that in that case $f$ attains its maximum on the boundary of $[0,1]$. Since $f(0) = f(1) = 0$, we infer that

$$h(t) \leq at(1-t) \qquad \forall\, t \in [0,1],$$

for every $a > c$. Letting $a \downarrow c$, we get (3.3.3). Finally, if (3.3.4) is satisfied, (3.3.3) implies (recall that $h$ is nonnegative)

$$\|h\|_\infty = \sup_{t\in[0,1]} |h(t)| \leq (\|h\|_\infty + \varepsilon) \sup_{t\in[0,1]} t(1-t) = (\|h\|_\infty + \varepsilon)/4.$$

The inequality (3.3.5) follows easily. □

We recall that given $v_0, v_1 \in \mathrm{I}(x)$, for every $t \in [0,1]$ we set

$$v_t := (1-t)v_0 + tv_1, \qquad y_t := \exp_x(v_t), \qquad \bar{q}_t := -d_{v_t} \exp_x(v_t).$$

In addition, whenever $y_t$ does not belong to $\mathrm{cut}(x)$ (or equivalently $x \notin \mathrm{cut}(y_t)$) we denote by $q_t$ the velocity in $\mathrm{I}(y_t)$ such that

$$\exp_{y_t}(q_t) = x \quad \text{ and } \quad |q_t|_{y_t} = d(x, y_t).$$

The following results follow respectively from [48, Lemma B.2] and [49, Proposition 6.1] and does not need the non focality assumption.

**Lemma 3.3.5.** *Let $x \in M$ and $v_0, v_1 \in \mathrm{I}(x)$ be fixed. Then, up to slightly perturbing $v_0$ and $v_1$, we can assume that $v_0, v_1 \in \mathrm{I}(x)$, and that the semiconvex function $h : [0,1] \to \mathbb{R}$ defined as*

$$h(t) := \frac{|v_t|_x^2}{2} - \frac{d(x, y_t)^2}{2} \qquad \forall\, t \in [0,1],$$

*is $C^2$ outside a finite set of times $0 < t_1 < \ldots < t_N < 1$ and not differentiable at $t_i$ for $i = 1, \ldots, N$.*

**Lemma 3.3.6.** *Let $x \in M$ and $v_0, v_1 \in \mathrm{I}(x)$. Assume that the function $h$ defined above is $C^2$ outside a finite set of times $0 < t_1 < \ldots < t_N < 1$, and it is not differentiable at $t_i$ for*

$i = 1, \ldots, N$. *Furthermore, suppose that* $[\overline{q}_t, q_t] \subset \mathrm{NF}(y_t)$ *for all* $t \in [0,1]$. *Then, for every* $t \in [0,1] \setminus \{t_1, \ldots, t_N\}$ *we have*

$$\dot{h}(t) = \langle q_t - \overline{q}_t, \dot{y}_t \rangle_{y_t}, \tag{3.3.6}$$

$$\ddot{h}(t) = \frac{2}{3} \int_0^1 (1-s)\, \overline{\mathfrak{S}}_{(y_t,(1-s)\overline{q}_t + s q_t)}(\dot{y}_t, q_t - \overline{q}_t)\, ds. \tag{3.3.7}$$

We refer the reader to the Appendix for definitions and properties of semiconvex sets.

**Lemma 3.3.7.** *Taking $K$ larger if necessary, the following properties are satisfied for any $x \in M$:*

*(i) Assume there are constants $\omega > 0$ and $\kappa \in (0, \overline{\mu})$ such that*

$$\forall\, v_0, v_1 \in I(x), \quad |v_1 - v_0|_x \leq \omega \quad \Longrightarrow \quad \sup_{q \in [\overline{q}_t, q_t]} \big\{ \rho_{y_t}\big(q, \mathrm{I}(y_t)\big) \big\} \leq \kappa.$$

*Then $\bar{\mathrm{I}}(x)$ is $(K\kappa)$-semiconvex.*

*(ii) Assume there are constants $\omega, \alpha, \varepsilon \geq 0$ such that*

$$\forall\, v_0, v_1 \in I(x), \quad |v_1 - v_0|_x \leq \omega$$
$$\Longrightarrow \quad \sup_{q \in [\overline{q}_t, q_t]} \big\{ \rho_{y_t}\big(q, \mathrm{I}(y_t)\big) \big\} \leq \min \left\{ \alpha \left( \frac{|v_t|_x^2}{2} - \frac{d(x, y_t)^2}{2} \right) + \varepsilon, \overline{\mu} \right\}.$$

*Then $\bar{\mathrm{I}}(x)$ is $(K\varepsilon)$-semiconvex.*

*Proof of Lemma 3.3.7.* We first prove assertion (i). We need to show that there is a uniform constant $K > 0$ and $\nu > 0$ sufficiently small (see Appendix) such that, for any $v_0, v_1 \in \mathrm{I}(x)$ with $|v_0 - v_1|_x < \nu$,

$$\rho_x\big(v_t, \mathrm{I}(x)\big) \leq K\kappa t(1-t) \big|v_0 - v_1\big|^2 \qquad \forall\, t \in [0,1].$$

As above, we set

$$h(t) := \frac{|v_t|_x^2}{2} - \frac{d(x, y_t)^2}{2} \qquad \forall\, t \in [0,1].$$

Let $v_0, v_1 \in \mathrm{I}(x)$ and $\nu > 0$ with $|v_1 - v_0|_x < \nu \leq \omega$ be fixed. By Lemma 3.3.5, up to slightly perturbing $v_0, v_1$ we may assume that $h : [0,1] \to \mathbb{R}$ is semiconvex, $C^2$ outside a finite set of times $0 < t_1 < \ldots < t_N < 1$, and not differentiable at $t_i$ for $i = 1, \ldots, N$. By Lemmas 3.3.2 and 3.3.6 (observe that $\kappa < \overline{\mu}$ and $\mathrm{I}^{\overline{\mu}}(y_t) \subset \mathrm{NF}(y_t)$), for every $t \in [0,1] \setminus \{t_1, \ldots, t_N\}$ it holds

$$\ddot{h}(t) \geq -C|\dot{h}(t)||\dot{y}_t|_{y_t}|q_t - \overline{q}_t|_{y_t} - D \max_{q \in [\overline{q}_t, q_t]} \{\rho_{y_t}(q, \mathrm{I}(y_t))\} \, |\dot{y}_t|_{y_t}^2 |q_t - \overline{q}_t|_{y_t}^2 \qquad \forall\, t \in [0,1].$$

Moreover, by compactness of $M$, there is a uniform constant $E > 0$ such that

$$\big|\dot{y}_t\big|_{y_t} \leq E\big|v_0 - v_1\big|_x \quad \text{and} \quad \big|q_t - \overline{q}_t\big|_{y_t} \leq E.$$

Hence

$$\ddot{h}(t) \geq -CE^2|\dot{h}(t)|\big|v_1 - v_0\big|_x - DE^4\kappa\big|v_1 - v_0\big|_x^2. \tag{3.3.8}$$

Taking $\nu \in (0,\omega)$ small enough yields

$$\ddot{h}(t) \geq -|\dot{h}(t)| - DE^4\kappa|v_1 - v_0|^2 \qquad \forall\, t \in [0,1] \setminus \{t_1,\ldots,t_N\}.$$

So Lemma 3.3.4 gives

$$h(t) \leq DE^4\kappa\, t(1-t)|v_1 - v_0|_x^2 \qquad \forall\, t \in [0,1],$$

which shows that $\mathrm{I}(x)$ is $(K\kappa)$-semiconvex where $K > 0$ is a uniform constant.
To prove (ii) we note that (3.3.8) implies

$$\ddot{h}(t) \geq -CE^2|\dot{h}(t)|\big|v_1 - v_0\big|_x - DE^4\alpha|h(t)|\big|v_1 - v_0\big|_x^2 - DE^4\varepsilon\big|v_1 - v_0\big|_x^2, \tag{3.3.9}$$

which (by choosing $\nu \in (0,\omega)$ sufficiently small) gives

$$\ddot{h}(t) \geq -|\dot{h}(t)| - \|h\|_\infty - DE^4\varepsilon|v_1 - v_0|_x^2 \qquad \forall\, t \in [0,1] \setminus \{t_1,\ldots,t_N\}.$$

Hence, by the second part of Lemma 3.3.4 we obtain

$$\|h\|_\infty \leq \frac{DE^4}{3}\varepsilon|v_1 - v_0|_x^2.$$

Plugging this information back into (3.3.9) gives, for $\nu$ sufficiently small,

$$\ddot{h}(t) \geq -CE^2|\dot{h}(t)|\big|v_1 - v_0\big|_x - 2DE^4\varepsilon\big|v_1 - v_0\big|_x^2.$$

We conclude as in the first part of the proof. □

Returning to the proof of Theorem 3.1.6, we say that $P(r)$ is satisfied if for any $x \in M$ the set $B_x(r) \cap \mathrm{I}(x)$ is convex (here $B_x(r)$ denotes the unit open ball in $T_xM$ with respect to $|\cdot|_x$). If $P(r)$ is satisfied for any $r \geq 0$, then all the injectivity domains of $M$ are convex. Since $r_0 := \inf_{x\in M,\, v\in \mathrm{TCL}(x)} |v|_x$ is strictly positive, $P(r)$ is true for any $r \leq r_0$. Therefore, the set of $r \geq 0$ such that $P(r)$ is satisfied is an interval $J$ with positive length. Moreover, since the convexity property is closed, $J$ is closed. Consequently, in order to prove that $J = [0,\infty)$, it is sufficient to show that $J$ is open.

**Lemma 3.3.8.** *The set of $r$ for which $P(r)$ holds is open in $[0,\infty)$.*

*Proof of Lemma 3.3.8.* Assume that $P(r)$ holds. We want to prove that, if $\beta > 0$ is sufficiently small then also $P(r+\beta)$ holds.
The proof is divided in two steps: first we will show that, for any $\beta \in (0, \overline{\mu}/(2K))$ (here $\mu$ and $K$ are as in Lemma 3.3.7), the sets $B_x(r+\beta) \cap \mathrm{I}(x)$ are $(K\beta)$-semiconvex for any $x \in M$. Then, in Step 2 we show the following "bootstrap-type" result: if the sets $B_x(r+\beta) \cap \mathrm{I}(x)$ are $A$-semiconvex for all $x \in M$, then they are indeed $(A/2)$-semiconvex. The combination of Steps 1 and 2 proves that, for any $x \in M$ and $\beta > 0$ small, the sets $B_x(r+\beta) \cap \mathrm{I}(x)$ are $(K\beta/2^k)$-semiconvex for any $k \in \mathbb{N}$, hence they are convex as desired.

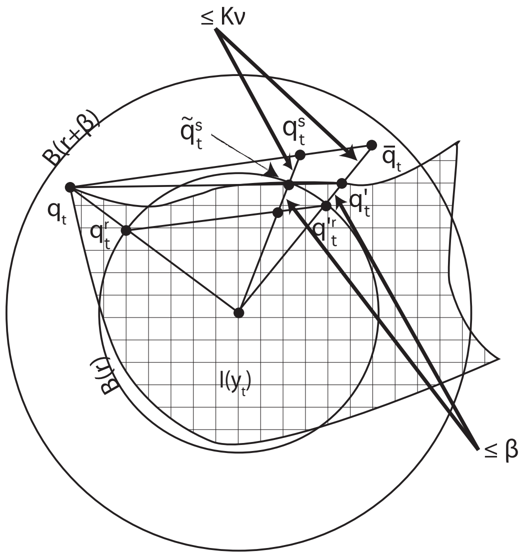



Figure 3.3.1: Definitions

**Step 1:** $\mathrm{I}(x) \cap B_x(r+\beta)$ *is* $(K\beta)$*-semiconvex for any* $\beta \in (0, \overline{\mu}/(2K))$. Assume that $P(r)$ holds, and fix $x \in M$ and $\nu > 0$. Thanks to Lemma 3.3.1, for any $v_0, v_1 \in \mathrm{I}(x)$ with $|v_0 - v_1|_x < \nu$ we have

$$v_t \in \mathrm{I}^{K\nu}(x) \quad \text{ and } \quad \overline{q}_t \in \mathrm{I}^{K\nu}(y_t).$$

Let $\beta > 0$ and $v_0, v_1 \in B_x(r+\beta) \cap \mathrm{I}(x)$ be fixed. By construction

$$|\overline{q}_t|_{y_t} = |v_t|_x < r + \beta, \quad |q_t|_{y_t} \leq |v_t|_x < r + \beta, \quad q_t \in \mathrm{I}(y_t).$$

Since $\overline{q}_t \in \mathrm{I}^{K\nu}(y_t)$ we can find $q_t{}' \in \overline{\mathrm{I}(y_t)} \cap B_{y_t}(r+\beta)$ such that

$$\rho_{y_t}\big(\overline{q}_t, \mathrm{I}(y_t)\big) = |\overline{q}_t - q'{}_t| \leq K\nu.$$

Moreover, using that $\mathrm{I}(y_t)$ is starshaped and that $q_t, q_t' \in B_{y_t}(r+\beta)$, we can find $q_t^r,\ q'^r_t \in \overline{B}_{y_t}(r) \cap \overline{\mathrm{I}(y_t)}$ such that $\rho_{y_t}(q_t, q_t^r) \leq \beta$ and $\rho_{y_t}(q'{}_t, q'^r_t) \leq \beta$. Recalling that by assumption $P(r)$ holds, we have $[q_t^r, q'^r_t] \subset \overline{\mathrm{I}(y_t)}$, which implies (see Figure 1)

$$\begin{aligned} \max_{q \in [\overline{q}_t, q_t]} \big\{\rho_{y_t}\big(q, \mathrm{I}(y_t)\big)\big\} &\leq \max_{q \in [\overline{q}_t, q_t]} \big\{\rho_{y_t}\big(q, [q_t^r, q'^r_t]\big)\big\} \\ &= \max\big\{\rho_{y_t}\big(q_t, q_t^r\big), \rho_{y_t}\big(\overline{q}_t, q'^r_t\big)\big\} \\ &\leq \beta + K\nu, \end{aligned}$$

where at the second line we used that the maximum is attained at one of the extrema of the segment. Thus, Lemma 3.3.7(i) gives $B_x(r+\beta) \cap \mathrm{I}(x)$ is $(K\beta + K^2\nu)$-semiconvex for any $\beta, \nu > 0$ such that $\beta + K\nu < \overline{\mu}/K$, and we conclude letting $\nu \downarrow 0$.

We now show the following "improvement of semiconvexity" result.

**Step 2:** *if all* $\mathrm{I}(x) \cap B_x(r+\beta)$ *are* $A$*-semiconvex, then they are* $(A/2)$*-semiconvex.* We want

to prove that the following holds: there exists $\beta_0 > 0$ small such that, if for some $A > 0$ the sets $\mathrm{I}(x) \cap B_x(r+\beta)$ are $A$-semiconvex for all $x \in M$ and $\beta < \beta_0$, then they are indeed $(A/2)$-semiconvex.
To this aim, by the results in the Appendix, we need to prove that there exists $\nu > 0$ sufficiently small such, that for any $\beta \in (0, \beta_0)$ ($\beta_0$ to be fixed later, independently of $A$) and $v_0, v_1 \in B_x(r+\beta) \cap \mathrm{I}(x)$ with $|v_0 - v_1|_x < \nu$, we have

$$\rho_x\big(v_t, \mathrm{I}(x)\big) \leq \frac{A}{2} t(1-t)\big|v_0 - v_1\big|^2 \qquad \forall\, t \in [0,1].$$

Let $v_0, v_1 \in \mathrm{I}(x)$ and $\nu > 0$ with $|v_1 - v_0|_x < \nu$, and for $t, s \in [0,1]$ set $q_t^s := (1-s)\overline{q}_t + s q_t$ and denote by $\widetilde{q}_t^s$ the intersection of the segments $[0, q_t^s]$ and $[q_t, q_t']$ (see Figure 1). We have (by Lemmas 3.2.2 and 3.2.3)

$$\begin{aligned}
\rho_{y_t}\big(q_t^s, \mathrm{I}(y_t)\big) &\leq \rho_{y_t}\big(q_t^s, \widetilde{q}_t^s\big) + \rho_{y_t}\big(\widetilde{q}_t^s, \mathrm{I}(y_t)\big)\\
&\leq \rho_{y_t}\big(\overline{q}_t, q_t'\big) + \rho_{y_t}\big(\widetilde{q}_t^s, \mathrm{I}(y_t)\big)\\
&= \rho_{y_t}\big(\overline{q}_t, \mathrm{I}(y_t)\big) + \rho_{y_t}\big(\widetilde{q}_t^s, \mathrm{I}(y_t)\big)\\
&\leq K\rho_x\big(v_t, \mathrm{I}(x)\big) + \rho_{y_t}\big(\widetilde{q}_t^s, \mathrm{I}(y_t)\big)\\
&\leq K^2\left(\frac{|v_t|_x^2}{2} - \frac{d(x,y_t)^2}{2}\right) + \rho_{y_t}\big(\widetilde{q}_t^s, \mathrm{I}(y_t)\big).
\end{aligned}$$

Therefore, for every $t \in [0,1]$ we get

$$\max_{q \in [\overline{q}_t, q_t]} \left\{\rho_{y_t}\big(q, \mathrm{I}(y_t)\big)\right\} \leq K^2\left(\frac{|v_t|_x^2}{2} - \frac{d(x,y_t)^2}{2}\right) + \max_{\widehat{q} \in [q_t, q_t']} \left\{\rho_{y_t}\big(\widehat{q}, \mathrm{I}(y_t)\big)\right\}. \tag{3.3.10}$$

Set for every $t, s \in [0,1], \widehat{q}_t^s = (1-s)q_t' + s q_t$. By $A$-semiconvexity we have

$$\rho_{y_t}\big(\widehat{q}_t^s, \mathrm{I}(y_t)\big) \leq As(1-s)|q_t - q_t'|_{y_t}^2. \tag{3.3.11}$$

Then, we finally obtain for $\nu > 0$ small enough,

$$\sup_{q \in [\overline{q}_t, q_t]} \left\{\rho_{y_t}\big(q, \mathrm{I}(y_t)\big)\right\} \leq \min\Big\{K^2\left(\frac{|v_t|_x^2}{2} - \frac{d(x,y_t)^2}{2}\right) + A|q_t - q_t'|_{y_t}^2, \overline{\mu}\Big\},$$

for every $t \in [0,1]$. Two cases may appear:

*First case:* $|q_t - q_t'|_{y_t}^2 \leq 1/(2K)$.
Then by Lemma 3.3.7 (ii), we deduce that $\mathrm{I}(x) \cap B_x(r+\beta)$ is $(A/2)$-semiconvex.

*Second case:* $|q_t - q_t'|_{y_t}^2 > 1/(2K)$.
We work in the plane generated by $0, q_t, q_t'$ in $T_{y_t}M$. We define the curve $\gamma : [0,1] \to \mathrm{I}(y_t)$ by (see Figure 2)

$$\gamma(s) = w \quad \text{ where } \quad \rho_{y_t}\big(\widehat{q}_t^s, \mathrm{I}(y_t)\big) = |\widehat{q}_t^s - w|_{y_t} \qquad \forall\, s \in [0,1],$$

and we denote by $a = \gamma(s_a)$ the first point of $\gamma$ which enters $\overline{B}_{y_t}(r)$ and $b = \gamma(s_b)$ the last one

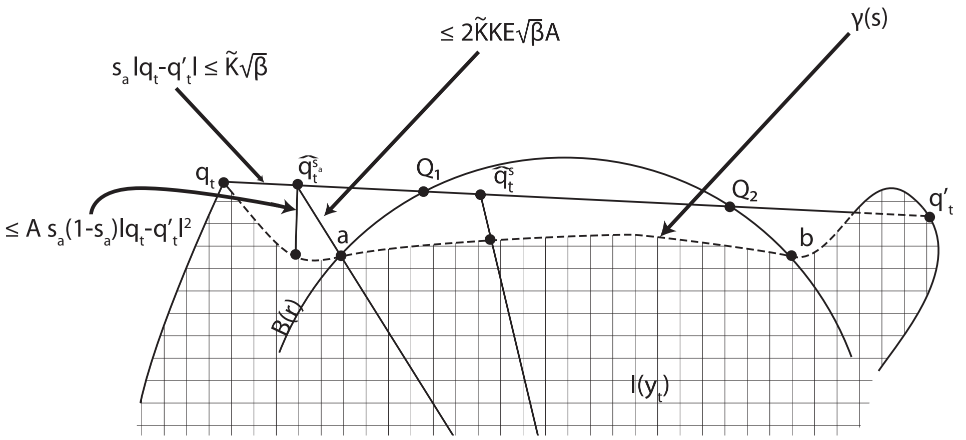



Figure 3.3.2: Estimations

(see Figure 2).

Since both $q_t, q_t'$ belong to $B_{y_t}(r+\beta)$ and $|q_t - q_t'|^2_{y_t} > 1/(2K)$, the intersection of the segment $[q_t, q_t']$ with $B_{y_t}(r)$ is a segment $[Q_1, Q_2]$ such that

$$\left|Q_1 - q_t\right|_{y_t}, \left|Q_2 - q_t'\right|_{y_t} \leq \widetilde{K}\sqrt{\beta},$$

for some uniform constant $\widetilde{K} > 0$ and $\beta > 0$ small enough. Since

$$\left|q_t - \widehat{q}_t^{s_a}\right|_{y_t} \leq \left|Q_1 - q_t\right|_{y_t} \quad \text{and} \quad \left|q_t' - \widehat{q}_t^{s_b}\right|_{y_t} \leq \left|Q_2 - q_t'\right|_{y_t},$$

this implies that both $s_a$ and $1 - s_b$ are bounded by $\dfrac{\widetilde{K}\sqrt{\beta}}{|q_t - q_t'|_{y_t}} < 2K\widetilde{K}\sqrt{\beta}$. Let us distinguish two cases:

- On $[s_a, s_b]$, $P(r)$ is true so $[a, b] \subset \overline{\mathrm{I}(y_t)}$. Hence

$$\sup_{q \in [\widetilde{q}_t^{s_a}, \widetilde{q}_t^{s_b}]} \left\{\rho_{y_t}\big(q, \mathrm{I}(y_t)\big)\right) \leq \max\left\{\rho_{y_t}\big(\widetilde{q}_t^{s_a}, \mathrm{I}(y_t)\big), \rho_{y_t}\big(\widetilde{q}_t^{s_b}, \mathrm{I}(y_t)\big)\right\}.$$

- On $[0, s_a]$ (similarly on $[1 - s_b, 1]$), the $A$-semiconvexity of $\overline{B}_{y_t}(r+\beta) \cap \overline{\mathrm{I}(y_t)}$ yields (by (3.3.11))

$$\rho_{y_t}\big(\widehat{q}_t^s, \mathrm{I}(y_t)\big) \leq As(1-s)|q_t - q_t'|^2_{y_t} \leq As_a|q_t - q_t'|^2_{y_t} \leq 2K\widetilde{K}E\sqrt{\beta}A,$$

where we used that $|q_t - q_t'|^2_{y_t} \leq E$ for some uniform constant $E > 0$. Recalling (3.3.10) we obtain

$$\sup_{q \in [q_t, \overline{q}_t]} \left\{\rho_{y_t}\big(q, \mathrm{I}(y_t)\big)\right\} \leq K^2\left(\frac{|v_t|_x^2}{2} - \frac{d(x, y_t)^2}{2}\right) + 2K\widetilde{K}E\sqrt{\beta}A.$$

Hence, if we choose $\beta_0$ sufficiently small so that $K^2\widetilde{K}E\sqrt{\beta_0} \leq 1/4$ we get

$$\sup_{q\in[q_t,\overline{q}_t]} \left\{\rho_{y_t}\big(q, \mathrm{I}(y_t)\big)\right\} \leq K^2\left(\frac{|v_t|_x^2}{2} - \frac{d(x,y_t)^2}{2}\right) + \frac{A}{2K},$$

and we conclude again by Lemma 3.3.7(ii).

As we explained above, combining Steps 1 and 2 we infer that, for $\beta > 0$ small enough, all the $\mathrm{I}(x) \cap B_x(r+\beta)$ are convex. This shows that the interval $J$ is open in $[0,\infty)$ and concludes the proof of Theorem 3.1.6. □

## 3.4 General version of the proof

To improve the result let us resume the proof. We can identify three acts.

1. For any $x \in M$ and $v_0, v_1 \in I(x)$ we define a function $h$ nonnegative, equal to 0 if and only if $[v_0, v_1] \in I(x)$. We compute the first and second derivatives of $h$ and find an inequality thanks to the extended Ma-Trudinger-Wang tensor (lemma 3.3.6). For this computation we need every points to be in the extended-tensor's domain of definition.

2. Then we show that this inequality implies small semiconvexity for all the possible $h$ functions (lemma 3.3.4).

3. Finally we have a bootstrap argument: the small semiconvexity of step 3 improve the estimation in step 2 which in turn improve the estimation of step 3 up to convexity. For this step we need the lemma 3.2.2,3.2.3.

This summary leads us to the following definitions. Let $Z$ be a set in $TM$ and $Z(x) = \Big\{v \in T_xM \text{ s.t. } (x,v) \in Z\Big\}$.

**Definition 3.4.1.** *We say that* $\overline{\mathbf{MTW}}(-D\rho, C)$ *holds on* $Z$ *if there are constants* $C, D > 0$ *such that, for any* $(x,v) \in TM$ *with* $v \in Z(x)$*, there holds*

$$\overline{\mathfrak{S}}_{(x,v)}(\xi,\eta) \geq -C\,|\langle\xi,\eta\rangle_x|\,|\xi|_x|\eta|_x - D\rho_x(v,\mathrm{I}(x))|\xi|_x^2|\eta|_x^2 \qquad \forall\,\xi,\eta \in T_xM.$$

**Definition 3.4.2** ($P(r)$)**.** *For* $M$ *a Riemannian manifold and* $r \in \mathbb{R}^+$ *we say that* $P(r)$ *is satisfied if for any* $x \in M$ *the set* $B_x(r) \cap \mathrm{I}(x)$ *is convex.*

where $B_x(r)$ is the ball of $T_xM$ with center 0 and radius $r$ with respect to $|\cdot|_x$). Finally we say that $M$ satisfies the convexity-condition if the following conditions are true.

**Definition 3.4.3** (convexity-condition)**.**
*For all* $0 < r <$*diam(M) such that* $P(r)$ *is true there exist a* $\overline{\beta} > 0$ *and a set* $Z \subset TM$*, such that for all* $\beta \leq \overline{\beta}$ *holds:*

1. $\forall x \in M$, $I(x) \cap B_x(r+\beta) \subset Z(x) \subset \overline{\mathrm{NF}}(x)$*, and* $Z(x)$ *is a radial set (cf. Definition A.1.2).*

2. *There are nonnegative constants* $C, D$ *such that* $\overline{\mathbf{MTW}}(-D\rho, C)$ *holds true on* $Z$*.*

3. *For any* $v_0, v_1 \in I(x) \cap B_x(r+\beta)$ *we have* $v_t \in Z(x)$ *and* $[q_t, \overline{q}_t] \subset Z(y_t)$

4. *there are* $\delta, K > 0$ *such that, for every* $v \in Z(x)$*, holds*

$$|v|_x^2 - d\big(x, \exp_x(v)\big)^2 \leq \delta \quad \Longrightarrow \quad \rho_x\big(v, \mathrm{I}(x)\big) \leq K^* \left(|v|_x^2 - d\big(x, \exp_x(v)\big)^2\right).$$

Now we can state a general version of Theorem 3.1.6.

**Theorem 3.4.4.** *Let* $(M, g)$ *be a Riemannian manifold satisfying the convexity-condition 3.4.3. Then all injectivity domains of* $M$ *are convex.*

Before proving this theorem we quickly give a refined version of Lemma 3.3.4.

**Lemma 3.4.5.** *Let* $h : [0,1] \to [0, \infty)$ *be a semiconvex function such that* $h(0) = h(1) = 0$ *and let* $c, C > 0$ *be two fixed constants. Assume that there are* $t_1 < \ldots, t_N \in (0,1)$ *such that* $h$ *is not differentiable at* $t_i$ *for* $i = 1, \ldots, N$*, is* $C^2$ *on* $(0,1) \setminus \{t_1, \ldots, t_N\}$*, and satisfies*

$$\ddot{h}(t) \geq -C|\dot{h}(t)| - c \qquad \forall\, t \in [0,1] \setminus \big\{t_1, \ldots, t_N\big\}. \tag{3.4.1}$$

*Then*

$$h(t) \leq 4ce^{(1+C)} t(1-t) \qquad \forall\, t \in [0,1]. \tag{3.4.2}$$

The main difference with 3.3.4 is that we do not need the constant $C$ to be less than 1 to obtain the $(Kc)$-semiconvexity.

*Proof of Lemma 3.4.5.* Given $\mu, \lambda > 0$, denote by $f_{\mu,\lambda} : [0,1] \to \mathbb{R}$ the semiconvex function defined by

$$f_{\mu,\lambda}(t) = h(t) - \mu \min\left\{1 - e^{-\lambda t}, 1 - e^{-\lambda(1-t)}\right\} \qquad \forall t \in [0,1].$$

Let $\bar{t}$ be a maximum point for $f_{\mu,\lambda}$. Since $f_{\mu,\lambda}$ is semiconvex, it has to be differentiable at $\bar{t}$, so $\bar{t} \neq 1/2$ and $\bar{t} \neq t_i$ for $i = 1, \ldots, N$. If $\bar{t} \in (0, 1/2)$, then there holds $\dot{f}_{\mu,\lambda}(\bar{t}) = 0$ and $\ddot{f}_{\mu,\lambda}(\bar{t}) \leq 0$. Then using (3.4.1), we get

$$|\dot{h}(\bar{t})| = \mu\lambda e^{-\lambda\bar{t}},$$

$$0 \geq \ddot{f}_{\mu,\lambda}(\bar{t}) = \ddot{h}(\bar{t}) + \mu\lambda^2 e^{-\lambda\bar{t}} \geq -C|\dot{h}(\bar{t})| - c + \mu\lambda^2 e^{-\lambda\bar{t}} \geq \mu\lambda(\lambda - C)e^{-\lambda/2} - c.$$

This yields a contradiction provided we choose $\lambda = 1 + C$ and $\mu = 2ce^{1+C}/(1+C)$ and implies that $f_{\mu,\lambda}$ attains its maximum at $t = 0$. Repeating the same argument on $[1/2, 1]$, since $f(0) = f(1) = 0$, we infer that

$$h(t) \leq 2ce^{(1+C)} \min\left\{\frac{1 - e^{-(1+C)t}}{1+C}, \frac{1 - e^{-(1+C)(1-t)}}{1+C}\right\} \qquad \forall t \in [0,1].$$

Noting that

$$\frac{1 - e^{-(1+C)t}}{1+C} \leq t \quad \text{ and } \quad \min\{t, 1-t\} \leq 2t(1-t) \qquad \forall t \in [0,1],$$

we get the result. □

Now we can give the proof of Theorem 3.4.4

*Proof of Theorem 3.4.4.*

**Act 1: Construction of $h$.**

Let $x \in M$ and $v_0, v_1 \in \mathrm{I}(x)$ be fixed. For every $t \in [0,1]$ we set

$$v_t := (1-t)v_0 + tv_1, \qquad y_t := \exp_x(v_t), \qquad \overline{q}_t := -d_{v_t}\exp_x(v_t).$$

In addition, whenever $y_t$ does not belong to $\mathrm{cut}(x)$ (or equivalently $x \notin \mathrm{cut}(y_t)$) we denote by $q_t$ the velocity in $\mathrm{I}(y_t)$ such that

$$\exp_{y_t}(q_t) = x \quad \text{and} \quad |q_t|_{y_t} = d(x, y_t).$$

By Lemma 3.3.5, up to slightly perturbing $v_0$ and $v_1$, we can assume that $v_0, v_1 \in \mathrm{I}(x)$ and the semiconvex function $h : [0,1] \to \mathbb{R}$ defined as

$$h(t) := \frac{|v_t|_x^2}{2} - \frac{d(x, y_t)^2}{2} \qquad \forall\, t \in [0,1],$$

is $C^2$ outside a finite set of times $0 < t_1 < \ldots < t_N < 1$ and not differentiable at $t_i$ for $i = 1, \ldots, N$. To compute it we use the following lemma.

**Lemma 3.4.6.** *Let $r > 0$ such that $P(r)$ is true, $\overline{\beta}$ given by 3.4.3, and consider a $\beta < \overline{\beta}$. For any $v_0, v_1 \in B_x(\beta) \cap I(x)$ and $t \in [0,1] \setminus \{t_1, \ldots, t_N\}$, we have:*

$$\dot{h}(t) = \langle q_t - \overline{q}_t, \dot{y}_t \rangle_{y_t}, \tag{3.4.3}$$

$$\ddot{h}(t) = \frac{2}{3}\int_0^1 (1-s)\, \overline{\mathfrak{S}}_{(y_t,(1-s)\overline{q}_t + sq_t)}(\dot{y}_t, q_t - \overline{q}_t)\, ds. \tag{3.4.4}$$

*Proof.* The first and third condition of definition 3.4.3 guarantee that $[q_t, \overline{q}_t] \subset \overline{\mathrm{NF}}(x)$. In this case [49, Proposition 6.1] or 3.3.6 gives the result. □

**Act 2: Estimations 3.4.6 implies semiconvexity.**

In this part we make a link with lemma 3.4.6 and semiconvexity. It can bee seen as a variation of 3.3.7.

**Lemma 3.4.7.** *Let $r > 0$ such that $P(r)$ is true, $\overline{\beta}$ given by 3.4.3. There exist $\overline{K}, \overline{\kappa} > 0$, such that for any $\beta < \overline{\beta}$, $x \in M$, $v_0, v_1 \in I(x) \cap B_x(r + \beta)$, $y_t, q_t, \overline{q}_t$ the associated vectors and $\kappa \leq \overline{\kappa}$ if*

$$\sup_{q \in [\overline{q}_t, q_t]} \left\{ \rho_{y_t}\big(q, \mathrm{I}(y_t)\big) \right\} \leq \kappa,$$

*then $\overline{\mathrm{I}}(x) \cap B_x(r + \beta)$ is $(\kappa\overline{K})$-semiconvex.*

*Proof of Lemma 3.4.7.* We need to show that, for any $v_0, v_1 \in \mathrm{I}(x) \cap B_x(r+\beta)$,

$$\rho_x\big(v_t, \mathrm{I}(x) \cap B_x(r+\beta)\big) \le \kappa \overline{K} t(1-t)\big|v_0 - v_1\big|^2 \qquad \forall\, t \in [0,1].$$

As above, we set

$$h(t) := \frac{|v_t|_x^2}{2} - \frac{d(x,y_t)^2}{2} \qquad \forall\, t \in [0,1].$$

Let $v_0, v_1 \in \mathrm{I}(x) \cap B_x(r+\beta)$. As in Act 1, up to slightly perturbing $v_0, v_1$ we may assume that $h : [0,1] \to \mathbb{R}$ is semiconvex, $C^2$ outside a finite set of times $0 < t_1 < \ldots < t_N < 1$, and not differentiable at $t_i$ for $i = 1, \ldots, N$. Moreover by lemma (3.4.4) and the second condition in definition 3.4.3 we have

$$\ddot{h}(t) \ge -C|\dot{h}(t)||\dot{y}_t|_{y_t}|q_t - \overline{q}_t|_{y_t} - D \max_{q \in [\overline{q}_t, q_t]} \{\rho_{y_t}(q, \mathrm{I}(y_t))\} \, |\dot{y}_t|^2_{y_t} |q_t - \overline{q}_t|^2_{y_t} \qquad \forall\, t \in [0,1].$$

Since by compactness of $M$, there is a uniform constant $E > 0$ such that

$$\big|\dot{y}_t\big|_{y_t} \le E\big|v_0 - v_1\big|_x \quad \text{and} \quad \big|q_t - \overline{q}_t\big|_{y_t} \le E.$$

Then, we get

$$\ddot{h}(t) \ge -CE^2|\dot{h}(t)|\big|v_1 - v_0\big|_x - DE^4\kappa\big|v_1 - v_0\big|_x^2. \tag{3.4.5}$$

Thus, Lemma 3.4.5 gives

$$h(t) \le 4e^{(1+CE^2)} DE^4 \kappa \, t(1-t)|v_1 - v_0|_x^2 \qquad \forall\, t \in [0,1],$$

which shows, using the condition (4) of 3.4.3, that $\mathrm{I}(x) \cap B(r)$ is $(\kappa\overline{K})$-semiconvex with $\overline{K} = K^* 4e^{(1+CE^2)} DE^4$ and $\overline{\kappa} = \delta \dfrac{K^*}{\overline{K}}$ two constant not depending on $\beta$.

□

**Act 3: Bootstrap.**
Returning to the proof of Theorem 3.4.4, we recall that $P(r)$ is satisfied if for any $x \in M$ the set $B_x(r) \cap \mathrm{I}(x)$ is convex. If $P(r)$ is satisfied for any $r \ge 0$, then all the injectivity domains of $M$ are convex. Since $r_0 := \inf_{x \in M, \, v \in \mathrm{TCL}(x)} |v|_x$ is strictly positive, $P(r)$ is true for any $r \le r_0$. Therefore, the set of $r \ge 0$ such that $P(r)$ is satisfied contains an interval $J$ with positive length. Moreover, since the convexity property is closed, $J$ is closed. Consequently, in order to prove that $J = [0, \infty)$, it is sufficient to show that $J$ is open.

**Lemma 3.4.8.** *The set of $r$ for which $P(r)$ holds is open in $[0, \infty)$.*

*Proof of Lemma 3.4.8.* The proof is divided in two steps: first we will show that there is a $\beta_0$ and $K$ positive such that for any $\beta \in (0, \beta_0)$, the sets $B_x(r+\beta) \cap \mathrm{I}(x)$ are $((K+1)\overline{K}\beta)$-semiconvex for any $x \in M$. Then, in Step 2, we show the following "bootstrap-type" result: if the sets $B_x(r+\beta) \cap \mathrm{I}(x)$ are $A$-semiconvex for all $x \in M$, then they are indeed $(A/2)$-semiconvex. The combination of Steps 1 and 2 proves that, for any $x \in M$ and $\beta > 0$ small, the sets $B_x(r+\beta)\cap \mathrm{I}(x)$ are $((K+1)\overline{K}\beta/2^k)$-semiconvex for any $k \in \mathbb{N}$, hence they are convex as desired.

**Step 1:** $\mathrm{I}(x)\cap B_x(r+\beta)$ *is* $(K+1)\overline{K}\beta)$*-semiconvex for any* $\beta \in (0,\beta_0)$. Assume that $P(r)$ holds, fix $x \in M$ and $\overline{\beta} > \beta > 0$. As $B_x(r+\beta)\cap \mathrm{I}(x)$ is starshaped we can find $v_0^r, v_1^r \in \mathrm{I}(x)\cap B(r)$ with, for $i=(0,1)$, $\rho_x(v_i,v_i^r)\le \beta$. Thuse $P(r)$ implies that for all $t\in[0,1]$, $\rho_x(v_t, I(x))\le\beta$, in other words:

$$v_t \in \mathrm{I}^{\beta}(x).$$

And by lemma 3.2.3

$$\overline{q}_t \in \mathrm{I}^{K\beta}(y_t).$$

By construction we also have

$$|\overline{q}_t|_{y_t} = |v_t|_x < r+\beta, \quad |q_t|_{y_t} \le |v_t|_x < r+\beta, \quad q_t \in \mathrm{I}(y_t).$$

Since $\overline{q}_t \in \mathrm{I}^{K\beta}(y_t)$ we can find ${q_t}' \in \overline{\mathrm{I}(y_t)}\cap B_{y_t}(r+\beta)$ such that

$$\rho_{y_t}\big(\overline{q}_t, \mathrm{I}(y_t)\big) = |\overline{q}_t - {q'}_t| \le K\beta.$$

Moreover, using that $\mathrm{I}(y_t)$ is starshaped and that $q_t, q_t' \in B_{y_t}(r+\beta)$, we can find $q_t^r$, ${q'}_t^r \in \overline{B}_{y_t}(r)\cap \overline{\mathrm{I}(y_t)}$ such that $\rho_{y_t}(q_t,q_t^r)\le\beta$ and $\rho_{y_t}({q'}_t,{q'}_t^r)\le\beta$. $P(r)$ again implies that $[q_t^r,{q'}_t^r]\subset\overline{\mathrm{I}(y_t)}$, so (see Figure 1)

$$\begin{aligned}
\max_{q\in[\overline{q}_t,q_t]}\big\{\rho_{y_t}\big(q,\mathrm{I}(y_t)\big)\big\} &\le \max_{q\in[\overline{q}_t,q_t]}\big\{\rho_{y_t}\big(q,[q_t^r,{q'}_t^r]\big)\big\}\\
&= \max\big\{\rho_{y_t}\big(q_t,q_t^r\big),\rho_{y_t}\big(\overline{q}_t,{q'}_t^r\big)\big\}\\
&\le \beta + K\beta,
\end{aligned}$$

where at the second line we used that the maximum is attained at one of the extrema of the segment. Thus, Lemma 3.4.7 gives $B_x(r+\beta)\cap\mathrm{I}(x)$ is $((K+1)\overline{K}\beta)$-semiconvex for any $\beta_0 = \min(\frac{\overline{K}}{(K+1)},\beta)\ge\beta>0$. We impose a small semiconvexity parameter, taking $((K+1)\overline{K}\beta)\le 1$.

We now prove the following "improvement of semiconvexity" result.

**Step 2:** *if all* $\mathrm{I}(x)\cap B_x(r+\beta)$ *are* $A$*-semiconvex, then they are* $(A/2)$*-semiconvex.* We want to prove that the following holds: there exists $\overline{\beta}_0 > 0$ small such that, if for some $A > 0$ the sets $\mathrm{I}(x)\cap B_x(r+\beta)$ are $A$-semiconvex for all $x \in M$ and $\beta < \overline{\beta}_0$, then they are indeed $(A/2)$-semiconvex. We need to prove that for any $\beta\in(0,\overline{\beta}_0)$ and $v_0,v_1\in B_x(r+\beta)\cap\mathrm{I}(x)$, we have

$$\rho_x\big(v_t,\mathrm{I}(x)\big) \le \frac{A}{2}t(1-t)\big|v_0-v_1\big|^2 \qquad \forall\, t\in[0,1]. \tag{3.4.6}$$

Let $v_0,v_1\in\mathrm{I}(x)\cap B_x(r+\beta)$, we work in the plane generated by $0,v_0,v_1$ in $T_xM$. We define the curve $\gamma:[0,1]\to\mathrm{I}(x)$ by (see Figure 2)

$$\gamma(t) = w \quad \text{where} \quad \rho_x\big(v_t,\mathrm{I}(x)\big) = |v_t - w|_x \qquad \forall\, t\in[0,1],$$

we denote by $a = \gamma(t_a)$ the first point of $\gamma$ which enters $\overline{B}_x(r)$ and $b = \gamma(t_b)$ the last one. Since both $v_0, v_1$ belong to $B_x(r+\beta)$ and $B_x(r) \cap \mathrm{I}(x)$ is convex, the intersection of the segment $[v_0, v_1]$ with $B_x(r)$ is a segment $[Q_1, Q_2]$ such that

$$\left|Q_1 - v_0\right|, \left|Q_2 - v_1\right|_x \leq \widetilde{K}\sqrt{\beta},$$

for some uniform constant $\widetilde{K} > 0$ and $\beta > 0$ small enough. As

$$\left|v_{t_a} - v_0\right|_x \leq \left|Q_1 - v_0\right|_x \quad \text{and} \quad \left|v_{t_b} - v_1\right|_x \leq \left|Q_2 - v_1\right|_x,$$

it implies that both $t_a$ and $1 - t_b$ are bounded by $\dfrac{\widetilde{K}\sqrt{\beta}}{|v_0 - v_1|_x}$. Let us distinguish two cases:

- On $[t_a, t_b]$, $P(r)$ is true so $[a, b] \subset \overline{\mathrm{I}(x)}$. Hence

$$\sup_{v \in [v_{t_a}, v_{t_b}]} \left\{\rho_x\big(v, \mathrm{I}(x)\big)\right) \leq \max\left\{\rho_x\big(v_{t_a}, \mathrm{I}(x)\big), \rho_x\big(v_{t_b}, \mathrm{I}(x)\big)\right\}.$$

- On $[0, t_a]$ (similarly on $[1 - t_b, 1]$), $\overline{B}_x(r+\beta) \cap \overline{\mathrm{I}(x)}$ is $A$ semiconvex so

$$\rho_x\big(v_t, \mathrm{I}(x)\big) \leq A \max(t, 1-t)|v_1 - v_0|_x^2 \leq AE\widetilde{K}\sqrt{\beta}.$$

Combining the two estimations we get for all $t \in [0, 1]$:

$$\rho_x\big(v_t, \mathrm{I}(x)\big) \leq A \max(t, 1-t)|v_1 - v_0|_x^2 \leq AE\widetilde{K}\sqrt{\beta}.$$

Then we define as above $q'_t$ such that $\rho_{y_t}\big(\overline{q}_t, \mathrm{I}(y_t)\big) = |\overline{q}_t - q'_t|$. By lemma 3.2.3 we get :

$$\begin{aligned}
\max_{q \in [\overline{q}_t, q_t]} \left\{\rho_{y_t}\big(q, \mathrm{I}(y_t)\big)\right\} &\leq \rho_{y_t}\big(\overline{q}_t, \mathrm{I}(y_t)\big) + \max_{\widehat{q} \in [q'_t, q_t]} \left\{\rho_{y_t}\big(\widehat{q}, \mathrm{I}(y_t)\big)\right\} \\
&\leq KAE\widetilde{K}\sqrt{\beta} + \max_{\widehat{q} \in [q'_t, q_t]} \left\{\rho_{y_t}\big(\widehat{q}, \mathrm{I}(y_t)\big)\right\},
\end{aligned}$$

As $\overline{B}_x(r+\beta) \cap \overline{\mathrm{I}(x)}$ is $A$ semiconvex for every $x \in M$ the same argument for the estimation of $[v_0, v_1]$ is also valid on each $[q'_t, q_t]$. It yields:

$$\max_{q \in [\overline{q}_t, q_t]} \left\{\rho_{y_t}\big(q, \mathrm{I}(y_t)\big)\right\} \leq KAE\widetilde{K}\sqrt{\beta} + AE\widetilde{K}\sqrt{\beta}$$

Hence, if we choose $\overline{\beta}_0$ sufficiently small such that $A(K+1)\widetilde{K}E\sqrt{\overline{\beta}_0} \leq \min(\kappa, \dfrac{A}{2\overline{K}}, \overline{\beta})$ we get

$$\sup_{q \in [q_t, \overline{q}_t]} \left\{\rho_{y_t}\big(q, \mathrm{I}(y_t)\big)\right\} \leq \frac{A}{2\overline{K}},$$

and obtain the inequation 3.4.6 by Lemma 3.4.7.

Note that the definition of $\overline{\beta}_0$ does not depends on $A$ but on an upper bound of the semiconvexity parameter and in step 1 we fixed it less than 1. So $\overline{\beta}_0$ does not decrease during the bootstrap argument. □

The proof of lemma 3.4.8 concludes the proof of Theorem 3.4.4. □

We can see Theorem 3.1.6 as a consequence of Theorem 3.4.4.

**Claim 1.** *Let $(M, g)$ be a nonfocal Riemannian manifold satisfying* **(MTW)**. *The set $Z = \mathrm{I}^{\overline{\mu}}$ define in section 3.3 satisfies the convexity-condition 3.4.3 for any $r$.*

The first condition of 3.4.3 is ok as $\forall x \in M$, $I(x) \subset \mathrm{I}^{\overline{\mu}} \subset \overline{\mathrm{NF}}(x)$.
The second condition was checked with lemma 3.3.2.
By lemma 3.2.2 the fourth condition is true on $TM$, hence on $Z$. The third condition was used only for small speed at the end of step 1 in the proof of lemma 3.3.8. To check it for all speed we perform the same estimation as in step 1 in the proof of lemma 3.4.8 and get the inequality 3.4.6 which is exactly, $[\overline{q}_t, q_t] \subset I^{\beta+K\beta}$. For $\beta$ small enough we get $[\overline{q}_t, q_t] \subset I^{\overline{\mu}}$. On his side $[v_0, v_1] \subset I^{\overline{\mu}}$ as $\beta \leq \overline{\mu}$.

## 3.5 Proof of Theorem 3.1.7: convexity in the analytic two dimensionnal case

Let $M$ be an analytic Riemannian manifold of dimension 2. The goal of this section is to construct a set $Z$ satisfying the convexity-condition 3.4.3. Thus Theorem 3.4.4 will prove Theorem 3.1.7. We start with preliminary lemma on the structure of TCL near TCL $\cap$ TFL.

### 3.5.a Structure near TCL $\cap$ TFL

**Lemma 3.5.1.** *Let $M$ be an analytic Riemannian manifold. For any $x \in M$, $\mathrm{TCL}(x) \cap \mathrm{TFL}(x)$ is a closed set with a finite number of connected component. Each of them is either an arc or an isolated points of $\mathrm{I}(x)$. We denote them $A_i$, $i \in \{1, \dots, n_x\}$.*
*Moreover there is a $a, N > 0$ such that $\forall x \in M, d_h(A_i, A_j) \geq 2a$, $n_x \leq N$.*

Here $d_h$ is the Hausdorff distance. We refer to the appendix $A$ and picture A.2 for a definition of an arc.

*Proof.* The sets $\mathrm{TCL}(x)$ and $\mathrm{TFL}(x)$ can be written as the projection of sub analytic sets, it leads the conclusion. For the definition and rigorous proof of this implication we refer to [4, 17]. □

**Lemma 3.5.2.** *There exists $\overline{\varepsilon} > 0$ such that, for any $\varepsilon \in (0, \overline{\varepsilon})$, $x \in M$, $i \in \{1, \dots, n_x\}$, $A_i(\varepsilon) = B(A_i, \varepsilon) \cap \partial \mathrm{I}(x)$ is a strictly convex arc of $\mathrm{I}(x)$. Moreover we can impose $\forall i, j \in \{1, \dots, n_x\}$, $d_h(A_i(\overline{\varepsilon}), A_j(\overline{\varepsilon})) \geq a$,*

where $B(A_i, \varepsilon)$ denotes $\cup_{v \in A_i} B_x(v, \varepsilon)$.

*Proof.* Let $x \in M$, as proved by Figalli Rifford and Villani in [47] (proposition 3.1) $\mathbf{MTW}(0)$ implies that the curvature of $\mathrm{TFL}(x)$ and TCL at any point $v \in$ TFL $\cap$ TCL is nonnegative. By analyticity we deduce that the curvature is positive on a small neighbourhood of $v$ and thus

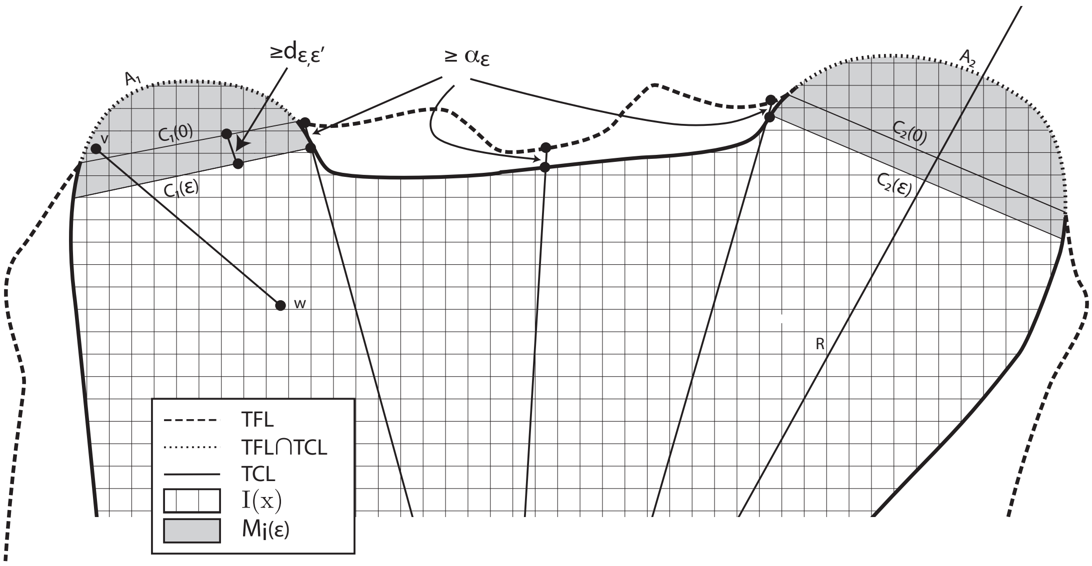



Figure 3.5.1: Definition

for any $i \in [1, n_x]$ on a small neighbourhood of $A_i = A_i(0)$. By compactness we can choose this neighbourhood such that it does not depends on $x$ and $i$. In addition we fix $\bar{\varepsilon} \leq a$ thus lemma 3.5.1 concludes the proof. □

**Definition 3.5.3.** *For any $x \in M$, $i \in \{1, \dots, n_x\}$, $\varepsilon \leq \bar{\varepsilon}$ we consider the arc $A_i(\varepsilon)$ and define $M_i(\varepsilon)$ as its convex hull. We also define the chord $C_i(\varepsilon)$ by $\partial M_i(\varepsilon) = A_i(\varepsilon) \sqcup C_i(\varepsilon)$ and note that $C_i(\varepsilon)$ is a segment (see figure 3).*
*If one $M_i(\varepsilon)$ contains 0 we perform the small surgery drawn in figure 3 to split it in three smaller convex set convex.*

An intersection between a segment drawn in $I$ and a set $M_i$ is quite rigid. We explain it in the following lemma.

**Lemma 3.5.4.** *Let $v, w \in \mathrm{I}(x)$ such that $[v, w] \subset \mathrm{I}(x)$. For any $\varepsilon \leq \bar{\varepsilon}$ and $i \in \{1, \dots, n_x\}$ we consider the intersection between $[v, w]$ and $M_i(\varepsilon)$. Three cases may appear:*

- *if $v, w \in M_i(\varepsilon)$ then $[v, w] \subset M_i(\varepsilon)$,*
- *if $v \in M_i(\varepsilon)$ and $w \notin M_i(\varepsilon)$ then there exist $z \in [v, w] \cap C_i(\varepsilon)$ such that $[v, w] \cap M_i(\varepsilon) = [v, z]$,*
- *if $v, w \notin M_i(\varepsilon)$ then $[v, w] \cap M_i(\varepsilon) = \emptyset$.*

*In particular if $w = \lambda v$ with $\lambda > 1$ and $v \in M_i(\varepsilon)$ then $w \in M_i(\varepsilon)$.*

*Proof.* If $v, w \in M_i(\varepsilon)$ then by convexity of $M_i(\varepsilon)$ we have $[v, w] \subset M_i(\varepsilon)$. If $v \in M_i(\varepsilon)$ and $w \notin M_i(\varepsilon)$ the convexity of $M_i(\varepsilon)$ again implies that $[v, w] \cap M_i(\varepsilon)$ is a segment say $[v, z]$, moreover $z$ cannot be in $A_i(\varepsilon)$ since $[v, w] \in \mathrm{I}(x)$. Last case, if neither $v$ nor $w$ lies in $M_i(\varepsilon)$ and

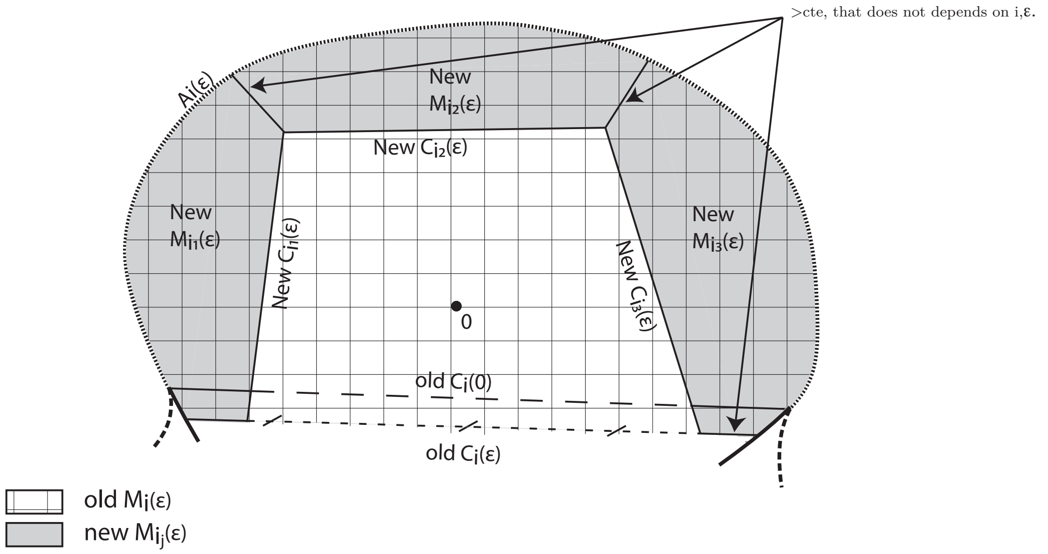



Figure 3.5.2: Surgery

$[v,w] \cap M_i(\varepsilon)$ is not empty then there exist $a \neq b$ with $[v,w] \cap M_i(\varepsilon) = [a,b]$, as above since $[v,w] \in \mathrm{I}(x)$ we have $a,b \in C_i(\varepsilon)$ therefore $[a,b] \in C_i(\varepsilon)$ and consequently $[v,w] \in C_i(\varepsilon)$ ($C_i(\varepsilon)$ is maximal in $\mathrm{I}(x)$ ) which is a contradiction with $v,w \notin M_i(\varepsilon)$. We deduce that $[v,w] \cap M_i(\varepsilon) = \emptyset$. In particular let $w$ and $v$ be colinear ($w = \lambda v$ with $\lambda > 1$ ) and suppose that $v \in M_i(\varepsilon)$. Since $\mathrm{I}(x)$ is starshaped the semi line $R$ colinear to $v$ and $w$ ( $t \in \mathbb{R}^+ \mapsto tv$) crosses $C_i(\varepsilon)$ one and only one time. As $0 \in I(x) \setminus M_i(\varepsilon)$ for any $t > 1$ such that $tv \in I(x)$ where are in the second of the three cases described in the lemma ( the third is excluded since $v \in [0, tv]$) it gives that $tv \in M_i(\varepsilon)$.

□

The main idea of the construction is that $\{\underset{i}{\cup} M_i(\varepsilon)\}^c$ is far from TFL. The next lemma quantify this property.

**Lemma 3.5.5.** *For any $\varepsilon \in (0, \overline{\varepsilon})$ there exist $\alpha_\varepsilon > 0$ such that for all $x \in M$ and $i \in \{1, \ldots, n_x\}$*

1. $\rho_x\big(\{\underset{i}{\cup} M_i(\varepsilon)\}^c \cap I(x), \mathrm{TFL}(x)\big) \geq \alpha_\varepsilon$
2. $\forall \varepsilon' \leq \varepsilon$, $\exists d_{\varepsilon',\varepsilon} > 0$ *such that* $\rho_x\left(M_i(\varepsilon'), (M_i(\varepsilon))^c\right) > d_{\varepsilon,\varepsilon'}$.

*Proof.* For the first statement we remark that $N(\varepsilon) = \big\{\bigcup_{x \in M} \bigcup_{i \in \{1,\ldots,n_x\}} M_i(\varepsilon)\big\}^c$ is compact and $\rho_x(\cdot, \mathrm{TFL}(x))$ is continuous on $TM$ therefore the minimum of $\rho\big(N(\varepsilon), \mathrm{TFL}\big)$ is achieved. As $\big(\cup_{i \in \{1,\ldots,n_x\}} B(V_i, \varepsilon)\big)^c$ is at distance at least $\varepsilon$ from $\mathrm{TCL}(x) \cap \mathrm{TFL}(x)$ the minimum is positive, we choose $\alpha_\varepsilon$ smaller.
As $M_i(\varepsilon') \cap (M_i(\varepsilon))^c = \emptyset$, the second statement is a straightforward compactness argument. □

We have now enough tools to prove the main lemma of this section.

**Lemma 3.5.6.** *Under the condition $B(r) \cap I(x)$ convex, for any $\mu > 0$ there exist $\overline{\beta}, p > 0$ such that for any $\beta \leq \overline{\beta}$, $(v_0, v_1) \in B(r+\beta) \cap I(x)$ then*

1. $\rho_x([v_0, v_1], \mathrm{I}(x)) \leq \mu$ *i.e.* $[v_0, v_1] \subset \mathrm{I}^\mu(x)$,

2. $\rho_x([q_t, \overline{q}_t], \mathrm{I}(y_t)) \leq \mu$ *i.e.* $[q_t, \overline{q}_t] \subset \mathrm{I}^\mu(y_t)$

3. $\rho_x([v_0, v_1] \cap \overline{\mathrm{I}}^c, \mathrm{TFL}(x)) \geq p$ *i.e.* $[v_0, v_1] \cap \overline{\mathrm{I}}^c \subset \mathrm{NF}^p(x)$.

4. $\rho_x([q_t, \overline{q}_t] \cap \mathrm{I}(y_t)^c, \mathrm{TFL}(y_t)) \geq p$ *i.e.* $[q_t, \overline{q}_t] \cap \mathrm{I}(y_t)^c \subset \mathrm{NF}^p(y_t)$,

*where as usual* $v_t = (1-t)v_0 + tv_1$, $y_t = \exp_x(v_t)$, $\overline{q}_t = -d_{v_t}\exp_x(v_t)$.

*Proof of lemma 3.5.6.* .

**Proof of 1, 2**

Let $\beta > 0$ and $v_0, v_1 \in B(r+\beta) \cap I(x)$. As $\mathrm{I}(x)$ is starshaped, we can find $v_0^r, v_1^r \in B(r) \cap I(x)$ with $\rho_x(v_i^r, v_i) \leq \beta$ for $i = 1, 2$. Since $B(r) \cap I(x)$ is convex $[v_0^r, v_1^r] \subset \mathrm{I}(x)$ thus for any $v \in [v_0, v_1]$ we have $\rho_x(v, I(x)) \leq \rho_x(v, [v_0^r, v_1^r]) \leq \beta$ which gives the estimation 1, provided $\beta \leq \mu$. For the last inequality we said that the radial distance between $[v_0, v_1]$ and $[v_0^r, v_1^r]$ is achieved at the endpoints.
The estimation $\rho_x(v, I(x)) \leq \beta$, we've just proved, combine to lemma 3.2.3 implies for all $t \in [0, 1]$, $\overline{q}_t \in \mathrm{I}^{K\beta}(y_t) \cap B(r+\beta)$. As $q_t \in \mathrm{I}(y_t) \cap B(r+\beta)$, mimicking the proof above we get $\rho_{y_t}([q_t, \overline{q}_t], \mathrm{I}(y_t)) \leq \max(\beta, K\beta)$. Taking $\max(\beta, K\beta) \leq \mu$ does the job for both 1 and 3.

**Proof of 3**

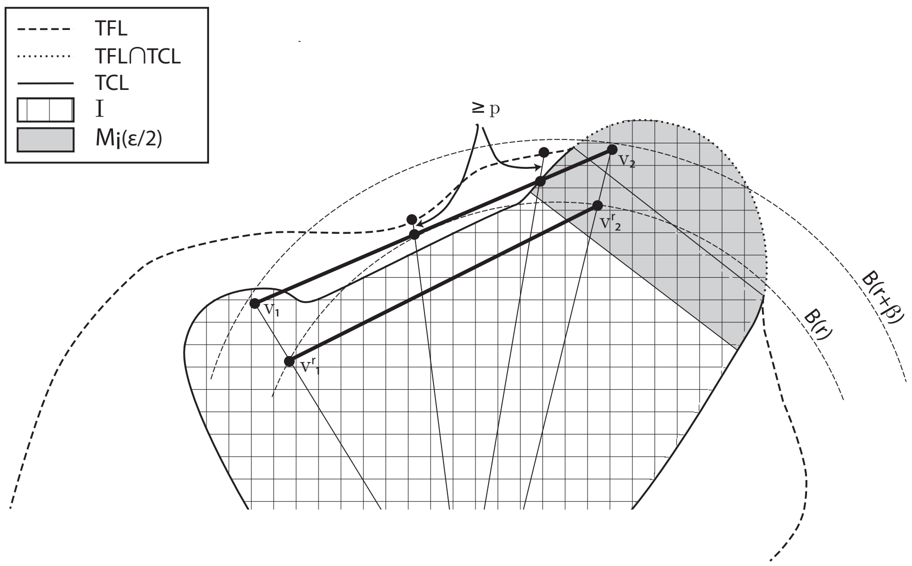
For this estimation, we just need to consider $v_0, v_1 \in B(r+\beta) \cap \mathrm{TCL}(x)$ such that $]v_0, v_1[ \subset \overline{\mathrm{I}}^c(x)$ (as we look for an estimation on $[v_0, v_1] \cap I^c(x)$). We define as above $v_0^r, v_1^r \in B(r) \cap \mathrm{I}(x)$ such that for $i = 1, 2$, $\rho_x(v_i^r, v_i) \leq \beta$ and consider two cases regarding where $v_0^r$ and $v_1^r$ are.
**Case 1,** $v_0^r, v_1^r \notin \bigcup_{i \in \{1,\dots,n_x\}} M_i(\frac{\overline{\varepsilon}}{2})$**.**
If $v_0^r, v_1^r \notin \bigcup_{i \in \{1,\dots,n_x\}} M_i(\frac{\overline{\varepsilon}}{2}) = M(\frac{\overline{\varepsilon}}{2})$ by 3.5.4 we get:

$$[v_0^r, v_1^r] \subset M(\frac{\overline{\varepsilon}}{2})^c \cap \mathrm{I}(x). \tag{3.5.1}$$

For any $v \in [v_0, v_1]$ we note $v^r = [v_0^r, v_1^r] \cap [0, v]$. The estimations 3.5.1 and 1 of lemma 3.5.5 implies:

$$\rho_x(v, \mathrm{TFL}(x)) \geq \rho_x(v^r, \mathrm{TFL}(x)) - \rho_x(v^r, v)) \geq \alpha_{\frac{\overline{\varepsilon}}{2}} - \beta.$$

Taking $\overline{\beta}$ such that $\alpha_{\frac{\overline{\varepsilon}}{2}} - \overline{\beta} \geq \frac{1}{2}\alpha_{\frac{\overline{\varepsilon}}{2}}$ implies that $p = \frac{1}{2}\alpha_{\frac{\overline{\varepsilon}}{2}}$ does the job.
**Case 2,** $v_0^r$ **or** $v_1^r$ **is in** $\bigcup_{i \in \{1,\dots,n_x\}} M_i(\frac{\overline{\varepsilon}}{2})$**.**
• Without lost of generality we suppose that $v_0^r \in M_{i_0}\left(\frac{\overline{\varepsilon}}{2}\right)$, for $i_0 \in \{1, \dots, n_x\}$. In this configuration since $v_0 \in \mathrm{TCL}(x)$ the colinear case in Lemma 3.5.4 implies that $v_0 \in A_{i_0}\left(\frac{\overline{\varepsilon}}{2}\right)$.




• We can now notice that $v_1$ cannot be close to $v_0$. Indeed let $d_{\frac{\overline{\varepsilon}}{2},\overline{\varepsilon}}$ be the constant defined by 2 of lemma 3.5.5, if $|v_0 - v_1| \leq d_{\frac{\overline{\varepsilon}}{2},\overline{\varepsilon}}$ then by definition of $d_{\frac{\overline{\varepsilon}}{2},\overline{\varepsilon}}$ we have $v_1 \in M_i(\overline{\varepsilon})$. Therefore lemma 3.5.4 implies $]v_0, v_1[\subset M_i(\overline{\varepsilon}) \subset I(x)$ which is in contradiction with $]v_0, v_1[\subset \overline{\mathrm{I}}^c(x)$.
• It left us the case when $v_0 \in A_{i_0}\left(\frac{\overline{\varepsilon}}{2}\right)$, $]v_0, v_1[\subset \overline{\mathrm{I}}^c(x)$ and $|v_0 - v_1| \geq d_{\frac{\overline{\varepsilon}}{2},\overline{\varepsilon}} = d$. We are going to show that this case is impossible for $\overline{\beta}$ small enough. If for any $\overline{\beta} > 0$ there exist $x \in M$, $i_0 \in \{1, \ldots, n_x\}$, $v_0 \in A_{i_0}\left(\frac{\overline{\varepsilon}}{2}\right) \cap B(r + \overline{\beta})$ and $v_1 \in \mathrm{TCL}(x) \cap B(r + \overline{\beta})$ such that $|v_0 - v_1| \geq d$ and $]v_0, v_1[\subset \overline{\mathrm{I}}^c(x) \cap B(r + \overline{\beta})$ then by compactness up to subsequences we can find sequences such that:

- ⋄ $\beta_n \to 0$
- ⋄ $x_n \to x_\infty$, $i_n \to i_\infty \in \{1, \ldots, n_{x^\infty}\}$,
- ⋄ $v_0^n \to v_0^\infty \in A_{i_\infty}(\frac{\overline{\varepsilon}}{2}) \cap B(r)$,
- ⋄ $v_1^n \to v_1^\infty \in \mathrm{TCL}(x_\infty) \cap B(r)$,
- ⋄ with $]v_0^n, v_1^n[\subset \overline{\mathrm{I}}^c(x_n) \cap B(r + \beta_n)$ and $|v_0^n - v_1^n| \geq d$.

Passing to the limit when $n$ goes to $+\infty$ we get $v_0^\infty \neq v_1^\infty$ and $]v_0^\infty, v_1^\infty[\subset \overline{\overline{\mathrm{I}}^c}(x_\infty) \cap B(r)$. Since $B(r) \cap \mathrm{I}(x_\infty)$ is convex and $v_0^\infty, v_1^\infty \in \mathrm{TCL}(x_\infty) \cap B(r)$ we obtain $[v_0^\infty, v_1^\infty] \subset \mathrm{TCL}(x_\infty)$. This is a contradiction with the strict convexity of $A_{i_\infty}(\frac{\overline{\varepsilon}}{2})$ at $v_0^\infty$ ( $\mathrm{TCL}(x_\infty)$ cannot be flat at $v_0^\infty$ ). It gives that there is no such points $v_0, v_1$ for $\overline{\beta}$ small enough. These paths are in fact automatically in $\overline{I}(x)$ for $\overline{\beta}$ small enough.

All in one we proved that for $\overline{\beta}$ small enough and $p = \frac{1}{2}\alpha_{\frac{\overline{\varepsilon}}{2}}$ the condition 3 is satisfied.

**Proof of 4**

Thanks to 3 we know that either $v_t \in I(x)$ and we are done ($q_t = \overline{q}_t$) or $v_t \in \mathrm{NF}^p(x)$. In this last situation lemma 3.2.3 implies $\overline{q}_t \in \mathrm{NF}^{\frac{p}{K}}(y_t) \setminus \mathrm{I}(y_t)$, on the other side $q_t \in I(y_t)$. We also recall that $v_t \in B_x(r+\beta)$ and thus $q_t, \overline{q}_t$ are in $B_{y_t}(r+\beta)$. Moreover as we look for an estimation on $[\overline{q}_t, q_t] \cap \overline{\mathrm{I}}^c(y_t)$ we can restrict ourselves to path $[a,b]$ with $a \in \mathrm{TCL}(y_t) \cap B_{y_t}(r+\beta)$ and $b \in B_{y_t}(r+\beta) \cap \mathrm{NF}^{\frac{p}{K}}(y_t)$ and using 3 we can also suppose that $]a,b[\subset \overline{\mathrm{I}}^c(y_t)$. Following the proof of 3 we can define $a^r \in B_{y_t}(r) \cap I(y_t)$ and $b^r \in B_{y_t}(r) \cap \mathrm{I}(y_t)$ with $\rho_{y_t}(a,a^r) \leq \beta$ and $\rho_{y_t}(b,b^r) \leq \max(K\beta,\beta)$. We next remark that there exist $\varepsilon_0$ such that $b^r \notin \bigcup_{i\in\{1,\dots,n_{y_t}\}} M_i(\varepsilon_0)$. Indeed $\rho_{y_t}([b,b^r] \cap \mathrm{TCL}(y_t), \mathrm{TFL}(y_t)) \geq \frac{p}{K} - \max(\beta, K\beta) \geq \frac{p}{2K}$ for $\beta$ small enough. Hence (lemma 3.5.2) there exist $\varepsilon_0$ such that $[b,b^r]\cap\mathrm{TCL}(y_t) \notin \bigcup_{i\in\{1,\dots,n_{y_t}\}} M_i(\varepsilon_0)$ consequently $b^r \notin \bigcup_{i\in\{1,\dots,n_{y_t}\}} M_i(\varepsilon_0)$.

We split the rest of the proof in two cases.

**Case 1,** $a^r \notin \bigcup_{i\in\{1,\dots,n_{y_t}\}} M_i(\frac{\varepsilon_0}{2})$**.** As $b \notin \bigcup_{i\in\{1,\dots,n_{y_t}\}} M_i(\frac{\varepsilon_0}{2})$ this case is similar as the first case of the proof of 3. We repeat this proof and get for any $q \in [a,b]$

$$\rho_{y_t}(q, \mathrm{TFL}(y_t)) \geq \alpha_{\frac{\varepsilon_0}{2}} - \max(K\beta,\beta).$$

Taking $\alpha_{\frac{\varepsilon_0}{2}} - \max(K\beta,\beta) \geq \frac{1}{2}\alpha_{\frac{\varepsilon}{2}}$ is enough.

**Case 2,** $a^r \in \bigcup_{i\in\{1,\dots,n_{y_t}\}} M_i(\frac{\varepsilon_0}{2})$**.**

In this case lemma 3.5.4 implies $a \in \bigcup_{i\in\{1,\dots,n_{y_t}\}} A_i(\frac{\varepsilon_0}{2})$. Since $b^r \notin \bigcup_{i\in\{1,\dots,n_{y_t}\}} M_i(\varepsilon_0)$ lemma 3.5.5 gives $|b-a| \geq |b^r - a^r| \geq d_{\frac{\varepsilon_0}{2},\overline{\varepsilon}}$. We continue by contradiction. If the result was false up to subsequences we could find sequences $\beta_n \to 0, p_n \to 0$, $y_n \to y_\infty$, $a^n \to a^\infty \in \bigcup_{i\in\{1,\dots,n_{x_\infty}\}} A_i(\varepsilon_0) \cap B(r)$, $b^n \to b^\infty \in \mathrm{TCL}(x_\infty) \cap B(r)$ with $]a^n, b^n[\subset \overline{\mathrm{I}}^c(y_n) \cap B(r+\beta_n)$, and $q_n \to q_\infty$ such that $\rho_{y_n}(q_n, \mathrm{TFL}(y_n)) \leq p_n$. We conclude with the exact same contradiction made in the second case of the proof of 3.

□

### 3.5.b Construction of a control domain Z

Finally we define $Z \subset TM$ such that for all $x \in M$, $Z(x) = Z_1(x) \cup Z_2(x)$ with

$$Z_1(x) = \cup_{i=1}^{n_x} M_i(\varepsilon_0) \cap \mathrm{I}(x),$$

$$Z_2(x) = (I^\mu(x) \cap \mathrm{NF}^p)(x)$$

where $\mu$ and $p$ are those define in lemma 3.5.6. We recall that by definition

$$\mathrm{I}^\mu(x) = \{v \in T_xM \mid \rho_x(v, \mathrm{I}(x)) \leq \mu\},$$

$$\mathrm{NF}^p(x) = \{v \in \mathrm{NF}(x) \,|\, \rho_x(v, \mathrm{TFL}(x)) \geq p\}\,.$$

By Theorem 3.4.4, in order to prove Theorem 3.1.7, we just have to show that $Z$ satisfies the convexity-condition 3.4.3.

### 3.5.c $Z$ satisfies the convexity-condition

We prove that taking $\overline{\beta}$ as define in 3.5.6, $Z$ satisfies the convexity-condition in definition 3.4.3.

**Conditions (1)**

By definition $Z_1(x) \subset \mathrm{I}(x) \subset \mathrm{NF}(x)$, $Z_2(x) \subset \mathrm{NF}^p)(x) \subset \mathrm{NF}(x)$. By definition of $\varepsilon_0, p$ and $\overline{\beta}$ (see 3.5.6). For any $v \in B(r+\beta) \cap I(x)$, $\beta \leq \overline{\beta}$ either $v \in Z_1(x)$ either $v \in \mathrm{I}(x) \cap \mathrm{NF}^p(x) \subset Z_2(x)$.

**Conditions (2)**

By hypotheses **MTW** holds on $M$, thus [47] gives a $C > 0$ such that $\overline{\mathbf{MTW}(-D\rho, C)}$ holds on $Z_1(x)$ for any constant $D \geq 0$ (we can reduce a bit $\varepsilon_0$ if needed). As $\delta(Z_2) > 0$ lemma 3.3.3 gives constant $C, D$ for $Z_2$.

**Conditions (3)**

The definition of $Z_1(x)$, $Z_2(x)$ together with lemma 3.5.6 is exactly the third condition. When a speed is not in $Z_1$ it lies in $\mathrm{NF}^p \cap \mathrm{I}^\mu$.

**Conditions (4)**

The fourth condition is trivially true on $Z_1(x) \subset I(x)$. Again as $\delta(Z_2) > 0$ lemma 3.2.2 gives the result.
Finally we proved that $Z$ satisfies 3.4.3; therefore Theorem 3.1.7 is true.

## 3.6 Conclusion and perspectives

This result is very interesting as it makes the link between the **MTW** tensor and the geometry of the manifold. Moreover we can go a further to cover entirely the results obtained in [83], modifying just a bit lemma 3.4.5 we can prove that **MTW**(**K**) for $K > 0$ gives uniform convexity.

**Lemma 3.6.1.** *[Modified lemma] Let $h : [0,1] \to [0,\infty)$ be a semiconvex function such that $h(0) = h(1) = 0$ and let $c, C > 0$ be two fixed constants. Assume that there are $t_1 < \ldots, t_N \in (0,1)$ such that $h$ is not differentiable at $t_i$ for $i = 1, \ldots, N$, is $C^2$ on $(0,1) \setminus \{t_1, \ldots, t_N\}$, and satisfies*

$$\ddot{h}(t) \geq -|\dot{h}(t)| + c \qquad \forall\, t \in [0,1] \setminus \{t_1, \ldots, t_N\}. \tag{3.6.1}$$

*Then*

$$h(t) \leq -4ce^{(1+C)}t(1-t) \qquad \forall\, t \in [0,1]. \tag{3.6.2}$$

With this lemma we indeed obtained that the path $t \mapsto v_t$ is strictly inside $I(x)$, and we get a lower bound of the curvature of $\mathrm{TCL}(x)$. The rest of the proof goes easily since once the convexity achieved there is no more problem of path definition and we can run the exact same argument with lemma 3.6.1. It leads for example to the following theorems.

**Theorem 3.6.2.** *Let $(M, g)$ be a nonfocal Riemannian manifold satisfying* **(MTW(K))**. *Then there exist $\kappa > 0$ such that all injectivity domains of $M$ are $\kappa$ uniformly convex.*

**Theorem 3.6.3.** *Let $(M, g)$ be a compact, analytic two-dimensional Riemannian manifold satisfying* **(MTW(K))**.*Then there exist $\kappa > 0$ such that all injectivity domains of $M$ are $\kappa$ uniformly convex.*

Theorem 3.4.4 on his side is very general, can be extended to $\kappa$ uniform convexity, and let us the possibility to use it for the non-focal case in any dimension. We only need to find a domain satisfying the convexity conditions 3.4.3. For this construction we face two difficulties located around the purely focal points, the first one is to give a sign to the extended tensor near these points, The second one is to isolate them. To be done we need to better understand the repartition of purely focal points, and the behaviour of the tensor near them. For example in dimension 2 we do not used everything we knew, near the purely focal points the tangent injective domain is in fact round. If one succeed in proving this, it will give a very nice formulation for the (**TCP**) condition. To go for a $TC^{\infty}P$ condition one also need to catch the behaviour near purely focal points.

# Appendix A

# Reminders

## A.1 Semiconvexity

We talk about semiconvexity in $\mathbb{R}^n$ but several things can be done on a Riemannian manifold..

**Definition A.1.1** (Semiconvexity). *Let $O$ be a convex subset of $\mathbb{R}^n$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be $\delta$ semiconvex if equivalently, for any $x$, $y$ in $\mathbb{R}^n$ and $t$ in $[0,1]$,*

*(i)* $f((1-t)x + t(y)) \leq (1-t)f(x) + tf(y) + \delta t(1-t)\dfrac{|x-y|^2}{2}$,

*(ii)* $f + \dfrac{|x|^2}{2}$, *is convex*

*(iii)* $\nabla^2 f \geq -\delta$.

Here (iii) has to be understand in a distributional sense where $f$ is not differentiable. $(ii)$ equivalent to $(iii)$ is a classical convex result, $(i)$ equivalent to $(ii)$ is easy to check. $(iii)$ also tells us that as convexity, semiconvexity is a local property. when $(i)$ makes us thing that semiconvexity is a one-dimensional property.

**Definition A.1.2.** *Let $V$ be an open set of $R^{n+1}$. We say that $V$ is a radial set it its boundary is parametrized by a Lipschitz function defined on $S^n$. In particular $V$ is starshaped. $V$ is said to be $\delta$ distance-semiconvex if for any $x$, $y$ in $V$ then $h(t) = \operatorname{dist}((1-t)x + ty, \overline{V})$, is a semiconvex function on $[0,1]$. Or $dist(x, \overline{V})$ is semiconvex. In this case $V$ will be a sublevel of a semiconvex function define on $\mathbb{R}^{n+1}$.*

For a general semiconvex set (define as a sublevel) the second fundamental form as to be bounded below by $-\kappa$, but in general this bound does not imply the $\kappa$ semiconvexity of the distance as in the uniformly convex case. It also depends on the global shape of the domain. The $\kappa$ semiconvexity holds only locally.
But in the case of a domain with $S^n$ lip boundary the result is true.
For a radial set we also define $\rho_0$ the radial distance as in section 1.

**Proposition A.1.3.** *If a radial set $V$ is $\delta$ semiconvex then $\rho_o$ (the radial distance) is $K\delta$ semiconvex. Reciprocally if $\rho_0$ is $\delta$ semiconvex then $V$ is $\delta$ semiconvex.*

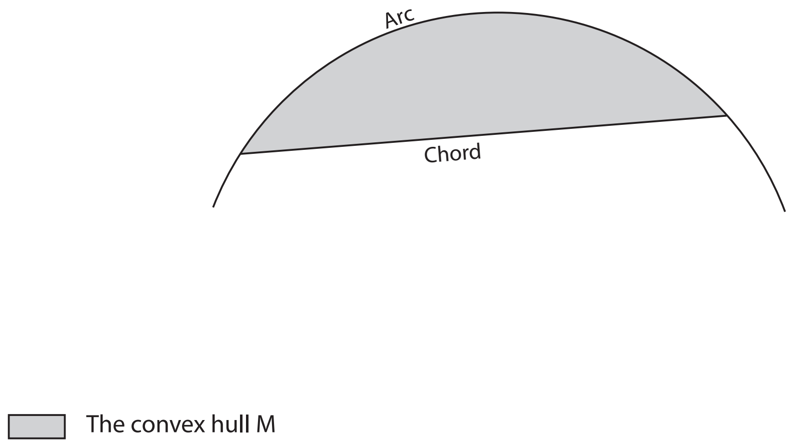



Figure A.2.1: Arc, chord. The arc is not necessary convex.

*Proof.* By compactness there is a $\eta > 0$ such that the angle between the radius and the tangent at a point $x$ in $\partial A$ is bigger than $\eta$. This gives the equivalence of the two distance. □

**Proposition A.1.4.** *Let $V$ be a radial locally semiconvex set, then $V$ is semiconvex*

*Proof.* As $V$ is radial we can consider $\Phi$ the jauge of $V$, the local $\delta$ semiconvexity of the distance gives the $\delta$ semiconvexity of $\Phi$ on the sublevel $\Phi = 1$. By homogeneity it is $\delta$-semiconvex in the classical way. Let $a, b \in V$, we have to show that $\text{dist}((1-t)x + ty, \overline{V}) \leq \delta t(1-t)|x-y|^2$, the orthogonal projection is locally unique but with the radial assumption it is globally unique, it gives a $C^0$ curve $\gamma$ drawn in $\partial V$ such that $\text{dist}((1-t)x + ty, \overline{V}) = |(1-t)x + ty - \gamma(t)|$, the property of the second fundamental form on $\partial V$ gives the result. □

## A.2 Arc

**Definition A.2.1.** *Let $V \subset R^{n+1}$ starshaped with respect to 0. An arc $A$ of $V$ is a connected component (in the $S^n$ parameter) of $\partial V$. Such an arc $A$ is said to be convex (resp. strictly convex) if $\forall a, b \in A$, $t \in [0,1]$, $(1-t)a + tb \in V$ (resp. $(1-t)a + tb \in V \setminus \partial V$).*

As in A.1 we can define the semiconvexity for an arc. The local semiconvexity of a radial set $V$ can be read on the semiconvexity of small arc.

## A.3 Riemannian Settings

### A.3.a definitions

- Let $x \in M$ we denote the tangent space above $x$ by $T_xM$ and consider $v \in T_xM$. The geodesic distance is noted $d$, and the quadratic geodesic distance associated to the Action of the Lagrangian $\frac{1}{2}|v|^2$ is noted $\frac{d^2}{2}$.

- Let $\gamma : TM \times I \to \mathbb{R}$ be the unique geodesic path with initial position $x$ and initial speed $v$, then the exponential map from $TM$ onto $M$ is defined by:

$$(x, v) \mapsto \exp_x(v) = \gamma(x, v, 1).$$

  Moreover the path $t \mapsto \exp_x(tv)$ is geodesic.

- Related to the exponential map we define the injectivity domain

$$\mathrm{I}(x) = \{v \in T_xM \,|\, \exists\, t > 1 \text{ s.t. } d(x, \exp_x(tv)) = |tv|_x\},$$

  and the non focal domain

$$\mathrm{NF}(x) = \{v \in T_xM \,|\, d_{p=v} \exp_x \text{ is invertible }\}.$$

### A.3.b Properties

We consider here a Lagrangian function $L(x, v) = \frac{1}{2}|v|_x^2 + V(x)$, where V is a smooth potential. Note the for $V = 0$ we have the Lagrangian leading to the quadratic geodesic cost. Let $\gamma_0 : [0, 1] \to M$ be a path of least action for $L$, and Let $\gamma : [0, 1] \times [0, 1] \to M$ be a $C^2$ function such that $\gamma(\cdot, 0) = \gamma_0$. The $L$-Jacobi Field above $\gamma_0$ associated to $\gamma$ is defined by

$$J(u) = \left.\frac{d}{ds}\right|_{s=0} \gamma(u, s).$$

We choose an orthonormal basis of $T_xM$: $B = (v, e_2, ..., e_i, ..., e_n)$ and define by parallel transport an orthonormal basis of $T_{\gamma_0(u)}M$: $B(u) = (e_1(u), e_2(u), ..., e_i(u), ..., e_n(u))$. We identify $T_{\gamma_0(u)}M$ with $\mathbb{R}^n$ thanks to the basis $B(u)$. In this $L$-Jacobi Field $J$ satisfies the following linear equation:

$$\ddot{J}(u) + R(u)J(u) + \nabla^2 V u(J(u)) = 0.$$

Noting *Riem* the Riemannian tensor, the matrix $R$ is defined by for all $i, j \in [1, n]$:

$$R_{ij}(u) = \left\langle Riem_{\gamma_0(u)}\left(\dot{\gamma}_0(u), e_i(u)\right)\dot{\gamma}_0(u), e_j(u)\right\rangle_{\gamma_0(u)}.$$

In particular for $V = 0$ we find that a Jacobi Field is solution of

$$\ddot{J}(u) + R(u)J(u) = 0.$$

The proof is quite simple it suffices to differentiate two times the Jacobi field and use the definition of the Riemannian tensor, see for example [53, 86, 110].
The identification of $T_{\gamma_0(u)}M$ with $\mathbb{R}^n$ through the basis $B(u)$ plus the fact that the Jacobi fields are $C^\infty$ justifies the limited development done in chapter one.

# Part II

# Particle approximation of the one dimensional Keller-Segel equation, stability and rigidity of the blow-up

**Abstract**

This part is devoted to the study of the behaviour of a Keller-Segel solution in the super critical case [39]. In particular we are interested in the mass quantization problem [104], that is to quantify the mass aggregated when the blow-up occurs. To study this behaviour we consider a particle approximation of a Keller-Segel type equation in dimension 1. To define this approximation we use the gradient flow interpretation of the Keller-Segel equation and the peculiar structure of the Wasserstein space in dimension 1. We show two kinds of results the first one is a theorem of stability for the blow-up mechanism. The second one is a sort of Liouville theorem, or a rigidity theorem for the blow-up mechanism. The appendix B can be read first as an introduction to the Keller-Segel equation in dimension 2.

# Chapter 4

# Introduction

## 4.1 Modelling cell to cell interaction

### 4.1.a The Keller-Segel model

In this chapter we study a chemotaxis related problem. We consider particles with density $\rho(t,x)$ at time $t \in \mathbb{R}_+$ and position $x \in \mathbb{R}^n$. Each particle produces a chemical substance with concentration $c(t,x)$ which in turn attracts the cells. We suppose that the substance instantly reaches the equilibrium state and does not degenerate. The resulting attraction-diffusion model is the so-called Keller-Segel ( or Patlak-Keller-Segel ) model in its simplified form [65]

$$\partial_t \rho(t,x) - \Delta \rho(t,x) + \chi \nabla \cdot (\rho(t,x) \nabla c(t,x)) = 0, \qquad (t,x) \in \mathbb{R}_+ \times \mathbb{R}^n \qquad (4.1.1a)$$

$$-\Delta c(t,x) = \rho(t,x), \qquad (t,x) \in \mathbb{R}_+ \times \mathbb{R}^n, \qquad (4.1.1b)$$

where $\chi$ is a fixed positive parameter quantifying the intensity of the attraction. We immediately see that at any time $t \in \mathbb{R}_+$, the total mass $M = \displaystyle\int \rho(t,x)\,dx$ is constant, and so is the first moment $\displaystyle\int x\rho(t,x)\,dx$ (and the center of mass $\dfrac{1}{M}\displaystyle\int x\rho(t,x)\,dx$). Equation (4.1.1a) shows a competition between the attraction (due to the concentration gradient $\nabla c$) and the diffusion.

This model was introduced independently by Patlak in [95] and Keller and Segel in [69], [70], [71]. The well-posedness issue for (4.1.1) strongly depends on the space dimension. In dimension $n = 1$ there exist global solutions for any mass $M$ [91], whereas in dimension $n = 2$ solutions develop singularities (blow-up of the density in $L^\infty$ norm) in finite time when the mass $M$ is too large. This phenomenon was first postulated by Childress and Percus [31]. The existence of solutions for small mass in dimension two was proved by Jäger and Luckhaus [65], then by Biler and Nadzieja [7]. The occurrence of blow-up for large mass was proved by Nagai [91]. He obtained the correct threshold for non-existence of global solutions, namely $\chi M > 8\pi$. On the other hand the right threshold for global existence, $\chi M < 8\pi$ was obtained independently by Diaz, Nagai and Rakotoson [38] and Blanchet, Dolbeault and Perthame [12]. Explicit radially symmetric, blowing-up solutions have been constructed by Herrero and Velázquez [58, 59]. In higher dimension $n \geq 3$, not only the mass determines whether solutions blow-up or not, but also the initial distribution of the density. In particular there is not such a clear dichotomy between global existence and blow-up. We refer to Biler [5] Corrias-Perthame-Zaag [32] and [24] for precise results and discussion. In the radially symmetric case, Brenner et al. [15] exhibit two kinds of possible blow-up ; and Herrero, Medina and Velázquez built explicit blowing-up solutions in dimension three [57]. For an extensive discussion about the well-posedness issue, we refer to the review articles by Hortsmann [62], Perthame [96], Hillen and Painter [61], and the books by Suzuki [104] and Perthame [97]. For a short summary and proof of the main results we also refer to the appendix B.

### 4.1.b Self interacting and diffusing particles

Equation (4.1.1) may be rewritten so as to see it as a particular case of a larger class of problems. Let $\kappa$ be the Green kernel of $-\Delta$ in dimension $n$; thus $c\,(t,.) = \kappa\star\rho\,(t,.)$ (note that the convolution is only in the space variable). Therefore $\nabla c = (\nabla\kappa \star \rho)$ and (4.1.1) becomes:

$$\partial_t\rho\,(t,x) - \Delta\rho\,(t,x) + \chi\nabla\cdot\left(\rho\,(t,x)\,(\nabla\kappa\star\rho)\right) = 0 \quad (t,x) \in \mathbb{R}_+\times\mathbb{R}^n. \tag{4.1.2}$$

We may then construct a family of similar equations where $\kappa$ is no more related to the Laplacian.

### 4.1.c Optimal transport point of view

A third way of writing the problem is to see it in its divergence form. Let $\kappa$ be a symmetric convolution kernel; Problem (4.1.2) may be rewritten as:

$$\partial_t\rho\,(t,x) = \nabla\cdot\left[\rho\,(t,x)\,\nabla\left(\log\left(\rho\,(t,x)\right) - \chi\nabla c\,(t,x)\right)\right] \qquad (t,x)\in\mathbb{R}_+\times\mathbb{R}^n, \tag{4.1.3a}$$

$$c\,(t,y) = \left[\kappa\star\rho\,(t,\cdot)\right](y) \qquad (t,y)\in\mathbb{R}_+\times\mathbb{R}^n. \tag{4.1.3b}$$

For such an equation we define the free energy $E$ by

$$E\left(\rho(t,\cdot)\right) = \int \rho\,(t,x)\log\left(\rho\,(t,x)\right)dx - \frac{\chi}{2}\int \kappa\,(x-y)\,\rho\,(t,x)\,\rho\,(t,y)\,dxdy. \tag{4.1.4}$$

The remarkable fact is that for any value of $\chi$ the free energy is nonincreasing as a function of $t$. Indeed using the symmetry of $\kappa$ and an integration by parts, we get:

$$\begin{aligned}\frac{d}{dt}E\left(\rho(t,\cdot)\right) &= \int \partial_t\rho\left(\log\left(\rho\right)+1\right)dx - \chi\int \kappa\star\rho\partial_t\rho dxdy\\ &= -\int \rho\left[\nabla\log(\rho) - \chi\nabla c\right]\nabla\log\left(\rho\right)dx + \chi\int \nabla c\rho\left[\nabla\log(\rho)-\chi\nabla c\right]dx\\ &= -\int \rho\left[\nabla\log(\rho)-\chi\nabla c\right]^2 dx.\end{aligned}$$

As we shall see later the free energy $E$ is a powerful tool to study the Keller-Segel equation. Indeed, $E$ is not only nonincreasing along the flow but thanks to this formulation, Equation (4.1.3) can formally be seen as a gradient flow for $E$ in the Wasserstein space ($W_2^{ac}$); this point

of view can be formalized in some particular cases: Carrillo Villani and McCann introduce and take advantage of the gradient flow interpretation to prove some dissipative energy estimate in [29] and [28]. In dimension one we will in chapter 5 that the gradient flow in Wasserstein is in fact a real gradient flow in $L^2$.
In the particular case of the dimension two the interaction kernel reads $\kappa(z) = -\frac{1}{2\pi}\log|z|$. Therefore we have the following scaling property for the energy for any $\lambda > 0$:

$$E\left(\lambda^2\rho\left(t,\lambda\cdot\right)\right) = E\left(\rho\left(t,\cdot\right)\right) + M\left(2 - \frac{M\chi}{4\pi}\right)\log\left(\lambda\right). \tag{4.1.5}$$

## 4.2 The blow-up phenomenon in dimension two

The behaviour of the solution of Problem (4.1.1) in dimension two depends on the product $M\chi$. For $M\chi$ smaller than the critical value $8\pi$, the diffusion dominates and the solution goes to 0. The critical threshold was obtained successively by [91, 38, 52, 39]. In the critical case $M\chi = 8\pi$, the solution converges to a Dirac mass as $t \to +\infty$ if the second moment is initially finite [11, 6]. The situation is more complicated when the second moment is infinite, see [9, 27]. In the super critical case $M\chi > 8\pi$ the solution blows-up in finite time if the second moment is initially finite.
We read the critical parameter on the free energy. When $M\chi < 8\pi$ the free energy goes to $-\infty$ as $\lambda \to 0$. This corresponds to a spreading of the density $\rho$. Since the energy is decreasing along the solution, we expect the latter to decay in a self-similar fashion.
Finally, when $M\chi > 8\pi$ the energy goes to $-\infty$ as $\lambda \to \infty$. This corresponds to a contraction of the density $\rho$. We expect the density to concentrate.

The proof that solutions cannot exist globally when $M\chi > 8\pi$ is usually done as followed. We introduce the second moment of the density $\Pi_2\left(t\right) = \int |x|^2\rho\left(t,x\right)dx$. We impose $\Pi_2\left(0\right) < +\infty$ and compute formally that

$$\frac{d}{dt}\Pi_2\left(t\right) = M\left(4 - \frac{M\chi}{2\pi}\right). \tag{4.2.1}$$

Indeed integration by parts and symmetrization give successively:

$$\begin{aligned}
\frac{d}{dt}\Pi_2\left(t\right) &= \int |x|^2\partial_t\rho\left(t,x\right)dx = \int |x|^2\left(\Delta\rho\left(t,x\right) - \chi\nabla\cdot\left(\rho\left(t,x\right)\nabla c\left(t,x\right)\right)\right)dx \\
&= -\int 2x\nabla\rho\left(t,x\right)dx + \chi\iint 2x\left(\rho\left(t,x\right)\frac{-1}{2\pi}\frac{1}{x-y}\rho\left(t,y\right)\right)dxdy \\
&= 4\int\rho\left(t,x\right)dx - \frac{\chi}{2\pi}\left(\int\rho\left(t,x\right)\right)^2 = M\left(4 - \frac{\chi M}{2\pi}\right)
\end{aligned}$$

In the supercritical case, $M\chi > 8\pi$, the second moment decreases linearly. Since the second moment is positive there is clearly an obstruction to global existence of smooth solutions. In fact the density blows-up in $L^\infty$ norm [65]. In this case the attraction term dominates. When $M\chi < 8\pi$, the solution exists for all time [39], but this cannot be proved directly using the second moment. Nevertheless the second moment increases linearly, and it can be proved that the density converges to zero with a diffusive self-similar decay [12].

The critical constant $8\pi$ leads us to two different interpretations. If the sensibility to the substance $\chi$ is fixed we naturally define the critical mass $M_\chi$ by

**Definition 4.2.1** (Critical mass)**.**

$$M_\chi = \frac{8\pi}{\chi}.$$

This point of view is natural in the biological context. The sensibility parameter $\chi$ is fixed. If there are enough cells a blow-up occurs; otherwise the diffusion wins.
The other point of view is to fix the total mass $M$ for example equal to one. It leads us to the definition of the critical parameter $\chi_M$ by:

**Definition 4.2.2** (Critical parameter)**.**

$$\chi_M = \frac{8\pi}{M}.$$

*In the particular case $M = 1$ we note $\chi_M = \chi_c$.*

This point of view will be the one used in our discrete model. From now on **the total mass $M$ is set equal to one.**

*Remark.* Variants of the Keller-Segel equation with nonlinear diffusion and nonlinear sensitivity have been studied [60, 94, 63, 84, 10]. For a particular choice of the nonlinearities it is possible to reproduce the peculiar critical mass phenomenon occurring in the two-dimensional Keller-Segel equation. For instance, Blanchet, Carrillo and Laurençot replaces the linear diffusion with a Porous-Medium type diffusion $\Delta\rho^m$ for the cell density. They prove that the critical exponent $m(n) = 2\dfrac{n-1}{n}$ yields a critical mass phenomenon.

### 4.2.a The mass quantization problem

One of the challenging question concerning the Keller-Segel problem is the structure of the blow-up, and in particular, the amount of mass concentrated at one point when blow-up occurs. This problem is called the mass quantization problem [104]. The main conjecture is that the blow-up is the formation of a Dirac mass that contains exactly the critical mass.

**Problem 4.2.3** (Picolo Graal)**.** *Let $\chi$ be fixed. In the supercritical regime $M > M_\chi$ the first singular point contains, generically, exactly the critical mass.*

This claim is proved by Suzuki in [104]. The approach is based on a suitable rescaling of the solution and estimates on the partial second moment. This conjecture has been tackled by many authors with a radial initial condition. Herrero and Velázquez exhibit in [58, 59] some radial solutions blowing-up with exactly the critical mass, Senba gives some sufficient conditions in [100]. Ohtsuka, Senba and Suzuki in [92] give a similar result for a system close to equation (4.1.1).
Another way to prove the conjecture would be to give sense to global measure solutions of the problem (4.1.1) in dimension 2. Velázquez [109] and Dolbeault-Schmeiser [40] give different candidates for measure-valued solutions extending (4.1.1) beyond the blow-up time. Dolbeault and

Schmeiser have developed Poupaud's framework on defect measures to the Keller-Segel equation. The strategy in any case is to regularise the equation and then pass to the limit. However the limit system depends upon the regularization procedure. We emphasize that Dolbeault and Schmeiser keep the gradient flow structure through the regularization procedure. The theory works very well, but in any case there is yet no answer to the mass quantization problem with this approach, except the easy side of the inequality: the singular point contains at least the critical mass.

Haškovec and Schmeiser propose in [55] a stochastic approximation of (4.1.1) defined beyond the blow-up time.

## 4.3 Log interaction in dimension one

Here we aim to simplify the problem in order to tackle Problem 4.2.3. We seek a simpler equation sharing similar features with the Keller-Segel equation (4.1.1) in dimension 2. Recall that in two space dimensions, the Green kernel is given by $\kappa(\cdot) = -\dfrac{1}{2\pi}\log(|\cdot|)$. We saw that the logarithmic interaction is critical for the homogeneity of the free energy (4.1.5). This motivates the definition of the $n$ dimensional log-interaction problem as follows:

$$\partial_t \rho(t,x) = \nabla \cdot \left[\rho(t,x)\nabla\left(\log(\rho(t,x)) - \chi\nabla\kappa \star \rho(t,\cdot)(t,x)\right)\right], (t,x) \in \mathbb{R}_+ \times \mathbb{R}^n, \tag{4.3.1}$$

where $\kappa(\cdot) = -\dfrac{1}{2\pi}\log(|\cdot|)$. The associated free energy is still log homogeneous and nonincreasing along the flow. Moreover in any dimension the same argumentation as in the previous section holds true, changing only the critical number $8\pi$ to $4n\pi$ [25]. The equation for $n \geq 3$ has no reason to be simpler than the case $n = 2$, so we shall consider the simpler case $n = 1$.

**Definition 4.3.1** (The one dimensional log interaction problem.)**.**

$$\begin{cases} \partial_t \rho(t,x) = \partial_x.\left(\rho(t,x)\partial_x \log(\rho(t,x)) - \chi\rho(t,x)\partial_x c(t,x)\right) & (t,x) \in \mathbb{R}_+ \times \mathbb{R}, \\ c(t,x) = -\dfrac{1}{2\pi}\displaystyle\int_{\mathbb{R}} \log|x-y|\rho(t,y)\,dy & (t,x) \in \mathbb{R}_+ \times \mathbb{R}. \end{cases} \tag{4.3.2}$$

*associated to the free energy*

$$E(\rho(t,\cdot)) = \int_{\mathbb{R}} \rho(t,x)\log(\rho(t,x))\,dx + \frac{\chi}{4\pi}\int_{\mathbb{R}^2} \log|x-y|\rho(t,x)\rho(t,x). \tag{4.3.3}$$

The one dimensional problem behaves like the two dimensional Keller-Segel equation, with the critical number $\chi = 4\pi$. However it is simpler to study. Indeed, it is generally not easy to prove the convergence or contraction of (4.1.3) in the Wasserstein space. In dimension one, Calvez and Carrillo [23] prove it for Equation (4.3.2) in the subcritical regime. In chapter 6, we prove it with similar techniques, and obtain that the free energy increases along the geodesic starting from a minimum.

## 4.4 Discrete gradient flows, deterministic particles scheme

In order to obtain a better understanding of the one dimensional log interaction problem (4.3.2), we study a particle model scheme in view of its approximation. This approach is used for the two dimensional Keller-Segel problem by Haškovec and Schmeiser in [55] and [56] where they study extensively the case of two and three particles. They define a particle model directly on (4.1.1), and use a Brownian motion to modelize the diffusion. In the one dimensional case we can take advantage of the explicit gradient flow structure (see Section 5.4). The basic idea is to first rewrite (4.3.2) with the inverse of the repartition function, denoted by $X(t,\cdot)$. Then we shall use a particle scheme: let $h_N = \dfrac{1}{N+1}$ be a space discretization step, then for $i = 0, \dots, N+1$ we define $X_i(t) = X\left(t, ih_N\right)$, with the convention $X_0(t) = -\infty$ and $X_{N+1}(t) = +\infty$. Alternatively speaking, we assign equal fractions of the mass to each particle.
It leads us to the following discrete gradient flow equation on the vector valued function $X(t) = (X_1(t), \dots, X_N(t))$. The discrete energy writes

$$\mathbb{E}\left(X\right) = -\sum_{i=1}^{N-1} \log\left(X_{i+1} - X_i\right) + \chi h_N \sum_{1 \leq i \neq j \leq N} \log|X_i - X_j|. \tag{4.4.1}$$

**Definition 4.4.1** (Gradient flow equation.)**.**

$$\begin{cases} \dot{X}\left(t\right) = -\nabla \mathbb{E}\left(X\left(t\right)\right) & t \in \mathbb{R} \\ X\left(0\right) = X_0 & X^0 \in \mathbb{R}^N, \end{cases} \tag{4.4.2}$$

This is the approach followed by Blanchet, Calvez and Carrillo in [8] and Devys in [37]. The former is mainly concerned with the subcritical regime. The authors prove the convergence of the scheme and describe the long time asymptotic behaviour of the solution. The latter builds a particle scheme to extend the solution beyond the blow-up time. We will study here the super critical case with the mass quantization problem in mind.
Remark that we do not describe the diffusion part of the flux using a Brownian motion. In dimension one, using a Brownian motion, the particles will surely cross in finite time and the interaction would be too singular. Then we should truncate the interaction kernel, and we would face the same troubles as for the continuous Keller-Segel equation beyond blow-up.
We emphasize the work of Filbet [51] who analyses a finite volume scheme for (4.1.1).

## 4.5 Our concerns

Now that we have discretized the problem using a particle scheme, the mass quantization problem translates into a problem of counting the number of particles. For this peculiar numerical scheme there is a minimal number of particles which is necessary for BU to occur. We ask whether the blow-up point contains the minimal number of particles, or more than the minimal number.
Since each particle carries a mass $\dfrac{1}{h_N}$ the critical mass problem can be formulated as follows:

**Problem 4.5.1** (Discrete mass quantization problem.)**.** *Let* $\chi_N^k = \dfrac{N+1}{k}$*. When* $\chi_N^k < \chi <$

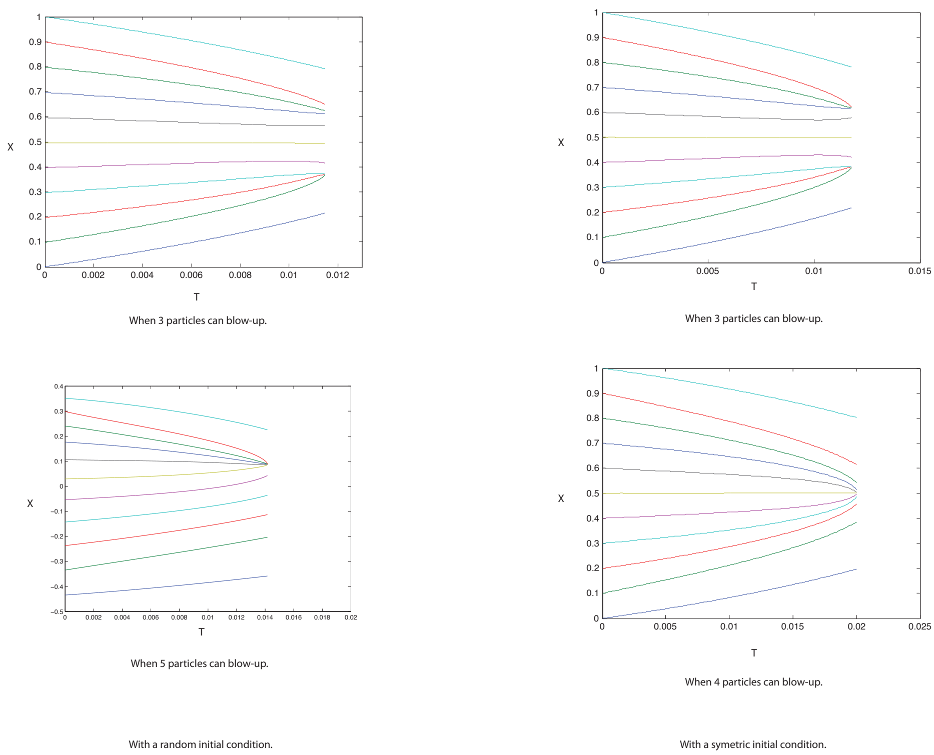



Figure 4.5.1: Numerical simulations of the discrete gradient flow (4.4.2) with 11 particles. (Left) The initial data is a perturbation of a symmetric configuration. We observe that the blow-up occurs merging the minimal number of particles, resp. 3 (top left) and 5 (bottom left). (Right) The initial data is symmetric. The minimal number of particles is resp. 3 (top right) and 4 (bottom right). We observe in the latter case that the first blow-up point contains 5 particles.

*$\chi_N^{k-1}$, the blow-up aggregates exactly $k$ particles.*

This claim calls for comments. The definition of $\chi_N^k$ is natural, it corresponds to the computation of some partial second moment (Equation (7.1.8)). The claim is wrong full generality. Indeed some non generic symmetric cases blow-up with too many particles; for example when $N = 3$ and $X_1 = -X_3$, $X_2 = 0$ the blow-up aggregates the three particles for any parameter $\chi$. Another numerical example is depicted in Figure 4.5.1. Our results concerning the discrete mass quantization problem are of two kinds.

1. **Stability of the profile.** We start by exhibiting some basins of attraction $A_\chi^N$ such that if a solutions enter in one of this basin of attraction then the blow-up will occur with exactly the critical number of particles. This stability result is the object of Section 7.4.
   A stability result is obtained for example by Velázquez in [108]. The author shows that a small perturbation of the initial data in the two dimensional Keller-Segel equation leads to a singularity which is close in time and location.

2. **Rigidity of the blow-up.** When a solution lying in a basin of attraction blows-up we get a very precise estimate on the behaviour which allows a parabolic rescaling of the solution. Then we prove that our rescaled solution is a very particular solution of the new system

and converges to a critical point of the partial $^{th}$ energy functional:

$$\mathbb{E}_k(Y) = -\sum_{i\in\mathcal{I}\backslash\{l+k-1\}} \log(Y_{i+1}-Y_i) + \chi h_N \sum_{(i,j)\in\mathcal{I}\times\mathcal{I}\backslash\{i\}} \log|Y_i - Y_j| - \frac{\alpha}{2}\sum_{i\in\mathcal{I}} |Y_i|^2 .$$

It means that the blow-up profile involves only the $k$ particles contributing to the blow-up. It is the concern of Section 7.5.

The parabolic rescaling is a very classical tool for the analysis of blow-up in the nonlinear heat equation (see e.g. [54] [89]). Senba and Suzuki use it in [101] to prove the convergence of some particular radial solution to a Dirac mass having equal to $8\pi$.

# Chapter 5

# The gradient flow point of view

We first recall the standard gradient flow equation and the definition of the Wasserstein space. Then we give an idea of how to define a gradient flow equation in a metric space, for example in $W_2^{ac}(\mathbb{R}^n)$ and especially in $W_2^{ac}(\mathbb{R})$.

**Some basics on the gradient flow equation.** A gradient flow equation is always related to an energy $E$ and a metric $d$. It is the steepest descent equation to a local minimum of the energy.

**Definition 5.0.1.** *Let $H$ be an Hilbert space and $E$: $H \to \mathbb{R}$ a $C^1$ energy functional. The gradient-flow equation reads:*

$$\dot{X}(t) = -\nabla E(X(t)), \quad t \in \mathbb{R}_+ \tag{5.0.1a}$$

$$X(0) = X_0, \quad X_0 \in H. \tag{5.0.1b}$$

Such an equation comes with a lot of structure. Some are encapsulated in the following proposition.

**Proposition 5.0.2.** *Let $X : \mathbb{R}_+ \to H$ be a solution of the gradient flow equation* (5.0.1) *in the sense of Definition 5.0.1 then:*

1. *$t \mapsto E(X(t))$ is nonincreasing.*
2. *Moreover if $E$ is strictly convex and achieves its unique minimum at the point $\overline{X}$, then $t \mapsto X(t)$ exists for all time, is unique and the function $t \mapsto ||X(t) - \overline{X}||_H$ is nonincreasing.*
3. *If in addition $E$ is $C^2$ and $\nabla^2 V(\xi, \xi) \geq \alpha \|\xi\|^2$ for some $\alpha > 0$ and for all $\xi$ then $||X(t) - \overline{X}||_H$ converges exponentially to $0$ with rate $\alpha$.*

*Proof.* The local existence is given by the Cauchy-Lipschitz theorem, thus $X$ is $C^1$ on its domain of definition. Therefore $\dfrac{d}{dt}E(X(t)) = \nabla E(X(t)) \cdot \dot{X}(t) = -||\dot{X}(t)||_H^2 \leq 0$, and therefore the

function $t \mapsto E(X(t))$ is nonincreasing.
If $E$ is strictly convex the minimizer is unique. Thanks to the convexity of $E$, we get that

$$\frac{d}{dt}||X(t) - \overline{X}||_H^2 = -\langle X(t) - \overline{X}, \nabla E(X(t)) - \nabla E(\overline{X})\rangle_H \leq 0.$$

The equality occurs if and only if $X(t) = \overline{X}$. In particular, $||X(t) - \overline{X}||_H^2$ stays in a compact subset fixed by $X_0$, and consequently exists for all time.
If, in addition $E$ is $C^2$, a Taylor expansion yields that $\frac{d}{dt}||X(t) - \overline{X}||_H^2 \leq -\alpha||X(t) - \overline{X}||_H^2$. Thus we get the exponential convergence of the gradient flow to the equilibrium state $\overline{X}$. □

The goal when defining a gradient flow for a metric space is to have the same kind of result as Proposition 5.0.2 in the metric space setting. In particular, we wish to keep the steepest descent idea.

## 5.1 The Wasserstein space

We recall that given $\mu$ and $\nu$ two absolute continue probability measures on $\mathbb{R}^n$, the optimal transport $T$ from $\mu$ to $\nu$, for the quadratic cost $c(x, y) = |y - x|^2$, is given by the gradient of a convex function $\varphi : \mathbb{R}^n \mapsto \mathbb{R}$ (see [3] [110] ). This function $T = \nabla\varphi$ is called the Brenier map. We also refer to chapter one for more details about the optimal transport and the Brenier map.

**Definition 5.1.1** (Wasserstein space)**.** *The Wasserstein space $W_2(\mathbb{R}^n)$ is the space of probability measures on $\mathbb{R}^n$ with finite second moment.*

$$W_2(\mathbb{R}^n) = \{\mu \in \mathbb{P}(\mathbb{R}^n) \ \textit{such that} \int |x|^2 d\mu(x) < +\infty\}.$$

*We will work only in the subset composed with absolutely continuous measure (with respect to the Lebesgue measure).*

$$W_2^{ac}(\mathbb{R}^n) = \{\mu \in \mathbb{P}_{ac}(\mathbb{R}^n) \ \textit{such that} \int |x|^2 \mu(x)dx < +\infty\}.$$

*For any $\mu, \nu \in W_2^{ac}(\mathbb{R}^n)$ we define the Wasserstein distance $d_W(\mu, \nu)$ by the optimal transport for the quadratic cost $c(x, y) = |y - x|^2$:*

$$d_W^2(\mu, \nu) = \inf_{\mathbf{t}} \left\{ \int |\mathbf{t}(x) - x|^2 d\mu(x); \quad \mathbf{t}_{\#}\mu = \nu \right\} = \int |T(x) - x|^2 d\mu(x),$$

*where the infimum is taken over all transport map pushing forward $\mu$ onto $\nu$. The infimum is reached for $T$ the Brenier's map for the transport. This distance is in fact well defined on $W_2(\mathbb{R}^n)$, if we consider the Kantorovich formulation of the optimal transport. Equipped with this distance, $W_2^{ac}(\mathbb{R}^n)$ is a metric space [110].*

Thanks to the work of Otto [93], Villani [110], Sturm [103, 102], followed by Ambrosio, Gigli and Savare [1], we can equip $W_2^{ac}(\mathbb{R}^n)$ with a Riemannian structure where the optimal transport paths are the geodesic paths. We present here the formal formulation due to Otto.

Let $\rho, \mu \in W_2^{ac}$ we define

$$d_W^2(\mu, \rho) = \inf \left\{ \int_0^1 ||\dot{\mu}(s)||^2_{\mu(s)}\, ; \quad \mu(0) = \mu,\ \mu(1) = \rho \right\},$$

where the infimum is taken over all absolute continuous paths $\mu(s)$ joining $\mu$ and $\rho$, and the Lagrangian action is given by

$$||\dot{\mu}||^2_\mu = \inf_\xi \left\{ \int ||\xi||^2 \mu(x)dx; \quad \dot{\mu} + \nabla \cdot (\xi\mu) = 0 \right\}, \tag{5.1.1}$$

where the infimum is taken over all vector fields $\xi : (0,1) \times \mathbb{R}^n$ such that the path $\mu$ satisfies the continuity equation $\partial_t \mu + \nabla \cdot (\xi\mu) = 0$. The vector field $\xi$ encodes the speed of the particles.
The Benamou-Brenier formula [3] shows that the two above definitions for $d_W^2$ coincide. The first one is the Lagrangian viewpoint (points are transported), whereas the second one is the eulerian viewpoint (densities are moving).
Moreover when $\xi$ realizes the minimum in (5.1.1) we see that there exists $v$, such that $\nabla v = \xi$. It leads us to the definition of the metric tensor $g$. For any two tangent vectors $\dot{\mu}_a$, $\dot{\mu}_b$ at $\mu$, let $v_a, v_b$ such that $\dot{\mu}_a = -\nabla \cdot (\mu \nabla v_a)$, $\dot{\mu}_b = -\nabla \cdot (\mu \nabla v_b)$ then

$$g_\mu(\dot{\mu}_a, \dot{\mu}_b) = \int \mu(x) \nabla v_a(x) \nabla v_b(x) dx.$$

A remarkable fact is the link with optimal transport: the optimal paths for the quadratic cost are the geodesic paths.

**Definition 5.1.2.** *Let $\rho_0, \rho_1 \in W_2^{ac}(\mathbb{R}^n)$ and $T$ the optimal transport from $\rho_0$ to $\rho_1$. We define the optimal transport path: $s \in [0,1] \mapsto \rho_s$ by*

$$\rho_s = ((1-s)\, Id + sT)_\# \,\rho_0 = (T_s)_\# \rho_0 = [\rho_0, \rho_1]_s.$$

This path is the McCann interpolation introduced in [87]. For $\rho_s$ we have the following properties

1. $\sqrt{d_W^2(\rho_0, \rho_s)} = s\sqrt{d_W^2(\rho_0, \rho_1)}$.
2. $s \in [0,1] \mapsto \rho_s$ is a geodesic path.

The first property is a straightforward consequence of the fact that $T_s$ is the Brenier map for the transport from $\rho_0$ to $\rho_s$. For the second we need to differentiate $s \in [0,1] \mapsto \rho_s$ and identify the tangent vector. For any smooth test function with compact support $\varphi$: $\mathbb{R}^n \to \mathbb{R}$ we have by definition of $T_s$:

$$\begin{aligned} \frac{d}{ds} \int \varphi(y) \rho_s(y) dy = \frac{d}{ds} \int \varphi(T_s(x)) \rho_0(x) dx &= \int \nabla\varphi(T_s(x)) (T(x) - x) \rho_0(x) dx \\ &= \int \nabla\varphi(y) \left((T - Id)(T_s^{-1})(y)\right) \rho_s(y) dx \\ &= -\int \varphi(y) \nabla \cdot \left((T - Id)(T_s^{-1})(y) \rho_s(y)\right) dy. \end{aligned}$$

Thus $\dot\rho_s = (T - Id) \circ T_s^{-1}$ and $g_{\rho_s}(\dot\rho_s, \dot\rho_s) = \displaystyle\int \rho_0(x)\,(T(x) - x)^2\,dx$. We deduce that $\rho_s$ is a geodesic path.

## 5.2 Gradient flow in metric spaces

### 5.2.a A direct approach

The Riemannian interpretation of the Wasserstein space allows us to differentiate and thus define the notion of gradient. Let $E : W_2^{ac} \to \mathbb{R}$ be an energy functional. We define $\nabla_W E$ considering all the absolute continuous paths $s \in [0,1] \mapsto \rho(t)$ and using the following identity for any $s \in ]0,1[$:

$$g_{\rho(s)}\left(\nabla_W E(\rho(s)), \partial_s \rho(s)\right) = \frac{d}{ds} E(\rho)(s). \tag{5.2.1}$$

We then formally define the gradient flow equation by

$$\partial_t \rho = -\nabla_W E(\rho).$$

In the Wasserstein case it leads us to the following definition of $\nabla_W E(\rho)$.

**Proposition 5.2.1.** *Let $E : W_2^{ac}(\mathbb{R}^n) \to \mathbb{R}$ be an energy functional and $\rho \in W_2^{ac}$ then*

$$-\nabla_W E(\rho) = \nabla \cdot \left(\rho \nabla \frac{\delta E}{\delta \rho}\right),$$

*where $\dfrac{\delta E}{\delta \rho}$ is the formal $L^2$ gradient of $E$.*

*Proof.* Let $s \in [0,1] \mapsto \rho(s)$ be an absolute continuous path, we note $v$ the vector such that $\partial_s \rho = -\nabla \cdot (\rho \nabla v)$, we need to formally identify a vector $w$ such that

$$\nabla_W E(\rho) = -\nabla \cdot (\rho \nabla w). \tag{5.2.2}$$

On one hand, by definition of $\dfrac{\delta E}{\delta \rho}$ we have

$$\begin{aligned}\frac{d}{ds} E(\rho)(s) &= \int \frac{\delta E}{\delta \rho} \partial_s \rho = -\int \frac{\delta E}{\delta \rho} \nabla \cdot (\rho \nabla v) \\ &= \int \nabla \frac{\delta E}{\delta \rho} (\rho \nabla v).\end{aligned}$$

On the other hand, by definition of $g_\rho$, and since $w$ is sought so as to satisfy (5.2.2), we have

$$g_\rho\left(\nabla_W E(\rho), \partial_s \rho\right) = \int \rho \nabla v \nabla w.$$

By (5.2.1) we may then identify $w = \dfrac{\delta E}{\delta \rho}$ and therefore

$$\nabla_W E(\rho) = -\nabla \cdot \left(\rho \nabla \frac{\delta E}{\delta \rho}\right).$$

□

This proposition leads to the following definition of a gradient flow equation.

**Definition 5.2.2.** *Let $E : W_2^{ac}(\mathbb{R}^n) \to \mathbb{R}$ be an energy functional and $T \in \mathbb{R}^+$. The function $t \in [0, T[ \mapsto \rho(t) \in W_2^{ac}$ is a solution of the gradient flow associated to $E$ if for any $t \in [0, T[$:*

$$\frac{\partial \rho}{\partial t} = -\nabla_W E(\rho(t)) = \nabla \cdot \left(\rho(t) \nabla \frac{\delta E}{\delta \rho}(\rho(t))\right), \quad t, x \in [0, T[ \times \mathbb{R}^n \tag{5.2.3a}$$

$$\rho(0, x) = \rho_0(x), \;\; x \in \mathbb{R}^n. \tag{5.2.3b}$$

Since we define a gradient flow we hope that the energy is nonincreasing along the flow. This is indeed the case, as stated in the following proposition.

**Proposition 5.2.3.** *Let $\rho : [0, T[ \to W_2^{ac}(\mathbb{R}^n)$ be a a gradient flow equation in the sense of Definition 5.2.2, then the function $t \mapsto E(\rho(t, \cdot))$ is nonincreasing.*

*Proof.*

$$\begin{aligned} \frac{d}{dt} E(\rho(t, \cdot)) &= \int \partial_t \rho(t, x) \frac{\delta E}{\delta \rho} = \int \nabla \cdot \left(\rho(t, x) \nabla \frac{\delta E}{\delta \rho}\right) \frac{\delta E}{\delta \rho} \\ &= -\int \rho(t, x) \left(\nabla \frac{\delta E}{\delta \rho}\right)^2 \leq 0. \end{aligned}$$

□

### 5.2.b A metric approach

We give another point of view to define a gradient flow in metric spaces, referring to [110] for details. The idea is to find a formulation which does not use the Hilbert structure. This is the aim of the following lemma.

**Lemma 5.2.4.** *Let $H$ be an Hilbert space, and let $E$: $H \to \mathbb{R}$ be an energy functional; then the function $t \in [0, 1] \mapsto X(t)$ is a gradient flow for $E$ (that is, a solution of* (5.0.1a)*) if and only if*

$$\forall x \in H, \forall t \in [0, 1], \frac{d}{dt} \frac{1}{2} ||X(t) - x||^2 = \left.\frac{d}{ds}\right|_{s=0} E((1 - s)X(t) + sx).$$

The proof is an easy computation. Be careful the right hand side is a derivative along a geodesic path, whereas the left hand side is a derivative along the flow. It leads us to another definition for a gradient flow equation, for example in $W_2^{ac}(\mathbb{R}^n)$.

**Definition 5.2.5.** *Let $E$: $W_2^{ac}(\mathbb{R}^n) \to \mathbb{R}$ be the energy. We say that $t \in [0, T[ \mapsto \rho_t$ is a solution of the gradient flow equation relative to $E$ if for any $\mu \in W_2^{ac}(\mathbb{R}^n)$ and $t \in [0, T[$:*

$$\frac{d}{dt}\frac{1}{2}d_W^2(\mu, \rho_t) = \left.\frac{d}{ds}\right|_{s=0} E\left([\rho_t, \mu]_s\right), \tag{5.2.4}$$

*where $[\rho_t, \mu]_s = (T_s)_{\#}\rho_t$ is the displacement interpolation defined by* (5.1.2). *In particular $(T_1)_{\#}\rho_t = \mu$.*

Note that this latter definition does not require the notion of gradient; it gives in $W_2^{ac}$ the same definition than that of Definition 5.2.2.

**Lemma 5.2.6.** *The definitions 5.2.2 and 5.2.5 are equivalent.*

*Proof.* Let us compute $\frac{d}{dt}d_W^2(\mu, \rho_t)$. The first variation formula in a Riemannian manifold (2.2.1) gives

$$\frac{d}{dt}\frac{1}{2}d_W^2(\mu, \rho_t) = -g_{\rho_t}\left(\left.\frac{d}{ds}\right|_{s=0}[\rho_t, \mu]_s, \frac{d}{dt}\rho_t\right) = -\int \rho_t(x)\left(T_1(x) - x\right)\cdot \xi_t(x)dx, \tag{5.2.5}$$

where $\xi_t$ satisfies $\frac{d}{dt}\rho_t = -\nabla(\rho_t\xi_t)$ and $(Id + T_1 - Id)_{\#}\,\rho_t = \mu$. We recall that we compute the derivative of a geodesic path in Section 5.1. Note that (5.1) is merely the equivalent of the classical formula

$$\frac{d}{dt}\frac{1}{2}||x - X(t)||^2 = -\left\langle x - X(t), \dot{X}(t)\right\rangle$$

in a Riemannian setting. The geodesic path in this case is $X(t) + s(x - X(t))$.
Let us turn to the energy; by definition we have

$$\begin{aligned}\left.\frac{d}{ds}\right|_{s=0} E\left([\rho_t, \mu]_s\right) &= \int \frac{\delta E}{\delta\rho}\left.\frac{d}{ds}\right|_{s=0}\left([\rho_t, \mu]_s\right) = -\int \frac{\delta E}{\delta\rho}\nabla\left(\rho_t(x)\left(T_1(x) - x\right)\right)dx \\ &= \int \nabla\frac{\delta E}{\delta\rho}\cdot\left(T_1(x) - x\right)\rho_t(x))dx.\end{aligned}$$

If we take Definition 5.2.5 then we can do the computation for any $\mu$, thus $\xi_t = -\nabla\frac{\delta E}{\delta\rho}$ which is exactly Definition 5.2.2. If we take the Definition 5.2.2 then for any $\mu = (T_1)_{\#}(\rho_t)$, we obtain (5.2.4). □

**Remark 5.2.7** (Rigorous definition of the gradient flow)**.** *It is not so simple to give a rigorous sense to Definition 5.2.5. In particular we have avoided here the regularity issue. For example the Wasserstein distance and the energy functional are usually not plainly differentiable. Therefore the definition needs to be modified using some limsup and inequalities. For details we refer for instance to [110],[1].*

### 5.2.c A numerical approach

A third way to define gradient flow in a metric space consists in another reformulation using only the metric structure. This approach is based on a time implicit Euler scheme. We fix a constant time step $\tau$ and discretize the gradient flow equation (5.0.1) with an Euler implicit scheme:

$$\frac{X_\tau^{n+1} - X_\tau^n}{\tau} = -\nabla E\left(X_\tau^{n+1}\right), \tag{5.2.6a}$$

$$X_\tau^0 = X_0. \tag{5.2.6b}$$

Under suitable conditions on $E$, for instance $C^2$, the solution $X_\tau$ of (5.2.6) converges when $\tau$ goes to 0 to the unique solution of (5.0.1). In order to generalize the gradient flow to general metric spaces, we need a formulation which does not involve the gradient. This is obtained thanks to the following result, which gives a minimization interpretation of the scheme (5.2.6a).

**Lemma 5.2.8.** *Let $E \in C^1(H,\mathbb{R})$. A solution of*

$$X_\tau^{n+1} = \underset{X\in H}{\operatorname{argmin}} \left\{E\left(X\right) + \frac{1}{2\tau}||X - X_\tau^n||_H^2\right\}, \tag{5.2.7a}$$

$$X^0 = X_0. \tag{5.2.7b}$$

*is a solution of* (5.2.6).

*Proof.* When the function $X \mapsto E\left(X\right) + \frac{1}{2\tau}||X - X_n||_H^2$ reaches its minimum, we have $\nabla E\left(X\right) = -\frac{1}{\tau}\left(X - X_n\right)$ which is exactly (5.2.6) since at the minimum $X = X_\tau^{n+1}$. □

Reciprocally we have the following lemma.

**Lemma 5.2.9.** *Let $E \in C^2(H,\mathbb{R})$ with $||E||_{C^2} < +\infty$. There exists $\overline{\tau}$ such that for any $\tau \leq \overline{\tau}$ a solution of* (5.2.6) *is a solution of* (5.2.7).

*Proof.* For $\tau$ small enough the function $X \mapsto E\left(X\right) + \frac{1}{2\tau}||X - X_n||_H^2$ is uniformly strictly convex. Indeed the Hessian is given by $\nabla^2 E + \frac{1}{\tau}Id$. Thus the minimum exist and is achieved only once at $X_\tau^{n+1}$. This point must satisfies $\nabla E\left(X_\tau^{n+1}\right) = -\frac{1}{\tau}\left(X_\tau^{n+1} - X_n\right)$ which is exactly (5.2.6). □

Thanks to this remark, Jordan, Kinderleher, and Otto [66] proposed a scheme to minimize an energy in $W_2^{ac}$.

**Lemma 5.2.10** (JKO)**.** *Let $E : W_2^{ac}\left(\mathbb{R}^n\right) \to \mathbb{R}$, for $\tau \in \mathbb{R}^+$ we define $\rho_\tau$ by the steepest descent scheme:*

$$\rho_\tau^{n+1} = \underset{\rho\in W_2^{ac}}{\operatorname{argmin}} \left\{E\left(\rho\right) + \frac{1}{2\tau}d^2_{W_2^{ac}}\left(\rho,\rho_\tau^n\right)\right\},, \tag{5.2.8a}$$

$$\rho_\tau^0 = \rho_0, \quad \rho_0 \in W_2^{ac}\left(\mathbb{R}^n\right). \tag{5.2.8b}$$

The next step is to find a solution defined on a time interval. For simplicity in the passage to the limit when $\tau$ tends to 0, we extend $\rho_\tau$ on $[n\tau,(n+1)\tau]$ thanks to the optimal transport. Let $T_n$ be the Brenier map from $\rho_\tau^n$ to $\rho_\tau^{n+1}$, then, for any $s \in [n\tau,(n+1)\tau]$ we define $\rho_\tau(s)$ by

$$\rho_\tau(s) = \left((1-s)Id + sT_n\right)_\# \rho_\tau^n = \left[\rho_\tau^n, \rho_\tau^{n+1}\right]_s.$$

Therefore $\rho_\tau \in W_2^{ac}(\mathbb{R}^n)$. When $\lim_{\tau\to 0} \rho_\tau$ exists, we denote it by $\rho_\infty$, and then $\rho_\infty$ is the solution of the formal gradient flow equation for the energy $E$

$$\partial_t \rho = -\text{“}\nabla_{W_2^{ac}}\text{”} E, \tag{5.2.9}$$

where "$\nabla_{W_2^{ac}}$"$E$ is obtained by passing to the limit in the JKO scheme. This strategy is the original approach to define a gradient flow in general metric spaces. The definitions here are only formal; in each particular case, each step must be proven to be well defined and convergence must be proven. The rigorous analysis was for instance performed for the Patlak-Keller-Segel equation by Blanchet Calvez and Carillo in [8]. Let us finally formally check that Definition 5.2.8 is equivalent to Definition 5.2.2. By (5.2.5), we know how to differentiate $d_W^2$; therefore at the minimum $\rho$, we exactly find the time implicit Euler scheme of Definition 5.2.2.

## 5.3 Examples

For a first example we can take $E(\rho) = \displaystyle\int \rho \log \rho$, which allows to see the standard heat equation $\partial_t \rho = \Delta\rho$ as a gradient flow for the Boltzmann entropy.

We give in the next paragraph a very particular example illuminating the different points of view.

### Eulerian or Lagrangian ?

We consider the Eulerian equation

$$\partial_t \rho = \nabla \cdot (\rho_t \nabla V), \tag{5.3.1}$$

where $V$ satisfies $\nabla \cdot \nabla V = 0$ ( incompressible hypothesis). We construct the characteristics $t \mapsto X(t)$ of (5.3.1) to find the Lagrangian formulation:

$$\dot{X}(t) = -\nabla V(X(t)). \tag{5.3.2}$$

In the Lagrangian point of view we follow the particles, whereas in the Eulerian case we study the flow at a fixed point. To go from one equation to the other we use the formula

$$\rho(t, X(t)) = \rho_0 (X(0)).$$

We remark that in the Lagrangian setting, we have a usual gradient flow equation. The Otto computation transfers this structure to the Eulerian formulation; it can be seen as the Eulerian gradient flow . The Otto interpretation allows a definition of the trajectory of the particles thanks to the optimal transport.

### Displacement convexity

A large class of examples, and in particular the Keller-Segel free energy, is given by energies of the form

$$E(\rho) = \int U(\rho(x))\,dx + \int V(x)\,\rho(x)\,dx + \frac{1}{2}\int W(x-y)\,\rho(x)\,\rho(y)\,dxdy, \tag{5.3.3}$$

where $U : \mathbb{R}_+ \mapsto \mathbb{R}$ is the internal energy, $V : \mathbb{R}^n \mapsto \mathbb{R}$ a potential and $W : \mathbb{R}^n \mapsto \mathbb{R}$ a symmetric interaction potential. In this case we can compute $\dfrac{\delta E}{\delta \rho}$ thanks to the derivative of the function $s \mapsto E(\rho(s))$, for any path $s \mapsto \rho(s)$. We first remark that

$$\begin{aligned}
&\frac{d}{ds}\int U(\rho(s,x))\,dx = \int U'(\rho(s,x))\,\partial_s\rho(s,x)\,dx,\\
&\frac{d}{ds}\int V(x)\,\rho(s,x)\,dx = \int V(x)\,\partial_s\rho(s,x)\,dx,\\
&\frac{d}{ds}\frac{1}{2}\int W(x-y)\,\rho(s,x)\,\rho(s,y)\,dxdy = \int (W\star\rho(s,\cdot))(x)\partial_s\rho(s,x)\,dx,
\end{aligned}$$

using the symmetry of $W$ in the last equation. Thus

$$\frac{\delta E}{\delta\rho} = U'(\rho) + V + W\star\rho.$$

And the gradient flow equation (5.2.3) may be rewritten

$$\frac{\partial\rho}{\partial t} = \nabla\cdot\big(\rho\nabla U'(\rho) + \rho\nabla V + \rho(\nabla W\star\rho)\big).$$

For such a functional McCann introduced in [87] the concept of displacement convexity.

**Definition 5.3.1** (Displacement convexity). *Let $\rho_0, \rho_1 \in W_2^{ac}(\mathbb{R}^n)$, $\rho_0 \neq \rho_1$, and $T$ be the optimal transport map from $\rho_0$ to $\rho_1$. We define the geodesic path $s \in [0,1] \mapsto \rho_s$ by*

$$\rho_s = ((1-s)\,Id + sT)_{\#}\,\rho_0 = (T_s)_{\#}\rho_0 = [\rho_0,\rho_1]_s.$$

*An energy functional for example* (5.3.3) *is said to be displacement convex (resp. strictly displacement convex) if it is convex (resp. strictly convex) along the geodesic $s \mapsto \rho_s$:*

$$E(\rho_s) \leq (1-s)\,E(\rho_0) + sE(\rho_1) \quad (\textit{resp. } E(\rho_s) < (1-s)\,E(\rho_0) + sE(\rho_1)).$$

*If moreover $E$ is $C^2$ then $E$ is said to be $\lambda$ uniformly displacement convex (for some $\lambda > 0$) if for any $\rho_0, \rho_1 \in W_2^{ac}(\mathbb{R}^n)$, $\rho_0 \neq \rho_1$, $s \in (0,1)$:*

$$\frac{d^2}{ds^2}(E(\rho))(s) \geq \lambda d_W^2(\rho_0,\rho_1).$$

According to this definition McCann proved the following useful theorem [87].

**Theorem 5.3.2.** *Let $E$ an energy functional defined by* (5.3.3). *Assume that*

1. $\lambda \mapsto \lambda^n U\left(\lambda^{-n}\right)$ *is convex nonincreasing on* $\mathbb{R}_+$.

2. $V$ *is convex.*

3. $W$ *is symmetric and convex.*

*Then the energy* (5.3.3) *is displacement convex; more precisely, each of the three terms defining* $E$ *in* (5.3.3) *is displacement convex.*

The converse is also true, in the sense that if one of the three terms in (5.3.3) is displacement convex then the corresponding assumption in Theorem 5.3.2 holds true. We have an analogue result with strict convexity.

*Proof.* We are going to prove that each of the three terms of $E$ are displacement convex. The key tool is the change of variable formula given by the Monge-Ampère equation

$$\det \nabla T(x) = \frac{\mu(x)}{\nu(T(x))}. \tag{5.3.4}$$

Let $s \mapsto \rho_s$ be a geodesic path.
**The potential term**: Thanks to (5.3.4),

$$\int V(x)\,\rho_s(x)\,dx = \int V(T_s(x))\,\rho_0(x)\,dx.$$

By convexity of $V$,

$$V(T_s(x)) = V((1-s)\,x + sT(x)) \leq (1-s)\,V(x) + sV(T(x)).$$

Thus

$$\begin{aligned}\int V(x)\,\rho_s(x)\,dx &\leq (1-s)\int V(x)\,\rho_0(x)\,dx + s\int V(T(x))\,\rho_0(x)\,dx \\ &\leq (1-s)\int V(x)\,\rho_0(x)\,dx + s\int V(y)\,\rho_1(y)\,dy.\end{aligned}$$

**The internal energy term**: Let $\lambda(s) = (\det \nabla T_s(x))^{\frac{1}{n}}$. Again using the Monge-Ampère equation (5.3.4),

$$\begin{aligned}\int U(\rho_s(x))\,dx &= \int \lambda^n(s)\,U\left(\lambda^n(s)\,\frac{1}{\lambda^n(s)}\rho_s(T(x))\right)dx \\ &= \int \lambda^n(s)\,U\left(\rho_0(x)\,\frac{1}{\lambda^n(s)}\right)dx.\end{aligned}$$

Since $\det^{\frac{1}{n}}$ is a concave function, we remark that $\lambda(s) \geq (1-s) + s\,(\det \nabla T(x))^{\frac{1}{n}} = (1-s)\,\lambda(0) + s\lambda(1)$. Since $\lambda \mapsto \lambda^n U\left(\lambda^{-n}\right)$ is nonincreasing and convex, we get

$$\lambda^n(s)\,U\left(\rho_0(x)\,\frac{1}{\lambda^n(s)}\right) \leq (1-s)\,\lambda^n(0)\,U\left(\rho_0(x)\,\frac{1}{\lambda^n(0)}\right) + s\lambda^n(1)\,U\left(\rho_0(x)\,\frac{1}{\lambda^n(1)}\right).$$

As $\lambda^n(0) = 1$, a backward change of variable on the last term yields:

$$\int U(\rho_s(x))\,dx \le (1-s)\int U(\rho_0(x))\,dx + s\int U(\rho_1(x))\,dx.$$

**The interaction potential term**: The displacement convexity of this last term is again a straightforward consequence of the convexity.

$$\begin{aligned}\int W(x-y)\,\rho_s(x)\,\rho_s(y)\,dxdy &= \int W((1-s)(x-y)+s(T(x)-T(y)))\,\rho_0(x)\,\rho_0(y)\,dxdy\\ &\le (1-s)\int W((x-y))\,\rho_0(x)\,\rho_0(y)\,dxdy + s\int W((T(x)-T(y)))\,\rho_0(x)\,\rho_0(y)\\ &\le (1-s)\int W((x-y))\,\rho_0(x)\,\rho_0(y)\,dxdy + s\int W((x-y))\,\rho_1(x)\,\rho_1(y).\end{aligned}$$

□

This theorem is very important, since displacement convexity is the equivalent in the Wasserstein case of the usual convexity. For example, we have the following proposition, analogue to Proposition 5.0.2 in a Hilbert space.

**Proposition 5.3.3.** *Let $\rho : \mathbb{R}_+ \to W_2^{ac}(\mathbb{R}^n)$ be a gradient flow in the sense of Definition 5.2.2 then:*

1. *$t \mapsto E(\rho(t,\cdot))$ is nonincreasing.*
2. *Moreover if $E$ is strictly displacement convex, the minimum can be achieved only once, say at $\overline{\rho}$. In this case the function $t \mapsto d^2_{W_2}(\rho(t,\cdot),\overline{\rho})$ is nonincreasing.*
3. *If in addition $E$ is $\lambda$ displacement convex for some $\lambda > 0$, then $d^2_{W_2^{ac}}(\rho(t,\cdot),\overline{\rho})$ converges exponentially to $0$.*

Again this is a formal proposition, and we forget about regularity or existence issues in the proof.

*Proof.* We already saw in Proposition 5.2.3 that the energy is nonincreasing. If $E$ is strictly displacement convex and reaches his minimum in two different points $\rho_0, \rho_1$ then for any $s \in ]0,1[$,

$$E([\rho_0,\rho_1]_s) < \min E.$$

This is a contradiction, and therefore the minimum is unique.
In order to show that $t \mapsto d^2_{W_2}(\rho(t,\cdot),\overline{\rho})$ is nonincreasing and converges exponentially to 0 if $E$ is $\lambda$ displacement convex, it is more convenient to work with Definition 5.2.5. We only consider here the case $E$ $\lambda$ displacement convex, the other part of the proof is similar.
Since $E$ is $\lambda$ is assumed to be displacement convex,

$$\frac{d}{dt}\frac{1}{2}d^2_W(\overline{\rho},\rho_t) \le \left.\frac{d}{ds}\right|_{s=0} E([\rho_t,\overline{\rho}]_s) = -\int_0^1 \frac{d^2}{ds^2}E([\rho_t,\overline{\rho}]_s)\,ds \le -\lambda d^2_W(\overline{\rho},\rho_t).$$

Therefore we obtain the exponential convergence (using the inequality $f' \le -2\lambda f$ for $f = d^2_W$). □

**Remark 5.3.4.** *In general, other gradient flow equations may be well defined without satisfying* (5.3.2). *Moreover we can have the $\lambda$ contraction without $\lambda$ displacement convexity: when $\mu$ is an equilibrium state for $E$ and $\rho_s = [\rho, \mu]_s$ we only need that $\int_0^1 \frac{d^2}{ds^2} E(\rho_s)\, ds \geq \lambda d_w^2(\rho, \mu)$. For example the function can be only locally displacement convex around $\mu$, the "BMX handlebar" depicted in the opposite figure is a good example of such a shape for $E$. The BMX handlebar is not $\lambda$ displacement convex, it is the case only near the minimum $\mu$. However if $\rho_t$ is solution of the associated gradient flow then $d_W^2(\mu, \rho_t)$ converges exponentially to zero. We conjecture that this is the case for the Keller-Segel energy.*

The BMX handlebar.

Let us check that the Keller-Segel energy is of the form (5.3.3); indeed, expressing $c$ with the fundamental solution of $-\Delta$. We set $U(\rho) = \rho \log(\rho)$, $V = 0$ and $W$ the fundamental solution of $-\Delta$. In dimension 1, with the log-interaction we get

$$\begin{cases} \partial_t \rho(t,x) = \nabla.\left(\rho(t,x) \nabla \log(\rho(t,x)) - \chi \rho(t,x) \nabla c(t,x)\right) & (t,x) \in \mathbb{R}^2, \\ c(t,x) = -\frac{1}{2\pi} \int_{\mathbb{R}} \log|x-y| \rho(t,y)\, dy & (t,x) \in \mathbb{R}^2. \end{cases} \tag{5.3.5}$$

associated to the free energy

$$E = \int_{\mathbb{R}} \rho \log(\rho) + \frac{\chi}{4\pi} \int_{\mathbb{R}^2} \log|x-y| \rho\rho. \tag{5.3.6}$$

**Remark 5.3.5.** *For simplicity we worked with $\mathbb{R}^n$, in general one can take a Riemannian manifold $M$ instead or $\mathbb{R}^n$ and define gradient flow equation in $W_2(M)$ see for instance [110].*

## 5.4 The particular one dimensional case

We show in this section that we can express a gradient flow in $W_2^{ac}(\mathbb{R})$ as a real gradient flow in $L^2$ by a suitable change of variable. The key point is the following lemma.

**Lemma 5.4.1.** *There is an isometric injection $i$ from $W_2^{ac}(\mathbb{R})$ into $L^2(0,1)$. Moreover $i(W_2^{ac}(\mathbb{R}))$ is a convex subset of $L^2(0,1)$.*

*Proof.* Let $\mu \in W_2^{ac}(\mathbb{R})$ we define the cumulative mass $M_\mu : \mathbb{R} \to [0,1]$ by

$$M_\mu(x) = \mu((-\infty, x)). \tag{5.4.1}$$

The function $M_\mu$ is obviously a non decreasing function, and we can thus define the pseudo–inverse $X_\mu : [0,1] \to \mathbb{R}$ by

$$X_\mu(m) = \inf\{x \in \mathbb{R} \text{ such that } M_\mu(x) > m\}. \tag{5.4.2}$$

By definition, the function $X_\mu$ is a right semi continuous non decreasing function and $X_\mu(m)$ is the position at which we obtain a cumulative mass $m$ for $\mu$. Note that $X_\mu$ also gives a change of variable formula.

**Lemma 5.4.2.** *Let $\mu$ be fixed and $X_\mu$ the pseudo–inverse of the cumulative mass. Let $h:\mathbb{R}\to\mathbb{R}$ then*

$$\int_{X_\mu(a)}^{X_\mu(b)} h(x)\,d\mu(x) = \int_a^b h(X_\mu(m))\,dm$$

*Proof.* For convenience we suppose that $\mu$ is given by a $C^\infty$ positive density function $f:\mathbb{R}\to\mathbb{R}^{*+}$, that is $\mu(x) = f(x)\,dx$. Therefore the function $M_\mu$ defined by (5.4.1) is a $C^\infty$ primitive of $f$ and $X$ defined by (5.4.2) is the usual inverse; thus $X_\mu'(M_\mu(x)) = \dfrac{1}{f(x)}$, it gives (5.4.2) in this case. The result follows by a density argument. □

Since $\mu$ has a finite second moment, Lemma 5.4.2 applied with $h = |.|^2$, $a = -\infty = -b$ gives $X_\mu \in L^2(0,1)$. We may then define the injection map $i: W_2^{ac}(\mathbb{R}) \to L^2(0,1)$ by $i(\mu) = X_\mu$. It remains to prove that $i$ is isometric.
Let $\mu = fdx$ and $\nu = gdx$ be two $C^\infty$ positive density measures. We have

$$d_W^2(\mu,\nu) = \int_{-\infty}^{\infty} |x - T(x)|^2 d\mu(x)$$

where $T$ is the optimal transport from $\mu$ to $\nu$. In dimension 1 the transport $T$ is easy to expresses with the cumulative mass and its inverse see [3].

$$T(x) = X_\nu(M_\mu(x)) = i(\nu)\left[i(\mu)^{-1}(x)\right].$$

To prove it one can see that $T_\#\mu = \nu$ and $T$ is nonincreasing, therefore by uniqueness $T$ is the optimal transport. Applying Lemma 5.4.2, we get that

$$\begin{aligned} d_W^2(\mu,\nu) &= \int_0^1 |i(\mu)(m) - T(i(\mu)(m))|^2 dm = \int_0^1 |i(\mu)(m) - i(\nu)(m)|^2 dm \\ &= ||i(\mu)(m) - i(\nu)(m)||_{L^2} = ||X_\mu(m) - X_\nu(m)||_{L^2}. \end{aligned}$$

□

We can go further and identify the tangent vectors:

**Lemma 5.4.3.** *Let $s \mapsto \mu(s)$, be a smooth path, $X_s = i(\mu(s))$ and $\xi_s$ such that $\partial_s\mu(s) = -\nabla(\mu(s)\xi_s)$ then $\partial_s X_s = \xi_s(X_s)$. Moreover the formal gradient $\nabla_W$ is send to the usual $L^2$ gradient.*

*Proof.* By definition $M_s(X_s(m)) = m$ where, for $x\in\mathbb{R}$, $M_s(x) = \int_{-\infty}^x \mu(s,y)\,dy$ is the cumulative mass. Thus

$$\partial_s X_s(m) = -\frac{\partial_s M_s(X_s(m))}{\partial_x M_s(X_s(m))} = -\frac{\int_{-\infty}^{X_s(m)} \partial_s\mu(s,y)\,dy}{\mu(s,X_s(m))}.$$

By definition of $\xi_s$ we have $\partial_s X_s(m) = \xi_s(X_s(m))$.
Moreover, for any other tangent vector $\dot{\mu}$ at $\mu(s)$ with $\dot{\mu} = -\nabla(\mu(s)\eta)$ we have:

$$g_\mu(\dot{\mu}, \partial_s\mu(s)) = \int \xi_s(x) \cdot \eta(x)\mu(s,x)dx$$
$$= \int \xi_s(X_s(m)) \cdot \eta(X_s(m))dm = \langle \xi_s(X_s(m)), \eta(X_s(m)) \rangle_{L^2}.$$

□

The following result is a direct consequence of Lemma 5.4.3. .

**Proposition 5.4.4** (Wasserstein and classical gradient flows). *The injection $i$ transforms the Wasserstein gradient flow given in Definition 5.2.2 into the $L^2$ gradient flow defined on $i(W_2^{ac}(\mathbb{R}))$ by the energy $E_L$:*

$$E_L(X) = E\left(i^{-1}(X)\right),$$

*and the gradient flow* (5.2.3a) *reads:*

$$\dot{X}(t,m) = -\nabla_{L^2} E_L(X(t,m)).$$

This injection is a way of constructing characteristics for the system with $X$ the position of the particle. In chapter 7 we rewrite the one dimensional log interaction equation (4.3.2) thanks to this injection.

# Chapter 6

# The one dimensional, subcritical regime

This section takes its inspiration in the work of Calvez and Carrillo [23]. We consider the log interaction equation in dimension one and we want to take advantage of the gradient flow structure. The additional result here with respect to [23] concerns the geometry of the functional: the energy is nonincreasing along the geodesic path. We considered the log-interaction form of (4.1.2) in dimension 1:

$$\begin{aligned}
&\frac{\partial \rho}{\partial t} = \partial_x \left(\rho \partial_x \log\left(\rho\right)\right) - \partial_x(\chi \rho \partial_x c), \quad t > 0, x \in R, \qquad (6.0.1)\\
&\rho(0,\cdot) = \rho_0, \quad \rho_0 \in W_2^{ac}(\mathbb{R}).\\
&c(t,x) = -\frac{1}{2\pi}\int_{\mathbb{R}} \log|x-y|\rho(y)dy\\
&E\left(\rho(t,\cdot)\right) = \int_{\mathbb{R}} \rho\left(t,x\right)\log\left(\rho\left(t,x\right)\right)dx - \frac{\chi}{2}\int_{\mathbb{R}} c\left(t,x\right)\rho\left(t,x\right)dx,
\end{aligned}$$

where as usual, the center of mass is fixed equal to 0 ($\int x\rho(x)dx = 0$) and the mass equal to 1. Moreover we are concerned here with the subcritical case so we fix $\chi < \chi_c$. In this case, the energy is not bounded from below and in order to catch the profile we rescale the equation, with a confinement potential [12]. It leads to:

$$\begin{aligned}
&\frac{\partial \rho}{\partial t} = \nabla \cdot \rho \nabla \log\left(\rho\right) - \nabla \cdot \left(\chi \rho \nabla c\right) + \nabla \cdot x\rho, \quad t > 0, x \in R, \qquad (6.0.2)\\
&\rho(0,\cdot) = \rho_0, \quad \rho_0 \in W_2^{ac}(\mathbb{R}).\\
&c(t,x) = -\frac{1}{2\pi}\int_{R^n} \log|x-y|\rho(y)dy\\
&E\left(\rho(t,\cdot)\right) = \int_{\mathbb{R}} \rho\left(t,x\right)\log\left(\rho\left(t,x\right)\right)dx - \frac{\chi}{2}\int_{\mathbb{R}} c\left(t,x\right)\rho\left(t,x\right)dx + \int \frac{|x|^2}{2}\rho(t,x)dx.
\end{aligned}$$

By [12] we know that a steady state $\mu$ of (6.0.2) exists. We take advantage of the gradient flow formulation to prove the three followings fact:

1. For any $\rho \in W_2^{ac}(\mathbb{R})$ such that $\int x\rho(x)dx = 0$, $F(\rho) \geq F(\mu)$.
2. Let $s \mapsto [\mu,\rho]_s$ be a geodesic path, then $s \mapsto F([\mu,\rho]_s)$ is nonincreasing.

3. Let $t \mapsto \rho_t$ be a solution of (6.0.2) then there exists $\lambda > 0$ with $d_W^2(\mu, \rho_t) \leq e^{-\lambda t} d_W^2(\mu, \rho_0)$.

But in the present case, the energy is not displacement convex and we cannot apply the result of McCann [87]. The first item in our case would prove the classical logarithmic Hardy Littlewood Sovolev (HLS) inequality. The strategy is the following. For the sake of clarity, we first study the critical case, in a formal way since there is no steady state for (6.0.1). Anyway it will be illuminating when looking at the rescaled equation.

## 6.1 The critical case as a formal example

In the weak formulation a steady state, $\mu$, of (6.0.1) is characterized by

$$\int \theta''(x)\mu(x)dx = \frac{\chi}{4\pi} \iint \frac{(\theta'(x) - \theta'(y))}{x - y} \mu(x)\mu(y)dxdy, \quad \forall \theta \in C_b((\mathbb{R}, \mathbb{R}). \qquad (6.1.1)$$

Since we want our results to be valid in any dimension, we will try not to use the injection of $W_2^{ac}(R)$ into $L^2(0,1)$. We start by showing that a steady state is minimal.

**Theorem 6.1.1.** *If there exists a steady state $\mu \in W_2^{ac}(\mathbb{R})$ of (6.0.1), then for any $\rho \in W_2^{ac}(\mathbb{R})$ with $\int x\rho(x)dx = 0$ we have $E(\rho) \geq E(\mu)$. Moreover $E(\rho) = E(\mu)$ if and only if $\rho$ is equal to $\mu$ up to a dilatation and a translation.*

**Remark 6.1.2.** *A consequence of this theorem is that the existence of an equilibrium point of the energy implies the existence of a global minimizer.*

*Proof.* We take $\rho$ such that $W_2^{ac}(\rho, \mu) < \infty$, we suppose that the steady state $\mu \in W_2^{ac}(\mathbb{R})$ exists and we consider the optimal transport map from $\mu$ to $\rho$ and the quadratic cost given by the semi-convex function $\varphi$: $T(x) = x + \varphi'(x)$ [3]. This transport gives, thanks to the Monge-Ampère equation (5.3.4), a change of variable formula:

$$1 + \varphi''(x) = \frac{\mu(x)}{\rho(x + \varphi'(x))}.$$

We apply it to get:

$$E(\rho) = \int \mu(x) \log(\mu(x)/(1 + \varphi''(x)))dx + \chi \iint \log|T(x) - T(y)|\mu(x)\mu(y)dxdy,$$

which is

$$E(\rho) = E(\mu) - \int \mu(x) \log(1 + \varphi''(x))dx + \frac{\chi}{4\pi} \iint \log\left|\frac{T(x) - T(y)}{x - y}\right| \mu(x)\mu(y)dxdy. \qquad (6.1.2)$$

Since $T$ is non decreasing we can dismiss the norm in the last term. On one hand, with (6.1.2) we get:

$$E(\rho) - E(\mu) = -\int \mu(x) \log(1 + \varphi''(x))dx + \frac{\chi}{4\pi} \iint \log\left(\frac{T(x) - T(y)}{x - y}\right) \mu(x)\mu(y)dxdy,$$

On the other hand, let $\theta$ be a solution of $\theta'' = \log(1+\varphi'')$ and test the steady state equation (6.1.1) against $\theta$ ( We can compute $\theta$ thanks to the kernel of the Laplacian in dimension one.) It gives

$$\int \theta(x)''\mu(x)dx = \frac{\chi}{4\pi}\iint \frac{\theta'(x)-\theta'(y)}{x-y}\mu(x)\mu(y)dxdy$$

The two above equations lead to

$$E(\rho)-E(\mu) = \frac{\chi}{4\pi}\left[\iint \log(\frac{T(x)-T(y)}{x-y})\mu(x)\mu(y)dxdy - \iint \frac{\theta'(x)-\theta'(y)}{x-y}\mu(x)\mu(y)dxdy\right]. \quad (6.1.3)$$

Since log is concave, and according to the definition of $\theta$ we can apply lemma 6.1.3 below to the right hand side of (6.1.3). The equality in lemma 6.1.3 i.e. in Jensen's inequality implies $\rho$ is a dilatation, translation of $\mu$. □

**Lemma 6.1.3.** *Let $g$ be a concave function defined on an interval $D$, let $h \in C^2(\mathbb{R},\mathbb{R}))$ with $Im(h'') \in D$ and let $\theta$ be a solution of $\theta'' = g(h'')$ then*

$$\frac{\theta'(x)-\theta'(y)}{x-y} \leq g\left(\frac{h'(x)-h'(y)}{x-y}\right)$$

*Proof.* The proof of lemma 6.1.3 is just an application of Jensen inequality:

$$\theta'(x)-\theta'(y) = \int_y^x \theta''(u)du = \int_y^x g(h''(u))du \leq (x-y)g(\int_y^x h''(u)\frac{du}{x-y}).$$

We can be more specific on the structure of $E$ on $W_2^{ac}(\mathbb{R})$: □

**Theorem 6.1.4.** *Let $\mu \in W_2^{ac}(\mathbb{R})$ be a steady-state of (6.0.1) and $\rho \in W_2^{ac}(\mathbb{R})$ be such that $\rho$ is not obtained by dilation or translation of $\mu$. Then if $\rho_t$ is a geodesic path from $\mu$ to $\rho$, that is $\rho_t = [\mu,\rho]_t$, then $E(\rho_t)$ is increasing on $[0,1]$.*

**Remark 6.1.5.** *Since $E$ is monotone increasing is along all the geodesic paths starting from $\mu$, we also get the uniqueness of the minimizer. This was not obvious since $E$ is not displacement convex*

*Proof.* Let $T_t(x) = x + t\varphi'(x)$ be the transport from $\mu$ to $\rho_t$. For any $t \in [0,1]$ and $h \leq 1-t$, we consider the quantity $E(\rho_{t+h}) - E(\rho_t)$. The equation (6.1.2) applies for $\rho = \rho_{t+h}$ and $\rho = \rho_t$ implies

$$E(\rho_{t+h}) - E(\rho_t) = \frac{\chi}{4\pi}\iint \log\left(\frac{x-y+(t+h)(\varphi'(x)-\varphi'(y))}{x-y+t(\varphi'(x)-\varphi'(y))}\right)\mu(x)\mu(y) - \int \log\frac{1+t\varphi''(x)+h\varphi''(x)}{1+t\varphi''(x)}\mu(x)dx \quad (6.1.4)$$

We test the steady-state equation against $\theta$ such that $\theta'' = \log(1 + h\frac{\varphi''}{1+t\varphi''})$, plugging this into (6.1.4) we obtain

$$E(\rho_{t+h}) - E(\rho_t) = \iint \log\left(1 + h\frac{\varphi'(x)-\varphi'(y)}{x-y+t(\varphi'(x)-\varphi'(y))}\right)\mu(x)\mu(y) - \iint \frac{\theta'(x)-\theta'(y)}{x-y}\mu(x)\mu(y) \tag{6.1.5}$$

The function $u \in [-1, +\infty[ \mapsto \log(1+h\frac{u}{1+tu})$ is concave, thus lemma 6.1.3 together with (6.1.5) gives $E(\rho_{t+h}) - E(\rho_t) > 0$. We indeed obtain a positivity since by hypothesis we avoid the case of equality in lemma 6.1.3.

□

We now have formally a better comprehension of the energy and proved, up to the equality case of the Jensen inequality, the uniqueness of a minimizer for $E$. The next step is to show the decay of the $W_2^{ac}$ distance to a steady-state.

**Claim 2.** *Let $\mu \in W_2^{ac}(\mathbb{R})$ be a steady-state of* (6.0.1) *and $\rho_0 \in W_2^{ac}(\mathbb{R})$. Let $\rho_t$ be a smooth solution of* 6.0.1 *then $\frac{d}{dt}d_w^2(\rho_t, \mu) \leq 0$.*

Unfortunately this claim is empty since there is no steady state of (6.0.1) with finite second moment [9] [27]. The demonstration remains very instructive for the future.

**Remark 6.1.6.** *The metric definition of a gradient flow equation combine to Theorem 6.1.4 gives a direct formal proof of Theorem 2. The main restriction is the smooth solution, we are not sure that the solution $\rho_t$ is smooth enough to perform the computation done in the proof.*

*Proof.* We call $\psi_t$ the convex Brenier map given by the optimal transport from $\rho_t$ to $\mu$ then by 5.2.5 we have:

$$\frac{d}{dt}\frac{1}{2}d_w^2(\rho_t, \mu) = \int (\psi_t'(x) - x)\rho'(x)dx + 2\iint (\psi_t'(x) - x)\frac{x-y}{|x-y|^2}\rho(y)\rho(x)dxdy.$$

We symmetrize the second term. We perform an integration by parts on the first one it gives

$$\frac{d}{dt}\frac{1}{2}d_w^2(\rho_t, \mu) = -\int (\psi_t''(x))\rho(x)dx + \iint ((\psi_t'(x) - \psi_t'(y))\frac{1}{x-y}\rho(y)\rho(x)dxdy,$$

we perform the usual change of variable given by the transport from $\mu$ to $\rho_t$ ( we denote $\phi' = \psi_t'^{-1}$ ). Thus:

$$\frac{d}{dt}\frac{1}{2}d_w^2(\rho_t, \mu) = -\int \frac{1}{\phi''(a)}\mu(a)dx + \iint \frac{a-b}{\phi'(a)-\phi'(b)}\mu(a)\mu(b)dadb,$$

Again we test the steady state equation against $\theta$ define by $\theta'' = \frac{1}{\phi''}$ and use lemma 6.1.3 with $u \in \mathbb{R}^{*+} \mapsto \frac{1}{u}$ convex to finish the proof. □

## 6.2 The subcritical case

In the subcritical we are going to mimic the proof of the critical case, but with this time enough regularity to valid the computation see [12] or appendix B. In the weak formulation the steady state equation for equation (6.0.2) is

$$\int \theta''(x)\mu(x)dx = \frac{\chi}{4\pi}\iint \frac{(\theta'(x)-\theta'(y))}{x-y}\mu(x)\mu(y)dxdy + \int \theta'(x)x\mu(x)dx, \quad \forall\theta \in C_b((\mathbb{R},\mathbb{R}).$$

Since the center of mass of $\mu$ is equal to 0 we can double the variable and get the equivalent formulation, $\forall\theta \in C_b((\mathbb{R},\mathbb{R})$,

$$\int \theta''(x)\mu(x)dx = \frac{\chi}{4\pi}\iint \frac{(\theta'(x)-\theta'(y))}{x-y}\mu(x)\mu(y)dxdy + \frac{1}{2}\int (\theta'(x)-\theta'(y))(x-y)\mu(x)dx. \tag{6.2.1}$$

We can set the following theorem.

**Theorem 6.2.1.** *Let $\mu$ be a solution of (6.2.1) then for any $\rho \in W_2^{ac}(\mathbb{R})$ with $\int x\rho(x)dx = 0$ then*

1. *$F(\rho) \geq F(\mu)$.*

2. *Let $s \mapsto [\mu,\rho]_s$ be a geodesic path, then $s \mapsto F([\mu,\rho]_s)$ is nonincreasing.*

3. *Let $t \mapsto \rho_t$ be a solution of (6.0.2) then there exists $\lambda > 0$ with $d_W^2(\mu,\rho_t) \leq e^{-\lambda t}\, d_W^2(\mu,\rho_0)$.*

**Remark 6.2.2.** *The second item is a geometric information about the energy functional $F$. It says that $F$ has the bmx handlebar shape.*

*Proof.* The proof is the same as the one done for the critical case: we consider the optimal transport map from $\mu$ to $\rho$ and the quadratic cost given by the semi-convex function $\varphi$: $T(x) = x + \varphi'(x)$ and along a geodesic $T_s(x) = x + s\varphi'(x)$, $s \in [0,1]$. It gives for any $s$ a change of variable formula. Using it we get:

1.

$$F(\rho) - F(\mu) \geq -\int \mu(x)\log(1+\varphi''(x))dx + \frac{\chi}{4\pi}\iint \log\left|\frac{T(x)-T(y)}{x-y}\right|\mu(x)\mu(y)dxdy + \frac{1}{4}\int |T(x)-T(y)|^2\mu(x)\mu(y)dxdy - \frac{1}{4}\int |x-y|^2\mu(x)\mu(y)dxdy. \tag{6.2.2}$$

2. Let $s \mapsto [\mu, \rho]_s$ be a geodesic path, then

$$
\begin{aligned}
F(\rho(s+h) - F(\rho(s)) \geq \frac{\chi}{4\pi} \iint \log\left(\frac{x-y+(s+h)(\varphi'(x)-\varphi'(y))}{x-y+s(\varphi'(x)-\varphi'(y))}\right)\mu(x)\mu(y) \\
- \int \log\frac{1+s\varphi''(x)+h\varphi''(x)}{1+s\varphi''(x)}\mu(x)dx \\
+ \frac{1}{4}\int \left|x-y+(s+h)(\varphi'(x)-\varphi'(y))\right|^2 \mu(x)\mu(y)dxdy \\
- \frac{1}{4}\int \left|x-y+s(\varphi'(x)-\varphi'(y))\right|^2 \mu(x)\mu(y)dxdy. \qquad (6.2.3)
\end{aligned}
$$

3. And using (5.2.5) (with the notation of the critical case):

$$
\begin{aligned}
d_W^2(\mu, \rho_t) \leq - \int \frac{1}{\phi''(a)}\mu(a)dx + \iint \frac{a-b}{\phi'(a)-\phi'(b)}\mu(a)\mu(b)dadb \\
+ \iint \frac{(a-b}{\phi'(a)-\phi'(b)}\mu(a)\mu(b)dadb, \quad (6.2.4)
\end{aligned}
$$

where we double the variables. The next step is to use the steady state equation. We test the steady state equation against:

1. $\theta'' = \log(1+\varphi'')$ for (6.2.2).
2. $\theta'' = \log(1+h\dfrac{\varphi''}{1+s\varphi''})$ for (6.2.3).
3. $\theta'' = \dfrac{1}{\phi''}$ for (6.2.4).

It leads us to three different inequality. We deal with them with the following lemma replacing lemma (6.1.3).

**Lemma 6.2.3.** *1. Let $\theta$ solution of $\theta'' = \log(1+\varphi'')$ then for any $\alpha > 0$, $\beta > 0$ we have*

$$
\alpha \log\left(\frac{T(x)-T(y)}{x-y}\right) + \beta\left(\frac{T(x)-T(y)}{x-y}\right)^2 \geq (\alpha+2\beta)\left(\frac{\theta'(x)-\theta'(y)}{x-y}\right) + \beta. \qquad (6.2.5)
$$

*2. Let $\theta$ solution of $\theta'' = \log(1+h\dfrac{\varphi''}{1+s\varphi''}) = g(\varphi'')$ then*

$$
\begin{aligned}
\alpha g\left(\frac{\varphi'(x)-\varphi'(y)}{x-y}\right) + \beta\left(1+(t+h)\left(\frac{\varphi'(x)-\varphi'(y)}{x-y}\right)\right)^2 \\
\geq (\alpha+2\beta)\left(\frac{\theta'(x)-\theta'(y)}{x-y}\right) + \beta\left(1+t\left(\frac{\varphi'(x)-\varphi'(y)}{x-y}\right)\right)^2. \quad (6.2.6)
\end{aligned}
$$

*3. Let $\theta$ solution of $\theta'' = \dfrac{1}{\phi''}$ then*

$$
\alpha\left(\frac{\phi'(x)-\phi'(y)}{x-y}\right)^{-1} - \beta\left(\frac{\phi'(x)-\phi'(y)}{x-y}\right) \geq (\alpha+\beta)\left(\frac{\theta'(x)-\theta'(y)}{x-y}\right) - 2\beta. \qquad (6.2.7)
$$

*Proof.* We only do the proof for (6.2.6), taking $s = 0$ gives (6.2.5) the case of (6.2.7) is similar. By definition

$$\frac{\varphi'(x) - \varphi'(y)}{x - y} = \int_0^1 \varphi''\left((1-s)\,y + sx\right) ds,$$

and

$$\frac{\theta'(x) - \theta'(y)}{x - y} = \int_0^1 \theta''\left((1-s)\,y + sx\right) ds = \int_0^1 g(\varphi'')\left((1-s)\,y + sx\right) ds.$$

Therefore the Jensen's inequality (or directly Lemma 6.1.3) applied to

$$(\alpha + 2\beta) \int_0^1 g(\varphi'')\left((1-s)\,y + sx\right) ds$$

implies

$$\begin{aligned}
&(\alpha + 2\beta)\frac{\theta'(x) - \theta'(y)}{x - y} + \beta\left(1 + t\left(\frac{\varphi'(x) - \varphi'(y)}{x - y}\right)\right)^2 \\
&\leq \alpha g\left(\frac{\varphi'(x) - \varphi'(y)}{x - y}\right) + \beta\left(1 + (t+h)\left(\frac{\varphi'(x) - \varphi'(y)}{x - y}\right)\right)^2 \\
&+ \left[2\beta g\left(\frac{\varphi'(x) - \varphi'(y)}{x - y}\right) + \beta\left(1 + t\left(\frac{\varphi'(x) - \varphi'(y)}{x - y}\right)\right)^2\right. \\
&\left.-\beta\left(1 + (t+h)\left(\frac{\varphi'(x) - \varphi'(y)}{x - y}\right)\right)^2\right].
\end{aligned}$$

There remains to check that the quantity between the bracket is nonpositive. It is equivalent to see that

$$\frac{2}{1 + tz}\log(Z) + 1 - Z^2 \leq 0, \tag{6.2.8}$$

with $z = \dfrac{\varphi'(x) - \varphi'(y)}{x - y}$ and $Z = 1 + h\dfrac{z}{1 + tz}$. Since we always have $Z^2 \geq 1 + 2\log(Z)$, (6.2.8) is equivalent to

$$\log(z)\left(\frac{1}{1 + tz} - 1\right) \leq 0. \tag{6.2.9}$$

Two cases exist: either $z \geq 0$ then $Z > 1$ and (6.2.9) is true or $z < 0$ then $Z \leq 1$ and (6.2.9) is also true. We recall that by construction $1 + tz$ and $Z - 1$ are almost everywhere positive. □

The lemma directly implies theorem 6.2.1 with $\alpha = \dfrac{\xi}{4\pi}$, $\beta = \dfrac{|x-y|^2}{4}$ in the two first cases and $\alpha = \dfrac{\xi}{4\pi}$, $\beta = \dfrac{|x-y|^2}{2}$ in the last one. □

## Remarks

- Lemma 6.2.3 is not completely satisfactory; indeed, a unified version for any convex or concave function in the spirit of Lemma 6.1.3 would be preferable.
- As in [23] everything works in the radial two dimensional case, in particular it gives a new proof of the logarithmic HLS inequality, $F(\rho) \geq F(\rho^\star) \geq F(\mu)$. For a former proof see [26] and

[2] or the appendix B.

• An open question is to extend these computations to the $N$ dimensional case.

# Chapter 7

# Extensive analysis of the particle scheme

## 7.1 Setting of the numerical scheme

### 7.1.a Discretization of the free energy

We consider the problem (4.3.2). As we saw in section 5.4 in the particular case of dimension 1 there is an isometry from $W_2^{ac}(\mathbb{R})$ into $L^2(\mathbb{R})$, given by the pseudo-inverse $(X)$ of the repartition function $(M)$: for $\rho \in W_2^{ac}$ we have $M(x) = \int_{-\infty}^{x} \rho(x)\,dx$ and we associate $X(\cdot) = M(\cdot)^{-1}$. We rewrite the problem with this notation, the energy (4.3.3) becomes for any $X \in L^2(0,1)$:

$$E(X(\cdot)) = -\int \log\left(X'(m)\right) dm + \frac{\chi}{4\pi}\int \log|X(m) - X(p)|dmdp. \tag{7.1.1}$$

We are interested in the supercritical case of (4.3.2). Since the mass is conserved we fix it to one: $\int \rho = 1$. We then consider a discrete formulation splitting uniformly the mass of $\rho$. Let the discretization step $h_N$ be fixed ($h_N = \dfrac{1}{N+1}$ ), then for $i = 0..N+1$ we define $X_i = X(ih_N)$. Note that $X_0 = -\infty, X_{N+1} = +\infty$. For more convenience we also restate $\chi := \dfrac{\chi}{4\pi}$. We want to define a particle scheme therefore we replace for $m \in [ih_N, (i+1)h_N]$, $\dfrac{d}{dm}X(m)$ by $\dfrac{X_{i+1} - X_i}{h_N}$ and $X(m)$ by $X_i$, we get

$$E_{h_N}(X) = -h_N \sum_{i=1}^{N} \log(X_{i+1} - X_i) + \log h_N + \chi h_N^2 \sum_{0 \le i \ne j \le N+1} \log|X_i - X_j|.$$

This energy is not defined because of $X_0$ and $X_1$, we dismiss them and consider the discrete energy $\mathbb{E} : \mathbb{R}^N \to \mathbb{R}$:

$$\mathbb{E}(X) = -\sum_{i=1}^{N-1} \log(X_{i+1} - X_i) + \chi h_N \sum_{1 \le i \ne j \le N} \log|X_i - X_j|, \tag{7.1.2}$$

and define the associate gradient flow equation:

**Definition 7.1.1** (Gradient flow equation.).

$$\begin{cases} \dot{X}(t) = -\nabla \mathbb{E}(X(t)) & t \in \mathbb{R} \\ X(0) = X^0 & X^0 \in \mathbb{R}^N. \end{cases}$$

*We can write it explicitly :*

$$\begin{aligned} \dot{X}_1 &= -\frac{1}{X_2 - X_1} + 2\chi h_N \sum_{j \neq 1} \frac{1}{X_j - X_1} \\ \dot{X}_i &= -\frac{1}{X_{i+1} - X_i} + \frac{1}{X_i - X_{i-1}} + 2\chi h_N \sum_{j \neq i} \frac{1}{X_j - X_i} \\ \dot{X}_N &= \frac{1}{X_N - X_{N-1}} + 2\chi h_N \sum_{j \neq N} \frac{1}{X_i - X_N}. \end{aligned} \tag{7.1.3}$$

For simplicity, we use the convention $\frac{1}{X_1 - X_0} = \frac{1}{X_{N+1} - X_N} = 0$, which is coherent with $X_0 = -\infty$ and $X_1 = +\infty$.
Since the center of mass is conserved $\left(\int x\rho dx\right)$ we set $\sum_{1 \leq i \leq N} X_i = 0$. Thus without loss of generality we may alternatively work with the $N-1$ differences $u_i = X_{i+1} - X_i$. the energy becomes

$$\mathbb{E}_u = -\sum_{i=1}^{N-1} \log(u_i) + 2\chi h_N \left[ \sum_{1 \leq i < j \leq N-1} \log\left( \sum_{k=i}^{j-1} u_k \right) \right] \tag{7.1.4}$$

and the scheme may be rewritten

$$\begin{aligned} \dot{u}_1 &= \frac{2}{u_1} - \frac{1}{u_2} + 2\chi h_N \left[ \sum_{j \neq 2} \frac{1}{X_j - X_2} - \sum_{j \neq 1} \frac{1}{X_j - X_1} \right] \\ \dot{u}_i &= \frac{2}{u_i} - \frac{1}{u_{i-1}} - \frac{1}{u_{i+1}} + 2\chi h_N \left[ \sum_{j \neq i+1} \frac{1}{X_j - X_{i+1}} - \sum_{j \neq i} \frac{1}{X_j - X_i} \right] \\ \dot{u}_n &= \frac{2}{u_{n-1}} - \frac{1}{u_{n-2}} + 2\chi h_N \left[ \sum_{j \neq N} \frac{1}{X_j - X_N} - \sum_{j \neq N-1} \frac{1}{X_j - X_{N-1}} \right]. \end{aligned} \tag{7.1.5}$$

Note that the scheme (7.1.5) is not a gradient flow equation for the energy (7.1.4).
For the gradient flow equation defined in 7.1.1 we give two definitions of the blow-up.

**Definition 7.1.2** (Blow-up). *Let $X$ be a solution of 7.1.1 defined on $[0,T]$. Let $\mathcal{I} \subset [1,N]$ a connected set of indices we say that $\mathcal{I}$ weakly blows up if*

$$\forall (i, i+1) \in \mathcal{I} \times \mathcal{I} \quad \liminf_{t \to T} (X_{i+1} - X_i) = 0. \tag{7.1.6}$$

*We say that $I$ strongly blows-up if*

$$\forall (i, i+1) \in \mathcal{I} \times \mathcal{I} \quad \lim_{t \to T} (X_{i+1} - X_i) = 0. \tag{7.1.7}$$

*In any case, when the set $\mathcal{I}$ is maximal for the inclusion, under the blow-up condition, we call it a blow-up set. If there exists a blow-up set we say that $X$ blows up.*

In the supercritical case, thanks to the second moment computation (see (7.1.8) below ), we are sure that a blow-up exists. One of our goal is to provide natural and robust conditions under which the largest set $\mathcal{I}$ contributing to the blow up will carry generically only the critical number of particles. Indeed in some case the symmetry will force the blow-up to aggregates more than the critical number of particles.

The main difference between a weak and a strong blow-up set concerns the boundary behaviour. Indeed when $\mathcal{I} = [q, p]$ is a weak blow-up set, then $\liminf_{t \to T} (X_{p+1} - X_p)$ and $\liminf_{t \to T} (X_q - X_{q-1})$ are positive. When $\mathcal{I}$ is a strong blow-up set then $(X_{p+1} - X_p)$ and $(X_q - X_{q-1})$ may have no limit in $T$, and in fact $[q-1, p+1]$ can even weakly blow up. This possibility of oscillations at the boundary of a strong blow-up set will be a major issue in the sequel.

### 7.1.b Critical Parameter

We recall that a gradient flow equation takes the steepest descent path for an energy $\mathbb{E}$. Thus it is natural to study the lower bound of $\mathbb{E}$. For $\lambda \in \mathbb{R}_+^*$ we have

$$\mathbb{E}(\lambda X) = \mathbb{E}(X) - \log(\lambda)\left[(N-1) - \chi h_N N (N-1)\right].$$

The outcome of this computation is that $[(N-1) - \chi h_N N (N-1)]$ changes sign depending on $\chi$ thus we define the critical parameter $\chi_N$:

**Definition 7.1.3** (Critical parameter)**.**

$$\chi_N = \frac{1}{h_N N} = 1 + \frac{1}{N} = \frac{N+1}{N}.$$

The same interpretation as in chapter 4 works: if $\chi < \chi_N$ then the energy is not bounded from below when $\lambda$ goes to infinity, which means a dilatation of the set of particles $(X_i)$. It is the relaxation regime.

If $\chi > \chi_N$ the energy is not bounded from below when $\lambda$ goes to 0 which corresponds to a contraction of the set of particles $(X_i)$. Moreover the computation of the second moment $\Pi^2(X) = \sum_{i=1}^{N} X_i^2$ gives ( see (7.1.8) below )

$$\frac{1}{2}\frac{d}{dt}\Pi_k^2 = (N-1) - \chi h_N N (N-1) < 0.$$

Since $\Pi^2$ is positive we know that in finite time this computation fails. It means that there exists $(T, i) \in \mathbb{R}^+ \times [1, N]$ such that $X_{i+1}(T) = X_i(T)$. The set $[i, i+1]$ strongly blows-up. In the supercritical case $\chi > \chi_N$ the blow-up occurs in finite time.

We identified a critical parameter but this is not enough. Indeed we want to show that a blow-up set contains only the critical mass. To guess the eventual critical parameter $\chi_N^k$ for $k$ particles out of $N$ to blow-up we can compute the second moment of an isolated subset of $k$ particles, neglecting the influence of the others. We find with $\Pi_2^k = \sum_{j=0}^{k-1} X_{j+i}^2$:

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\Pi_2^k &= \sum_{i=1}^{k} X_i \dot{X}_i = -\sum_{i=1}^{k} \frac{X_i}{X_{i+1}-X_i} + \sum_{i=1}^{k} \frac{X_i}{X_i - X_{i-1}} + 2\chi h_N \sum_{i=1}^{k}\sum_{j\neq i} \frac{X_i}{X_j - X_i} \\ &= \sum_{i=1}^{k-1} \frac{X_{i+1}-X_i}{X_{i+1}-X_i} + \chi h_N \sum_{i=1}^{k}\sum_{j\neq i} \frac{X_i - X_j}{X_j - X_i} \\ &= (k-1) - \chi h_N k\,(k-1) = (k-1)\,(1-\chi h_N k)\,,\end{aligned} \tag{7.1.8}$$

where we have neglected the boundary terms $\frac{1}{X_{k+1}-X_k}$ and $\frac{1}{X_1 - X_0} = 0$.
This motivates the definition of the $^{th}$ critical parameter for $k$ particles to blow-up:

**Definition 7.1.4** (The $^{th}$ critical parameter)**.**

$$\chi_N^k = \frac{N+1}{k}.$$

If $\chi > \chi_N^k$ the second moment of $k$ particles decreases with a constant speed, thus the solution exists only for a finite time, and $k$ particles can be enough to create the blow-up. In the case of $k = 2$ and $N = 3$ we denote the 2 critical parameter $\chi_3^{(2)}$.

## 7.2 The case of three particles as a toy problem

We first study the particular case of three particles. In this case we can perform a complete analysis. We have two possible cases: three or two particles collapse, depending on the value of $\chi$. In any case we rescale the solution and thanks to a Liouville-type result we catch the asymptotic behaviour. For three particles the energy (7.1.2) and the equation (7.1.3) becomes

$$\begin{aligned} G(X) = &-\log(X_2 - X_1) - \log(X_3 - X_2) \\ &+ 2\chi h_3 \left[\log(X_2 - X_1) + \log(X_3 - X_1) + \log(X_3 - X_2)\right] \end{aligned} \tag{7.2.1}$$

$$G_u(u) = -\log(u_1) - \log(u_2) + 2\chi h_3 h \left[\log(u_1) + \log(u_2) + \log(u_1 + u_2)\right] \tag{7.2.2}$$

and

$$\begin{aligned} \dot{X}_1 &= -\frac{1}{X_2 - X_1} + 2\chi h_3 \left[\frac{1}{X_2 - X_1} + \frac{1}{X_3 - X_1}\right] \\ \dot{X}_2 &= -\frac{1}{X_3 - X_2} + \frac{1}{X_2 - X_1} + 2\chi h_3 \left[\frac{1}{X_3 - X_2} - \frac{1}{X_2 - X_1}\right] \\ \dot{X}_3 &= \frac{1}{X_3 - X_2} + 2\chi h_3 \left[-\frac{1}{X_3 - X_2} - \frac{1}{X_3 - X_1}\right]. \end{aligned} \tag{7.2.3}$$

$$\begin{aligned} \dot{u}_1 &= \frac{2}{u_1} - \frac{1}{u_2} - 2\chi h_3 \left[\frac{2}{u_1} - \frac{1}{u_2} + \frac{1}{u_1 + u_2}\right] \\ \dot{u}_2 &= \frac{2}{u_2} - \frac{1}{u_1} - 2\chi h_3 \left[\frac{2}{u_2} - \frac{1}{u_1} + \frac{1}{u_1 + u_2}\right]. \end{aligned} \tag{7.2.4}$$

Let $X$ be a solution of equation (7.2.3) and suppose that $\chi > \chi_3$. As usual we define the second moment by $\Pi^2(X) = X_1^2 + X_2^2 + X_3^2$, a direct computation gives the following result.

**Proposition 7.2.1.** *For $\chi > \chi_3$ the function $t \mapsto \Pi_2(X(t))$ decreases and*

$$\frac{d}{dt}\Pi_2 = 4(1 - 3h_3\chi).$$

This ensures that there is a blow-up ( $u_1$, $u_2$ or both equal to 0) in finite time.

**Remark 7.2.2.** *In the rest of this section we assume without loss of generality that $u_2 \geq u_1$. Indeed if there exists $t > 0$ such that $u_2(t) = u_1(t)$ then for any $s > t$ we have $u_2(s) = u_1(s)$. Therefore a solution of (7.2.3) cannot cross the line $u_2 = u_1$.*
*This remark rules out the possibility of oscillations ($X_2$ alternatively near $X_1$ then $X_3$).*

### 7.2.a Three particles collapse

In this case: $\chi_3 < \chi < \chi_3^{(2)}$, three particles are required for the blow-up.

**Proposition 7.2.3.** *Let $T$ be the blow-up time we have $u_1(t), u_2(t) \to 0$ as $t \to T$.*

*Proof.* Since the second moment decreases (7.2.1), we already know that $T$ exists. The solution cannot cross the line $u_2 = u_1$ (if $u_2 = u_1$ once then $u_2 = u_1$ forever). Therefore the computation of the second moment (7.2.1) implies that $u_1$ goes to 0 as $t$ goes to $T$.

Moreover there exists $a > 0$ such that if $\frac{u_2}{u_1} \geq a$ then $\frac{u_2}{u_1}$ increases; indeed starting with (7.2.4) we get:

$$\begin{aligned}\frac{d}{dt}\frac{u_2}{u_1} &= \frac{\dot{u}_2 u_1 - \dot{u}_1 u_2}{u_1^2} \\ &\leq \frac{1}{u_1^2}\left[2\left(\frac{u_1}{u_2} - \frac{u_2}{u_1}\right)\left(1 - \frac{\chi}{\chi_3^{(2)}}\right) + \chi_3^{(2)}\frac{u_2 - u_1}{u_2 + u_1}\right] \\ &\leq \frac{1}{u_1^2}\left[-a\left(1 - \frac{\chi}{\chi_3^{(2)}}\right) + \chi_3^{(2)}\right] < 0,\end{aligned}$$

thanks to the fact that $\left(1 - \frac{\chi}{\chi_3^{(2)}}\right) > 0$, and taking $a$ large enough ( $a \geq \frac{\chi_3^{(2)}}{\left(1 - \frac{\chi}{\chi_3^{(2)}}\right)}$). Thus $\frac{u_2}{u_1}$ is bounded and the proposition follows. □

In order to catch the blow-up profile we rescale the solution in this case, it is easy since we learnt from proposition 7.2.3 that the blow-up occurs when the second moment is equal to 0.

**Parabolic rescaling**

By proposition 7.2.1, $\Pi_2$ is decreasing linearly with speed

$$4\left(1 - \chi 3 h_3\right) = -2\alpha = -4\left(\frac{\chi}{\chi_3} - 1\right).$$

We rescale the solution of (7.2.3), in order to fix the second moment equal to one and get a solution defined for all time:

$$Y\left(\tau\left(t\right)\right) = \frac{X\left(t\right)}{R\left(t\right)}, \tag{7.2.5}$$

where $R\left(t\right) = \left|X\left(t\right)\right| = \sqrt{\left|X\left(0\right)\right|^2 - 2\alpha t} = \sqrt{2\alpha\left(T - t\right)}$ and $\tau\left(t\right) = -\frac{1}{\alpha}\log\left(\frac{R\left(t\right)}{R(0)}\right)$.

The gradient flow (7.2.3) may then be written as

$$\dot{Y} = -\nabla\mathbb{E}_r\left(Y\right) = -\nabla\mathbb{E}\left(Y\right) + \alpha Y,$$

where

$$\begin{aligned}\mathbb{E}_r\left(Y\right) = &-\log\left(Y_2 - Y_1\right) - \log\left(Y_3 - Y_2\right) \\ &+ 2\chi h_3\left[\log\left(Y_2 - Y_1\right) + \log\left(Y_3 - Y_1\right) + \log\left(Y_3 - Y_2\right)\right] - \frac{\alpha}{2}|Y|^2,\end{aligned} \tag{7.2.6}$$

or:

$$\begin{aligned}
\dot{Y_1} &= -\frac{1}{Y_2-Y_1} &- 2\chi h_3\left(-\frac{1}{Y_2-Y_1}-\frac{1}{Y_3-Y_1}\right)+\alpha Y_1\\
\dot{Y_2} &= -\frac{1}{Y_3-Y_2}+\frac{1}{Y_2-Y_1} &- 2\chi h_3\left(-\frac{1}{Y_3-Y_2}+\frac{1}{Y_2-Y_1}\right)+\alpha Y_2\\
\dot{Y_3} &= \frac{1}{Y_3-Y_2} &- 2\chi h_3\left(+\frac{1}{Y_3-Y_2}+\frac{1}{Y_3-Y_1}\right)+\alpha Y_3
\end{aligned} \tag{7.2.7}$$

We still have $Y_1+Y_2+Y_3=0$ but now $\Pi^2(Y)=1$. We define $v_1=Y_2-Y_1$ and $v_2=Y_3-Y_2$ to get :

$$\begin{aligned}
\dot{v_1} &= \frac{2}{v_1}-\frac{1}{v_2} &- 2\chi h_3\left(\frac{2}{v_1}-\frac{1}{v_2}+\frac{1}{v_1+v_2}\right)+\alpha v_1\\
\dot{v_2} &= \frac{2}{v_2}-\frac{1}{v_1} &- 2\chi h_3\left(\frac{2}{v_2}-\frac{1}{v_1}+\frac{1}{v_1+v_2}\right)+\alpha v_2
\end{aligned} \tag{7.2.8}$$

and, noting that

$$Y_1^2+Y_2^2+Y_3^2=\frac{2}{3}(v_1^2+v_2^2+v_1v_2), \tag{7.2.9}$$

yields

$$\begin{aligned}
\mathbb{E}_{r,u}(v_1,v_2) = -\log(v_1)-\log(v_2)+&\\
2\chi h_3\left[\log(v_1)+\log(v_1+v_2)+\log(v_2)\right]-\frac{\alpha}{3}\left(v_1^2+v_2^2+v_1v_2\right)&
\end{aligned} \tag{7.2.10}$$

In order to prove (7.2.9) one has to remark that $0=(Y_1+Y_2+Y_3)^2=Y_1^2+Y_2^2+Y_3^2+2(Y_1Y_2+Y_3Y_2+Y_1Y_3)$.

The next proposition will help us to prove that the rescaled solution of (7.2.3) is a very particular solution of (7.2.7).

**Proposition 7.2.4.** *Let $X$ be a solution of* (7.2.3) *then the rescaled solution $Y$, defined with* (7.2.5)*, is a bounded solution of* (7.2.7)*. Moreover $Y$ is defined for all time $t>0$.*

*Proof.* In order to prove this proposition we need to show that there exists a $A>0$ such that

$$\frac{1}{A}\le\frac{X_2-X_1}{|X|},\frac{X_3-X_2}{|X|}\le A. \tag{7.2.11}$$

The upper bound is easy to obtain with the second moment:

$$\frac{X_2-X_1}{|X|}=\frac{u_1}{|X|}\le\frac{u_2}{|X|}=\frac{X_3-X_2}{|X|}\le\frac{\sqrt{X_3^2}+\sqrt{X_2^2}}{|X|}\le 2\sqrt{\Pi^2(Y)}=2.$$

For the lower bound we need to work a bit. Starting with (7.2.9) we have:

$$1=Y_1^2+Y_2^2+Y_3^2=\frac{2}{3}(v_1^2+v_2^2+v_1v_2)\le 2v_2^2+\frac{1}{2}v_1^2+\frac{1}{2}v_2^2\le 3v_2^2,$$

which is exactly $\dfrac{X_3 - X_2}{|X|} \geq \dfrac{1}{\sqrt{3}}$. For $v_1$ let us examine the derivative $\dot{v}_1$ knowing that $v_2 \geq \dfrac{1}{\sqrt{3}}$.

$$\begin{aligned}\dot{v_1} &= \frac{2}{v_1} - \frac{1}{v_2} - 2\chi h_3 \left(\frac{2}{v_1} - \frac{1}{v_2} + \frac{1}{v_1+v_2}\right) + \alpha v_1 \geq \left(\frac{2}{v_1} - \frac{1}{v_2}\right)\left(1 - \frac{\chi}{\chi_3^{(2)}}\right) - \frac{\chi}{\chi_3^{(2)}}\sqrt{3} \\ &\geq \left(\frac{2}{v_1} - \sqrt{3}\right)\left(1 - \frac{\chi}{\chi_3^{(2)}}\right) - \frac{\chi}{\chi_3^{(2)}}\sqrt{3} \quad (> 0 \textit{ for } v_1 \textit{ small enough}). \end{aligned} \tag{7.2.12}$$

Thus there exists $a$ such that if $v_1 \leq a$, thanks to (7.2.12), $v_1$ increases. We deduce

$$\frac{X_2 - X_1}{|X|} = v_1 \geq \min\left(v_1(0), a\right).$$

□

**Remark 7.2.5.** *We refer to Section 7.5 for the proof with N particles, one can remark that our two steps here corresponds to the descent and the reinitialization step in the N particles case. In the particular case of* 3 *particles we are able to give a very precise description of the dynamics. Nonetheless Proposition 7.2.4 ensures a rigidity theorem: a solution defined for all time $t > 0$ and bounded, is unique, see Figure 7.2.a.*

### The blow-up profile

Now we want to describe the explosion behaviour exactly. Let us start by classifying the solutions of (7.2.8).

**Proposition 7.2.6.** *We consider the set $Y \in \mathbb{R}^3$ such that $\Pi_2(Y) = 1$. Let $\overline{\chi} = \dfrac{16}{9} = (\chi_3)^2$, $\chi_3^{(2)} \geq \overline{\chi} \geq \chi_3$. If $\chi \leq \overline{\chi}$ then there is a unique attractive point for the equation* (7.2.8)*: $\left(\dfrac{\sqrt{2}}{2}, \dfrac{\sqrt{2}}{2}\right)$.*
*Otherwise there are two symmetric attractive points $(\overline{v}_1(\chi), \overline{v}_2(\chi))$ and $(\overline{v}_2(\chi), \overline{v}_1(\chi))$ with $(\overline{v}_1(\chi), \overline{v}_2(\chi)) \to \left(0, \dfrac{\sqrt{3}}{2}\right)$ when $\chi \to \chi_3^{(2)}$.*

**Theorem 7.2.7.** *Let $\chi \leq \overline{\chi}$.*

1. *There exists $(V_1, V_2)$ solution of* (7.2.8) *satisfying $V_1^2 + V_2^2 + V_1V_2 = \dfrac{3}{2}$ , defined on $\mathbb{R}$, with*

$$\lim_{t\to-\infty} \left(V_1(t), V_2(t)\right) = \left(0, \sqrt{\frac{3}{2}}\right).$$

2. *Moreover if $(v_1, v_2)$, $v_2 \geq v_1$, is solution of* (7.2.8) *satisfying $v_1^2 + v_2^2 + v_1v_2 = \dfrac{3}{2}$, there exists $s \in \mathbb{R}$ such that for any $t > 0$:*

$$\left(v_1(t), v_2(t)\right) = \left(V_1(t+s), V_2(t+s)\right).$$

*We can set a similar result for $\chi \geq \overline{\chi}$: There exist two functions $V_g$ and $V_d$, one starting at $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ for $t = -\infty$ and going to $(\overline{v}_1(\chi), \overline{v}_2(\chi))$ as $t$ goes to $+\infty$. The other one starting at $\left(0, \sqrt{\frac{3}{2}}\right)$ for $t = -\infty$ and going to $(\overline{v}_1(\chi), \overline{v}_2(\chi))$ as $t$ goes to $+\infty$. Both are solution of (7.2.8), satisfying $V_1^2 + V_2^2 + V_1V_2 = \frac{3}{2}$. Then for any $v = (v_1, v_2)$, $v_2 \geq v_1$, solution of (7.2.8) satisfying $v_1^2 + v_2^2 + v_1v_2 = \frac{3}{2}$, there exists $s \in \mathbb{R}$ such that for any $t > 0$: $v(t) = V_g(t+s)$ or$v(t) = V_d(t+s)$.*

**Remark 7.2.8.** *Another way of thinking is to see $V = V_1, V_2$ as a parametric one dimensional manifold. The condition $v_1^2 + v_2^2 + v_1v_2 = \frac{3}{2}$ forced any solution to stay on this manifold.*

In figure 7.2.a we show the different configurations one can find for the rescaled equation (7.2.8).

*Proof of proposition 7.2.6.* We start with solutions of (7.2.8) satisfying

$$v_1^2 + v_2^2 + v_1v_2 = \frac{3}{2}, \tag{7.2.13}$$

it means that we restrict the energy to this variety defined by $|Y| = 1$. We look for critical points on this curve. We easily find that $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ is one of them, attractive if it is alone, repulsive otherwise. Moreover it is the only one satisfying $v_1 = v_2$. To find other critical points we look when the necessary condition $\frac{d}{dt}(v_2 - v_1) = 0$ is true. By symmetry we suppose $v_2 > v_1$. The equation (7.2.8) gives

$$0 = \left(\frac{3}{v_2} - \frac{3}{v_1}\right)(1 - 2\chi h_3) + \alpha(v_2 - v_1) = 0.$$

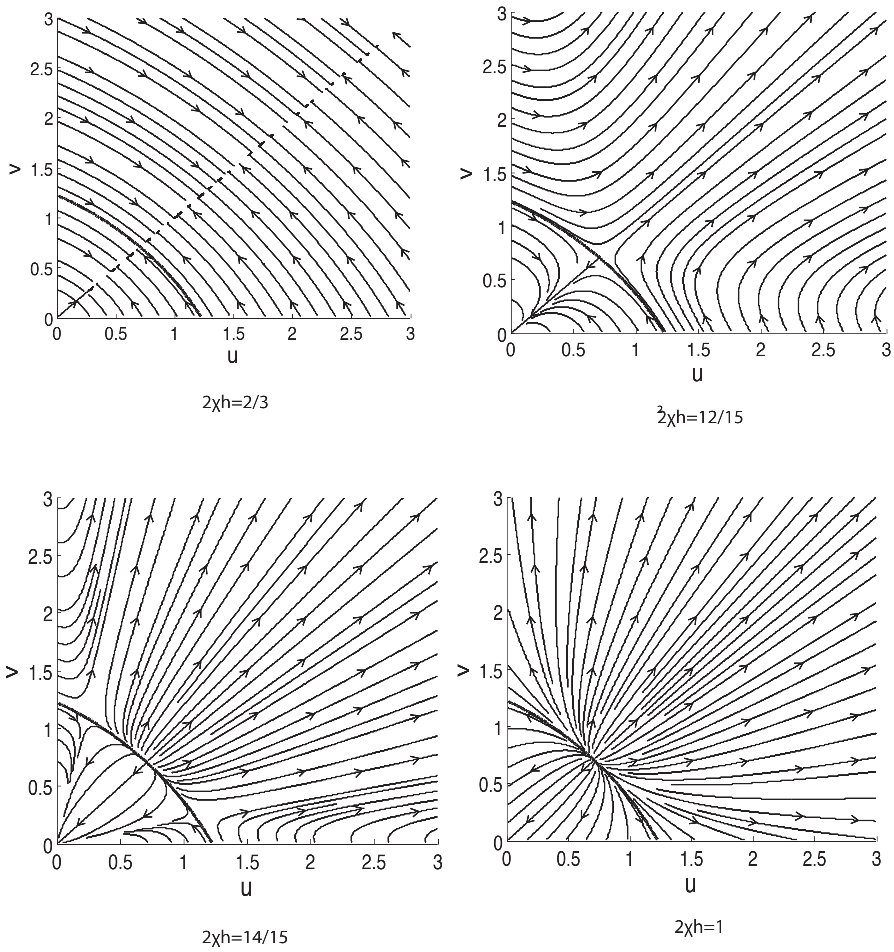
Since $v_2 - v_1 \neq 0$ we find

$$v_1v_2 = 3\frac{1 - 2\chi h_3}{\alpha}.$$

By definition $v_1v_2$ is non-negative, since $\chi \leq \chi_3^2$ the above equation makes sense. Moreover (7.2.13) imposes that $\frac{3}{2} - 9\frac{1 - 2\chi h_3}{\alpha} = (v_2 - v_1)^2 > 0$, thus for $v_1, v_2$ to exist we need $2\chi h_3 > \frac{8}{9}$, that is $\chi \geq \frac{16}{9} = \overline{\chi}$. In this case (7.2.13) gives

$$v_2 + v_1 = \sqrt{\frac{3}{2} + v_1v_2} = \sqrt{\frac{3}{2}\left(1 + 2\frac{1 - 2\chi h_3}{\alpha}\right)}$$

$$v_2 - v_1 = \sqrt{\frac{3}{2} - 3v_1v_2} = \sqrt{\frac{3}{2}\left(1 - 6\frac{1 - 2\chi h_3}{\alpha}\right)}$$




The rescaled equation for $1/3 \leq \chi \leq 2/3$.

From there we get $\overline{v}_1(\chi)$ and $\overline{v}_2(\chi)$

$$\overline{v}_1(\chi) = \frac{1}{2}\left(\sqrt{\frac{3}{2}\left(1+2\frac{1-2\chi h_3}{\alpha}\right)} - \sqrt{\frac{3}{2}\left(1-6\frac{1-2\chi h_3}{\alpha}\right)}\right)$$

$$\overline{v}_2(\chi) = \frac{1}{2}\left(\sqrt{\frac{3}{2}\left(1+2\frac{1-2\chi h_3}{\alpha}\right)} + \sqrt{\frac{3}{2}\left(1-6\frac{1-2\chi h_3}{\alpha}\right)}\right)$$

When $2\chi h_3$ goes to 1, $(\overline{v}_1(\chi), \overline{v}_2(\chi))$ goes to $(0, \sqrt{\frac{3}{2}})$.

Since $\mathbb{E}_r(v_1, v_2) \to \infty$ as $(v_1, v_2) \to (0, \sqrt{\frac{3}{2}})$ or $(0, \sqrt{\frac{3}{2}})$ We know that $(\overline{v}_1(\chi), \overline{v}_2(\chi))$ and $(\overline{v}_2(\chi), \overline{v}_1(\chi))$ are attractive if we stay on the curve, repulsive otherwise. □

*Proof of theorem 7.2.7.* Let $V = (V_1, V_2)$ be a maximal solution of (7.2.8) considered with positive and negative time. Therefore $V$ is define on $]-\infty, +\infty[$ with $\lim_{t\to-\infty} (V_1(t), V_2(t)) = \left(0, \sqrt{\frac{3}{2}}\right)$ and $\lim_{t\to+\infty} (V_1(t), V_2(t)) = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$. Thus $V$ describe all the curve (7.2.13) satisfying $v_2 > v_1$.
Consequently for $(v_1, v_2)$ solution of (7.2.8), define on $[0, +\infty]$ and satisfying (7.2.13) there exists $s$ such that $(v_1(0), v_2(0)) = (V_1(s), V_2(s))$. By uniqueness of the solution for any $t > 0$, $(v_1(t), v_2(t)) = (V_1(t+s), V_2(t+s))$.
We can do exactly the same construction when $\chi \geq \overline{\chi}$. □

**Remark 7.2.9.** *We can summarize this theorem regarding the number of degrees of freedom. Let us start with $u_1, u_2$ solution of (7.2.4). We have two degrees of freedom: the initial conditions. When we rescale u the theorem 7.2.7 says only one degree of freedom is left: that is the shift from the solution $V$. The second degree of of freedom is the blow-up time.*

$$(u_1(t), u_2(t)) = \sqrt{2\alpha(T-t)}\,(v_1(\tau(t)), v_2(\tau(t))) = \sqrt{2\alpha(T-t)}\,(V_1(\tau(t)+s), V_2(\tau(t)+s))\,,$$

*where $T$ is the blow-up time. The function $\tau$ depends only on $T$ on $\alpha$. In other words we parametrize the set of the solution of (7.2.4) thanks to $(T, s)$: the blow-up time and the shift instead of the initial conditions $(u_1(0), u_2(0))$.*

### Come back to the initial problem

Returning to the $Y$ variable we get the following limit

$$\begin{aligned} Y_1^\infty &= \frac{-2\overline{v}_1(\chi) - \overline{v}_2(\chi)}{3} \\ Y_2^\infty &= \frac{\overline{v}_1(\chi) - \overline{v}_2(\chi)}{3} \\ Y_3^\infty &= \frac{\overline{v}_1(\chi) + 2\overline{v}_2(\chi)}{3} \end{aligned} \tag{7.2.14}$$

In the $X$ variables it gives an equivalent as $t \to T$.

$$
\begin{aligned}
X_1 &\sim \frac{-2\overline{v}_1(\chi) - \overline{v}_2(\chi)}{3}\sqrt{2\alpha(T-t)} \\
X_2 &\sim \frac{\overline{v}_1(\chi) - \overline{v}_2(\chi)}{3}\sqrt{2\alpha(T-t)} \\
X_3 &\sim \frac{\overline{v}_1(\chi) + 2\overline{v}_2(\chi)}{3}\sqrt{2\alpha(T-t)}
\end{aligned}
\tag{7.2.15}
$$

Notice that thanks to Theorem 7.2.7, we know that the function for any $t \in [0, T]$ up to two parameters $T$ the blow up time and $s$ the shift. The important comments here are about the qualitative behaviour of the solution. Near $\chi_3$ the solution goes to a Dirac in a perfect symmetric profile. It suggest that the limit profile is symmetric around the Dirac mass. Then for $\chi$ near $\chi_3^2$ the solution looses his symmetry. It explains why after $\chi_3^{(2)}$ only two particles will participate to the blow and which one is rejected.

### 7.2.b Two particles collapse

In this case we assume $\chi \geq \chi_3^{(2)}$. Let $(u_1, u_2)$ be a solution of (7.2.4). In this section we want to show the following theorem.

**Theorem 7.2.10.**

1. *If $u_2(0) > u_1(0)$ the blow-up involves $X_1$ and $X_2$ only.*
2. *If $u_1(0) > u_2(0)$ the blow-up involves $X_2$ and $X_3$ only.*

**Remark 7.2.11.** *The non generic case $u_1(0) = u_2(0)$ proves that even if $\chi \geq \chi_3^{(2)}$ the blow-up can aggregates three particles. Indeed the equation keeps the symmetry.*

We suppose without lost of generality that $u_2(0) \geq u_1(0)$. The computation for the second moment $X_1^2 + X_2^2 + X_3^2$ is still valid, thus we know that the blow-up time $T < +\infty$ exist, and we can rescale the solution.

#### Parabolic rescaling

As in the previous case we perform the parabolic rescaling,

$$
Y(\tau(t)) = \frac{X(t)}{R(t)},
$$

with

$$
\begin{aligned}
R(t) &= \sqrt{2\overline{\alpha}(T-t)}, \\
\overline{\alpha} &= -2\left(1 - \frac{\chi}{\chi_3^2}\right) > 0, \\
\tau(t) &= -\frac{1}{\alpha}\log(R(t)) + \frac{1}{\alpha}\log(R(0)).
\end{aligned}
$$

We define $v_1 = Y_2 - Y_1, v_2 = Y_3 - Y_2$, the rescaled equations are still (7.2.7) and (7.2.8), the energy(7.2.6) and (7.2.10) with $\overline{\alpha}$ instead of $\alpha$.
We start with a proposition that one should compare to (7.2.11) and in the case of $N$ particles to theorem 7.5.1. It will allow us to track our solution in the rescaled system.

**Proposition 7.2.12.** *There exists $A > 0$ such that for any $t < T$:*

1. $\dfrac{1}{A} \leq \dfrac{X_2(t) - X_1(t)}{\sqrt{2\overline{\alpha}(T-t)}} \leq A.$
2. $\dfrac{X_2(t) - X_1(t)}{\sqrt{2\overline{\alpha}(T-t)}} \sim 1,$ *as* $t \to T$.
3. $\dfrac{1}{A\sqrt{2\overline{\alpha}(T-t)}} \leq \dfrac{X_3(t) - X_2(t)}{\sqrt{2\overline{\alpha}(T-t)}} = Y_3 - Y_2.$

*Proof.* We consider $\liminf_{t\to T} u_1 = 0$ without loss of generality. We first show the third estimate, that is $u_2$ is bounded from below. The equation (7.2.4) gives

$$\dot{u}_2 - \dot{u}_1 = \left(\frac{3}{u_2} - \frac{3}{u_1}\right)(1 - 2\chi h_3) = \left(\frac{3}{u_2} - \frac{3}{u_1}\right)\left(1 - \frac{\chi}{\chi_3^{(2)}}\right).$$

Since $u_2(0) > u_1(0)$, and $\left(1 - \dfrac{\chi}{\chi_3^{(2)}}\right) \leq 0$ we deduce $u_2 - u_1$ increases. In particular for any $t \in [0, T)$:

$$u_2(t) \geq u_2(0) - u_1(0) + u_1(t) \geq u_2(0) - u_1(0). \tag{7.2.16}$$

Taking $A \geq \dfrac{1}{u_2(0) - u_1(0)}$ proves the third item of proposition 7.2.12.
For the first estimate we start with the non-rescaled equation:

$$\dot{u}_1 = \frac{1}{u_1}2(1 - 2\chi h_3) - \frac{1}{u_2}(1 - 2\chi h_3) - \frac{2\chi h_3}{u_1 + u_2},$$

Since $u_1 \leq u_2$ we get

$$2u_1\dot{u}_1 = -2\overline{\alpha} + \frac{u_1}{u_2}\overline{\alpha} - \frac{2u_1\chi h_3}{u_1 + u_2} \leq -\overline{\alpha}. \tag{7.2.17}$$

It shows that $u_1^2$ decreases therefore $u_1 \underset{t\to T}{\longrightarrow} 0$. Thus we can set $u_1(T) = 0$, since $u_2$ is bounded from below we deduce that for any $\varepsilon > 0$ there exists $t_\varepsilon$ such that for all $t \in [t_\varepsilon, T)$:

$$\left|\frac{u_1}{u_2}\overline{\alpha} - \frac{2u_1\chi h_3}{u_1 + u_2}\right| \leq 2\overline{\alpha}\varepsilon. \tag{7.2.18}$$

Plugging (7.2.18) in (7.2.17) we obtain:

$$-2\overline{\alpha}(1 + \varepsilon) \leq 2u_1\dot{u}_1 \leq -2\overline{\alpha}(1 - \varepsilon).$$

we integrate between $t$ and $T$ to get

$$2\overline{\alpha}(T-t)(1-\varepsilon) \leq u^2(t) \leq 2\overline{\alpha}(T-t)(1+\varepsilon).$$

Letting $\varepsilon$ going to 0 we prove the second estimate of proposition 7.2.12. Taking for instance $\varepsilon = \frac{1}{2}$ proves the first one. $\square$

The second estimate of this proposition can be seen as a particular case of the following weak Liouville theorem.

**Theorem 7.2.13.** *Let* $(v_1, v_2)$ *be a solution of 7.2.8 defined* $[0, +\infty)$*. Assume that* $v_1$ *is bounded from above and* $\lim_{t\to+\infty} v_2 = +\infty$ *then* $\lim_{t\to+\infty} v_1 = 1$.

*Proof.* We perform the exact same proof as above but in this case with (7.2.8); we get

$$2v_1\dot{v}_1 - 2\overline{\alpha}v_1^2 = -2\overline{\alpha} + \frac{v_1}{v_2}\overline{\alpha} - \frac{2v_1\chi h_3}{v_1 + v_2}.$$

Since $\lim_{t\to+\infty} v_2 = +\infty$ and $v_1$ is bounded from above for any $\varepsilon > 0$ there exists a time $t_\varepsilon$ such that on $[t_\varepsilon, +\infty)$:

$$-2\overline{\alpha}(1+\varepsilon)e^{-2\overline{\alpha}t} \leq \left(2v_1\dot{v}_1 - 2\overline{\alpha}v_1^2\right) e^{-2\overline{\alpha}t} \leq -2\overline{\alpha}(1-\varepsilon)e^{-2\overline{\alpha}t}.$$

Again since $v_1$ is bounded from above we have $\lim_{t\to+\infty} v_1^2 e^{-2\overline{\alpha}t} = 0$ , thus we integrate between $t$ and $+\infty$ to get:

$$\forall t \geq t_\varepsilon \quad e^{-2\overline{\alpha}t}(1-\varepsilon) \leq v_1^2\,(t)\, e^{-2\overline{\alpha}t} \leq e^{-2\overline{\alpha}t}(1+\varepsilon).$$

that is

$$\forall t \geq t_\varepsilon \quad (1-\varepsilon) \leq v_1^2\,(t) \leq (1+\varepsilon).$$

Letting $\varepsilon$ goes to 0 we get $\lim_{t\to+\infty} v_1(t) = 1$.

$\square$

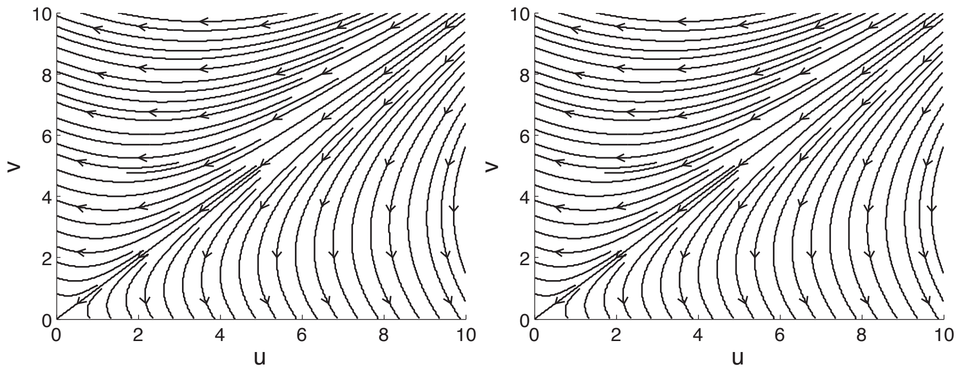
This theorem says that under the conditions 1 and 3 of proposition 7.2.12 we get the second condition of proposition 7.2.12. It is more general since it applies on any solution of (7.2.8) not only the one coming from (7.2.4). In this case it seems a bit trivial but in the case of $N$ particles we will hardly show a similar theorem. In our present case with three particles we can go a bit further and prove a real Liouville theorem.
The next theorem proves the existence and uniqueness of a maximal solution, $\left(\overline{V}_1, \overline{V}_2\right)$, of 7.2.8 satisfying the conditions in proposition 7.2.12.

**Theorem 7.2.14.** *There exists* $\left(\overline{V}_1, \overline{V}_2\right)$*, defined on* $\mathbb{R}$*, solution of* (7.2.8) *such that: for any* $(v_1, v_2)$*,* $v_2 \geq v_1$*, solution of* (7.2.8)*, defined on* $[0, +\infty[$*and satisfying* $v_1 \in L^\infty$*,* $\lim_{t\to+\infty} v_2 = +\infty$ *there exists* $s \geq 0$ *with* $(v_1\,(t)\,, v_2\,(t)) = \left(\overline{V}_1\,(t+s)\,, \overline{V}_2\,(t+s)\right)$.

On figure 7.2.b there are 4 figures to show the different configurations one can find on the rescaled equation (7.2.8) and the original equation (7.2.b).

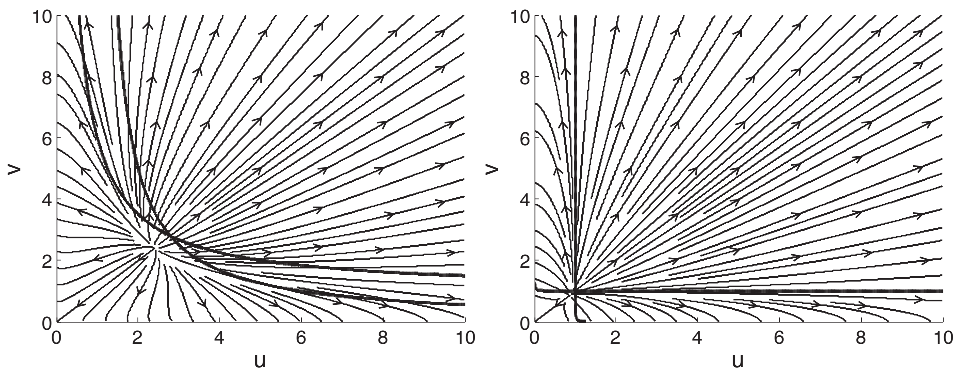
*Proof.* We work with the variable $(\xi, \eta) = \left(v_1 - 1, \dfrac{1}{v_2}\right)$ and we linearise (7.2.8) near the critical

The super critical case.




$2\chi h$=1.05 $2\chi h$=100

The rescaled super critical case.




$2\chi h$=1.05 $2\chi h$=100

point $1, 0$, we get $\left(\dot{\xi}, \dot{\eta}\right) = L\left(\xi, \eta\right) + f\left(\xi, \eta\right)$ with $|f| \leq \left(\xi^2 + \eta^2\right)$ and

$$L = \begin{pmatrix} 2\overline{\alpha} & -1 \\ 0 & -\overline{\alpha} \end{pmatrix}.$$

As $v_1$ is bounded and far from 0 we know that $(\xi)$ is bounded and far from $-1$ therefore the solution $(v_1, v_2)$ must asymptotically satisfies $2\overline{\alpha}\xi - \eta = 0$ that is:

$$2\overline{\alpha}v_1 = \frac{1}{v_2}.$$

Therefore define $\left(\overline{V}_1, \overline{V}_2\right)$ thanks to a shooting method. We find that the function $\left(\overline{V}_1, \overline{V}_2\right)$ is define on $\mathbb{R}$ with $\lim_{t \to -\infty} V_1 = \lim_{t \to -\infty} V_2 = \sqrt{\frac{3\chi h_3 - 1}{\overline{\alpha}}}$.
By definition of $\left(\overline{V}_1, \overline{V}_2\right)$ there exists $s$ such that

$$\left(\overline{V}_1, \overline{V}_2\right)(s) = (v_1, v_2)(0).$$

Since, up to the initial condition, the solution of (7.2.8) is unique we find for any $t > 0$:

$$\left(\overline{V}_1, \overline{V}_2\right)(t + s) = (v_1, v_2)(t).$$

□

**Remark 7.2.15.** *Again we change the two initial conditions of a solution of* (7.2.4) *into two parameters: the blow-up time $T$ and the shift $s$. In other words, when we rescale our solution with $T$, we fix a curve where the solution must stay. There remains only one parameter: where do we start on this curve. We can express $s$ thanks to the formula:*

$$s = V_1^{-1}\left(\frac{u_0}{\sqrt{2\overline{\alpha}T}}\right).$$

### In the original variable

We can come back to the original variable with an exact formula:

$$u_1 = \sqrt{2\overline{\alpha}\left(T - t\right)}\overline{V}_1\left(\tau(t) + s\right)$$
$$u_2 = \sqrt{2\overline{\alpha}\left(T - t\right)}\overline{V}_2\left(\tau(t) + s\right)$$

so

$$\begin{aligned} X_1 - \frac{X1 + X_2}{2} &= -\frac{1}{2}\sqrt{2\overline{\alpha}\left(T - t\right)}\overline{V}_1\left(\tau(t) + s\right) &\sim_{t=T} -\frac{1}{2}\sqrt{2\overline{\alpha}\left(T - t\right)} \\ X_2 - \frac{X1 + X_2}{2} &= \frac{1}{2}\sqrt{2\overline{\alpha}\left(T - t\right)}\overline{V}_1\left(\tau(t) + s\right) &\sim_{t=T} \frac{1}{2}\sqrt{2\overline{\alpha}\left(T - t\right)}, \end{aligned}$$

where

$$\frac{X1 + X_2}{2} = -\frac{\sqrt{2\overline{\alpha}\left(T - t\right)}\overline{V}_1\left(\tau(t) + s\right) + 2\sqrt{2\overline{\alpha}\left(T - t\right)}\overline{V}_2\left(\tau(t) + s\right)}{6}.$$

It shows that $X_3$ has a role to play, which is to fix where the blow-up occurs. To see if the blow–up is symmetric one should take a deeper look at $\left(\overline{V}_1, \overline{V}_2\right)$.

## 7.3 The case of N particles

Now let us take a look at the N particles problem given by the system of equation 7.1.5. It can be written

$$\dot{U} = A_\chi^N \left(U\right).$$

The strategy is to find $N-1$ different behaviours for the particles, but each of them is of the same type. We will divide $U$ in two subset say $U_+, U_-$ such that as long as $(U_+, U_-)$ is include in a domain $D_\chi^\varepsilon$ the system can be rewritten as the following

$$\dot{U}_+ = A_\chi^+ \left(U_+\right)\left[1+\varepsilon\right]$$

$$\dot{U}_- = A_\chi^- \left(U_+\right)\left[1+\varepsilon\right]$$

where "$A_\chi^+ \left(U_+\right) \leq 0$". Hence the set $U_+$ will be the one involving in the explosion. The set $U_-$ will smoothly converge to a limit vector.

The second step is to prove that $D_\chi^\varepsilon$ is stable and the equation inside is attractive with only one limit point. It is easier to study the behaviour of the inequality with $\varepsilon = 0$ and asymptotically we can change our equation with this one and track our solution. This is the spirit of our demonstration. To do this approximation we have to rescale our system, and identify the good solution in the new system. The solution will act, for $t$ large enough like the unique one of the following system:

$$\dot{U}_+ = A_\chi^+ \left(U_+\right).$$

In other words we can forget all the vector who doesn't participate to the blow-up. Then we have to check that the limit profile is the one we are looking for. ($A_\chi^+$ is in fact $A_{f(\chi)}^k$ if $U_+$ has $k$ vector ).

A way to interpret is: as soon as the mechanism of explosion start it will be quicker and quicker and nothing can stop it. Thus every particle not in the package $U_+$ will not be able to participate to the blow-up. This explain why the blow-up will only concern a fixed amount of mass.

In the next section we will follow this strategy. To fix some ideas we first give a quick look at the easy example when $\chi \geq \chi_N^2$. We follow with a detailed proof for all cases.

## 7.4 Stability

### 7.4.a The case of 2 blowing-up particles

Here we take $\chi \geq \chi_N^2$ and we exhibit the attractive set $D_\chi^\delta$.

**Definition 7.4.1.** *Let $\delta, t > 0$ and $X$ be a solution of* (7.1.3)*. We say that $X(t) \in D_\chi^\delta$ if there exists $l$ such that* $u_l(t) \leq \delta \dfrac{2\left(2\chi h_N - 1\right)}{\left(N\chi 2 h_N\right)} \min\left(u_{l-1}(t), u_{l+1}(t)\right)$.

**Proposition 7.4.2.** *There exists $\overline{\delta}$ such that for any $\delta \leq \overline{\delta}$, $D_\chi^\delta$ is stable. That is if $X(t) \in D_\chi^\delta$ then for any $s \in [t,T]$, $X(s) \in D_\chi^\delta$, where $T \in \mathbb{R} \cap \{+\infty\}$ and $[0,T)$ is the definition domain for $X$.*

*Proof.* let $u_l, u_{l-1}, u_{l+1}$ and $t > 0$ such that $u_l(t) \leq \delta \frac{2\left(2\chi h_N - 1\right)}{\left(N\chi 2h_N\right)} \min\left(u_{l-1}(t), u_{l+1}(t)\right)$, By (7.1.5) we get

$$\dot{u}_l \leq \frac{2}{u_l}\left(1 - 2\chi h_N\right)\left[1 + \left(\frac{u_l}{u_{l-1}} + \frac{u_l}{u_{l+1}}\right)\frac{\left(N2\chi h_N\right)}{2\left(1 - 2\chi h_N\right)}\right]$$

and $\left(\frac{u_l}{u_{l-1}} + \frac{u_l}{u_{l+1}}\right)\frac{\left(N2\chi h_N\right)}{2\left(\chi h_N - 1\right)} \leq 2\delta < 1$, for $\delta$ small enough. Thus $u_l$ decrease (note that $\left(1 - 2\chi h_N\right) = 1 - \frac{\chi}{\chi_N^2} < 0$). To prove that $D_\chi^\delta$ is stable it remains to check that $u_{l-1}, u_{l+1}$ increase. We proceed in the same way, starting from (7.1.5) we get:

$$\dot{u}_{l-1} \geq -\frac{1}{u_l}\left(1 - 2\chi h_N\right) - 4N\chi h_N\left[\frac{1}{u_{l-1}}\right]$$

so

$$\dot{u}_{l-1} \geq -\frac{1}{u_l}\left(1 - 2\chi h_N\right)\left[1 + 4\frac{u_l}{u_{l-1}}\frac{\left(N\chi h_N\right)}{\left(1 - 2\chi h_N\right)}\right]$$

For $\delta$ small enough, $1 \geq 4\delta \geq \left(\frac{u_l}{u_{l-1}} + \frac{u_l}{u_{l+1}}\right)\frac{\left(N2\chi h_N\right)}{\left(1 - 2\chi h_N\right)}$. We see that $u_{l-1}$ and by symmetry $u_{l+1}$ increase. The proposition is proved. □

**Remark 7.4.3.** *We can find some bound on the speed and show that we strictly win on $\delta$. In brief not only $D_\chi^\delta$ is stable but it is decreasing: if $u(t) \in D_\chi^\delta$ then $u(t + \varepsilon) \in D_\chi^{\delta - \delta(\varepsilon)}$.*
*As $u_l$ goes to $0$ in finite time ($\dot{u}_l^2 \leq -c < 0$) whereas $\min(u_{l-1}, u_{l+1}) > 0$ we showed that the blow-up concerns only two particles, that is what we expected. Anyway the blow-up can happen before for another $i$ in the same disposition, to be sure to catch the first blow-up we need to consider the 4 uplet of particles with the smallest $\delta$ possible.*

### 7.4.b The case of $k$ blowing-up particles: basins of stability

In this section we fix $\chi_N^k \leq \chi \leq \chi_N^{k-1}$. Then we exhibit stable sets of $k$ blowing-up particles. In order to catch the structure of the discrete Keller-Segel equation we define three important quantities.

**Definition 7.4.4.** *Let $\mathcal{I}$ be a connected set of indices (the* inner set*), e.g. $\mathcal{I} = [l, l+p]$, and $\mathcal{O} = [1,N] \setminus \mathcal{I}$ (the* outer set*). The standard deviation of the family $(X_\mathcal{I}) = \{X_l, ... X_{l+p}\}$ is defined as follows*

$$\Pi_\mathcal{I}^2 = \sum_{i \in \mathcal{I}}\left(X_i - \overline{X}_\mathcal{I}\right)^2, \quad \text{where} \quad \overline{X}_\mathcal{I} = \frac{1}{|\mathcal{I}|}\sum_{i \in \mathcal{I}} X_i\,. \tag{7.4.1}$$

*We also define a variant of this quantity: for a given $\overline{X} \in \mathbb{R}$ (e.g. the blow-up point ) we define the squared distance to $\overline{X}$ by*

$$\overline{\Pi}_\mathcal{I}^2 = \sum_{i \in \mathcal{I}}\left(X_i - \overline{X}\right)^2\,. \tag{7.4.2}$$

*The exterior interaction potential*

$$H_{\mathcal{IO},2} = \sum_{j\in\mathcal{O}} \sum_{i\in\mathcal{I}} \frac{1}{(X_j - X_i)^2}\,.$$

*More generally we define* $H_{\mathcal{IO},m}$ *by*

$$H_{\mathcal{IO},m} = \sum_{j\in\mathcal{O}} \sum_{i\in\mathcal{I}} \frac{1}{(X_j - X_i)^m}\,.$$

**Remark 7.4.5.** *A set of indices* $\mathcal{I}$ *strongly blow-up, at time* $T$*, if and only if*

$$\lim_{t\to T} \Pi_{\mathcal{I}}^2 = 0.$$

*It weakly blow-up if and only if*

$$\varliminf_{t\to T} \Pi_{\mathcal{I}}^2 = 0.$$

*The same is true for* $\overline{\Pi}_{\mathcal{I}}^2$ *taking* $\overline{X}$ *the blow-up point.*

We are able to close a system of inequalities controlling the growth of these three quantities.

**Lemma 7.4.6.** *The following estimates for the evolution of* $\Pi_{\mathcal{I}}^2$*,* $\overline{\Pi}_{\mathcal{I}}^2$ *and* $H_{\mathcal{IO},2}$ *hold true,*

$$\left|\frac{d}{dt}\Pi_{\mathcal{I}}^2 - p\left(1 - \frac{\chi}{\chi_N^{p+1}}\right)\right| \leq \left(2 + \frac{2\chi}{\sqrt{N}}\right)\sqrt{\Pi_{\mathcal{I}}^2 H_{\mathcal{IO},2}}\,, \tag{7.4.3}$$

$$\left|\frac{d}{dt}\overline{\Pi}_{\mathcal{I}}^2 - p\left(1 - \frac{\chi}{\chi_N^{p+1}}\right)\right| \leq \left(2 + \frac{2\chi}{\sqrt{N}}\right)\sqrt{\overline{\Pi}_{\mathcal{I}}^2 H_{\mathcal{IO},2}}\,, \tag{7.4.4}$$

$$\frac{d}{dt}H_{\mathcal{IO},2} \leq \left(12 + 14\chi + 4N^{1/4}\right) H_{\mathcal{IO},4}\,, \tag{7.4.5}$$

$$\frac{d}{dt}H_{\mathcal{IO},2} \leq C_{4,2}\left(12 + 14\chi + 4N^{1/4}\right) H_{\mathcal{IO},2}^2\,. \tag{7.4.6}$$

**Remark 7.4.7.** *We start by a formal calculus with continuous variables just to understand the respective signs and singularities of the various quantities. In continuous variable Definition 7.4.4 is analogue to the following quantities.*

$$\overline{\Pi} = \int_{\mathcal{I}} \left(x - \overline{X}\right)^2 u(x)dx \text{ which corresponds to } \overline{\Pi}_{\mathcal{I}}^2,$$

$$H = \int_{\mathcal{I}}\int_{\mathcal{O}} \frac{1}{(y-x)^2} u(y)u(x)dydx \text{ which corresponds to } H_{\mathcal{IO},2},$$

$$H_4 = \int_{\mathcal{I}}\int_{\mathcal{O}} \frac{1}{(y-x)^4} u(y)u(x)dydx \text{ which corresponds to } H_{\mathcal{IO},4}.$$

*For simplicity we omit to mention the* $t$ *variable.*

**Standard deviation evolution.** *We start by studying for the standard deviation successively the diffusion and the contraction contribution.*

**The diffusion contribution:** *we consider $u$ as a solution of $\frac{d}{dt}u = \Delta u$ to understand the diffusion contribution.*
*We compute forgetting the boundary term when an IPP is done.*

$$\begin{aligned}
\frac{d}{dt}\overline{\Pi} &= \int_{\mathcal{I}} \left(x - \overline{X}\right)^2 \Delta u(x)dx \\
&= -2\int_{\mathcal{I}} \nabla \left(x - \overline{X}\right) \nabla u(x)dx \\
&= 2\int_{\mathcal{I}} u(x)dx = ||u||_{L^1(\mathcal{I})}.
\end{aligned}$$

*As we expected (the diffusion spreads the solution) the diffusion contribution is positive. However we are going to show that under suitable condition the contraction term control the diffusion.*

**The contraction contribution:** *We consider $u$ as a solution of $\frac{d}{dt}u = -\chi\nabla\left(u\nabla\left(\kappa \star u\right)\right)$ to understand the contraction contribution.*

$$\begin{aligned}
\frac{d}{dt}\overline{\Pi} &= -\chi\int_{\mathcal{I}} \left(x - \overline{X}\right)^2 \nabla\left(u(x)\nabla\left(\kappa \star u\right)\right)(x)dx \\
&= 2\chi\int_{\mathcal{I}} \left(x - \overline{X}\right) u(x)\left(\nabla\kappa\right) \star u(x)dx \\
&= -2\chi\int_{\mathcal{I}} \left(x - \overline{X}\right) u(x)\int \frac{1}{x-z}u(z)dzdx \\
&= -2\chi\int_{\mathcal{I}}\int_{\mathcal{I}} \frac{x-\overline{X}}{x-z}u(x)u(z)dzdx + 2\chi\int_{\mathcal{I}}\int_{\mathcal{O}} \frac{x-\overline{X}}{x-z}u(x)u(z)dzdx \qquad = T_1 + T_2.
\end{aligned}$$

*We cannot identify a sign for $T_2$ but we can bound it, thanks to Cauchy-Scharwz, by $2\chi\sqrt{\overline{\Pi}}H$. For $T_1$ we use the symmetric role of $x$ and $z$ to get*

$$\begin{aligned}
T_1 &= -\chi\int_{\mathcal{I}}\int_{\mathcal{I}} \frac{x-\overline{X}}{x-z}u(x)u(z)dzdx + \chi\int_{\mathcal{I}}\int_{\mathcal{I}} \frac{z-\overline{X}}{x-z}u(x)u(z)dzdx \\
&= -\chi\int_{\mathcal{I}}\int_{\mathcal{I}} \frac{x-z}{x-z}u(x)u(z)dzdx = -\chi||u||^2_{L^1(\mathcal{I})}.
\end{aligned}$$

*We combine $T_1$ with the diffusion contribution to get the linear contribution*

$$||u||_{L^1(\mathcal{I})}(1 - \chi||u||_{L^1(\mathcal{I})}).$$

*It is negative since by definition the segment $\mathcal{I}$ contains at least the critical mass $(\frac{1}{\chi})$.*
*Finally for a solution of the one dimensional log-interaction equation we can expect an estimate of the following form:*

$$\left|\frac{1}{2}\frac{d}{dt}\overline{\Pi} - ||u||_{L^1(\mathcal{I})}(1 - \chi||u||_{L^1(\mathcal{I})})\right| \leq 2\chi\sqrt{\overline{\Pi}}H.$$

**Exterior potential evolution.** *We now focus on the exterior potential $H$.*

**The diffusion contribution:** *We consider $u$ as a solution of $\frac{d}{dt}u = \Delta u$ to understand the diffusion contribution. We perform an integration by parts forgetting the boundary terms.*

$$\begin{aligned}
\frac{d}{dt}H &= \int_{\mathcal{I}}\int_{\mathcal{O}} \frac{1}{(y-x)^2}\left[u(y)\Delta u(x) + u(x)\Delta u(y)\right] dydx \\
&= -\int_{\mathcal{I}}\int_{\mathcal{O}} \nabla_x \frac{1}{(y-x)^2} u(y)\nabla u(x) dydx - \int_{\mathcal{I}}\int_{\mathcal{O}} \nabla_y \frac{1}{(y-x)^2} u(x)\nabla u(y) dydx \\
&= -2\int_{\mathcal{I}}\int_{\mathcal{O}} \frac{1}{(y-x)^3} u(y)\nabla u(x) dydx + 2\int_{\mathcal{I}}\int_{\mathcal{O}} \frac{1}{(y-x)^3} u(x)\nabla u(y) dydx \\
&= 6\int_{\mathcal{I}}\int_{\mathcal{O}} \frac{1}{(y-x)^4} u(y)u(x) dydx + 6\int_{\mathcal{I}}\int_{\mathcal{O}} \frac{1}{(y-x)^4} u(x)u(y) dydx \\
&= A + B = 12H_4
\end{aligned}$$

*The diffusion contribution is positive however we can estimate it with $H_4$ or $H^2$.*

**The contraction contribution:** *We consider $u$ as a solution of $\frac{d}{dt}u = -\chi\nabla\left(u\nabla\left(\kappa \star u\right)\right)$ to understand the contraction contribution. We perform an integration by parts forgetting the boundary terms, then we split the convolution regarding the variable's position.*

$$\begin{aligned}
\frac{d}{dt}H &= -\chi\int_{\mathcal{I}}\int_{\mathcal{O}} \frac{1}{(y-x)^2}\left[u(y)\nabla\left(u(x)\nabla\left(\kappa\star u\right)(x)\right) + u(x)\nabla\left(u(y)\nabla\left(\kappa\star u\right)(y)\right)\right] dydx \\
&= \chi\int_{\mathcal{I}}\int_{\mathcal{O}} \nabla_x \frac{1}{(y-x)^2} u(y)u(x)\left(\left(\nabla\kappa\right)\star u(x)\right) dydx \\
&\qquad\qquad + \chi\int_{\mathcal{I}}\int_{\mathcal{O}} \nabla_y \frac{1}{(y-x)^2} u(x)u(y)\left(\left(\nabla\kappa\right)\star u(y)\right) dydx \\
&= -2\chi\int_{\mathcal{I}}\int_{\mathcal{O}} \frac{1}{(y-x)^3} u(y)u(x)\int \frac{1}{x-z}u(z)dzdydx \\
&\qquad\qquad + 2\chi\int_{\mathcal{I}}\int_{\mathcal{O}} \frac{1}{(y-x)^3} u(x)u(y)\int \frac{1}{y-z}u(z)dzdydx \\
&= -2\chi\int_{\mathcal{I}}\int_{\mathcal{O}}\int_{\mathcal{I}} \frac{1}{(y-x)^3}\frac{1}{x-z} u(y)u(x)u(z)dzdydx \\
&\qquad\qquad - 2\chi\int_{\mathcal{I}}\int_{\mathcal{O}}\int_{\mathcal{O}} \frac{1}{(y-x)^3}\frac{1}{x-z} u(y)u(x)u(z)dzdydx \\
&\qquad\qquad + 2\chi\int_{\mathcal{I}}\int_{\mathcal{O}}\int_{\mathcal{I}} \frac{1}{(y-x)^3}\frac{1}{y-z} u(x)u(y)u(z)dzdydx \\
&\qquad\qquad + 2\chi\int_{\mathcal{I}}\int_{\mathcal{O}}\int_{\mathcal{O}} \frac{1}{(y-x)^3}\frac{1}{y-z} u(x)u(y)u(z)dzdydx \\
&= C_1 + C_2 + D_1 + D_2.
\end{aligned}$$

*The contribution given by $D_1$ is positive. Indeed for $D_1$ $y \in \mathcal{O}$ and $(x,y) \in \mathcal{I}\times\mathcal{I}$ we have $(y-x)(y-z)$ positive. However we easily bound $D_1$ by $H_4$ thanks to an Hölder inequality with* *parameter* $p = 4/3, q = 4$.

$$D_1 \le$$
$$2\chi \left(\int_{\mathcal{I}}\int_{\mathcal{O}}\int_{\mathcal{I}} \frac{1}{(y-x)^4} u(x)u(y)u(z)dzdydx\right)^{\frac{3}{4}} \left(\int_{\mathcal{I}}\int_{\mathcal{O}}\int_{\mathcal{I}} \frac{1}{(y-z)^4} u(x)u(y)u(z)dzdydx\right)^{\frac{1}{4}}$$
$$\le 2\chi \left(||u||_{L^1(\mathcal{I})} H_4\right)^{3/4} \left(||u||_{L^1(\mathcal{I})} H_4\right)^{1/4} \le 2\chi H_4.$$

*Away from the boundary* $\mathrm{I} \cap \mathcal{O}$, $C_2$ *does not present any singularity, we bound it by* $H_4$ *thanks to an Hölder estimate.*

$$C_2 \le 2\chi H_4$$

*For* $C_1$ *and* $D_2$ *there is a singularity when* $x = z$ *regarding* $C_1$ *and* $y = z$ *regarding* $D_2$. *We get rid of those singularities thanks to symmetry.*

$$\begin{aligned}
C_1 &= -2\chi \int_{\mathcal{I}}\int_{\mathcal{O}}\int_{\mathcal{I}} \frac{1}{(y-x)^3}\frac{1}{x-z} u(y)u(x)u(z)dzdydx \\
&= \chi \int_{\mathcal{I}}\int_{\mathcal{O}}\int_{\mathcal{I}} \frac{1}{x-z}\left[\frac{1}{(y-z)^3} - \frac{1}{(y-x)^3}\right] u(y)u(x)u(z)dzdydx \\
&= \chi \int_{\mathcal{I}}\int_{\mathcal{O}}\int_{\mathcal{I}} \frac{1}{x-z}\frac{(y-x)^3-(y-z)^3}{(y-z)^3(y-x)^3} u(y)u(x)u(z)dzdydx \\
&= -\chi \int_{\mathcal{I}}\int_{\mathcal{O}}\int_{\mathcal{I}} \frac{(y-x)^2+(y-z)^2+(y-z)(y-x)}{(y-z)^3(y-x)^3} u(y)u(x)u(z)dzdydx \\
&= -\chi \int_{\mathcal{I}}\int_{\mathcal{O}}\int_{\mathcal{I}} \left[\frac{1}{(y-z)^3(y-x)} + \frac{1}{(y-z)(y-x)^3} \right.\\
&\qquad\qquad \left. + \frac{1}{(y-z)^2(y-x)^2}\right] u(y)u(x)u(z)dzdydx
\end{aligned}$$

*The singularity disappears; moreover* $(y-z)(y-x)$ *is always positive when* $y \in \mathcal{O}$ *and* $x, z \in \mathcal{I}$. *Therefore* $C_1$ *is non positive and we do not need to estimate it more precisely. We deal with* $D_2$ *in the same way.*

$$\begin{aligned}
D_2 &= 2\chi \int_{\mathcal{I}}\int_{\mathcal{O}}\int_{\mathcal{O}} \frac{1}{(y-x)^3}\frac{1}{y-z} u(y)u(x)u(z)dzdydx \\
&= -\chi \int_{\mathcal{I}}\int_{\mathcal{O}}\int_{\mathcal{O}} \frac{1}{y-z}\left[\frac{1}{(z-x)^3} - \frac{1}{(y-x)^3}\right] u(y)u(x)u(z)dzdydx \\
&= -\chi \int_{\mathcal{I}}\int_{\mathcal{O}}\int_{\mathcal{O}} \frac{1}{y-z}\frac{(y-x)^3-(z-x)^3}{(z-x)^3(y-x)^3} u(y)u(x)u(z)dzdydx \\
&= -\chi \int_{\mathcal{I}}\int_{\mathcal{O}}\int_{\mathcal{O}} \frac{(y-x)^2+(z-x)^2+(z-x)(y-x)}{(z-x)^3(y-x)^3} u(y)u(x)u(z)dzdydx \\
&= -\chi \int_{\mathcal{I}}\int_{\mathcal{O}}\int_{\mathcal{O}} \left[\frac{1}{(z-x)^3(y-x)} + \frac{1}{(z-x)(y-x)^3} \right.\\
&\qquad\qquad \left. + \frac{1}{(z-x)^2(y-x)^2}\right] u(y)u(x)u(z)dzdydx.
\end{aligned}$$

*This time the contribution is sometimes positive. We bound each of the three term by $H_4$ thanks to some Hölder inequalities. We get*

$$D_2 \leq 3\chi H_4.$$

*Considering now $u$ as a solution of the one dimensional log interaction equation we can expect an estimate of the form:*

$$\frac{d}{dt}H \leq (12 + 7\chi)\, H_4.$$

*This formal computation will help us for the similar computation done with the particles model. In the same time we understand the pertinence of the particles model to deal with the $(\mathcal{I}, \mathcal{O})$ boundary singularity.*

*Proof.* We do the discrete proof following the spirit of the formal one done above. In addition we have to deal with the boundary terms. We start with the evolution of $\Pi_{\mathcal{I}}^2$, recalling that $X$ satisfies the differential equation 7.1.3.

$$\frac{1}{2}\frac{d}{dt}\Pi_{\mathcal{I}}^2 = \sum_{i\in\mathcal{I}}\left[-\frac{X_i - \overline{X}_{\mathcal{I}}}{X_{i+1} - X_i} + \frac{X_i - \overline{X}_{\mathcal{I}}}{X_i - X_{i-1}} + 2\chi h_N \sum_{j\neq i}\frac{X_i - \overline{X}_{\mathcal{I}}}{X_j - X_i}\right] - \sum_{i\in\mathcal{I}}\left(X_i - \overline{X}_{\mathcal{I}}\right)\frac{d}{dt}\overline{X}_{\mathcal{I}} \tag{7.4.7}$$

$$= \sum_{i=\mathcal{I}\setminus\{l+p\}}\left[-\frac{X_i - X_{i+1}}{X_{i+1} - X_i}\right] - \frac{X_{l+p} - \overline{X}_{\mathcal{I}}}{X_{l+p+1} - X_{l+p}} + \frac{X_l - \overline{X}_{\mathcal{I}}}{X_l - X_{l-1}} + 2\chi h_N \sum_{i\in\mathcal{I}}\left[\sum_{j\neq i}\frac{X_i - \overline{X}_{\mathcal{I}}}{X_j - X_i}\right] \tag{7.4.8}$$

$$= p - \frac{X_{l+p} - \overline{X}_{\mathcal{I}}}{X_{l+p+1} - X_{l+p}} + \frac{X_l - \overline{X}_{\mathcal{I}}}{X_l - X_{l-1}} + 2\chi h_N \sum_{i\in\mathcal{I}}\left[\sum_{j\neq i}\frac{X_i - \overline{X}_{\mathcal{I}}}{X_j - X_i}\right].$$

We used $\sum_{i\in\mathcal{I}} X_i = |\mathcal{I}|\overline{X}_{\mathcal{I}}$. For a moment we just look at the contraction term:

$$T = \sum_{i\in\mathcal{I}}\left[\sum_{j\neq i}\frac{X_i - \overline{X}_{\mathcal{I}}}{X_j - X_i}\right] = \sum_{i\in\mathcal{I}}\left[\sum_{j\in\mathcal{I}\setminus\{i\}}\frac{X_i - \overline{X}_{\mathcal{I}}}{X_j - X_i}\right] + \sum_{i\in\mathcal{I}}\left[\sum_{j\in\mathcal{O}}\frac{X_i - \overline{X}_{\mathcal{I}}}{X_j - X_i}\right]$$
$$= T_1 + T_2.$$

The Cauchy-Schwarz inequality on the last term implies

$$T_2 \leq \sqrt{N\Pi_{\mathcal{I}}^2 H_{\mathcal{I}\mathcal{O},2}}.$$

Using the symmetry we simplify $T_1$:

$$T_1 = \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{I}\setminus\{i\}}\left[\frac{X_i-\overline{X}_{\mathcal{I}}}{X_j-X_i}\right] = \frac{1}{2}\left(\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{I}\setminus\{i\}}\left[\frac{X_i-\overline{X}_{\mathcal{I}}}{X_j-X_i}\right]+\sum_{r\in\mathcal{I}}\sum_{s\in\mathcal{I}\setminus\{r\}}\left[\frac{X_r-\overline{X}_{\mathcal{I}}}{X_s-X_r}\right]\right)$$
$$= \frac{1}{2}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{I}\setminus\{i\}}\left[\frac{X_i-\overline{X}_{\mathcal{I}}}{X_j-X_i}\right]+\frac{1}{2}\sum_{j\in\mathcal{I}}\sum_{i\in\mathcal{I}\setminus\{j\}}\left[\frac{X_j-\overline{X}_{\mathcal{I}}}{X_i-X_j}\right] = \frac{1}{2}\sum_{(i,j)\in\mathcal{I}\times\mathcal{I}\setminus\{i=j\}}\frac{X_j-X_i}{X_i-X_j}$$
$$= -\frac{p\,(p+1)}{2}.$$

All in one we obtain

$$2\chi h_N T \leq -2\chi h_N\frac{p\,(p+1)}{2} + 2\chi h_N\sqrt{N\Pi_{\mathcal{I}}^2 H_{\mathcal{IO},2}}.$$

Coming back to $\frac{1}{2}\frac{d}{dt}\Pi_{\mathcal{I}}^2$ we get

$$\frac{1}{2}\frac{d}{dt}\Pi_{\mathcal{I}}^2 \leq p\left(1-\frac{\chi}{\chi_N^{p+1}}\right)+\left(2+\frac{2\chi}{\sqrt{N}}\right)\sqrt{\Pi_{\mathcal{I}}^2 H_{\mathcal{IO},2}}.$$

Similarly

$$\frac{1}{2}\frac{d}{dt}\Pi_{\mathcal{I}}^2 \geq p\left(1-\frac{\chi}{\chi_N^{p+1}}\right)-\left(2+\frac{2\chi}{\sqrt{N}}\right)\sqrt{\Pi_{\mathcal{I}}^2 H_{\mathcal{IO},2}},$$

where we used the trivial inequalities

$$\max\left(|\frac{X_{l+p}-\overline{X}_{\mathcal{I}}}{X_{l+p+1}-X_{l+p}}|,|\frac{X_l-\overline{X}_{\mathcal{I}}}{X_l-X_{l-1}}|\right) \leq \sqrt{\Pi_{\mathcal{I}}^2 H_{\mathcal{IO},2}}.$$

The demonstration for $\overline{\Pi}_{\mathcal{I}}^2$ is exactly the same but this time $\sum_{i\in\mathcal{I}}\left(X_i-\overline{X}\right)\frac{d}{dt}\overline{X} = 0$ because $\overline{X}=0$.

We now look for the $H_{\mathcal{IO},2}$ time derivative and again we refer to the formal proof to well understand the structure of the computation.

$$\frac{d}{dt}H_{\mathcal{IO},2} = -2\sum_{j\in\mathcal{O}}\sum_{i\in\mathcal{I}}\frac{1}{(X_j-X_i)^3}\Bigg[$$
$$\frac{1}{(X_{i+1}-X_i)}-\frac{1}{(X_i-X_{i-1})}-\frac{1}{(X_{j+1}-X_j)}+\frac{1}{(X_j-X_{j-1})}$$
$$+2\chi h_N\left(-\sum_{k\neq i}\frac{1}{(X_k-X_i)}+\sum_{k\neq i}\frac{1}{(X_k-X_j)}\right)\Bigg].$$

We split the right hand side into four terms:

$$\frac{d}{dt}H_{\mathcal{IO},2} = A+B+C+D, \tag{7.4.9}$$

where

$$A = -2\sum_{j\in\mathcal{O}}\sum_{i\in\mathcal{I}} \frac{1}{(X_j - X_i)^3}\left[\frac{1}{(X_{i+1}-X_i)} - \frac{1}{(X_i - X_{i-1})}\right],$$

$$B = -2\sum_{j\in\mathcal{O}}\sum_{i\in\mathcal{I}} \frac{1}{(X_j - X_i)^3}\left[-\frac{1}{(X_{j+1}-X_j)} + \frac{1}{(X_j - X_{j-1})}\right],$$

$$C = 4\chi h_N \sum_{j\in\mathcal{O}}\sum_{i\in\mathcal{I}} \frac{1}{(X_j - X_i)^3}\left[\sum_{k\neq i}\frac{1}{(X_k - X_i)}\right],$$

$$D = -4\chi h_N \sum_{j\in\mathcal{O}}\sum_{i\in\mathcal{I}} \frac{1}{(X_j - X_i)^3}\left[\sum_{k\neq i}\frac{1}{(X_k - X_j)}\right].$$

The strategy is to bound each term from above with $H_{\mathcal{IO},4}$.
A discrete integration by parts on $B$ gives

$$\begin{aligned} B = & -2\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{O}} \frac{1}{(X_j - X_{j-1})}\left[\frac{1}{(X_j - X_i)^3} - \frac{1}{(X_{j-1}-X_i)^3}\right] \\ & + 2\sum_{i\in\mathcal{I}} \frac{1}{(X_l - X_{l-1})}\frac{1}{(X_{l-1}-X_i)^3} - 2\sum_{i\in\mathcal{I}}\frac{1}{(X_{l+p+1}-X_{l+p})}\frac{1}{(X_{l+p+1}-X_i)^3}. \end{aligned} \tag{7.4.10}$$

Since $l-1 < i < l+p+1$ we have

$$(X_l - X_{l-1})(X_{l-1}-X_i) \leq 0 \text{ and } (X_{l+p+1}-X_{l+p})(X_{l+p+1}-X_i) \geq 0.$$

Therefore the boundary terms, *i.e.* the two last terms in (7.4.10), are nonpositive and we can dismiss them for the upper bound of $\frac{d}{dt}H_{IO,2}$. There remains to treat the first term of the right hand side of (7.4.10). In the following computation the summation over $i$ and $j$ is taken for $i \in \mathcal{I}$ and $j \in \mathcal{O}$.

$$\begin{aligned} & -2\sum_{i,j} \frac{1}{(X_j - X_{j-1})}\left[\frac{1}{(X_j - X_i)^3} - \frac{1}{(X_{j-1}-X_i)^3}\right] \\ & = -2\sum_{i,j}\frac{1}{(X_j - X_{j-1})}\left[\frac{(X_{j-1}-X_i)^3 - (X_j - X_i)^3}{(X_j - X_i)^3(X_{j-1}-X_i)^3}\right] \\ & = -2\sum_{i,j}\left[\frac{(X_{j-1}-X_j)\Big((X_{j-1}-X_i)^2 + (X_j - X_i)^2 + (X_{j-1}-X_i)(X_j - X_i)\Big)}{(X_j - X_{j-1})(X_j - X_i)^3(X_{j-1}-X_i)^3}\right] \\ & = 2\sum_{i,j}\left[\frac{(X_{j-1}-X_i)^2 + (X_j - X_i)^2 + (X_{j-1}-X_i)(X_j - X_i)}{(X_j - X_i)^3(X_{j-1}-X_i)^3}\right] \\ & = 2\sum_{i,j}\left[\frac{1}{(X_j - X_i)^3(X_{j-1}-X_i)} + \frac{1}{(X_j - X_i)(X_{j-1}-X_i)^3} + \frac{1}{(X_j - X_i)^2(X_{j-1}-X_i)^2}\right]. \end{aligned}$$

This contribution is always positive. The Hölder inequality applied on each of the three terms, with coefficient $(4/3, 4)$, $(4, 4/3)$ and $(2, 2)$ gives

$$2\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{O}}\frac{(X_{j-1}-X_i)^2+(X_j-X_i)^2+(X_{j-1}-X_i)(X_j-X_i)}{(X_j-X_i)^3(X_{j-1}-X_i)^3}\leq 6H_{\mathcal{IO},4}.$$

Coming back to $B$ we get

$$B\leq 6H_{\mathcal{IO},4}. \tag{7.4.11}$$

A discrete integration by parts on $A$ gives something similar to $B$.

$$\begin{aligned}A=-2\sum_{i,i-1\in\mathcal{I}}\sum_{j\in\mathcal{O}}&\frac{1}{(X_i-X_{i-1})}\left[\frac{1}{(X_j-X_{i-1})^3}-\frac{1}{(X_j-X_i)^3}\right]\\&+2\sum_{j\in\mathcal{O}}\frac{1}{(X_l-X_{l-1})}\frac{1}{(X_j-X_l)^3}-2\sum_{j\in\mathcal{O}}\frac{1}{(X_{l+p+1}-X_{l+p})}\frac{1}{(X_j-X_{l+p})^3}.\end{aligned} \tag{7.4.12}$$

This time, the boundary terms, *i.e.* the last two terms of the right hand side of (7.4.12), has no sign, and we have to control it. Since $j\in\mathcal{O}$ and $l+p\in\mathcal{I}$, the Hölder inequality applied to the last term of (7.4.12) with coefficient $p=4$ and $q=4/3$ implies

$$2\left|\sum_{j\in\mathcal{O}}\frac{1}{(X_{l+p+1}-X_{l+p})}\frac{1}{(X_j-X_{l+p+1})^3}\right|\leq 2N^{1/4}H_{\mathcal{IO},4}.$$

Similarly, the second term of the r.h.s. of (7.4.12) satisfies

$$\left|2\sum_{j\in\mathcal{O}}\frac{1}{(X_l-X_{l-1})}\frac{1}{(X_j-X_l)^3}\right|\leq 2N^{1/4}H_{\mathcal{IO},4}.$$

There remains to deal with the first term of the right hand side of (7.4.12), the core of the integration by parts. We follow here the proof done for $B$ to avoid the singularity.

$$\begin{aligned}&-2\sum_{i,i-1\in\mathcal{I}}\sum_{j\in\mathcal{O}}\frac{1}{(X_i-X_{i-1})}\left[\frac{1}{(X_j-X_{i-1})^3}-\frac{1}{(X_j-X_i)^3}\right]\\&=2\sum_{i,i-1\in\mathcal{I}}\sum_{j\in\mathcal{O}}\left[\frac{1}{(X_j-X_{i-1})(X_j-X_i)^3}+\frac{1}{(X_j-X_{i-1})^3(X_j-X_i)}\right.\\&\qquad\qquad\left.+\frac{1}{(X_j-X_{i-1})^2(X_j-X_i)^2}\right].\end{aligned}$$

This contribution is positive. We control it with three Hölder inequalities. All in one we find the following estimate for $A$.

$$A\leq\left(6+4N^{1/4}\right)H_{\mathcal{IO},4}. \tag{7.4.13}$$

We now take a look at $D$ and keep in mind that the main idea is to isolate the contribution due to the inside particles ($\mathcal{I}$) and the other ($\mathcal{O}$) considered as a perturbation. It leads us to the definition of $D_1$ and $D_2$ by

$$\begin{aligned}
D &= -4\chi h_N \sum_{j\in\mathcal{O}}\sum_{i\in\mathcal{I}} \frac{1}{(X_j - X_i)^3}\left[\sum_{k\neq j}\frac{1}{(X_k - X_j)}\right] \\
&= \sum_{j\in\mathcal{O}}\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{I}} \frac{-4\chi h_N}{(X_k - X_j)(X_j - X_i)^3} + \sum_{j\in\mathcal{O}}\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{O}\setminus\{j\}} \frac{-4\chi h_N}{(X_k - X_j)(X_j - X_i)^3} \\
&= D_1 + D_2.
\end{aligned}$$

Since $j \in \mathcal{O}$ and $i, k \in \mathcal{I}$, the contribution of $D_1$ is positive. The Hölder inequality with $p = 4$, $q = 4/3$ gives:

$$\begin{aligned}
D_1 &= -4\chi h_N \sum_{j\in\mathcal{O}}\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{I}} \frac{1}{(X_k - X_j)(X_j - X_i)^3} \\
&\leq 4\chi h_N \left[\sum_{j\in\mathcal{O}}\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{I}} \frac{1}{(X_k - X_j)^4}\right]^{1/4} \left[\sum_{j\in\mathcal{O}}\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{I}} \frac{1}{(X_j - X_i)^4}\right]^{3/4} \\
&\leq 4\chi h_N N^{1/4} (H_{\mathcal{IO},4})^{1/4} N^{3/4} (H_{\mathcal{IO},4})^{3/4} \\
&\leq 4\chi h_N N H_{\mathcal{IO},4} \\
&\leq 4\chi H_{\mathcal{IO},4}.
\end{aligned}$$

For $D_2$ we follow the formal proof and use the symmetric role of $j$ and $k$.

$$\begin{aligned}
D_2 &= -4\chi h_N \sum_{j\in\mathcal{O}}\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{O}\setminus\{j\}} \frac{1}{(X_k - X_j)(X_j - X_i)^3} \\
&= -2\chi h_N \sum_{j\in\mathcal{O}}\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{O}\setminus\{j\}} \left[\frac{1}{(X_k - X_j)(X_j - X_i)^3} - \frac{1}{(X_k - X_j)(X_k - X_i)^3}\right] \\
&= -2\chi h_N \sum_{j\in\mathcal{O}}\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{O}\setminus\{j\}} \left[\frac{(X_k - X_i)^3 - (X_j - X_i)^3}{(X_k - X_j)(X_j - X_i)^3 (X_k - X_i)^3}\right] \\
&= -2\chi h_N \sum_{j\in\mathcal{O}}\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{O}\setminus\{j\}} \left[\frac{(X_k - X_i)^2 + (X_j - X_i)^2 + (X_k - X_i)(X_j - X_i)}{(X_j - X_i)^3 (X_k - X_i)^3}\right] \\
&= -2\chi h_N \sum_{j\in\mathcal{O}}\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{O}\setminus\{j\}} \left[\frac{1}{(X_j - X_i)^3 (X_k - X_i)} + \frac{1}{(X_j - X_i)(X_k - X_i)^3} \right. \\
&\qquad\qquad \left. + \frac{1}{(X_j - X_i)^2 (X_k - X_i)^2}\right].
\end{aligned}$$

Wee see that this contribution is negative when $j, k \geq i$ or $j, k \leq i$, positive elsewhere. We estimate it in all cases with an Hölder estimate on each of the three term. The parameters are respectively $p = 4/3$, $q = 4$ then $p = 4$, $q = 4/3$ and $p = 2$, $q = 2$). We obtain

$$D_2 \leq 6\chi h_N N H_{\mathcal{IO},4} \leq 6\chi H_{\mathcal{IO},4}. \tag{7.4.14}$$

Getting back to $D$ we find

$$D \leq 10\chi H_{\mathcal{IO},4}. \tag{7.4.15}$$

In a similar way we estimate $C$.

$$\begin{aligned}
C &= 4\chi h_N \sum_{j\in\mathcal{O}} \sum_{i\in\mathcal{I}} \frac{1}{(X_j - X_i)^3} \left[ \sum_{k\neq i} \frac{1}{(X_k - X_i)} \right] \\
&= 4\chi h_N \sum_{j\in\mathcal{O}} \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{I}\setminus\{i\}} \frac{1}{(X_j - X_i)^3} \frac{1}{(X_k - X_i)} + 4\chi h_N \sum_{j\in\mathcal{O}} \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{O}} \frac{1}{(X_j - X_i)^3} \frac{1}{(X_k - X_i)} \\
&= C_1 + C_2.
\end{aligned}$$

The sign of $C_2$ depends on the relative position of the indices. In any case since $j, k \in \mathcal{O}$ and $i \in \mathcal{I}$, the Hölder inequality with parameters $p = 4/3$, $q = 4$ gives, as for $D_1$,

$$C_2 \leq 4\chi H_{\mathcal{IO},4}.$$

To understand $C_1$ we use the symmetric role of $i$ and $k$. A similar computation as the one done for $D_2$ gives

$$\begin{aligned}
C_1 &= 4\chi h_N \sum_{j\in\mathcal{O}} \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{I}\setminus\{i\}} \frac{1}{(X_j - X_i)^3} \frac{1}{(X_k - X_i)} \\
&= 2\chi h_N \sum_{j\in\mathcal{O}} \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{I}\setminus\{i\}} \left[ \frac{1}{(X_j - X_i)^3} \frac{1}{(X_k - X_i)} - \frac{1}{(X_j - X_k)^3} \frac{1}{(X_k - X_i)} \right] \\
&= 2\chi h_N \sum_{j\in\mathcal{O}} \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{I}\setminus\{i\}} \left[ \frac{(X_j - X_k)^3 - (X_j - X_i)^3}{(X_k - X_i)(X_j - X_i)^3 (X_j - X_k)^3} \right] \\
&= -2\chi h_N \sum_{j\in\mathcal{O}} \sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{I}\setminus\{i\}} \left[ \frac{1}{(X_j - X_i)^3 (X_j - X_k)} + \frac{1}{(X_j - X_i)(X_j - X_k)^3} \right. \\
&\qquad\qquad \left. + \frac{1}{(X_j - X_i)^2 (X_j - X_k)^2} \right].
\end{aligned}$$

Therefore the contribution is nonpositive and we do not need to control it. Coming back to $C$ we find

$$C \leq 4\chi H_{\mathcal{IO},4}. \tag{7.4.16}$$

Together 7.4.11, 7.4.13, 7.4.15 and 7.4.16 in 7.4.9 implies

$$\begin{aligned}
\frac{d}{dt} H_{\mathcal{IO},2} &\leq \left(12 + 14\chi + 4N^{1/4}\right) H_{\mathcal{IO},4} && (7.4.17) \\
&\leq C_{4,2} \left(12 + 14\chi + 4N^{1/4}\right) H_{\mathcal{IO},2}^2, && (7.4.18)
\end{aligned}$$

where $C_{4,2}$ is the sharpest constant such that $||\cdot||_4 \leq C_{4,2}||\cdot||_2$, which we know exists since we consider a finite number of particles. □

We now give some result about the structure of blow-up.

**A first look on the blow-up structure** We recall that we still work with $\chi_N^k \leq \chi \leq \chi_N^{k-1}$. We explain here that a maximal-connected set $\mathcal{I}$ need at least the critical mass to blow-up.

**Claim 3.** *A weak blow-up set contains at least k particles. In other words a weak blow up needs at least the critical mass.*

*Proof.* We suppose that the blow-up use $p+1$ particles: $X_l, ..., X_{l+p}$, we note $\mathcal{I} = [l, l+p]$ and $\mathcal{O} = [1, N] \setminus [l, l+p]$. Since $\mathcal{I}$ is a blow-up set we have

$$\min\left(\liminf_{t\to T}\left(X_l - X_{l-1}\right), \liminf_{t\to T}\left(X_{l+p+1} - X_{l+p}\right)\right) > 0.$$

Therefore there exists $c > 0$ such that for any $j \in \mathcal{O}$, $i \in \mathcal{I}$ and $s \in [0, T]$, $|X_j(s) - X_i(s)| \geq \dfrac{1}{c}$, consequently $H_{\mathcal{IO},2} \leq c^2$ and the estimate 7.4.3 implies

$$\frac{d}{dt}\Pi_{\mathcal{I}}^2 \geq p\left(1 - \frac{\chi}{\chi_N^{p+1}}\right) - \frac{1}{c}\left(2 + \frac{2\chi}{\sqrt{N}}\right)\sqrt{\Pi_{\mathcal{I}}^2}. \tag{7.4.19}$$

If $p+1 < k$ then $\chi_N^{p+1} \geq \chi_N^{k-1} \geq \chi$ and $p\left(1 - \dfrac{\chi}{\chi_N^{p+1}}\right) > 0$. Thus as soon as

$$\frac{1}{c}\left(2 + \frac{2\chi}{\sqrt{N}}\right)\sqrt{\Pi_{\mathcal{I}}^2} \leq p\left(1 - \frac{\chi}{\chi_N^{p+1}}\right),$$

we have $\dfrac{d}{dt}\Pi_{\mathcal{I}}^2 > 0$. We deduce that

$$\varliminf_{t\to T}\Pi_{\mathcal{I}}^2 \geq c^2 p^2 \left(\frac{1 - \frac{\chi}{\chi_N^{p+1}}}{2 + \frac{2\chi}{\sqrt{N}}}\right)^2 > 0.$$

This is a contradiction with $\mathcal{I}$ being a blow-up set.

□

We can run the same argument with $k$ particles to get the intuition that the blow-up only aggregate the critical mass.

**Claim 4.** *A weak blow-up set may contains k particles.*

*Proof.* We again note $\mathcal{I}$ the particles contributing to the blow-up and $\mathcal{O}$ the other. Following the proof of the previous claim 3 we see that there exists $c > 0$ such that $H_{\mathcal{IO},2} \leq c^2$. Moreover $\liminf_{t\to T}\Pi_{\mathcal{I}}^2 = 0$ so there exists $t > 0$ such that

$$\Pi_{\mathcal{I}}^2 \leq -\frac{1}{2}c^2 p^2 \left(\frac{1 - \frac{\chi}{\chi_N^k}}{2 + \frac{2\chi}{\sqrt{N}}}\right)^2.$$

The equation 7.4.3, with this time $p+1=k$ and $\left(1-\frac{\chi}{\chi_N^k}\right)\leq 0$, implies for all $s>t$

$$\frac{1}{2}\frac{d}{dt}\Pi_{\mathcal{I}}^2 \leq \frac{k-1}{2}\left(1-\frac{\chi}{\chi_N^k}\right) < 0. \tag{7.4.20}$$

The second moment decreases and go to zero. □

The next theorem proves this result in the more general context of a strong blow up. The strong blow-up is indeed more general as we don't know anything about the particles at the boundary of the blow-up set $\mathcal{I}$ . The bad case would be the presence of oscillations. More precisely we are going to find basins of attractions where $k$ particles only will be aggregated. We define for $\chi_N^{k+1} < \chi < \chi_N^k$

$$D_{N,\chi}^{\varepsilon,\frac{c}{\varepsilon}} = \{X \in \mathbb{R}^N \text{ such that } \exists \ \ln > 0 \text{ with } \Pi_{\mathcal{I}}^2 \leq \varepsilon, \text{ and } H_{\mathcal{IO},2} < \frac{c}{\varepsilon}\}., \tag{7.4.21}$$

where $\mathcal{I}=[l,l+k-1]$ and $\mathcal{O}=[1,N]\setminus[l,l+k-1]$.
This set corresponds to $k$ particles really close on from each other, and all the other one far. Furthermore let

$$C_N \leq \min\left(\frac{(k-1)\left(\frac{\chi}{\chi_N^k}-1\right)}{8C_{4,2}\left(12+14\chi+4N^{1/4}\right)}, \frac{1}{4}\frac{(k-1)\left(\frac{\chi}{\chi_N^k}-1\right)}{\left(2+\frac{2\chi}{\sqrt{N}}\right)}\right). \tag{7.4.22}$$

**Theorem 7.4.8.** *If there exists $t\in[0,T]$ such that $X(t)\in D_{N,\chi}^{\varepsilon,\frac{C_N}{\varepsilon}}$ then we have a strong blow-up set with only $k$ particles. That is to say there is a strong blow-up aggregating only the critical mass.*

The idea of the proof is to show that we control the growth of $H_{\mathcal{IO},2}$ long enough to be sure that the blow-up effectively happen.

*Proof.* For simplicity we define $\alpha=-(k-1)\left(1-\frac{\chi}{\chi_N^k}\right)>0$. To prove this theorem we show that $\Pi_{\mathcal{I}}^2$ decreases and reaches 0 whereas $H_{\mathcal{IO},2}$ remains bounded. Our starting point is the equation 7.4.3 of lemma 7.4.6.

$$\frac{1}{2}\frac{d}{dt}\Pi_{\mathcal{I}}^2 \leq -\alpha + \left(2+\frac{2\chi}{\sqrt{N}}\right)\sqrt{\Pi_{\mathcal{I}}^2 H_{\mathcal{IO},2}}.$$

Thus as long as

$$\sqrt{\Pi_{\mathcal{I}}^2 H_{\mathcal{IO},2}} \leq \frac{1}{2}\frac{\alpha}{\left(2+\frac{2\chi}{\sqrt{N}}\right)}, \tag{7.4.23}$$

we get

$$\frac{1}{2}\frac{d}{dt}\Pi_{\mathcal{I}}^2 \leq -\frac{\alpha}{2}. \tag{7.4.24}$$

Integrating 7.4.24 from $t$ to $t+s$ we get

$$0 \leq \Pi^2 (t+s) \leq \Pi^2 (t) - \frac{\alpha}{2}s \leq \varepsilon - \frac{\alpha}{2}s.$$

Therefore under the condition 7.4.23 we find an upper bound for the blow-up time $T$.

$$T \leq t + \frac{2}{\alpha}\varepsilon. \tag{7.4.25}$$

Naturally the next step is to prove that starting at time $t$ with $X(t) \in D_{N,\chi}^{\varepsilon, \frac{C_N}{\varepsilon}}$ the estimate 7.4.23 is true for any $s \in [t, T]$. We already know that under the condition 7.4.23 the second moment decreases, so it suffices to prove that $H_{\mathcal{IO},2}$ remains bounded by $\frac{1}{\varepsilon}\frac{1}{2}\frac{\alpha}{\left(2+\frac{2\chi}{\sqrt{N}}\right)}$ up to $T$.

Since $X(t) \in D_{N,\chi}^{\varepsilon, \frac{C_N}{\varepsilon}}$ we have

$$H_{\mathcal{IO},2}^2(t) \leq \frac{C_N}{\varepsilon} \leq \frac{2C_N}{\varepsilon}.$$

Moreover thanks to the equation 7.4.6 of lemma 7.4.6 we control the growth of $H_{\mathcal{IO},2}$.

$$\frac{d}{dt}H_{\mathcal{IO},2} \leq C_{4,2}\left(12 + 14\chi + 4N^{1/4}\right) H_{\mathcal{IO},2}^2.$$

Therefore as long as $H_{\mathcal{IO},2} \leq \frac{2C_N}{\varepsilon}$ we have

$$\frac{d}{dt}H_{\mathcal{IO},2} \leq 4C_{4,2}\left(12 + 14\chi + 4N^{1/4}\right)\frac{C_N^2}{\varepsilon^2}. \tag{7.4.26}$$

Integrating 7.4.26 between $t$ and $t+s$ we find

$$H_{\mathcal{IO},2}(t+s) \leq H_{\mathcal{IO},2}2(t) + s4C_{4,2}\left(12 + 14\chi + 4N^{1/4}\right)\frac{C_N^2}{\varepsilon^2} \tag{7.4.27}$$

$$\leq \frac{C_N}{\varepsilon} + s4C_{4,2}\left(12 + 14\chi + 4N^{1/4}\right)\frac{C_N^2}{\varepsilon^2}. \tag{7.4.28}$$

Consequently for any $s > 0$ such that

$$\frac{C_N}{\varepsilon} + s4C_{4,2}\left(12 + 14\chi + 4N^{1/4}\right)\frac{C_N^2}{\varepsilon^2} \leq \frac{2C_N}{\varepsilon} \tag{7.4.29}$$

the inequality 7.4.26 holds and $H_{\mathcal{IO},2} \leq \frac{2C_N}{\varepsilon}$. We can rewrite 7.4.29 to obtain the condition

$$s \leq \frac{\varepsilon}{4C_{4,2}\left(12 + 14\chi + 4N^{1/4}\right)C_N}. \tag{7.4.30}$$

To conclude the proof we just have to be sure that 7.4.30 and 7.4.23 hold up to the blow-up. That is to say

$$\sqrt{\Pi_{\mathcal{I}}^2 H_{\mathcal{IO},2}} \leq 2C_N \leq \frac{1}{2}\frac{\alpha}{\left(2+\frac{2\chi}{\sqrt{N}}\right)}.$$

and

$$s \leq T - t \leq \frac{2}{\alpha}\varepsilon \leq \frac{\varepsilon}{4C_{4,2}\left(12 + 14\chi + 4N^{1/4}\right)C_N}.$$

Or equivalently

$$C_N \leq \frac{\alpha}{8C_{4,2}\left(12 + 14\chi + 4N^{1/4}\right)}.$$

It suffices to take

$$C_N \leq \min\left(\frac{(k-1)\left(\frac{\chi}{\chi_N^k} - 1\right)}{8C_{4,2}\left(12 + 14\chi + 4N^{1/4}\right)}, \frac{1}{4}\frac{(k-1)\left(\frac{\chi}{\chi_N^k} - 1\right)}{\left(2 + \frac{2\chi}{\sqrt{N}}\right)}\right). \tag{7.4.31}$$

That's exactly what we did in 7.4.22. Thus the standard deviation of the $\mathcal{I}$ particles reaches 0 whereas the exterior potential stays bounded from above, that is the definition of $\mathcal{I}$ being a strong blow-up set. In this case letting $\overline{X}$ be the blow up point and adapting the demonstration we easily prove that there exists $\bar{t}$ such that on $[\bar{t}, T]$ the squared distance $\overline{\Pi}_{\mathcal{I}}^2$ decreases and reaches 0 at time $T$. To summarize we proved that on $[t, T]$:

1. $\Pi_{\mathcal{I}}^2 \leq \varepsilon$.
2. $H_{\mathcal{IO},2} \leq \dfrac{2C_N}{\varepsilon}$.
3. $\dfrac{1}{2}\dfrac{d}{dt}\Pi_{\mathcal{I}}^2 \leq -\dfrac{\alpha}{2}$.
4. $\dfrac{d}{dt}H_{\mathcal{IO},2} \leq \dfrac{\beta}{\varepsilon^2}$.
5. On $[\bar{t}, T]$, $\dfrac{1}{2}\dfrac{d}{dt}\overline{\Pi}_{\mathcal{I}}^2 \leq -\dfrac{\alpha}{2}$,

where $\beta = 4C_{4,2}\left(12 + 14\chi + 4N^{1/4}\right)C_N^2$. It allows us to give a slightly more precise theorem.

**Theorem 7.4.9.** *If there exists* $t \in [0, T]$ *such that* $X(t) \in D_{N,\chi}^{\varepsilon, \frac{C_N}{\varepsilon}}$ *then for any* $s \in \left[0, \frac{2}{\alpha}\varepsilon\right]$,

$$X(t+s) \in D_{N,\chi}^{\varepsilon - s\frac{\alpha}{2}, \frac{C_N}{\varepsilon} + s\frac{\beta}{\varepsilon^2}}.$$

*In particular there exists a strong blow-up set $\mathcal{I}$ which aggregates only k particles.*

With this theorem we can see the suite $D_{N,\chi}^{\varepsilon - s\frac{\alpha}{2}, \frac{C_N}{\varepsilon} + s\frac{\beta}{\varepsilon^2}}$ as a Lyapunov function. The set $D_{N,\chi}^{\varepsilon, \frac{C_N}{\varepsilon}}$ are basins of attraction. □

## 7.5 Rigidity

In this Section we demonstrate that inside the stability set, the blow-up process is rigid in the following sense: particles in the inner set $\mathcal{I}$ blow-up with the same rate, whereas particles in the outer set $\mathcal{O}$ stay away from the blow-up point.

**Theorem 7.5.1.** *Let $X$ be a solution of 7.1.1. Assume there exists $t_0 > 0$ such that $X(t_0) \in D_{N,\chi}^{\varepsilon,\frac{C_N}{\varepsilon}}$ then let $\mathcal{I} = [l, l+k-1]$ be the inner set and $(X_i)_{i\in\mathcal{I}}$ be the particles inside. We denote by $T$ the blow-up time and $\overline{X}$ the blow-up point. Then there exists $A > 0$, depending only on $t_0, \varepsilon$ and $N$, such that*

1. $\dfrac{1}{A^2} \leq \dfrac{\overline{\Pi}_{\mathcal{I}}^2(t)}{2\alpha(T-t)} \leq A^2$, *furthermore we have* $\overline{\Pi}_{\mathcal{I}}^2(t) \sim 2\alpha(T-t)$ *as* $t \to T$.
   $\dfrac{1}{A^2} \leq \dfrac{\Pi_{\mathcal{I}}^2(t)}{2\alpha(T-t)} \leq A^2$, *furthermore we have* $\Pi_{\mathcal{I}}^2(t) \sim 2\alpha(T-t)$ *as* $t \to T$.

2. $\forall i \in \mathcal{I} \quad Y_i(\tau(t)) = \dfrac{|X_i(t) - \overline{X}|}{\sqrt{2\alpha(T-t)}} \leq A,$

3. $\forall (i,j) \in \mathcal{I}\times\mathcal{I} \quad \dfrac{1}{A} \leq \dfrac{|X_i(t) - X_j(t)|}{\sqrt{2\alpha(T-t)}} \leq A,$

4. $\forall j \in \mathcal{O} \quad Y_j(\tau(t)) = \dfrac{|X_j(t) - \overline{X}|}{\sqrt{2\alpha(T-t)}} \geq \dfrac{1}{A\sqrt{2\alpha(T-t)}},$

where $\tau(t) = -\dfrac{1}{\alpha}\log\left(\dfrac{R(t)}{R(0)}\right)$ and $R(t) = \sqrt{2\alpha(T-t)}$.

This theorem means that if we rescale the equation around $\overline{X}$ with $\sqrt{2\alpha(T-t)}$ then the new solution $Y$ exists for all time, is bounded for the blow-up indices $\mathcal{I}$ and not bounded for the other one $\mathcal{O}$.

**Remark 7.5.2.** *Statements of Theorem 7.5.1 are stronger than the max. principle. In the continuous case the maximum principle implies, for $\rho$ solution of $\frac{d}{dt}\rho = \Delta\rho - \chi\nabla\rho\nabla\kappa \star \rho$, that*

$$\frac{d}{dt}\left(\max_{x\in\mathbb{R}} \rho\right) \leq \left(\max_{x\in\mathbb{R}} \rho\right)^2.$$

*Integrating from 0 to $t$ we obtain*

$$\frac{1}{\rho} \geq \frac{1}{\rho_0} - \chi t. \tag{7.5.1}$$

*Integrating from $t$ to $T$ we get*

$$\frac{1}{\rho} \leq \chi(T-t). \tag{7.5.2}$$

*In the discrete case the analogue computation is to consider $X^m = \min_{i\in[1,N]}(X_{i+1} - X_i)$. For this quantity we can bound the way it decrease.*

$$X^m \dot{X}^m \geq -4\chi h_N.$$

*We deduce*

$$(X^m)^2 \geq (X^m)^2(0) - \frac{\chi}{\chi_N^2}t = \gamma^2(\beta T - t), \tag{7.5.3}$$

*and*

$$(X^m)^2 \leq \frac{\chi}{\chi_N^2}(T-t) = \gamma^2(T-t), \tag{7.5.4}$$

*where $T$ is the blow-up time. The first estimation 7.5.3 is not enough to rescale the equation, as the solution can blow-up a long time after $\beta T$. Moreover it does not give any information about the number of particles contributing to the blow-up. The second equation is nearly empty, it would have been better to have a $\max_{i\in\mathcal{I}}$. For all those reasons we have to be more accurate on our estimates. Again the well adapted tools are the second moment and the exterior potential defined in 7.4.4.*

*Proof.* We split the proof into several estimates, corresponding to the different items of Theorem 7.5.1.

**Estimate 1- The squared distance to the blow-up point is estimated from above and below.** We start by the first estimate. By 7.4.4 of lemma 7.4.6, with $p+1=k$ and $\alpha = -(k-1)\left(1-\frac{\chi}{\chi_N^k}\right)$ we have

$$|\frac{1}{2}\frac{d}{dt}\overline{\Pi}_{\mathcal{I}}^2 + \alpha| \leq \left(2+\frac{2\chi}{\sqrt{N}}\right)\sqrt{\overline{\Pi}_{\mathcal{I}}^2 H_{\mathcal{I}\mathcal{O},2}}.$$

Since $H_{\mathcal{I}\mathcal{O},2}$ is bounded and $\overline{\Pi}_{\mathcal{I}}^2 \to 0$ as $t\to T$, there exists $t_{1/2}$ such that for any $s\in[t_{1/2},T]$

$$-2\alpha \leq \frac{1}{2}\frac{d}{dt}\overline{\Pi}_{\mathcal{I}}^2(s) \leq -\frac{\alpha}{2}. \tag{7.5.5}$$

Moreover there exists $t_\varepsilon$ such that for any $s\in[t_\varepsilon,T]$

$$-\alpha(1-\varepsilon) \leq \frac{1}{2}\frac{d}{dt}\overline{\Pi}_{\mathcal{I}}^2(s) \leq \alpha(1+\varepsilon). \tag{7.5.6}$$

As $\overline{\Pi}_{\mathcal{I}}^2=0$, integrating from $t\geq t_{1/2}$ to $T$ 7.5.5 (*resp.* $t\geq t_\varepsilon$ *7.5.6*) we obtain:

$$\frac{\alpha}{2}(T-t) \leq \frac{1}{2}\overline{\Pi}_{\mathcal{I}}^2(t) \leq 2\alpha(T-t), \tag{7.5.7}$$

and

$$\alpha(1-\varepsilon)(T-t) \leq \frac{1}{2}\overline{\Pi}_{\mathcal{I}}^2(t) \leq \alpha(1+\varepsilon)(T-t). \tag{7.5.8}$$

The equation 7.5.7 gives a $A^2\geq\sqrt{2}$ such that for any $s\in[0,T]$

$$\frac{1}{A^2} \leq \frac{\overline{\Pi}_{\mathcal{I}}^2(s)}{2\alpha(T-t)} \leq A^2.$$

Letting $\varepsilon$ going to 0 7.5.8 gives $\overline{\Pi}_{\mathcal{I}}^2(t)\sim 2\alpha(T-t)$ as $t\to T$.
The proof is exactly the same when we change $\overline{\Pi}_{\mathcal{I}}^2$ by $\Pi_{\mathcal{I}}^2$.

**Estimate 2- In the blow-up set the rescaled solution is bounded from above.** This estimate proves that, for the inner set the rescaled solution is bounded from above. It is a straightforward consequence of 7.5.8, for any $i \in \mathcal{I}$

$$\frac{|X_i - \overline{X}|}{\sqrt{2\alpha\,(T-t)}} \leq \sqrt{\frac{\overline{\Pi}^2_{\mathcal{I}}(s)}{2\alpha\,(T-t)}} \leq A. \tag{7.5.9}$$

**Estimate 4- the rescaled particles in the outer set go to infinity.** We prove estimate 4 now as we need it in the proof of the third estimate. The key tool here is the upper bound on $H_{\mathcal{IO},2}$. The constant $A$ is not fixed and can be taken larger when needed. By hypothesis $X\,(t_0) \in D^{\varepsilon,\frac{C_N}{\varepsilon}}_{N,\chi}$ thus theorem 7.4.9 tells us that $H_{\mathcal{IO},2}$ is bounded above, say by $A^2$. We deduce a lower bound for $|Y_l - Y_{l-1}|$ and $|Y_{l+k} - Y_{l+k-1}|$.

$$\begin{aligned}\max\left(\frac{1}{|Y_{l+k} - Y_{l+k-1}|^2}, \frac{1}{|Y_l - Y_{l-1}|^2}\right) &\leq \max\left(\frac{2\alpha\,(T-t)}{|X_{l+k} - X_{l+k-1}|^2}, \frac{2\alpha\,(T-t)}{|X_l - X_{l-1}|^2}\right)\\ &\leq 2\alpha\,(T-t)\,H_{\mathcal{IO},2} \leq 2\alpha\,(T-t)\,A^2.\end{aligned}$$

Thus

$$\min\left(|Y_{l+k} - Y_{l+k-1}|, |Y_l - Y_{l-1}|\right) \geq \frac{1}{\sqrt{2\alpha\,(T-t)}A}. \tag{7.5.10}$$

In particular 7.5.10 says that both $|Y_{l+k} - Y_{l+k-1}|$ and $|Y_l - Y_{l-1}|$ are bounded from below. This remark will be useful during the proof of the third estimate. Anyway we already proved the second estimate is implies $\max\left(Y_{l+k-1}, Y_l\right) \leq A$. So taking $A$ larger if we need we find

$$\forall j \in \mathcal{O} \quad Y_j \geq \min\left(|Y_{l+k}|, |Y_l|\right) \geq \frac{1}{\sqrt{2\alpha\,(T-t)}A}.$$

In particular all the rescaled particles in $\mathcal{O}$ are sent to infinity.

**Estimate 3a- the rescaled relative distances in the inner set are bounded from above.** This time the estimate comes from 7.5.9, for any $(i,j) \in \mathcal{I} \times \mathcal{I}$

$$\frac{|X_i - X_j|}{\sqrt{2\alpha\,(T-t)}} \leq \frac{|X_i - \overline{X}|}{\sqrt{2\alpha\,(T-t)}} + \frac{|X_j - \overline{X}|}{\sqrt{2\alpha\,(T-t)}} \leq 2A. \tag{7.5.11}$$

This estimate is exactly one of those we failed to obtain with the maximal principle.

**Estimate 3b- the relative distances in the inner set are bounded from below.** This estimate is the core of our rigidity Theorem. Together with the estimate 3a it expresses that the particles blow-up with the same rate, homogeneously inside the inner set. This has an important consequence: the free energy of the inner set in the rescaled frame is bounded from below, therefore the particles converge towards a critical point of the free energy. We begin with some useful definitions

**Definition 7.5.3.** *For any $q, p \in \mathcal{I}$, $q < p$ we define the average $\overline{Y}_{q,p} = \dfrac{1}{p-q+1}\displaystyle\sum_{i=q}^{p} Y_i$, the pseudo inner set $\mathcal{I}^p_q = [q,p]$, the pseudo exterior set $\mathcal{O}^q_p = [1,N] \setminus [q,p]$ and the corresponding* *standard deviation and exterior squared potential by*

$$P_{q,p}^2 = \sum_{i\in\mathcal{I}_q^p}\left(\frac{X_i}{\sqrt{2\alpha(T-t)}} - \overline{Y}_{q,p}\right)^2 = \sum_{i\in\mathcal{I}_q^p}\left(Y_i - \overline{Y}_{q,p}\right)^2.$$

$$\mathrm{H}_{q,p,2} = \sum_{j\in\mathcal{O}_p^q}\sum_{i\in\mathcal{I}_q^p}\frac{2\alpha(T-t)}{(X_j-X_i)^2} = \sum_{j\in\mathcal{O}_p^q}\sum_{i\in\mathcal{I}_q^p}\frac{1}{(Y_j-Y_i)^2}$$

*For* $p = l+k-1$, $q = l$ *we have* $P_{q,p}^2 = P_{\mathcal{I}}^2$.

We proceed by induction on $q < p$ to control the partial standard deviation of all subsets of inner particles. The basic idea is the following. We face the alternative: either the standard deviation of all but the right-most particle is large, and we are done; or it is small, and the two right-most particles are far from each other. The last statement implies that the standard deviation of all but the right-most particle increases. Consequently the partial standard deviation cannot be too small.

To make this argument clear, we compute the evolution of the various quantities which are involved in the induction. This is the purpose of the next lemma, which is very similar to Lemma 7.4.6.

**Lemma 7.5.4.** *Let* $\alpha_q^p = (p-q)\left(1-\chi\frac{p-q+1}{N+1}\right) = (p-q)\left(1-\frac{\chi}{\chi_{N+1}^{p-q+1}}\right)$, *we have*

$$\alpha_q^p + \alpha P_{q,p}^2 - \left(2+\frac{2\chi}{\sqrt{N}}\right)\sqrt{P_{q,p}^2\mathrm{H}_{q,p,2}} \le \frac{d}{dt}P_{q,p}^2 \le \alpha_q^p + \alpha P_{q,p}^2 + \left(2+\frac{2\chi}{\sqrt{N}}\right)\sqrt{P_{q,p}^2\mathrm{H}_{q,p,2}}. \quad (7.5.12)$$

**Corollary 7.5.5.** *We deduce two different estimates regarding the number of particles* $p-q+1$.

1. *If* $p-q+1 \le k-1$ *i.e.* $\alpha_q^p > 0$ *and* $\sqrt{P_{q,p}^2\mathrm{H}_{q,p,2}} \le \frac{\alpha_q^p}{2\left(2+\frac{2\chi}{\sqrt{N}}\right)}$ *then*

$$\frac{d}{dt}P_{q,p}^2 \ge \frac{\alpha_q^p}{2} + \alpha P_{q,p}^2.$$

2. *If* $p-q+1 \ge k$ *i.e.* $\alpha_q^p < 0$ *and* $\sqrt{P_{q,p}^2\mathrm{H}_{q,p,2}} \le -\frac{\alpha_q^p}{2\left(2+\frac{2\chi}{\sqrt{N}}\right)}$ *then*

$$\frac{d}{dt}P_{q,p}^2 \le -\frac{\alpha_q^p}{2} + \alpha P_{q,p}^2.$$

*Proof.* The proof is a direct computation similar as the one done for the proof of Lemma 7.4.6. The only difference is that a new term pops up for $\frac{d}{dt}P_{q,p}^2$ : $\alpha\sum_{i\in\mathcal{I}_q^p}\left(Y_i-\overline{Y}_{q,p}\right)Y_i$. To deal with it we remark that

$$\begin{aligned}\sum_{i\in\mathcal{I}_q^p}\left(Y_i-\overline{Y}_{q,p}\right)Y_i &= \sum_{i\in\mathcal{I}_q^p}\left(Y_i-\overline{Y}_{q,p}\right)^2+\sum_{i\in\mathcal{I}_q^p}\left(Y_i-\overline{Y}_{q,p}\right)\overline{Y}_{q,p}\\ &= \sum_{i\in\mathcal{I}_q^p}\left(Y_i-\overline{Y}_{q,p}\right)^2+\overline{Y}_{q,p}\left(-\left(p-q+1\right)\overline{Y}_{q,p}+\sum_{i\in\mathcal{I}_q^p}Y_i\right)\\ &= \sum_{i\in\mathcal{I}_q^p}\left(Y_i-\overline{Y}_{q,p}\right)^2=P_{q,p}^2,\end{aligned}$$

which explains the additional term $\alpha P_{q,p}^2$ in Lemma 7.5.12 and Corollary 7.5.5 with respect to Lemma 7.4.6. □

The next proposition is the core of the induction, we identify two cases corresponding to two different steps.

**Proposition 7.5.6** (Induction)**.** *Let $q,p\in\mathcal{I}$. If there exists $B>0$ satisfying*

$$\begin{cases}P_{q,p}^2\geq\dfrac{1}{B^2}\\ |Y_{q-1}-Y_q|\geq\dfrac{1}{B}\end{cases}\quad\textit{(descent step)}$$

*or*

$$\begin{cases}|Y_p-Y_{p-1}|\geq\dfrac{1}{B}\\ |Y_{q-1}-Y_q|\geq\dfrac{1}{B}\end{cases}\quad\textit{(reinitialization step)}$$

*then there exists $B_s>0$ such that $P_{q,p-1}^2>\dfrac{1}{B_s^2}$.*

To illustrate the proof in both case we refer to figures 7.5 and 7.5.

*Proof.* We distinguish between the descent case and the reinitialization step.

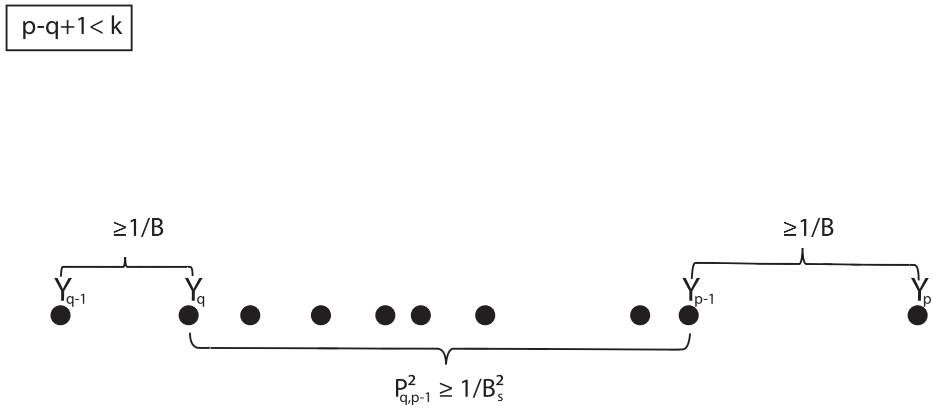
**1- The reinitialization step.** In this case we can bound from above the exterior potential $\mathrm{H}_{q,p-1,2}$ and deduce that the standard deviation $P_{q,p-1}^2$ stays away from 0.

$$\mathrm{H}_{q,p-1,2}=\sum_{j\in\mathcal{O}_{p-1}^q}\sum_{i\in\mathcal{I}_q^{p-1}}\frac{1}{\left(Y_j-Y_i\right)^2} \tag{7.5.13}$$

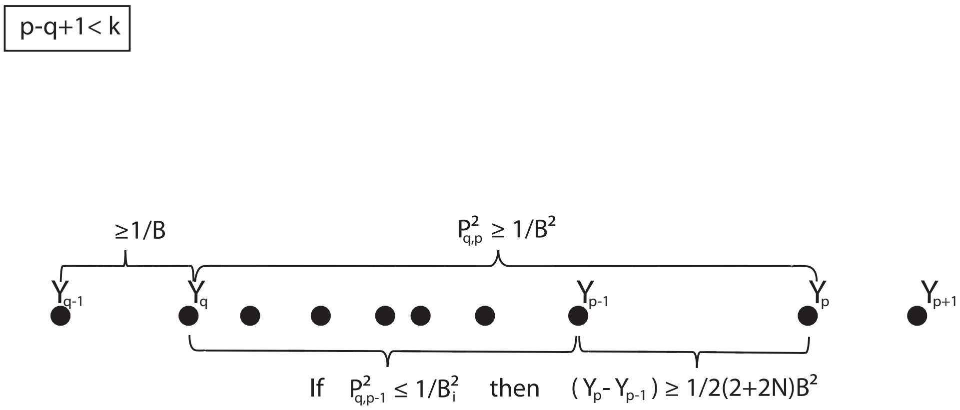
$$\leq N^2\min\left(\frac{1}{|Y_{q-1}-Y_q|^2},\frac{1}{|Y_p-Y_{p-1}|^2}\right)\leq N^2B^2. \tag{7.5.14}$$

Furthermore as long as

$$P_{q,p-1}^2\leq\frac{1}{N^2B^2}\left(\frac{\alpha_q^{p-1}}{2\left(2+\frac{2\chi}{\sqrt{N}}\right)}\right)^2$$




In the reinitialisation case, $P^2_{q,p-1}$ cannot be too small




In the descent case, $P^2_{q,p-1}$ cannot be too small

we have

$$\sqrt{P^2_{q,p-1}\mathrm{H}_{q,p-1,2}} \leq \frac{\alpha_q^p - 1}{2\left(2+\frac{2\chi}{\sqrt{N}}\right)}$$

Plugging this into Corollary 7.5.5 together with $p-q \leq k-1$ we get that the standard deviation $P^2_{q,p-1}$ increase:

$$\frac{d}{dt}P^2_{q,p-1} \geq \frac{\alpha_q^{p-1}}{2} + \alpha P^2_{q,p-1} > 0.$$

We easily deduce the existence of $B_s$.

$$P^2_{q,p-1} \geq \min\left(P^2_{q,p-1}(0), \frac{1}{N^2B^2}\left(\frac{\alpha_q^{p-1}}{2\left(2+\frac{2\chi}{\sqrt{N}}\right)}\right)^2\right) = \frac{1}{B_s^2}$$

**2- The descent step.** In this case we are not able to bound a priori $\mathrm{H}_{q,p-1,2}$ from above. Anyway we will show that under the condition $P^2_{q,p-1}$ small and $P^2_{q,p}$ large we obtain this bound. We start by making a link between $P^2_{q,p-1}$, $P^2_{q,p}$ and $(Y_p - Y_{p-1})^2$.

$$\begin{aligned} P^2_{q,p} &= \sum_{i\in\mathcal{I}^p_q} \left(Y_i - \overline{Y}_{q,p}\right)^2 = \left(Y_p - \overline{Y}_{q,p}\right)^2 + \sum_{i=q}^{p-1} \left(Y_i - \overline{Y}_{q,p-1} + \overline{Y}_{q,p-1} - \overline{Y}_{q,p}\right)^2 \\ &= \left(Y_p - \overline{Y}_{q,p}\right)^2 + \sum_{i\in\mathcal{I}^{p-1}_q} \left(Y_i - \overline{Y}_{q,p-1}\right)^2 + \sum_{i\in\mathcal{I}^{p-1}_q} \left(\overline{Y}_{q,p-1} - \overline{Y}_{q,p}\right)^2 \\ &+ \sum_{i\in\mathcal{I}^{p-1}_q} \left(Y_i - \overline{Y}_{q,p-1}\right)\left(\overline{Y}_{q,p-1} - \overline{Y}_{q,p}\right) \\ &= \left(Y_p - \overline{Y}_{q,p}\right)^2 + P^2_{q,p-1} + (p-q)\left(\overline{Y}_{q,p-1} - \overline{Y}_{q,p}\right)^2 \end{aligned} \tag{7.5.15}$$

By convexity we have

$$\begin{aligned} \left(\overline{Y}_{q,p} - \overline{Y}_{q,p-1}\right)^2 \le \left(Y_p - \overline{Y}_{q,p-1}\right)^2 &\le \left(Y_p - Y_{p-1} + Y_{p-1} - \overline{Y}_{q,p-1}\right)^2 \\ &\le 2\left(Y_p - Y_{p-1}\right)^2 + 2\left(Y_{p-1} - \overline{Y}_{q,p-1}\right)^2 \\ &\le 2\left(Y_p - Y_{p-1}\right)^2 + 2P^2_{q,p-1}. \end{aligned} \tag{7.5.16}$$

and

$$\left(Y_p - \overline{Y}_{q,p}\right)^2 \le \left(Y_p - \overline{Y}_{q,p-1}\right)^2 \le 2\left(Y_p - Y_{p-1}\right)^2 + 2P^2_{q,p-1}. \tag{7.5.17}$$

Plugging 7.5.16 and 7.5.17 in 7.5.15 we obtain

$$\frac{1}{B^2} \le P^2_{q,p} \le (3+2N)\,P^2_{q,p-1} + (2+2N)\left(Y_{p+1} - Y_p\right)^2 \tag{7.5.18}$$

With 7.5.18 we see that $P^2_{q,p-1}$ and $(Y_{p+1} - Y_p)^2$ cannot be small at the same time. Precisely for any $B_i > 0$ two case may appear: either $P^2_{q,p-1} \ge \frac{1}{B_i^2}$ or $P^2_{q,p-1} \le \frac{1}{B_i^2}$.

**In the last case** the equation 7.5.18 gives a lower bound for $(Y_p - Y_{p-1})^2$.

$$\left(Y_p - Y_{p-1}\right)^2 \ge \frac{1}{2+2N}\left(\frac{1}{B^2} - \frac{3+2N}{B_i^2}\right).$$

Taking $B_i$ large enough, for example $B_i^2 \ge 2\,(3+2N)\,B^2$, we obtain

$$\left(Y_p - Y_{p-1}\right)^2 \ge \frac{1}{2(2+2N)}\frac{1}{B^2}.$$

On the other side of the pseudo inner set $\mathcal{I}^{p-1}_q$ the hypothesis is $(Y_q - Y_{q-1})^2 \ge \frac{1}{B^2}$. Together with the equation7.5 we deduce an upper bound for $\mathrm{H}_{q,p-1,2}$ and therefore an upper bound for

$\sqrt{P^2_{q,p-1}\mathrm{H}_{q,p-1,2}}$.

$$\begin{aligned}\mathrm{H}_{q,p-1,2} &= \sum_{j\in\mathcal{O}_q^{p-1}}\sum_{i\in\mathcal{I}_q^{p-1}}\frac{1}{(Y_j-Y_i)^2}\\ &\le \sum_{j<q}\sum_{i\in\mathcal{I}_q^{p-1}}\frac{1}{(Y_j-Y_i)^2}+\sum_{j>p-1}\sum_{i\in\mathcal{I}_q^{p-1}}\frac{1}{(Y_j-Y_i)^2}\\ &\le N^2\left(B^2+2(2+2N)B^2\right)\le N^2\left(5+4N\right)B^2.\end{aligned}$$

Then taking $B_i$ larger if we need, such that $B_i \ge N\sqrt{5+4N}B\dfrac{2\left(2+\frac{2\chi}{\sqrt{N}}\right)}{\alpha_q^{p-1}}$:

$$\begin{aligned}\sqrt{P^2_{q,p-1}\mathrm{H}_{q,p-1,2}} &\le \frac{1}{B_i}N\sqrt{5+4N}B\\ &\le \frac{\alpha_q^{p-1}}{2\left(2+\frac{2\chi}{\sqrt{N}}\right)}.\end{aligned} \tag{7.5.19}$$

We thus fulfil the hypotheses of Corollary 7.5.5 and get that $P^2_{q,p-1}$ increase.
**In any case** either $P^2_{q,p-1}$ is large or $P^2_{q,p-1}$ increase. We therefore deduce a lower bound for $P^2_{q,p-1}$:

$$\begin{aligned}P^2_{q,p-1} &\ge \min\left(P^2_{q,p-1}(0),\frac{1}{B_i^2}\right)\\ &\ge \min\left(P^2_{q,p-1}(0),\frac{1}{2\left(3+2N\right)B^2},\frac{\alpha_q^{p-1}}{4\left(2+\frac{2\chi}{\sqrt{N}}\right)^2N^2\left(5+4N\right)B^2}\right)=\frac{1}{B_s^2}.\end{aligned}$$

It proves the descent case of Proposition 7.5.6 and finishes the proof of this proposition.

□

Armed with the induction proposition 7.5.6 we are now ready to prove the estimate 3b. The goal is to bound from below all the relative distances in the inner set. The strategy is to isolate the left most relative distance with the descent case of proposition 7.5.6, this is the local induction. Then we exclude the left most particles with the reinitialization case and repeat the first step. One by one we bound from below every relative distance. We recall that $\mathcal{I}=[l,l+k-1]$.

**Step 1- A lower bound for $|Y_{l+1}-Y_l|$: the local induction.** The first estimate of Theorem 7.5.1 says that $P^2_{\mathcal{I}}$ is bounded from below by $\dfrac{1}{A^2}$. In the proof of the forth estimate we showed that *(equation 7.5.10)* taking $A$ larger if we need $|Y_l-Y_{l-1}|\ge\dfrac{1}{A}$. With $q=l$ and $p=l+k-1$ we are exactly in the descent case of Proposition 7.5.6. It gives us $B_{s_1}>B>0$ such that $P^2_{l,l+k-2}\ge\dfrac{1}{B^2_{s_1}}$. Since $|Y_l-Y_{l-1}|\ge\dfrac{1}{A}\ge\dfrac{1}{B_s}$ we again apply the descent case of proposition 7.5.6 with this time $q=l$ and $p=l+k-2$ to win one more notch on the $p$ index. We repeat the same argument for $p$ down to $p=l+2$ and obtain with the last iteration a lower bound, say

$\dfrac{1}{B_s^2}$, for $P^2_{l,l+1}$. It gives us a lower bound for $|Y_{l+1} - Y_l|$:

$$\begin{aligned}(Y_{l+1} - Y_l)^2 &\geq \left(Y_{l+1} - \overline{Y}_{q,p}\right)^2 = P^2_{l,l+1} \\ &\geq \frac{1}{B_s^2}.\end{aligned} \tag{7.5.20}$$

**Step 2- Not so fast: reinitialization.** After the first step we naturally want to exclude the left most particle: $(X_l)$ and start over the local induction. Unfortunately it is not so simple as we know nothing about $P^2_{l+1,l+k-1}$. This is why a reinitialization step is needed.

Once again in the proof of the forth estimate we showed that *(equation 7.5.10)* $|Y_{l+k} - Y_{l+k-1}|$ is bounded from below. Few lines above we showed in 7.5.20 that $|Y_{l+1} - Y_l|$ is also bounded from below. Therefore we fulfil the hypotheses of the reinitialization case of proposition 7.5.6 with $q = l+1$ and $p = l+k-1$. The conclusion is exactly what we were looking for: a lower bound $\dfrac{1}{B_s^2}$ for $P^2_{l+1,l+k-1}$.

**Step 3- Yes we can: The global induction.** We explain here the global induction step. Once the reinitialization step done we can run a local induction argument similar to step 1. That is to say for any $q \in \mathcal{I}$ such that $P^2_{q,l+k-1}$ and $|Y_q - Y_{q-1}|$ are bounded from below we prove, with the local induction, that $|Y_{q+1} - Y_q|$ is bounded from below. We follow with a reinitialization step: since $|Y_{q+1} - Y_q|$ and $|Y_{l+k} - Y_{l+k-1}|$ are bounded from below we know that $P^2_{q+1,l+k-1}$ is also bounded from below. And so forth starting with $q = l$ we can go up to $q = l+k-2$. In doing so we proved that there exists $B > 0$ such that:

$$\forall i \in \mathcal{I} \setminus \{l+k-1\}, \quad |Y_{i+1} - Y_i| \geq \frac{1}{B}.$$

It trivially implies the estimate 3b and concludes the proof of Theorem 7.5.1.

□

### 7.5.a The rescaled syteme

We want to catch the profile of the blow-up. According to theorem 7.5.1 we perform the following parabolic rescaling:

$$Y(\tau(t)) = \frac{X(t) - \overline{X}}{R(t)}, \tag{7.5.21}$$

where $T$ is the blow-up time, $\overline{X}$ the blow-up point, $R(t) = \sqrt{2\alpha(T-t)}$ and $\tau(t) = -\dfrac{1}{\alpha}\log\left(\dfrac{R(t)}{R(0)}\right)$. The definition 7.1.1 becomes

$$\begin{cases} \dot{Y}(t) = -\nabla \mathbb{E}_r(Y(t)) & t \in \mathbb{R}_+ \\ Y(0) = Y^0 & Y^0 \in \mathbb{R}^N. \end{cases} \tag{7.5.22}$$

where

$$\mathbb{E}_r(Y) = -\sum_{i=1}^{N-1} \log(Y_{i+1} - Y_i) + \chi h_N \sum_{1 \leq i \neq j \leq N} \log|Y_i - Y_j| - \frac{\alpha}{2}|Y|^2. \tag{7.5.23}$$

We can write it explicitly:

$$\begin{array}{rcccl}
\dot{Y}_1 = & -\dfrac{1}{Y_2 - Y_1} & + & 2\chi h_N \displaystyle\sum_{j \neq 1} \frac{1}{Y_j - Y_1} & +\alpha Y_1 \\
\dot{Y}_i = & -\dfrac{1}{Y_{i+1} - Y_i} + \dfrac{1}{Y_i - Y_{i-1}} & + & 2\chi h_N \displaystyle\sum_{j \neq i} \frac{1}{Y_j - Y_i} & +\alpha Y_i \\
\dot{Y}_N = & \dfrac{1}{Y_N - Y_{N-1}} & + & 2\chi h_N \displaystyle\sum_{j \neq N} \frac{1}{Y_i - Y_N} & +\alpha Y_N.
\end{array} \tag{7.5.24}$$

The center of mass $c_y = \sum_1^n Y_i$ satisfies $\dot{c}_y = \alpha c_y$ but we cannot fix $c_y(0) = 0$ this quantity is given by $\overline{X}$. We can consider the variables $v_i = Y_{i+1} - Y_i$ but we need to be careful coming back to $Y$. The goal now is to well understand the structure of the blow-up. The idea is that the solutions of 7.5.22 coming from 7.1.1 are special, existing for all non-negative times and bounded in some variables. To track them we would like to use a Liouville theorem on 7.5.24. We first recall the specificity of a rescaled solution for 7.5.22.

**Definition 7.5.7.** *[The rescaled condition.] Let $Y$ be a solution of the differential equation 7.5.24. We say that $Y$ satisfies the rescaled condition if there exists $A > 0$ and a set $\mathcal{I}$ with $|\mathcal{I}| = k$ such that the following conditions hold true.*

1. *$Y$ is define for all nonnegative time.*
2. *$\forall i \in \mathcal{I} \quad Y_i \leq A$.*
3. *$\forall (i, i+1) \in \mathcal{I} \times \mathcal{I} \quad (Y_{i+1} - Y_i) \geq \frac{1}{A}$.*
4. *$\forall j \in [1, N] \setminus \mathcal{I} = \mathcal{O} \quad |Y_i| \underset{t \to +\infty}{\longrightarrow} +\infty$.*
5. *$\forall t \in \mathbb{R}^+ \quad \mathrm{H}_{\mathcal{IO},2}(t) \leq A^2 e^{-2\alpha t}$.*

The rescaled conditions are directly inspired by Theorem 7.5.1. Moreover we can state the following proposition.

**Proposition 7.5.8.** *Let $X$ be a solution of the differential equation 7.1.1 and let $Y$ be the rescaled solution of $X$ defined by 7.5.21. Then $Y$ satisfies the rescaled condition 7.5.7.*

*Proof.* The theorem 7.5.1 immediately implies the rescaled condition 7.5.7. □

Under this condition we would like to prove the following conjecture.

**Conjecture 7.5.9.** *[Liouville theorem for 7.5.24.] Let $\chi$ near $\chi_N^k$. There exists $Y_\infty \in \mathbb{R}^k$ such that for any $Y$ solution of the differential equation 7.5.24 satisfying the rescaled condition 7.5.7 we have $(Y_i)_{i \in \mathcal{I}} \to Y_\infty$ as $t \to \infty$.*

We are unfortunately unable to prove this conjecture at this stage; nevertheless the following proposition is a step in this direction.
We first need to introduce the $^{th}$ energy functional.

**Definition 7.5.10.** *[The $^{th}$ energy functional.] As usual we fix an inner set $\mathcal{I} = [l, l+k-1]$. We define $\mathbb{E}_k$ by:*

$$\mathbb{E}_k(Y) = -\sum_{i\in\mathcal{I}\setminus\{l+k-1\}} \log(Y_{i+1} - Y_i) + \chi h_N \sum_{(i,j)\in\mathcal{I}\times\mathcal{I}\setminus\{i\}} \log|Y_i - Y_j| - \frac{\alpha}{2}\sum_{i\in\mathcal{I}} |Y_i|^2.$$

This energy as to be understand as the energy of the inner set particles. Under the rescaled condition 7.5.7, we observe that the $^{th}$ energy is bounded from above and below.
We also define the gradient flow equation associated to $\mathbb{E}_k$:

$$\begin{cases} \dot{Z}(t) = -\nabla\mathbb{E}_k(Z(t)) & (t, Z) \in \mathbb{R}_+ \times \mathbb{R}^k \\ Z(0) = Z^0 & Z^0 \in \mathbb{R}^N. \end{cases} \tag{7.5.25}$$

Our first proposition is to remark that a rescaled solution behave almost like a solution of the $^{th}$ gradient flow.

**Proposition 7.5.11.** *For $(Y)$ solution of the differential equation 7.5.24 satisfying the rescaled condition 7.5.7 there exists $C > 0$ such that*

$$\left|\nabla\mathbb{E}_k((Y_i)_{i\in\mathcal{I}}) - (\nabla_i\mathbb{E}_r(Y))_{i\in\mathcal{I}}\right| \le Ce^{-\alpha t}.$$

*Proof.* We use the rescale condition 7.5.7 to obtain $A$ such that $\mathrm{H}_{\mathcal{IO},2}(t) \le A^2 e^{-2\alpha t}$. Then we compute for any $i \in \mathcal{I} = [l, l+k-1]$:

$$\begin{aligned} |\nabla\mathbb{E}_k(Y_i) - \nabla_i\mathbb{E}_r(Y)| &= \left| -\frac{\delta_{i,l}}{Y_l - Y_{l-1}} + \frac{\delta_{i,l+k}}{Y_{l+k+1} - Y_{l+k}} - 2\chi h_n \sum_{k\in\mathcal{O}} \frac{1}{Y_k - Y_i} \right| \\ &\le \left(2 + \frac{2\chi}{\sqrt{N}}\right) Ae^{-\alpha t}. \end{aligned}$$

We take $C \ge \left(2 + \frac{2\chi}{\sqrt{N}}\right) A.$ □

This proposition is not enough, we want to show that $Y$ goes as $t$ goes to $\infty$ to a critical point of $\mathbb{E}_k$. The next proposition is a good step to achieve this goal. To this purpose we have to introduce a technical condition.

**Definition 7.5.12** (Multiple blolw-up condition.).

1. *Let $X$ be solution of* (7.1.1). *We say that $X$ satisfies the multiple blow-up condition if for any blow-up set $\mathcal{I}$ for $X$ there exist $t_0$ such that $X(t_0) \in D_{N,\chi}^{\varepsilon, \frac{C_N}{\varepsilon}}$, where the set inner set in $D_{N,\chi}^{\varepsilon, \frac{C_N}{\varepsilon}}$ is exactly $\mathcal{I}$.*
2. *Let $Y$ be a solution of* (7.5.22). *We say that $Y$ satisfies the multiple blow-up condition if there exist a $A > 0$ such that for any $i \neq j$, $|Y_i - Y_j| \ge \frac{1}{A}$.*

This condition means that if there are multiple blow-up for $X$, we can control each of them. This condition seems generically true. Moreover if $X$ solution of (7.1.1) satisfies the multiple blow-up condition then it rescaled solution $Y$ satisfies the multiple blow-up condition. Indeed for any indices in the outer set either they do not contribute to any blow-up and therefore the difference is far from 0 or they are involved in another blow-up but in this case the theorem 7.5.1 applied for this other blow-up set gives the needed bound.

**Proposition 7.5.13.** *Let $Y$ be a solution of the differential equation 7.5.24 satisfying the rescaled condition 7.5.7, and the multiple blow-up condition 7.5.12 then*

- $\dot{Y}(t) \to 0$ *as* $t \to \infty$.
- $\mathbb{E}_k(Y(t))$ *converges to a limit noted* $E_\infty$ *as* $t \to \infty$.
- $(\nabla \mathbb{E}_k)(Y(t)) \to 0$ *as* $t \to \infty$.

*Proof.* This proof is divided in two step we first prove an independent lemma which gives regularity estimates for such a solution and then deduce the proposition.

**Lemma 7.5.14.** *Let $Y$ be a solution of the differential equation 7.5.24 satisfying the rescaled condition 7.5.7, and the multiple blow-up condition 7.5.12 then there exists $C > O$ such that*

1. $\left|\dot{Y}_i\right|_{i\in\mathcal{I}} \leq C$,
2. $|Y_j|_{j\in\mathcal{O}} \leq Ce^{\alpha t}$,
3. $\left|\dot{Y}_j\right|_{j\in\mathcal{O}} \leq Ce^{\alpha t}$,
4. $\left|\ddot{Y}_i\right|_{i\in\mathcal{I}} \leq C$.

*Proof.* We start with the estimation of $(\dot{Y}_i)_{i\in[1,N]}$. We use the rescaled condition 7.5.7 and do not try to be sharp. If $i \in \mathcal{I}$

$$\begin{aligned}
\left|\dot{Y}_i\right| &= \left| -\frac{1}{Y_{i+1}-Y_i} + \frac{1}{Y_i - Y_{i-1}} + 2\chi h_N \sum_{k\neq i} \frac{1}{Y_k - Y_i} + \alpha Y_i \right| \\
&\leq 2A + 2\chi h_N \sum_{k\in\mathcal{I}\setminus\{i\}} \frac{1}{Y_k - Y_i} + 2\chi h_N \sum_{k\in\mathcal{O}} \frac{1}{Y_k - Y_i} + \alpha A \\
&\leq (2 + \alpha + 2\chi h_N (k-1))\, A + Ae^{-\alpha t}.
\end{aligned}$$

Taking $C \geq C_1 = (3 + \alpha + 2\chi h_N (k-1))\, A$ we have $\left|\dot{Y}_i\right|_{i\in\mathcal{I}} \leq C$.
For $j \in \mathcal{O}$

$$\begin{aligned}
\left|\dot{Y}_j - \alpha Y_j\right| &= \left| -\frac{1}{Y_{j+1}-Y_j} + \frac{1}{Y_j - Y_{j-1}} + 2\chi h_N \sum_{k\neq j} \frac{1}{Y_k - Y_j} \right| \\
&\leq 2A + 2\chi h_N \sum_{k\in\mathcal{O}\setminus\{j\}} \frac{1}{|Y_k - Y_j|} + 2\chi h_N \sum_{k\in\mathcal{I}} \frac{1}{|Y_k - Y_j|} \\
&\leq (2 + \alpha + 2\chi h_N (N-k+1))\, A + Ae^{-\alpha t}.
\end{aligned}$$

Taking $C$ large enough we get $|Y_j| \leq Ce^{\alpha t}$ and by triangular inequality $\left|\dot{Y}_j\right| \leq Ce^{\alpha t}$.
We now compute $\ddot{Y}_i$ for any $i \in \mathcal{I}$.

$$
\begin{aligned}
\left|\ddot{Y}_i\right| = \left|\frac{d}{dt}\dot{Y}_i\right| &= \left|\frac{d}{dt}\left(-\frac{1}{Y_{i+1}-Y_i} + \frac{1}{Y_i - Y_{i-1}} + 2\chi h_N \sum_{k\neq i}\frac{1}{Y_k - Y_i} + \alpha Y_i\right)\right| \\
&= \left|\frac{\dot{Y}_{i+1}-\dot{Y}_i}{(Y_{i+1}-Y_i)^2} - \frac{\dot{Y}_i - \dot{Y}_{i-1}}{(Y_i - Y_{i-1})^2} - 2\chi h_N \sum_{k\neq i}\frac{\dot{Y}_k - \dot{Y}_i}{(Y_k - Y_i)^2} + \alpha \dot{Y}_i\right| \\
&\leq C_1\left(\left(4+\alpha+2\chi h_N(k-1)\right)A^2 + A^2 e^{-2\alpha t}\right) + \\
&\left|\frac{\dot{Y}_{l+k}}{(Y_{l+k}-Y_{l+k-1})^2}\right| + \left|\frac{\dot{Y}_{l-1}}{(Y_l - Y_{l-1})^2}\right| + 2\chi h_N \sum_{k\neq i}\left|\frac{\dot{Y}_k}{(Y_k - Y_i)^2}\right| \\
&\leq C + \left(2 + \frac{2\chi}{\sqrt{N}}\right) Ce^{\alpha t}A^2 e^{-2\alpha t} \leq C,
\end{aligned}
$$

where $C$ is still a floating constant depending only on $N, \chi$ and $A$. It proves the lemma. □

For the next step we compute the derivative of $\mathbb{E}_k(Y)$ using 7.5.24, the idea is that $Y$ is almost a solution for the gradient flow equation related to the $^{th}$ energy. The key tools are again discrete integration by parts and symmetry.

$$
\begin{aligned}
\frac{d}{dt}\mathbb{E}_k(Y(t)) &= -\sum_{i\in\mathcal{I}\setminus\{l+k-1\}}\frac{\dot{Y}_{i+1}-\dot{Y}_i}{Y_{i+1}-Y_i} + \chi h_N \sum_{(i,j)\in\mathcal{I}\times\mathcal{I}\setminus\{i\}}\frac{\dot{Y}_j - \dot{Y}_i}{Y_j - Y_i} - \alpha\sum_{i\in\mathcal{I}} Y_i\dot{Y}_i \\
&= \sum_{i\in\mathcal{I}}\dot{Y}_i\left(\frac{1}{Y_{i+1}-Y_i} - \frac{1}{Y_i - Y_{i-1}}\right) - \frac{\dot{Y}_{l+k-1}}{Y_{l+k}-Y_{l+k-1}} + \frac{\dot{Y}_l}{Y_l - Y_{l-1}} \\
&\qquad - 2\chi h_N \sum_{(i,j)\in\mathcal{I}\times\mathcal{I}\setminus\{i\}}\frac{\dot{Y}_i}{Y_j - Y_i} - \alpha\sum_{i\in\mathcal{I}}Y_i\dot{Y}_i \\
&= \sum_{i\in\mathcal{I}}\dot{Y}_i\left(\frac{1}{Y_{i+1}-Y_i} - \frac{1}{Y_i - Y_{i-1}} - 2\chi h_N\sum_{j\in[1,N]\setminus\{i\}}\frac{\dot{Y}_i}{Y_j - Y_i} - \alpha Y_i\right) \\
&\qquad - \frac{\dot{Y}_{l+k-1}}{Y_{l+k}-Y_{l+k-1}} + \frac{\dot{Y}_l}{Y_l - Y_{l-1}} - 2\chi h_N\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{O}}\frac{\dot{Y}_i}{Y_j - Y_i} \qquad (7.5.26) \\
&\leq -||\dot{Y}||^2_{l^2(\mathcal{I})} + \left(2 + \frac{2\chi}{\sqrt{N}}\right)||\dot{Y}||_{l^2(\mathcal{I})}\sqrt{\mathrm{H}_{\mathcal{I}\mathcal{O},2}} \\
&\leq -||\dot{Y}||^2_{l^2(\mathcal{I})}\left(1 - \underbrace{\left(2 + \frac{2\chi}{\sqrt{N}}\right)A}_{A_1}e^{-\alpha t}\right) + A_1 e^{-\alpha t} \qquad (7.5.27)
\end{aligned}
$$

Here we apply a Cauchy-Schwarz inequality on each term of (7.5.26), we used the notation $||\dot{Y}||^2_{l^2(\mathcal{I})} = \sum_{i\in\mathcal{I}}\dot{Y}_i^2$ and the inequality for any $a > 0$, $a \leq 1 + a^2$. We deduce from (7.5.27) the integrability of $||Y||_{l^2(\mathcal{I})}$: we chose $t_0$ such that $1 - A_1 e^{-\alpha t} > \dfrac{1}{2}$ then

$$\begin{aligned}\int_{t_0}^{\infty} ||\dot{Y}||^2_{l^2(\mathcal{I})} &\leq \limsup_{t\to+\infty} \left( \int_{t_0}^{t} -2\frac{d}{dt}\mathbb{E}_k\left(Y(t)\right)\right) + A_1 e^{-\alpha t} dt \\ &\leq 2(M-m) + \frac{A_1}{\alpha},\end{aligned} \tag{7.5.28}$$

where $M$ and $m$ are respectively the upper and lower bound of $G(Y)$, depending only on $N, \chi, A$. Since by (7.5.14) $\left|\ddot{Y}_i\right|_{i\in\mathcal{I}} \leq C$, the inequality (7.5.28) implies that $||\dot{Y}|| \to 0$ as $t \to \infty$.
We now prove the convergence of $\mathbb{E}_k(Y)$. The equality 7.5.26 implies:

$$\begin{aligned}\left|\frac{d}{dt}\mathbb{E}_k\left(Y(t)\right)\right| &\leq ||\dot{Y}||^2_{l^2(\mathcal{I})} + \left(2 + \frac{2\chi}{\sqrt{N}}\right) ||\dot{Y}||_{l^2(\mathcal{I})} \sqrt{\mathrm{H}_{\mathcal{I}\mathcal{O},2}} \\ &\leq \frac{3}{2}||\dot{Y}||^2_{l^2(\mathcal{I})} + \frac{A_1^2}{2} e^{-2\alpha t}.\end{aligned}$$

Therefore $\dfrac{d}{dt}\mathbb{E}_k(Y)$ is integrable and there exist $E_\infty$ such that $\mathbb{E}_k(Y) \to E_\infty$ as $t \to +\infty$.
The last estimate is a straightforward consequence of $\dot{Y} = \left(\nabla \mathbb{E}_k\right)(Y)$. □

With this proposition we are very close to the conjecture 7.5.9, the best theorem we can give is the following.

**Theorem 7.5.15** (Sub Liouville.)**.** *Let $Y$ be a solution of the differential equation 7.5.24 satisfying the rescaled condition 7.5.7, and the multiple blow-up condition 7.5.12, we denote $Y_{\mathcal{I}} = (Y_i)_{i\in\mathcal{I}}$, then there exists $E_\infty \in \mathbb{R}$ such that for any time sequence we can extract a subsequence $(t_{n,s})_{n\in N}$ with*

1. *$Y_{\mathcal{I}}(t_{n,s}) \to Y_{\infty,s} \in \mathbb{R}^{|\mathcal{I}|}$ as $t \to +\infty$.*
2. *$\mathbb{E}_k\left(Y_{\infty,s}\right) = E_\infty$.*
3. *$\left(\nabla \mathbb{E}_k\right)\left(Y_{\infty,s}\right) = 0$*

*Proof.* This theorem is a direct consequence of proposition 7.5.13. □

### 7.5.b Conclusions and perspectives

The theorem 7.5.15 is really close to what we are looking for, we only miss the uniqueness of the limit $Y_{+\infty,s}$. To prove it we need for example to show that the critical points of $\mathbb{E}_k$ are isolated. We really think it is the case but the computation of the Hessian of $\mathbb{E}_k$ did not give anything yet. Remark that the theorem 7.5.15 gives the existence of at least one critical point, this critical point is probably a minimum on the sub variety define by all the $Y$ satisfying 7.5.7 but it is not a global minimum and has no reason to be a local minimum either. Anyway we have a strong rigidity of the blow since the outter set contributes only to fix the blow-up time $T$ and the blow-up point $\overline{X}$, not the limit profile.
The theorem 7.5.11 on his side is very powerful, for example we can rewrite the proof of proposition 7.5.13 with it. It also allows to transfer any properties of the $^{th}$ energy functional to a

rescaled solution. For example suppose that $\mathbb{E}_k$ has a unique minimizer $Z^\infty$ and is $\beta$ BMX handlebar, that is to say there exists $\beta > 0$ such that for any $Z \in \mathbb{R}^k$:

$$- \langle Z - Z^\infty, \nabla \mathbb{E}_k(Z) \rangle \leq -\beta \left| Z - \overline{Z} \right|^2 .$$

Then for $(Y)$ solution of the differential equation 7.5.24 satisfying the rescaled condition 7.5.7 we have that $(Y_i)_{i \in \mathcal{I}}$ converges exponentially to $Z^\infty$.

*Proof.* For simplicity, we re-index $Z^\infty$ by $(Z_i^\infty)_{i \in \mathcal{I}}$ and $\nabla \mathbb{E}_k$ by $(\nabla_i \mathbb{E}_k)_{i \in \mathcal{I}}$.

$$\begin{aligned}
\frac{d}{dt} \sum_{i \in \mathcal{I}} |Y_i - Z_i^\infty|^2 = -2 \sum_{i \in \mathcal{I}} &- (Y_i - Z_i^\infty) \nabla_i \mathbb{E}_r(Y) \\
&- 2 \sum_{i \in \mathcal{I}} (Y_i - Z_i^\infty) \nabla_i \mathbb{E}_k(Y) - 2 \sum_{i \in \mathcal{I}} (Y_i - Z_i^\infty) \left( \nabla \mathbb{E}_r(Y)_i - \nabla \mathbb{E}_k(Y)_i \right) \\
&\leq -\beta \left| Y - \overline{Z} \right|^2 + C e^{-\alpha t}.
\end{aligned}$$

We conclude with the Gronwall inequality. □

This computation is not so silly as we do think that the energy $\mathbb{E}_k$ is almost BMX handlebar. We proved it for the subcritical case in chapter 6.

There are various way to continue this work, an idea is to extend our scheme beyond the blow-up time, in this case we consider particles carrying different mass, we need to write down the equation of the mass evolution. It would be interesting to compare this approach with the measure solutions of Dolbeault and Schmeiser [40]. Another direction is to change the log interaction kernel by others for example $x \mapsto |x|^{-\alpha}$, many things work similarly in these cases. The question of the convergence of our particles scheme when the number of particles goes to $+\infty$ is also interesting. The final result we are looking for is the convergence to a unique auto similar profile for any solution of the log interaction equation in dimension 1 in the supercritical case.

# Appendix B

# L'équation de Keller-Segel

This appendix is my M2 report. It deals with the Keller-Segel equation in dimension 2, in particular the existence of the threshold number $8\pi$ and the existence, uniqueness and regularity of the solution in the subcritical case. We follow the free energy method. We find here for example a positivity lemma for the Keller-Segel equation.

## Sommaire

## B.1 Introduction

Ce stage de M2, réalisé à l'Ecole Normale de Lyon sous la direction de Cédric Villani, s'est basé sur l'article de Jean Dolbeault et Benoît Perthame publié en 2004, *Optimal critical mass in the two dimensional Keller-Segel model in* $\mathbb{R}^2$ [39].

Il a consisté en l'étude d'un problème d'attraction-diffusion de cellules, la chemotaxis. On considère ainsi l'évolution dans le plan de la densité $n$ de cellules émettant une substance $c$ et étant attirés par cette dernière.

Le modèle simplifié dit de Keller-Segel de ce problème s'écrit comme suit :

$$\begin{array}{ll} \frac{\partial n}{\partial t}(x,t) = \Delta n(x,t) - \chi \nabla.\big(n(x,t)\nabla c(x,t)\big) & \forall x \in \mathbb{R}^2, \forall t > 0 \text{ et avec } \chi > 0 \\ -\Delta c(x,t) = n(x,t) & \forall x \in \mathbb{R}^2, \forall t > 0, \\ n(x,t=0) = n_0(x) & \forall x \in \mathbb{R}^2, \text{ avec } n_0 \in L^1(\mathbb{R}^2) \geq 0 \end{array} \tag{B.1.1}$$

où $\chi$ est un coefficient de sensibilité à la substance : plus il est grand, plus les cellules sont attirées par la substance.

Comme on connait le noyau du laplacien, on peut écrire que :

$$\forall x \in \mathbb{R}^2, t > 0, c(x,t) = -\frac{1}{2\pi}(\log(|.|) * n(.,t))(x)$$

Ici $*$ est l'opérateur de convolution en espace.

La masse des cellules est l'intégrale de leur densité sur l'espace total. Nous verrons que cette masse reste constante dans le temps. L'apport de l'article étudié est de montrer que $\frac{8\pi}{\chi}$ est une masse critique, c'est-à-dire qu'au dessus de cette masse les cellules se concentrent en un ou plusieurs points.

## B.2 Étude du système ; estimations a priori

### B.2.a Définition de solutions

#### Solutions au sens fort

Les solutions fortes sont les solutions $n$ telles que :

$$\begin{array}{l} \forall x \in \mathbb{R}^2, n(x,.) \in C^1([0,T]) \\ \forall t \in [0,T], n(.,t) \in C^2(\mathbb{R}^2) \\ n \in L^1(\mathbb{R}^2 \times [0,T]) \end{array} \tag{B.2.1}$$

**Solutions au sens faible**

Pour définir les solutions au sens faible, on teste l'équation (1) avec $\psi \in C_c^\infty(\mathbb{R}^2)$.

$$\begin{aligned}
&\forall t > 0, \\
&\int_{\mathbb{R}^2} \frac{\partial n}{\partial t}(x,t)\psi(x)dx = \int_{\mathbb{R}^2} \Delta n(x,t)\psi(x)dx - \chi \int_{\mathbb{R}^2} \nabla.\big(n(x,t)\nabla c(x,t)\big)\psi(x)dx \\
&\Rightarrow \frac{d}{dt}\int_{\mathbb{R}^2} n(x,t)\psi(x)dx = \int_{\mathbb{R}^2} n(x,t)\Delta\psi(x)dx + \chi \int_{\mathbb{R}^2} n(x,t)\nabla c(x,t)\nabla\psi(x)dx
\end{aligned} \tag{B.2.2}$$

Or :

$$\begin{aligned}
&\forall x \in \mathbb{R}^2, t > 0, \\
&c(x,t) = -\frac{1}{2\pi}(\log(|.|) * n(.,t))(x) = -\frac{1}{2\pi}\int_{\mathbb{R}^2} \log(|x-y|)n(y,t)dy \\
&\Rightarrow \nabla c(x,t) = -\frac{1}{2\pi}\int_{\mathbb{R}^2} \nabla \log(|x-y|)n(y,t)dy \\
&\Rightarrow \nabla c(x,t) = -\frac{1}{2\pi}\int_{\mathbb{R}^2} \frac{x-y}{|x-y|^2}n(y,t)dy
\end{aligned} \tag{B.2.3}$$

Donc $\forall t > 0$ :

$$\frac{d}{dt}\int_{\mathbb{R}^2} n(x,t)\psi(x)dx = \int_{\mathbb{R}^2} n(x,t)\Delta\psi(x)dx + \frac{\chi}{2\pi}\int\int_{\mathbb{R}^2\times\mathbb{R}^2} \frac{x-y}{|x-y|^2}n(y,t)n(x,t)\nabla\psi(x)dydx \tag{B.2.4}$$

Par symétrie on peut écrire $\forall t > 0$ :

$$\begin{aligned}
&\frac{d}{dt}\int \psi(x)n(x,t)dx \\
&\qquad = \int \Delta\psi(x)n(x,t)dx + \frac{\chi}{4\pi}\int\int \frac{x-y}{|x-y|^2}(\nabla\psi(x) - \nabla\psi(y))n(x,t)n(y,t)dydx
\end{aligned} \tag{B.2.5}$$

On contrôle ainsi la singularité en $\frac{1}{x-y}$ car $\frac{1}{|x-y|}(\nabla\psi(x) - \nabla\psi(y))$ est bornée.
On définit donc une solution faible de la façon suivante :

**Définition B.2.1.** *$n$ est solution faible de* (1) *si $n$ appartient à $L^\infty(\mathbb{R}^+; L^1(\mathbb{R}^2))$ et si de plus :* $\forall \psi \in C_c^\infty(\mathbb{R}^2)$,

$$\begin{aligned}
&\frac{d}{dt}\int \psi(x)n(x,t)dx \\
&\qquad = \int \Delta\psi(x)n(x,t)dx + \frac{\chi}{4\pi}\int\int \frac{x-y}{|x-y|^2}(\nabla\psi(x) - \nabla\psi(y))n(x,t)n(y,t)dydx
\end{aligned} \tag{B.2.6}$$

Comme souvent, une solution forte est aussi une solution faible.

Une autre façon de voir des solutions au sens des distributions est de remarquer que :

$$\Delta n - \chi\nabla.(n\nabla c) = \nabla.[n(\nabla \log n - \chi\nabla c)]$$

Si on appele $n(\nabla \log n - \chi\nabla c)$ le flux. On voit que la solution est bien définie dès que le flux est dans $L^1(\mathbb{R}^+_{loc} \times \mathbb{R}^2)$. C'est ce que l'on va essayer de prouver en se basant sur l'inégalité suivante :

$$\int\int_{[0,T]\times\mathbb{R}^2} n|n(\nabla \log n - \chi\nabla c)|dxdt \\ \leq (\int\int_{[0,T]\times\mathbb{R}^2} ndxdt)^{1/2}(\int\int_{[0,T]\times\mathbb{R}^2} n|n(\nabla \log n - \chi\nabla c)|^2 dxdt)^{1/2} \quad \text{(B.2.7)}$$

Et en trouvant les estimations qui conviennent.

### B.2.b Conservation de la masse

Un premier théorème utile :

**Théorème B.1.** *Si n est solution faible de* (1), *alors la masse est conservée dans le temps, c'est-à-dire :*
$\forall t > 0, \int n(x,t)dx = \int n_0(x)dx.$

Démonstration : Soit $\psi \in C_c^\infty(\mathbb{R}^+)$ tel que $\psi(r) = 1$ si $r < 1/2$ et $\psi(r) = 0$ si $r > 1$. On note alors, pour tout $R > 0$ et pour tout x appartenant à $\mathbb{R}^2$, $\psi_R(x) = \psi(\frac{|x|^2}{R^2})$.
On a alors : $\nabla\psi_R(x) = \frac{2x}{R^2}\psi'(\frac{|x|^2}{R^2})$
et donc : $\Delta\psi_R(x) = \frac{4}{R^2}\psi'(\frac{|x|^2}{R^2}) + \frac{4}{R^2}|x|^2\psi''(\frac{|x|^2}{R^2})$
Soit, puisque $\psi$ et donc ses dérivées sont nulles pour $\frac{|x|^2}{R^2} \geq 1$,

$$|\int_{\mathbb{R}^2} \Delta\psi_R(x)n(x,t)dx| \leq \frac{C}{R^2}\int_{\mathbb{R}^2} |n(x,t)|dx \underset{R\to\infty}{\longrightarrow} 0 \quad \text{(B.2.8)}$$

Et comme $2x\psi'(\frac{|x|^2}{R^2})$ est différentiable (et nulle loin de 0), on a :

$$\int\int_{\mathbb{R}^2\times\mathbb{R}^2} \frac{x-y}{|x-y|^2}\nabla(\psi_R(x) - \psi_R(y))n(x,t)n(y,t)dydx \\ \leq \frac{C}{R^2}\int\int_{\mathbb{R}^2\times\mathbb{R}^2} |n(x,t)n(y,t)|dxdy \underset{R\to\infty}{\longrightarrow} 0 \quad \text{(B.2.9)}$$

De plus, comme $n \in L^1(\mathbb{R}^2)$,

$$\frac{d}{dt}\int \psi_R(x)n(x,t)dx \underset{R\to\infty}{\longrightarrow} \frac{d}{dt}\int n(x,t)dx \quad \text{(B.2.10)}$$

et donc

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

On a donc bien montré que la masse est conservée.

### B.2.c Principe de positivité

Étant donné que l'on veut représenter une densité de particules, il semble important d'établir un principe de positivité :

**Théorème B.2.** *On considère $n$ une solution faible de* (1)*, avec pour tout $x$ appartenant à $\mathbb{R}^2$, $n_0(x) > 0$. Alors $n$ est positive.*

DÉMONSTRATION :
Nous allons démontrer le principe de positivité pour une équation de forme plus générale :

$$\begin{aligned} &\partial_t n - \Delta n + \operatorname{div}(n.V) = f \\ &n(x,t=0) = n_0(x) \qquad\qquad \forall x \in \mathbb{R}^2, \text{ avec } n_0 \in L^1(\mathbb{R}^2) \leq 0 \end{aligned} \tag{B.2.11}$$

avec $f \in L^\infty(\mathbb{R}^2 \times (0,T)) \leq 0$ (on va donc montrer que la solution reste négative), $V \in L^b(\mathbb{R}^2)$ pour un $b > 2$, et $n \in L^1(\mathbb{R}^2 \times (0,T))$ avec $T \geq 0$.

Nous démontrerons tout d'abord ce principe pour le cas stationnaire afin de simplifier l'écriture, puis nous rajouterons dans un deuxième temps le terme d'évolution.
Ce dernier résultat, plus général, sera utile par la suite pour résoudre le probléme régularisé linéaire.

#### Principe de positivité pour le problème stationnaire

On veut établir ici le principe de positivité pour le cas stationnaire. On s'intéresse donc à l'équation :

$$-\Delta n + \nabla(n.V) = f, f \in L^\infty(\mathbb{R}^2) \leq 0 \tag{B.2.12}$$

Pour $\eta$ et $\varepsilon$ strictement positifs, on définit $T_{\varepsilon,\eta}$ de $\mathbb{R}$ dans $\mathbb{R}^+$ par $T_{\varepsilon,\eta}(z) = 0$ si $z \leq \eta$, $T_{\varepsilon,\eta}(z) = \varepsilon$ si $z \geq \eta + \varepsilon$ et $T_{\varepsilon,\eta}(z) = z - \eta$ si $z \in [\eta, \eta + \varepsilon]$.
On teste alors l'équation contre $T_{\varepsilon,\eta}(n)$.
On obtient par IPP :

$$\int_\Omega \nabla(n).\nabla T_{\varepsilon,\eta}(n) - \int_\Omega nV.\nabla T_{\varepsilon,\eta}(n) = \int_\Omega f T_{\varepsilon,\eta}(n) \tag{B.2.13}$$

Avec ici $\Omega = \mathbb{R}^2$. Comme $f \leq 0$ et $T_{\varepsilon,\eta} \geq 0$, le second membre est négatif.
De plus, $\nabla T_{\varepsilon,\eta}$ est nulle hors de $[\eta, \eta + \varepsilon]$, et est linéaire de pente 1 sur cet intervalle, donc on peut écrire :

$$\int_\Omega \nabla(n).\nabla T_{\varepsilon,\eta}(n) = \int_{\Omega_{\varepsilon,\eta}} |\nabla T_{\varepsilon,\eta}(n)|^2, \text{ avec } \Omega_{\varepsilon,\eta} = \{n/\eta \leq n \leq \eta + \varepsilon)\} \tag{B.2.14}$$

On veut maintenant majorer $\int_\Omega nV.\nabla T_{\varepsilon,\eta}(n)$, c'est-à-dire $\int_{\Omega\varepsilon,\eta} nV.\nabla T_{\varepsilon,\eta}(n)$.
Pour $q$ et $r$ tels que $1/q + 1/r + 1/2 = 1$, on a par Holder :

$$\int_{\Omega\varepsilon,\eta} nV.\nabla T_{\varepsilon,\eta}(n) \leq ||\nabla T_{\varepsilon,\eta}(n)||_{L^2(\Omega_{\varepsilon,\eta})}||V||_{L^r(\Omega_{\varepsilon,\eta})}||n||_{L^q(\Omega_{\varepsilon,\eta})} \tag{B.2.15}$$

On remarque que $q > 2$, donc q est le $p^*$ d'un $p < 2$.
On a donc la suite de majoration suivante :
Par les inégalités de Sobolev on a :

$$||n||_{L^q(\Omega_{\varepsilon,\eta})} \leq ||\nabla(n)||_{L^p(\Omega_{\varepsilon,\eta})} \tag{B.2.16}$$

et on a déjà vu que :

$$||\nabla(n)||_{L^p(\Omega_{\varepsilon,\eta})} = ||\nabla T_{\varepsilon,\eta}(n)||_{L^p(\Omega_{\varepsilon,\eta})} \tag{B.2.17}$$

et comme la mesure des $\{x \in \mathbb{R}^2/n(x) \geq \eta\}$ est finie ($n \in L^1(\mathbb{R}^2)$), on a de plus :

$$||\nabla T_{\varepsilon,\eta}(n)||_{L^p(\Omega_{\varepsilon,\eta})} \leq C_{\eta,p}||\nabla T_{\varepsilon,\eta}(n)||_{L^2(\Omega_{\varepsilon,\eta})} \tag{B.2.18}$$

Au final on obtient donc :

$$||\nabla T_{\varepsilon,\eta}(n)||^2_{L^2(\Omega_{\varepsilon,\eta})} - C_{\eta,p}||V||_{L^r(\Omega_{\varepsilon,\eta})}||\nabla T_{\varepsilon,\eta}(n)||^2_{L^2(\Omega_{\varepsilon,\eta})} \leq 0 \tag{B.2.19}$$

Comme pour tout $\eta$ fixé on peut choisir un $\varepsilon$ tel que $C_{\eta,p}||V||_{L^r(\Omega_{\varepsilon,\eta})} \leq 1/2$ (par intégrabilité), on obtient que $\Omega_{\varepsilon,\eta}$ est vide, ce qui force $n \leq \eta$ pour tout $\eta$, et donc $n \leq 0$.
On a donc montré que $f \leq 0$ entraîne que $n \leq 0$. On a de même le cas positif.

**Principe de positivité avec terme d'évolution**

$$\begin{aligned} &\partial_t n - \Delta n + \operatorname{div}(n.V) = f \\ &n(x, t = 0) = n_0(x) \qquad\qquad n_0 \in L^1(\mathbb{R}^2) \leq 0, \forall x \in \mathbb{R}^2 \end{aligned} \tag{B.2.20}$$

avec $f \in L^\infty(\mathbb{R}^2 \times (0, T)) \leq 0$, $V \in L^b(\mathbb{R}^2)$ pour un $b > 2$, et $n \in L^1(\mathbb{R}^2 \times (0, T))$ avec $T \geq 0$.
Pour rajouter le terme d'évolution, on remarque que la majoration des $\{x \in \mathbb{R}^2/n(x, t) \geq \eta\}$ est indépendante du temps (conservation de la masse).
L'autre terme qui s'ajoute est :

$$\begin{aligned} &\int_0^T (\int_{\mathbb{R}^2} (\partial_t n) T_{\varepsilon,\eta}(n)) \\ &= \int_0^T (\int_{\mathbb{R}^2} (\partial_t (T_{\varepsilon,\eta}(n)) T_{\varepsilon,\eta}(n)) \\ &= \int_0^T \frac{d}{dt} (\int_{\mathbb{R}^2} (T_{\varepsilon,\eta}(n))^2/2) \\ &= \int_{\mathbb{R}^2} (T_{\varepsilon,\eta}(n(T, .))^2/2) - \int_{\mathbb{R}^2} (T_{\varepsilon,\eta}(n_0(.))^2/2) \end{aligned}$$

$n_0$ étant négative, $T_{\varepsilon,\eta}(n_0(.))^2/2 = 0$, donc la quantité rajoutée est bien positive, ce qui permet de faire le même raisonnement.

On a donc également le principe de positivité dans le cas d'évolution, et avec $V = \nabla c$ pour le cas particulier de notre problème.

### B.2.d Un résultat d'explosion

On note $M = \int_{\mathbb{R}^2} n(x,t)dx$ (conservation de la masse), et de plus on suppose que :

$$\begin{array}{l} n_0 \in L^1(\mathbb{R}^2, (1+|x|^2)dx) \\ n_0 \log(n_0) \in L^1(\mathbb{R}^2, dx) \end{array} \tag{B.2.21}$$

On s'intéresse au moment d'ordre 2 de $n$. On a alors le lemme suivant :

**Lemme B.2.2.**
*On considère une solution faible positive $n$ de* (1) *sur un intervalle* $[0,T]$ *dont le moment d'ordre 2 est borné, qui vérifie* (1.19) *et est telle que* $(x,t) \to \int_{\mathbb{R}^2} \frac{1+|x|}{|x-y|} n(y,t)dy$ *est bornée dans* $L^\infty(\mathbb{R}^2 \times [0,T])$.
*On a alors :*

$$\frac{d}{dt}\int_{\mathbb{R}^2} |x|^2 n(x,t)dx = 4M(1 - \frac{\chi M}{8\pi}).$$

Démonstration :
On considère une suite $\varphi_\varepsilon$ de fonctions $C_c^\infty(\mathbb{R}^2)$ qui converge vers $|x|^2$ quand $\varepsilon$ tend vers 0.
On a :

$$\frac{d}{dt}\int_{\mathbb{R}^2} n(x,t)\varphi_\varepsilon dx = \int_{\mathbb{R}^2} \varphi_\varepsilon(\Delta n - \chi\nabla.(n\nabla c))dx$$

$$= \int_{\mathbb{R}^2} \Delta\varphi_\varepsilon n - \frac{\chi}{2\pi}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2} n(x,t)n(y,t)\frac{(\nabla\varphi_\varepsilon(x) - \nabla\varphi_\varepsilon(y)).(x-y)}{|x-y|^2}dxdy$$

En passant à la limite par convergence monotone, car $\Delta\varphi_\varepsilon$ et $\frac{(\nabla\varphi_\varepsilon(x) - \nabla\varphi_\varepsilon(y)).(x-y)}{|x-y|^2}$ sont bornées et $n \in L^1(\mathbb{R}^2)$, on a :

$$\begin{aligned} &\frac{d}{dt}\int_{\mathbb{R}^2} n(x,t)\varphi_\varepsilon dx \\ &= 4M - \frac{\chi}{2\pi}M^2 \\ &= 4M(1 - \frac{\chi M}{8\pi}). \end{aligned}$$

On voit ici la compétition entre le terme de diffusion qui régularise la solution et le terme d'attraction qui tend à la faire exploser.
On a donc :
- Si $\chi M > 8\pi$, le second moment est strictement décroissant, donc la solution explose en temps fini (le second moment ne peut pas être negatif) .
- Si $\chi M < 8\pi$ et que la solution assez régulière, alors le second moment est borné dans $L^\infty([0,T]; L^1(\mathbb{R}^2)))$.
Il ne reste donc plus qu'à démontrer l'existence d'une solution dans le cas où $\chi M < 8\pi$.

### B.2.e Le résultat classique d'existence

L'idée est de calculer l'entropie de $\int_{\mathbb{R}^2} n\log(n)$. On s'intéresse donc à :

$$\frac{d}{dt}\int_{\mathbb{R}^2} n\log(n)dx = -4\int_{\mathbb{R}^2}|\nabla\sqrt{n}|^2dx + \chi\int_{\mathbb{R}^2}\nabla n.\nabla c dx$$

$$= -4\int_{\mathbb{R}^2}|\nabla\sqrt{n}|^2dx + \chi\int_{\mathbb{R}^2} n^2dx.$$

Grâce à l'inégalité de Gagliardo-Nirenberg-Sobolev avec $u=\sqrt{n}$ et $p=4$, on obtient :

$$\int_{\mathbb{R}^2} n^2dx \leq (C^{(4)}_{GNS})^{-2}M\int_{\mathbb{R}^2}|\nabla\sqrt{n}|^2dx \tag{B.2.22}$$

On a ainsi la décroissance de l'entropie lorsque $\chi M \leq 4(C^{(4)}_{GNS})^{-2}$. Numériquement, on sait que cette valeur est inférieure à $8\pi$, mais cela ne couvre pas entièrement le second cas.

Rappel de l'inégalité de Gagliardo-Nirenberg-Sobolev :

Pour la norme $L^2$ :

$$||u||_p^2 \leq C^{(p)}_{GNS}||\nabla u||^{\frac{4}{p}}_{L^2(\mathbb{R}^2)}||u||^{2-\frac{4}{p}}_{L^2(\mathbb{R}^2)}\ \forall u\in H^1(\mathbb{R}^2),\ \forall p\in[2,\infty) \tag{B.2.23}$$

### B.2.f Estimation des normes $L^p$

Beckner et Lauckhauss ont montré que les normes $L^p$ se transmettent, c'est-à-dire que si $n_0 \in L^p(\mathbb{R}^2)$, alors pour tout $t$, $n(t)\in L^p(\mathbb{R}^2)$. Ce résultat est basé sur l'estimation suivante :

On note $pG_p(t) = \int n(x,t)^p dx$. Vu l'équation, et par IPP successives :

$$G'_p(t) = \int n(x,t)^{(p-1)}\frac{\partial n(x,t)}{\partial t}dx = \int n(x,t)^{(p-1)}(\Delta n - \chi\nabla.(n\nabla c))$$

$$= -(p-1)\int|\nabla n|^2n^{p-2}dx + (p-1)\chi\int(\nabla n)n^{p-1}(\nabla c)dx$$

$$= -\frac{4(p-1)}{p^2}\int|\nabla n^{\frac{p}{2}}|^2dx + \frac{\chi(p-1)^2}{p}\int(\nabla n^p)(\nabla c)dx$$

$$= -\frac{4(p-1)}{p^2}\int|\nabla n^{\frac{p}{2}}|^2dx - \frac{\chi(p-1)^2}{p}\int n^p(\Delta c)dx$$

soit, comme $-\Delta c = n$ ,

$$-\frac{4(p-1)}{p^2}\int|\nabla n^{\frac{p}{2}}|^2dx + \frac{\chi(p-1)^2}{p}\int n^{p+1}dx$$

Or thanks to the Galiardo-Nirenberg-Sobolev inequation :

$$\int n^{p+1} \leq C(d,p)\int|\nabla n^{\frac{p}{2}}|dx||n||_{L^1}$$

Ce qui donne au final :

$$G_p'(t) \leq (\int |\nabla n^{\frac{p}{2}}|^2 dx)(-\frac{4(p-1)}{p^2} + C(d,p)\frac{\chi(p-1)^2}{p}||n||_{L^1})$$

On voit encore ici la compétition entre le terme de diffusion qui régularise la solution et le terme d'attraction qui tend à la faire exploser.

Cette estimation permet effectivement de transmettre les normes $L^p$, mais nécessite que la norme $L^1$, c'est-à-dire $M$, soit plus petite qu'une constante dépendant de $C(d,p) = C_p$. Ce n'est pas une bonne condition, on va donc s'en débarasser.

Pour cela, on va "tronquer" $n$ et refaire le calcul. On s'intéresse donc à $G_p(t) = \int (n(x,t) - K)_+^p dx$.

On calcule :

$$G_p'(t) = p(\int (n(x,t) - K)_+^{(p-1)} \frac{\partial n(x,t)}{dt} dx) = p(\int (n(x,t) - K)_+^{(p-1)} (\Delta n - \chi\nabla.(n\nabla c)))$$

$$= p(-(p-1)\int |\nabla n|^2 (n(x,t) - K)_+^{p-2} dx - \chi \int (n(x,t) - K)_+^{p-1} \nabla(n(\nabla c))dx)$$

$$= p(-\frac{4(p-1)}{p^2}\int |\nabla(n(x,t) - K)_+^{\frac{p}{2}}|^2 dx - \chi \int (n(x,t) - K)_+^{p-1} \nabla(n(\nabla c))dx)$$

On ne s'intéresse plus qu'au second terme :

$$-p\chi \int (n(x,t) - K)_+^{p-1} \nabla(n(\nabla c))dx = -p\chi \int (n(x,t) - K)_+^{p-1} (\nabla(n-K).\nabla c + n\Delta c)dx$$

$$= -\chi \int (n(x,t) - K)_+^{p-1} ((n-K)_+(-\Delta c) - pn(-\Delta c))dx$$

$$= (p-1)\chi \int (n(x,t) - K)_+^{p+1} dx + (2p-1)\chi K \int (n(x,t) - K)_+^{p} dx + p\chi K^2 \int (n(x,t) - K)_+^{p-1} dx$$

Pour le terme à la puissance $p-1$, on fait comme suit :

$$\int (n(x,t) - K)_+^{p-1} dx = \int_{K<n\leq K+1} (n-K)_+^{p-1} dx + \int_{n>K+1} (n-K)_+^{p-1} dx$$

or :

$$\int_{K<n\leq K+1} (n-K)_+^{p-1} dx \leq \int_{K<n\leq K+1} 1 dx \leq \frac{1}{K}\int_{K<n\leq K+1} n dx \leq \frac{M}{K}$$

et pour le second on utilise le fait que la différence est plus grande que 1, et donc :

$$\int_{n>K+1} (n-K)_+^{p-1} dx \leq \int (n-K)_+^p dx.$$

On est donc bien parti pour appliquer Gronwall. Il reste à faire disparaître (c'est-à-dire à rendre

négatif) le dernier terme : $-\frac{4(p-1)}{p}\int|\nabla(n(x,t)-K)_+^{\frac{p}{2}}|^2dx+(p-1)\chi\int(n(x,t)-K)_+^{p+1}dx$.
Pour cela on utilise G.N.S avec $(n-K)_+$. Le seul changement est que la norme $L^1$ de $n$ est changée par celle de $n-K$, que l'on peut rendre aussi petite que l'on veut, quitte à choisir $K$ grand.

Au final, on obtient une inégalité de la forme :

$$\frac{d}{dt}\int(n-K)_+^p dx\leq A\int(n-K)_+^p dx+B$$

On a donc via Gronwall une borne pour tout temps fini.

On sent bien que l'on contrôle les normes $L^p$ (car $n$ est positive, régulière et de masse finie), mais précisons tout de même le raisonnement :

On partage l'intégrale en deux parties :

$$\int n^p dx\leq\int_{n\leq K}n^p dx+\int_{n\geq K}n^p dx$$

La première partie se traite facilement avec la masse de $n$ :

$$\int_{n\leq K}n^p dx\leq K^{p-1}M$$

Pour la seconde, on remarque que pour $1<\lambda\leq x$ :

$$x^p\leq(\frac{\lambda}{\lambda-1})^{p-1}(x-1)^p$$

Puis on écrit :

$$\int_{K\leq n}n^p dx\leq\int_{K\leq n\leq\lambda}n^p dx+\int_{n>\lambda K}n^p$$
$$\leq(\lambda K)^{p-1}M+(\frac{\lambda}{\lambda-1})^{p-1}K^p\int_{K\leq n\leq\lambda}(\frac{n}{K}-1)_+^p dx$$
$$\leq(\lambda K)^{p-1}M+(\frac{\lambda}{\lambda-1})^{p-1}\int(n-K)_+^p dx$$

C'est donc gagné, on a propagé, du moins a priori, les normes $L^p$.

### B.2.g Stratégie

Ici une petite explication de la stratégie qui va être utilisée semble la bienvenue.

L'entropie n'étant pas assez précise, on définira une notion plus proche du système : l'énergie libre. Ensuite, nous allons donner quelques estimations a priori pour des solutions classiques. La suite de la démonstration se fera par une régularisation du système via la régularisation du noyau de convolution, la résolution par un point fixe de Banach de ces problèmes et enfin un argument de compacité pour passer à la limite.

### B.2.h L'énergie libre : pour en savoir un peu plus

L'entropie, quantité utilisée précédemment, n'est pas assez précise. On définit donc une quantité plus proche du problème, l'énergie libre, par :

$$F(t) = \int_{\mathbb{R}^2} n\log(n)dx - \frac{\chi}{2}\int_{\mathbb{R}^2} ncdx$$

On a alors le lemme suivant :

**Lemme B.2.3.** *On considère une solution de* (1) *n positive, continue, à valeur* $L^1(\mathbb{R}^2)$*. Si* $n(1+|x|^2)$ *et* $n\log(n)$ *sont bornées dans* $L^\infty_{loc}(\mathbb{R}^+, L^1(\mathbb{R}^2))$*, et si* $\nabla\sqrt{n}$ *est bornée dans* $L^1_{loc}(\mathbb{R}^+, L^2(\mathbb{R}^2))$ *et* $\nabla c$ *dans* $L^\infty_{loc}(\mathbb{R}^+ \times \mathbb{R})$*, alors*

$$\frac{d}{dt}F(t) = -\int_{\mathbb{R}^2} n|\nabla(\log(n)) - \chi\nabla c|^2 dx.$$

L'énergie libre a donc vocation à être décroissante.
Démonstration :
Pour le premier terme :

$$\frac{d}{dt}\int_{\mathbb{R}^2} n|\nabla(\log(n))dx = \int_{\mathbb{R}^2}\frac{d}{dt}n(\log(n)+1)dx$$

$$= \int(\log(n)+1)(\Delta n - \chi\nabla.(n\nabla c))dx$$

$$= -\int \nabla(\log(n))(\nabla n - \chi(n\nabla c))dx$$

$$= -\int n\nabla(\log(n))(\nabla\log(n) - \chi(\nabla c))dx.$$

Pour le second :

$$\frac{d}{dt}\frac{\chi}{2}\int ncdx = \frac{\chi}{2}(\int n\frac{d}{dt}c + c\frac{d}{dt}n)dx$$

Mais comme $c(x,t) = -\frac{1}{2\pi}(\log(|.|)) * n(.,t)$, on a :

$$\frac{d}{dt}c(x,t) = -\frac{1}{2\pi}(\log(|.|)) * \frac{d}{dt}n(.,t)$$

soit :

$$\frac{d}{dt}\frac{\chi}{2}\int ncdx = \frac{\chi}{2}(\int\int -\frac{1}{2\pi}n(x,t)\log(|x-y|)\frac{d}{dt}n(y,t)dxdy + \int c(x,t)\frac{d}{dt}n(x,t)dx)$$

Étant donné la symétrie du noyau du laplacien, on peut écrire par Fubini :

$$\frac{d}{dt}\frac{\chi}{2}\int ncdx = \frac{\chi}{2}(\int(\frac{d}{dt}n(y,t))(-\frac{1}{2\pi}(\log(|.|)) * n(.,t))dy + \int c(x,t)\frac{d}{dt}n(x,t)dx)$$

donc

$$= \chi(\int c(y,t)\frac{d}{dt}n(y,t)dy)$$

soit

$$= (\int \chi c(\Delta n - \chi \nabla.(n \nabla c)) dx)$$

$$= -(\int \chi \nabla c(\nabla n - \chi n \nabla c) dx)$$

et donc

$$= -\int \chi \nabla c(\nabla n - \chi n \nabla c) dx$$

$$= -\int n \chi \nabla c(\nabla \log(n) - \chi \nabla c) dx$$

On somme alors les deux et on obtient :

$$\frac{d}{dt} F(t) = -\int n \nabla(\log(n))(\nabla \log(n) - \chi(\nabla c)) dx + \int n \chi \nabla c(\nabla \log(n) - \chi \nabla c) dx$$

$$\frac{d}{dt} F(t) = -\int n(\nabla \log(n) - \chi(\nabla c))(\nabla(\log(n)) - \chi \nabla c) dx$$

et donc le résultat :

$$\frac{d}{dt} F(t) = -\int n|(\nabla \log(n) - \chi(\nabla c))|^2 dx.$$

$F$ est donc décroissante.

On va maintenant chercher à en déduire une borne supérieure et inférieure sur l'entropie ce qui nous permettra d'appliquer la démonstration classique (on obtient l'équiintégrabilité avec le second moment).

On a donc $F(t) \leq F(0)$, soit :

$$\int n \log(n) dx \leq F(0) + \int \frac{\chi}{2} n c dx$$

que l'on écrit :

$$\int n \log(n) dx \leq F(0) - \frac{\chi}{4\pi} \int \int n(x,t) n(y,t) \log|x-y| dx dy$$

On utilise alors l'inégalité de Hardy-Littlewood-Sobolev logarithmique (voir plus loin) qui dit ici :

$$M(1 + \log(\pi) + \log(M)) \leq \int n \log(n) dx + \frac{2}{M} \int \int n(x,t) n(y,t) \log|x-y| dx dy$$

Avec $M = \int n(x,t) dx$ (on a déja vu que c'était constant), et en notant $C(M) = \frac{M^2}{2}(1 + log(\pi) + log(M))$, on a :

$$\int n \log(n) dx \leq F(0) - \frac{\chi}{4\pi} C(M) - \frac{M}{2} \int n \log(n)$$

ce qui donne une borne supérieure de l'entropie.

On aurait également pu procéder comme suit :

$$F : t \to (1-\theta)\int_{\mathbb{R}^2} n(x,t)\log(n(x,t))dx$$
$$+\theta(\int_{\mathbb{R}^2} n(x,t)\log(n(x,t))dx + \frac{\chi}{4\pi\theta}\int\int_{\mathbb{R}^2\times\mathbb{R}^2} n(x,t)n(y,t)\log(|x-y|dxdy) \quad \text{(B.2.24)}$$

est bornée par sa valeur en 0, puisqu'elle est décroissante.
On choisit alors $\theta = \dfrac{\chi M}{8\pi}$ et on applique Hardy-Littlewood-Sobolev logarithmique :

$$(1-\theta)\int_{\mathbb{R}^2} n(x,t)\log(n(x,t))dx - \theta C(M) \leq F(0)$$

si $\chi M \leq 8\pi$, alors $\theta \leq 1$ et donc :

$$\int_{\mathbb{R}^2} n(x,t)\log(n(x,t))dx \leq \frac{F(0)+\theta C(M)}{1-\theta}$$

On obtient ainsi une borne supérieure de $\displaystyle\int n\log(n)$.
Pour la borne inférieure on utilise :

$$\frac{1}{1+t}\int_{\mathbb{R}^2} |x|^2 n(x,t)dx \leq K \ \forall t > 0$$

ce qui est vrai par hypothèse.
On en déduit que :

$$\int_{\mathbb{R}^2} n(x,t)\log(n(x,t))dx \geq \int_{\mathbb{R}^2} \frac{n(x,t)}{\mu(x,t)}\log(\frac{n(x,t)}{\mu(x,t)})\mu(x,t)dx - M\log[\pi(1+t)] - K$$

avec $\mu(x,t) = \dfrac{1}{\pi(1+t)}\exp(-\dfrac{x^2}{1+t})$
Par Jensen on obtient :

$$\int_{\mathbb{R}^2} \frac{n(x,t)}{\mu(x,t)}\log(\frac{n(x,t)}{\mu(x,t)})\mu(x,t)dx \geq X\log X \text{ où } X = \int_{\mathbb{R}^2} \frac{n(x,t)}{\mu(x,t)}\mu(x,t)dx = M$$

L'entropie $\displaystyle\int_{\mathbb{R}^2} n\log(n)$ est bornée inférieurement.
Il suffit donc de montrer des résultats de régularité pour obtenir l'existence.
La démonstration d'existence se déroule en deux temps : régulariser le problème afin de le résoudre avec un point fixe de Banach, puis obtenir des estimations a priori suffisantes pour avoir de la conpacité et pour passer à la limite.

## B.3 Démonstration du théorème ; existence et unicité

Oui mais de quel théorème ? Il faudrait le donner.
Donc le voilà : (C'est une phrase sans verbe. En voici une autre.)

**Théorème B.3.** *Si* $\chi M \leq 8\pi$ *et qu'on a les conditions (1.19) sur* $n_0$*, alors le problème* (1) *a une solution globale positive, avec :*

$$(1+|x|^2+|\log n|)n \in L^\infty_{loc}(\mathbb{R}^+, L^1(\mathbb{R}^2))$$

On a déjà vu que si une solution existait elle était positive, et que la condition sur $M$ était nécessaire. On a aussi vu que les normes $L^p$ se transmettaient. On verra que l'on a même mieux grâce à des propriétés d'hypercontractivité.

### B.3.a Régularisation

On approche le noyau du laplacien par la fonction $\kappa^\varepsilon(z) = -\frac{1}{2\pi}log|z|$ si $|z| > \varepsilon$ et $\kappa^\varepsilon(z) = -\frac{1}{2\pi}log|\varepsilon|$ sinon, avec $\varepsilon > 0$. On a donc le problème suivant (on oubliera d'écrire le $\varepsilon$ partout ) :

$$\begin{array}{ll} \frac{\partial n}{\partial t}(x,t) = \Delta n(x,t) - \chi\nabla.(n(x,t)\nabla c(x,t)) & \forall x \in \mathbb{R}^2, \quad t>0, \\ c(x,t) = (\kappa(.) * n(.,t))(x) & \forall x \in \mathbb{R}^2, \quad t>0, \\ n(x,t=0) = n_0(x) \geq 0 & \forall x \in R^2 \end{array} \tag{B.3.1}$$

avec $n^0 \in L^1(\mathbb{R}^2) > 0$. De plus : $\forall z/|z| >> \varepsilon, \nabla\kappa(z) = -\frac{1}{2\pi}\frac{z}{|z|^2}$

On va résoudre ce problème avec un point fixe de Banach.

On notera $E$ l'espace $C((0,T); L^1(R^2))$ muni de sa norme naturelle de Banach. On regarde alors l'application de $E$ dans $E$ qui à $m \in E$ associe la solution (dont on dira un mot plus tard, mais admise) du problème linéaire correspondant, c'est-à-dire :

$$\begin{array}{ll} \frac{\partial n}{\partial t}(x,t) = \Delta n(x,t) - \chi\nabla.(n(x,t)\nabla c(x,t)) & \forall x \in \mathbb{R}^2, \quad t>0, \\ c(x,t) = (\kappa(.) * m(.,t))(x) & \forall x \in \mathbb{R}^2, \quad t>0, \\ n(x,t=0) = n_0(x) \geq 0 & \forall x \in R^2 \end{array} \tag{B.3.2}$$

Le principe de positivité s'applique également sur ce problème. On peut donc se restreindre à l'espace où $m$ et donc $n$ sont positives. On notera $B^+$ cet espace, qui est toujours de Banach. On peut également se restreindre aux $m$ de masse $M$ égale à celle de $n_0$.

On s'attache donc ici à démontrer le caractère contractant de notre application. Pour ce faire, prenons $m1$ et $m2$ dans $B^+$ et leurs solutions correspondantes $n1$ et $n2$, puis calculons $n1-n2$. Pour cela, il est utile d'introduire le noyau de la chaleur : $Q(x,t) = \frac{1}{\pi t}e^{-\frac{|x|^2}{t}}$.

On peut alors écrire, par la formule de Duhamel :

$$n = n_0(.) * Q(.,t) - \int_0^T Q(.,t-s) * div(n\nabla c)(.,s)ds$$

que l'on réecrit :

$$n = n_0(.) * Q(.,t) + \int_0^T \nabla Q(.,t-s) * (n\nabla c)(.,s)ds$$

On peut donc écrire (car $n_0$ ne dépend pas de $m$) :

$$||n1 - n2||_{L^1} \leq \int_0^T ||\nabla Q(.,t-s) * ((n1\nabla c1)(.,s) - (n2\nabla c2)(.,s))||_{L^1} ds$$

Remarquons que $\nabla Q(x,t) = \dfrac{-2x}{t\pi t} e^{-\frac{|x|^2}{t}}$, donc $||\nabla Q||_{L^1} \leq \dfrac{A}{\sqrt{t}}$, et donc :

$$||n1 - n2||_{L^1} \leq \int_0^T \frac{A}{\sqrt{t-s}} ||((n1\nabla c1)(.,s) - (n2\nabla c2)(.,s))||_{L^1} ds$$

soit

$$||n1 - n2||_{L^1} \leq \int_0^T \frac{A}{\sqrt{t-s}} (||n1||_{L^1} ||(\nabla c1)(.,s) - (\nabla c2)(s,.))||_{L^\infty} + ||n1 - n2||_{L^1} ||(\nabla c2)(.,s))||_{L^\infty} ds \quad \text{(B.3.3)}$$

On note alors $M$ la masse constante des solutions. Par simple calcul, on a :

$$||n1 - n2||_{L^1} \leq A\sqrt{t}(M||\nabla\kappa||_{L^\infty} sup_{0,T}(||m1 - m2||_{L^1}) + sup_{0,T}(||n1 - n2||_{L^1})||\nabla\kappa||_{L^\infty} M$$

En passant au sup, en mettant à gauche ce qui est en $n$ et en supposant $t$ assez petit, on obtient alors :

$$||n1 - n2||_E \leq \frac{A\sqrt{t}M||\nabla\kappa||_{L^\infty}}{1 - A\sqrt{t}M||\nabla\kappa||_{L^\infty}} ||m1 - m2||_E = B||m1 - m2||_E$$

avec $B < 1$.

L'application est donc contractante, le théorème du point fixe de Banach nous livre alors une solution.

### B.3.b Un mot sur le problème linéaire

Pour l'existence du problème linéaire, j'ai été orienté vers une méthode de type équation de renouvellement, où les estimations découlent d'inégalités d'entropie et d'inégalités de type Poincaré. Le problème est que cette méthode nécessite que le noyau soit indépendant du temps, ce qui n'est pas le cas ici (et je m'en suis rendu compte tard).

Je ne suis pas allé plus loin dans la recherche de cette solution, qui parait-il est classique (il est peut être possible d'utiliser les méthodes de la chaleur).

Par contre, on a l'unicité de cette solution grâce au théorème de positivité demontré dans la première partie.

### B.3.c Passage à la limite

Revenons à la démonstration. Il faut maintenant obtenir des estimations pour passer à la limite. Pour cela on va utiliser le théorème d'Aubin, que je vais brièvement justifier.

**Lemme d'Aubin**

**Lemme B.3.1.** *Soit $T > 0$, $p \in (1, \infty)$ et $f_n$ une famille de fonctions bornées dans $L^p((0,T);H)$, òu $H$ est un espace de Banach. De plus, soit $V$ un espace s'injectant compactement dans $H$ et $V'$ son dual.*
*Si les $f_n$ sont bornées dans $L^p((0,T);V)$ et que les $\frac{\partial}{\partial t} f_n$ sont bornées uniformément par rapport à $n$ dans $L^p((0,T);V')$, alors $f_n$ est relativement compacte dans $L^p((0,T),H)$.*

Le lemme d'Aubin nous permet de récupérer de la compacité en temps à partir de la compacité en espace. Dans notre cas, on va prendre $H = L^2(\mathbb{R}^2)$, $V = \{u \in H^1(\mathbb{R}^2), \sqrt{|x|}u \in L^2(\mathbb{R}^2)\}$ et $p = 2$.
Il faut donc montrer les estimations permettant de justifier ces espaces, et notament de transformer nos estimations $L^1$ en estimations $L^2$.

Digression, justification

On aurait pu dans le même esprit se borner aux estimations $L^1$. Je vais expliquer ce point par un petit calcul qui en même temps justifiera le lemme d'Aubin (moralement en tout cas, car on verra comment de la compacité en espace peut donner de la compacité en temps).

**Lemme B.3.2.** *Soit $n_k(x,t)$ des fonctions telles que :*
*$\omega(\beta) = sup_{1 \leq k, 0 \leq t \leq T} \int_B |n_k(x+\beta,t) - n_k(x,t)|dx \underset{\beta \to 0}{\longrightarrow} 0$ ($\approx$ les $n_k$ sont compacts en $x$ uniformément en temps), $\frac{\partial}{\partial t} n_k = D^\alpha \phi_k(x,t)$ et $||\phi_k||_{L^\infty} < \infty$*
*alors $n_k$ est compacte dans $C((0,T), L^1(B))$*

Démonstration :

On prend une approximation de l'unité $\rho_\varepsilon(x)$ telle que :
$\rho_\varepsilon(x) = \frac{1}{\varepsilon^d}\rho(\frac{x}{\varepsilon})$ avec $\rho \in D(\mathbb{R}^d)$, $\rho \geq 0$ et $\int \rho = 1$
On convole alors l'équation dans le but de mettre les dérivées sur $\rho$ :

$$\frac{\partial}{\partial t} n_k(x,t) * \rho_\varepsilon(x) = \phi * D^\alpha \rho_\varepsilon$$

On réecrit alors l'équation (en oubliant à partir de maintenant les $k$) :

$$\frac{\partial}{\partial t} n(x,t) = \frac{\partial}{\partial t}(n(x,t) - n(x,t) * \rho_\varepsilon(x)) + \phi * D^\alpha \rho_\varepsilon$$

Et on s'intéresse logiquement à une différence d'incrémentation en $t$ en intégrant l'équation entre $t$ et $t+h$ :

$$n(x,t+h) - n(x,t) = [\int n(x,t) - n(y,t)\rho_\varepsilon(x-y)dy]_t^{t+h} + \int_t^{t+h} \phi * D^\alpha \rho_\varepsilon ds$$

Calcul à verifier ici.

On intègre sur la boule $B$.

$$||n(x,t+h)-n(x,t)||_{L^1(B)} \leq 2sup_{(t,t+h)}\Big(\int \rho_\varepsilon(z)\int |n(x+z,t)-n(x,t)|dxdz + \int_t^{t+h} ||\phi||_{L^1(2B)}\frac{c}{\varepsilon^\alpha}ds \qquad \text{(B.3.4)}$$

$$\leq 2\omega_\varepsilon + \frac{c_\phi h}{\varepsilon^\alpha}$$

Il suffit alors de prendre $\varepsilon^\alpha = \sqrt{h}$ et la borne devient $2\omega_\varepsilon + c_\phi\sqrt{h}$. On passe alors au sup sur $t$, et on obtient :

$$||n(.+h,x)-n(.,x)||_{C((0,T),L^1(B))} \leq 2\omega_\varepsilon + c_\phi\sqrt{h}$$

La borne étant aussi petite que l'on veut quand $h$ tend vers 0, on obtient bien la compacité en temps.
Remarque : d'autres conditions auraient donné la compacité $L^1$.
Ce résultat en tant que tel ne sera pas appliqué, mais il permet de rendre un peut moins abstrait le lemme d'Aubin.

#### Estimations

Je donne ici une série d'estimations a priori qui permettent de passer à la limite, la difficulté étant le terme non linéaire pour lequel on ne peut pas se contenter d'une limite distribution.

**Lemme B.3.3.** *Les solutions du problème régularisé vérifient les assertions suivantes, uniformément par rapport à $\varepsilon$ et avec des bornes dépendant uniquement de $\int(1+|x|^2)n_0dx$ et $\int n_0log(n_0)dx$ :*
*(i) $|x|^2n^\varepsilon(x,t)$ est bornée dans $L^\infty(\mathbb{R}^+_{loc};L^1(\mathbb{R}^2))$*
*(ii) $\int n^\varepsilon(x,t)\log n^\varepsilon(x,t)dx$ et $\int n^\varepsilon(x,t)c^\varepsilon(x,t)$ sont bornées dans $\mathbb{R}$*
*(iii) $n^\varepsilon(x,t)\log n^\varepsilon(x,t)$ est bornée dans $L^\infty(\mathbb{R}^+_{loc};L^1(\mathbb{R}^2))$*
*(iv) $\nabla\sqrt{n}^\varepsilon(x,t)$ est bornée dans $L^2(\mathbb{R}^+_{loc}\times\mathbb{R}^2)$*
*(v) $n^\varepsilon(x,t)$est bornée dans $L^2(\mathbb{R}^+_{loc}\times\mathbb{R}^2)$*
*(vi) $n^\varepsilon(x,t)\Delta c^\varepsilon(x,t)$ est bornée $L^1(\mathbb{R}^+_{loc}\times\mathbb{R}^2)$*
*(vii) $\sqrt{n}^\varepsilon(x,t)\nabla c^\varepsilon(x,t)$ est bornée $L^2(\mathbb{R}^+_{loc}\times\mathbb{R}^2)$.*

Commençons par montrer ces estimations, nous les exploiterons après.
Démonstration : (i) La même démonstration que pour le lemme (1.1) s'applique, c'est-à-dire :

$$\frac{d}{dt}\int_{\mathbb{R}^2}|x|^2n^\varepsilon(x,t)dx = \int_{\mathbb{R}^2}|x|^2(\Delta n^\varepsilon - \chi\nabla.(n^\varepsilon\nabla c^\varepsilon))dx$$

$$= 4M + 2\chi\int_{\mathbb{R}^2}\int_{\mathbb{R}^2} n^\varepsilon(x,t)n^\varepsilon(y,t)(\nabla\kappa^\varepsilon(x)-\nabla\kappa^\varepsilon(y)).(x-y)$$

$$\leq 4M - \frac{\chi}{2\pi}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\frac{n^\varepsilon(x,t)n^\varepsilon(y,t)}{|x-y|} \leq 4M.$$

Pour (ii), on calcule et on trouve :

$$\frac{d}{dt}\Big(\int n^{\varepsilon}(x,t)\log n^{\varepsilon}(x,t)dx - \frac{\chi}{2}\int n^{\varepsilon}(x,t)c^{\varepsilon}(x,t)dx\Big)$$
$$= -\int n^{\varepsilon}(x,t)|\nabla(\log(n^{\varepsilon}(x,t)) - \chi\nabla c^{\varepsilon}(x,t)|^2 dx \quad \text{(B.3.5)}$$

Puis le même calcul que dans le lemme (1.2) nous permet d'obtenir les estimations (ii).
Pour (iii) démontrons le lemme (d'équiintégrabilité) suivant :

**Lemme B.3.4.** *Si $u \in L^1_+$ et que l'entropie et le second moment sont bornés, alors $u\log u$ est uniformément borné dans $L^{\infty}(\mathbb{R}^+_{loc}; L^1(\mathbb{R}^2))$*

La seule difficulté est quand $u \leq 1$ :

$$\int_{u\leq 1} n|\log n| \leq \int_{e^{-|x|^2}\leq u\leq 1} n|\log n| + \int_{u\leq e^{-|x|^2}, u\leq \frac{1}{e}} n|\log n| + \int_{\frac{1}{e}\leq u\leq e^{-|x|^2}} n|\log n|$$

Comme $|\log|$ est décroissant sur $[0,1]$, on a $\displaystyle\int_{e^{-|x|^2}\leq u\leq 1} n|\log n| \leq \int |x|^2 u dx$.
Sur $[0,\frac{1}{e}]$, $.|\log .|$ est croissante, donc $\displaystyle\int_{u\leq e^{-|x|^2}, u\leq \frac{1}{e}} n|\log n| \leq \int |x|^2 e^{-|x|^2}$ qui est fini Il reste le dernier morceau, mais on intègre sur un ensemble de mesure finie une quantité bornée, donc il est fini, ce qui prouve le lemme, et donc le (iii).
Passons au (iv)
On a déjà vu qu'un simple calcul donne :

$$\frac{d}{dt}\int n\log n dx \leq -4\int |\nabla\sqrt{n}|dx + \chi n(-\Delta c)dx$$

Comme on a déjà une borne pour $\displaystyle\int n\log n$, il reste à estimer le dernier terme :

$$\int n(-\Delta c)dx = \int n(-\Delta(\kappa^{\varepsilon} * n)dx = (1) + (2) + (3)$$

avec $(1) = \displaystyle\int_{n<K} n(-\Delta(\kappa^{\varepsilon} * n))dx$, $(2) = \displaystyle\int_{K\leq n} n(-\Delta(\kappa^{\varepsilon} * n))dx - (3)$ et $(3) = \displaystyle\int_{K\leq n} |n|^2 dx$.
Pour simplifier les calculs, on écrit $-\Delta\kappa^{\varepsilon} = \frac{1}{\varepsilon^2}\phi_1(\frac{1}{\varepsilon})$. Par définition, on sait que $\phi_1 > 0$. On peut alors écrire :

$$(1) = \int_{n<K} n\int \frac{1}{\varepsilon^2}\phi_1(\frac{x-y}{\varepsilon})n(y,t)dx \leq KM$$

On espère contrôler (2) car $-\Delta(\kappa^{\varepsilon})$ tend faiblement vers un dirac. Précisemment :

$$(2) = \int_{K\leq n} n(x,t)\int (n(x-\varepsilon y,t) - n(x,t))\phi_1(y)dydx \leq$$
$$\int_{K\leq n}\int (\sqrt{n(x-\varepsilon y,t)} - \sqrt{n(x,t)})\sqrt{\phi_1(y)})(\sqrt{n(x-\varepsilon y,t)} - \sqrt{n(x,t)} + 2\sqrt{n(x,t)}\sqrt{\phi_1(y)})dydx \quad \text{(B.3.6)}$$

On utilise alors Cauchy-Schwartz et l'ingalité $(a+2b)^2 \leq 2a^2 + 8b^2$, ce qui donne :

$$\leq \int_{K\leq n} (||\phi 1||_{L^\infty}^{\frac{1}{2}} \int_{\frac{1}{2}\leq y\leq 2} |(\sqrt{n(x-\varepsilon y,t)} - \sqrt{n(x,t)})|^2 dy)^{\frac{1}{2}})$$
$$\cdot (\int |((\sqrt{n(x-\varepsilon y,t)} - \sqrt{n(x,t)})|^2 + 8|n(x,t)|)\phi_1(y)dy)^{\frac{1}{2}})dx \quad \text{(B.3.7)}$$

On utilise alors l'inégalité de Poincaré, ce qui donne :

$$(2) \leq \int_{K\leq n} n||\phi 1||_{L^\infty}^{\frac{1}{2}} C_p ||\nabla\sqrt{n}||_{L^2(\mathbb{R}^2)} \times [\sqrt{2}||\phi 1||_{L^\infty}^{\frac{1}{2}} C_p ||\nabla\sqrt{n}||_{L^2(\mathbb{R}^2)} + 2\sqrt{2}\sqrt{|n|}||\phi_1||_{L^1(\mathbb{R}^2)}^{\frac{1}{2}}]dx$$

Faisons une légère pose pour contrôler (3). On réutilise GNS dans ce cas :

$$\int_{K\leq n} |n|^2 dx \leq C_{GNS}^2 \int_{K\leq n} ||\nabla\sqrt{n}||_{L^2(\mathbb{R}^2)}^2 dx \int_{K\leq n} ndx$$

On a déjà vu que l'on peut rendre $\int_{K\leq n} ndx$ aussi petit que l'on veut (avec $K$ assez grand). On majore donc ce terme par un $\eta(K)$, et donc :

$$\int_{K\leq n} |n|^2 dx \leq C\eta(K) \int_{K\leq n} ||\nabla\sqrt{n}||_{L^2(\mathbb{R}^2)}^2 dx$$

Par Cauchy-Schwarz on a alors :

$$\int_{K\leq n} |n|^{\frac{3}{2}} dx \leq (\int_{K\leq n} |n| dx)^{\frac{1}{2}} (\int_{K\leq n} |n|^2 dx)^{\frac{1}{2}}$$
$$\leq \eta(K) C_{GNS} ||\nabla\sqrt{n}||_{L^2(\mathbb{R}^2)}$$

Donc en sommant (2) + (3) on obtient :

$$(2) + (3) \leq B\eta(K) ||\nabla\sqrt{n}||_{L^2(\mathbb{R}^2)}^2$$

En revenant au début, en notant $X(t) = ||\nabla\sqrt{n}||_{L^2(\mathbb{R}^2)}^2$ et en prenant $\eta(K) < \frac{4}{B}$, on a :

$$\frac{d}{dt} \int n \log n dx \leq MK + (-4 + B\eta(K))X(t)$$

Et donc en intégrant on obtient la borne voulu.

Ouf on passe maintenant à (v) qui découle de l'inégalité de Gagliardo-Niremberg-Sobolev :

$$||u||_p^2 \leq C_{GNS}^{(p)} ||\nabla u||_{L^2}^{\frac{4}{p}} ||u||_{L^2}^{2-\frac{4}{p}} \; u = \sqrt(n), p = 4. \quad \text{(B.3.8)}$$

Pour (vi) : on l'a vu dans (iv).

Reste (vii) qui est l'estimation importante pour le passage à la limite. Calculons alors :

$$\frac{d}{dt}\int \frac{1}{2} ncdx = \int c(\Delta n - \chi\nabla(n\nabla c))dx$$
$$= \int n\Delta cdx + \chi \int n|\nabla c|^2 dx$$

On peut écrire

$$\int_{(0,T)\times\mathbb{R}^2} n|\nabla c|^2 dxdt \leq \frac{1}{2\chi}|\int ncdx - \int n_0(\kappa * n_0)dx| + \frac{1}{\chi}\int_0^T \int_{\mathbb{R}^2} n(-\Delta c)dx.$$

(vi) et (ii) donnent les bornes des termes de droite.

On remarque que ces majorations sont uniformes en $\varepsilon$ (c'est facile sans l'écrire). Le lemme d'Aubin donne l'existence d'un $n$ limite. Il faut alors passer à la limite dans l'équation et surtout dans le terme non linéaire. Au passage une chose importante à voir pour passer à la limite est que $n|\nabla(\log n) - \chi\nabla c|$ est bornée dans $L^1$ en temps et espace.

La difficulté se trouve donc dans le terme $n^\varepsilon(x,t)\nabla c^\varepsilon(x,t)$ (terme non linéaire). On remarque qu'il est borné dans $L^1((0,T)\times\mathbb{R}^2)$ uniformément par rapport à $\varepsilon$. En effet :

$$(\int\int_{[0,T]\times\mathbb{R}^2} n^\varepsilon|\nabla c^\varepsilon|dxdt)^2 \leq$$
$$(\int\int_{[0,T]\times\mathbb{R}^2} n^\varepsilon dxdt)(\int\int_{[0,T]\times\mathbb{R}^2} n^\varepsilon|\nabla c^\varepsilon|^2 dxdt) = MT(\int\int_{[0,T]\times\mathbb{R}^2} n^\varepsilon|\nabla c^\varepsilon|^2 dxdt) \quad \text{(B.3.9)}$$

Et par (vii) on contrôle ce dernier terme. Par contre, à cause de la non-linearité, cela ne suffit par pour que $n$ soit solution. Pour cela, il faut recupérer une limite forte. On remarque alors grâce a l'inégalitée G.N.S que pour $p > 2$ et $t \in \mathbb{R}^+$, on a :

$$\int_{\mathbb{R}^2} |n^\varepsilon|^{p/2} \leq (C_{GNS}^{(p)} M \int_{\mathbb{R}^2} |\nabla\sqrt{n}^\varepsilon|^2 dx)^{p/2-1}$$

Ce qui donne $n^\varepsilon \in L^q(\mathbb{R}^+_{loc}\times\mathbb{R}^2)$ pour tout $q \in [1,\infty)$.
On a donc a une extraction près (on ne change pas la notation) la convergence faible de $n^\varepsilon$ vers $n$ dans tout ces espaces. Reste a évaluer $\nabla c^\varepsilon$ :

$$\nabla c^\varepsilon - \nabla c =$$
$$-\frac{1}{2\pi}\int_{\mathbb{R}^2}\frac{x-y}{|x-y|^2}(n^\varepsilon(y,t)-n(y,t))dy + \int_{|x-y|\leq 2\varepsilon}(\frac{1}{\varepsilon}\nabla\kappa(\frac{x-y}{\varepsilon}) + \frac{x-y}{2\pi|x-y|^2})n^\varepsilon(y,t)dy. \quad \text{(B.3.10)}$$

On a donc convergence p.p.t. $(x,t)$ vers 0 Ce qui maintenant suffit pour dire que $n^\varepsilon\nabla c^\varepsilon$ converge vers $n\nabla c$ au sens des distributions. De plus les estimations (v), (vii), et (iv) passent alors par semi continuité à la limite. On en déduit via

$$\int\int_{[0,T]\times\mathbb{R}^2} n|(\nabla \log n - \chi\nabla c)|^2 dxdt =$$
$$4\int\int_{[0,T]\times\mathbb{R}^2} |\nabla\sqrt{n^\varepsilon}|^2 dxdt - 2\chi\int\int_{[0,T]\times\mathbb{R}^2} |n^\varepsilon|^2 dxdt + \chi^2\int\int_{[0,T]\times\mathbb{R}^2} n^\varepsilon|\nabla c^\varepsilon|^2 dxdt \quad \text{(B.3.11)}$$

et via l'inégalité de Cauchy-Schwarz :

$$\int\int_{[0,T]\times\mathbb{R}^2} n|(\nabla \log n - \chi\nabla c)| dxdt \leq$$
$$(\int\int_{[0,T]\times\mathbb{R}^2} ndxdt)^{1/2}(\int\int_{[0,T]\times\mathbb{R}^2} n|(\nabla \log n - \chi\nabla c)|^2 dxdt)^{1/2} \quad \text{(B.3.12)}$$

que le flux est bien dans $L^1([0,T]\times\mathbb{R}^2)$, et donc l'existence de la solution.

### B.3.d Ultracontractivité

La transmission des normes $L^p$ telle qu'on l'a montrée nécessite une borne à l'instant initial. Cette condition initiale n'est en fait pas nécessaire. En effet, il existe un résultat d'hypercontractivité qui dit que pour tout temps $t > 0$, la norme $L^p$ est finie. Pour l'exprimer en termes mathématiques, on peut écrire :

**Théorème B.4.** *On considère une solution de (1) avec les hypothèses (1.19) et $\chi M \leq 8\pi$. Alors pour tout $p \in (1, +\infty)$, il existe une fonction notée $h_p$, continue sur $(0, +\infty)$, telle que pour presque tout $t$, $||n(.,t)||_{L^p} \leq h_p(t)$.*

On voit que l'on autorise la valeur infinie en 0.
Je ne tape pas la preuve ici. Elle mélange un peu tout ce qui a déjà été utilisé. Le calcul se rapproche de la méthode de Jager et Luckhaus (avec les $(n-K)_+$). La différence (quand même importante) est que l'exposant $p$ n'est pas fixe : on prend pour tout $t$ une fonction affine $p(s)$ telle que $p(0) = 1$ (on sait que la norme $L^1$ de $n$ est finie en $t = 0$) et $p(t) = P$. Cela rajoute quelques termes en plus dans la dérivation. Les nouveaux termes sont alors contrôlés avec une inégalité de Sobolev logarithmique.

En réalité le problème n'est pas complètement résolu, il semblerait que la distance de Waserstein aux solutions constantes soit en fait la bonne estimation du problème.

## B.4 Inégalité de Gagliardo-Niremberg-Sobolev

**Théorème B.5.** $\forall u \in H^1(\mathbb{R}^2)$, $\forall r \in [2,\infty)$, *on a :*

$$||u||^2_{\mathbb{R}} \leq C^{(r)}_{GNS}||\nabla u||^{\frac{4}{r}}_{L^2(\mathbb{R}^2)}||u||^{2-\frac{4}{r}}_{L^2(\mathbb{R}^2)}$$

Démonstration :

Cette équation est une amélioration des inégalités de Sobolev traditionnelles avec de l'interpolation. En effet, pour $u \in S(\mathbb{R}^d)$ (les inégalités générales s'en déduisent par densité), l'inégalité de Sobolev s'écrit :

$$||u||_{L^{p^*}} \leq ||\nabla u||_{L^p}$$

avec $\frac{1}{p^*} = \frac{dp}{d-p}$ et $1 \leq p < d$. En particulier on n'a pas d'inégalité lorsque $p = d$. Comme de plus :

$$||u||_{L^q} \leq ||u||_{L^q}$$

pour $1 \leq q \leq \infty$, par interpolation on a pour $\frac{1}{r} = \frac{\theta}{q} + \frac{1-\theta}{p^*}, \theta \in [0,1]$ :

$$||u||_{L^r} \leq ||u||_{L^q}^{\theta} ||\nabla u||_{L^p}^{1-\theta}$$

Remarque : c'est plus simplement Holder puis Sobolev.
On remarque que la condition sur les indices peut sécrire $\frac{1}{r} = \frac{\theta}{q} + \frac{(1-\theta)(d-p)}{dp}$, ce qui laisse à penser que l'on peut prendre $d = p$. Et bien c'est vrai.
Et c'est celle-là que l'on utilise. En effet, pour $p = q = d = 2$ et $\theta = \frac{2}{r}$, on a :

$$||u||_r^2 \leq C_{GNS}^{(r)} ||\nabla u||_{L^2(R^2)}^{\frac{4}{r}} ||u||_{L^2(R^2)}^{2-\frac{4}{r}}$$

## B.5 Inégalité de Hardy-Littlewood-Sobolev

Pour ces inégalité les références sont [78, 79]

### B.5.a Tout d'abord la normale

**Théorème B.6.** *Soit $n \in \mathbb{N}$, $0 < \lambda < n$, $p > 1$ et $r > 1$ tels que $\frac{1}{p} + \frac{\lambda}{n} + \frac{1}{r} = 2$.*
*Alors il existe $C$ ne dépendant que de $n, p$ et $\lambda$ tel que pour tout $f \in L^p(\mathbb{R}^n)$ et $g \in L^r(\mathbb{R}^n)$ on ait :*

$$\int_{\mathbb{R}^N} \int_{\mathbb{R}^N} f(x)|x-y|^{-\lambda} g(y) dx dy \leq C ||f||_p ||g||_r$$

*De plus, si on note encore $C$ la constante optimale, on a :*

$$C \leq \frac{n}{n-\lambda} \frac{|S^{n-1}|}{n}^{\frac{\lambda}{n}} \frac{1}{rp} ((\frac{\frac{\lambda}{n}}{1-\frac{1}{r}})^{\lambda} n + (\frac{\frac{\lambda}{n}}{1-\frac{1}{p}})^{\lambda} n)$$

### B.5.b Démonstration

Par homogénéité on prend $f$ et $g$ normées à un (pour leur norme associée). On utilisera également des fonctions positives ($|\int| < \int ||$). On notera $\chi$ les fonctions caractéristiques d'ensemble.
On a alors :

$$f(x) = \int_0^{\infty} \chi_{0 \leq a \leq f(x)} da$$

$$g(y) = \int_0^\infty \chi_{0 \leq b \leq g(y)} db$$
$$|x-y|^\lambda = \int_0^\infty \lambda . c^{-\lambda-1} \chi_{B(0,c)}(x-y) dc$$

On réecrit alors le membre de gauche de l'inégalité, on applique Fubini-Tonelli (tout est positif) et on obtient :

$$K = \int_0^\infty \int_0^\infty \int_0^\infty \lambda I(a,b,c) dadbdc$$

avec

$$I(a,b,c) = \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \chi_{0 \leq a \leq f(x)} \chi_{0 \leq b \leq g(y)} c^{-\lambda-1} \chi_{B(0,c)}(x-y) dxdy.$$

Le but est donc de majorer $I$ de la façon la plus faible possible. On va donc majorer par 1 le terme le moins souvent nul (un terme en moins suffit car on peut alors réaliser les intégrations). Par commodité notons :

$$u(a) = \int_{\mathbb{R}^N} \chi_{0 \leq a \leq f(x)} dx$$
$$v(b) = \int_{\mathbb{R}^N} \chi_{0 \leq b \leq g(y)} dy$$
$$w(c) = \int_{\mathbb{R}^N} \chi_{B(0,c)}(z) dz = c^n \frac{|S^{n-1}|}{n}$$

On a alors, avec le changement de variable $z = x-y$ lorsque c'est nécessaire : $I \leq c^{-\lambda-1} min(u(a)v(b), u(a)w(c)$
De facon plus interessante :

$$I \leq c^{-\lambda-1} \frac{u(a)v(b)w(c)}{max(u(a), v(b), w(c))}$$

Il faut ensuite intégrer trois fois de 0 à l'infini. Notons pour la suite la remarque sur les normes suivantes (directe avec Fubini) :

$$||f||_p^p = \int_0^\infty pa^{p-1} u(a) da = 1$$
$$||g||_r^r = \int_0^\infty rb^{r-1} v(b) db = 1$$

On peut remarquer que $a$ et $b$ jouent des rôles symétriques et que l'on sait faire un peu plus de choses sur $c$. On va donc commencer par lui.
On suppose ici que $u(a) < v(b)$. Alors :

$$\int_0^\infty I dc = \int_{w(c) \leq u(a)} c^{-\lambda-1} w(c) v(b) dc + \int_{w(c) > u(a)} c^{-\lambda-1} u(a) v(b) dc.$$

On note $I1$ et $I2$ les deux intégrales obtenues. On voit que ces deux quantités sont faciles à calculer, il suffit de voir que $w(c) = u(a)$ entraîne $c = (\frac{nu(a)}{S^{n-1}})^{\frac{1}{n}}$ (ici $w$ est croissant et $u(a) fixe$).

On a donc $w(c) \leq u(a)$ tant que $c$ est plus petit que cette valeur limite. On a donc :

$$I1 = v(b)\frac{1}{n-\lambda}(\frac{S^{n-1}}{n})^{\frac{\lambda}{n}})u(a)^{1-\frac{\lambda}{n}}$$

$$I2 = (\frac{S^{n-1}}{n})^{\frac{\lambda}{n}}\frac{1}{\lambda}v(b)u(a)^{1-\frac{\lambda}{n}}.$$

La valeur en 0 de $I1$ est nulle car $n-\lambda > 0$, et la valeur en l'infini de $I2$ est nulle car $\lambda > 0$. Ainsi :

$$\int_0^\infty Idc \leq (\frac{S^{n-1}}{n})^{\frac{\lambda}{n}}\frac{n}{\lambda(n-\lambda)}v(b)u(a)^{1-\frac{\lambda}{n}}.$$

et par symétrie, si $u(a) \geq v(b)$, on a :

$$\int_0^\infty Idc \leq (\frac{S^{n-1}}{n})^{\frac{\lambda}{n}}\frac{n}{\lambda(n-\lambda)}u(a)v(b)^{1-\frac{\lambda}{n}}.$$

Il faut maintenant intégrer suivant $a$ et $b$, mais on ne sait pas donner une description explicite en fonction de $a$ et $b$ des domaines mis en jeu ($f$ et $g$ étant quelconques). On va donc partitionner différement. Pour cela remarquons le fait suivant (qui est du au bon choix de majoration) :

$$\int_0^\infty \lambda Idc \leq (\frac{S^{n-1}}{n})^{\frac{\lambda}{n}}\frac{n}{(n-\lambda)}min(u(a)v(b)^{1-\frac{\lambda}{n}}, v(b)u(a)^{1-\frac{\lambda}{n}}. \tag{B.5.1}$$

En effet, si $v(b) > u(a)$, alors $v(b)u(a)^{1-\frac{\lambda}{n}} > u(a)v(b)^{1-\frac{\lambda}{n}}$. Cela nous permet de choisir n'importe quelle partition de $\mathbb{R}^2$ et le terme à intégrer dessus.
On note $m(a,b)$ ce minimum, et on prend alors la partition définie par $A = \{a, b \in R^2, a^p < b^r\}$, $B = \{a, b \in R^2, a^p \geq< b^r\}$.
Sur A on choisi de majorer $m$ par $v(b)u(a)^{1-\frac{\lambda}{n}}$, l'interet étant d'avoir des bornes finies pour $u$, qui contient les difficultées. On fait l'inverse sur $B$. On s'intéresse donc à :

$$J1 = \int_0^\infty v(b)(\int_0^{b^{\frac{r}{p}}} u(a)^{1-\frac{\lambda}{n}}da)db,$$

$$J2 = \int_0^\infty u(a)(\int_0^{b^{\frac{r}{p}}} v(b)^{1-\frac{\lambda}{n}}db)da.$$

et plus particulierement à $J11 = \int_0^{b^{\frac{r}{p}}} u(a)^{1-\frac{\lambda}{n}}da$.
Étant donné la remarque sur l'expression de la norme de $f$ faite précedement, on a envie d'appliquer Holder de façon à faire apparaître $u(a)a^{p-1}$. Cela nous incite à poser $\frac{1}{k} = 1 - \frac{\lambda}{n}$ et $\frac{1}{s} = \frac{\lambda}{n}$, puis à multiplier et diviser à l'intérieur de l'intgrale par $a^{\frac{p-1}{k}}$.
On applique alors l'inégalité de Holder car $f \in L^p$ (ce qui est une condition suffisante pour que le premier terme soit $L^k$ du segment voulu) et $a^{-\frac{p-1}{k}}$ est dans $L^s$ puisque $-s.\frac{p-1}{k} = -(p-1)\frac{n-\lambda}{\lambda} > -1$ vu la relation sur $n, p, \lambda, r$ et le fait que $r > 1$. On remarque ici l'intérêt de la borne finie de l'intégrale.

Bref comme disait pepin on obtient (en majorant la borne supérieure par l'infini pour $f$) :

$$J11 \leq (\frac{1}{p}||f||)^{\frac{1}{k}} (\int_0^{b^{\frac{r}{p}}} (a^{-(p-1)\frac{n-\lambda}{\lambda}})^{\frac{\lambda}{n}}$$

Soit après calcul facile mais desagréable, on obtient :

$$J11 \leq (\frac{1}{p}(\frac{\frac{\lambda}{n}}{1-\frac{1}{r}}))^{\frac{\lambda}{n}} p^{\frac{\lambda}{n}-1} b^{r-1}.$$

De même avec des notations évidentes :

$$J21 \leq (\frac{1}{r}(\frac{\frac{\lambda}{n}}{1-\frac{1}{p}}))^{\frac{\lambda}{n}} r^{\frac{\lambda}{n}-1} a^{p-1}.$$

Il suffit pour finir de réintégrer $J11$ par rapport à $b$. Cela se résume à une constante près à :

$$\int_0^{\infty} v(b) b^{r-1} db = \frac{1}{r}$$

ceci toujours grâce à la mème remarque sur les normes et grâce au fait que celle de $g$ vaut 1.
Le calcul est analogue sur $J21$.
On a ainsi démontré l'inégalité de Hardy-Littlewood-Sobolev avec la constante :

$$\frac{n}{n-\lambda}(\frac{|S^{n-1}|}{n})^{\frac{\lambda}{n}} \frac{1}{rp}((\frac{\frac{\lambda}{n}}{1-\frac{1}{r}})^{\frac{\lambda}{n}} + (\frac{\frac{\lambda}{n}}{1-\frac{1}{p}})^{\frac{\lambda}{n}})$$

annoncé comme majorante de l'optimale.

### B.5.c Inégalité de Hardy-Littlewood-Sobolev

**Théorème B.7.** $\int_{\mathbb{R}^d} f \log(f) dx + \frac{d}{M} \int \int_{\mathbb{R}^d \times \mathbb{R}^d} f(x) f(y) \log|x-y| dxdy \geq -C(M)$, *avec* $C(M) = M(1 + \log(\pi) - \log(M))$

Pour la démonstration, on considère les inégalités "normales" comme une famille de fonctionelles $\Phi_\lambda$ positives, définies sur les fonctions infiniment dérivables à support compact et indexée par le paramètre $\lambda$ positif, avec $\Phi_0 = 0$ (en 0 on a égalité). Ce faisant on peut dériver cette famille en $\lambda = 0$ et comme $0 \leq \frac{\Phi_\lambda}{\lambda}$ on obtient que $\Phi_0' \geq 0$. Cette méthode donne des inégalités que l'on peut qualifier d'extrèmales.
Dans notre cas on obtient :

$$\Phi_\lambda = C||f||_p||g||_r - \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} f(x)|x-y|^{-\lambda} g(y) dxdy$$

Il reste alors à estimer la constante, et c'est là le point le plus dur. Il existe deux méthodes, une de Carlen et Loss, et l'autre de Beckner. Je n'ai pas refait toutes les étapes de ces démonstrations point par point, c'est pourquoi je me contente d'en expliquer seulement l'idée ici.

Nous avons besoin de l'inégalité sur l'espace entier, alors que dans les démonstrations sont énoncées sur une boule. Cependant, il est précisé que via la projection stéréographique, on passe facilement d'une boule a l'espace.

### B.5.d Carlen et Loss

La démonsration est basée sur les fonctions qui rendent optimales les inégalités normales [26]. Ces fonctions sont obtenues par une méthode de compétition symétrique. La méthode de base est d'utiliser l'invariance par transformation conforme de l'inégalité et de réaranger alors la masse des fonctions avec des réarangements symétriques. Ces transformations (réalisables uniquement pour $p = r$) permettent d'optimiser l'inégalité.

### B.5.e Beckner

La méthode de Beckner est basée sur une décomposition du noyau en série de fonctions sphériques harmoniques, ce qui permet de calculer ensuite [2].

# Bibliography


[1] 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 Verlag, Basel, second edition, 2008.

[2] W. Beckner. Sharp Sobolev inequalities on the sphere and the Moser-Trudinger inequality. *Ann. of Math. (2)*, 138(1):213–242, 1993.

[3] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.

[4] E. Bierstone and P. D. Milman. Semianalytic and subanalytic sets. *Inst. Hautes Études Sci. Publ. Math.*, (67):5–42, 1988.

[5] P. Biler. Existence and nonexistence of solutions for a model of gravitational interaction of particles. III. *Colloq. Math.*, 68(2):229–239, 1995.

[6] P. Biler, G. Karch, P. Laurençot, and T. Nadzieja. The $8\pi$-problem for radially symmetric solutions of a chemotaxis model in the plane. *Math. Methods Appl. Sci.*, 29(13):1563–1583, 2006.

[7] P. Biler and T. Nadzieja. Existence and nonexistence of solutions for a model of gravitational interaction of particles. I. *Colloq. Math.*, 66(2):319–334, 1994.

[8] A. Blanchet, V. Calvez, and J. A. Carrillo. Convergence of the mass-transport steepest descent scheme for the subcritical Patlak-Keller-Segel model. *SIAM J. Numer. Anal.*, 46(2):691–721, 2008.

[9] A. Blanchet, E. A. Carlen, and J. A. Carrillo. Functional inequalities, thick tails and asymptotics for the critical mass Patlak-Keller-Segel model. *J. Funct. Anal.*, 262(5):2142–2230, 2012.

[10] A. Blanchet, J. A. Carrillo, and P. Laurençot. Critical mass for a Patlak-Keller-Segel model with degenerate diffusion in higher dimensions. *Calc. Var. Partial Differential Equations*, 35(2):133–168, 2009.

[11] A. Blanchet, J. A. Carrillo, and N. Masmoudi. Infinite time aggregation for the critical Patlak-Keller-Segel model in $\mathfrak{R}^2$. *Comm. Pure Appl. Math.*, 61(10):1449–1481, 2008.

[12] A. Blanchet, J. Dolbeault, and B. Perthame. Two-dimensional Keller-Segel model: optimal critical mass and qualitative properties of the solutions. *Electron. J. Differential Equations*, pages No. 44, 32 pp. (electronic), 2006.

[13] B. Bonnard, J.-B. Caillau, and L. Rifford. Convexity of injectivity domains on the ellipsoid of revolution: the oblate case. *C. R. Math. Acad. Sci. Paris*, 348(23-24):1315–1318, 2010.

[14] Y. Brenier. Décomposition polaire et réarrangement monotone des champs de vecteurs. *C. R. Acad. Sci. Paris Sér. I Math.*, 305(19):805–808, 1987.

[15] M. P. Brenner, P. Constantin, L. P. Kadanoff, A. Schenkel, and S. C. Venkataramani. Diffusion, attraction and collapse. *Nonlinearity*, 12(4):1071–1098, 1999.

[16] H. Brezis. *Analyse fonctionnelle*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1983. Théorie et applications. [Theory and applications].

[17] M. A. Buchner. Simplicial structure of the real analytic cut locus. *Proc. Amer. Math. Soc.*, 64(1):118–121, 1977.

[18] L. A. Caffarelli. Boundary regularity of maps with convex potentials. *Comm. Pure Appl. Math.*, 45(9):1141–1151, 1992.

[19] L. A. Caffarelli. The regularity of mappings with a convex potential. *J. Amer. Math. Soc.*, 5(1):99–104, 1992.

[20] L. A. Caffarelli. Boundary regularity of maps with convex potentials. II. *Ann. of Math. (2)*, 144(3):453–496, 1996.

[21] L. A. Caffarelli. The Monge Ampere equation and optimal transportation. In *Recent advances in the theory and applications of mass transport*, volume 353 of *Contemp. Math.*, pages 43–52. Amer. Math. Soc., Providence, RI, 2004.

[22] L. A. Caffarelli and X. Cabré. *Fully nonlinear elliptic equations*, volume 43 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1995.

[23] V. Calvez and J. A. Carrillo. Refined asymptotics for the subcritical Keller-Segel system and related functional inequalities. *Proc. Amer. Math. Soc.*, 140(10):3515–3530, 2012.

[24] V. Calvez, L. Corrias, and M. A. Ebde. Blow-up, concentration phenomenon and global existence for the Keller-Segel model in high dimension. *Comm. Partial Differential Equations*, 37(4):561–584, 2012.

[25] V. Calvez, B. Perthame, and M. Sharifi tabar. Modified Keller-Segel system and critical mass for the log interaction kernel. In *Stochastic analysis and partial differential equations*, volume 429 of *Contemp. Math.*, pages 45–62. Amer. Math. Soc., Providence, RI, 2007.

[26] E. Carlen and M. Loss. Competing symmetries, the logarithmic HLS inequality and Onofri's inequality on $S^n$. *Geom. Funct. Anal.*, 2(1):90–104, 1992.

[27] E. A. Carlen and A. Figalli. Stability for a gns inequality and the log-hls inequality, with application to the critical mass keller-segel equation. *Duke Math. J.*, To appear, 2012.

[28] J. A. Carrillo, R. J. McCann, and C. Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, 19(3):971–1018, 2003.

[29] J. A. Carrillo, R. J. McCann, and C. Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.*, 179(2):217–263, 2006.

[30] M. Castelpietra and L. Rifford. Regularity properties of the distance functions to conjugate and cut loci for viscosity solutions of Hamilton-Jacobi equations and applications in Riemannian geometry. *ESAIM Control Optim. Calc. Var.*, 16(3):695–718, 2010.

[31] S. Childress and J. K. Percus. Nonlinear aspects of chemotaxis. *Math. Biosci.*, 56(3-4):217–237, 1981.

[32] L. Corrias, B. Perthame, and H. Zaag. Global solutions of some chemotaxis and angiogenesis systems in high space dimensions. *Milan J. Math.*, 72:1–28, 2004.

[33] P. Delanoë. Classical solvability in dimension two of the second boundary-value problem associated with the Monge-Ampère operator. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 8(5):443–457, 1991.

[34] P. Delanoë and Y. Ge. Locally nearly spherical surfaces are almost-positively $c$-curved. 09 2010.

[35] P. Delanoë and Y. Ge. Regularity of optimal transport on compact, locally nearly spherical, manifolds. *J. Reine Angew. Math.*, 646:65–115, 2010.

[36] P. Delanoë and G. Loeper. Gradient estimates for potentials of invertible gradient-mappings on the sphere. *Calc. Var. Partial Differential Equations*, 26(3):297–311, 2006.

[37] A. Devys. *Modélisation, analyse mathématique et simulation numérique de problèmes issus de la biologie.* PhD thesis, Lille 1, 2010.

[38] J. I. Diaz, T. Nagai, and J.-M. Rakotoson. Symmetrization techniques on unbounded domains: application to a chemotaxis system on $\mathbf{R}^N$. *J. Differential Equations*, 145(1):156–183, 1998.

[39] J. Dolbeault and B. Perthame. Optimal critical mass in the two-dimensional Keller-Segel model in $\mathfrak{R}^2$. *C. R. Math. Acad. Sci. Paris*, 339(9):611–616, 2004.

[40] J. Dolbeault and C. Schmeiser. The two-dimensional Keller-Segel model after blow-up. *Discrete Contin. Dyn. Syst.*, 25(1):109–121, 2009.

[41] A. Figalli, Y.-H. Kim, and R. J. McCann. Continuity and injectivity of optimal maps for non-negatively cross-curved costs. 11 2009.

[42] A. Figalli, Y.-H. Kim, and R. J. McCann. Regularity of optimal transport maps on multiple products of spheres. 06 2010.

[43] A. Figalli, Y.-H. Kim, and R. J. McCann. Hölder continuity and injectivity of optimal maps. 07 2011.

[44] A. Figalli, Y.-H. Kim, and R. J. McCann. When is multidimensional screening a convex program? *J. Econom. Theory*, 146(2):454–478, 2011.

[45] A. Figalli and G. Loeper. $C^1$ regularity of solutions of the Monge-Ampère equation for optimal transport in dimension two. *Calc. Var. Partial Differential Equations*, 35(4):537–550, 2009.

[46] A. Figalli and L. Rifford. Continuity of optimal transport maps and convexity of injectivity domains on small deformations of $\mathfrak{S}^2$. *Comm. Pure Appl. Math.*, 62(12):1670–1706, 2009.

[47] A. Figalli, L. Rifford, and C. Villani. On the Ma-Trudinger-Wang curvature on surfaces. *Calc. Var. Partial Differential Equations*, 39(3-4):307–332, 2010.

[48] A. Figalli, L. Rifford, and C. Villani. Necessary and sufficient conditions for continuity of optimal transport maps on Riemannian manifolds. *Tohoku Math. J. (2)*, 63(4):855–876, 2011.

[49] A. Figalli, L. Rifford, and C. Villani. Nearly round spheres look convex. *Amer. J. Math.*, 134(1):109–139, 2012.

[50] A. Figalli and C. Villani. An approximation lemma about the cut locus, with applications in optimal transport theory. *Methods Appl. Anal.*, 15(2):149–154, 2008.

[51] F. Filbet. A finite volume scheme for the Patlak-Keller-Segel chemotaxis model. *Numer. Math.*, 104(4):457–488, 2006.

[52] H. Gajewski and K. Zacharias. Global behaviour of a reaction-diffusion system modelling chemotaxis. *Math. Nachr.*, 195:77–114, 1998.

[53] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian geometry*. Universitext. Springer-Verlag, Berlin, third edition, 2004.

[54] Y. Giga and R. V. Kohn. Asymptotically self-similar blow-up of semilinear heat equations. *Comm. Pure Appl. Math.*, 38(3):297–319, 1985.

[55] J. Haškovec and C. Schmeiser. Stochastic particle approximation for measure valued solutions of the 2D Keller-Segel system. *J. Stat. Phys.*, 135(1):133–151, 2009.

[56] J. Haškovec and C. Schmeiser. Convergence of a stochastic particle approximation for measure solutions of the 2D Keller-Segel system. *Comm. Partial Differential Equations*, 36(6):940–960, 2011.

[57] M. A. Herrero, E. Medina, and J. J. L. Velázquez. Finite-time aggregation into a single point in a reaction-diffusion system. *Nonlinearity*, 10(6):1739–1754, 1997.

[58] M. A. Herrero and J. J. L. Velázquez. Chemotactic collapse for the Keller-Segel model. *J. Math. Biol.*, 35(2):177–194, 1996.

[59] M. A. Herrero and J. J. L. Velázquez. Singularity patterns in a chemotaxis model. *Math. Ann.*, 306(3):583–623, 1996.

[60] T. Hillen and K. Painter. Global existence for a parabolic chemotaxis model with prevention of overcrowding. *Adv. in Appl. Math.*, 26(4):280–301, 2001.

[61] T. Hillen and K. J. Painter. A user's guide to PDE models for chemotaxis. *J. Math. Biol.*, 58(1-2):183–217, 2009.

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

[63] D. Horstmann and M. Winkler. Boundedness vs. blow-up in a chemotaxis system. *J. Differential Equations*, 215(1):52–107, 2005.

[64] J.-i. Itoh and M. Tanaka. The Lipschitz continuity of the distance function to the cut locus. *Trans. Amer. Math. Soc.*, 353(1):21–40, 2001.

[65] W. Jäger and S. Luckhaus. On explosions of solutions to a system of partial differential equations modelling chemotaxis. *Trans. Amer. Math. Soc.*, 329(2):819–824, 1992.

[66] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.

[67] L. V. Kantorovich. On a problem of Monge. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 312(Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 11):15–16, 2004.

[68] L. V. Kantorovich. On mass transportation. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 312(Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 11):11–14, 2004.

[69] E. F. Keller and L. Segel. Initiation of slime mold aggregation viewed as an instability. *J. Theor. Biol.*, 26:399–415, 1970.

[70] E. F. Keller and L. Segel. Model for chemotaxis. *J. Theo. Biol.*, 30:225–234, 1971.

[71] E. F. Keller and L. Segel. Traveling bands of chemotactic bacteria: a theoretical analysis. *J. Theor. Biol.*, 30:235–248, 1971.

[72] Y.-H. Kim. Counterexamples to continuity of optimal transport maps on positively curved Riemannian manifolds. *Int. Math. Res. Not. IMRN*, pages Art. ID rnn120, 15, 2008.

[73] Y.-H. Kim and R. J. McCann. On the cost-subdifferentials of cost-convex functions. 06 2007.

[74] Y.-H. Kim and R. J. McCann. Continuity, curvature, and the general covariance of optimal transportation. *J. Eur. Math. Soc. (JEMS)*, 12(4):1009–1040, 2010.

[75] Y.-H. Kim and R. J. McCann. Towards the smoothness of optimal maps on Riemannian submersions and Riemannian products (of round spheres in particular). *J. Reine Angew. Math.*, 664:1–27, 2012.

[76] P. W. Y. Lee and R. J. McCann. The Ma-Trudinger-Wang curvature for natural mechanical actions. *Calc. Var. Partial Differential Equations*, 41(1-2):285–299, 2011.

[77] Y. Li and L. Nirenberg. The distance function to the boundary, Finsler geometry, and the singular set of viscosity solutions of some Hamilton-Jacobi equations. *Comm. Pure Appl. Math.*, 58(1):85–146, 2005.

[78] E. H. Lieb. Sharp constants in the Hardy-Littlewood-Sobolev and related inequalities. *Ann. of Math. (2)*, 118(2):349–374, 1983.

[79] E. H. Lieb and M. Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1997.

[80] J. Liu. Hölder regularity of optimal mappings in optimal transportation. *Calc Var. Partial Differential Equations*, 34:435–451, 2009.

[81] J. Liu, N. S. Trudinger, and X.-J. Wang. Interior $C^{2,\alpha}$ regularity for potential functions in optimal transportation. *Comm. Partial Differential Equations*, 35(1):165–184, 2010.

[82] G. Loeper. On the regularity of solutions of optimal transportation problems. *Acta Math.*, 202(2):241–283, 2009.

[83] G. Loeper and C. Villani. Regularity of optimal transport in curved geometry: the nonfocal case. *Duke Math. J.*, 151(3):431–485, 2010.

[84] S. Luckhaus and Y. Sugiyama. Large time behavior of solutions in super-critical cases to degenerate Keller-Segel systems. *M2AN Math. Model. Numer. Anal.*, 40(3):597–621, 2006.

[85] X.-N. Ma, N. S. Trudinger, and X.-J. Wang. Regularity of potential functions of the optimal transportation problem. *Arch. Ration. Mech. Anal.*, 177(2):151–183, 2005.

[86] M. Mazzucchelli. *Critical Point Theory for Lagrangian Systems*, volume 293 of *Progress in Mathematics*. Birkhäuser, 2012.

[87] R. J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.

[88] R. J. McCann. Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.*, 11(3):589–608, 2001.

[89] F. Merle and H. Zaag. Stability of the blow-up profile for equations of the type $u_t = \Delta u + |u|^{p-1}u$. *Duke Math. J.*, 86(1):143–195, 1997.

[90] G. Monge. *Mémoire sur la théorie des deblais et des remblais*, pages 666–704. Histoire de l'académie royale des sciences de Paris. 1781.

[91] T. Nagai. Blow-up of radially symmetric solutions to a chemotaxis system. *Adv. Math. Sci. Appl.*, 5(2):581–601, 1995.

[92] H. Ohtsuka, T. Senba, and T. Suzuki. Blowup in infinite time in the simplified system of chemotaxis. *Adv. Math. Sci. Appl.*, 17(2):445–472, 2007.

[93] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

[94] K. J. Painter and T. Hillen. Volume-filling and quorum-sensing in models for chemosensitive movement. *Can. Appl. Math. Q.*, 10(4):501–543, 2002.

[95] C. S. Patlak. Random walk with persistence and external bias. *Bull. Math. Biophys.*, 15:311–338, 1953.

[96] B. Perthame. PDE models for chemotactic movements: parabolic, hyperbolic and kinetic. *Appl. Math.*, 49(6):539–564, 2004.

[97] B. Perthame. *Transport equations in biology.* Frontiers in mathematics. Birkhäuser, 2006.

[98] T. Sakai. *Riemannian geometry*, volume 149 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1996. Translated from the 1992 Japanese original by the author.

[99] T. Sei. A jacobian inequality for gradient maps on the sphere and its application to directional statistics. 06 2009.

[100] T. Senba. Blowup of radial solutions to a parabolic-elliptic system related to chemotaxis. *Bull. Kyushu Inst. Technol. Pure Appl. Math.*, (58):1–11, 2011.

[101] T. Senba and T. Suzuki. Chemotactic collapse of radial solutions to Jäger-Luckhaus system. *Adv. Math. Sci. Appl.*, 14(1):241–250, 2004.

[102] K.-T. Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1):65–131, 2006.

[103] K.-T. Sturm. On the geometry of metric measure spaces. II. *Acta Math.*, 196(1):133–177, 2006.

[104] T. Suzuki. *Free energy and self-interacting particles.* Progress in Nonlinear Differential Equations and their Applications, 62. Birkhäuser Boston Inc., Boston, MA, 2005.

[105] N. S. Trudinger and X.-J. Wang. On the second boundary value problem for Monge-Ampère type equations and optimal transportation. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 8(1):143–174, 2009.

[106] J. Urbas. On the second boundary value problem for equations of Monge-Ampère type. *J. Reine Angew. Math.*, 487:115–124, 1997.

[107] J. Urbas. The second boundary value problem for a class of Hessian equations. *Comm. Partial Differential Equations*, 26(5-6):859–882, 2001.

[108] J. J. L. Velázquez. Stability of some mechanisms of chemotactic aggregation. *SIAM J. Appl. Math.*, 62(5):1581–1633 (electronic), 2002.

[109] J. J. L. Velázquez. Point dynamics in a singular limit of the Keller-Segel model. II. Formation of the concentration regions. *SIAM J. Appl. Math.*, 64(4):1224–1248 (electronic), 2004.

[110] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 2009. Old and new.