

Ecole Normale Supérieure  
Magistère de Mathématiques

— o —

Autour de l'équation de Boltzmann

Mémoire

de

François Bolley

soutenu le 26 juin 2001 devant le jury composé de  
B. Perthame, Professeur à l'Ecole Normale Supérieure  
C. Villani, Professeur à l'Ecole Normale Supérieure de Lyon  
S. Dumitrescu, Caïman à l'Ecole Normale Supérieure



Ce dossier présente les études et travaux que j'ai faits dans le cadre du Magistère de Mathématiques de l'École Normale Supérieure. Ces travaux m'ont conduit principalement à l'étude de certains aspects de l'équation de Boltzmann, ainsi qu'à une ouverture sur des applications des transports de mesures qui devraient faire l'objet de recherches ultérieures.

Je remercie les Professeurs B. Perthame et C. Villani de m'avoir proposé ces sujets d'étude et de m'avoir aidé tout au long de la réalisation de ces travaux.

### Table des matières

. Curriculum vitæ Magistère	p. 4
. Introduction aux équations de Boltzmann	p. 5
. Méthode des moments de Levermore (Mémoire de stage de Cursus mixte d'après des travaux de C.D. Levermore et de M. Junk)	p. 13
. Théorème de Tanaka (Mémoire de stage de D.E.A. d'après des travaux de C. Villani et de H. Tanaka)	p. 29
. Projets de recherche	p. 73

# Curriculum vitæ Magistère

— o —

## Formation dans le cadre du Magistère

Première année 1998–1999: Licence de Physique, Licence et Maîtrise de Mathématiques dans le cadre du Cours mixte Mathématiques et Physique.

Cours validés: *Analyse fonctionnelle* (F. Béthuel et Y. Ekeland)

*Analyse complexe* (J. Faraut)

*Intégration* (F. Comets)

*Equations aux dérivées partielles* (H. Berestycki et B. Perthame)

*Mouvement brownien et fonctions harmoniques* (J-F. Le Gall)

*Mécanique quantique* (S. Haroche)

*Physique statistique* (B. Roulet)

*Relativité et électromagnétisme* (J-M. Raymond)

Mémoire: *Méthode des moments de Levermore pour l'Equation de Boltzmann*, sous la direction de B. Perthame.

Deuxième année 1999-2000: Agrégation de Mathématiques.

Troisième année 2000-2001: DEA d'Analyse numérique (Equations aux dérivées partielles) de l'Université de Paris VI.

Cours validés: *Théorie des équations d'évolution* (J-Y. Chemin)

*Systèmes hyperboliques* (G. Allaire)

*Calcul des variations et équations elliptiques* (F. Béthuel)

*Modélisation en météorologie* (C. Basdevant)

*Introduction aux équations cinétiques* (B. Perthame)

*Méthodes numériques pour les équations cinétiques* (S. Cordier)

Mémoire: *Transport de mesures et Equation de Boltzmann: le théorème de Tanaka*, sous la direction de C. Villani.

## Séminaires et séjours

Au cours de l'année 2000-2001, j'ai assisté d'une part à des exposés sur divers thèmes dans le cadre du séminaire EDP du DMI, et d'autre part à des exposés plus spécialisés en liaison avec mon mémoire de DEA; ainsi j'ai suivi deux mini-cours donnés par les Professeurs Desvillettes et Villani lors des Rencontres Mathématiques de l'ENS Lyon sur les équations cinétiques (mars 2001) et assisté aux exposés donnés à Paris-Dauphine par le Professeur Gangbo (Atlanta) sur la distance de Wasserstein (juin 2001).

Pendant un séjour d'un mois à la Scuola Normale Superiore (Pise, avril 2001), j'ai commencé à travailler mon mémoire de DEA et rencontré le Professeur Ambrosio et deux de ses élèves avec qui j'ai pu discuter de questions relatives à mon travail. J'ai également assisté chaque semaine au Colloquim De Giorgi portant sur les EDP et Calcul des Variations.

# Introduction aux Equations de Boltzmann

— o —

## 1 - Origine des équations.

On s'intéresse à la modélisation mathématique d'un gaz formé d'un très grand nombre  $n$  de particules identiques en interaction. En négligeant les effets de bord, on suppose que le gaz occupe l'espace  $\mathbb{R}^N$ , le cas  $N = 3$  correspondant au cadre physique le plus naturel.

Si l'on suppose que la position  $x_i \in \mathbb{R}^N$  et la vitesse  $v_i \in \mathbb{R}^N$  suffisent à décrire la particule  $p_i$  pour  $i = 1, \dots, n$ , l'état du gaz peut alors à chaque instant  $t \geq 0$  être décrit de façon *microscopique* par un point  $((x_1(t), v_1(t)), \dots, (x_n(t), v_n(t)))$  de représentation dans l'espace des phases  $(\mathbb{R}^N \times \mathbb{R}^N)^n$ . On peut écrire les lois de Newton sous la forme d'un (très grand) système de  $2n$  équations du premier ordre faisant intervenir les positions  $x_i(t)$  et les vitesses  $v_i(t)$  de toutes les particules, ainsi que les forces d'interaction et extérieures agissant sur ces particules. Un exemple typique est le modèle des *sphères dures* dans lequel les particules rebondissent les unes sur les autres comme les boules d'un billard; le flot associé aux équations de Newton est alors bien défini pour presque toute condition initiale en  $t = 0$  et détermine ainsi l'évolution des particules au cours du temps.

Cependant cette description microscopique présente beaucoup d'inconvénients: on ne connaît pas les données initiales à l'instant  $t = 0$ , les calculs sont difficiles car ils font intervenir un très grand nombre de paramètres, la connaissance des positions  $x_i(t)$  et des vitesses  $v_i(t)$  ne donne pas les grandeurs *macroscopiques* physiques intéressantes que sont les densité  $\rho(t, x)$ , vitesse  $u(t, x)$  et température  $T(t, x)$  moyennes du gaz à l'instant  $t \geq 0$  et au point  $x \in \mathbb{R}^N$ .

C'est pourquoi on préfère remplacer cette description microscopique exacte de chaque particule dans un espace des phases très grand par une description dans un espace des phases restreint, et dans laquelle une particule donnée n'aura plus une position et une vitesse données, mais différentes probabilités d'avoir différentes positions et vitesses.

Ainsi en considérant que le gaz se rapproche d'un continuum puisqu'il est formé d'un très grand nombre  $n$  de particules, l'état du gaz peut être décrit dans un *modèle cinétique* par une densité de probabilité  $f(t, x, v)$  représentant au temps  $t \geq 0$  la densité de présence de l'ensemble des particules dans l'espace des phases  $\mathbb{R}^N \times \mathbb{R}^N$ . A chaque instant  $t \geq 0$ ,  $f(t, \dots)$  est une mesure de probabilité dans  $\mathbb{R}^N \times \mathbb{R}^N$ ,  $x \in \mathbb{R}^N$  représente la position,  $v \in \mathbb{R}^N$  représente la vitesse et  $f(t, x, v) dx dv$  représente la quantité de particules dans l'élément de volume  $dx dv$  centré en  $(x, v)$  dans  $\mathbb{R}^N \times \mathbb{R}^N$ .

Cette représentation est intéressante car elle permet en particulier d'atteindre des quantités *macroscopiques observables* telles que la densité locale  $\rho$ , la vitesse macroscopique locale  $u$  et la température  $T$  du gaz à l'instant  $t \geq 0$  et au point  $x \in \mathbb{R}^N$  par les formules

$$\begin{aligned}\rho(t, x) &= \int_{\mathbb{R}^N} f(t, x, v) dv \\ \rho u(t, x) &= \int_{\mathbb{R}^N} v f(t, x, v) dv\end{aligned}$$

$$\rho T(t, x) = \frac{1}{N} \int_{\mathbb{R}^N} |v - u(t, x)|^2 f(t, x, v) dv.$$

## 2 - Description des interactions.

En l'absence de force, les particules se déplacent en ligne droite à vitesse constante d'après le principe de Newton, de sorte que  $f$  est solution de l'équation de transport libre

$$\partial_t f + (v, \nabla_x f) = 0$$

où  $\nabla_x$  est l'opérateur de gradient par rapport à la variable  $x$  et  $(\cdot, \cdot)$  le produit scalaire dans l'espace euclidien  $\mathbb{R}^N$ .

Par contre quand il y a une force agissant sur les particules, leurs trajectoires ne sont plus rectilignes et cette équation doit être corrigée. Selon les forces considérées, on aboutit à des équations cinétiques variées, les plus connues étant celles de Boltzmann et celles de Vlasov.

En 1872, L. Boltzmann développe les travaux de J.C. Maxwell (1866) et établit une équation permettant de décrire l'évolution d'un gaz raréfié en l'absence de forces extérieures. Plus précisément on fait les hypothèses suivantes:

1. les particules interagissent via des *collisions binaires*: une collision est le résultat de l'interaction microscopique de deux particules qui passent très près l'une de l'autre, ce qui entraîne une forte déviation de leurs trajectoires en un temps très court. On suppose de plus que ce processus ne met en jeu que deux particules à la fois, c'est-à-dire que l'hypothèse de raréfaction permet de négliger les collisions faisant intervenir plus de deux particules.

2. les collisions sont *instantanées* en temps et *localisées* en espace: elles se déroulent sur des échelles de temps et d'espace très inférieures aux échelles typiques de description.

3. les collisions sont *élastiques*: la quantité de mouvement et l'énergie cinétique sont préservées dans le processus de collision. Par conséquent, si  $v$  et  $v_*$  désignent les vitesses de deux particules entrant en collision et si  $v'$  et  $v'_*$  désignent leurs vitesses juste après la collision, alors on a

$$\begin{aligned} v' + v'_* &= v + v_*, \\ |v'|^2 + |v'_*|^2 &= |v|^2 + |v_*|^2. \end{aligned}$$

Ces relations sont résolues sous la forme

$$\begin{aligned} v' &= \frac{v + v_*}{2} + \frac{|v - v_*|}{2} \sigma, \\ v'_* &= \frac{v + v_*}{2} - \frac{|v - v_*|}{2} \sigma, \end{aligned}$$

où le paramètre  $\sigma$  décrit la sphère unité  $S^{N-1}$ .

L'angle  $\theta \in [0, \pi]$  de déviation défini par

$$\cos \theta = (k, \sigma)$$

où  $k$  est le vecteur unitaire

$$k = \frac{v - v_*}{|v - v_*|}$$

est l'angle entre les vitesses pré et post-collisionnelles.

4. les collisions sont *microréversibles*: d'un point de vue probabiliste, la probabilité que les vitesses  $(v', v'_*)$  soient changées en  $(v, v_*)$  dans une collision est égale à la probabilité que les vitesses  $(v, v_*)$  soient changées en  $(v', v'_*)$ .

5. les collisions satisfont l'hypothèse de *chaos moléculaire*: les vitesses de deux particules qui vont entrer en collision ne sont pas corrélées.

Sous ces cinq hypothèses, L. Boltzmann montre alors en 1872 que l'équation de transport libre doit être modifiée sous la forme

$$\partial_t f + (v, \nabla_x f) = Q(f, f)$$

dans laquelle l'opérateur  $Q$  modélise les interactions de collisions entre les particules.

Ce terme de collision est un opérateur quadratique agissant seulement sur la dépendance en  $v$  de la forme

$$Q(f, f) = \int_{\mathbb{R}^N} dv_* \int_{S^{N-1}} B(v - v_*, \sigma) [f' f'_* - f f_*] d\sigma$$

où  $f_*$ ,  $f'$ ,  $f'_*$  désignent respectivement  $f(t, x, v_*)$ ,  $f(t, x, v')$ ,  $f(t, x, v'_*)$  et où  $B$  est une fonction positive appelée la *section efficace*; cette section efficace tient compte du fait que toutes les valeurs de  $\sigma \in S^{N-1}$  n'interviennent pas de façon identique dans une collision. De plus pour des raisons de symétrie, cette section efficace dépend seulement du module  $|v - v_*|$  de la vitesse relative et du cosinus de l'angle  $\theta$  de déviation défini précédemment, et  $Q$  peut s'écrire sous la forme

$$Q(f, f) = \int_{\mathbb{R}^N} dv_* \int_{S^{N-1}} b(|v - v_*|, (k, \sigma)) [f' f'_* - f f_*] d\sigma.$$

La section efficace peut être explicitée dans certains cas. Ainsi dans le modèle étudié par Boltzmann, les particules sont assimilées à des boules de billard (cas des sphères dures) et la section efficace est de la forme

$$b(|v - v_*|, \cos \theta) = c_0 |v - v_*|$$

où  $c_0$  est une constante, alors que si la force d'interaction entre deux particules dérive d'un potentiel de la forme générale  $\phi(r) = r^{1-s}$  où  $s \geq 2$ , la section efficace peut être explicitée sous la forme

$$b(|v - v_*|, \cos \theta) = |v - v_*|^\gamma \tilde{b}(\cos \theta)$$

avec  $\gamma = \frac{s - (2N - 1)}{s - 1}$ . Le cas de l'interaction coulombienne correspond à  $s = 2$  et  $N = 3$ , et la section efficace est alors donnée par la formule de Rutherford

$$b(|v - v_*|, \cos \theta) = |v - v_*|^{-3} \sin^{-4} \frac{\theta}{2}.$$

Si on suppose que la section efficace ne dépend que du cosinus de l'angle de déviation, le terme quadratique de collision s'écrit alors

$$Q(f, f) = \int_{\mathbb{R}^N} dv_* \int_{S^{N-1}} b\left(\left(\frac{v - v_*}{|v - v_*|}, \sigma\right)\right) [f' f'_* - f f_*] d\sigma$$

où  $b$  est définie sur  $[-1, +1]$ . Dans ce cas les particules sont dites *Maxwelliennes*. Cette situation modélise par exemple des particules qui interagissent suivant des forces dérivant d'un potentiel du type  $\phi(r) = r^{1-2N}$ .

Pour terminer, il est important de noter que le modèle de Boltzmann décrit précédemment est obtenu comme résultat de l'hypothèse d'interaction localisée: il ne tient pas compte en particulier de possibles interactions lointaines. De façon générale, si une force macroscopique  $F$  est aussi présente, l'équation de Boltzmann doit être corrigée sous la forme

$$\partial_t f + (v, \nabla_x f) + (F, \nabla_v f) = Q(f, f).$$

En particulier les interactions lointaines peuvent résulter d'un champ moyen de forces, typiquement de la forme

$$F(t, x) = -\nabla \Phi(t, x)$$

avec  $\Phi = \rho * \phi$  (avec une convolution en  $x$ ) où  $\phi$  est le potentiel d'interaction et  $\rho$  la densité locale. Ce modèle où l'interaction est décrite par une telle force couplée est appelé le *modèle de Vlasov*. Pour une force d'interaction avec un potentiel  $\phi(r) = r^{1-s}$ , le terme de Boltzmann domine le terme de champ moyen si  $s > 2$ , alors qu'il est négligeable devant le terme de champ moyen si  $s < 2$ . Par contre si  $s = 2$ , les deux termes sont d'importance équivalente et il faut tenir compte des deux contributions.

### 3 - Lois de conservation et entropie.

En intégrant l'équation de Boltzmann avec un poids  $\varphi(t, x, v) dx dv$  et en utilisant la forme particulière du terme de collision  $Q$ , on obtient (au moins formellement) deux propriétés importantes des solutions  $f$  de l'équation de Boltzmann:

1. *lois de conservation globale*: pour tout  $t \geq 0$

$$\frac{d}{dt} \int_{\mathbb{R}^N \times \mathbb{R}^N} \varphi(v) f(t, x, v) dx dv = 0$$

pour  $\varphi(v) = 1$ ,  $v_i$  ( $i = 1, \dots, N$ ) et  $|v|^2$ , ce qui signifie que la masse, la quantité de mouvement et l'énergie cinétique globales sont conservées au cours du temps.

2. *théorème H de Boltzmann*: pour tout  $t \geq 0$

$$\frac{d}{dt} \int_{\mathbb{R}^N \times \mathbb{R}^N} f(t, x, v) \log f(t, x, v) dx dv \leq 0,$$

ce qui signifie que l'entropie

$$H(f) = \int f \log f dx dv$$

est décroissante au cours du temps. Ce théorème de Boltzmann correspond à la seconde loi (de Clausius) de la thermodynamique qui établit que l'entropie physique  $-H(f)$  est croissante au cours du temps. En particulier il traduit une propriété d'irréversibilité dans le modèle de Boltzmann.



## 4 - Equilibre thermodynamique local. Méthode des moments.

Dans le cas d'un gaz en équilibre thermodynamique, la distribution  $f(t, \cdot, \cdot)$  est donnée par une Maxwellienne locale de la forme

$$f(t, x, v) = \frac{\rho(t, x)}{(2\pi T(t, x))^{\frac{N}{2}}} \exp\left(-\frac{|v - u(t, x)|^2}{2T(t, x)}\right)$$

à partir de quantités macroscopiques  $\rho = \rho(t, x)$ ,  $u = u(t, x)$  et  $T = T(t, x)$  représentant respectivement la densité, la vitesse moyenne et la température à l'instant  $t \geq 0$  et au point  $x \in \mathbb{R}^N$ , dont l'évolution est gouvernée par le système des trois équations d'Euler compressibles traditionnellement écrites sous la forme

$$\begin{cases} \partial_t \rho + (\nabla_x, \rho u) = 0 \\ \partial_t(\rho u) + (\nabla_x, \rho u \otimes u + p) = 0 \\ \partial_t E + (\nabla_x, E u + p u) = 0 \end{cases}$$

où l'énergie totale  $E$  est donnée par

$$E = \rho|u|^2 + N\rho T$$

et la pression  $p$  est donnée par la loi des gaz parfaits

$$p = \rho T.$$

Ces équations d'Euler peuvent être obtenues directement à partir de considérations de Mécanique des Fluides, mais aussi par certains passages à la limite de l'équation de Boltzmann. Elles peuvent alors être retrouvées à partir de l'équation de Boltzmann en prenant les trois moments par rapport à 1,  $v$  et  $|v|^2$ .

Le lemme de Gibbs assure que cette Maxwellienne est la distribution qui minimise l'entropie du système sous la contrainte que ses trois moments par rapport à 1,  $v$  et  $|v|^2$  soient prescrits, les moments correspondant étant ici les combinaisons  $\rho$ ,  $\rho u$  et  $E$ .

Près d'un équilibre, une telle Maxwellienne peut être considérée comme une première approximation de la distribution  $f$ . De façon plus générale, les déviations par rapport à un équilibre local peuvent être précisées par des méthodes de développement introduites par Hilbert et par Chapman-Enskog sous la forme

$$f_\varepsilon = M \left( 1 + \varepsilon f^{(1)} + \varepsilon^2 f^{(2)} + \dots \right)$$

où le terme principal  $M$  est une Maxwellienne.

L'approximation de Navier-Stokes consiste à ne conserver que la correction à l'ordre 1. La correction  $f^{(1)}$  en chaque  $(t, x)$  est alors exprimée en fonction des variables  $\rho$ ,  $u$  et  $T$ , l'évolution de ces quantités macroscopiques étant gouvernée par les trois équations de Navier-Stokes compressibles qui contiennent des termes correctifs par rapport aux trois équations d'Euler.

Cependant ces équations d'Euler - Navier-Stokes ne sont valables que pour de petites déviations autour d'un équilibre et ne le sont plus dans un régime de transition où les grandeurs macroscopiques deviennent comparables au libre parcours moyen, c'est-à-dire lorsque  $\varepsilon$  est de l'ordre de 1.

Pour décrire des gaz dont l'état est alors plus éloigné d'un équilibre, C.D. Levermore a proposé une nouvelle description en utilisant une méthode qui généralise la propriété fondamentale de Gibbs: elle consiste à chercher à décrire la distribution  $f$  en l'approchant à l'aide de minima d'entropie sous contraintes de moments, mais en faisant intervenir un plus grand nombre de moments afin d'obtenir une structure plus détaillée de la distribution, c'est-à-dire en introduisant de nouvelles grandeurs macroscopiques en plus de  $\rho, u, \theta$ .

Cette étude est l'objet du mémoire présenté p.12.

## 5 - Problème de convergence vers l'équilibre.

Comme le minimum de l'entropie est réalisé par une Maxwellienne, il est naturel de penser que l'effet des collisions va amener la distribution  $f(t, x, v)$  à se rapprocher de plus en plus d'une Maxwellienne quand le temps grandit.

En dimension  $N = 3$ , pour un gaz spatialement homogène formé de particules Maxwelliennes, la fonction de distribution  $f = f(t, v)$ , qui est indépendante de  $x$  dans ce cas, solution de l'équation de Boltzmann

$$\partial_t f = Q(f, f)$$

et correspondant à une donnée initiale  $f_0 = f_0(v)$  de densité  $\rho \geq 0$ , de vitesse moyenne  $u \in \mathbb{R}^3$  et de température  $T \geq 0$ , converge, quand le temps  $t$  tend vers  $+\infty$  et dans un sens que l'on va préciser, vers la solution Maxwellienne d'équilibre  $M$  correspondant à ces données, définie par

$$M(v) = \frac{\rho}{(2\pi T)^{\frac{3}{2}}} e^{-\frac{|v-u|^2}{2T}}.$$

Si tous les moments de  $f_0$  sont finis, alors  $f(t, \cdot)$  converge vers  $M$  au sens faible\* des mesures dans  $\mathbb{R}^3$  quand  $t$  tend vers  $+\infty$ .

En fait on peut préciser cette convergence à l'aide d'une distance sur l'espace des probabilités dans  $\mathbb{R}^3$ , appelée distance de Wasserstein. C'est l'objet du théorème de Tanaka qui est le résultat principal du mémoire présenté p.28.

## 6 - Transport de mesures. Distance de Wasserstein.

Etant données deux mesures de probabilité  $f$  et  $g$  dans  $\mathbb{R}^N$  dont les moments d'ordre 2 sont finis, on note  $\mathcal{T}_2(f, g)$  le *coût de transport optimal* entre  $f$  et  $g$  associé à la fonction de coût  $|x - y|^2$  dans  $\mathbb{R}^N$ , c'est-à-dire

$$\mathcal{T}_2(f, g) = \inf_{\pi \in \Pi(f, g)} \int_{\mathbb{R}^N \times \mathbb{R}^N} |x - y|^2 d\pi(x, y)$$

où  $\Pi(f, g)$  désigne l'ensemble des mesures de probabilité sur  $\mathbb{R}^N \times \mathbb{R}^N$  de marginales  $f$  et  $g$ .

Cette borne inférieure définissant  $\mathcal{T}_2$  est atteinte. Plus précisément  $\pi \in \Pi(f, g)$  réalise le minimum si et seulement s'il existe une fonction  $\varphi$  semicontinue inférieurement et convexe sur  $\mathbb{R}^N$  telle que le support de  $\pi$  soit contenu dans le graphe de  $\partial\varphi$ , ce qui signifie que  $y \in \partial\varphi(x)$  pour presque tout  $(x, y)$  pour la mesure  $d\pi$ .

Plus particulièrement si  $g$  est absolument continue par rapport à la mesure de Lebesgue, le minimum  $\pi$  est unique et est donné par

$$d\pi(x, y) = \delta(\{x = \nabla\varphi(y)\}) dg(y)$$

où  $\nabla\varphi$  est l'unique gradient d'une fonction convexe  $\varphi$  envoyant  $g$  sur  $f$ , c'est-à-dire telle que pour toute fonction  $\psi$  continue bornée

$$\int \psi(y) df(y) = \int \psi(\nabla\varphi(x)) dg(x).$$

Ce coût optimal peut être représenté à l'aide de variables aléatoires. Si  $f$  et  $g$  sont deux mesures de probabilité dans  $\mathbb{R}^N$  dont les moments d'ordre 2 sont finis, alors

$$\mathcal{T}_2(f, g) = \inf_{X, Y} E|X - Y|^2$$

où la borne inférieure est prise sur l'ensemble des variables aléatoires  $X$  et  $Y$  de loi respective  $f$  et  $g$ , et il existe un couple  $(X, Y)$  de telles variables aléatoires tel que

$$\mathcal{T}_2(f, g) = E|X - Y|^2.$$

Cette application de coût optimal est convexe et est continue dans le sens suivant: soient  $f$  et  $g$ ,  $f_n$  et  $g_n$  pour  $n \geq 0$  des mesures de probabilité dans  $\mathbb{R}^N$  dont les moments d'ordre 2 sont finis. Si les suites  $(f_n)_n$  et  $(g_n)_n$  convergent au sens faible\* vers  $f$  et  $g$  respectivement, alors

$$\mathcal{T}_2(f, g) \leq \liminf_{n \rightarrow +\infty} \mathcal{T}_2(f_n, g_n),$$

et si de plus les suites des moments d'ordre 2 des  $f_n$  et des  $g_n$  convergent respectivement vers les moments d'ordre 2 de  $f$  et de  $g$ , alors

$$\mathcal{T}_2(f, g) = \lim_{n \rightarrow +\infty} \mathcal{T}_2(f_n, g_n).$$

Enfin l'application

$$W := \mathcal{T}_2^{\frac{1}{2}}$$

définit une métrique sur l'ensemble des mesures de probabilité dans  $\mathbb{R}^N$  dont les moments d'ordre 2 sont finis, appelée *distance de Wasserstein*.

## 7 - Le théorème de Tanaka.

On suppose que le gaz est spatialement homogène dans  $\mathbb{R}^3$  et formé de particules Maxwelliennes, et que la section efficace  $b$  vérifie la condition

$$\int_0^\pi b(\cos \theta)(1 - \cos \theta) \sin^{N-2} \theta d\theta < +\infty.$$

Si  $f_0$  est une mesure de probabilité vérifiant les conditions de normalisation

$$\int_{\mathbb{R}^3} f_0 dv = 1 \quad \int_{\mathbb{R}^3} f_0 v dv = 0 \quad \int_{\mathbb{R}^3} f_0 |v|^2 dv = 3,$$

alors il existe alors une unique solution faible de l'équation de Boltzmann

$$\partial_t f = \int_{\mathbb{R}^N} dv_* \int_{S^2} b\left(\frac{v-v_*}{|v-v_*|}, \sigma\right) [f' f'_* - f f_*] d\sigma$$

avec donnée initiale  $f_0$ .

Si tous les moments de  $f_0$  sont finis, on a vu que la solution  $f(t)$  converge vers  $M$  au sens faible\* des mesures quand  $t$  tend vers  $+\infty$ , où  $M$  est la Maxwellienne ayant les mêmes trois premiers moments que  $f_0$ , c'est-à-dire définie par

$$M(v) = \frac{1}{(2\pi)^{\frac{3}{2}}} e^{-\frac{|v|^2}{2}}.$$

Comme de plus les moments d'ordre 2 de  $f(t)$  sont constants d'après les lois de conservation globale la continuité de la distance de Wasserstein assure que

$$\lim_{t \rightarrow +\infty} W(f(t), M) = 0.$$

Le théorème de H. Tanaka (1978) donne le résultat plus précis suivant: en supposant seulement que les trois premiers moments des données initiales sont finis, il assure une propriété de contraction de la distance de Wasserstein le long du flot associé à l'équation de Boltzmann et précise que la distance de Wasserstein entre deux solutions  $f(t, \cdot)$  et  $g(t, \cdot)$  de l'équation de Boltzmann est une fonction décroissante de  $t$  qui tend vers 0 quand  $t$  tend vers  $+\infty$ .

Plus précisément si  $f_0$  et  $g_0$  sont deux mesures de probabilité vérifiant les conditions de moyenne nulle et de température unité et si  $f(t)$  et  $g(t)$  sont les solutions associées de l'équation de Boltzmann, alors la fonction  $W(f(t), g(t))$  est continue décroissante de  $[0, +\infty[$  sur  $]0, W(f_0, g_0)[$ .

## 8 - Autres exemples.

La distance de Wasserstein d'ordre 2 utilisée pour l'équation de Boltzmann dans le théorème de Tanaka est associée à la fonction de coût quadratique  $|x - y|^2$  dans  $\mathbb{R}^3$ . Elle peut être associée à d'autres fonctions dans d'autres situations.

Ainsi par exemple, l'ensemble  $\mathcal{M}$  des mesures dans  $\mathbb{R}^3$  de masse totale finie est un espace complet pour la distance de Wasserstein  $W_1$  d'ordre 1 associée à la fonction de coût  $|x - y|$ , et une méthode de point fixe permet de montrer que l'équation de Vlasov

$$\partial_t f + (v, \nabla_x f) + (F, \nabla_v f) = 0$$

où

$$F(t, x) = \int \nabla U(x - x') f(t, x', v') dx' dv'$$

avec un potentiel  $U$  dont le gradient est borné et lipschitzien, a une solution et une seule dans des espaces associés à  $\mathcal{M}$ . De plus il y a aussi un contrôle de cette distance de Wasserstein le long du flot associé à l'équation de Vlasov: ainsi si  $f(t)$  et  $g(t)$  sont les solutions de l'équation de Vlasov associées à des conditions initiales  $f_0$  et  $g_0$ , alors pour tout  $t$  on a

$$W_1(f(t), g(t)) \leq e^{ct} W_1(f_0, g_0)$$

où  $c$  est une constante ne dépendant que de  $U$ .

**Ecole Normale Supérieure**  
**Cursus mixte Mathématiques-Physique**

Année 1998-99

— o —

**Méthode des moments de Levermore**

— o —

Mémoire de stage dirigé par Benoît Perthame,  
Professeur à l'Ecole Normale Supérieure

## Résumé.

La description statistique d'un gaz raréfié constitué de particules ponctuelles repérées à l'instant  $t$  par leur position  $x \in \mathbb{R}^N$  et leur vitesse  $v \in \mathbb{R}^N$  est faite à l'aide d'une fonction de distribution  $f = f(t, x, v)$  dont l'évolution est régie par l'équation de Boltzmann qui dans le cas modèle hypothétique d'un gaz unidimensionnel ( $N = 1$ ) est de la forme

$$\partial_t f + v \partial_x f = Q(f) \quad t \in \mathbb{R}_+, x, v \in \mathbb{R}.$$

Le second membre  $Q(f)$ , appelé terme de collision, représente un bilan statistique des collisions entre les particules. L'opérateur  $Q$  possède diverses propriétés importantes dont celle de conservation de la masse, du moment et de l'énergie et celle, donnée par le théorème H de Boltzmann, de dissipation de l'entropie.

Cette équation de Boltzmann (en  $(t, x, v)$ ) est très difficile à étudier tant du point de vue théorique que du point de vue numérique. C'est pourquoi, lorsque cela est possible, on préfère passer à des modèles dits macroscopiques (en  $(t, x)$ ).

Dans le cas d'un gaz en équilibre thermodynamique, la distribution  $f$  est donnée par une Maxwellienne de la forme

$$f(t, x, v) = \frac{\rho}{(2\pi T)^{1/2}} \exp\left(-\frac{|v - u|^2}{2T}\right) \quad t \in \mathbb{R}_+, x, v \in \mathbb{R}$$

à partir de quantités macroscopiques  $\rho = \rho(t, x)$ ,  $u = u(t, x)$  et  $T = T(t, x)$  représentant respectivement la densité, la vitesse et la température à l'instant  $t$  et au point  $x$ , dont l'évolution est gouvernée par le système des trois équations d'Euler compressibles traditionnellement écrites sous la forme

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0 \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p) = 0 \\ \partial_t E + \partial_x(E u + p u) = 0 \end{cases} \quad t \in \mathbb{R}_+, x, v \in \mathbb{R}$$

où l'énergie totale  $E$  et la pression  $p$  sont reliées à  $u$ ,  $\rho$  et  $T$ .

Le point important est que cette Maxwellienne minimise l'entropie du système sous la contrainte que les trois moments de  $f$  par rapport à  $1$ ,  $v$  et  $v^2$  sont prescrits, les moments correspondant étant ici les combinaisons  $\rho$ ,  $\rho u$  et  $\rho u^2 + \rho T$ .

Ces équations d'Euler peuvent être obtenues directement à partir de considérations de Mécanique des Fluides, mais aussi par certains passages à la limite de l'équation de Boltzmann. Elles peuvent alors être retrouvées à partir de l'équation de Boltzmann en prenant les trois moments par rapport à  $1$ ,  $v$  et  $v^2$ .

L'idée fondamentale de C.D. Levermore [L] est que pour un gaz dont l'état n'est pas un équilibre, on peut encore chercher à décrire la distribution  $f$  en l'approchant à l'aide de minima d'entropie sous contrainte de moments, mais en faisant intervenir un plus grand nombre de moments afin d'obtenir une structure plus détaillée de la distribution, c'est-à-dire en introduisant de nouvelles grandeurs macroscopiques en plus de  $\rho$ ,  $u$ ,  $T$ .

En prenant les moments successifs de l'équation de Boltzmann par rapport à la suite  $(1, v, v^2, \dots)$ , on obtient un système infini d'équations hydrodynamiques couplées de proche en proche, ce système étant formellement équivalent à l'équation de Boltzmann. Pratiquement, on ne garde qu'un nombre fini  $M$  d'équations en prenant les moments par rapport au vecteur  $m(v) = (1, v, \dots, v^{M-1})$ , mais le système obtenu

$$\partial_t \int m f dv + \partial_x \int v m f dv = \int m Q(f) dv \quad t \in \mathbb{R}_+, x \in \mathbb{R}.$$

est indéterminé et on cherche à fermer ce système en exprimant les vecteurs  $\int m f dv$ ,  $\int v m f dv$  et  $\int m Q(f)$  à l'aide d'une fonction de  $M$  variables, traditionnellement le vecteur moment  $\rho = \rho(t, x)$  lui-même défini par

$$\rho = \int m f dv.$$

Mais cette fermeture doit être faite de telle façon que le système correspondant en  $\rho$ , qu'on écrira alors sous la forme

$$\partial_t \rho + \partial_x G(\rho) = r(\rho) \quad t \in \mathbb{R}_+, x \in \mathbb{R}$$

soit d'une part bien posé mathématiquement et d'autre part fournisse des moments  $\rho$  qui soient réalisables par une densité non négative admissible physiquement.

Pour y parvenir, il faut utiliser des approximations. Le procédé de fermeture choisi par C.D. Levermore se veut valable en toutes circonstances: il est basé sur le principe de minimisation de l'entropie qui consiste à chercher la distribution  $f$  qui minimise l'entropie sous la contrainte que ses moments relativement au vecteur  $m$  soient fixés. (Notons qu'il existe des fermetures utilisées en Physique (introduites par Grad, Müller et Ruggeri) s'appuyant sur d'autres principes). Cependant de façon générale, il existe des moments  $\rho$  qui sont bien définis mais pour lesquels le problème de minimisation n'a pas de solution. Plus précisément, si le minimum est atteint, nécessairement les moments de contrainte  $\rho$  sont ceux d'une densité exponentielle de la forme

$$\mathcal{M}(\alpha) = \exp(\alpha \cdot m)$$

où  $\alpha = (\alpha_0, \dots, \alpha_{M-1})$  et  $\alpha_i = \alpha_i(t, x)$  est une fonction à valeurs réelles vérifiant certaines propriétés, et où  $\alpha \cdot m$  est le produit scalaire dans  $\mathbb{R}^M$  des vecteurs  $\alpha$  et  $m$ . Et inversement si les moments de contrainte  $\rho$  sont de ce type, alors le minimum de l'entropie sous cette contrainte est atteint, et est atteint par une densité de ce type exponentiel.

La fermeture choisie consiste alors à considérer des fonctions  $f$  de la forme exponentielle  $\mathcal{M}(\alpha)$ . Ces fonctions permettent d'approcher les solutions de l'équation de Boltzmann et l'ordre  $M$  sera choisi d'autant plus grand que l'on veut plus de précision. Le système aux moments pour une telle fonction  $\mathcal{M}(\alpha)$  conduit alors à un système de  $M$  équations aux  $M$  inconnues  $\alpha_i$ . En utilisant la correspondance biunivoque entre le vecteur  $\alpha$  de cette exponentielle  $\mathcal{M}(\alpha)$  et son vecteur moment  $\rho = \int m \mathcal{M}(\alpha) dv$ , le vecteur flux  $\int v m \mathcal{M}(\alpha) dv$  peut être écrit sous la forme d'un terme  $G(\rho)$  et le terme de collision sous la forme d'un  $r(\rho)$ . Ainsi pour  $f$  de la forme  $\mathcal{M}(\alpha)$ , le système aux moments peut alors bien se mettre sous la forme

$$\partial_t \rho + \partial_x G(\rho) = r(\rho) \quad t \in \mathbb{R}_+, x \in \mathbb{R}.$$

Ce système qui est alors une extension du système d'Euler, est en particulier de type symétrique hyperbolique, donc bien posé mathématiquement.

Il est donc fondamental de connaître le domaine de définition de ce système, c'est-à-dire l'ensemble des moments  $\rho = \rho(t, x)$  pour lesquels la minimisation de l'entropie sous la contrainte  $\rho$  est réalisable. Ce domaine a été analysé par M. Junk dans [J1], [J2]; en particulier, et c'est ce que l'on présente essentiellement dans ce mémoire, M. Junk caractérise dans [J1] cet ensemble dans le cas de la dimension 1 d'espace ( $x, v \in \mathbb{R}$ ) et pour le système des cinq moments ( $m(v) = (1, \dots, v^4)$ ). Ce domaine n'est pas convexe; dans un cadre normalisé dans lequel on peut fixer les trois premières composantes des moments à  $\rho_0 = 1, \rho_1 = 0, \rho_2 = 1$ , il s'agit plus précisément d'un convexe défini par  $\rho_4 > 1 + \rho_3^2$  privé de la demi-droite  $\rho_4 \geq 3$ , et cette coupure qui correspond donc aux moments de contrainte pour lesquels le problème de minimisation n'a pas de solution est constituée des points en lesquels la fonction d'entropie perd sa stricte convexité et le flux  $G$  devient singulier.

Cette étude peut être vue comme un problème modèle pour le cas de dimensions supérieures (en espace et en moments), et il faut noter que C.D. Levermore présente dans [L] une hiérarchie de systèmes aux moments fermés dans un cadre général et que M. Junk dans [J2] s'intéresse à l'existence de solutions du problème de minimisation de l'entropie cinétique pour une large classe de problèmes aux moments.

## 1 - Hiérarchie de modèles aux moments pour l'équation de Boltzmann.

### 1.1 - L'équation de Boltzmann.

Si le mouvement des particules composant par exemple un gaz peut être considéré comme régi par les équations de Newton, leur grand nombre empêche d'utiliser des équations différentielles couplées pour suivre l'évolution de chacune d'elles et une description simplifiée statistique s'impose. Cette description introduite il y a plus d'un siècle par L. Boltzmann et J.C. Maxwell fait intervenir la fonction de distribution, ou densité,  $f = f(t, x, v)$  du gaz constitué de particules ponctuelles repérées à l'instant  $t$  par leur position  $x \in \mathbb{R}^N$

et leur vitesse  $v \in \mathbb{R}^N$ . Le modèle physique correspond à la dimension  $N = 3$ , mais pour simplifier on prendra souvent le modèle hypothétique d'un gaz unidimensionnel avec  $N = 1$ .

L'équation de Boltzmann régit l'évolution de la fonction de distribution  $f = f(t, x, v)$  d'un gaz raréfié et est de la forme

$$\partial_t f + v \cdot \nabla_x f = Q(f) \quad t \in \mathbb{R}_+, x, v \in \mathbb{R}^N. \quad (1.1)$$

Le terme  $Q(f) = Q(f)(t, x, v)$  appelé terme de collision, traduit le type d'interaction entre les particules et donc le type de collisions, et représente leur bilan statistique. L'opérateur  $Q$  qui agit sur les fonctions  $g = g(v)$  de la variable  $v$ , est de type quadratique et vérifie les propriétés importantes suivantes:

- conservation de la masse, du moment et de l'énergie durant les collisions dans le sens

$$\langle Q(g) \rangle = 0 \quad \langle v Q(g) \rangle = 0 \quad \langle |v|^2 Q(g) \rangle = 0, \quad (1.2)$$

où l'on note de façon générale

$$\langle g \rangle = \int_{\mathbb{R}^N} g(v) dv,$$

- dissipation de l'entropie (donnée par le théorème H de Boltzmann) dans le sens

$$\langle \ln(g) Q(g) \rangle \leq 0, \quad (1.3)$$

les *équilibres locaux*  $g$  de  $Q$  étant caractérisés par les trois propriétés équivalentes suivantes

$$\begin{aligned} (i) \quad & Q(g) = 0 \\ (ii) \quad & \langle \ln(g) Q(g) \rangle = 0 \\ (iii) \quad & g \text{ est une densité Maxwellienne,} \end{aligned} \quad (1.4)$$

c'est-à-dire  $g$  est de la forme

$$g(v) = \frac{\rho}{(2\pi T)^{N/2}} \exp\left(-\frac{|v-u|^2}{2T}\right)$$

pour des paramètres  $(\rho, u, T) \in \mathbb{R}_+ \times \mathbb{R}^N \times \mathbb{R}_+$ .

Les relations (1.2) impliquent que les solutions  $f$  de l'équation cinétique (1.1) satisfont formellement les lois macroscopiques de conservation de la masse  $\langle f \rangle$ , du moment  $\langle v f \rangle$  et de l'énergie cinétique  $\langle \frac{1}{2}|v|^2 f \rangle$  sous la forme

$$\begin{cases} \partial_t \langle f \rangle + \nabla_x \cdot \langle v f \rangle = 0 \\ \partial_t \langle v f \rangle + \nabla_x \cdot \langle v \otimes v f \rangle = 0 \\ \partial_t \langle \frac{1}{2}|v|^2 f \rangle + \nabla_x \cdot \langle \frac{1}{2}|v|^2 v f \rangle = 0, \end{cases} \quad t \in \mathbb{R}, x \in \mathbb{R}^N$$

et la relation (1.3) assure que ces solutions  $f$  satisfont formellement la loi macroscopique de dissipation locale de son entropie

$$H(f) = \langle f \ln(f) - f \rangle$$



sous la forme

$$\partial_t H(f) + \nabla_x \cdot \langle v(f \ln(f) - f) \rangle = \langle \ln(f) Q(f) \rangle \leq 0.$$

A l'équation d'évolution (2.1), il convient d'adjoindre une condition initiale

$$f(0, x, v) = f^0(x, v) \quad x, v \in \mathbb{R}^N$$

où la donnée  $f^0$  décrit l'état des particules en position et vitesse à l'instant initial  $t = 0$ .

## 1.2 - Equilibre local et déviations.

L'équation microscopique de Boltzmann régit donc la distribution de vitesse  $f$  des particules d'un gaz raréfié, peu dense, ce qui correspond à un fort déséquilibre cinétique. Si le libre parcours moyen, c'est-à-dire la distance moyenne parcourue par les particules entre deux collisions, est petit devant certaines grandeurs macroscopiques du problème, ce qui correspond à un faible déséquilibre cinétique, alors le gaz va approcher un état d'équilibre local et son évolution est décrite par les équations macroscopiques d'Euler compressibles pour un équilibre, ou par celles de Navier-Stokes compressibles pour de petites déviations autour d'un équilibre.

Plus précisément dans le cas d'un équilibre, la distribution de vitesse  $f(t, x, v)$  est donnée par une densité particulière, appelée distribution *Maxwellienne*, du type

$$\mathcal{E}(t, x, v) = \frac{\rho(t, x)}{2\pi T(t, x)} \exp\left(-\frac{|v - u(t, x)|^2}{2T(t, x)}\right) \quad (1.5)$$

où les grandeurs macroscopiques  $(\rho, u, T) \in \mathbb{R}_+ \times \mathbb{R}^N \times \mathbb{R}_+$  représentent respectivement la densité du gaz, la vitesse moyenne et la température à l'instant  $t$  et au point  $x$ , leur évolution étant gouvernée par les équations d'Euler compressibles (ou de la dynamique des gaz)

$$\begin{cases} \partial_t \rho + \nabla_x \cdot (\rho u) = 0 \\ \partial_t (\rho u) + \nabla_x \cdot (\rho u \otimes u) + \nabla_x (\rho T) = 0 \\ \partial_t \left(\frac{1}{2} \rho |u|^2 + \frac{N}{2} \rho T\right) + \nabla_x \cdot \left(\frac{1}{2} \rho |u|^2 u + \frac{N+2}{2} \rho T u\right) = 0. \end{cases} \quad t \in \mathbb{R}, x \in \mathbb{R}^N$$

Ces équations d'Euler peuvent être obtenues directement à partir de considérations de la Mécanique des Fluides, mais aussi par certains passages à la limite de l'équation de Boltzmann. Ainsi l'idée fondamentale du passage à la limite appelé limite hydrodynamique, consiste à étudier le comportement du gaz raréfié lorsque les collisions deviennent de plus en plus fréquentes. Pour cela on considère l'équation de Boltzmann avec un paramètre d'échelle  $\varepsilon$

$$\partial_t f_\varepsilon + v \cdot \nabla_x f_\varepsilon = \frac{1}{\varepsilon} Q(f_\varepsilon), \quad (1.6)$$

ce paramètre  $\varepsilon$  étant un nombre sans dimension, de l'ordre du rapport entre le libre parcours moyen et certaines grandeurs macroscopiques.

Si les premiers membres de (1.6) restent bornés (par exemple par les lois de conservation) on en déduit formellement que  $Q(f_\varepsilon)$  tend vers 0 lorsque  $\varepsilon$  tend vers 0, et si  $f_\varepsilon$

tend vers une distribution  $f$ , alors  $Q(f) = 0$ , si bien que d'après le théorème H sur les équilibres de  $Q$  alors  $f$  est nécessairement une Maxwellienne du type (1.5).

Les équations macroscopiques d'Euler peuvent aussi être retrouvées à partir de l'équation microscopique de Boltzmann en prenant les moments par rapport à  $1$ ,  $v$  et  $|v|^2$ , et en fermant ces équations aux moments en cherchant la distribution  $f$  sous forme Maxwellienne.

Près d'un équilibre, une telle Maxwellienne (1.5) peut être considérée comme une première approximation de la distribution des vitesses  $f$ . De façon plus générale, les déviations par rapport à un équilibre local peuvent être précisées par des méthodes de développement introduites par Hilbert et par Chapman-Enskog. On développe formellement la distribution  $f_\varepsilon$  de l'équation (1.6) en puissances de  $\varepsilon$  sous la forme

$$f_\varepsilon = \mathcal{E} \left( 1 + \varepsilon f^{(1)} + \varepsilon^2 f^{(2)} + \dots \right) \quad (1.7)$$

où le terme principal  $\mathcal{E}$  est une maxwellienne du type (1.5).

L'approximation de Navier-Stokes consiste à ne conserver que la correction à l'ordre 1, c'est-à-dire à considérer l'approximation de  $f_\varepsilon$  par  $\mathcal{E} (1 + \varepsilon f^{(1)})$ . En reportant (1.7) dans l'équation (1.6), on vérifie que la correction  $f^{(1)}$  en chaque  $(t, x)$  est alors exprimée en fonction des variables  $\rho$ ,  $u$  et  $T$ , l'évolution de ces quantités macroscopiques étant gouvernées par les équations de Navier-Stokes compressibles

$$\begin{cases} \partial_t \rho + \nabla_x \cdot (\rho u) = 0 \\ \partial_t (\rho u) + \nabla_x \cdot (\rho u \otimes u) + \nabla_x (\rho T) + \nabla_x \cdot \Sigma = 0 \\ \partial_t \left( \frac{1}{2} \rho |u|^2 + \frac{N}{2} \rho T \right) + \nabla_x \cdot \left( \frac{1}{2} \rho |u|^2 u + \frac{N+2}{2} \rho T u + \Sigma u + q \right) = 0 \end{cases} \quad t \in \mathbb{R}, x \in \mathbb{R}^N$$

dans lesquelles le terme des contraintes  $\Sigma$  et le flux de chaleur  $q$  s'écrivent alors

$$\begin{cases} \Sigma = -\mu (\nabla_x u + (\nabla_x u)^T) - \frac{2}{N} (\nabla_x \cdot u) I \\ q = -K \nabla_x T \end{cases}$$

où  $\mu = \mu(\rho, T)$  est la viscosité et  $K = K(\rho, T)$  est la conductivité thermique.

Ces équations de Navier-Stokes peuvent aussi être obtenues en prenant les moments de l'équation de Boltzmann par rapport à  $1$ ,  $v$  et  $|v|^2$ , et en fermant les équations aux moments correspondantes en cherchant la distribution  $f_\varepsilon$  sous la forme  $\mathcal{E} (1 + \varepsilon f^{(1)})$ , ce qui dans cette fermeture de Navier-Stokes, fixe  $\Sigma$  et  $q$  en fonction de  $\rho$ ,  $u$  et  $T$  sous la forme précédente.

Notons cependant que cette fermeture ne convient que lorsque  $\varepsilon$  tend vers 0; au contraire quand les grandeurs macroscopiques deviennent comparables au libre parcours moyen, c'est-à-dire lorsque  $\varepsilon$  est de l'ordre de 1, les termes  $\Sigma$  et  $q$  des équations de Navier-Stokes ne représentent plus la physique et la dynamique des fluides obtenue n'est alors plus valable.

Les équations d'Euler - Navier-Stokes n'étant donc plus valables dans ce régime de transition, et pour décrire des gaz dont l'état est plus éloigné d'un équilibre, C.D. Levermore a proposé de trouver de nouvelles équations en développant une méthode qui généralise le fait que dans le cas d'un équilibre local, la maxwellienne  $\mathcal{E}$  minimise l'entropie

$H(f) = \langle f \ln f - f \rangle$  sous la contrainte que les moments de  $f$  par rapport à 1,  $v$  et  $|v|^2$  soient prescrits. Cette méthode consiste encore à chercher à décrire la distribution de vitesse  $f$  des particules par la minimisation d'une fonction entropie, mais avec un nombre de contraintes sur les moments d'autant plus grand qu'on voudra avoir plus de détails sur la structure de la distribution de vitesse.

### 1.3 - Fermeture des modèles aux moments.

La construction des modèles aux moments introduite par H. Grad [G], puis développée par C.D. Levermore [L], débute par le choix d'un espace  $\mathcal{M}$  de dimension finie  $M$  de fonctions de la variable  $v$  de  $\mathbb{R}^N$  et à valeurs réelles,  $\mathcal{M}$  étant habituellement un espace de polynômes. En considérant les moments de l'équation de Boltzmann par rapport au vecteur  $m = m(v)$  des fonctions de base de  $\mathcal{M}$ , on est conduit à la construction d'un système de  $M$  équations aux moments

$$\partial_t \langle m f \rangle + \nabla_x \cdot \langle m v f \rangle = \langle m Q(f) \rangle. \quad (1.8)$$

Le choix d'une relation supplémentaire est alors nécessaire pour fermer les modèles aux moments que l'on souhaite construire. Le choix n'est pas unique et n'est soumis a priori qu'à la positivité de la distribution de vitesse  $f$  du modèle que l'on souhaite construire. La fermeture du modèle aux moments est obtenue en cherchant la distribution qui minimise l'entropie sous la contrainte de réaliser les moments engendrés sur l'espace  $\mathcal{M}$ .

On peut donner une idée formelle conduisant à la structure de ces densités optimales. La condition nécessaire d'Euler-Lagrange portant sur la dérivée première en un tel minimum donne l'existence d'un multiplicateur de Lagrange  $\alpha \in \mathbb{R}^M$  tel que

$$0 = \ln f - \alpha \cdot m(v).$$

Ainsi formellement, les densités optimales  $f$  sont nécessairement de la forme

$$f = \mathcal{M}(\alpha) := \exp(\alpha \cdot m).$$

Les distributions présentes dans le système aux moments (1.8) seront donc prises de la forme  $f = \mathcal{M}(\alpha)$  pour un  $\alpha = \alpha(t, x)$  à valeurs dans  $\mathbb{R}^M$  tel que  $m f$  soit intégrable en  $v$ . On obtient ainsi un système de  $M$  équations à  $M$  inconnues  $\alpha = (\alpha_0, \dots, \alpha_{M-1})$

$$\partial_t \langle m \mathcal{M}(\alpha) \rangle + \nabla_x \cdot \langle m v \mathcal{M}(\alpha) \rangle = \langle m Q(\mathcal{M}(\alpha)) \rangle. \quad (1.9)$$

Introduisons alors le potentiel densité  $h^*$  et le potentiel flux  $j^*$  définis par

$$h^*(\alpha) = \langle \mathcal{M}(\alpha) \rangle \quad j^*(\alpha) = \langle v \mathcal{M}(\alpha) \rangle.$$

En les différentiant par rapport à  $\alpha$ , on reconnaît les termes du système (1.9) qui se met sous la forme suivante, appelée forme potentielle de Godunov

$$\partial_t h_\alpha^*(\alpha) + \nabla_x \cdot j_\alpha^*(\alpha) = r^*(\alpha). \quad (1.10)$$

Ce modèle aux moments possède des propriétés structurelles importantes d'hyperbolicité et d'entropie, l'hyperbolicité du système assurant en particulier que celui-ci est bien posé mathématiquement.

**Proposition 1.1** *Le système  $\partial_t h_\alpha^*(\alpha) + \nabla_x \cdot j_\alpha^*(\alpha) = r^*(\alpha)$  possède les deux propriétés suivantes:*

- 1 . *Il est hyperbolique symétrisable.*
- 2 . *Il dissipe localement l'entropie  $S$  définie par*

$$S(\alpha) = \langle \mathcal{M}(\alpha) \ln(\mathcal{M}(\alpha)) - \mathcal{M}(\alpha) \rangle .$$

**Démonstration.** 1 . Le système (1.10) peut être réécrit sous la forme

$$h_{\alpha\alpha}^*(\alpha) \partial_t \alpha + j_{\alpha\alpha}^*(\alpha) \nabla_x \alpha = r^*(\alpha) \quad (1.11)$$

dans lequel la matrice hessienne  $h_{\alpha\alpha}^*$  de  $h^*$  est symétrique définie positive et la matrice hessienne  $j_{\alpha\alpha}^*(\alpha)$  de  $j^*$  est symétrique. Ainsi le système (1.10) peut être réécrit sous la forme d'un système hyperbolique symétrique en  $\alpha$ .

2 - En prenant le produit scalaire de (1.11) à gauche par le vecteur  $\alpha$  on obtient

$$\partial_t (\alpha \cdot h_\alpha^*(\alpha) - h^*(\alpha)) + \nabla_x (\alpha \cdot j_\alpha^*(\alpha) - j^*(\alpha)) = \alpha \cdot r^*(\alpha), \quad (1.12)$$

où

$$\alpha \cdot r^*(\alpha) = \langle \alpha \cdot m Q(\mathcal{M}(\alpha)) \rangle = \langle \ln(\mathcal{M}(\alpha)) Q(\mathcal{M}(\alpha)) \rangle .$$

La relation de dissipation locale (1.3) que vérifie l'opérateur  $Q$  implique que

$$\alpha \cdot r^*(\alpha) \leq 0, \quad (1.13)$$

tandis que la caractérisation (1.4) des équilibres de  $Q$  conduit à l'équivalence des trois propriétés suivantes portant sur les équilibres de  $r^*$

$$\begin{aligned} (i) \quad & r^*(\alpha) = 0 \\ (ii) \quad & \alpha \cdot r^*(\alpha) = 0 \\ (iii) \quad & \alpha \cdot m \text{ est engendré par } 1, v, v^2. \end{aligned} \quad (1.14)$$

Or les densités dans (1.12) peuvent être réécrites sous la forme

$$\alpha \cdot h_\alpha^*(\alpha) - h^*(\alpha) = \langle \mathcal{M}(\alpha) \ln(\mathcal{M}(\alpha)) - \mathcal{M}(\alpha) \rangle \quad (1.15)$$

que l'on a noté  $S(\alpha)$ , et

$$\alpha \cdot j_\alpha^*(\alpha) - j^*(\alpha) = \langle v (\mathcal{M}(\alpha) \ln(\mathcal{M}(\alpha)) - \mathcal{M}(\alpha)) \rangle$$

que l'on notera  $\eta(\alpha)$ .

Ainsi le système (1.10) vérifie la relation

$$\partial_t S(\alpha) + \nabla_x \cdot \eta(\alpha) \leq 0,$$

c'est-à-dire qu'il dissipe localement l'entropie  $S(\alpha)$  associée à la densité  $\mathcal{M}(\alpha)$  ■

Traditionnellement, et c'est ce qu'on utilisera dans toute la suite, le système (1.10) est exprimé en terme des moments notés  $\rho$  définis par

$$\rho(t, x) = \langle m \mathcal{M}(\alpha(t, x)) \rangle ,$$

c'est-à-dire que  $\rho$  est aussi défini comme une fonction de  $\alpha$  par

$$\rho = h_\alpha^*(\alpha). \quad (1.16)$$

Cette relation peut être inversée et il y a une correspondance biunivoque entre le coefficient  $\alpha = \alpha(t, x)$  d'une densité exponentielle  $f = \mathcal{M}(\alpha)$  et ses moments  $\rho = \rho(t, x) = \langle m \mathcal{M}(\alpha(t, x)) \rangle$ . Plus précisément, considérons la transformée de Legendre  $h = h(\rho)$  de la fonction strictement convexe  $h^*$  définie par

$$h(\rho) + h^*(\alpha) = \alpha \cdot \rho \quad (1.17)$$

$\rho$  et  $\alpha$  étant reliés par (1.16). Alors  $\alpha$  peut être exprimé comme fonction de  $\rho$  par

$$\alpha = h_\rho(\rho). \quad (1.18)$$

En comparant (1.15) et (1.17),  $h(\rho)$  s'identifie à la quantité  $S(\alpha)$  pour  $\alpha$  et  $\rho$  reliés par (1.16), et apparaît comme une entropie. Le système (1.10) peut alors être formulé en terme de moments  $\rho$  sous la forme

$$\partial_t \rho + \nabla_x \cdot G(\rho) = r(\rho) \quad (1.19)$$

où le vecteur flux  $G(\rho)$  est défini par

$$G(\rho) = \langle v m(v) \mathcal{M}(\alpha(\rho)) \rangle = j_\alpha^*(h_\rho(\rho)).$$

Notons que les deux propriétés structurelles d'hyperbolicité et de dissipation locale de l'entropie  $S(\alpha)$  remarquées sur le système (1.10) (en la fonction  $\alpha$ ) se retrouvent bien sur cette formulation (1.19) (en la fonction  $\rho$ ), en particulier pour la dissipation de l'entropie  $h(\rho)$ .

## 2 - Etude du modèle aux cinq moments en dimension 1 d'espace.

On a déjà noté que les équations d'Euler peuvent être retrouvées à partir de l'équation de Boltzmann en prenant les moments par rapport à 1,  $v$ ,  $v^2$  et en supposant que la

distribution de vitesse est de la forme  $f(t, x, v) = \exp(\sum_{k=0}^2 \alpha_k(t, x) v^k)$ .

Le système qui suit dans la hiérarchie est obtenu de manière analogue mais à partir du vecteur

$$m(v) = (1, v, v^2, v^3, v^4)$$

et d'une densité de la forme

$$f(t, x, v) = \exp(\sum_{k=0}^4 \alpha_k(t, x) v^k).$$

Dans ce qui suit, et en posant  $\alpha = (\alpha_0, \dots, \alpha_4)$ , une telle densité sera notée

$$\mathcal{M}(\alpha) = \exp(\alpha \cdot m).$$

Il s'agit donc du système (1.19) à l'ordre  $M = 5$  et c'est celui que l'on va considérer désormais. On en a déjà donné deux propriétés importantes dans la proposition 1.1 concernant l'hyperbolicité et l'entropie. Maintenant, on va étudier plus particulièrement le domaine de définition du vecteur flux  $G$ . Par sa définition,  $G(\rho)$  est bien défini pour un  $\rho$  donné si un coefficient  $\alpha$  peut être déterminé, et alors il est unique d'après (1.16) et (1.18), de telle sorte que

$$\rho = \mu(\mathcal{M}(\alpha))$$

où de façon générale  $\mu$  est l'opérateur des cinq moments défini par

$$\mu(f) = \langle m f \rangle .$$

On s'intéressera aux distributions de type exponentiel de l'ensemble

$$E = \{\mathcal{M}(\alpha); \alpha \in \mathcal{A}\}$$

où

$$\mathcal{A} = \{\alpha \in \mathbb{R}^5; \alpha_4 < 0\} \cup \{\alpha \in \mathbb{R}^5; \alpha_2 < 0, \alpha_3 = \alpha_4 = 0\}.$$

En conséquence  $G$  est défini sur l'ensemble

$$\mu(E) = \{\rho = \mu(\mathcal{M}(\alpha)); \alpha \in \mathcal{A}\}.$$

On va décrire cet ensemble  $\mu(E)$  qui, comme on va le voir, joue un rôle important dans le problème de minimisation de l'entropie de fonctions  $f$  sous une contrainte  $\mu(f) = \rho$  donnée.

## 2.1 - Minimisation de l'entropie H.

Plus généralement les cinq moments sont bien définis sur l'ensemble  $D$  suivant:

$$D = \{f \geq 0; (1 + v^4)f(v) \in L^1(\mathbb{R}), f \neq 0\}.$$

Etant donné un vecteur admissible  $\rho \in \mu(D)$ , le problème des moments, à savoir trouver  $f \in D$  tel que  $\mu(f) = \rho$ , est résoluble par définition. Cependant cette solution n'est pas unique et on cherche une solution particulière compatible avec l'information donnée  $\rho$ , mais contenant, si possible, une information supplémentaire. Le critère retenu est celui de minimiser la fonction entropie  $H$  définie par  $H(g) = \langle g \ln(g) - g \rangle$ , et on cherche donc  $f$  tel que

$$\begin{cases} f \in D, \mu(f) = \rho \\ H(f) = \inf\{H(g); g \in D, \mu(g) = \rho\} \end{cases}$$

Tout d'abord il faut noter que pour  $\rho \in \mu(D)$  fixé, la borne inférieure de l'ensemble  $\{H(g); g \in D, \mu(g) = \rho\}$  est bien définie dans  $\mathbb{R}$ . Par contre il existe des moments admissibles  $\rho \in \mu(D)$  pour lesquels le problème des moments est résoluble, mais pour lesquels le problème de minimisation de l'entropie  $H$  n'a pas de solution, c'est-à-dire pour lesquels cette borne inférieure n'est pas atteinte. On verra plus précisément qu'il existe une solution d'entropie minimale sous la contrainte des moments  $\rho$  si et seulement si  $\rho \in \mu(E)$ .

Au lieu d'utiliser directement la fonctionnelle d'entropie  $H$ , il est pratique de considérer l'entropie relative  $H(f, \hat{f})$  associée à la maxwellienne normalisée

$$\hat{f}(v) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}v^2}$$

(c'est-à-dire  $\hat{f} = \mathcal{M}(\hat{\alpha})$  correspondant au coefficient  $\hat{\alpha} = -\frac{1}{2}(\ln(2\pi), 0, 1, 0, 0)$  et aux moments  $\hat{\rho} = (1, 0, 1, 0, 3)$ ) et définie par

$$\begin{aligned} H(f, \hat{f}) &= \langle f \ln\left(\frac{f}{\hat{f}}\right) \rangle \\ &= H(f) + \frac{1}{2} [(2 + \ln(2\pi))\mu_0(f) + \mu_2(f)]. \end{aligned}$$

Sous une contrainte  $\mu(f) = \rho$  fixée, minimiser l'entropie  $H(f)$  est donc équivalent à minimiser l'entropie relative  $H(f, \hat{f})$ .

On considère alors la fonction d'entropie  $S$  définie pour  $\rho \in \mu(D)$  par

$$S(\rho) = \inf\{H(f, \hat{f}); f \in D, \mu(f) = \rho\}.$$

Cette borne inférieure  $S(\rho)$  est bien définie sur  $\mu(D)$  comme on l'a déjà noté, et elle a la propriété suivante:

**Proposition 2.1** *La fonction entropie  $S$  est une fonction convexe sur  $\mu(D)$ , strictement convexe sur  $\mu(E)$  et telle que  $\frac{\partial S}{\partial \rho_4} < 0$  dans l'intérieur de  $\mu(E)$ .*

De plus, le résultat suivant caractérise les moments admissibles  $\rho$  pour lesquels le problème de minimisation de l'entropie sous la contrainte  $\rho$  est résoluble.

**Théorème 2.2** *Pour  $\mu \in \rho(D)$ , les deux propositions suivantes sont équivalentes:*

1 -  $\rho \in \mu(E)$ .

2 - La borne inférieure  $S(\rho)$  est atteinte.

Plus précisément si  $\rho = \mu(\exp(\alpha \cdot m))$  pour un (unique)  $\alpha \in \mathcal{A}$ , alors  $S(\rho)$  est atteinte par la fonction (unique)  $\exp(\alpha \cdot m)$ .

On peut noter que cette fonctionnelle  $S$  est en relation avec l'entropie  $h$  du système hyperbolique (1.19) définie par

$$h(\rho) = \alpha(\rho) \cdot \rho - \rho_0.$$

Plus précisément pour  $\rho \in \mu(E)$ , on a

$$S(\rho) = \frac{1}{2} ((\ln(2\pi) + 2)\rho_0 + \rho_2) + h(\rho).$$

## 2.2 - Caractérisation du domaine de définition $\mu(E)$ .

Pour caractériser les ensembles  $\mu(D)$  et  $\mu(E)$ , il peut être pratique de normaliser les densités  $f$  en fixant les trois premiers moments. Pour  $f \in D$ , on définit une densité macroscopique  $n(f)$ , une vitesse moyenne  $u(f)$  et une température  $T(f)$  par

$$n = \mu_0, \quad u = \frac{\mu_1}{\mu_0}, \quad T = \frac{\mu_2}{\mu_0} - \left(\frac{\mu_1}{\mu_0}\right)^2$$

et on lui associe la densité

$$f^*(w) = \frac{\sqrt{T(f)}}{n(f)} f\left(\sqrt{T(f)}w + u(f)\right).$$

Alors  $\mu_0(f^*) = \mu_2(f^*) = 1$  et  $\mu_1(f^*) = 0$ . Et inversement on peut retrouver la densité  $f$  à partir de la densité normalisée  $f^*$  et de  $n, u$  et  $T$  par

$$f(v) = \frac{n(f)}{\sqrt{T(f)}} f^*\left(\frac{v - u(f)}{\sqrt{T(f)}}\right).$$

L'ensemble des moments admissibles  $\mu(D)$  peut aussi être retrouvé par des formules explicites à partir de l'ensemble des moments normalisés  $\mu(D^*)$  où  $D^* = \{f^*; f \in D\}$ . Ainsi, caractériser l'ensemble  $\mu(D)$  est équivalent à caractériser l'ensemble  $\mu(D^*)$ . Or l'ensemble de ces moments normalisés  $\mu(D^*)$  peut être caractérisé simplement et on a les propriétés suivantes:

**Proposition 2.3** 1 -  $\mu(D^*) = \{(1, 0, 1, \rho_3, \rho_4); \rho_4 > 1 + \rho_3^2\}$ .

2 - La fonction entropie  $S$  est décroissante par rapport à  $\rho_4$  dans  $\mu(D^*)$ .

De même le sous-ensemble  $\mu(E)$  peut être retrouvé à partir de l'ensemble normalisé  $\mu(E^*)$ . Or cet ensemble  $\mu(E^*)$  peut être caractérisé simplement et on a les propriétés suivantes:

**Théorème 2.4** 1 -  $\mu(E^*) = \mu(D^*) \setminus \{(1, 0, 1, 0, \rho_4); \rho_4 > 3\}$ .

2 - La fonction entropie  $S$  est strictement convexe sur  $\mu(E^*)$  et est égale à son minimum absolu  $S(\hat{\rho}) = 0$  sur la coupure  $\{(1, 0, 1, 0, \rho_4); \rho_4 \geq 3\}$ .

**Démonstration.** On donne les idées de la démonstration de ce théorème en utilisant certaines propriétés de l'entropie relative que l'on admettra.

Soit  $\rho \in \mu(D^*)$  fixé et soit le convexe  $F = \{f \in D^*; \mu(f) = \rho\}$ . Par la définition de  $S(\rho)$ , soit  $(f_n) \subset F$  une suite minimisante telle que

$$S(\rho) = \lim_{n \rightarrow \infty} H(f_n, \hat{f}).$$

Par des propriétés de l'entropie relative  $H$  sur l'ensemble  $F$ , on montre que la suite  $(f_n)$  converge vers une fonction  $f$  dans  $L^1(\mathbb{R})$ . D'une part cette limite  $f$  appartient à  $D$  car par le lemme de Fatou

$$\mu_4(f) = \langle v^4 f \rangle \leq \lim_{n \rightarrow +\infty} \langle v^4 f_n \rangle = \rho_4,$$



donc  $v^4 f \in L^1$ , et comme  $f \in L^1$ , alors  $f \in D$ . D'autre part pour tout  $R > 0$  on a

$$\int_{|v|>R} |v|^i f_n(v) dv \leq \int_{|v|>R} \left(\frac{|v|}{R}\right)^{4-i} |v|^i f_n(v) dv \leq R^{i-4} \rho_4 \quad \text{pour } i = 0, 1, 2, 3, 4,$$

$$\int_{|v|\leq R} v^i f(v) dv = \lim \int_{|v|\leq R} v^i f_n(v) dv = \rho_i + \mathcal{O}\left(\frac{1}{R}\right). \quad \text{pour } i = 0, 1, 2, 3.$$

De là on déduit que

$$\mu_i(f) = \rho_i \quad i = 0, 1, 2, 3 \quad \text{et} \quad \mu_4(f) \leq \rho_4.$$

En particulier  $f \in D^*$ .

Or comme  $S(\rho)$  est décroissante par rapport à  $\rho_4$ , on en déduit que

$$S(\mu(f)) \geq S(\rho),$$

et comme l'entropie relative est semi-continue inférieurement, on a

$$H(f, \hat{f}) \leq \liminf_{n \rightarrow +\infty} H(f_n, \hat{f}) = S(\rho).$$

Il en résulte alors que

$$S(\mu(f)) = H(f, \hat{f})$$

et ainsi  $f$  réalise le minimum de l'entropie relative  $H(\cdot, \hat{f})$  sous la contrainte  $\mu(f)$ . Par le Théorème 2.2, on en déduit que  $\mu(f) \in \mu(E^*)$  et que  $f = \exp(\alpha \cdot m) = \mathcal{M}(\alpha)$  pour un certain  $\alpha \in \mathcal{A}$

Par conséquent pour  $\rho \in \mu(D^*)$  quelconque, la suite minimisante  $(f_n)$  converge toujours vers une densité exponentielle, mais on ne sait pas si la contrainte  $\mu(f) = \rho$  est satisfaite, c'est-à-dire si  $\mu_4(f) = \rho_4$  (pour pouvoir dire que  $\rho \in \mu(E^*)$  par le Théorème 2.2). Tout ce que l'on peut dire d'après ce qui précède est que  $\mu_4(f) \leq \rho_4$ .

Considérons alors la coupure  $L$  de  $\mu(D^*)$  définie par

$$L = \{(1, 0, 1, 0, \rho_4) ; \rho_4 > 3\},$$

qui est une demi-droite de  $\mu(D^*)$  issue du point  $\hat{\rho} = \mu(\hat{f})$ . Tout d'abord comme  $S$  est décroissante par rapport à  $\rho_4$ , on a

$$S(\rho) \leq S(\hat{\rho}) \quad \forall \rho \in L,$$

et comme  $H(g, \hat{f}) \geq 0$  pour tout  $g \in D^*$ , avec égalité si et seulement si  $g = \hat{f}$  d'après des propriétés de l'entropie relative, on a

$$S(\hat{\rho}) = H(\hat{f}, \hat{f}) = 0.$$

Par suite

$$S(\rho) = 0 \quad \forall \rho \in L.$$

En particulier si la contrainte  $\rho \in L$ , alors la densité limite  $f = \mathcal{M}(\alpha)$  vérifiera

$$0 = S(\rho) = S(\mu(f)) = H(f, \hat{f})$$

et donc  $f = \hat{f}$ . Par suite  $\mu_4(f) = \mu_4(\hat{f}) = 3 < \rho_4$  et la limite  $f$  ne vérifie donc pas la condition des moments. Ce n'est donc pas le maximum de l'entropie sous la contrainte  $\rho$  et par conséquent, par le théorème 2.2, on a  $\rho \notin \mu(E^*)$ . Ainsi

$$\mu(E^*) \subset \mu(D^*) \setminus L.$$

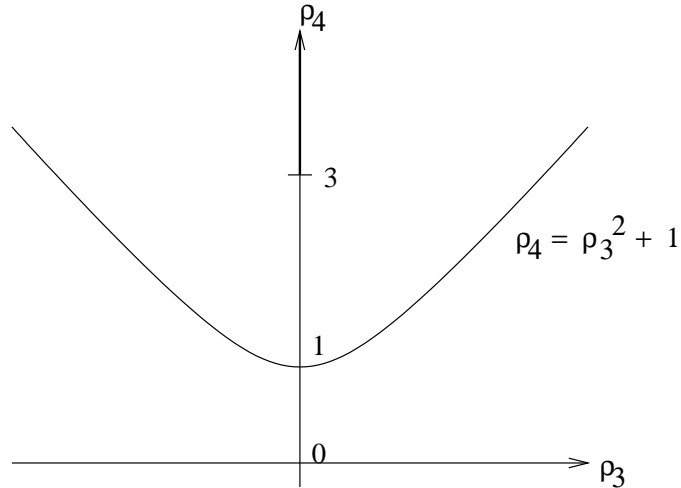
Inversement si  $\rho \notin L$ , la fonction limite  $f = \mathcal{M}(\alpha)$  associée à cette contrainte  $\rho$  vérifie  $\mu(f) + (0, \dots, 0, \delta) = \rho$  pour un certain  $\delta \geq 0$ . Deux cas sont alors à examiner suivant le coefficient  $\alpha$ :

- si  $\alpha_4 = 0$ , alors  $\alpha$  appartient à  $\partial\mathcal{A}$  et  $f$  est la Maxwellienne normalisée  $\hat{f}$  pour laquelle  $\mu(f) = \hat{\rho}$ . Le moment prescrit  $\rho$  est donc nécessairement de la forme

$$\rho = (1, 0, 1, 0, 3 + \delta)$$

avec  $\delta \geq 0$ , et donc  $\delta = 0$  car  $\rho \notin L$ . Par suite  $\rho = \mu(f)$  et donc  $\rho \in \mu(E^*)$ .

- si  $\alpha_4 < 0$ , alors  $\rho$  appartient à l'intérieur de  $\mu(E)$  sur lequel  $S$  est localement strictement décroissante en  $\rho_4$ . Ainsi  $S(\rho) = S(\mu(f) + (0, \dots, 0, \delta)) < S(\mu(f))$  si  $\delta > 0$ , ce qui contredit le fait que nécessairement  $S(\rho) = S(\mu(f))$  pour la fonction limite  $f$ . Par conséquent  $\delta = 0$  et donc  $\rho = \mu(f) \in \mu(E^*)$  ■



Représentation de  $\mu(E^*)$  dans le plan  $(\rho_3, \rho_4)$ .

### 2.3 - Caractérisation de la coupure L.

On complète<sup>1</sup> la caractérisation de  $\mu(E^*)$  donnée par M. Junk dans le théorème 2.4, en donnant une description de la coupure  $L = \{(1, 0, 1, 0, \rho_4); \rho_4 > 3\}$  de  $\mu(E^*)$  dans  $\mu(D^*)$  à l'aide des fonctions

$$f_{\alpha a}(v) = e^{-\alpha|v|^a}.$$

En notant  $f_{\alpha a}^*$  la densité normalisée associée à la densité exponentielle  $f_{\alpha a}$ , on a:

**Proposition 2.5** *Pour tout  $\alpha > 0$ , on a  $L = \{\mu(f_{\alpha a}^*); 0 < a < 2\}$ .*

<sup>1</sup> Ce résultat n'est pas donné dans les travaux de M. Junk. Il s'agit d'une réponse donnée à une question posée par B. Perthame.

**Démonstration.** Tout d'abord par des calculs élémentaires sur l'intégrale eulérienne

$$\Gamma(x) = \int_0^{+\infty} e^{-t} t^{x-1} dt$$

on obtient que

$$\mu_0(f_{\alpha a}) = 2\alpha^{-1/a} \frac{1}{a} \Gamma\left(\frac{1}{a}\right), \quad \mu_1(f_{\alpha a}) = 0, \quad \mu_2(f_{\alpha a}) = \frac{2}{a} \alpha^{-3/a} \Gamma\left(\frac{3}{a}\right)$$

et par normalisation

$$\mu_0(f_{\alpha a}^*) = \mu_2(f_{\alpha a}^*) = 1, \quad \mu_1(f_{\alpha a}^*) = \mu_3(f_{\alpha a}^*) = 0, \quad \mu_4(f_{\alpha a}^*) = \frac{\Gamma\left(\frac{1}{a}\right) \Gamma\left(\frac{5}{a}\right)}{\left(\Gamma\left(\frac{3}{a}\right)\right)^2}.$$

Ensuite en utilisant en particulier la caractérisation de  $\Gamma(x)$  par le produit infini de Weierstrass, on montre que la fonction  $g$  définie par

$$g(x) = \frac{\Gamma(x) \Gamma(5x)}{(\Gamma(3x))^2}$$

est continue et strictement croissante de  $]0, +\infty[$  sur  $]5/9, +\infty[$  avec  $g(1/2) = 3$ .

De ces deux étapes, on en déduit le résultat annoncé ■

## 2.4 - Sur le problème de Cauchy.

Le problème de Cauchy pour le système des cinq moments peut donc être écrit sous la forme

$$\begin{cases} \partial_t \rho + \partial_x G(\rho) = r(\rho) & t \in \mathbb{R}_+, x \in \mathbb{R} \\ \rho(0, x) = \rho^0(x) & x \in \mathbb{R} \end{cases} \quad (2.1)$$

et d'après ce qui a été vu précédemment, ce système est symétrique hyperbolique et le vecteur flux  $G$  est défini sur la famille  $\mu(E)$  des moments des exponentielles par  $G(\rho) = \langle v m \mathcal{M}(\alpha(\rho)) \rangle$ .

On va d'abord montrer que  $G(\rho)$  n'est pas prolongeable sur  $\mu(D)$  tout entier. En effet pour sa cinquième composante

$$G_4(\rho) = \langle v^5 \mathcal{M}(\alpha(\rho)) \rangle$$

on a l'estimation suivante:

**Proposition 2.6** *Pour tout  $\rho \in \mu(E^*)$  on a*

$$\rho_3 G_4(\rho) \geq (\rho_4 - 3)(\rho_4 - 1) + 4\rho_3^2$$

*avec égalité si et seulement si  $\rho = \hat{\rho} (= (1, 0, 1, 0, 3))$ .*

Cette estimation permet en particulier de préciser le comportement de  $G(\rho)$  près de la coupure L. En effet pour  $\varepsilon$  donné et  $\rho_4 \geq 3 + \varepsilon$  fixé, alors

$$G_4(\rho) \geq \frac{\varepsilon(2 + \varepsilon)}{\rho_3} + 4\rho_3 \text{ pour } \rho_3 > 0, \quad G_4(\rho) \leq \frac{\varepsilon(2 + \varepsilon)}{\rho_3} + 4\rho_3 \text{ pour } \rho_3 < 0.$$

Ainsi  $G_4(\rho)$  tend vers  $+\infty$  si  $\rho_3$  tend vers 0,  $\rho_3 > 0$  et  $G_4(\rho)$  tend vers  $-\infty$  si  $\rho_3$  tend vers 0,  $\rho_3 < 0$ . Autrement dit le terme de flux  $G$  devient singulier sur la coupure  $L$  et par conséquent il n'est pas prolongeable à l'ensemble  $\mu(D)$  tout entier.

Pour que le problème (2.1) soit bien posé, on doit donc déjà s'assurer que la condition initiale  $\rho^0(x)$  appartient à  $\mu(E)$  pour tout  $x \in \mathbb{R}$ . Si par exemple  $\rho^0(\mathbb{R})$  est contenu dans un compact situé dans l'intérieur de  $\mu(E)$  et si  $\rho^0$  est suffisamment régulier, alors le problème de Cauchy hyperbolique (2.1) admet une solution classique pour  $t$  petit. Cependant il se peut qu'en temps fini, la solution perde de la régularité ou bien qu'elle ne reste pas confinée dans un compact de  $\mu(E)$  fixé à l'avance. Dans ce contexte la solution  $\rho$  a alors tendance à se rapprocher de la frontière de  $\mu(E)$  quand  $t$  grandit.

La question de savoir si la solution  $\rho = \rho(t, x)$  reste dans  $\mu(E)$  durant son évolution est une question difficile.

On peut donner une réponse positive à cette question dans le cas homogène par rapport à la variable d'espace  $x$ . Dans ce cas, le système (2.1) devient un problème de Cauchy dans la seule variable  $t$

$$\begin{cases} \frac{d\rho}{dt} = r(\rho) & t \in \mathbb{R}_+ \\ \rho(0) = \rho^0 \end{cases} \quad (2.2)$$

où la donnée initiale est une constante  $\rho^0$  de  $\mu(E)$ .

Le membre de droite  $r(\rho)$  peut être explicité facilement si on utilise par exemple un opérateur  $Q$  de type BGK proposé par [L], les quantités  $n, u, T$  ne dépendant alors pas de  $t$  car elles sont calculées à partir des moments constants  $\rho_0, \rho_1, \rho_2$  et les fonctions  $\rho_3$  et  $\rho_4$  sont solutions d'un système d'équations découplées.

Pour savoir si ces solutions  $\rho = \rho(t)$  restent à l'intérieur du domaine  $\mu(E)$ , on considère les moments  $\rho^*$  normalisés qui s'écrivent alors sous la forme :

$$\begin{aligned} \rho_3^*(t) &= \rho_3^*(0) e^{-\lambda t}, \\ \rho_4^*(t) &= \rho_4^*(0) e^{-\lambda t} + 3(1 - e^{-\lambda t}). \end{aligned}$$

Ainsi dans le plan  $\mu(E^*)$ , le graphe des solutions est une ligne droite joignant l'état initial  $\rho^*(0)$  à l'état  $\hat{\rho}$  de la maxwellienne  $\hat{f}$  qui est atteint quand  $t$  tend vers  $+\infty$ .

D'après la structure de  $\mu(E^*)$ , la solution de (2.2) existe donc pour tout  $t$ .

## Bibliographie.

- [G] H. Grad . *On the kinetic theory of rarefied gases*, Comm. Pure Appl. Math., 2 (1949) p.331-407.
- [J1] M. Junk . *About the domain of definition of Levermore's Five-Moment System*, J. Stat. Phys., 93 (1998).
- [J2] M. Junk . *Maximum entropy for reduced moment problems*. Preprint (1998).
- [L] C. D. Levermore . *Moment Closure Hierarchies for Kinetic Theories*, J. Stat. Phys., 83 (1996) p.1021-1065.

Université de Paris VI  
DEA d'Analyse Numérique

Année 2000-2001

— o —

Transport de mesures et Equation de Boltzmann  
Le théorème de Tanaka

— o —

Mémoire de stage dirigé par Cédric Villani,  
Professeur à l'Ecole Normale Supérieure de Lyon

## Introduction.

Pour un gaz spatialement homogène, la fonction de distribution  $f = f(t, v)$ ,  $v \in \mathbb{R}^N$  solution de l'équation de Boltzmann et correspondant à une donnée initiale  $f_0 = f_0(v)$  de densité  $\rho \geq 0$ , de vitesse moyenne  $u \in \mathbb{R}^N$  et de température  $T \geq 0$ , converge vers la solution Maxwellienne d'équilibre  $M$  correspondant à ces données, définie par

$$M(v) = \frac{\rho}{(2\pi T)^{\frac{N}{2}}} e^{-\frac{|v-u|^2}{2T}}$$

quand le temps  $t$  tend vers  $+\infty$ .

En supposant de plus que le gaz est constitué de particules maxwelliennes, on peut préciser la convergence de  $f(t, \cdot)$  vers  $M$  à l'aide d'une certaine distance sur l'espace des probabilités sur  $\mathbb{R}^N$ , appelée la distance de Wasserstein.

Le théorème de Tanaka présenté ici assure en dimension  $N = 3$  une propriété de contraction de la distance de Wasserstein le long du flot associé à l'équation de Boltzmann et précise que la distance de Wasserstein de deux solutions  $f(t, \cdot)$  et  $g(t, \cdot)$  de l'équation de Boltzmann correspondant à des probabilités initiales  $f_0$  et  $g_0$  de moyenne nulle et de température unité, est une fonction décroissante de  $t$  qui tend vers 0 quand  $t$  tend vers  $+\infty$ .

La méthode de Tanaka développée dans le paragraphe 3 à partir des travaux de [T] et [V2], utilise une approche probabiliste du problème; elle fait appel en particulier à différents résultats de stabilité et de contrôle des solutions de l'équation de Boltzmann présentés dans le paragraphe 1 et à différents résultats algébriques et topologiques de la distance de Wasserstein présentés dans le paragraphe 2, certains de ces résultats auxiliaires étant valables dans un cadre plus large que celui adopté ici.

# 1 - L'équation de Boltzmann.

## 1.1 - Le modèle.

La description statistique d'un gaz raréfié constitué de particules ponctuelles repérées à l'instant  $t \geq 0$  par leur position  $x \in \mathbb{R}^N$  et leur vitesse  $v \in \mathbb{R}^N$  est faite à l'aide d'une fonction de distribution  $f = f(t, x, v)$  dont l'évolution est régie par l'équation de Boltzmann

$$\partial_t f + (v, \nabla_x f) = Q(f, f)$$

dans laquelle le second membre  $Q(f, f)$  appelé terme de collision représente un bilan statistique des collisions entre particules.

Une collision est le résultat d'une interaction microscopique de deux particules qui passent très près l'une de l'autre. Si  $v$  et  $v_*$  désignent les vitesses de deux particules qui viennent juste d'entrer en collision, on note  $v'$  et  $v'_*$  leur vitesse juste après la collision. Puisque les collisions sont élastiques et que toutes les particules sont identiques, les lois de conservation de la quantité de mouvement et de l'énergie cinétique s'écrivent sous la forme

$$\begin{aligned} v' + v'_* &= v + v_*, \\ |v'|^2 + |v'_*|^2 &= |v|^2 + |v_*|^2. \end{aligned}$$

Ces relations sont résolues en introduisant deux paramètres: la direction  $\sigma \in S^{N-1}$  telle que

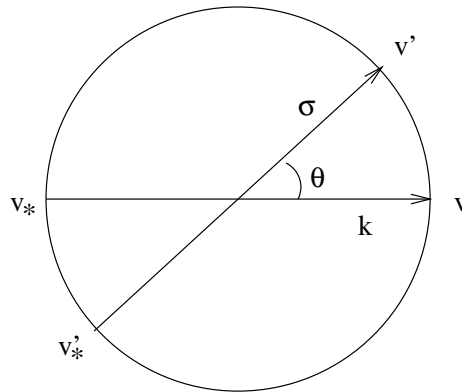
$$\begin{aligned} v' &= \frac{v + v_*}{2} + \frac{|v - v_*|}{2} \sigma, \\ v'_* &= \frac{v + v_*}{2} - \frac{|v - v_*|}{2} \sigma, \end{aligned}$$

et l'angle  $\theta \in [0, \pi]$  de déviation défini par

$$\cos \theta = (k, \sigma)$$

où  $k$  est le vecteur unitaire

$$k = \frac{v - v_*}{|v - v_*|}.$$



Une collision binaire élastique

Le terme de collision est un opérateur quadratique agissant seulement sur la dépendance en  $v$  de la forme

$$Q(f, f) = \int_{\mathbb{R}^N} dv_* \int_{S^{N-1}} B(v - v_*, \sigma) [f' f'_* - f f_*] d\sigma$$

où  $f_*$ ,  $f'$ ,  $f'_*$  désignent respectivement  $f(t, x, v_*)$ ,  $f(t, x, v')$ ,  $f(t, x, v'_*)$  et où  $B$  est une fonction positive appelée la *section efficace* ; cette section efficace tient compte du fait que toutes les valeurs de  $\sigma \in S^{N-1}$  n'interviennent pas de façon identique dans une collision. De plus pour des raisons de symétrie, cette section efficace dépend seulement du module  $|v - v_*|$  de la vitesse relative et du cosinus de l'angle  $\theta$  de déviation défini précédemment, et  $Q$  peut s'écrire sous la forme

$$Q(f, f) = \int_{\mathbb{R}^N} dv_* \int_{S^{N-1}} b(|v - v_*|, (k, \sigma)) [f' f'_* - f f_*] d\sigma.$$

Avec les notations précédentes, on utilisera le changement de variables involutif (dit pré-postcollisionnel)

$$(v, v_*, \sigma) \rightarrow (v', v'_*, k)$$

qui permettra, sous des conditions d'intégrabilité sur une fonction  $F$ , d'écrire

$$\begin{aligned} & \int F(v, v_*, v', v'_*) b(|v - v_*|, (k, \sigma)) dv dv_* d\sigma \\ &= \int F(v, v_*, v', v'_*) b(|v - v_*|, (k, \sigma)) dv' dv'_* dk \\ &= \int F(v', v'_*, v, v_*) b(|v - v_*|, (k, \sigma)) dv dv_* d\sigma. \end{aligned}$$

Si on suppose de plus que la section efficace ne dépend que du cosinus de l'angle de déviation, le terme quadratique de collision s'écrit alors

$$Q(f, f) = \int_{\mathbb{R}^N} dv_* \int_{S^{N-1}} b\left(\left(\frac{v - v_*}{|v - v_*|}, \sigma\right)\right) [f' f'_* - f f_*] d\sigma \quad (1.1)$$

où  $b$  est définie sur  $[-1, +1]$ . Dans ce cas les particules considérées sont dites *Maxwelliennes*. Cette situation modélise par exemple des particules qui interagissent suivant des forces en  $1/r^{2N-1}$ . C'est un modèle théorique pour l'étude d'un gaz constitué d'une seule espèce de particules; par contre il rend compte de l'interaction entre un ion et une particule neutre.

Dans toute la suite on se place dans le cadre particulier où l'on suppose que le gaz est composé de particules Maxwelliennes et est spatialement homogène: sa fonction de distribution  $f$  ne dépend pas de la position  $x$  et se réduit à une mesure  $f(t, v)$  par rapport à  $v$  dans  $\mathbb{R}^N$ , dépendant du temps  $t \geq 0$ . L'équation de Boltzmann s'écrit alors dans ce cas

$$\partial_t f = \int_{\mathbb{R}^N} dv_* \int_{S^{N-1}} b\left(\left(\frac{v - v_*}{|v - v_*|}, \sigma\right)\right) [f' f'_* - f f_*] d\sigma \quad (1.2)$$

pour  $t \in [0, +\infty[$  et  $v \in \mathbb{R}^N$ .

## 1.2 - Le problème de Cauchy.

**Définition 1.1** *L'opérateur de collision  $Q$  est défini au sens faible par*

$$\int Q(f, f) \phi \, dv = \frac{1}{2} \int (\phi' + \phi'_* - \phi - \phi_*) b\left(\left(\frac{v - v_*}{|v - v_*|}, \sigma\right)\right) f f_* \, dv \, dv_* \, d\sigma \quad (1.3)$$

pour  $\phi \in \mathcal{C}_0^\infty(\mathbb{R}^N)$ .

**Remarque.** Grâce au changement de variables pré-post collisionnel, cette définition faible de  $Q$  est formellement équivalente à celle donnée en (1.1), mais elle est intéressante car la quantité (1.3) peut être définie sans que (1.1) le soit.

**Définition 1.2** *L'opérateur de collision étant ainsi défini,  $f$  est une solution faible de l'équation (1.2) de donnée initiale  $f_0$  si pour toute fonction  $\phi(v)$  appartenant à  $\mathcal{C}_0^\infty(\mathbb{R}^N)$  on a*

$$\frac{d}{dt} \int_{\mathbb{R}^N} f \phi \, dv = \int_{\mathbb{R}^N} Q(f, f) \phi \, dv$$

où la dérivation en temps est prise au sens des distributions, et

$$f(0, v) = f_0(v)$$

pour  $v \in \mathbb{R}^N$ .

La donnée  $f_0$  décrit l'état des particules à l'instant  $t = 0$ ; il s'agit d'une fonction ou plus généralement d'une mesure dans  $\mathbb{R}^N$ .

**Convention.** Si  $f(v)$  est une fonction de distribution, les quantités

$$\int_{\mathbb{R}^N} f_0 \, dv \quad \int_{\mathbb{R}^N} f_0 v \, dv \quad \int_{\mathbb{R}^N} f_0 |v|^2 \, dv$$

seront appelées ses *trois premiers moments*.

Concernant l'existence et l'unicité, on a d'après [V3] le

**Théorème 1.3** *Soit  $f_0$  une mesure de probabilité vérifiant les conditions*

$$\int_{\mathbb{R}^N} f_0 \, dv = 1 \quad \int_{\mathbb{R}^N} f_0 v \, dv = 0 \quad \int_{\mathbb{R}^N} f_0 |v|^2 \, dv = N \quad (1.4)$$

et  $b$  une section efficace vérifiant la condition

$$\int_0^\pi b(\cos \theta) (1 - \cos \theta) \sin^{N-2} \theta \, d\theta < +\infty. \quad (1.5)$$

Alors il existe une unique solution faible de l'équation de Boltzmann (1.2) avec donnée initiale  $f_0$ . De plus cette solution conserve ses trois premiers moments.

Si en outre le noyau est intégrable, c'est-à-dire

$$\int_0^\pi b(\cos \theta) \sin^{N-2} \theta \, d\theta < +\infty, \quad (1.6)$$



alors  $f$  est de classe  $C^1$  dans  $[0, +\infty[$  à valeurs dans l'espace des mesures.

**Remarque.** Dans ce cas maxwellien, on note que la solution satisfait automatiquement la propriété de conservation des trois premiers moments. Ce n'est pas vrai de façon générale, car dans certains cas on peut construire des solutions (physiquement non acceptables) qui font croître l'énergie.

On donne maintenant un résultat concernant la stabilité par rapport à la donnée initiale (cf. [V3]).

**Proposition 1.4** *Soit  $b$  une section efficace vérifiant (1.5),  $f_0$  et  $f_{0n}$  pour  $n$  entier  $\geq 0$  des mesures de probabilité vérifiant (1.4),  $f$  et  $f_n$  les solutions de l'équation de Boltzmann (1.2) avec données initiales respectives  $f_0$  et  $f_{0n}$ .*

*Si la suite des données initiales  $(f_{0n})_n$  converge vers  $f_0$  au sens faible\* des mesures de probabilité dans  $\mathbb{R}^N$ , alors pour tout temps  $t \geq 0$  la suite des solutions  $(f_n(t, \cdot))_n$  converge vers  $f(t, \cdot)$  au sens faible\* des mesures de probabilités dans  $\mathbb{R}^N$ .*

On a de plus le résultat de régularisation

**Proposition 1.5** *Si  $M$  est une Maxwellienne et  $f$  est solution de l'équation (1.2) de donnée initiale  $f_0$ , alors  $f * M$  est une fonction de classe  $C^\infty(\mathbb{R}^+ \times \mathbb{R}^N)$ , solution de l'équation pour la donnée initiale  $f_0 * M$ .*

Ce résultat est une conséquence du

**Lemme 1.6 (de Bobylev)** *Pour toute distribution Maxwellienne de probabilité  $M$ , alors*

$$Q(f, f) * M = Q(f * M, f * M).$$

Concernant la stabilité par rapport à la section efficace, on va montrer la

**Proposition 1.7** *Soit  $f_0$  une mesure de probabilité vérifiant (1.4),  $b$  et  $b_n$  pour  $n$  entier  $\geq 0$  des sections efficaces vérifiant (1.5),  $f$  et  $f_n$  les solutions de l'équation de Boltzmann du type (1.2) associées respectivement aux sections efficaces  $b$  et  $b_n$  et de donnée initiale  $f_0$ .*

*Si la suite des sections efficaces  $(b_n(\cos \theta) \sin^{N-2} \theta d\theta)_n$  converge vers la section efficace  $b(\cos \theta) \sin^{N-2} \theta d\theta$  au sens faible\* des mesures dans  $]0, \pi]$  et si de plus la suite  $\left(\int \theta^2 b_n(\cos \theta) \sin^{N-2} \theta d\theta\right)_n$  converge vers  $\int \theta^2 b(\cos \theta) \sin^{N-2} \theta d\theta$ , alors pour tout temps  $t \geq 0$  la suite des solutions  $(f_n(t, \cdot))_n$  converge vers  $f(t, \cdot)$  au sens faible\* des probabilités dans  $\mathbb{R}^N$ .*

**Démonstration.** Les propriétés de conservation de l'équation de Boltzmann assurent que  $\int f_n(t) dv = \int f_0 dv$  et  $\int f_n(t) |v|^2 dv = \int f_0 |v|^2 dv$  pour tout  $t \geq 0$  et tout entier  $n$ . De ces deux propriétés on déduit par le théorème de Prokhorov (cf. [D]) que chaque suite  $(f_n(t))_n$  admet une sous-suite faiblement convergente, c'est-à-dire que pour tout  $t \geq 0$ , il existe une mesure de probabilité  $\tilde{f}(t)$  et une fonction d'extraction  $\psi_t$  telles que

$$f_{\psi_t(n)}(t) \rightharpoonup \tilde{f}(t).$$

Ici la fonction d'extraction dépend de  $t$ , et dans une première étape on va montrer que l'on a ce résultat pour une sous-suite indépendante de  $t$ . On considère pour cela une suite  $(t_j)_j$  de temps dense dans  $\mathbb{R}_+$ . Par extraction diagonale, on déduit du point précédent qu'il existe une sous-suite de la suite  $(f_n)_n$ , que l'on notera encore  $(f_n)_n$ , convergeant faiblement vers  $\tilde{f}$  en tous les  $t_j$ .

Soit maintenant  $\phi$  une fonction test, et montrons que la suite  $(\int f_n(t)\phi dv)_n$  est uniformément (équi)continue en  $t$ . Pour cela on écrit le terme de collision  $Q_n(f_n, f_n)$  associé au noyau  $b_n$  sous la forme

$$\int \phi Q_n(f_n, f_n)(\tau) dv = \frac{1}{2} \int [\phi' + \phi'_* - \phi - \phi_*] f_n f_{n*}(\tau) b_n(\cos \theta) \sin^{N-2} \theta d\theta d\sigma_{N-2} dv dv_*$$

avec  $\sup \int \theta^2 b_n(\cos \theta) \sin^{N-2} \theta < \infty$  et en adaptant la technique de [V1] donnée dans le cas  $N = 3$ , on va d'abord montrer qu'il existe une constante  $C > 0$  telle que pour tous  $\phi \in W^{2,\infty}(\mathbb{R}^N)$  et  $\theta$  on ait

$$|\int_{S^{N-2}} (\phi' + \phi'_* - \phi - \phi_*) d\sigma_{N-2}| \leq C \theta^2 |v - v_*|^2 \|\phi\|_{W^{2,\infty}}.$$

**Remarques.** 1. On montre en fait:

$$|\int_{S^{N-2}} (\phi' + \phi'_* - \phi - \phi_*) d\sigma_{N-2}| \leq C \theta^2 |v - v_*|^2 \|D^2 \phi\|_{L^\infty(B[v, v_*])}$$

où  $B[v, v_*]$  désigne la boule de diamètre  $[v, v_*]$  dans  $\mathbb{R}^N$ .

2. Si  $b$  est une section efficace vérifiant (1.5) et  $g$  est une mesure ayant ses trois premiers moments finis, on en déduit

$$|\int \phi(v) Q(g, g) dv| \leq \int |v - v_*|^2 g g_* dv dv_* \int \theta^2 b(\cos \theta) \sin^{N-2} \theta d\theta \|\phi\|_{W^{2,\infty}} \leq C \|\phi\|_{W^{2,\infty}},$$

ce qui assure que  $Q(g, g)(t) \in W^{-2,1}(\mathbb{R}^N)$ .

**Suite de la démonstration.** En rappelant que l'on a posé

$$v' = \frac{v + v_*}{2} + \frac{|v - v_*|}{2} \sigma \quad v'_* = \frac{v + v_*}{2} - \frac{|v - v_*|}{2} \sigma$$

avec  $\sigma \in S^{N-1}$ ,

$$k = \frac{v - v_*}{|v - v_*|} \quad \cos \theta = (k, \sigma),$$

et en posant maintenant

$$\sigma = k - 2(\omega, k) \omega$$

avec  $\omega \in S^{N-1}$ , on introduit une nouvelle représentation équivalente

$$v' = v - (v - v_*, \omega) \omega \quad v'_* = v_* + (v - v_*, \omega) \omega.$$

Par suite

$$\phi(v') - \phi(v) =$$

$$\begin{aligned}
& -(v - v_*, \omega) (\nabla \phi(v), \omega) + (v - v_*, \omega)^2 \sum_{|\alpha|=2} \int_0^1 \partial^\alpha \phi(v + t(v - v_*, \omega)) \omega^\alpha (1 - t) dt, \\
& \phi(v'_*) - \phi(v_*) = \\
& +(v - v_*, \omega) (\nabla \phi(v_*), \omega) + (v - v_*, \omega)^2 \sum_{|\alpha|=2} \int_0^1 \partial^\alpha \phi(v - t(v - v_*, \omega)) \omega^\alpha (1 - t) dt,
\end{aligned}$$

et de là on a donc pour une constante  $C$  ne dépendant que de  $N$

$$\begin{aligned}
& \left| \int_{S^{N-2}} (\phi' + \phi'_* - \phi - \phi_*) d\sigma_{N-2} \right| \leq \\
& \left| \int_{S^{N-2}} (v - v_*, \omega) (\nabla \phi(v_*) - \nabla \phi(v), \omega) d\sigma_{N-2} \right| + C \|D^2 \phi\|_{L^\infty(B[v, v_*])} \int_{S^{N-2}} (v - v_*, \omega)^2 d\sigma_{N-2}.
\end{aligned}$$

Pour majorer chacun des termes du second membre, on se place en coordonnées sphériques d'axes  $(k_1, k_2, \dots, k_N)$  où on a choisi comme premier axe

$$k_1 = k = \frac{v - v_*}{|v - v_*|}.$$

Alors  $\sigma \in S^{N-1}$  s'écrit sous la forme

$$\sigma = \sum_{i=1}^N x_i k_i$$

où

$$\begin{cases}
x_1 = \cos \theta \\
x_2 = \sin \theta \cos \varphi_1 \\
x_3 = \sin \theta \sin \varphi_1 \cos \varphi_2 \\
\cdots \\
x_{N-1} = \sin \theta \sin \varphi_1 \cdots \sin \varphi_{N-2} \cos \varphi_{N-1} \\
x_N = \sin \theta \sin \varphi_1 \cdots \sin \varphi_{N-2} \sin \varphi_{N-1}
\end{cases}$$

qu'on écrit en abrégé

$$\sigma = \cos \theta k + \sin \theta x.$$

Comme  $\sigma \in S^{N-1}$  et  $\omega \in S^{N-1}$  sont liés par

$$\sigma = k - 2(\omega, k)\omega$$

la décomposition de  $\omega$  suivant  $k$  et  $x$  s'écrit sous la forme

$$\omega = \alpha k + \beta x$$

avec  $\alpha$  et  $\beta$  tels que

$$\cos \theta k + \sin \theta x = k - 2\alpha(\alpha k + \beta x),$$

soit

$$\cos \theta = 1 - 2\alpha^2 \quad \sin \theta = -2\alpha \beta,$$

c'est-à-dire

$$\alpha = \sin \frac{\theta}{2} \quad \beta = -\cos \frac{\theta}{2}.$$

Ainsi

$$\omega = \sin \frac{\theta}{2} k - \cos \frac{\theta}{2} x.$$

En particulier on a donc

$$(v - v_*, \omega) = \sin \frac{\theta}{2} |v - v_*|$$

et par suite

$$\int_{S^{N-2}} (v - v_*, \omega) (\nabla \phi(v_*) - \nabla \phi(v), \omega) d\sigma_{N-2} = \sin \frac{\theta}{2} |v - v_*| \int_{S^{N-2}} (\nabla, \omega) d\sigma_{N-2}$$

où on a noté  $\nabla = \nabla \phi(v_*) - \nabla \phi(v)$ . En décomposant ce vecteur sous la forme  $\nabla = (\nabla, k)k + \nabla^\perp$  où  $\nabla^\perp$  est orthogonal à  $k$ , on a donc

$$(\nabla, \omega) = (\nabla, k) (k, \omega) - \frac{1}{2 \sin \frac{\theta}{2}} (\nabla^\perp, \sum_{i=2}^N x_i k_i).$$

Par suite

$$\begin{aligned} \int_{S^{N-2}} (\nabla, \omega) d\sigma_{N-2} &= \int_{S^{N-2}} (\nabla, k) (k, \omega) d\sigma_{N-2} - \frac{1}{2 \sin \frac{\theta}{2}} \int_{S^{N-2}} (\nabla^\perp, \sum_{i=2}^N x_i k_i) d\sigma_{N-2} \\ &= (\nabla, k) \int_{S^{N-2}} \sin \frac{\theta}{2} d\sigma_{N-2} - \frac{1}{2 \sin \frac{\theta}{2}} (\nabla^\perp, \int_{S^{N-2}} (\sum_{i=2}^N x_i k_i) d\sigma_{N-2}) \\ &= (\nabla, k) \int_{S^{N-2}} \sin \frac{\theta}{2} d\sigma_{N-2} \end{aligned}$$

car  $\int_{S^{N-2}} (\sum_{i=2}^N x_i k_i) d\sigma_{N-2} = 0$  d'après les définitions précédentes des  $x_i$ . Ainsi pour une constante  $C$  ne dépendant que de  $N$  on a

$$|\int_{S^{N-2}} (\nabla, \omega) d\sigma_{N-2}| = C (\nabla, k) \sin \frac{\theta}{2}.$$

Comme de plus  $|(\nabla, k)| \leq \|\nabla\| = \|\nabla \phi(v_*) - \nabla \phi(v)\| \leq |v - v_*| \|\phi\|_{L^\infty(B[v, v_*])}$  et comme on l'a vu précédemment  $(v - v_*, \omega) = \sin \frac{\theta}{2} |v - v_*|$ , on en déduit que

$$|\int_{S^{N-2}} (v - v_*, \omega) (\nabla \phi(v_*) - \nabla \phi(v), \omega) d\sigma_{N-2}| \leq C \sin^2 \frac{\theta}{2} |v - v_*|^2 \|\phi\|_{L^\infty(B[v, v_*])}.$$

Ainsi il existe une constante  $C$  telle que pour tous  $\phi$  et  $\theta$ , on ait

$$|\int_{S^{N-2}} (\phi' + \phi'_* - \phi - \phi_*) d\sigma_{N-2}| \leq C \|\phi\|_{L^\infty(B[v, v_*])} \theta^2 |v - v_*|^2.$$

De là on en déduit alors qu'il existe une constante  $C(\phi)$  dépendant de  $\|\phi\|_{W^{2, \infty}}$  et de  $\sup_n \int \theta^2 b_n(\cos \theta) \sin^{N-2} \theta d\theta (< \infty)$  telle que pour tout  $n$  et pour tout  $\tau$

$$|\int \phi Q_n(f_n, f_n)(\tau) dv| \leq C(\phi) \int f_n(\tau) f_{n*}(\tau) (1 + |v - v_*|^2) dv dv_*.$$

Comme pour tout  $n$ , la fonction  $f_n$  est solution de l'équation de Boltzmann, on a

$$\int f_n(t) \phi dv - \int f_n(s) \phi dv = \int_s^t d\tau \int \phi Q_n(f_n, f_n)(\tau) dv$$

et donc pour différentes constantes  $C(\phi)$  dépendant de  $\|\phi\|_{W^{2,\infty}}$  et de  $\sup_n \int \theta^2 b_n(\cos \theta) \sin^{N-2} \theta d\theta$  on en déduit que pour tous  $n, t$  et  $s$

$$\begin{aligned} \left| \int f_n(t) \phi dv - \int f_n(s) \phi dv \right| &\leq |t - s| C(\phi) \int f_n(\tau) f_{n*}(\tau) (1 + |v - v_*|^2) dv dv_* \\ &\leq C(\phi) |t - s| \left( \int (1 + |v|^2) f_0 dv \right)^2 \\ &\leq C(\phi) |t - s|, \end{aligned}$$

ce qui donne l'uniforme (équi)continuité des fonctions  $\int f_n(t) \phi dv$ .

Cette propriété et la densité de la suite  $(t_j)_j$  permettent alors de montrer qu'à  $t$  fixé, la suite  $(\int f_n(t) \phi dv)_n$  est une suite de Cauchy. En effet pour tous  $t_j, n$  et  $m$ , on peut écrire

$$\begin{aligned} &\left| \int f_n(t) \phi dv - \int f_m(t) \phi dv \right| \leq \\ &\left| \int f_n(t) \phi dv - \int f_n(t_j) \phi dv \right| + \left| \int f_n(t_j) \phi dv - \int f_m(t_j) \phi dv \right| + \left| \int f_m(t_j) \phi dv - \int f_m(t) \phi dv \right| \\ &\leq 2C|t - t_j| + \left| \int f_n(t_j) \phi dv - \int f_m(t_j) \phi dv \right|. \end{aligned}$$

Soit alors  $\varepsilon > 0$  fixé. Il existe un  $j$  tel que  $|t - t_j| \leq \frac{\varepsilon}{4C}$ , puis un  $n_0$  tel que pour tous  $n, m \geq n_0$ ,  $\left| \int f_n(t_j) \phi dv - \int f_m(t_j) \phi dv \right| \leq \frac{\varepsilon}{2}$ , si bien que

$$\left| \int f_n(t) \phi dv - \int f_m(t) \phi dv \right| \leq \varepsilon.$$

Or il existe une sous-suite de cette suite qui converge vers  $\int \tilde{f}(t) \phi dv$  d'après la conséquence du théorème de Prokhorov. Par conséquent toute la suite  $(\int f_n(t) \phi dv)_n$  converge vers  $\int \tilde{f}(t) \phi dv$  et en particulier pour tous  $s$  et  $t$

$$\int f_n(t) \phi dv - \int f_n(s) \phi dv \rightarrow \int \tilde{f}(t) \phi dv - \int \tilde{f}(s) \phi dv.$$

Dans une deuxième étape on va montrer que  $\tilde{f}$  est solution de l'équation. Pour cela il reste à étudier le passage à la limite dans le terme de collision, c'est-à-dire que l'on doit montrer que  $\int_s^t \int [\phi' + \phi'_* - \phi - \phi_*] b_n(\cos \theta) \sin^{N-2} \theta d\theta d\sigma_{N-2} f_n(\tau) f_{n*}(\tau) dv dv_* d\tau$  tend vers  $\int_s^t \int [\phi' + \phi'_* - \phi - \phi_*] b(\cos \theta) \sin^{N-2} \theta d\theta d\sigma_{N-2} \tilde{f}(\tau) \tilde{f}_*(\tau) dv dv_* d(\tau)$  quand  $n$  tend vers  $+\infty$ .

Soit alors  $\varepsilon > 0$  fixé et  $\theta_0$  tel que

$$\int_0^{\theta_0} \theta^2 b(\cos \theta) \sin^{N-2} \theta d\theta \leq \varepsilon.$$

D'une part  $b_n(\cos \theta) \sin^{N-2} \theta d\theta \rightarrow b(\cos \theta) \sin^{N-2} \theta d\theta$  sur  $[\theta_0, \pi]$  si bien que

$$\int_{\theta_0}^{\pi} \theta^2 b_n(\cos \theta) \sin^{N-2} \theta d\theta \rightarrow \int_{\theta_0}^{\pi} \theta^2 b(\cos \theta) \sin^{N-2} \theta d\theta,$$

et d'autre part on a supposé que

$$\int_0^{\pi} \theta^2 b_n(\cos \theta) \sin^{N-2} \theta d\theta \rightarrow \int_0^{\pi} \theta^2 b(\cos \theta) \sin^{N-2} \theta d\theta.$$

Par conséquent, en omettant l'élément d'intégration  $\sin^{N-2} \theta d\theta$ , on a

$$\begin{aligned} & \left| \int_0^{\theta_0} \theta^2 b_n(\cos \theta) \right| \leq \\ & \left| \int_0^{\pi} \theta^2 b_n(\cos \theta) - \int_0^{\pi} \theta^2 b(\cos \theta) \right| + \left| \int_0^{\theta_0} \theta^2 b(\cos \theta) \right| + \left| \int_{\theta_0}^{\pi} \theta^2 b(\cos \theta) - \int_{\theta_0}^{\pi} \theta^2 b_n(\cos \theta) \right| \end{aligned}$$

qui est majoré par  $3\varepsilon$  pour  $n$  assez grand.

En notant  $\Phi(v, v_*, \theta) = \int [\phi' + \phi'_* - \phi - \phi_*] d\sigma_{N-2}$ , on a d'après l'étape précédente  $|\Phi| \leq C \inf(\theta^2 |v - v_*|^2, 1)$ .

Donc pour une constante  $C$  indépendante de  $n$  et de  $\tau$

$$\left| \int \Phi f_n f_{n*} dv dv_* \right| \leq C \int \theta^2 |v - v_*|^2 f_n f_{n*} dv dv_* \leq C\theta^2 \quad (1.7)$$

et par suite pour  $n$  assez grand

$$\left| \int_0^{\theta_0} b_n(\cos \theta) \sin^{N-2} \theta d\theta \int \Phi f_n f_{n*} dv dv_* \right| \leq C\varepsilon.$$

De même

$$\left| \int_0^{\theta_0} b(\cos \theta) \sin^{N-2} \theta d\theta \int \Phi \tilde{f} \tilde{f}_* dv dv_* \right| \leq C\varepsilon.$$

D'autre part  $f_n f_{n*} b_n(\cos \theta) \sin^{N-2} \theta d\theta \rightarrow \tilde{f} \tilde{f}_* b(\cos \theta) \sin^{N-2} \theta d\theta$  sur  $[\theta_0, \pi] \times \mathbb{R}^3 \times \mathbb{R}^3$  comme produit tensoriel de mesures convergeant faiblement (cf. [S] par exemple) et  $(v, v_*, \theta) \mapsto \Phi(v, v_*, \theta)$  est continue et bornée, si bien que

$$\int_{\theta_0}^{\pi} b_n(\cos \theta) \sin^{N-2} \theta d\theta \int \Phi f_n f_{n*} dv dv_* \rightarrow \int_{\theta_0}^{\pi} b(\cos \theta) \sin^{N-2} \theta d\theta \int \Phi \tilde{f} \tilde{f}_* dv dv_*.$$

Ainsi

$$\int_0^{\pi} b_n(\cos \theta) \sin^{N-2} \theta d\theta \int \Phi f_n(\tau) f_{n*}(\tau) dv dv_* \rightarrow \int_0^{\pi} b(\cos \theta) \sin^{N-2} \theta d\theta \int \Phi \tilde{f}(\tau) \tilde{f}_*(\tau) dv dv_*.$$

Comme de plus ces intégrales sont toutes majorées par une constante indépendante de  $\tau$  et de  $n$  d'après (1.7), on en déduit que la suite des intégrales en  $\tau$  sur  $]s, t[$  converge et de

là que  $\tilde{f}$  est solution faible de l'équation de Boltzmann. Le résultat d'unicité du théorème 1.3 assure alors que  $\tilde{f} = f$ .

Enfin étant donnée une sous-suite quelconque de  $(f_n)_n$ , le raisonnement précédent montre qu'elle contient une sous-suite qui converge vers  $f (= \tilde{f})$ , donc toute la suite  $(f_n)_n$  converge vers  $f$ .

On a ainsi approché au sens faible la solution  $f$  de l'équation associée à la section efficace  $b$  par la suite des solutions  $f_n$  de l'équation associée à la section  $b_n$  pour la même donnée initiale  $f_0$  •

### 1.3 - Equilibre.

Dans le cas d'un gaz en équilibre thermodynamique, la solution distribution  $f$  est donnée par une Maxwellienne qui, compte tenu des conditions (1.4) imposées aux trois premiers moments par rapport à 1,  $v$  et  $|v|^2$  se réduit à la Maxwellienne  $M$  suivante qui est la Gaussienne classique centrée en 0

$$M(v) = \frac{1}{(2\pi)^{\frac{N}{2}}} e^{-\frac{|v|^2}{2}}. \quad (1.8)$$

Soit  $f(t, \cdot)$  la solution de l'équation de Boltzmann associée à une probabilité initiale  $f_0$  de vitesse moyenne 0, de température 1 et d'entropie  $H(f_0)$  finie. Les trois premiers moments sont conservés au cours de l'évolution, et son entropie  $H(f(t))$  décroît. Comme la Maxwellienne minimise l'entropie parmi les distributions de mêmes trois premiers moments, on peut espérer la convergence de la solution  $f(t)$  vers cette Maxwellienne  $M$  quand le temps  $t$  tend vers  $+\infty$ .

Une question importante est de préciser le sens de la convergence de  $f(t, \cdot)$  vers  $M$ . C'est ce qui sera fait dans le paragraphe 3 dans le cas particulier des particules maxwelliennes, c'est-à-dire pour l'équation (1.2).

On donne tout d'abord un résultat concernant le comportement en temps des moments de la solution  $f$ ; en dimension  $N = 3$  on a la

**Proposition 1.8** *Soient  $f_0$  une mesure de probabilité sur  $\mathbb{R}^3$  vérifiant (1.4) et dont les moments d'ordre  $P$  sont finis, et  $f$  la solution de l'équation de Boltzmann (1.2) dans  $\mathbb{R}^3$  associée à  $f_0$ . Alors tous ses moments jusqu'à l'ordre  $P$  convergent exponentiellement vers ceux de la Maxwellienne  $M$  définie en (1.8).*

**Démonstration.** Pour cela on introduit comme dans [IT] la notion de moments harmoniques associés à des polynômes harmoniques dans  $\mathbb{R}^3$ .

Pour tout entier  $k \geq 0$ , on choisit  $2k+1$  polynômes harmoniques de degré  $k$ , linéairement indépendants et homogènes, notés  $(\zeta_k^l)_{|l| \leq k}$  (cf. [DL] par exemple). Pour  $n = (r, k, l)$  avec  $r = 0, 1, \dots$ ,  $k = 0, 1, \dots$  et  $|l| \leq k$ , on pose alors  $\zeta_n(v) = |v|^{2r} \zeta_k^l(v)$  avec en particulier  $\zeta_{(0,0,0)}(v) = 1$ . On note que  $\zeta_n$  est un polynôme homogène de degré  $|n| = 2r + k$ .

Etant donnée une mesure  $h$ , on définit ses moments harmoniques par

$$\mu_n(h) = \int \zeta_n(v) h(v) dv.$$

Par les propriétés des polynômes harmoniques, pour chaque  $n$  tel que  $|n| \leq P$ , il existe d'après [IT] des coefficients  $\beta_{n_1, n_2}^n$  pour les paires  $(n_1, n_2)$  telles que  $|n_1| + |n_2| = |n|$  et  $|n_1|, |n_2| \geq 1$ , et un coefficient  $\beta_n > 0$  tels que

$$\frac{d}{dt} \mu_n(h(t)) = \sum \beta_{n_1, n_2}^n \mu_{n_1}(h(t)) \mu_{n_2}(h(t)) - \beta_n \mu_n(h(t))$$

pour toute solution  $h(t)$  de l'équation de Boltzmann dont la condition initiale vérifie la condition (1.4).

A partir de cette formule, on démontre alors par récurrence sur  $|n|$  pour  $|n| \leq P$  que si  $M$  est la Maxwellienne définie précédemment en (1.8), il existe des constantes  $C_n$  et  $k_n > 0$  telles que pour tout  $t \geq 0$  on ait

$$|\mu_n(f(t)) - \mu_n(M)| \leq C_n e^{-k_n t}.$$

En effet tout d'abord pour  $n = (0, 0, 0)$  le résultat est vrai puisque  $\mu_0(f(t)) = \mu_0(M) = 1$ . Ensuite étant donné un  $n$  avec  $1 \leq |n| < P$ , on suppose le résultat vrai pour tout  $m$  tel que  $|m| \leq |n|$  et on le montre pour  $n$ .

Par la formule précédente appliquée aux solutions  $f$  et  $M$  de l'équation de Boltzmann dont les conditions initiales ont les mêmes trois premiers moments 1, 0 et 3, on a par différence

$$\begin{aligned} & \frac{d}{dt} (\mu_n(f(t)) - \mu_n(M)) = \\ & \sum \beta_{n_1, n_2}^n (\mu_{n_1}(f(t)) \mu_{n_2}(f(t)) - \mu_{n_1}(M) \mu_{n_2}(M)) - \beta_n (\mu_n(f(t)) - \mu_n(M)) \end{aligned}$$

soit

$$\frac{d}{dt} (e^{\beta_n t} (\mu_n(f(t)) - \mu_n(M))) = e^{\beta_n t} \sum \beta_{n_1, n_2}^n (\mu_{n_1}(f(t)) \mu_{n_2}(f(t)) - \mu_{n_1}(M) \mu_{n_2}(M)).$$

Or par hypothèse de récurrence pour tout  $j$  tel que  $|n_j| < |n|$ , il existe des constantes  $C_{n_j}$  et  $k_{n_j} > 0$  telles que pour  $t \geq 0$  on ait

$$|\mu_{n_j}(f(t)) - \mu_{n_j}(M)| \leq C_{n_j} e^{-k_{n_j} t}.$$

On note  $K_n = \inf_{|n_j| < |n|} k_{n_j}$ , puis  $L_n = \inf(K_n, \frac{\beta_n}{2})$ , si bien que  $0 < L_n < \beta_n$  et pour une constante  $\alpha_n$

$$\left| \frac{d}{dt} (e^{\beta_n t} (\mu_n(f(t)) - \mu_n(M))) \right| \leq \alpha_n e^{(\beta_n - L_n)t}.$$

De là par intégration:

$$\begin{aligned} e^{\beta_n t} (\mu_n(f(t)) - \mu_n(M)) - (\mu_n(f(0)) - \mu_n(M)) & \leq \alpha_n \int_0^t e^{(\beta_n - L_n)s} ds \\ & = \frac{\alpha_n}{\beta_n - L_n} (e^{(\beta_n - L_n)t} - 1), \end{aligned}$$

soit

$$\mu_n(f(t)) - \mu_n(M) \leq (\mu_n(f(0)) - \mu_n(M)) e^{-\beta_n t} + \frac{\alpha_n}{\beta_n - L_n} (e^{-L_n t} - e^{-\beta_n t}),$$



d'où pour des constantes  $C_n$  et  $k_n > 0$  on a

$$|\mu_n(f(t)) - \mu_n(M)| \leq C_n e^{-k_n t}.$$

Comme tout polynôme homogène de degré  $P$  peut se décomposer en une combinaison linéaire des  $(\zeta_n)_{|n|=P}$ , on en déduit en particulier le résultat annoncé •

**Corollaire 1.9** *Soit  $f_0$  une mesure de probabilité sur  $\mathbb{R}^3$  vérifiant (1.4) et dont tous les moments sont finis. Alors la solution  $f(t)$  de l'équation de Boltzmann (1.2) dans  $\mathbb{R}^3$  de donnée initiale  $f_0$  converge, quand  $t$  tend vers  $+\infty$ , vers  $M$  au sens faible\* des mesures, où  $M$  est la Maxwellienne ayant les mêmes trois premiers moments que  $f_0$*

**Corollaire 1.10** *Soit  $f_0$  une mesure de probabilité sur  $\mathbb{R}^3$  vérifiant (1.4) et dont les moments d'ordre  $P$  sont finis. Alors la solution  $f$  de l'équation de Boltzmann (1.2) associée à  $f_0$  est telle que*

$$\sup_{t \geq 0} \int_{\mathbb{R}^3} f(t, v) |v|^P dv < +\infty.$$

**Remarque.** Ce résultat peut être montré directement (cf. [E] par exemple).

On aura besoin de la variante suivante du corollaire 1.9:

**Proposition 1.11** *Etant données une mesure de probabilité  $g_0$  dans  $\mathbb{R}^3$  vérifiant (1.4) et telle que*

$$\int_{\mathbb{R}^3} |v|^{13} g_0 dv < +\infty$$

*et une Maxwellienne  $G$  de masse 1, soit  $f_0$  la mesure de probabilité définie par*

$$f_0 = g_0 * G$$

*et  $f(t)$  la solution de l'équation de Boltzmann (1.2) de donnée initiale  $f_0$ .*

*Alors  $f(t)$  converge vers  $M$  au sens faible\* des probabilités sur  $\mathbb{R}^3$  quand  $t$  tend vers  $+\infty$ , où  $M$  est la Maxwellienne ayant les mêmes trois premiers moments que  $f_0$ .*

**Remarque.** La condition imposée sur les moments de  $g_0$  peut facilement être affaiblie.

**Démonstration.** On note tout d'abord que les résultats concernant l'existence et l'unicité (théorème 1.3) et la stabilité par rapport à la donnée initiale (proposition 1.4) restent vrais lorsque les trois premiers moments de  $f_0$  sont seulement supposés finis, ce qui est bien le cas pour la donnée initiale  $f_0 = g_0 * G$ .

D'après la proposition 1.5, la solution  $f$  est une fonction de classe  $\mathcal{C}^\infty(\mathbb{R}^+ \times \mathbb{R}^3)$  et est donnée par la convolution en  $v$  dans  $\mathbb{R}^3$  sous la forme

$$f(t) = g(t) * G$$

où  $g(t)$  est la solution de l'équation de Boltzmann (1.2) de donnée initiale  $g_0$ .

De la conservation des trois premiers moments de la solution  $f(t)$  on déduit par le théorème de Prokhorov (cf [D]) que toute suite  $(t_n)_n$  tendant vers  $+\infty$  contient une sous-suite, encore notée  $(t_n)_n$  pour simplifier, telle que la suite des probabilités  $(f(t_n))_n$  converge au sens faible\* vers une probabilité notée  $f_\infty$ . Cette convergence a lieu en particulier au sens des distributions dans  $\mathbb{R}^3$ .

Pour  $n \in \mathbb{N}$  et  $t \geq 0$  on pose

$$f_n(t) = f(t_n + t)$$

et on note  $f_\infty(t)$  la solution de l'équation de Boltzmann (1.2) de donnée initiale  $f_\infty$ . La proposition 1.4 sur la stabilité par rapport à la donnée initiale assure en particulier que pour tout temps  $t \geq 0$  la suite  $(f_n(t))_n$  converge vers  $f_\infty(t)$  au sens faible\*, donc au sens des distributions dans  $\mathbb{R}^3$ .

**1.** On précise d'abord la régularité de  $f$  et de  $f_\infty$ .

**1.1.** Tout d'abord  $f \in \mathcal{C}^\infty(\mathbb{R}^+ \times \mathbb{R}^3)$  et  $\int f(t, v) dv = 1$ , donc  $f \in L^1([0, T] \times \mathbb{R}^3)$  pour tout  $T > 0$ .

**1.2** Ensuite pour tout entier  $m$  et tout  $t \geq 0$ , alors  $f(t) \in H_1^m(\mathbb{R}^3)$  où  $H_1^m(\mathbb{R}^3) := \{u \in L^2(\mathbb{R}^3), v^\gamma \partial_v^\delta u \in L^2(\mathbb{R}^3), |\gamma| \leq 1, |\delta| \leq m\}$  muni de sa norme naturelle, avec

$$\sup_{t \geq 0} \|f(t)\|_{H_1^m(\mathbb{R}^3)} < +\infty.$$

En effet pour des constantes  $\alpha$  et  $\beta > 0$

$$f(t, v) = \alpha \int_{\mathbb{R}^3} e^{-\beta|v-y|^2} dg(t, y), \quad (1.9)$$

donc par dérivation sous le signe d'intégration, pour tous  $\gamma, \delta \in \mathbb{N}^3$  il existe un polynôme  $P_\delta$  tel que

$$v^\gamma \partial_v^\delta f(t, v) = v^\gamma \int P_\delta(v - y) e^{-\beta|v-y|^2} dg(t, y).$$

Comme  $g(t)$  est une mesure de probabilité, l'inégalité de Jensen permet d'écrire que

$$|v^\gamma \partial_v^\delta f(t, v)|^2 \leq \int |v|^{2|\gamma|} P_\delta^2(v - y) e^{-2\beta|v-y|^2} dg(t, y),$$

puis par le théorème de Fubini, on a pour tout  $t \geq 0$

$$\begin{aligned} \int_{\mathbb{R}^3} |v^\gamma \partial_v^\delta f(t, v)|^2 &\leq \int_{\mathbb{R}^3} \left( \int_{\mathbb{R}^3} |v|^{2|\gamma|} P_\delta^2(v - y) e^{-2\beta|v-y|^2} dg(t, y) \right) dv \\ &\leq 2 \int_{\mathbb{R}^3} |w|^{2|\gamma|} P_\delta^2(w) e^{-2|w|^2} dw \int_{\mathbb{R}^3} dg(t, y) \\ &\quad + 2 \int_{\mathbb{R}^3} P_\delta^2(w) e^{-2|w|^2} dw \int_{\mathbb{R}^3} |y|^{2|\gamma|} dg(t, y) \end{aligned}$$

qu'on peut majorer par une constante finie indépendante de  $t$  puisque les moments d'ordre  $\leq 2$  de  $g(t)$  sont constants.

**1.3.** On en déduit que

$$\sup_{t \geq 0} \int_{\mathbb{R}^3} f(t, v) (1 + |v|^2 + |\log f(t, v)|) dv < +\infty.$$

En effet par l'inégalité de Jensen appliquée à la formule (1.9)

$$\log f(t, v) \geq \log \alpha - \int_{\mathbb{R}^3} \beta |v - y|^2 dg(t, y)$$

et comme la solution  $g(t)$  a les mêmes trois premiers moments que sa condition initiale  $g_0$ , il reste

$$\log f(t, v) \geq \log \alpha - \beta|v|^2 - 3\beta.$$

Ainsi pour une constante  $C$  on a pour tous  $t$  et  $v$

$$f(t, v) \geq \log f(t, v) \geq -C(1 + |v|^2),$$

donc

$$f(t, v)^2 \geq f(t, v) \log f(t, v) \geq -Cf(t, v)(1 + |v|^2)$$

et par suite

$$|f(t, v) \log f(t, v)| \leq f(t, v)^2 + Cf(t, v)(1 + |v|^2).$$

Comme la solution  $f(t)$  a les mêmes trois premiers moments que sa donnée initiale  $f_0$  (qui sont finis) et comme  $\sup_{t \geq 0} \|f(t)\|_{L^2(\mathbb{R}^3)} < +\infty$  d'après 1.1, on a l'estimation annoncée.

**1.4.** On déduit aussi que

$$\sup_{t \geq 0} \int_{\mathbb{R}^3} f_\infty(t, v)(1 + |v|^2 + |\log f_\infty(t, v)|) dv < +\infty.$$

Pour cela on montre par exemple que l'injection de  $H_1^1(\mathbb{R}^3)$  dans  $L^2(\mathbb{R}^3)$  est compacte. En effet soit  $(u_n)_n$  une suite bornée de  $H_1^1(\mathbb{R}^3)$ . Elle contient une sous-suite, encore notée  $(u_n)_n$ , qui converge faiblement dans  $H_1^1(\mathbb{R}^3)$  vers une fonction  $u$  que l'on peut supposer nulle par différence. Pour montrer que cette convergence est en fait forte dans  $L^2(\mathbb{R}^3)$ , on décompose

$$\begin{aligned} \|u_n\|_{L^2(\mathbb{R}^3)} &\leq \|u_n\|_{L^2(B_R)} + \|u_n\|_{L^2(\mathbb{R}^3 \setminus B_R)} \\ &\leq \|u_n\|_{L^2(B_R)} + \frac{C}{R} \end{aligned}$$

où  $B_R = \{v \in \mathbb{R}^3, |v| \leq R\}$  et  $C = \sup_n \|u_n\|_{H_1^1(\mathbb{R}^3)}$ .

Etant donné  $\varepsilon > 0$ , on fixe  $R$  tel que  $\frac{C}{R} < \varepsilon$ . L'injection de  $H^1(B_R)$  dans  $L^2(B_R)$  étant compacte et  $(u_n)_n$  convergeant faiblement vers 0 dans  $H^1(B_R)$ , elle converge fortement vers 0 dans  $L^2(B_R)$ . D'où l'on déduit le résultat.

Par conséquent pour tout  $t \geq 0$  fixé, la suite  $(f_n(t))_n$  étant bornée dans  $H_1^1(\mathbb{R}^3)$ , elle contient une sous-suite qui converge dans  $L^2(\mathbb{R}^3)$  vers une limite  $\tilde{f}(t)$ . En particulier cette sous-suite converge vers  $\tilde{f}(t)$  au sens des distributions dans  $\mathbb{R}^3$ . Or  $(f_n(t))_n$  converge au sens faible\* des probabilités, donc des distributions dans  $\mathbb{R}^3$  vers  $f_\infty(t)$ , donc  $\tilde{f}(t) = f_\infty(t)$ . Quitte à extraire une nouvelle sous-suite, cette convergence a lieu presque partout dans  $\mathbb{R}^3$ . Or d'après 1.3 on a

$$\sup_{t \geq 0} \sup_{n \geq 0} \int_{\mathbb{R}^3} f_n(t, v)(1 + |v|^2 + |\log f_n(t, v)|) dv < +\infty,$$

donc à la limite

$$\sup_{t \geq 0} \int_{\mathbb{R}^3} f_\infty(t, v)(1 + |v|^2 + |\log f_\infty(t, v)|) dv < +\infty.$$

**1.5.** On déduit enfin de 1.2 que  $f_\infty(t) \in C^\infty(\mathbb{R}^3)$  pour tout  $t \geq 0$ .

Soit en effet  $t \geq 0$  et  $m \in \mathbb{N}$  fixés. Comme d'après 1.1 on a par exemple

$$\sup_{n \geq 0} \|f_n(t)\|_{H^m(\mathbb{R}^3)} < +\infty,$$

la suite  $(f_n(t))_n$  est bornée dans  $H^m(\mathbb{R}^3)$ , et par conséquent contient une sous-suite qui converge faiblement vers une fonction dans  $H^m(\mathbb{R}^3)$ . Comme précédemment il en résulte que  $f_\infty(t) \in H^m(\mathbb{R}^3)$ .

Ceci étant vrai pour tout entier  $m$ , on a le résultat annoncé.

**2.** On montre maintenant que pour tout  $T > 0$ , la suite  $(F_n)_n$  des produits  $F_n(t, v, v_*) := f_n(t, v) f_n(t, v_*)$  converge faiblement dans  $L^1(]0, T[ \times \mathbb{R}^3 \times \mathbb{R}^3)$  vers le produit  $F_\infty(t, v, v_*) := f_\infty(t, v) f_\infty(t, v_*)$ .

**2.1.** Tout d'abord  $(F_n)_n$  converge vers  $F_\infty$  au sens des distributions dans  $]0, T[ \times \mathbb{R}^3 \times \mathbb{R}^3$ .

En effet soit  $\varphi \in C_0^\infty(]0, T[ \times \mathbb{R}^3 \times \mathbb{R}^3)$  fixée. Pour tout  $t \geq 0$ , on a la convergence simple

$$\int_{\mathbb{R}^3 \times \mathbb{R}^3} f_n(t, v) f_n(t, v_*) \varphi(t, v, v_*) dv dv_* \rightarrow \int_{\mathbb{R}^3 \times \mathbb{R}^3} f_\infty(t, v) f_\infty(t, v_*) \varphi(t, v, v_*) dv dv_*$$

puisque la suite  $(f_n(t))_n$  convergeant vers  $f_\infty(t)$  au sens faible\* des probabilités sur  $\mathbb{R}^3$ , la suite des produits tensoriels  $(f_n(t, v) f_n(t, v_*))_n$  converge donc vers le produit tensoriel  $f_\infty(t, v) f_\infty(t, v_*)$  (cf. [S] par exemple), et pour tous  $t \geq 0$  et  $n \in \mathbb{N}$  on a la domination

$$\left| \int_{\mathbb{R}^3 \times \mathbb{R}^3} f_n(t, v) f_n(t, v_*) \varphi(t, v, v_*) dv dv_* \right| \leq \sup |\varphi|.$$

Le théorème de Lebesgue permet alors de conclure.

**2.2.** De plus la suite  $(F_n)_n$  est relativement compacte dans  $L^1(]0, T[ \times \mathbb{R}^3 \times \mathbb{R}^3)$  pour la topologie faible  $\sigma(L^1, L^\infty)$ .

Pour cela on utilise le théorème de Dunford-Pettis (cf. par exemple [Gr]) qui donne une condition suffisante (et nécessaire) pour que la suite  $(F_n)_n$  soit faiblement relativement compacte dans  $L^1(]0, T[ \times \mathbb{R}^3 \times \mathbb{R}^3)$ . Cette condition suffisante est en particulier vérifiée (cf. par exemple [Ge]) s'il existe  $G \in \mathcal{C}(\mathbb{R}^+, \mathbb{R}^+)$  et  $w \in L_{loc}^\infty(\mathbb{R}^7, \mathbb{R}^+)$  telles que  $\frac{G(s)}{s}$  tend vers  $+\infty$  quand  $t$  tend vers  $+\infty$  et  $w(t, v, v_*)$  tend vers  $+\infty$  quand  $|(t, v, v_*)|$  tend vers  $+\infty$ , et pour lesquelles la relation

$$\sup_n \int G(|F_n(t, v, v_*)|) + |F_n(t, v, v_*)|(1 + w(t, v, v_*)) dt dv dv_* < +\infty$$

est vérifiée.

Dans le cas présent on prend  $G(s) = s \log s$  et  $w(t, v, v_*) = (1 + t^2 + |v|^2)(1 + t^2 + |v_*|^2)$  sur  $]0, T[ \times \mathbb{R}^3 \times \mathbb{R}^3$  et il s'agit de vérifier que

$$\sup_n \int_0^T \int_{\mathbb{R}^3 \times \mathbb{R}^3} f_n(t, v) f_n(t, v_*) \left( |\log f_n(t, v)| + |\log f_n(t, v_*)| + (1 + t^2 + |v|^2)(1 + t^2 + |v_*|^2) \right) dt dv dv_*$$

est fini. Pour cela on examine chaque terme. Ainsi

$$\sup_n \int_0^T \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} f_n(t, v) f_n(t, v_*) |\log f_n(t, v)| dt dv dv_*$$

$$\begin{aligned} &\leq T \sup_n (\sup_t \int_{\mathbb{R}^3} f_n(t, v) |\log f_n(t, v)| dv) (\sup_t \int_{\mathbb{R}^3} f_n(t, v_*) dv_*) \\ &= T \sup_n (\sup_t \int_{\mathbb{R}^3} f_n(t, v) |\log f_n(t, v)| dv) < +\infty \end{aligned}$$

d'après 1-3.

De même

$$\sup_n \int_0^T \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} f_n(t, v) f_n(t, v_*) |\log f_n(t, v_*)| dt dv dv_* < +\infty.$$

Enfin

$$\begin{aligned} &\sup_n \int_0^T \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} f_n(t, v) f_n(t, v_*) (1 + t^2 + |v|^2) (1 + t^2 + |v_*|^2) dt dv dv_* \\ &\leq T \sup_n (\sup_t \int_{\mathbb{R}^3} f_n(t, v) (1 + T^2 + |v|^2) dv) (\sup_t \int_{\mathbb{R}^3} f_n(t, v_*) (1 + T^2 + |v_*|^2) dv_*) < +\infty \end{aligned}$$

d'après 1-3.

**2.3.** On en déduit que  $(F_n)_n$  converge vers  $F_\infty$  dans  $L^1([0, T] \times \mathbb{R}^3 \times \mathbb{R}^3)$  faible.

En effet grâce à 2.2 et au théorème de Smulian (cf. [Gr] par exemple), de toute sous-suite de  $(F_n)_n$  on extrait par compacité faible une sous-suite qui converge faiblement dans  $L^1([0, T] \times \mathbb{R}^3 \times \mathbb{R}^3)$ , donc dans  $\mathcal{D}'([0, T] \times \mathbb{R}^3 \times \mathbb{R}^3)$ . Par conséquent d'après 2.1, cette limite est  $F_\infty$  et ne dépend pas de la sous-suite. D'où le résultat.

**3.** D'après 1.3, l'entropie  $H(f(t))$  de la solution  $f(t)$  est bien définie (et finie) pour tout  $t \geq 0$ . On montre alors que l'application  $t \mapsto H(f(t)) = \int_{\mathbb{R}^3} f(t, v) \ln f(t, v) dv$  est dérivable sur  $[0, +\infty[$ , de dérivée

$$\frac{dH(f(t))}{dt} = -D(f(t))$$

avec

$$D(f(t)) = \frac{1}{4} \int_{\mathbb{R}^3 \times \mathbb{R}^3 \times S^2} [f' f'_* - f f_*] [\ln f' f'_* - \ln f f_*] b\left(\frac{v - v_*}{|v - v_*|}, \sigma\right) dv dv_* d\sigma.$$

**3.1.** On montre que pour cela il suffit d'établir l'existence d'une fonction  $g$  sur  $\mathbb{R}^3$  telle que

$$\int_{\mathbb{R}^3} (1 + |v|^2) g(v) dv < +\infty$$

et

$$|Q(f, f)|(t, v) \leq g(v)$$

pour tous  $t \geq 0$  et  $v \in \mathbb{R}^3$ .

En effet pour tout  $v \in \mathbb{R}^3$  fixé,  $f(\cdot, v)$  est dérivable et  $f(t, v) > 0$  pour tout  $t$ , donc  $f(\cdot, v) \ln f(\cdot, v)$  est dérivable, avec

$$\frac{\partial}{\partial t} [f(t, v) \ln f(t, v)] = \frac{\partial f}{\partial t} [1 + \ln f] = Q(f, f) [1 + \ln f].$$

Or d'après 1.3,

$$|1 + \ln f(t, v)| \leq C[1 + |v|^2]$$

pour une constante  $C$  indépendante de  $t$ . Admettant alors la domination

$$|Q(f, f)|(t, v) \leq g(v)$$

avec

$$\int_{\mathbb{R}^3} (1 + |v|^2) g(v) dv < +\infty,$$

on obtient

$$\left| \frac{\partial}{\partial t} [f(t, v) \ln f(t, v)] \right| \leq (1 + |v|^2) g(v)$$

qui est une fonction de  $L^1(\mathbb{R}^3)$ , et ce pour tout  $t \geq 0$ . Cette majoration permet alors de dériver sous le signe d'intégration et d'écrire

$$\frac{dH(f(t))}{dt} = \int Q(f, f) \ln f(t, v) dv = -D(f(t)).$$

**3.2.** Dans cette partie, on utilisera parfois la notation

$$m_v(y) = \alpha e^{-\beta|v-y|^2}$$

pour différentes constantes  $\alpha$  et  $\beta$ .

Ainsi l'identité

$$\left(\frac{\beta}{\pi}\right)^{3/2} e^{-\beta|v-y|^2} = \int_{\mathbb{R}^3} \left(\frac{\gamma}{\pi}\right)^{3/2} e^{-\gamma|v-z|^2} \left(\frac{\delta}{\pi}\right)^{3/2} e^{-\delta|z-y|^2} dz$$

avec

$$\frac{1}{\beta} = \frac{1}{\gamma} + \frac{1}{\delta}$$

sera-t-elle écrite sous la forme

$$m_v(y) = \int_{\mathbb{R}^3} m_v(z) m_z(y) dz.$$

Avec cette notation, le lemme 1.6 (de Bobylev) sécrit:

$$\begin{aligned} Q(f, f)(v) &= \int_{\mathbb{R}^3} m_v(w) Q(g, g) dw \\ &= \int [m_v(w') + m_v(w'_*) - m_v(w) - m_v(w_*)] b g(w) g(w_*) dw dw_* d\sigma \\ &= \int [\int m_v(y) m_y(w') dy + \int m_v(y) m_y(w'_*) dy - \int m_v(y) m_y(w) dy \\ &\quad - \int m_v(y) m_y(w_*) dy] b d\sigma g g_* dw dw_* \\ &= \int_{\mathbb{R}^3} m_v(y) \int [m_y(w') + m_y(w'_*) - m_y(w) - m_y(w_*)] b g g_* dw dw_* d\sigma dy \\ &= \int_{\mathbb{R}^3} m_v(y) h_t(y) dy \end{aligned}$$

où l'on note

$$h_t(y) = \int_{\mathbb{R}^3} e^{-\delta|y-w|^2} Q(g, g)(w) dy$$

dans le sens donné par la formulation faible (1.3) de l'opérateur de collision.

**3.3.** La fonction  $h_t$  est continue bornée en  $y$ , uniformément en temps. De plus:

$$\sup_{t \geq 0} \int_{\mathbb{R}^3} |y|^6 |h_t(y)| dy < \infty.$$

On a ainsi réécrit la convolution

$$Q(f, f)(v) = \int_{\mathbb{R}^3} m_v(w) Q(g, g) dw$$

sous la forme

$$Q(f, f)(v) = \int_{\mathbb{R}^3} m_v(y) h_t(y) dy,$$

c'est-à-dire remplacé la distribution  $Q(g, g)$  par la fonction  $h_t$ , qui est régulière et a suffisamment de moments finis.

La remarque faite lors de la démonstration de la proposition 1.7 assure que  $Q(g, g) \in W^{-2,1}(\mathbb{R}^3)$ , d'où l'on déduit que  $h_t \in \mathcal{C}(\mathbb{R}^3)$ .

Comme  $\sup_{y \in \mathbb{R}^3} \|D^2 m_y\|_{L^\infty}(y)$ ,  $\int \theta^2 b(\cos \theta) \sin \theta d\theta$  et  $\sup_t \int_{\mathbb{R}^3} (1 + |v|^2) g(v) dv$  sont tous trois finis, on déduit de la même remarque que  $h_t$  est bien bornée sur  $\mathbb{R}^3$ , et ce uniformément en  $t$ .

Cette remarque assure plus précisément que

$$\begin{aligned} & \left| \int m_y(w') + m_y(w'_*) - m_y(w) - m_y(w_*) d\varphi \right| \\ & \leq c\theta^2 |w - w_*|^2 \left\| D^2 e^{-\delta|\cdot - y|^2} \right\|_{L^\infty(B[w, w_*])} \\ & \leq C\theta^2 |w - w_*|^2 \left\| [1 + |\cdot - y|^2] e^{-\delta|\cdot - y|^2} \right\|_{L^\infty(B[w, w_*])} \end{aligned}$$

où  $B([w, w_*])$  désigne la boule de diamètre  $[w, w_*]$  dans  $\mathbb{R}^3$ .

Par conséquent

$$\begin{aligned} \int |y|^6 |h_t(y)| dy & \leq C \int |y|^6 \int \theta^2 |w - w_*|^2 \left\| [1 + |\cdot - y|^2] e^{-\delta|\cdot - y|^2} \right\|_{L^\infty(\dots)} gg_* b \sin \theta d\theta dw dw_* dy \\ & = C \int_{\mathbb{R}^6} |w - w_*|^2 \left[ \int_{\mathbb{R}^3} |y|^6 \left\| [1 + |\cdot - y|^2] e^{-\delta|\cdot - y|^2} \right\|_{L^\infty(B[w, w_*])} dy \right] gg_* dw dw_* \end{aligned}$$

Si  $|y| \geq 2W$  avec  $W = \max(|w|, |w_*|)$ , alors pour tout  $\tilde{w} \in B([w, w_*])$ ,

$$|\tilde{w} - y| \geq |y| - |\tilde{w}| \geq |y| - W \geq \frac{|y|}{2},$$

si bien que

$$e^{-\delta|\tilde{w} - y|^2} \leq e^{-\frac{\delta}{2}|y|^2},$$

et

$$1 + |\tilde{w} - y|^2 \leq 1 + |W|^2 + |y|^2 \leq C[1 + |y|^2].$$

Par conséquent:

$$\int_{|y| \geq 2W} |y|^6 \left\| [1 + |\cdot - y|^2] e^{-\delta|\cdot - y|^2} \right\|_{L^\infty(B[w, w_*])} dy \leq C \int_{|y| \geq 2W} |y|^6 e^{-\frac{\delta}{2}|y|^2} [1 + |y|^2] dy < +\infty$$

et la contribution à l'intégrale correspondante est bornée par

$$C \int_{\mathbb{R}^6} |w - w_*|^2 gg_* dw dw_* \int_{|y| \geq 2W} |y|^6 e^{-\frac{\delta}{2}|y|^2} [1 + |y|^2] dy$$

qui est fini puisque les deux intégrales le sont.

Si  $|y| \leq 2W$ , alors

$$\sup_{\tilde{w} \in B(w, w_*)} \left| [1 + |\tilde{w} - y|^2] e^{-\delta|\tilde{w} - y|^2} \right| \leq (1 + W^2 + 4W^2).1 \leq C[1 + W^2]$$

si bien que

$$\begin{aligned} \int_{|y| \leq 2W} |y|^6 \left\| [1 + |\cdot - y|^2] e^{-\delta|\cdot - y|^2} \right\|_{L^\infty(B[w, w_*])} dy &\leq CW^9[1 + W^2] \\ &\leq C([|w| + |w_*|]^{11} + [|w| + |w_*|]^9) \\ &\leq C(|w|^9 + |w_*|^9 + |w|^{11} + |w_*|^{11}). \end{aligned}$$

La contribution à l'intégrale correspondante est alors bornée par

$$C \int_{\mathbb{R}^6} |w - w_*|^2 [|w|^9 + |w_*|^9 + |w|^{11} + |w_*|^{11}] gg_* dw dw_*.$$

Cette quantité est à son tour bornée uniformément en  $t$  dès que

$$\sup_t \int_{\mathbb{R}^3} |w|^{13} g(w) dw < +\infty,$$

ce qui est ici vérifié d'après la condition imposée sur  $g_0$  et le corollaire 1.10.

Ajoutant les deux contributions, il en résulte que

$$\sup_t \int_{\mathbb{R}^3} |y|^6 |h_t(y)| < +\infty.$$

**3.4.** De 3.3 on déduit en particulier que

$$\sup_t \int_{\mathbb{R}^3} (1 + |y|^6) |h_t(y)| < +\infty.$$

**3.5.** On montre que pour une constante  $C$  indépendante de  $t$ ;

$$|Q(f, f)(t, v)| \leq C[e^{-\frac{\gamma}{4}|v|^2} + \frac{1}{|v|^6} \wedge 1]$$

où  $a \wedge b = \min(a, b)$ .

D'après 3.2

$$Q(f, f)(t, v) = C \int_{\mathbb{R}^3} e^{-\gamma|v-y|^2} h_t(y) dy.$$

Soit  $v \in \mathbb{R}^3$  fixé et  $R = |v|$ .



Sur  $\{|v - y| \geq \frac{R}{2}\}$  on écrit:

$$\begin{aligned} \int_{|v-y| \geq \frac{R}{2}} e^{-\gamma|v-y|^2} |h_t(y)| dy &\leq \int_{|v-y| \geq \frac{R}{2}} e^{-\gamma \frac{R^2}{4}} |h_t(y)| dy \\ &\leq e^{-\gamma \frac{R^2}{4}} \int_{\mathbb{R}^3} |h_t(y)| dy \\ &\leq C e^{-\gamma \frac{R^2}{4}}. \end{aligned}$$

d'après 3.4.

Sur  $\{|v - y| \leq \frac{R}{2}\}$ , alors  $|v| - |y| \leq |v - y| \leq \frac{R}{2}$ , et donc

$$|y| \geq |v| - \frac{R}{2} = \frac{R}{2}.$$

Par conséquent la contribution à  $Q(f, f)$  correspondante est bornée par:

$$\begin{aligned} \int_{|v-y| \leq \frac{R}{2}} e^{-\gamma|v-y|^2} |h_t(y)| dy &\leq \int_{|v-y| \geq \frac{R}{2}} 1 |h_t(y)| dy \\ &\leq \int_{|y| \geq \frac{R}{2}} |h_t(y)| dy \\ &\leq C \left[ \frac{1}{R^6} \wedge 1 \right] \end{aligned}$$

d'après 3.4.

Ajoutant les deux contributions, on obtient bien l'estimation annoncée puisque toutes les constantes utilisées sont indépendantes de  $t$ .

**3.6.** Choissant

$$g(v) = e^{-\frac{\gamma}{4}|v|^2} + \frac{1}{|v|^6} \wedge 1$$

on a d'après 3.5:

$$|Q(f, f)(t, v)| \leq g(v)$$

pour tout  $t$ , et

$$\int_{\mathbb{R}^3} (1 + |v|^2) g(v) dv < +\infty,$$

ce qui conclut cette partie d'après 3.1.

**4.** On montre que  $f_\infty$  est la Maxwellienne  $M$  ayant les mêmes trois premiers moments que  $f_0$ .

**4.1.** On montre d'abord que

$$\int_0^T D(f_\infty(t)) dt \leq \liminf_{n \rightarrow +\infty} \int_0^T D(f_n(t)) dt.$$

Pour cela on vérifie en premier lieu qu'on peut écrire de façon générale

$$D(f(t)) = \frac{1}{4} \int_{\mathbb{R}^3 \times \mathbb{R}^3 \times S^2} A(F(t, v, v_*)) b\left(\left(\frac{v - v_*}{|v - v_*|}, \sigma\right)\right) dv dv_* d\sigma$$

où

$$F(t, v, v_*) = f(t, v) f(t, v_*),$$

$$A(F) = (F - F')(\log F - \log F'),$$

et on note  $J$  l'application définie sur  $L^1(]0, T[ \times \mathbb{R}^3 \times \mathbb{R}^3)$  à valeurs dans  $] -\infty, +\infty]$  par

$$J(F) = \int_{\mathbb{R}^3 \times \mathbb{R}^3 \times S^2 \times ]0, T[} A(F(t, v, v_*)) B(v - v_*, \sigma) dv dv_* d\sigma dt$$

que l'on notera aussi pour simplifier  $J(F) = \int A(F)$ .

D'une part cette application  $J$  est convexe. En effet l'application

$$(X, Y) \mapsto (X - Y)(\log X - \log Y)$$

est convexe dans  $\mathbb{R}^+ \times \mathbb{R}^+$  car sa hessienne a pour valeurs propres 0 et  $\frac{(X+Y)(X^2+Y^2)}{X^2Y^2}$  qui sont positives.

D'autre part  $J$  est semi-continue inférieurement dans  $L^1(]0, T[ \times \mathbb{R}^3 \times \mathbb{R}^3)$  muni de sa topologie forte, c'est-à-dire si une suite  $(F_n)_n$  converge fortement dans  $L^1(]0, T[ \times \mathbb{R}^3 \times \mathbb{R}^3)$  alors

$$J(F) \leq \liminf_{n \rightarrow +\infty} J(F_n).$$

En effet soit  $(F_{n_k})_k$  une sous-suite de  $(F_n)_n$  telle que

$$\liminf_{n \rightarrow +\infty} \int A(F_n) = \lim_{k \rightarrow +\infty} \int A(F_{n_k}).$$

De  $(F_{n_k})_k$  on extrait une sous-suite  $(F_{n_{k'}})_{k'}$  qui converge presque partout vers  $F$  dans  $[0, T] \times \mathbb{R}^3 \times \mathbb{R}^3$ , et donc  $(A(F_{n_{k'}}))_{k'}$  converge presque partout vers  $A(F)$ . Comme de plus  $A(F_{n_{k'}}) \geq 0$  car la fonction  $\log$  est croissante, le lemme de Fatou assure que

$$\int \liminf_{k' \rightarrow +\infty} A(F_{n_{k'}}) \leq \liminf_{k' \rightarrow +\infty} \int A(F_{n_{k'}}).$$

En regroupant ces différents résultats, on en déduit que

$$\int A(F) = \int \liminf_{k' \rightarrow +\infty} A(F_{n_{k'}}) \leq \liminf_{k' \rightarrow +\infty} \int A(F_{n_{k'}}) = \lim_{k' \rightarrow +\infty} \int A(F_{n_{k'}}) = \lim_{k \rightarrow +\infty} \int A(F_{n_k})$$

soit

$$J(F) \leq \liminf_{n \rightarrow +\infty} J(F_n).$$

$J$  étant ainsi convexe et semi-continue inférieurement dans  $L^1(]0, T[ \times \mathbb{R}^3 \times \mathbb{R}^3)$  fort, elle est semi-continue inférieurement dans  $L^1(]0, T[ \times \mathbb{R}^3 \times \mathbb{R}^3)$  faible, c'est-à-dire si une suite  $(F_n)_n$  converge faiblement vers  $F$ , alors

$$J(F) \leq \liminf_{n \rightarrow +\infty} J(F_n).$$

Or d'après l'étape 2, la suite  $(F_n)_n$  des produits  $F_n(t, v, v_*) = f_n(t, v) f_{n*}(t, v_*)$  converge faiblement vers le produit  $F_\infty(t, v, v_*) = f_\infty(t, v) f_{\infty*}(t, v_*)$  dans  $L^1(]0, T[ \times \mathbb{R}^3 \times \mathbb{R}^3)$ . Ce qui donne le résultat annoncé.

**4.2.** On en déduit que  $D(f_\infty(t)) = 0$  pour tout  $t \geq 0$ .

En effet pour tout  $n$  on écrit

$$\int_0^T D(f_n(t)) dt = H(f_n(0)) - H(f_n(T)) = H(f(t_n)) - H(f(t_n + T)).$$

Or  $H(f(t))$  a une limite finie quand  $t$  tend vers  $+\infty$  puisque cette fonction est décroissante et de plus bornée d'après 1.3. Par conséquent

$$\lim_{n \rightarrow +\infty} \int_0^T D(f_n(t)) dt = 0.$$

Comme de plus  $D(f_\infty(t)) \geq 0$ , on en déduit à l'aide de 3.1 le résultat annoncé.

**4.3.** Ainsi  $f_\infty(t)$  est une solution de l'équation de Boltzmann, de classe  $\mathcal{C}^\infty(\mathbb{R}^3)$  pour tout  $t \geq 0$  et telle que  $D(f_\infty(t)) = 0$ . Alors  $f_\infty(t)$  est constante et égale à une Maxwellienne  $M$ .

En effet, supposant par exemple que  $b(\cos \theta) > 0$  pour tout  $\theta$ , on déduit dans un premier temps que

$$f_\infty(t) f_{\infty*}(t) = f'_\infty(t) f'_{\infty*}(t)$$

pour tout  $t \geq 0$ , en particulier pour  $t = 0$ . Alors par transformation de Fourier en  $v$ , on montre, comme dans [P] par exemple, que  $f_\infty(0)$  est une Maxwellienne  $M$ , et par conséquent, par unicité, que  $f_\infty(t)$  est constante égale à cette Maxwellienne  $M$ .

On montre plus précisément que si

$$\int |v|^4 dg_0(v) < +\infty,$$

alors  $M$  est la Maxwellienne qui a les mêmes trois premiers moments que  $f_0$ .

Tout d'abord on note que sous cette hypothèse on a aussi

$$\int |v|^4 f_0(v) dv < +\infty$$

car

$$\begin{aligned} \int |v|^4 f_0(v) dv &\leq \alpha \int \left( \int |v|^4 e^{-\beta|v-y|^2} dg_0(y) \right) dv \\ &= \alpha \int \left( \int |v+y|^4 e^{-\beta|v|^2} dv \right) dg_0(y) \\ &\leq 8\alpha \left[ \int |v|^4 e^{-\beta|v|^2} dv \int dg_0(v) + \int e^{-\beta|v|^2} dv \int |y|^4 dg_0(y) \right] \\ &< +\infty. \end{aligned}$$

Par le corollaire 1.10, la solution  $f(t)$  vérifie alors

$$\sup_{t \geq 0} \int_{\mathbb{R}^3} |v|^4 f(t, v) dv < +\infty.$$

Par suite pour tous  $R \geq 0$  et  $t \geq 0$  on a

$$\int_{|v| \geq R} |v|^2 f(t, v) dv \leq \frac{1}{R^2} \sup_{t \geq 0} \int_{\mathbb{R}^3} |v|^4 f(t, v) dv$$

d'où l'on déduit un contrôle du deuxième moment de  $f(t)$  sous la forme

$$\lim_{R \rightarrow +\infty} \sup_{t \geq 0} \int_{|v| \geq R} |v|^2 f(t, v) dv = 0.$$

Soit alors  $\varepsilon > 0$  fixé. On décompose

$$\begin{aligned} \left| \int |v|^2 f_n(t, v) dv - \int |v|^2 M(v) dv \right| &\leq \int_{|v| \geq R} |v|^2 f_n(t, v) dv + \int_{|v| \geq R} |v|^2 M(v) dv \\ &\quad + \left| \int_{|v| \leq R} |v|^2 [f_n(t, v) - M(v)] dv \right|. \end{aligned}$$

On fixe un  $R$  tel que

$$\int_{|v| \geq R} |v|^2 M(v) dv \leq \varepsilon$$

et tel que, d'après le contrôle précédent

$$\sup_{t \geq 0} \int_{|v| \geq R} |v|^2 f(t, v) dv \leq \varepsilon.$$

Comme la suite  $(f_n)_n$  converge faiblement\* vers  $M$ , il existe  $N$  tel que pour tout  $n \geq N$

$$\left| \int_{|v| \leq R} |v|^2 [f_n(t, v) - M(v)] dv \right| \leq \varepsilon.$$

Ainsi

$$\lim_{n \rightarrow +\infty} \int_{\mathbb{R}^3} |v|^2 f_n(t, v) dv = \int_{\mathbb{R}^3} |v|^2 M(v) dv.$$

Or par conservation de l'énergie cinétique on a pour tout  $n$  et tout  $t$

$$\int_{\mathbb{R}^3} |v|^2 f(t + t_n, v) dv = \int_{\mathbb{R}^3} |v|^2 f_0(v) dv$$

donc

$$\int_{\mathbb{R}^3} |v|^2 f_n(t, v) dv = \int_{\mathbb{R}^3} |v|^2 f_0(v) dv.$$

Ainsi

$$\int_{\mathbb{R}^3} |v|^2 M(v) dv = \int_{\mathbb{R}^3} |v|^2 f_0(v) dv.$$

On montre de même l'égalité des deux premiers moments.

**5.** On en déduit que  $f(t) \rightharpoonup M$  quand  $t$  tend vers  $+\infty$ , au sens faible\* des mesures de probabilité sur  $\mathbb{R}^3$ .

En effet l'étude précédente montre que toute suite  $(t_n)_n$  tendant vers  $+\infty$  contient une sous-suite  $(t_{n_k})_k$  tendant vers  $+\infty$  telle que  $(f(t_{n_k}))_k$  converge vers la même Maxwellienne  $M$  •

## 2 - La distance de Wasserstein.

Dans toute cette section  $p$  désigne un entier  $\geq 1$ ,  $X$  une partie de  $\mathbb{R}^N$  et  $\mathcal{P}_p$  l'ensemble des mesures de probabilité  $\mu$  sur  $X$  de moment d'ordre  $p$  fini, c'est-à-dire telles que  $\int_X |x|^p d\mu(x) < +\infty$ .

**Définition 2.1** Etant données deux mesures  $f$  et  $g$  de  $\mathcal{P}_p$ , on note  $\mathcal{T}_p(f, g)$  le coût de transport optimal entre  $f$  et  $g$  associé à la fonction de coût  $|x - y|^p$  dans  $\mathbb{R}^N$ , c'est-à-dire

$$\mathcal{T}_p(f, g) = \inf_{\pi \in \Pi(f, g)} \int_{X \times X} |x - y|^p d\pi(x, y)$$

où  $\Pi(f, g)$  désigne l'ensemble des mesures de probabilité sur  $X \times X$  de marginales  $f$  et  $g$ .

On utilisera plusieurs fois le

**Lemme 2.2** Si  $(\pi_k)_k$  est une suite de mesures de probabilité sur  $X \times X$  convergeant au sens faible\* vers la mesure  $\pi$ , alors

$$\int_{X \times X} |x - y|^p d\pi(x, y) \leq \liminf_{k \rightarrow +\infty} \int_{X \times X} |x - y|^p d\pi_k(x, y).$$

**Démonstration.** On considère une suite croissante  $(\phi_n)_n$  de fonctions de  $C_0(X \times X)$  convergeant ponctuellement vers  $|x - y|^p$ . D'après le théorème de convergence monotone, on a alors

$$\begin{aligned} \int |x - y|^p d\pi(x, y) &= \lim_{n \rightarrow \infty} \int \phi_n(x, y) d\pi(x, y) \\ &= \lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} \int \phi_n(x, y) d\pi_k(x, y) \leq \liminf_{k \rightarrow \infty} \int |x - y|^p d\pi_k(x, y) \bullet \end{aligned}$$

**Proposition 2.3** La borne inférieure définissant  $\mathcal{T}_p$  est atteinte.

**Démonstration.** Soit  $f$  et  $g$  deux mesures de  $\mathcal{P}_p$  et  $(\pi_k)_k$  une suite minimisante de  $\mathcal{T}_p(f, g)$ .

Soit une suite croissante  $(\phi_n)_n$  de fonctions de  $C_0(X)$  convergeant ponctuellement vers  $|x|^p$  dans  $X$ . Pour tous  $k$  et  $n$  on a

$$\begin{aligned} \int_{X \times X} (\phi_n(x) + \phi_n(y)) d\pi_k(x, y) &= \int_X \phi_n(x) df(x) + \int_X \phi_n(y) dg(y) \leq \\ &= \int_X |x|^p df(x) + \int_X |y|^p dg(y) = C \end{aligned}$$

où  $C$  est une constante finie car  $f$  et  $g$  ont leur moment d'ordre  $p$  fini. Par le théorème de convergence monotone par rapport à  $n$  pour chaque  $k$ , on en déduit que pour chaque  $k$  on a

$$\int_{X \times X} (|x|^p + |y|^p) d\pi_k(x, y) \leq C.$$

Ainsi la suite  $(\pi_k)_k$  est tendue et le théorème de Prokhorov (cf.[D]) assure alors que cette suite admet une sous-suite convergente dont la limite  $\pi$  appartient à  $\Pi(f, g)$ . Et comme d'après le lemme précédent

$$\int_{X \times X} |x - y|^p d\pi(x, y) \leq \liminf_{k \rightarrow +\infty} \int_{X \times X} |x - y|^p d\pi_k(x, y) = \mathcal{T}_p(f, g),$$

$\pi$  est bien un minimiseur •

On donne maintenant une propriété de représentation du coût optimal à l'aide de variables aléatoires particulières.

**Proposition 2.4** *Soit  $f$  et  $g$  deux éléments de  $\mathcal{P}_p$  et  $\pi$  minimisant le transport entre  $f$  et  $g$ . Alors il existe deux couples  $(X, Y)$  et  $(X_*, Y_*)$  de variables aléatoires, indépendantes et de loi  $\pi$ . En particulier  $X$  et  $X_*$  sont de loi  $f$ ,  $Y$  et  $Y_*$  sont de loi  $g$  et  $E|X - Y|^p = E|X_* - Y_*|^p = \mathcal{T}_p(f, g)$ .*

**Démonstration.** On note d'abord qu'étant donnée une loi de probabilité  $\mu$  sur un ensemble  $E$ , on peut construire deux variables aléatoires indépendantes de loi  $\mu$ ; il suffit pour cela de considérer les variables aléatoires  $X_i(\omega) = \omega_i$  pour  $i = 1, 2$  sur l'espace probabilisé  $(E^2, \mathcal{B}^{\otimes 2}, \mu^{\otimes 2})$ .

En appliquant cette propriété au minimiseur  $\pi$  obtenu dans la proposition précédente, on obtient l'existence de deux variables aléatoires indépendantes  $Z = (X, Y)$  et  $Z_* = (X_*, Y_*)$  de loi  $\pi$ . Comme  $\pi$  admet pour marginales  $f$  et  $g$ , alors  $X$  et  $X_*$  sont de loi  $f$ ,  $Y$  et  $Y_*$  de loi  $g$ . De plus

$$E|X - Y|^p = \int_{X \times X} |x - y|^p d\pi(x, y) = \mathcal{T}_p(f, g)$$

et de même pour  $X_*$  et  $Y_*$ , ce qui montre la proposition •

**Remarque 2.5** Dans le cas où  $p = 2$  et par exemple  $g$  est une gaussienne, on peut préciser les propositions précédentes. D'après [V2], le minimum  $\pi$  est unique et est donné par

$$d\pi(x, y) = \delta(\{x = \nabla\varphi(y)\}) dg(y)$$

où  $\nabla\varphi$  est l'unique gradient d'une fonction convexe  $\varphi$  envoyant  $g$  sur  $f$ , c'est-à-dire telle que pour toute fonction  $\psi$  continue bornée

$$\int \psi(y) df(x) = \int \psi(\nabla\varphi(x)) dg(x).$$

Si  $(X, Y)$  est un couple de variables aléatoires de loi  $\pi$ , alors

$$E|X - \nabla\varphi(Y)| = \int |x - \nabla\varphi(y)| d\pi(x, y) = \int |x - \nabla\varphi(y)| \delta(\{x = \nabla\varphi(y)\}) dg(y) = 0$$

si bien que

$$X = \nabla\varphi(Y) \quad ps.$$

On donne une propriété de convexité du coût optimal.

**Proposition 2.6** Soit  $f_1, f_2, g_1$  et  $g_2$  des mesures de probabilité sur  $X$  dont les moments d'ordre  $p$  sont finis et soit  $\alpha \in [0, 1]$ . Alors

$$\mathcal{T}_p(\alpha f_1 + (1 - \alpha) f_2, \alpha g_1 + (1 - \alpha) g_2) \leq \alpha \mathcal{T}_p(f_1, g_1) + (1 - \alpha) \mathcal{T}_p(f_2, g_2).$$

**Démonstration.** Il suffit pour cela de remarquer que si  $\pi_1 \in \Pi(f_1, g_1)$  et  $\pi_2 \in \Pi(f_2, g_2)$ , alors  $\alpha\pi_1 + (1 - \alpha)\pi_2 \in \Pi(\alpha f_1 + (1 - \alpha)f_2, \alpha g_1 + (1 - \alpha)g_2)$  •

On donne une majoration du coût optimal pour deux mesures particulières dans les cas où  $X = \mathbb{R}^3$  et  $p = 2$ .

**Proposition 2.7** Si  $U_{c,r,k}$  désigne la distribution uniforme de probabilité sur le cercle  $C_{c,r,k}$  de centre  $c$ , rayon  $r$  et situé dans le plan orthogonal au vecteur unitaire  $k$ , alors

$$\mathcal{T}_2(U_{c_1, r_1, k_1}, U_{c_2, r_2, k_2}) \leq |c_1 - c_2|^2 + r_1^2 + r_2^2 - r_1 r_2 (1 + |(k_1, k_2)|).$$

**Démonstration.** Sans perte de généralité, on peut supposer par une rotation que  $k_1 = (0, 0, 1)$  et  $k_2 = (0, -\sin \gamma, \cos \gamma)$ . On introduit alors les variables aléatoires  $X$  et  $Y$  définies par

$$X = c_1 + r_1(\cos \omega, \sin \omega, 0) \quad Y = c_2 + r_2(\cos \omega, \sin \omega \cos \gamma, \sin \omega \sin \gamma)$$

où  $\omega$  décrit  $S^1$  suivant la loi de probabilité uniforme. Ces variables aléatoires ayant pour lois respectives  $U_{c_1, r_1, k_1}$  et  $U_{c_2, r_2, k_2}$ , la définition du coût optimal assure alors que

$$\mathcal{T}_2(U_{c_1, r_1, k_1}, U_{c_2, r_2, k_2}) \leq E|X - Y|^2.$$

Or

$$E|X - Y|^2 = \frac{1}{2\pi} \int_0^{2\pi} |X(\omega) - Y(\omega)|^2 d\omega = |c_1 - c_2|^2 + r_1^2 + r_2^2 - r_1 r_2 (1 + \cos \gamma)$$

d'où l'on déduit la majoration annoncée •

**Proposition 2.8** L'application  $W_p = \mathcal{T}_p^{\frac{1}{p}}$  définit une métrique sur l'ensemble  $\mathcal{P}_p$ , appelée distance de Wasserstein.

**Démonstration.** On donne par exemple la preuve de l'inégalité triangulaire.

Soit  $\mu_1, \mu_2$  et  $\mu_3$  dans  $\mathcal{P}_p$ , dont les supports seront notés  $X_1, X_2$  et  $X_3$ . On note  $\pi_{12}$  une mesure optimale entre  $\mu_1$  et  $\mu_2$  d'une part, et  $\pi_{23}$  une mesure optimale entre  $\mu_2$  et  $\mu_3$  d'autre part. D'après le lemme de Dudley donné ci-après, il existe une mesure  $\pi$  sur  $X_1 \times X_2 \times X_3$  de marginales  $\pi_{12}$  sur  $X_1 \times X_2$  et  $\pi_{23}$  sur  $X_2 \times X_3$ . Si  $\pi_{13}$  est la marginale de  $\pi$  sur  $X_1 \times X_3$ , alors  $\pi_{13} \in \Pi(\mu_1, \mu_3)$  et

$$\begin{aligned} W_p(\mu_1, \mu_3) &\leq \left( \int_{X_1 \times X_3} |x_1 - x_3|^p d\pi_{13}(x_1, x_3) \right)^{1/p} = \left( \int_{X_1 \times X_2 \times X_3} |x_1 - x_3|^p d\pi(x_1, x_2, x_3) \right)^{1/p} \\ &\leq \left( \int_{X_1 \times X_2 \times X_3} |x_1 - x_2|^p d\pi(x_1, x_2, x_3) \right)^{1/p} + \left( \int_{X_1 \times X_2 \times X_3} |x_2 - x_3|^p d\pi(x_1, x_2, x_3) \right)^{1/p} \\ &= \left( \int_{X_1 \times X_2} |x_1 - x_2|^p d\pi_{12}(x_1, x_2) \right)^{1/p} + \left( \int_{X_2 \times X_3} |x_2 - x_3|^p d\pi_{23}(x_2, x_3) \right)^{1/p} \end{aligned}$$

$$= W_p(\mu_1, \mu_2) + W_p(\mu_2, \mu_3) \bullet$$

Le lemme de Dudley s'énonce ainsi:

**Lemme 2.9 (de Dudley)** Soient  $\mu_1, \mu_2$  et  $\mu_3$  trois mesures de probabilité sur  $X_1, X_2$  et  $X_3$ , et soient  $\pi_{12} \in \Pi(\mu_1, \mu_2), \pi_{23} \in \Pi(\mu_2, \mu_3)$ . Alors il existe une mesure de probabilité  $\pi$  sur  $X_1 \times X_2 \times X_3$  de marginales  $\pi_{12}$  sur  $X_1 \times X_2$  et  $\pi_{23}$  sur  $X_2 \times X_3$ .

**Démonstration.** De façon générale le principe de désintégration des mesures (cf. [Ga]) assure que si  $\pi$  est une mesure de probabilité sur  $X \times Y$  de marginale  $\mu$  sur  $X$ , alors il existe une application mesurable  $x \mapsto \pi_x$  de  $X$  dans l'ensemble  $\mathcal{P}(Y)$  des mesures de probabilité sur  $Y$  telle que

$$\pi = \int_X (\delta_x \otimes \pi_x) d\mu(x).$$

En appliquant ce principe à  $\pi_{12}$  et  $\pi_{23}$  par rapport à leur marginale commune  $\mu_2$ , il existe des applications mesurables  $\pi_{12;2}$  et  $\pi_{23;2}$  de  $X_2$  dans  $\mathcal{P}(X_1)$  et  $\mathcal{P}(X_3)$  respectivement telles que

$$\pi_{12} = \int_{X_2} (\pi_{12;2} \otimes \delta_{x_2}) d\mu_2(x_2),$$

$$\pi_{23} = \int_{X_2} (\delta_{x_2} \otimes \pi_{23;2}) d\mu_2(x_2).$$

On définit alors  $\pi$  sur  $X_1 \times X_2 \times X_3$  par

$$\pi = \int_{X_2} (\pi_{12;2} \otimes \delta_{x_2} \otimes \pi_{23;2}) d\mu_2(x_2)$$

qui est une mesure de probabilité sur  $X_1 \times X_2 \times X_3$  de marginales  $\pi_{12}$  sur  $X_1 \times X_2$  et  $\pi_{23}$  sur  $X_2 \times X_3$  •

On donne enfin des propriétés topologiques de cette distance de Wasserstein.

**Théorème 2.10** Soit  $(f_n)_n$  une suite de mesures de probabilité de  $\mathcal{P}_p$  convergeant au sens faible\* vers  $f$  et telle que la suite  $\left( \int_X |x|^p df_n(x) \right)_n$  de leur moment d'ordre  $p$  converge vers  $\int_X |x|^p df(x)$ .

Alors la suite  $(W_p(f_n, f))_n$  converge vers 0.

**Démonstration.** 1 - On montre d'abord que la suite  $(f_k)_k$  vérifie la condition de tension

$$\lim_{R \rightarrow \infty} \sup_{k \geq 0} \int_{|x| \geq R} |x|^p df_k(x) = 0.$$

Pour cela on décompose  $\int_{|x| \geq R} |x|^p df_k(x)$  sous la forme

$$\begin{aligned} \int_{|x| \geq R} |x|^p df_k(x) &\leq \left| \int |x|^p df_k(x) - \int |x|^p df(x) \right| \\ &+ \left| \int_{|x| \leq R} |x|^p df(x) - \int_{|x| \leq R} |x|^p df_k(x) \right| + \int_{|x| \geq R} |x|^p df(x). \end{aligned}$$



Soit alors  $\varepsilon > 0$  fixé. Comme les moments d'ordre  $p$  convergent, il existe  $k_0$  tel que pour tout  $k \geq k_0$  on ait

$$\left| \int_X |x|^p df_k(x) - \int_X |x|^p df(x) \right| \leq \varepsilon,$$

comme  $f \in \mathcal{P}_p$ , il existe  $R_0$  tel que

$$\int_{|x| \geq R_0} |x|^p df \leq \varepsilon$$

et comme enfin  $(f_k)_k$  converge faiblement vers  $f$ , il existe  $k_1$  tel que pour tout  $k \geq k_1$  on ait

$$\left| \int_{|x| \leq R_0} |x|^p df_k(x) - \int_{|x| \leq R_0} |x|^p df(x) \right| \leq \varepsilon.$$

Par suite pour tout  $k \geq k_2 = \sup(k_0, k_1)$  on a

$$\int_{|x| \geq R_0} |x|^p df_k(x) \leq 3\varepsilon$$

et comme  $f_k \in \mathcal{P}_p$  pour  $k \leq k_2$ , il existe  $R_1$  tel que pour  $k \leq k_2$

$$\int_{|x| \geq R_1} |x|^p df_k(x) \leq \varepsilon.$$

De là on déduit que pour  $\varepsilon > 0$ , il existe  $R_2$  tel que

$$\sup_{k \geq 0} \int_{|x| \geq R_2} |x|^p df_k(x) \leq \varepsilon.$$

Comme de plus  $\sup_{k \geq 0} \int_{|x| \geq R} |x|^p df_k(x)$  est une fonction décroissante de  $R$ , on a le résultat annoncé.

2 - On montre maintenant qu'il suffit de vérifier le théorème quand  $X$  est compact.

Soit  $f_k$  et  $f$  vérifiant l'hypothèse du théorème et soit  $\varepsilon > 0$  fixé.

Pour tout  $\pi_k \in \Pi(f_k, f)$  et  $R \geq 0$  on décompose

$$\int_{X \times X} |x - y|^p d\pi_k = \int_{|x|^p + |y|^p \leq R^p} |x - y|^p d\pi_k + \int_{|x|^p + |y|^p > R^p} |x - y|^p d\pi_k.$$

D'une part on a

$$\int_{|x|^p + |y|^p \leq R^p} |x - y|^p d\pi_k(x, y) \leq \int_{\{|x| \leq R\} \times \{|y| \leq R\}} |x - y|^p d\pi_k(x, y).$$

D'autre part pour une constante  $C$  ne dépendant que de  $p$  on a

$$\begin{aligned} \int_{|x|^p + |y|^p \geq R^p} |x - y|^p d\pi_k &\leq \int_{|x| \geq R/2} |x - y|^p d\pi_k + \int_{|y| \geq R/2} |x - y|^p d\pi_k \\ &\leq C \int_{|x| \geq R/2} (|x|^p + |y|^p) d\pi_k + C \int_{|y| \geq R/2} (|x|^p + |y|^p) d\pi_k \\ &= C \int_{|x| \geq R/2} |x|^p df_k(x) + C f_k(\{|x| \geq R/2\}) \int_X |y|^p df_k(y) \\ &+ C \int_{|y| \geq R/2} |y|^p df_k(y) + C f_k(\{|y| \geq R/2\}) \int_X |x|^p df_k(x) \end{aligned}$$

et d'après l'étape précédente, il existe  $R$  tel que chacune de ces 4 quantités soit plus petite que  $\varepsilon$ .

Ainsi il existe  $R$  tel que pour tout  $\pi_k \in \Pi(\mu_k, \mu)$  on ait

$$\int_{X \times X} |x - y|^p d\pi_k \leq \int_{K \times K} |x - y|^p d\pi_k + \varepsilon$$

où  $K = \{x \in X; |x| \leq R\}$ .

Ce compact  $K$  étant fixé, on considère les mesures de probabilité  $\nu_k$  et  $\nu$  définies sur  $K$  par

$$\nu_k = \frac{f_{k|K}}{f_k(K)} \quad \nu = \frac{f|K}{f(K)}$$

auxquelles on associe, avec des notations naturelles,  $\pi_k^K \in \Pi^K(\nu_k, \nu)$  qui réalise  $\mathcal{T}_p^K(\nu_k, \nu)$ , c'est-à-dire

$$\int_{K \times K} |x - y|^p d\pi_k^K = \mathcal{T}_p^K(\nu_k, \nu).$$

On définit alors une nouvelle mesure  $\tilde{\pi}_k^K$  sur  $K \times K$  par

$$\tilde{\pi}_k^K = f_k(K) f(K) \pi_k^K,$$

qu'on prolonge ensuite en une mesure  $\pi_k$  sur  $X \times X$  définie par

$$\pi_k(F) = \tilde{\pi}_k^K(F \cap (K \times K)) + f_k \otimes \mu(F \cap (K \times K)^c).$$

Cette mesure  $\pi_k$  sur  $X \times X$  admet pour marginales  $f_k$  et  $f$ , et vérifie

$$\int_{K \times K} |x - y|^p d\pi_k = \int_{K \times K} |x - y|^p d\pi_k^K.$$

Compte-tenu du choix de  $K$ , c'est-à-dire de  $R$ , on a donc d'après les estimations précédentes valables pour tout  $\pi_k$

$$\int_{X \times X} |x - y|^p d\pi_k(x, y) \leq \mathcal{T}_p^K(\nu_k, \nu) + \varepsilon.$$

Or d'une part la suite  $(\nu_k)_k$  converge faiblement vers  $\nu$  sur  $K$ ; en effet  $(f_{k|K})_k$  converge faiblement vers  $f|K$  sur  $K$  et  $(f_k(K))_k$  converge vers  $f(K)$  puisque  $f_k$  converge faiblement vers  $f$  sur  $K$ . D'autre part les moments d'ordre  $p$  dans  $K$  convergent aussi. Par conséquent si l'on suppose le théorème vrai lorsque l'espace de définition est compact, alors  $\mathcal{T}_p^K(\nu_k, \nu)$  tend vers 0 quand  $k$  tend vers  $+\infty$ .

Par suite la majoration précédente permet d'en déduire que  $\int_{X \times X} |x - y|^p d\pi_k(x, y)$  tend vers 0 quand  $k$  tend vers  $+\infty$ . A fortiori  $\mathcal{T}(\mu_k, \mu)$  tend aussi vers 0.

3 - On est donc ramené à montrer le théorème dans le cas où  $X$  est compact. Il suffit dans ce cas de montrer que  $W_1(f_k, f)$  tend vers 0 car  $W_p \leq W_1^{\frac{1}{p}} \text{diam}(X)^{1-\frac{1}{p}}$  d'après l'inégalité de Hölder.

Or pour  $p = 1$  la formule de Kantorovitch-Rubinstein assure que

$$W_1(f_k, f) = \sup \left\{ \int \varphi d(f_k - f); \varphi \in C(X), \|\varphi\|_{Lip} \leq 1 \right\}$$

où

$$\|\varphi\|_{Lip} = \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{|x - y|}.$$

Comme l'ensemble  $Lip(X)$  des fonctions  $\varphi$  telles que  $\|\varphi\|_{Lip} < +\infty$  est dense dans  $C(X)$ , il suffit alors de montrer que

$$\sup\left\{\int \varphi d(f_k - f); \varphi \in Lip(X), \|\varphi\|_{Lip} \leq 1\right\}$$

et même, en remplaçant  $\varphi$  par  $\varphi - \varphi(0)$ , que

$$\sup\left\{\int \varphi d(f_k - f); \varphi \in Lip(X), \varphi(0) = 0, \|\varphi\|_{Lip} \leq 1\right\}$$

tend vers 0 quand  $k$  tend vers  $+\infty$ .

Pour simplifier on note  $\mathcal{F} = \{\varphi \in Lip(X); \varphi(0) = 0, \|\varphi\|_{Lip} \leq 1\}$  et  $L_k(\varphi) = \int \varphi d(f_k - f)$ .

Alors  $(L_k)_k$  est une suite de formes linéaires continues dans l'espace de Banach  $C(X)$  qui converge faiblement vers 0 et d'après le lemme d'Ascoli,  $\mathcal{F}$  est un sous-ensemble compact de  $C(X)$ .

On veut donc montrer que la suite  $\sup\{L_k(\varphi); \varphi \in \mathcal{F}\}$  tend vers 0 quand  $k$  tend vers  $+\infty$ . Pour cela il suffit de montrer que toute sous-suite contient une sous-suite qui tend vers 0.

Soit donc une sous-suite de la suite  $(L_k)_k$  que l'on note aussi  $(L_k)_k$  pour simplifier. Pour tout  $k$  il existe  $\varphi_k$  telle que

$$\sup\{L_k(\varphi); \varphi \in \mathcal{F}\} = L_k(\varphi_k)$$

puisque  $L_k$  est continue et  $\mathcal{F}$  compact. Or il existe une sous-suite  $(\varphi_{k_j})_j$  de la suite  $(\varphi_k)_k$  qui converge dans  $\mathcal{F}$  puisque  $\mathcal{F}$  est compact, et comme de plus  $(L_k)_k$  converge faiblement vers 0, il en résulte que la suite  $(L_{k_j}(\varphi_{k_j}))_j$  tend vers 0 quand  $j$  tend vers  $+\infty$ .

Ce qui termine la démonstration du théorème •

**Corollaire 2.11** *Soit  $f$  et  $g$ ,  $f_n$  et  $g_n$  pour  $n \geq 0$  des mesures de probabilité dans  $\mathbb{R}^N$  dont les moments d'ordre  $p$  sont finis.*

*Si les suites  $(f_n)_n$  et  $(g_n)_n$  convergent au sens faible\* vers  $f$  et  $g$  respectivement, alors*

$$W_p(f, g) \leq \liminf_{n \rightarrow +\infty} W_p(f_n, g_n),$$

*et si de plus les suites des moments d'ordre  $p$  des  $f_n$  et des  $g_n$  convergent respectivement vers les moments d'ordre  $p$  de  $f$  et de  $g$ , alors*

$$W_p(f, g) = \lim_{n \rightarrow +\infty} W_p(f_n, g_n).$$

**Corollaire 2.12** *Soit  $f_0$  une mesure de probabilité vérifiant (1.4) et dont tous les moments sont finis. Alors la solution  $f$  de l'équation de Boltzmann (1.2) de donnée initiale  $f_0$  converge pour la distance de Wasserstein, vers sa Maxwellienne associée.*

**Démonstration.** En effet il suffit de noter que le corollaire 1.9 assure que  $f(t)$  converge vers la Maxwellienne  $M$  au sens faible des mesures, et que son moment d'ordre 2 est constant, égal à celui de  $M$  •

Le théorème de Tanaka, qu'on va montrer dans la partie suivante, va donner un résultat plus précis, puisqu'il assure que  $W_2(f(t), M)$  décroît vers 0 dès que la donnée initiale  $f_0$  vérifie (1.4).

On notera dans la suite

$$W := W_2.$$

### 3 - Le théorème de Tanaka.

Dans toute cette partie on se place en dimension  $N = 3$ . On va supposer d'abord que la section efficace  $b$  vérifie l'hypothèse de type cutoff suivante

$$\int_{S^2} b((k, \sigma)) d\sigma < +\infty$$

pour  $k$  vecteur de norme unité, c'est-à-dire que

$$\int_0^\pi b(\cos \theta) \sin \theta d\theta < +\infty.$$

Par dilatation du temps, il est alors possible de se ramener au cas où

$$\int_{S^2} b((k, \sigma)) d\sigma = 1$$

pour  $k$  vecteur de norme unité, c'est-à-dire

$$\int_0^\pi b(\cos \theta) \sin \theta d\theta = \frac{1}{2\pi}. \quad (3.1)$$

C'est le cas en particulier lorsque la section efficace  $b$  est de type atomique, c'est-à-dire lorsque

$$b(\cos \theta) \sin \theta d\theta = \frac{1}{2\pi} \sum_{i=1}^M \alpha_i \delta_{\theta_i} \quad (3.2)$$

avec des  $\theta_i \in [0, \pi]$  et des  $\alpha_i \geq 0$  tels que  $\sum_{i=1}^M \alpha_i = 1$ .

L'équation de Boltzmann s'écrit alors dans le cas normalisé (3.1)

$$\partial_t f = Q^+(f, f) - f \quad (3.3)$$

avec un terme de gain donné par

$$Q^+(f, f) = \int_{\mathbb{R}^3} dv_* \int_{S^2} b\left(\left(\frac{v - v_*}{|v - v_*|}, \sigma\right)\right) f' f'_* d\sigma. \quad (3.4)$$

On note en particulier que si  $g$  est une maxwellienne normalisée par (1.4), alors  $Q^+(g, g) = g$ .

Sous cette hypothèse, on va montrer que le terme de gain  $Q^+$  a une propriété de contraction relativement à la distance de Wasserstein.

Pour cela on établit d'abord la formule de représentation de Tanaka de  $Q^+$ .

**Proposition 3.1** *On suppose que la section efficace  $b$  est de type atomique de la forme*

$$b(\cos \theta) \sin \theta d\theta = \frac{1}{2\pi} \delta_{\theta_0}$$

pour un  $\theta_0 \in [0, \pi]$ . Alors pour toute probabilité  $f$ , on a la formule de Tanaka

$$Q^+(f, f) = E\Pi_{X, X^*, \theta_0}$$

où  $X$  et  $X^*$  sont des variables aléatoires indépendantes quelconques de loi  $f$  et  $\Pi_{v, v_*, \theta_0}$  est la distribution uniforme sur le cercle  $C_{c, r, k}$  de centre

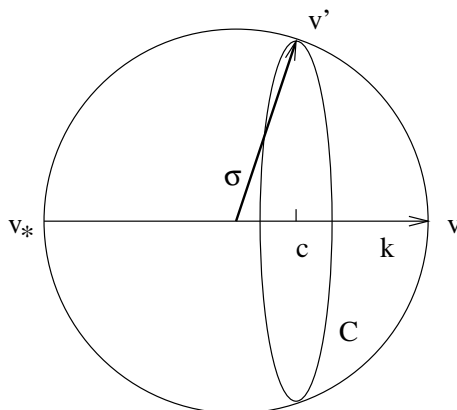
$$c = \frac{v + v_* + (v - v_*) \cos \theta_0}{2},$$

de rayon

$$r = \frac{|v - v_*|}{2} \sin \theta_0$$

et situé dans le plan orthogonal au vecteur unitaire

$$k = \frac{v - v_*}{|v - v_*|}.$$



Définition de  $C_{c, r, k}$

**Démonstration.** Etant données une probabilité  $f$  dans  $\mathbb{R}^3$  et une fonction  $\phi$  continue bornée dans  $\mathbb{R}^3$ , on a

$$\begin{aligned} \int_{\mathbb{R}^3} Q^+(f, f) \phi(v) dv &= \int_{S^2} \int_{\mathbb{R}^6} f' f_* \phi b\left(\left(\frac{v - v_*}{|v - v_*|}, \sigma\right)\right) d\sigma dv dv_* \\ &= \int f f_* \phi' b\left(\left(\frac{v - v_*}{|v - v_*|}, \sigma\right)\right) d\sigma dv dv_*. \end{aligned}$$

Or en représentation sphérique dans  $S^2$  on peut écrire

$$\int_{S^2} \phi(v') b\left(\left(\frac{v-v_*}{|v-v_*|}, \sigma\right)\right) d\sigma = \int_0^\pi \int_0^{2\pi} \phi(v'(v, v_*, \theta, \varphi)) b(\cos \theta) \sin \theta d\theta d\varphi,$$

soit si  $b$  est de type atomique concentré en un point  $\theta_0$  dans  $[0, \pi]$

$$\int_{S^2} \phi(v') b\left(\left(\frac{v-v_*}{|v-v_*|}, \sigma\right)\right) d\sigma = \frac{1}{2\pi} \int_0^{2\pi} \phi(v'(v, v_*, \theta_0, \varphi)) d\varphi.$$

Ainsi avec les notations introduites dans la proposition 3.1 on a donc

$$\int_{S^2} \phi(v') b\left(\left(\frac{v-v_*}{|v-v_*|}, \sigma\right)\right) d\sigma = \frac{1}{|C_{c,r,k}|} \int_{C_{c,r,k}} \phi(w) dw = \int_{\mathbb{R}^3} \phi(w) \Pi_{v,v_*,\theta_0}(dw).$$

Par conséquent

$$\begin{aligned} \int_{\mathbb{R}^3} Q^+(f, f) \phi(v) dv &= \int_{\mathbb{R}^3 \times \mathbb{R}^3} f f_* \int_{\mathbb{R}^3} \phi(w) \Pi_{v,v_*,\theta_0}(dw) dv dv_* \\ &= \int_{\mathbb{R}^3} \phi(w) \int_{\mathbb{R}^3 \times \mathbb{R}^3} f f_* \Pi_{v,v_*,\theta_0}(dw) dv dv_*. \end{aligned}$$

Ainsi pour toute probabilité  $f$ , le terme de gain  $Q^+(f, f)$  est représenté par la formule de Tanaka sous la forme

$$Q^+(f, f) = \int_{\mathbb{R}^6} f f_* \Pi_{v,v_*,\theta_0} dv dv_*,$$

c'est-à-dire aussi

$$Q^+(f, f) = E \Pi_{X, X^*, \theta_0}$$

où  $X$  et  $X^*$  sont des variables aléatoires quelconques de loi  $f$  •

On peut alors établir la propriété de contraction du terme de gain  $Q^+$ .

**Proposition 3.2** *On suppose que la section efficace  $b$  est de type atomique (3.2). Alors*

1. *pour toutes mesures de probabilité  $f$  et  $g$ , on a l'inégalité fonctionnelle*

$$W(Q^+(f, f), Q^+(g, g)) \leq W(f, g),$$

2. *si  $g$  est la distribution gaussienne ayant les mêmes trois premiers moments que la mesure de probabilité  $f$ , on a l'égalité fonctionnelle*

$$W(Q^+(f, f), Q^+(g, g)) = W(f, g)$$

*si et seulement si  $f = g$ .*

**Démonstration. 1.** On suppose d'abord que la section efficace  $b$  est de type atomique concentrée en un point  $\theta_0$ , c'est-à-dire vérifie  $b(\cos \theta) \sin \theta d\theta = \frac{1}{2\pi} \delta_{\theta_0}$ .

**1.1.** Etant données deux probabilités  $f$  et  $g$ , on va montrer l'inégalité fonctionnelle

$$W(Q^+(f, f), Q^+(g, g)) \leq W(f, g).$$

La formule de représentation de Tanaka et la convexité du carré de la distance de Wasserstein permettent d'écrire que

$$W(Q^+(f, f), Q^+(g, g))^2 = W(E\Pi_{X, X_*, \theta_0}, E\Pi_{Y, Y_*, \theta_0})^2 \leq EW(\Pi_{X, X_*, \theta_0}, \Pi_{Y, Y_*, \theta_0})^2$$

pour toutes variables aléatoires  $X$  et  $X_*$  de loi  $f$ ,  $Y$  et  $Y_*$  de loi  $g$ .

Or d'après la majoration de la proposition 2.7 appliquée à  $U_{c_1, r_1, k_1} = \Pi_{v, v_*, \theta_0}$  et  $U_{c_2, r_2, k_2} = \Pi_{w, w_*, \theta_0}$  où  $c_1$ ,  $r_1$  et  $k_1$  (resp.  $c_2$ ,  $r_2$  et  $k_2$ ) sont liés à  $v$ ,  $v_*$  et  $\theta_0$  (resp.  $w$ ,  $w_*$  et  $\theta_0$ ) comme dans la proposition 3.1, on a

$$\begin{aligned} & W(\Pi_{v, v_*, \theta_0}, \Pi_{w, w_*, \theta_0})^2 \\ & \leq \frac{1}{4} |(v + v_* + (v - v_*) \cos \theta_0) - (w + w_* + (w - w_*) \cos \theta_0)|^2 \\ & + \frac{|v - v_*|^2}{4} \sin^2 \theta_0 + \frac{|w - w_*|^2}{4} \sin^2 \theta_0 - \frac{|v - v_*||w - w_*|}{4} \sin^2 \theta_0 \left[1 + \frac{|(v - v_*, w - w_*)|}{|v - v_*||w - w_*|}\right] \\ & = \left| \frac{1 + \cos \theta_0}{2} (v - w) + \frac{1 - \cos \theta_0}{2} (v_* - w_*) \right|^2 \\ & + \frac{\sin^2 \theta_0}{4} [|v - v_*|^2 + |w - w_*|^2 - |v - v_*||w - w_*| - |(v - v_*, w - w_*)|] \\ & \leq \left| \frac{1 + \cos \theta_0}{2} (v - w) + \frac{1 - \cos \theta_0}{2} (v_* - w_*) \right|^2 + \frac{\sin^2 \theta_0}{4} |(v - v_*) - (w - w_*)|^2, \end{aligned}$$

car  $(v - v_*, w - w_*) \leq |v - v_*||w - w_*|$ .

On note que si les deux termes extrêmes des inégalités précédentes sont égaux alors nécessairement on a l'égalité  $(v - v_*, w - w_*) = |v - v_*||w - w_*|$ .

Si  $\pi$  minimise le transport entre  $f$  et  $g$ , on peut, d'après la proposition 2.4, choisir deux couples  $(X, Y)$  et  $(X_*, Y_*)$  de variables aléatoires de loi  $\pi$  telles que

1.  $X$  et  $X_*$  sont de loi  $f$ ,  $Y$  et  $Y_*$  sont de loi  $g$ ,
2.  $E|X - Y|^2 = E|X_* - Y_*|^2 = W(f, g)^2$ ,
3.  $(X, Y)$  et  $(X_*, Y_*)$  sont indépendants.

On déduit alors des inégalités précédentes et de 1. que

$$\begin{aligned} EW(\Pi_{X, X_*, \theta_0}, \Pi_{Y, Y_*, \theta_0})^2 & \leq E \left| \frac{1 + \cos \theta_0}{2} (X - Y) + \frac{1 - \cos \theta_0}{2} (X_* - Y_*) \right|^2 \\ & + \frac{\sin^2 \theta_0}{4} |(X - X_*) - (Y - Y_*)|^2, \end{aligned}$$

puis de 3. que

$$\begin{aligned} EW(\Pi_{X, X_*, \theta_0}, \Pi_{Y, Y_*, \theta_0})^2 & \leq \frac{(1 + \cos \theta_0)^2}{4} E(X - Y)^2 + \frac{(1 - \cos \theta_0)^2}{4} E(X_* - Y_*)^2 \\ & + \frac{\sin^2 \theta_0}{4} E(X - Y)^2 + \frac{\sin^2 \theta_0}{4} E(X_* - Y_*)^2 \\ & = E(X - Y)^2. \end{aligned}$$

Ainsi

$$W(Q^+(f, f), Q^+(g, g))^2 \leq E(X - Y)^2$$

et la condition 2. sur  $X$  et  $Y$  donnent alors l'inégalité fonctionnelle

$$W(Q^+(f, f), Q^+(g, g)) \leq W(f, g).$$

**Remarque.** Dans cette situation, on note en particulier que si on a l'égalité

$$W(Q^+(f, f), Q^+(g, g))^2 = E(X - Y)^2,$$

alors nécessairement on a l'égalité  $(X - X^*, Y - Y^*) = |X - X^*| |Y - Y^*|$ .

**1.2.** Soit maintenant  $f$  une probabilité et  $g$  la distribution gaussienne ayant les mêmes trois premiers moments que  $f$ . On va montrer que si on a l'égalité fonctionnelle

$$W(Q^+(f, f), Q^+(g, g)) = W(f, g)$$

alors

$$f = g.$$

On utilise maintenant les notations et les résultats de l'étape précédente.

Si  $(X, Y)$  et  $(X_*, Y_*)$  sont deux couples indépendants de variables aléatoires ayant pour loi le minimiseur  $\pi$  du transport entre  $f$  et  $g$ , alors d'après la remarque 2.5 on a  $X = u(Y)$  et  $X_* = u(Y_*)$  ps pour la même fonction  $u$ .

Si de plus on a l'égalité  $W(Q^+(f, f), Q^+(g, g)) = W(f, g)$ , alors d'après la remarque précédente, on a en particulier l'égalité

$$(X - X_*, Y - Y_*) = |X - X_*| |Y - Y_*|$$

soit

$$\frac{X - X_*}{|X - X_*|} = \frac{Y - Y_*}{|Y - Y_*|}.$$

Par conséquent

$$\frac{u(y) - u(x)}{|u(y) - u(x)|} = \frac{y - x}{|y - x|}$$

presque sûrement sur  $\mathbb{R}^3 \times \mathbb{R}^3$  pour la loi de  $(Y, Y_*)$ , puis pour la mesure de Lebesgue puisque  $Y$  et  $Y_*$  sont des variables aléatoires indépendantes de loi gaussienne. On considère alors un  $x_0$  tel que pour presque tout  $x$  dans  $\mathbb{R}^3$

$$u(x) = u(x_0) + \frac{|u(x) - u(x_0)|}{|x - x_0|} (x - x_0).$$

Alors pour presque tout  $x_1$  et  $x_2$  dans  $\mathbb{R}^3$

$$\frac{|u(x_1) - u(x_0)|}{|x_1 - x_0|} (x_1 - x_0) - \frac{|u(x_2) - u(x_0)|}{|x_2 - x_0|} (x_2 - x_0) = \frac{|u(x_1) - u(x_2)|}{|x_1 - x_2|} (x_1 - x_2).$$

Comme  $(x_0, x_1, x_2)$  est libre dans  $\mathbb{R}^3$  pour presque tout  $x_1$  et  $x_2$ , en particulier le coefficient de  $x_0$  doit être nul, c'est-à-dire

$$\frac{|u(x_1) - u(x_0)|}{|x_1 - x_0|} = \frac{|u(x_2) - u(x_0)|}{|x_2 - x_0|}.$$



Ainsi  $\frac{|u(x) - u(x_0)|}{|x - x_0|}$  est constante presque partout. Par conséquent il existe une constante  $\gamma$  telle que

$$u(x) = u(x_0) + \gamma(x - x_0),$$

soit pour des constantes  $\alpha$  et  $\beta$

$$X = \alpha + \beta Y.$$

Comme  $f$  et  $g$  ont leurs trois premiers moments égaux, on a

$$E(X) = E(Y) \quad E(X^2) = E(Y^2)$$

ce qui donne successivement  $\alpha = 0$ , puis  $\beta^2 = 1$ . Ainsi

$$X = \varepsilon Y \quad X^* = \varepsilon Y^*$$

pour le même  $\varepsilon = \pm 1$ . Or

$$(X - X_*, Y - Y_*) = |X - X_*||Y - Y_*|,$$

ce qui impose  $\varepsilon = +1$  et par conséquent

$$f = g.$$

**2.** On suppose maintenant que la section efficace  $b$  est de type atomique de la forme

$$b(\cos \theta) \sin \theta d\theta = \frac{1}{2\pi} \sum_{i=1}^M \alpha_i \delta_{\theta_i}.$$

Pour le noyau de collision particulier  $b_i$  concentré au point  $\theta_i$  de  $[0, \pi]$ , c'est-à-dire de la forme  $b_i(\cos \theta) \sin \theta d\theta = \frac{1}{2\pi} \delta_{\theta_i}$ , le terme de gain correspondant sera noté plus précisément  $Q_{\theta_i}^+$ .

Pour une mesure de probabilité  $f$ , on a donc

$$Q^+(f, f) = \sum_{i=1}^M \alpha_i Q_{\theta_i}^+(f, f),$$

et de plus  $Q^+(f, f)$  et les  $Q_{\theta_i}^+(f, f)$  sont aussi des mesures de probabilité.

Par conséquent pour des probabilités  $f$  et  $g$ , la convexité du carré de la distance de Wasserstein permet alors d'écrire que

$$\begin{aligned} W(Q^+(f, f), Q^+(g, g))^2 &= W\left(\sum_{i=1}^M \alpha_i Q_{\theta_i}^+(f, f), \sum_{i=1}^M \alpha_i Q_{\theta_i}^+(g, g)\right)^2 \\ &\leq \sum_{i=1}^M \alpha_i W(Q_{\theta_i}^+(f, f), Q_{\theta_i}^+(g, g))^2. \end{aligned}$$

**2.1.** La conclusion 1. montrée précédemment dans l'étape 1.1. pour chaque noyau de collision discret concentré en  $\theta_i$  permet alors d'en déduire que

$$W(Q^+(f, f), Q^+(g, g))^2 \leq \sum_{i=1}^M \alpha_i W(f, g)^2 = W(f, g).$$

**2.2.** - Si on a l'égalité

$$W(Q^+(f, f), Q^+(g, g)) = W(f, g),$$

alors pour tout  $\theta_i$  on a nécessairement l'égalité

$$W(Q_{\theta_i}^+(f, f), Q_{\theta_i}^+(g, g)) = W(f, g).$$

La conclusion 2. montrée précédemment dans l'étape 1.2. pour chaque noyau de collision discret concentré en  $\theta_i$  permet alors d'en déduire que  $f = g$  •

On peut maintenant établir le théorème de Tanaka qui assure en particulier une propriété de contraction de la distance de Wasserstein le long du flot associé à l'équation de Boltzmann. Plus précisément on a le

**Théorème 3.3 (de Tanaka)** *Soient  $b$  une section efficace vérifiant la condition (1.5),  $f_0$  et  $g_0$  deux mesures de probabilité vérifiant les conditions (1.4) de vitesse moyenne nulle et de température unité. Si  $f(t)$  et  $g(t)$  sont les solutions associées de l'équation de Boltzmann (1.2), alors*

1.  $W(f(t), g(t)) \leq W(f_0, g_0)$  pour tout  $t \geq 0$ ,
2.  $W(f(t), g(t)) \rightarrow 0$  quand  $t \rightarrow +\infty$ .

**Démonstration. 1.1.** On montre d'abord que si  $(b_n)_n$  est une suite de sections efficaces convergeant vers  $b$  au sens faible\* des mesures dans  $[-1, +1]$  et si la conclusion 1. du théorème de Tanaka est satisfaite pour les  $b_n$ , alors elle est aussi satisfaite pour  $b$ .

En effet si  $f_n$  est la solution de l'équation de Boltzmann associée à la section  $b_n$  pour la même donnée initiale  $f_0$ , alors d'après la proposition 1.7, la suite  $(f_n)_n$  approche la solution  $f$  au sens faible\*. On a un résultat analogue pour la solution  $g$ .

Le corollaire 1.11 assure alors que pour tout  $t \geq 0$

$$W(f(t), g(t)) \leq \liminf_{n \rightarrow +\infty} W(f_n(t), g_n(t)).$$

Si on suppose que la conclusion 1. du théorème de Tanaka est vraie pour chacune des solutions  $f_n$  et  $g_n$ , alors pour tout  $t \geq 0$  et tout  $n$  on a

$$W(f_n(t), g_n(t)) \leq W(f_n(0), g_n(0)) = W(f_0, g_0)$$

donc d'après l'inégalité précédente, on en déduit que tout  $t \geq 0$  on a aussi

$$W(f(t), g(t)) \leq W(f_0, g_0).$$

**1.2.** On suppose ici que la section efficace  $b$  est de type cutoff (3.1). On montre que si les solutions  $f$  et  $g$  de l'équation de Boltzmann (3.3) vérifient dans  $[0, +\infty[$  l'inéquation fonctionnelle

$$W(Q^+(f, f), Q^+(g, g)) \leq W(f, g),$$

alors elles satisfont la conclusion 1. du théorème de Tanaka.

Pour cela la convexité du carré de la distance de Wasserstein permet d'écrire que pour  $t > 0$  et  $s > 0$  on a

$$\frac{W(f(t+s), g(t+s))^2 - W(f(t), g(t))^2}{s} \leq W\left(f(t) + \frac{f(t+s) - f(t)}{s}, g(t) + \frac{g(t+s) - g(t)}{s}\right)^2 - W(f(t), g(t))^2.$$

Or d'une part  $f$  est dérivable d'après le théorème 1.3, donc on a

$$\frac{f(t+s) - f(t)}{s} \xrightarrow{s \rightarrow 0} \partial_t f(t)$$

quand  $s$  tend vers 0, et d'autre part par la conservation de l'énergie cinétique de la solution  $f$  on a pour tout  $s > 0$

$$\int \left( f(t, v) + \frac{f(t+s, v) - f(t, v)}{s} \right) |v|^2 dv = \int f(t, v) |v|^2 dv$$

$$\int (f(t, v) + \partial_t f(t, v)) |v|^2 dv = \int f(t, v) |v|^2 dv + \int Q(f, f)(t, v) |v|^2 dv = \int f(t, v) |v|^2 dv.$$

Comme on a un résultat analogue pour  $g$ , le corollaire 2.11 assure alors que

$$\lim_{s \rightarrow 0^+} W\left(f(t) + \frac{f(t+s) - f(t)}{s}, g(t) + \frac{g(t+s) - g(t)}{s}\right) = W(f(t) + \partial_t f(t), g(t) + \partial_t g(t)).$$

On en déduit donc que

$$\limsup_{s \rightarrow 0^+} \frac{W(f(t+s), g(t+s))^2 - W(f(t), g(t))^2}{s} \leq W(f(t) + \partial_t f(t), g(t) + \partial_t g(t))^2 - W(f(t), g(t))^2$$

et comme  $f$  et  $g$  sont des solutions de l'équation de Boltzmann (3.3), on a donc

$$= W(Q^+(f(t), f(t)), Q^+(g(t), g(t)))^2 - W(f(t), g(t))^2.$$

Si  $f$  et  $g$  vérifient l'inégalité fonctionnelle  $W(Q^+(f, f), Q^+(g, g)) \leq W(f, g)$ , il en résulte alors que pour tout  $t > 0$  on a

$$\limsup_{s \rightarrow 0^+} \frac{W(f(t+s), g(t+s))^2 - W(f(t), g(t))^2}{s} \leq 0.$$

Comme de plus d'après le corollaire 2.11, la fonction  $W(f, g)$  est continue en temps sur  $[0, +\infty[$ , la propriété précédente assure alors que la fonction  $W(f, g)^2$ , et donc aussi la fonction  $W(f, g)$  est décroissante en temps.

En effet de façon générale (cf. [Y] par exemple) si  $w$  est une fonction continue dans un intervalle  $[a, b]$  telle que pour tout  $t \in [a, b[$  on ait

$$\limsup_{s \rightarrow 0^+} \frac{w(t+s) - w(t)}{s} \leq 0$$

alors  $w(b) \leq w(a)$ . En effet si on suppose le contraire, on a  $w(b) - w(a) > \varepsilon(b - a)$  pour un  $\varepsilon > 0$  et la fonction  $v(t) = w(t) - w(a) - \varepsilon(t - a)$  vérifie

$$\limsup_{s \rightarrow 0^+} \frac{v(t+s) - v(t)}{s} = \limsup_{s \rightarrow 0^+} \frac{w(t+s) - w(t)}{s} - \varepsilon < 0$$

pour tout  $t \in [a, b[$ . Comme  $\limsup_{s \rightarrow 0^+} \frac{v(a+s) - v(a)}{s} < 0$  avec  $v(a) = 0$  il existe donc un  $t_0 > a$  tel que  $v(t_0) < 0$ , et comme  $v$  est continue avec  $v(b) > 0$ , il en résulte qu'il existe  $t_1$  avec  $t_0 < t_1 < b$  tel que  $v(t_1) = 0$  et  $v(t) > 0$  pour  $t_1 < t < b$ . On aurait donc  $\limsup_{s \rightarrow 0^+} \frac{v(t_1+s) - v(t_1)}{s} \geq 0$ , ce qui contredit l'hypothèse.

**Conséquence.** Grâce à la proposition 3.2 et aux deux étapes précédentes, la conclusion 1. du théorème de Tanaka est donc démontrée.

**Remarque.** La conclusion 1. du théorème de Tanaka est en fait vraie même si  $f_0$  et  $g_0$  n'ont pas le même moment d'ordre 2.

**2.1.** On montre d'abord que si  $f_0$  est une mesure de probabilité dans  $\mathbb{R}^3$  vérifiant les conditions (1.4) avec de plus

$$\int_{\mathbb{R}^3} |v|^{13} f_0(v) dv < +\infty,$$

alors la solution  $f(t)$  de l'équation de Boltzmann (1.2) de donnée initiale  $f_0$  vérifie

$$\lim_{t \rightarrow +\infty} W(f(t), M) = 0$$

où  $M$  est la Maxwellienne ayant les mêmes trois premiers moments que  $f_0$ , définie par

$$M(v) = (2\pi)^{-3/2} e^{-\frac{|v|^2}{2}}.$$

De façon générale les trois premiers moments de la Maxwellienne

$$G(v) = \left(\frac{\alpha}{\pi}\right)^{3/2} e^{-\alpha|v|^2}$$

où  $\alpha > 0$ , sont 1, 0 et  $\frac{3}{2\alpha}$ . En posant pour  $n \in \mathbb{N}_*$

$$G_n(v) = \left(\frac{n}{\pi}\right)^{3/2} e^{-n|v|^2}$$

$$f_{0n} = G_n * f_0$$

on vérifie que les trois premiers moments de  $f_{0n}$  sont 1, 0 et  $3(1 + \frac{1}{2n})$  et la Maxwellienne  $M_n$  ayant les mêmes trois premiers moments que  $f_{0n}$  est alors définie par

$$M_n(v) = \left(2 + \frac{1}{n}\right)^{-3/2} e^{-\frac{|v|^2}{2 + \frac{1}{n}}}.$$

D'une part la suite  $(M_n)_n$  converge vers  $M$  au sens faible\* des mesures dans  $\mathbb{R}^3$ , comme on le voit par le théorème de Lebesgue, et d'autre part la suite des moments d'ordre 2 de

$M_n$  converge vers le moment d'ordre 2 de  $M$ , comme on le voit par les formules explicites précédentes. Par conséquent, d'après le corollaire 2.12

$$\lim_{n \rightarrow +\infty} W(M_n, M) = 0.$$

De même

$$\lim_{n \rightarrow +\infty} W(f_{0n}, f_0) = 0.$$

Pour tout entier  $n \geq 1$  fixé, on note  $f_n(t)$  la solution de l'équation de Boltzmann (1.2) de donnée initiale  $f_{0n}$ ; la proposition 1.11 assure alors que  $f_n(t)$  converge vers  $M_n$  au sens faible\* quand  $t$  tend vers  $+\infty$ , et comme de plus par conservation de l'énergie cinétique on a

$$\int_{\mathbb{R}^3} |v|^2 f_n(t, v) dv = \int_{\mathbb{R}^3} |v| f_{0n}(v) dv = \int_{\mathbb{R}^3} |v|^2 M_n(v) dv,$$

on a aussi la convergence en distance de Wasserstein

$$\lim_{t \rightarrow +\infty} W(f_n(t), M_n) = 0.$$

Or d'après la première partie du théorème de Tanaka qui est en fait vraie même si les deux conditions initiales n'ont pas le même troisième moment, on peut écrire que pour tout  $n \geq 1$  et tout  $t \geq 0$

$$W(f(t), f_n(t)) \leq W(f_0, f_{0n}).$$

Ecrivant alors

$$\begin{aligned} W(f(t), M) &\leq W(f(t), f_n(t)) + W(f_n(t), M_n) + W(M_n, M) \\ &\leq W(f_0, f_{0n}) + W(f_n(t), M_n) + W(M_n, M), \end{aligned}$$

on en déduit que

$$\lim_{t \rightarrow +\infty} W(f(t), M) = 0.$$

**2.2.** Soit maintenant  $f_0$  une mesure de probabilité vérifiant seulement les conditions (1.4) de vitesse moyenne nulle et de température unité, et soit  $f(t)$  la solution de l'équation de Boltzmann (1.2) de donnée initiale  $f_0$ .

On va montrer que

$$\lim_{t \rightarrow +\infty} W(f(t), M) = 0.$$

On considère pour cela une suite  $(f_{0n})_n$  de données initiales convergeant au sens faible\* vers  $f_0$ , vérifiant les conditions (1.4) avec de plus

$$\int_{\mathbb{R}^3} |v|^{13} f_{0n}(v) dv < +\infty.$$

D'une part d'après le théorème 2.10 on a

$$\lim_{n \rightarrow +\infty} W(f_{0n}, f_0) = 0$$

et d'autre part, si  $f_n(t)$  est la solution associée à la donnée initiale  $f_{0n}$ , alors d'après l'étape précédente on a pour tout entier  $n$

$$\lim_{t \rightarrow +\infty} W(f_n(t), M) = 0.$$

Or par inégalité triangulaire et la conclusion 1. du théorème de Tanaka on a

$$\begin{aligned} W(f(t), M) &\leq W(f(t), f_n(t)) + W(f_n(t), M) \\ &\leq W(f_0, f_{0n}) + W(f_n(t), M). \end{aligned}$$

Soit alors  $\varepsilon > 0$  fixé. D'après ce qui précède, il existe  $n$  tel que

$$W(f_0, f_{0n}) \leq \varepsilon,$$

puis à ce  $n$  on associe  $t$  tel que

$$W(f_n(t), M) \leq \varepsilon.$$

La décroissance de  $W$  le long du flot assure alors que

$$\lim_{t \rightarrow +\infty} W(f(t), M) = 0.$$

**Conséquence.** Par l'étape précédente et par inégalité triangulaire, on a donc la conclusion 2. du théorème de Tanaka.

Le théorème peut aussi être écrit sous la forme du

**Corollaire 3.4** *Soient  $b$  une section efficace vérifiant la condition (1.5),  $f_0$  et  $g_0$  deux mesures de probabilité vérifiant les conditions (1.4) de moyenne nulle et de température unité. Si  $f(t)$  et  $g(t)$  sont les solutions associées de l'équation de Boltzmann (1.2), alors la fonction  $W(f(t), g(t))$  est continue décroissante de  $[0, +\infty[$  sur  $]0, W(f_0, g_0)[$ .*

## References

- [D] R. DUDLEY. *Probabilities and metrics: Convergence of laws on metric spaces, with a view to statistical testing*. Aarhus Universitet, (1976).
- [DL] R. DAUTRAY et J.L. LIONS. *L'opérateur de Laplace*. Analyse mathématique et calcul numérique, Chap. 2, Masson, Paris (1987).
- [E] T. ELMROTH. *Global Boundedness of moments of solutions of the Boltzmann equation for forces of infinite range*. Arch. Rational Mech. Anal. 82 (1983), p.1-12.
- [Ga] W. GANGBO. *The Monge Mass Transfer Problem and its Applications*, preprint.
- [Ge] P. GERARD. *Solutions globales du problème de Cauchy pour l'équation de Boltzmann (d'après R.J. DiPerna et P.L. Lions)*. Séminaire Bourbaki, Exp. 699 (1987-88).
- [Gr] A. GROTHENDIECK. *Espaces Vectoriels Topologiques*, Sao Paulo (1964).
- [IT] E. IKENBERRY and C. TRUESDELL. *On the pressures and the flux of energy in a gas according to Maxwell's kinetic theory I*. J. Rational Mech. Anal. 5 (1956), p.1-54.
- [P] B. PERTHAME. *Introduction to the collision models in Boltzmann's theory*. In Modeling of Collisions, P-A. Raviart, Ed., vol. 2 of *Series in Applied Mathematics*. Gauthier-Villars, Paris (1998).

- [S] L. SCHWARTZ. *Cours d'Analyse de l'Ecole Polytechnique*. Hermann. Paris
- [T] H. TANAKA. *Probabilistic treatment of the Boltzmann equation of Maxwellian molecules*. Z. Warsch. Verw. Gebiete 46, 1 (1978), p.67-105.
- [V1] C. VILLANI. *On a new class of weak solutions for the spatially homogeneous Boltzmann and Landau equations*. Arch. Rational Mech. Anal. 143, 3 (1998), p.273-307.
- [V2] C. VILLANI. *Mass transportation and partial differential equations*, à paraître dans un ouvrage en collaboration avec Y. Brenier et F. Otto.
- [V3] C. VILLANI. *A review of mathematical problems in collisional kinetic theory*, à paraître dans "Handbook of fluid mechanics", D. Serre et S. Friedlander, Eds.
- [Y] K. YOSIDA. *Functional Analysis*, Springer (1965).





## Projets de recherche

— o —

Dans ce qui précède nous avons étudié la convergence vers l'équilibre dans le cadre de l'équation de Boltzmann homogène, avec une section efficace Maxwellienne. Plus précisément le théorème de Tanaka assure la convergence de la solution vers la Maxwellienne ayant les mêmes trois premiers moments, et ce au sens de la distance de Wasserstein. La méthode employée est intéressante pour différentes raisons. D'une part les hypothèses sur la donnée initiale sont assez faibles puisqu'on se place dans le cadre très général d'une mesure de probabilité ayant ses trois premiers moments finis. D'autre part elle ne s'appuie pas de façon fondamentale sur la notion d'entropie: celle-ci a été introduite dans la démonstration du théorème uniquement pour vérifier que le résultat reste vrai sans l'hypothèse de cut-off, et il faut noter que dans l'article original Tanaka prouve le résultat sans utiliser l'entropie. En particulier la donnée initiale n'est pas supposée avoir une entropie finie. Le théorème de Tanaka nous rappelle ainsi que la convergence vers l'équilibre n'est pas uniquement liée au théorème H.

Il est à noter que dans ce même cadre de l'équation de Boltzmann homogène avec section efficace Maxwellienne, Toscani a introduit une distance, faisant intervenir la transformation de Fourier, possédant la même propriété de contraction le long du flot, et pour laquelle la solution converge vers l'équilibre.

De façon générale, l'étude de la convergence vers l'équilibre consiste en les problèmes suivants: étant donnée une équation décrivant un ensemble de particules (ou une classe de telles équations), existe-t-il un unique état d'équilibre, éventuellement sous certaines contraintes, de conservation de certaines quantités (masse, énergie,  $\dots$ ) par exemple? La solution converge-t-elle vers cet équilibre quand le temps tend vers l'infini? Dans quel sens et à quelle vitesse? Converge-t-elle en des temps comparables aux temps caractéristiques du système?

Notons que le théorème de Tanaka répond partiellement à ces questions, mais dans un cadre restreint. Depuis ces travaux, des méthodes liées à la notion d'**entropie** ont été développées pour donner des réponses plus globales et plus précises.

Dans le cas de l'équation de Boltzmann homogène, l'entropie de la solution  $f(t)$  a été définie par

$$H(f)(t) = \int_{\mathbb{R}^3} f(t, v) \ln f(t, v) dv.$$

Plus généralement, une fonction  $F$  de la solution  $f(t)$  est appelée une *entropie* du système si c'est une fonctionnelle de Lyapunov, c'est-à-dire si  $F(f)(t)$  est une fonction décroissante du temps.

Considérons par exemple un système de particules subissant un processus de diffusion, une force extérieure dérivant d'un potentiel  $V$  et une force intérieure dérivant d'un potentiel d'interaction interne  $W$ . Si  $\rho(t, x)$  désigne la densité de particules au point  $x$  et au temps  $t$ , alors  $\rho$  est gouverné par une équation de la forme

$$\partial_t \rho = \Delta P(\rho) + \nabla \cdot (\rho \nabla V) + \nabla \cdot (\rho \nabla (\rho *_x W)).$$

Notons que contrairement aux équations cinétiques, il n'y a pas de variable de vitesse dans l'espace des phases.

A cette équation nous associons une fonctionnelle, appelée énergie libre,

$$F(\rho) = \int U(\rho(x)) dx + \int \rho(x) V(x) dx + \frac{1}{2} \int \rho(x)\rho(y)W(x-y) dx dy,$$

qui s'écrit comme la somme d'une énergie interne, d'une énergie potentielle et d'une énergie d'interaction.  $U$  est liée à  $P$  par  $U''(\rho) = \frac{P'(\rho)}{\rho}$  et  $U(0) = 0$ ; dans le cas particulier de l'équation de MacKean-Vlasov où  $P(\rho) = \rho$ , on obtient  $U(\rho) = \rho \ln \rho$  et on reconnaît dans l'énergie interne l'entropie définie pour l'équation de Boltzmann. Suivant le choix de  $P, V$  et  $W$ , on reconnaît par exemple dans le modèle donné précédemment l'équation de la chaleur

$$\partial_t \rho = \Delta \rho,$$

l'équation des milieux poreux

$$\partial_t \rho = \Delta \rho^m,$$

ou l'équation de Fokker-Planck linéaire

$$\partial_t \rho = \Delta \rho + \nabla \cdot (\rho \nabla V).$$

Le point essentiel est le fait que  $F$  soit une fonctionnelle de Lyapunov pour l'équation considérée; l'étude du comportement asymptotique de la solution - qui nous intéresse ici - consiste alors en les questions suivante:

1.  $F$  admet-elle un unique minimum  $\rho_\infty$ ?
2. La solution  $\rho(t)$  converge-t-elle vers  $\rho_\infty$ , et en quel sens?
3. Avec quelle vitesse converge-t-elle?

Pour cela on pourra mesurer la distance entre  $\rho(t)$  et  $\rho_\infty$  de diverses manières, mais essentiellement

- via des normes  $L^p$ : par exemple une analyse spectrale d'équations de ce type pourrait conduire à des espaces  $L^2$  (avec poids). Cependant ce ne sont pas des espaces naturels pour des distributions de probabilité, et on considère plutôt  $L^1$ ;
- via la distance de Wasserstein  $W_2$ , avec un coût quadratique comme précédemment;
- via l'entropie: on introduit l'*entropie relative*

$$F(\rho|\rho_\infty) = F(\rho) - F(\rho_\infty)$$

et on s'intéresse à la convergence vers 0 de cette quantité.

Dans la suite de mon travail je compte considérer en particulier les deux derniers points. Pour cela les classes suivantes d'inégalités fonctionnelles seront utiles:

1. des inégalités entre l'entropie  $F$  et la dissipation d'entropie

$$D(\rho) = -\frac{d}{dt} F(\rho|\rho_\infty).$$

Ce sont des inégalités de la forme

$$D(\rho) \geq \Phi(F(\rho|\rho_\infty))$$

où  $\Phi$  est une fonction continue croissante vérifiant  $\Phi(0) = 0$ . Cette inégalité devra être vérifiée au moins pour une classe de densités  $\rho$  contenant les solutions de l'équation. Alors elle pourra assurer la convergence vers l'équilibre au sens de l'entropie, avec éventuellement une vitesse explicite: par exemple si  $\Phi(x) = Kx$  pour un  $K > 0$ , alors la convergence est au moins exponentielle.

2. des inégalités entre l'entropie et la distance de Wasserstein du type

$$W_2(\rho, \rho_\infty) \leq \Psi(F(\rho|\rho_\infty))$$

où de même  $\Psi$  est une fonction continue croissante vérifiant  $\Psi(0) = 0$ . Dès qu'une telle inégalité est vérifiée pour la solution (en tout temps), la convergence au sens de l'entropie (avec une vitesse de convergence) assure la convergence au sens de la distance de Wasserstein (avec éventuellement une vitesse explicite de convergence).

3. Le lien avec la norme  $L^1$  pourra alors être donné par des inégalités du type

$$\|\rho - \rho_\infty\|_{L^1} \leq \Lambda(F(\rho|\rho_\infty))$$

qui assureront à leur tour les mêmes résultats de convergence pour la norme  $L^1$ .

Dans le cas de l'équation de Fokker-Planck linéaire

$$\partial_t \rho = \Delta \rho + \nabla \cdot (\rho \nabla V),$$

l'entropie atteint son minimum en  $\rho_\infty = e^{-V}$ . En supposant  $\int e^{-V(x)} dx = 1$ , l'entropie relative

$$F(\rho, \rho_\infty) = H(\rho, \rho_\infty) = \int \rho \ln \frac{\rho}{e^{-V}} dx$$

s'appelle l'*entropie relative de Kullback*, et la dissipation d'entropie

$$D(\rho, \rho_\infty) = \int \rho \left| \nabla \ln \frac{\rho}{e^{-V}} \right|^2 dx$$

porte le nom d'*information de Fisher (relative)* et sera notée  $I(\rho, \rho_\infty)$ .

Les inégalités fonctionnelles mentionnées précédemment sont ici respectivement:

1. l'**inégalité de Sobolev logarithmique**:

$$\forall \rho \quad I(\rho, \rho_\infty) \geq 2\lambda H(\rho, \rho_\infty),$$

2. l'**inégalité de transport de Talagrand**:

$$\forall \rho \quad W_2(\rho, \rho_\infty) \leq \sqrt{\frac{2}{\lambda} H(\rho, \rho_\infty)},$$

3. l'**inégalité de Csiszár-Kullback**:

$$\forall \rho \quad \|\rho - \rho_\infty\|_{L^1} \leq \sqrt{2H(\rho, \rho_\infty)}.$$

La troisième inégalité est satisfaite pour tout  $\rho_\infty$ . Les deux premières sont vraies sous certaines conditions sur  $\rho_\infty$ , par exemple dès que  $\rho_\infty$  vérifie le critère de Bakry-Emery:  $D^2V \geq \lambda I_n$ ,  $\lambda > 0$  (cf. [BE], [C], [OV] et [BGL] pour différentes preuves). Il est à noter

que l'inégalité de Sobolev logarithmique implique l'inégalité de Talagrand (cf. [OV] et [BGL]). Si  $V$  satisfait le critère de Bakry-Emery, ces résultats assurent la convergence exponentielle de la solution de l'équation de Fokker-Planck, au sens de l'entropie, de la distance de Wasserstein et de la norme  $L^1$ .

Ces dernières années, certains des résultats exposés précédemment pour l'équation de Fokker-Planck ont été généralisés à d'autres équations, en particulier l'équation des milieux poreux (cf. [CT], [DPD] ou [O]) et des équations faisant intervenir une énergie d'interaction interne, sous des hypothèses d'uniforme convexité sur  $V$  et  $W$  (cf. [CMCV]).

De nouveaux points de vue ont également été donnés sur ce sujet et ont apporté des résultats et des perspectives très intéressants, en particulier la formulation de problèmes duaux et de liens avec les équations de Hamilton-Jacobi (cf [BGL]), l'élaboration d'une théorie de la perturbation, et l'utilisation d'équations différentielles stochastiques.

Cette dernière approche, développée en particulier par Malrieu dans [M], consiste dans un premier temps à interpréter la mesure de probabilité solution de l'équation de MacKean-Vlasov comme la loi d'un processus de Markov, solution d'une équation différentielle stochastique non linéaire. Dans un deuxième temps Malrieu approche ce processus non linéaire par un système de  $N$  processus linéaires couplés représentant  $N$  particules en interaction via leur mesure empirique.

Trois éléments montrent la force de cette méthode: d'une part le système de particules obtenu, dit *de champ moyen*, est plus facile à étudier que l'équation initiale, bien que la dimension de l'espace sur lequel il est défini soit grande. D'autre part, quand le nombre  $N$  de particules tend vers l'infini, la description du système de particules donnée par ce système de processus se rapproche de la description donnée par l'équation initiale, dans le sens où la loi (empirique) d'une particule suivant ce processus approché converge vers la loi d'une particule suivant le processus non linéaire original. Enfin, et surtout, l'utilisation d'inégalités de Sobolev logarithmiques permet de préciser cette convergence: au niveau du système de processus couplés, Malrieu obtient différentes estimations faisant intervenir des constantes *indépendantes* du nombre de particules, ce qui lui permet par la suite de donner une *vitesse explicite de convergence*. Sous certaines conditions de convexité sur les potentiels, il utilise finalement ce système de particules pour montrer un résultat de convergence vers l'équilibre pour l'équation de MacKean-Vlasov, au sens de la distance de Wasserstein et de la variation totale.

Nous avons déjà remarqué que la distance de Wasserstein s'est révélée être très bien adaptée à l'étude de la convergence vers l'équilibre dans le cadre de l'équation de Boltzmann - comme le met en évidence le théorème de Tanaka - ou de Fokker-Planck - comme nous l'avons vu précédemment. Les travaux de Malrieu - et avant eux ceux, précurseurs, de Sznitman (cf. [S]) - montrent à leur tour tout l'intérêt de cette distance dans l'étude de ce passage d'un système de particules (en très grand nombre) à un modèle continu.

Ces travaux ayant été réalisés sous des hypothèses assez contraignantes, pourrait-on maintenant généraliser cette approche, en affaiblissant les hypothèses et en dégagant les principes analytiques fondamentaux de cette étude?

## References

- [BE] D. BAKRY et M. EMERY. *Diffusions hypercontractives*. In Sém. Proba. XIX, n° 1123 in Lecture Notes in Math. Springer (1985).
- [BGL] S. BOBKOV, I. GENTIL et M. LEDOUX. *Hypercontractivity of Hamilton-Jacobi equations*. J. Math. Pures Appl. 80, 7 (2001).
- [BL] S. BOBKOV et M. LEDOUX. *From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities*. Geom. Funct. Anal. 10, 5 (2000).
- [C] L. A. CAFFARELLI. *Monotonicity properties of optimal transportation and the FKG and related inequalities*. Comm. Math. Phys. 214, 3 (2000).
- [CMCV] J. CARRILLO, R. MCCANN et C. VILLANI. *Kinetic equilibration rates for granular media*. Work in progress.
- [CT] J. CARRILLO et G. TOSCANI. *Asymptotic  $L^1$ -decay of solutions of the porous medium equation to self-similarity*. Indiana Univ. Math. J. 49, 1 (2000).
- [DPD] M. DELPINO et J. DOLBEAULT. *Generalized Sobolev inequalities and asymptotic behaviour in fast diffusion and porous medium problems*. Preprint Univ Paris IX-Dauphine, Ceremade, n° 9905 (1999).
- [M] F. MALRIEU. *Logarithmic Sobolev inequalities for some nonlinear PDEs*. Preprint.
- [O] F. OTTO. *The geometry of dissipative evolution equations: the porous medium equations*. Comm. PDE 26, 1-2 (2001).
- [OV] F. OTTO et C. VILLANI. *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*. J. Funct. Anal. 173, 2 (2000).
- [S] A.-S. SZNITMAN. *Topics in propagation of chaos*. In Ecole d'été de Probabilités de Saint-Flour XIX-1989. Lectures notes in Math. 1464, Springer (1991).