

(Ir)réversibilité et entropie

Cédric VILLANI
Université de Lyon
& Institut Henri Poincaré
11 rue Pierre et Marie Curie
75231 Paris Cedex 05, FRANCE

La cosa più meravigliosa è la felicità del momento

L. Ferré

La **flèche du temps** fait partie de notre expérience sensible et nous en faisons l'expérience chaque jour : les miroirs brisés ne se recollent pas, les êtres humains ne rajeunissent pas, les cernes croissent sans cesse dans les troncs des arbres, nous avons la mémoire des événements passés et pas des événements futurs. En somme, le temps s'écoule toujours dans le même sens ! Pourtant, les lois fondamentales de la physique classique ne privilégient aucune direction du temps et obéissent à une rigoureuse symétrie entre passé et futur. Il est possible, comme discuté dans l'article de T. Damour dans ce même volume, que l'irréversibilité soit inscrite dans d'autres lois de la physique, par exemple du côté de la relativité générale ou de la mécanique quantique. Depuis Boltzmann, la **physique statistique** avance une autre explication : la flèche du temps traduit un flot constant des événements moins probables vers les événements plus probables. Avant de continuer sur cette interprétation qui constitue le fil directeur de tout l'exposé, je noterai que l'écoulement du temps n'est pas forcément basé sur une explication unique.

Au premier examen, la suggestion de Boltzmann semble saugrenue : ce n'est pas parce qu'un événement est *probable* qu'il va se réaliser *effectivement*, or la flèche du temps semble inexorable et ne tolérer aucune exception. La réponse à cette objection tient dans un slogan : la **séparation d'échelles**. Si les lois fondamentales de la physique s'exercent au niveau microscopique, particulière (atomes, molécules...), les phénomènes que nous pouvons sentir ou mesurer mettent en jeu un nombre considérable de particules. L'effet de ce nombre est d'autant plus grand qu'il intervient dans des calculs combinatoires : si N , le nombre d'atomes participant à une expérience, est de l'ordre de 10^{10} , c'est déjà considérable, mais $N!$ ou 2^N sont des nombres surnaturellement grands, invincibles.

Les innombrables débats entre physiciens qui se sont ensuivis pendant plus d'un siècle, et se poursuivent encore aujourd'hui, témoignent de la subtilité et de la profondeur des arguments de Maxwell et Boltzmann, porte-drapeaux d'une petite révolution scientifique qui s'accomplit dans les années 1860 et 1870, et qui vit naître les fondements de la théorie cinétique des gaz moderne, le concept universel d'entropie statistique, et la notion d'irréversibilité macroscopique. À dire vrai, les arguments étaient si subtils que Maxwell et Boltzmann s'y sont eux-mêmes parfois perdus, hésitant sur certaines interprétations, alternant les erreurs naïves avec les concepts profonds; les plus grands scientifiques de la fin du dix-neuvième siècle, comme Poincaré ou Lord Kelvin, n'ont pas été en reste. On trouvera un aperçu de

ces atermoiements dans le texte de Damour déjà cité; pour ma part je me contenterai de présenter une version “décantée” de la théorie de Boltzmann. On évoquera à la fin de ce texte la façon dont Landau fit voler en éclat le paradigme de Boltzmann, découvrant une apparente irréversibilité là où il ne semblait pas y en avoir, et ouvrant une nouvelle mine de problèmes mathématiques.

En retraçant l’histoire de l’interprétation statistique de la flèche du temps, nous aurons l’occasion d’effectuer un voyage au cœur de problèmes profonds qui depuis plus d’un siècle agitent mathématiciens et physiciens.

Les notations utilisées dans cet exposé sont dans l’ensemble classiques; je noterai $\mathbb{N} = \{1, 2, 3, \dots\}$ et $\log =$ logarithme népérien.

1 Le royaume inaccessible de Newton

On va adopter ici une description purement classique de notre univers physique, selon les lois édictées par Newton : l’espace ambiant est euclidien, le temps absolu, et l’accélération est égale au produit de la masse par la résultante des forces.

Dans le cas de la description d’un gaz, ces hypothèses sont discutables : d’après E.G.D. Cohen, les fluctuations quantiques ne sont pas négligeables au niveau mésoscopique. La nature probabiliste de la mécanique quantique est toujours débattue; admettons cependant que l’incertitude accrue qui résulterait de la prise en compte de ces fluctuations ne puisse qu’arranger nos affaires, au moins qualitativement, et concentrons-nous donc sur des modèles classiques et déterministes, “à la Newton”.

1.1 Le modèle des sphères dures

Pour fixer les idées, considérons un système de particules sphériques idéales rebondissant les unes sur les autres : soient N particules dans une boîte Λ , on désigne par $X_i(t)$ la position au temps t du centre de la particule numérotée i . Les règles du mouvement s’énoncent comme suit :

- On suppose qu’initialement les particules sont bien séparées ($i \neq j \implies |X_i - X_j| > 2r$) et séparées de la paroi ($d(X_i, \partial\Lambda) > r$ pour tout i).
- Tant que ces conditions de séparation sont satisfaites, le mouvement est rectiligne uniforme : $\ddot{X}_i(t) = 0$ pour tout i , où l’on note $\ddot{X} = d^2X/dt^2$ l’accélération de X .
- Quand deux particules se rencontrent, leurs vitesses changent brutalement selon les lois de Descartes : si $|X_i(t) - X_j(t)| = 2r$, alors

$$\begin{cases} \dot{X}_i(t^+) = \dot{X}_i(t^-) - 2\langle \dot{X}_i(t^-) - \dot{X}_j(t^-), n_{ij} \rangle n_{ij}, \\ \dot{X}_j(t^+) = \dot{X}_j(t^-) - 2\langle \dot{X}_j(t^-) - \dot{X}_i(t^-), n_{ji} \rangle n_{ji}, \end{cases}$$

où l’on note $n_{ij} = (X_i - X_j)/|X_i - X_j|$ le vecteur unitaire joignant les centres des boules en collision.

- Quand une particule rencontre le bord, sa vitesse change aussi : si $|X_i - x| = r$ avec $x \in \partial\Lambda$, alors

$$\dot{X}_i(t^+) = \dot{X}_i(t^-) - 2\langle \dot{X}_i(t^-), n(x) \rangle n(x),$$

où $n(x)$ est la normale extérieure à Λ en x , supposée bien définie.

Ces règles ne sont pas suffisantes pour définir complètement la dynamique : on ne peut a priori exclure les possibilités de collision triple, collisions simultanées entre particules et le bord, ou encore d'une infinité de collisions se produisant en temps fini. Cependant ces événements sont de probabilité nulle si les conditions initiales sont tirées au hasard selon la mesure de Lebesgue (ou mesure de Liouville) dans l'espace des phases [40, Appendice 4.A]; on négligera donc ces éventualités. La dynamique ainsi définie, toute simple qu'elle soit, peut alors être considérée comme une caricature de notre univers complexe si le nombre N de particules est très grand. Étudiée depuis plus d'un siècle, cette caricature n'a pas encore livré tous ses secrets, il s'en faut de beaucoup.

1.2 Autres modèles newtoniens

À partir du modèle emblématique des sphères dures, on peut définir un certain nombre de variantes plus ou moins complexes :

- remplacer la dimension 3 par une dimension $d \geq 2$ arbitraire (la dimension 1 étant probablement pathologique);
- remplacer la condition aux limites (rebond élastique sur les parois) par une loi plus complexe [40, Chapitre 8];
- ou au contraire éliminer les bords, toujours délicats, en posant le système dans l'espace entier \mathbb{R}^d (mais on peut alors argumenter que le nombre de particules devrait être infini pour conserver une densité moyenne globale non nulle) ou dans un tore de côté L , $\mathbb{T}_L^d = \mathbb{R}^d / (L\mathbb{Z}^d)$, ce qui sera mon choix préférentiel dans la suite;
- remplacer l'interaction de contact des sphères dures par une autre interaction entre particules ponctuelles, par exemple associée à un potentiel d'interaction à deux corps $\phi(x - y) =$ potentiel exercé au point x par un point matériel situé en y .

Parmi les potentiels d'interaction notables, on mentionne, en dimension 3, à constante multiplicative près :

- le potentiel **coulombien** : $\phi(x - y) = 1/|x - y|$;
- le potentiel **newtonien** : $\phi(x - y) = -1/|x - y|$;
- le potentiel **maxwellien** : $\phi(x - y) = 1/|x - y|^4$.

L'interaction maxwellienne a été introduite artificiellement par Maxwell et Boltzmann dans le cadre de l'étude statistique des gaz; elle mène à d'importantes simplifications dans certaines formules. Il existe une zoologie d'autres potentiels (Lennard-Jones, Manev ...). Les sphères dures correspondent au cas limite d'un potentiel qui vaudrait 0 pour $|x - y| > r$ et $+\infty$ pour $|x - y| < 2r$.

Supposant plus généralement que l'interaction se fait sur une échelle d'ordre r et avec l'intensité a , on arrive à un **système de particules ponctuelles avec potentiel d'interaction** :

$$\ddot{X}_i(t) = -a \sum_{j \neq i} \nabla \phi \left(\frac{X_i - X_j}{r} \right), \quad (1)$$

pour tout $i \in \{1, \dots, N\}$; on suppose donc $X_i \in \mathbb{T}_L^d$. Ici encore, la dynamique est bien définie en-dehors d'un ensemble de conditions initiales exceptionnelles, et associée à un **flot newtonien** : \mathcal{N}_t qui à la configuration au temps s associe la position au temps $s + t$ ($t \in \mathbb{R}$ peut être positif ou négatif).

1.3 Fonctions de distribution

Même si l'on accepte le modèle newtonien (1), il nous reste *inaccessible* : d'abord parce que nous ne percevons pas les particules individuellement (trop petites), et parce que leur nombre N est considérable. Avec des expériences bien choisies, on peut mesurer la pression exercée sur une petite surface, la température autour d'un point, la densité moyenne, etc. Toutes ces quantités ne s'expriment pas directement en fonction des X_i , mais plutôt en fonction de *quantités moyennes*

$$\frac{1}{N} \sum_i \chi(X_i, \dot{X}_i), \quad (2)$$

où χ est une fonction scalaire.

La distinction peut sembler oiseuse : en particulierisant χ , en le concentrant près de la particule i , on retrouve l'information manquante. Mais bien évidemment, cela est impossible : en pratique χ est à variation *macroscopique*, par exemple de l'ordre de la taille de la boîte. En outre, l'information contenue dans les moyennes (2) ne distingue pas les particules, de sorte que l'on a remplacé le vecteur des (X_i, \dot{X}_i) par la **mesure empirique**

$$\widehat{\mu}_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{(X_i(t), \dot{X}_i(t))}. \quad (3)$$

La terminologie "empirique" est bien choisie : c'est la mesure que l'on observe au moyen (sans jeu de mots) de mesures.

Pour résumer : notre connaissance du système de particules s'effectue uniquement à travers le comportement de la mesure empirique dans une *topologie faible* qui modélise la limitation macroscopique de nos expériences – aussi bien expériences de laboratoire qu'expériences sensibles.

Très souvent, à notre échelle, la mesure empirique apparaît continue :

$$\widehat{\mu}_t^N(dx dv) \simeq f(t, x, v) dx dv.$$

On notera souvent $f(t, \cdot) = f_t$. La densité f est la **distribution cinétique** du gaz. L'étude de cette distribution constitue la théorie cinétique des gaz; le fondateur de cette science est sans doute D. Bernoulli (vers 1738), et les plus fameux contributeurs en sont Maxwell et Boltzmann. On trouvera une brève histoire de la théorie cinétique dans [40, Chapitre 1] et les références incluses.

Continuons l'étude du système newtonien. On peut imaginer que certaines expériences permettent des *mesures simultanées* des paramètres de plusieurs particules, donnant ainsi accès à des corrélations entre particules. Ceci nous mène à définir par exemple

$$\widehat{\mu}_t^{2;N}(dx_1 dv_1 dx_2 dv_2) = \frac{1}{N(N-1)} \sum_{i \neq j} \delta_{(X_{i_1}(t), \dot{X}_{i_2}(t), X_{i_2}(t), \dot{X}_{i_2}(t))},$$

ou plus généralement

$$\widehat{\mu}_t^{k;N}(dx_1 dv_1 \dots dx_k dv_k) = \frac{(N-k-1)!}{N!} \sum_{(i_1, \dots, i_k) \text{ distincts}} \delta_{(X_{i_1}(t), \dot{X}_{i_1}(t), \dots, X_{i_k}(t), \dot{X}_{i_k}(t))}.$$

Les approximations correspondantes sont les **fonctions de distribution à k particules** :

$$\widehat{\mu}_t^{k;N}(dx_1 dv_1 \dots dx_k dv_k) \simeq f^{(k)}(t, x_1, v_1, \dots, x_k, v_k).$$

Évidemment, en continuant jusqu'à $k = N$, on trouve une mesure $\widehat{\mu}^{N;N}(dx_1 \dots dv_N)$ concentrée sur le vecteur des positions et vitesses des particules (moyenné sur tous les choix de permutations des particules). Mais en pratique on ne va jamais jusqu'à $k = N$: k reste très petit (aller jusqu'à 3 serait déjà un exploit), alors que N est gigantesque.

1.4 Aléatoire microscopique

Malgré le déterminisme du modèle newtonien, on a déjà fait des hypothèses de nature probabiliste sur les données initiales, en supposant qu'elles ne sont pas configurées de sorte à aboutir à une catastrophe rare telle qu'une collision triple. On peut maintenant généraliser cette approche en considérant une distribution de probabilité sur l'ensemble des positions et vitesses initiales :

$$\mu_0^N(dx_1 dv_1 \dots dx_N dv_N),$$

que l'on appellera **mesure de probabilité microscopique**. Dans la suite on notera pour abréger

$$dx^N dv^N := dx_1 dv_1 \dots dx_N dv_N.$$

Il est naturel de choisir μ_0^N symétrique, c'est-à-dire invariante par permutation des coordonnées. La donnée μ_0^N remplace la mesure $\widehat{\mu}_0^{N;N}$ et la généralise. Elle donne lieu à un flot de mesures, obtenu par l'action du flot :

$$\mu_t^N = (\mathcal{N}_t)_\# \mu_0^N,$$

et des marginales

$$\mu_t^{k;N} = \int_{(x_1, v_1, \dots, x_k, v_k)} \mu_t^N.$$

Si le sens de la mesure empirique est transparent (c'est la "vraie" densité de particules), celui de la mesure de probabilité microscopique est moins évident. Imaginons que l'état initial a été préparé par un grand concours de circonstances dont nous ne savons pas grand chose : on peut seulement faire des suppositions et des devinettes. Ainsi μ_0^N est une mesure de probabilité sur l'ensemble des possibles configurations initiales. Un énoncé physique faisant intervenir μ_0^N n'aura donc guère de sens s'il utilise la forme précise de cette distribution (nous ne pourrions le vérifier, car nous n'observons pas μ_0^N); mais il en aura un s'il énonce une propriété μ_0^N -presque sûre, ou bien sûre avec μ_0^N -probabilité 0.99 ou davantage.

De même, la forme de $\mu_t^{1;N}$ n'a guère de sens physique. Mais si l'on a un phénomène de *concentration de la mesure* dû au gigantisme de N , alors on peut espérer que

$$\mu_0^N \left[\text{dist} \left(\widehat{\mu}_t^N, f_t(x, v) dx dv \right) \geq r \right] \leq \alpha(N, r),$$

où dist est une distance bien choisie sur l'espace des mesures et $\alpha(N, r) \rightarrow 0$ quand $r \rightarrow \infty$, d'autant plus vite que N est grand (par exemple $\alpha(N, r) = e^{-cNr}$). On

aura alors

$$\begin{aligned} \text{dist}(\mu_t^{1;N}, f_t(x, v) dx dv) &= \text{dist}\left(\int \widehat{\mu}_t^N d\mu_t^N, f_t(x, v) dx dv\right) \\ &\leq \int \text{dist}(\widehat{\mu}_t^N, f_t dx dv) \\ &\leq \int_0^\infty \alpha(N, r) dr =: \eta(N). \end{aligned}$$

Si $\eta(N) \rightarrow 0$ quand $N \rightarrow \infty$, il s'ensuivra que, avec très forte probabilité, $\mu_t^{1;N}$ est une excellente approximation de $f(t, x, v) dx dv$, qui lui-même est une bonne approximation de $\widehat{\mu}_t^N$.

1.5 Micromégas

Dans cette section on a introduit deux descriptions statistiques très différentes : la distribution macroscopique $f(t, x, v) dx dv$, et les probabilités microscopiques $\mu_t^N(dx^N dv^N)$. Bien sûr, la quantité d'information contenue dans μ_t^N est considérablement plus importante que celle qui est contenue dans la distribution macroscopique : cette dernière nous renseigne sur l'état d'une particule typique, alors qu'un tirage selon la distribution μ_t^N nous renseigne sur l'état de *toutes* les particules. Pensons que si l'on a 10^{20} degrés de liberté, il s'agit d'intégrer sur 99999999999999999999 d'entre eux. Pour manipuler des dimensions aussi vertigineuses, nous aurons besoin d'un concept fondamental : l'entropie.

2 Le monde entropique

Le concept et le nom d'entropie ont été introduits par Clausius en 1865 comme une partie de la théorie – alors en construction – de la thermodynamique. Quelques années plus tard, Boltzmann révolutionnait le concept en lui donnant une interprétation statistique basée sur la théorie atomique. En plus de cette section, on pourra consulter par exemple Balian [9, 10] sur la notion d'entropie en physique statistique.

2.1 Formule de Boltzmann

Soit un système physique, que l'on suppose complètement décrit par son état microscopique $z \in \mathcal{Z}$. L'expérience ne nous donne en général accès qu'à une description partielle de l'état, disons $\pi(z) \in \mathcal{Y}$, où \mathcal{Y} est un espace d'états macroscopiques. Je ne donnerai pas d'hypothèses précises sur les espaces \mathcal{Z} et \mathcal{Y} , mais dès que l'on introduira de la théorie de la mesure on supposera implicitement que ce sont des espaces polonais (métriques séparables complets).

Comment estimer la quantité d'information perdue quand on résume l'information microscopique par l'information macroscopique ? Supposons \mathcal{Y} et \mathcal{Z} dénombrables, alors il est naturel de penser que l'incertitude associée à un état $y \in \mathcal{Y}$ est fonction du cardinal de la pré-image, soit $\#\pi^{-1}(y)$.

Maintenant si l'on effectue deux mesures indépendantes de deux systèmes différents, on a envie de dire que les incertitudes s'ajoutent. Or, avec des notations évidentes, $\#\pi^{-1}(y_1, y_2) = (\#\pi_1^{-1}(y_1)) (\#\pi_2^{-1}(y_2))$. Pour passer de cette opération

multiplicative à une addition, on prendra un multiple du logarithme. On aboutit ainsi à la célèbre formule de Boltzmann, gravée sur sa tombe du Cimetière central de Vienne :

$$S = k \log W, \quad (4)$$

où $W = \#\pi^{-1}(y)$ est le nombre d'états microscopiques compatibles avec l'état macroscopique y observé. Ici k est une constante physique, plus précisément la constante de Planck.

Dans de nombreux cas, l'espace \mathcal{Z} des configurations microscopiques est continu, et pour appliquer la formule de Boltzmann il convient de remplacer la mesure de comptage par une mesure privilégiée : par exemple la mesure de Liouville si l'on s'intéresse à un système hamiltonien. Ainsi W dans (4) pourra être *volume* d'états microscopiques compatibles avec l'état macroscopique y .

Si l'espace \mathcal{Y} des configurations macroscopiques est également continu, alors cette notion de volume doit être maniée avec prudence : la fibre $\pi^{-1}(y)$ est typiquement de volume nul, ce qui n'a guère d'intérêt. On peut tenter de poser, pour une topologie donnée,

$$S(y) = \text{p.f.}_{\varepsilon \rightarrow 0} \log |\pi^{-1}(B_\varepsilon(y))|,$$

où $B_\varepsilon(y)$ est la boule de rayon ε centrée en y et p.f. désigne la partie finie, c'est à dire que l'on retranche la divergence en ε , si tant est qu'elle ait un comportement universel.

Si ce dernier point n'a en général rien d'évident, l'universalité est cependant vérifiée dans le cas particulier qui nous intéresse où l'état microscopique \mathcal{Z} est l'espace des configurations de N particules, soit \mathcal{Y}^N , et où l'on commence par prendre *la limite* $N \rightarrow \infty$. Dans cette limite, comme on va le voir, l'entropie moyenne par particule tend vers une valeur finie, et l'on pourra ensuite prendre la limite $\varepsilon \rightarrow 0$, qui correspond à une mesure *macroscopique* arbitrairement précise. Le résultat n'est autre que la fameuse *fonction* H de Boltzmann, à un signe près.

2.2 De l'entropie à la fonction H

Appliquons les considérations précédentes avec un espace macroscopique constitué de k états différents : un état macroscopique est donc un vecteur (f_1, \dots, f_k) de fréquences, avec bien sûr $f_1 + \dots + f_k = 1$. On suppose que la mesure est absolue (pas d'erreur) et que $Nf_j = N_j$ est entier pour tout j . Le nombre d'états microscopiques associés à cet état macroscopique vaut alors

$$W = \frac{N!}{N_1! \dots N_k!}.$$

(Si l'on prépare N_j emplacements dans l'état j et que l'on numérote tous les emplacements de 1 à N , alors on a $N!$ façons de ranger les N boules dans les N emplacements, et ensuite on est incapable de distinguer entre permutations à l'intérieur d'une même boîte.)

D'après la formule de Stirling, quand $N \rightarrow \infty$ on a $\log N! = N \log N - N + \log \sqrt{2\pi N} + o(1)$. Il s'ensuit facilement que

$$\begin{aligned} \frac{1}{N} \log W &= - \sum_i \frac{N_i}{N} \log \frac{N_i}{N} + O\left(\frac{k \log N}{N}\right) \\ &= - \sum f_i \log f_i + o(1). \end{aligned}$$

On note que l'on peut aussi aboutir au même résultat sans utiliser la formule de Stirling [41].

Si maintenant on augmente le nombre d'expériences, on pourra formellement faire tendre $k \rightarrow \infty$, tout en tenant compte de ce que k reste très inférieur à N . Imaginons que l'on dispose d'une mesure de référence ν sur l'espace macroscopique \mathcal{Y} , et que l'on puisse l'on puisse séparer cet espace \mathcal{Y} en "cellules" de volume (mesure) $\delta > 0$, correspondant aux différents états. Quand $\delta \rightarrow 0$, si le système a une distribution statistique $f(y)$ par rapport à la mesure ν , on peut raisonnablement penser que $f_i \simeq \delta f(y_i)$ où y_i est un point représentatif de la cellule numérotée i . Mais alors

$$\sum_i f_i \log \frac{f_i}{\delta} \simeq \delta \sum_i f(y_i) \log f(y_i) \simeq \int f \log f d\nu,$$

où la dernière approximation provient d'un argument de somme de Riemann.

Nous avons ainsi abouti à la **fonction H de Boltzmann** : étant donnée une mesure de référence ν sur un espace \mathcal{Y} , et μ une mesure de probabilité sur \mathcal{Y} ,

$$H_\nu(\mu) = \int f \log f d\nu, \quad f = \frac{d\mu}{d\nu}. \quad (5)$$

Si ν est une mesure de probabilité, ou plus généralement une mesure de masse finie, il est facile d'étendre cette formule à toutes les probabilités μ , en posant $H_\nu(\mu) = +\infty$ si μ n'est pas absolument continue par rapport à ν . Si ν est une mesure de masse infinie, plus de précautions sont de mise, on pourra requérir au minimum la finitude de $\int f (\log f)_- d\nu$.

On note ensuite que si l'espace macroscopique \mathcal{Y} porte une mesure ν , alors l'espace microscopique $\mathcal{Z} = \mathcal{Y}^N$ porte une mesure naturelle $\nu^{\otimes N}$.

Nous sommes maintenant mûrs pour énoncer la traduction mathématique précise de la formule de la fonction H : on se donne $\{\varphi_j\}_{j \in \mathbb{N}}$ une famille dense de fonctions bornées et uniformément continues, alors

$$\lim_{k \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \lim_{N \rightarrow \infty} \frac{1}{N} \log \nu^{\otimes N} \left[\left\{ (y_1, \dots, y_N) \in \mathcal{Y}^N; \quad \forall j \in \{1, \dots, k\}, \right. \right. \\ \left. \left. \left| \int \varphi_j d\mu - \frac{1}{N} \sum_i \varphi_j(y_i) \right| \leq \varepsilon \right\} \right] = -H_\nu(\mu). \quad (6)$$

On interprètera donc N comme le nombre de particules; les φ_j comme une suite d'observables dont on mesure la valeur moyenne; et ε comme la précision des mesures. Cette formule résume de manière concise l'information essentielle contenue dans la fonction H .

Si ν est une mesure de probabilité, l'énoncé (6) est connu sous le nom de **théorème de Sanov** [43] et c'est un résultat phare de la théorie des grandes déviations. Avant de donner l'interprétation de (6) dans cette théorie, je note qu'une fois que l'on sait traiter le cas où ν est une mesure de probabilité on en déduit facilement le cas où ν est une mesure de masse finie; en revanche je n'ai pas connaissance d'une discussion rigoureuse dans le cas où ν est de masse infinie, même si on s'attend à ce que le résultat reste vrai.

2.3 Grandes déviations

Soit ν une mesure de probabilité, et supposons que l'on tire des variables aléatoires y_j selon ν , de manière indépendante. La mesure empirique $\hat{\mu} = N^{-1} \sum \delta_{y_j}$ est alors une mesure aléatoire, convergeant presque sûrement vers ν quand $N \rightarrow \infty$ (c'est le théorème de Varadarajan, aussi appelé loi fondamentale de la statistique [49]). Il se peut bien sûr que les apparences soient trompeuses, et que l'on croie observer une mesure μ distincte de ν . Cette probabilité d'observation décroît exponentiellement avec le coefficient N , et elle est en gros proportionnelle à $\exp(-N H_\nu(\mu))$; autrement dit, l'entropie de Boltzmann dicte la rareté des conditions qui mènent à l'observation "inattendue" μ .

2.4 Information

La théorie de l'information est née en 1948 avec le remarquable traité de Shannon et Weaver [94] sur la "théorie de la communication". C'est maintenant un pilier de toute l'industrie de la transmission de l'information.

Dans la théorie de Shannon, quelque peu désincarnée pour pouvoir être reproduite et abordée froidement, on définit la quantité d'information apportée par le déchiffrement d'un signal aléatoire, en fonction de l'inverse de la probabilité de ce signal (ce qui est rare est précieux). L'usage du logarithme permet d'avoir la propriété d'additivité, et l'on obtient la formule de Shannon pour la quantité moyenne d'information gagnée au cours du déchiffrement : $\mathbb{E} \log(1/p(Y))$, où p est la loi de Y . Ceci redonne bien sûr la formule de Boltzmann !

2.5 Entropies à tous les étages

L'entropie n'est pas un concept intrinsèque, elle dépend de l'observateur et du degré de connaissance qu'il peut acquérir par des expériences et mesures. Il est commode de penser que l'entropie est un concept *anthropique*. Une conséquence est que la notion d'entropie va varier avec le degré de précision de la description.

L'entropie de Boltzmann, comme on l'a vu, nous informe sur la rareté de la fonction de distribution cinétique $f(x, v)$, et la quantité d'information microscopique qui reste à découvrir une fois que l'on connaît f .

Si au contraire on se donne l'état microscopique de toutes les particules microscopiques, il n'y a plus d'information cachée et donc plus d'entropie. Mais si l'on se donne une probabilité sur les configurations microscopiques, μ^N , alors le concept d'entropie prend à nouveau son sens : l'entropie sera d'autant plus basse que la probabilité μ^N sera concentrée et informative en elle-même. On se retrouve alors avec une notion d'**entropie microscopique** $S_N = -H_N$,

$$H_N = \frac{1}{N} \int f^N \log f^N dx^N dv^N,$$

qui est typiquement *conservée* par la dynamique newtonienne, comme conséquence du théorème de Liouville. On peut vérifier que

$$H_N \geq H(\mu^{1;N}),$$

avec égalité quand μ^N est un produit tensoriel et qu'il n'y a donc pas de corrélations entre particules. L'idée est que l'état des particules microscopiques est plus facile

à obtenir par mesures multiparticules, que particule par particule – sauf bien sûr si les particules sont indépendantes !

Dans l'autre sens, on peut aussi se donner une distribution *moins précise* que la distribution cinétique : il s'agit typiquement d'une description hydrodynamique, qui fait seulement intervenir des champs de densité $\rho(x)$, température $T(x)$ et vitesse moyenne $u(x)$. Le passage du formalisme cinétique au formalisme hydrodynamique se fait par les formules simples

$$\begin{aligned}\rho(x) &= \int f(x, v) dv; & u(x) &= \rho(x)^{-1} \int f(x, v) v dv; \\ T(x) &= \frac{1}{d \rho(x)} \int f(x, v) \frac{|v - u(x)|^2}{2} dv.\end{aligned}$$

À cette description est associée une notion d'**entropie hydrodynamique** :

$$S_h = - \int \rho \log \frac{\rho}{T^{d/2}}.$$

Cette information est toujours plus basse que l'information cinétique. Nous avons finalement une hiérarchie : d'abord l'information microscopique de bas niveau, puis l'information "mésoscopique" de la fonction de distribution à la Boltzmann, enfin l'information "macroscopique" apportée par la description hydrodynamique. Les proportions relatives de ces différentes entropies constituent d'excellents moyens d'apprécier l'état physique des systèmes considérés.

3 Ordre et chaos

Intuitivement, un système microscopique est ordonné si toutes les particules sont agencées de manière coordonnée, *corrélée*. Au contraire, il est chaotique si les particules, n'en faisant qu'à leur tête, agissent toutes indépendamment les unes des autres. Reformulons cette idée : une distribution de particules est chaotique si les particules s'ignorent toutes les unes les autres, au sens où un gain d'information obtenu sur une particule donnée n'apporte aucun gain d'information sur une autre particule. Cette notion simple, clé de l'équation de Boltzmann, présente des subtilités importantes que nous allons brièvement évoquer.

3.1 Chaos microscopique

Il est équivalent de dire de particules aléatoires qu'elles s'ignorent, ou que leur loi jointe est un produit tensoriel. Bien sûr, même si des particules s'ignorent au temps initial, elles vont tout de suite entrer en interaction, et la propriété d'indépendance sera détruite. Dans le cas du système des sphères dures, la situation est encore pire : les particules sont obligées de tenir compte les unes des autres puisque les sphères ne peuvent s'interpénétrer. L'indépendance est donc à comprendre en un sens asymptotique quand le nombre de particules devient très grand; et les expériences visant à mesurer le degré d'indépendance ne feront intervenir qu'un nombre fini de particules. Ceci mène naturellement à la définition qui suit.

Soit un espace macroscopique \mathcal{Y} et pour tout N une mesure de probabilité μ^N sur \mathcal{Y}^N , supposée symétrique (invariante par permutation des coordonnées). On dit

que (μ^N) est **chaotique** s'il existe une probabilité μ telle que $\mu^N \simeq \mu^{\otimes N}$ au sens de la topologie faible des mesures produits. Explicitement, cela signifie que pour tout $k \in \mathbb{N}$ et pour toutes fonctions $\varphi_1, \dots, \varphi_k$ continues bornées sur \mathcal{Y} ,

$$\int_{\mathcal{Y}^N} \varphi_1(y_1) \dots \varphi_k(y_k) \mu^N(dy_1 \dots dy_N) \xrightarrow{N \rightarrow \infty} \left(\int \varphi_1 d\mu \right) \dots \left(\int \varphi_k d\mu \right). \quad (7)$$

Bien sûr, la définition peut être quantifiée, en introduisant une notion de distance adéquate permettant de mesurer l'écart entre μ^N et $\mu^{\otimes N}$. On pourra ainsi dire qu'une distribution μ^N est plus ou moins chaotique. On répète encore une fois : ce qui compte c'est l'indépendance d'un petit nombre k de particules prises parmi un très grand nombre N .

On montre (voir l'argument dans [99]) qu'il est équivalent d'imposer la propriété (7) pour tout $k \in \mathbb{N}$, ou simplement pour $k = 2$. Ainsi, le chaos signifie exactement que *2 particules tirées au hasard parmi N sont asymptotiquement indépendantes quand $N \rightarrow \infty$* . La démonstration passe par l'observation des liens entre chaos et mesure empirique.

3.2 Chaos et mesure empirique

Par la loi des grands nombres, le chaos impliquera automatiquement un déterminisme asymptotique : avec très forte probabilité, la mesure empirique s'approchera de la distribution statistique d'une particule arbitraire quand le nombre total de particules deviendra gigantesque.

Il se trouve que réciproquement *les corrélations s'accommodent très mal d'une prescription de densité macroscopique*. Avant de donner un énoncé précis, on va illustrer ce concept dans un cadre simplissime. Soient une boîte à deux compartiments, dans laquelle on range un très grand nombre N de boules *indistinguables*. Un état très corrélé serait un état dans lequel toutes les particules occuperaient le même compartiment : si je tire deux boules au hasard, l'état de la première boule m'informe complètement sur l'état de la seconde. Mais bien sûr, à partir du moment où les quantités respectives de boules dans les compartiments sont fixées et toutes deux non nulles, un tel état de corrélation est impossible. En fait, si les particules sont indistinguables, quand on en tire deux au hasard, la seule information que l'on gagne est obtenue en exploitant le fait qu'elles sont distinctes, de sorte que la connaissance de l'état de la première particule réduit légèrement le nombre de possibilités pour l'état de la seconde. Ainsi, si la première particule occupe l'état 1, alors les chances de trouver la seconde particule dans l'état 1 ou 2 respectivement ne sont pas $f_1 = N_1/N$ et $f_2 = N_2/N$, mais $f'_1 = (N_1 - 1)/N$ et $f_2 = N_2/N$. La distribution jointe d'un couple de particules est donc très proche de la loi produit.

En développant l'argument précédent, on arrive à un résultat général élémentaire mais conceptuellement profond, dont on trouvera la démonstration dans le cours de Sznitman [99] (voir aussi [40, p. 91]) : *le chaos microscopique équivaut au déterminisme de la mesure empirique*. Plus précisément, les énoncés suivants sont équivalents :

- (i) (μ^N) est μ -chaotique;
- (ii) la mesure empirique $\hat{\mu}^N$ associée à μ^N converge en loi vers la mesure déterministe μ .

Par “mesure empirique $\widehat{\mu}^N$ associée à μ^N ” on entend la mesure image de μ^N par l’application $(y_1, \dots, y_N) \mapsto N^{-1} \sum \delta_{y_i}$; c’est une mesure de probabilité aléatoire. La convergence en loi signifie que pour toute fonction continue bornée Φ sur l’espace des mesures de probabilité, on a

$$\int \Phi \left(\frac{1}{N} \sum \delta_{y_i} \right) \mu^N(dy_1 \dots dy_N) \xrightarrow{N \rightarrow \infty} \Phi(\mu).$$

En langage informel, étant donnée une grandeur statistique faisant intervenir $\widehat{\mu}^N$, on peut en donner une excellente approximation pour N grand en remplaçant, dans l’expression de cette statistique, $\widehat{\mu}^N$ par μ .

La notion de chaos ainsi présentée est une notion faible, susceptible de nombreuses variantes; l’idée générale étant que μ^N devrait être proche de $\mu^{\otimes N}$. La notion plus forte de **chaos entropique** est introduite par Ben Arous et Zeitouni [13] : on y impose $H_{\mu^{\otimes N}}(\mu^N) = o(N)$. Une notion proche est développée par Carlen, Carvalho, Leroux, Loss et Villani [32] dans le cas où l’espace microscopique n’est pas un produit tensoriel, mais plutôt une sphère de grande dimension; on remplace alors la mesure $\mu^{\otimes N}$ par la mesure produit *restreinte* à la sphère. De nombreuses autres variantes sont encore à découvrir.

3.3 Le règne du chaos

Dans la théorie de Boltzmann, on postule que *le chaos est la règle* : quand on prépare un système, il est a priori dans un état chaotique. Voici quelques arguments possibles :

- si nous pouvons agir sur la configuration macroscopique, nous n’avons pas accès à la structure microscopique et il est très difficile d’imposer des corrélations;
- les lois sous-jacentes aux lois microscopiques nous sont inconnues et on peut supposer qu’elles font intervenir un très grand nombre de facteurs détruisant les corrélations;
- la mesure macroscopique observée en pratique semble toujours bien déterminée et pas aléatoire;
- si l’on fixe la distribution macroscopique, l’entropie d’une distribution microscopique chaotique est plus grande que l’entropie d’une distribution microscopique non chaotique.

Explicitons le dernier argument. Si l’on se donne une probabilité μ sur \mathcal{Y} , alors la probabilité produit $\mu^{\otimes N}$ est d’entropie maximale parmi toutes les probabilités symétriques μ^N sur \mathcal{Y}^N ayant μ pour marginale. Au vu des grands nombres N mis en jeu, cela représente une quantité de possibilités phénoménalement plus grande.

La mesure microscopique μ_0^N peut être considérée comme un objet de nature bayésienne, une probabilité a priori sur l’espace des observations possibles. Ce choix, en général arbitraire, est ici fait d’une manière canonique, par maximisation de l’entropie : en quelque sorte, on choisit la distribution qui laisse le plus de possibilités ouvertes, et rend l’observation la plus probable. On rejoint ainsi la démarche scientifique du maximum de vraisemblance, qui a prouvé sa robustesse et son efficacité — et l’on dépasse la traditionnelle querelle des fréquentistes contre les bayésiens !

Le problème de la **propagation du chaos** consiste à montrer que cette hypothèse de chaos, faite sur la donnée initiale (préparée on ne sait trop comment), est propagée

par la dynamique microscopique (qui elle est bien définie). La propagation du chaos est vitale à deux titres : d'abord elle montre que l'indépendance est asymptotiquement préservée, fournissant une information statistique sur la dynamique microscopique; ensuite, elle garantit que *la mesure empirique reste déterministe*, ce qui laisse espérer la possibilité d'une **équation macroscopique** gouvernant l'évolution de cette mesure empirique ou de son approximation.

3.4 Évolution de l'entropie

Un thème récurrent dans l'étude qualitative des systèmes dynamiques, au moins depuis Poincaré, est la recherche de mesures invariantes; l'exemple le plus connu est la mesure de Liouville pour les systèmes hamiltoniens. Cette mesure possède en outre la propriété remarquable de se tensoriser.

Supposons que l'on dispose d'une dynamique microscopique sur \mathcal{Y}^N , et d'une mesure ν sur l'espace \mathcal{Y} telle que $\nu^{\otimes N}$ est une mesure invariante pour la dynamique microscopique; ou plus généralement telle qu'il existe une mesure invariante ν -chaotique sur \mathcal{Y}^N . Que devient dans la limite $N \rightarrow \infty$ la préservation du volume microscopique ?

Une conséquence simple de cette préservation du volume est la *conservation de l'information microscopique* $H_{\nu^{\otimes N}}(\mu_t^N)$, où μ_t^N est la mesure image de μ_0^N par l'évolution microscopique. En effet, comme μ_t^N est préservée par le flot (par définition) et $\nu^{\otimes N}$ également, la densité $f^N(t, y_1, \dots, y_N)$ est constante le long des trajectoires du système, et il s'ensuit que $\int f^N \log f^N d\nu^{\otimes N}$ est également constante.

Les choses sont plus subtiles pour l'information macroscopique. Bien sûr, si les différentes particules évoluent indépendamment les unes des autres, la mesure μ_t^N reste factorisée pour tous les temps, et on en déduit facilement que l'entropie macroscopique reste constante. En général les particules interagissent les unes avec les autres, ce qui détruit l'indépendance; cependant s'il y a propagation du chaos en un sens suffisamment fort, l'indépendance est restaurée pour $N \rightarrow \infty$, et en conséquence on a toujours déterminisme de la mesure empirique. Toutes les configurations typiques pour la mesure microscopique initiale μ_0^N donnent lieu, après un temps t , à une mesure empirique $\widehat{\mu}_t^N \simeq \mu_t$, où μ_t est bien déterminée. *Mais* il est possible que *d'autres* configurations microscopiques soient compatibles avec l'état μ_t , configurations qui n'ont pas été obtenues par évolution des configurations typiques initiales.

En d'autres termes : si l'on a une propagation du déterminisme macroscopique entre le temps initial et le temps t , et que la dynamique microscopique préserve la mesure de référence produit, alors on s'attend à ce que le volume d'états microscopiques admissibles ne diminue pas entre le temps 0 et le temps t . Gardant en tête la définition de l'entropie, on aurait $e^{NS(t)} \geq e^{NS(0)}$, où $S(t)$ est la valeur de l'entropie au temps t . On s'attend donc à ce que l'entropie *ne diminue pas* au cours de l'évolution temporelle :

$$S(t) \geq S(0).$$

Mais alors, pourquoi ne pas renverser l'argument et dire que le chaos au temps t implique le chaos au temps 0, par réversibilité de la dynamique microscopique ? Cet argument est en général irrecevable tant que l'on ne précise pas la notion exacte de chaos qui est propagée. La donnée initiale, préparée "au hasard" avec juste une contrainte de distribution cinétique, est supposée chaotique en un sens très fort;

mais peut-être que le système au temps t est chaotique en un sens moins fort; ceci dépend de l'évolution microscopique.

La notion d'échelle d'interaction joue ici un rôle important. Certaines interactions s'effectuent à échelle macroscopique, d'autres à échelle microscopique, c'est-à-dire que tout ou partie de la loi d'interaction est codé dans des paramètres qui sont invisibles au niveau macroscopique. Dans ce dernier cas, la notion de chaos propice à la propagation de la dynamique risque de ne pas être visible à l'échelle macroscopique, et on peut s'attendre à une dégradation de la notion de chaos.

À partir de là, la discussion doit faire intervenir les détails de la dynamique, et nos pires ennuis commencent.

4 Équations cinétiques

Après l'introduction de l'entropie et du chaos, nous pouvons revenir aux systèmes newtoniens de la section 1, dont l'espace des phases est fait de positions et vitesses. Une **équation cinétique** est une équation d'évolution portant sur la fonction de distribution $f(t, x, v)$; le rôle important de la variable de vitesse v justifie la terminologie *cinétique*. Par extension, dans le cas où il y a des degrés de liberté externes (orientation de molécules par exemple), on parle encore d'équations cinétiques par extension.

Dans la lignée de Boltzmann, on se pose le problème de déduire une évolution macroscopique à partir du modèle microscopique sous-jacent. Ce problème est en général d'une difficulté considérable. Les équations fondamentales sont celles de Vlasov, Boltzmann, Landau et Balescu-Lenard, publiées respectivement en 1938, 1867, 1936 et 1960. (L'ordre plus ou moins logique de présentation de ces équations ne suit pas l'ordre dans lequel elles ont été découvertes...)

4.1 Équation de Vlasov

Aussi appelée équation de Boltzmann sans collisions, l'équation de Vlasov [111] est une équation de champ moyen, au sens où toutes les particules interagissent les unes avec les autres. Pour la déduire de la dynamique newtonienne, commençons par traduire l'équation de Newton sous forme d'équation sur la mesure empirique; pour cela on écrit l'équation d'évolution d'un observable arbitraire :

$$\begin{aligned} \frac{d}{dt} \frac{1}{N} \sum_i \varphi(X_i(t), \dot{X}_i(t)) \\ &= \frac{1}{N} \sum_i \left[\nabla_x \varphi(X_i, \dot{X}_i) \cdot \dot{X}_i + \nabla_v \varphi(X_i, \dot{X}_i) \cdot \ddot{X}_i \right] \\ &= \frac{1}{N} \sum_i \left[\nabla_x \varphi(X_i, \dot{X}_i) \cdot \dot{X}_i + \nabla_v \varphi(X_i, \dot{X}_i) \cdot \left(a \sum_j F(X_i - X_j) \right) \right]. \end{aligned}$$

Ceci peut se réécrire

$$\frac{\partial \widehat{\mu}^N}{\partial t} + v \cdot \nabla_x \widehat{\mu}_t^N + aN(F * \widehat{\mu}_t^N) \cdot \nabla_v \widehat{\mu}_t^N = 0. \quad (8)$$

Si maintenant on suppose que $aN \simeq 1$, et que l'on réalise l'approximation

$$\widehat{\mu}_t^N(dx dv) \simeq f(t, x, v) dx dv,$$

on obtient l'équation de Vlasov

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f + \left(F *_x \int f dv \right) \cdot \nabla_v f = 0. \quad (9)$$

Notons bien que $\widehat{\mu}_t^N$ dans (8) est une solution faible de l'équation de Vlasov, de sorte que le passage à la limite est conceptuellement très simple : il s'agit simplement d'un résultat de stabilité de l'équation de Vlasov.

Bien évidemment, je suis allé un peu vite en besogne car cette équation est non linéaire. Si $\widehat{\mu} \simeq f$ au sens de la topologie faible des mesures, alors $F * \widehat{\mu}$ converge vers $F * \int f dv$ dans une topologie déterminée par la régularité de F , et si cette topologie est plus faible que la convergence uniforme, rien ne garantit que $(F * \widehat{\mu})\widehat{\mu} \simeq (F * f)f$.

Si F est effectivement bornée et uniformément continue, alors l'argument ci-dessus peut être rendu rigoureux. Si F est en outre L -lipschitzienne, alors on peut faire mieux et établir une estimation de stabilité en topologie faible : si (μ_t) et (μ'_t) sont deux solutions faibles de l'équation de Vlasov, alors

$$W_1(\mu_t, \mu'_t) \leq e^{2 \max(1, L)|t|} W_1(\mu_0, \mu'_0),$$

où W_1 est la distance de Wasserstein d'ordre 1,

$$W_1(\mu, \nu) = \sup : \left\{ \int \varphi d\mu - \int \varphi d\nu; \quad \varphi \text{ 1-lipschitzienne} \right\}.$$

Des estimations de ce style se trouvent dans [95, Chapitre 5], et remontent aux années 1970 (Dobrushin [48], Braun et Hepp [24], Neunzert [87]). On peut aussi établir des estimations de grande déviation comme dans [20].

En revanche, pour des **interactions singulières**, le problème de la limite de Vlasov reste ouvert, à l'exception d'un résultat de Jabin et Hauray [64], qui suppose essentiellement que (a) $F(x - y) = O(|x - y|^{-s})$ avec $0 < s < 1$; et (b) les particules sont initialement bien séparées dans l'espace des phases, de sorte que

$$\inf_{j \neq i} \left(|X_i(0) - X_j(0)| + |\dot{X}_i(0) - \dot{X}_j(0)| \right) \geq \frac{c}{N^{\frac{1}{2d}}}.$$

Aucune de ces conditions n'est satisfaisante : la première manque le cas coulombien d'un ordre de singularité, tandis que la seconde exclut le cas de données aléatoires. Cependant il s'agit bien du seul résultat disponible à ce jour... Pour aller plus loin, on souhaiterait propager une notion de chaos suffisamment forte pour contrôler le nombre de paires (i, j) telles que $|X_i(t) - X_j(t)|$ est petit. En l'absence de tels contrôles, l'équation de Vlasov pour des interactions singulières reste un acte de foi.

Cet acte de foi est très efficace puisque c'est le modèle de Vlasov–Poisson, dans lequel $F = -\nabla W$, W solution fondamentale de $\pm\Delta$, est le modèle classique universellement accepté aussi bien en physique des plasmas [42, 71] qu'en astrophysique [15]. Dans le premier cas, une particule est un électron, dans le second c'est une étoile ! La seule différence réside dans le signe : interaction répulsive pour les électrons, attractive pour les étoiles. Il ne faut pas s'étonner de voir les étoiles

considérées ainsi comme des objets microscopiques : elles le sont effectivement à l'échelle d'une galaxie (qui peut compter 10^{12} étoiles...).

La théorie de l'équation de Vlasov–Poisson elle-même reste incomplète. On distingue actuellement deux théories principales, toutes deux développées dans l'espace entier. Celle de Pfaffelmoser, exposée par exemple dans [51], suppose que la donnée initiale f_i est C^1 à support compact; celle de Lions–Perthame, passée en revue dans [23], suppose seulement que $|f_i| + |\nabla f_i|$ décroît suffisamment vite à l'infini. La théorie de Pfaffelmoser a été adaptée dans un cadre périodique en espace [12], ce qui n'est pas le cas de la théorie de Lions–Perthame; cependant, l'hypothèse de support compact en vitesse dans la théorie de Pfaffelmoser est profondément insatisfaisante, et du travail reste encore à faire dans ce domaine.

4.2 Équation de Boltzmann

L'équation de Vlasov perd sa pertinence quand les interactions ont une courte portée. Un exemple typique est celui des gaz raréfiés, pour lesquels les interactions dominantes sont binaires et se produisent uniquement au cours de “collisions” entre particules.

L'équation de Boltzmann a été établie par Maxwell [80] et Boltzmann [21, 22]; elle décrit un régime où les interactions sont à courte portée et où chaque particule subit $O(1)$ choc par unité de temps. Beaucoup plus subtil que le régime de champ moyen de Vlasov, le régime de Boltzmann est cependant plus simple que le régime hydrodynamique où les particules subissent un très grand nombre de collisions par unité de temps.

Établissons pour commencer l'équation formellement. Le mouvement d'une particule est fait d'une alternance de trajectoires rectilignes et de chocs au cours desquels sa vitesse variera si brusquement que l'on peut considérer cet événement comme instantané et localisé en espace. Examinons d'abord le cas emblématique des sphères dures de rayon r : un choc se produit quand deux particules de positions respectives x et y , de vitesses respectives v et w , se retrouvent dans la configuration où $|x-y| = 2r$ et $(w-v) \cdot (y-x) < 0$. On parle alors de *configuration précollisionnelle*. On notera $\omega = (y-x)/|x-y|$.

Ici vient maintenant le point central de toute l'argumentation de Boltzmann : *quand deux particules se rencontrent, avec très forte probabilité elles ne sont (presque) pas corrélées* : penser à deux individus qui se rencontrent pour la première fois. On peut en conséquence appliquer l'hypothèse de chaos moléculaire à de telles particules, et on trouve que la probabilité d'une rencontre entre ces particules est proportionnelle

$$\begin{aligned} f^{2;N}(t, x, v, x + 2r\omega, w) &\simeq f^{1;N}(t, x, v) f^{1;N}(t, x + 2r\omega, w) \\ &\simeq f^{1;N}(t, x, v) f^{1;N}(t, x, w), \end{aligned}$$

à condition donc que $(w-v) \cdot \omega < 0$. Il faut également tenir compte des vitesses relatives pour évaluer l'influence des particules de vitesse w sur les particules de vitesse v : la probabilité de rencontrer une particule de vitesse w en une unité de temps est proportionnelle au produit de $|v-w|$ par la section efficace (en dimension 3 il s'agit de l'aire apparente des particules, soit πr^2) et par le cosinus de l'angle entre $v-w$ et ω (le cas extrême est celui où $v-w$ est orthogonal à ω , ce qui veut dire que les deux particules se frôlent, évidemment un événement de probabilité nulle).

Chacune de ces collisions fait disparaître une particule de vitesse v , et on a donc un terme négatif, *terme de perte*, proportionnel à

$$- \iint f(t, x, v) f(t, x, v_*) |(v - v_*) \cdot \omega| dv_* d\omega.$$

Les vitesses après le choc se calculent facilement :

$$v' = v - (v - v_*) \cdot \omega \omega; \quad v'_* = v_* + (v - v_*) \cdot \omega \omega. \quad (10)$$

Ces vitesses nous importent peu dans le bilan.

Cependant, il faut aussi tenir compte de toutes les particules de vitesse v qui ont été créées par collision entre particules de vitesses arbitraires. Par réversibilité microscopique, ces vitesses sont de la forme (v', v'_*) , et notre problème est de tenir compte de tous les couples possibles (v', v'_*) , qui dans ce problème de calcul du *terme de gain* sont des vitesses *pré-collisionnelles*. On applique donc à nouveau l'hypothèse de chaos pré-collisionnel, et on obtient finalement l'expression de l'**équation de Boltzmann pour les sphères dures** :

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f = Q(f, f), \quad (11)$$

$$\begin{aligned} Q(f, f)(t, x, v) \\ = B \int_{\mathbb{R}^3} \int_{S^2_-} |(v - v_*) \cdot \omega| \left(f(t, x, v') f(t, x, v'_*) - f(t, x, v) f(t, x, v_*) \right) dv_* d\omega, \end{aligned}$$

où S^2_- désigne les configurations pré-collisionnelles $\omega \cdot (v - v_*) < 0$, et B est une constante. En utilisant le changement de variable $\omega \rightarrow -\omega$ on peut symétriser cette expression, et arriver à l'expression finale

$$\begin{aligned} Q(f, f)(t, x, v) \\ = B \int_{\mathbb{R}^3} \int_{S^2} |(v - v_*) \cdot \omega| \left(f(t, x, v') f(t, x, v'_*) - f(t, x, v) f(t, x, v_*) \right) dv_* d\omega. \quad (12) \end{aligned}$$

L'opérateur (12) est l'**opérateur de collision de Boltzmann pour les sphères dures**. Le problème consiste maintenant à justifier cette approximation.

Pour ce faire, dans les années 1960, Grad proposa une limite mathématique précise : faire tendre r vers 0 et en même temps N vers l'infini, de sorte que $Nr^2 \rightarrow 1$, c'est-à-dire que la section efficace totale reste constante. Ainsi une particule donnée, se déplaçant parmi toutes les autres, en rencontrera typiquement un nombre fini par unité de temps. On part ensuite d'une densité de probabilité microscopique $f_0^N(x^N, v^N) dx^N dv^N$, que l'on fait évoluer par le flot newtonien \mathcal{N}_t , et l'on cherche à montrer que la première marginale $f^{1;N}(t, x, v)$ (obtenue en intégrant toutes les variables sauf la première variable de position et la première variable de vitesse) converge à la limite vers une solution de l'équation de Boltzmann.

La limite de Boltzmann-Grad est aussi souvent appelée **limite de faible densité** [40, p. 60] : en effet, si l'on part de la dynamique newtonienne et que l'on fixe la taille des particules, on va dilater l'échelle spatiale d'un facteur $1/\sqrt{N}$, et la densité sera d'ordre $N/N^{3/2} = N^{-1/2}$.

Au début des années 1970, Cercignani [37] montrait que le programme de Grad pouvait être mené à bien si l'on prouvait un certain nombre d'estimations *plausibles*; peu de temps après, indépendamment, Lanford [69] esquissait la preuve du résultat souhaité.

Le **théorème de Lanford** est peut-être le plus important résultat mathématique de la théorie cinétique. Dans ce théorème, on se donne des densités microscopiques f_0^N , telles que pour tout k les densités $f_0^{k;N}$ des marginales à k particules sont continues, satisfont des bornes gaussiennes aux grandes vitesses, et convergent uniformément en-dehors des configurations collisionnelles (celles où les positions de deux particules distinctes coïncident) vers leur limite $f_0^{\otimes k}$. La conclusion est qu'il existe un temps $t_* > 0$ tel que $f_t^{k;N}$ converge *presque partout* vers $f_t^{\otimes k}$, où f_t est solution de l'équation de Boltzmann, pour tout temps $t \in [0, t_*]$.

Les estimations de Lanford ont été plus tard réécrites par Spohn [95] et par Illner et Pulvirenti [61, 62] qui remplacèrent l'hypothèse de temps petit par une hypothèse de petitesse sur la donnée initiale, permettant de traiter l'équation de Boltzmann comme une perturbation du transport libre. Ces résultats sont passés en revue dans [40, 90, 95].

La technique employée par Lanford et ses successeurs passe par la **hiérarchie BBGKY** (Bogoljubov–Born–Green–Kirkwood–Yvon), méthode par laquelle on exprime l'évolution de la marginale à une particule $f^{1;N}$ en fonction de la marginale à deux particules $f^{2;N}$; l'évolution de la marginale à deux particules $f^{2;N}$ en fonction de la marginale à trois particules $f^{3;N}$, et ainsi de suite. Ce procédé est particulièrement peu économique (dans l'argument heuristique précédent, on a seulement utilisé $f^{1;N}$ et $f^{2;N}$, mais on ne lui connaît pas d'alternative).

On résout ensuite chacune des équations de la hiérarchie via la formule de Duhamel, en appliquant successivement les opérateurs de transport libre et de collision, et en sommant sur tous les histoires collisionnelles possibles. On exprime ainsi formellement, comme avec un opérateur exponentiel, la solution au temps t en fonction de la donnée initiale, et on peut appliquer l'hypothèse de chaos sur (f_0^N) .

Puis on passe à la limite $N \rightarrow \infty$ dans chacune des équations, après avoir vérifié que l'on peut négliger des situations "pathologiques" de "recollision" où une particule rencontre à nouveau une particule qu'elle a déjà rencontrée une première fois, et qui ne lui est donc pas inconnue. Ce point est subtil : on discute dans [40, Appendice 4.C] une dynamique a priori plus simple que celle des sphères dures, due à Uchiyama, avec seulement quatre vitesses dans le plan, pour laquelle les configurations recollisionnelles ne peuvent être négligées, et la limite cinétique n'existe pas.

Il reste à identifier le résultat avec la suite des produits tensoriels de la solution de l'équation de Boltzmann, et à conclure en utilisant un résultat d'unicité.

Spohn [95, Section 4.6] montre que l'on peut donner des informations plus précises sur la répartition microscopique des particules : à petite échelle, celle-ci suit une **loi de Poisson** homogène dans l'espace des phases. Ceci est cohérent avec l'idée intuitive du chaos moléculaire.

Le théorème de Lanford tranche une controverse qui durait depuis Boltzmann lui-même; mais il laisse de nombreuses questions en suspens. En premier lieu, il est limité à un intervalle de temps petit (sur lequel seulement 1/6 environ des particules ont eu le temps d'entrer en collision... mais l'impact conceptuel du théorème n'en est pas moins important). La variante d'Illner et Pulvirenti lève cette restriction de

temps petit, mais la preuve ne s'adapte pas à une géométrie bornée. Quant à lever la restriction de petitesse, c'est pour l'instant un rêve très lointain.

Ensuite, le théorème n'est à ce jour démontré que pour un système de sphères dures; les interactions à longue portée ne sont pas couvertes. Cercignani [36] note que la limite de Boltzmann–Grad pour de telles interactions pose des problèmes subtils, même d'un point de vue formel.

Enfin, le plus frustrant est peut-être que Lanford évite la discussion du **chaos pré-collisionnel**, cette notion selon laquelle des particules qui sont sur le point de se rencontrer ne sont pas corrélées. Cette notion est très subtile ! Car juste après le choc, des corrélations auront fatalement eu lieu. En d'autres termes, nous avons *chaos pré-collisionnel, mais pas post-collisionnel*.

Que signifie exactement le chaos pré-collisionnel ? Nous n'en avons pas pour l'instant de définition précise. C'est certainement une notion plus forte que le chaos au sens habituel; elle fait intervenir en plus une hypothèse de décorrélation qui se voit sur un ensemble de codimension 1, à savoir les configurations menant à des chocs. On aimerait en déduire que c'est une notion de chaos où l'on a remplacé la topologie faible par une topologie uniforme; mais cela ne peut être aussi simple, car le chaos en topologie uniforme impliquerait aussi le chaos post-collisionnel, et ce dernier est incompatible avec le chaos pré-collisionnel ! En effet, la continuité de la marginale à deux particules le long d'un choc impliquerait

$$\begin{aligned} f(t, x, v) f(t, x, v_*) &\simeq f^{(2;N)}(t, x, v, x + 2r\omega, v_*) \\ &= f^{(2;N)}(t, x, v', x + 2r\omega, v'_*) \simeq f^{(1;N)}(t, x, v') f^{(1;N)}(t, x, v'_*). \end{aligned}$$

Passant à la limite on en déduirait

$$f(t, x, v') f(t, x, v'_*) = f(t, x, v) f(t, x, v_*),$$

et comme on le verra dans la section 5.3 ceci implique que f est gaussienne dans la variable de vitesse, en général faux bien sûr ! Un autre raisonnement pour montrer que le chaos post-collisionnel doit être incompatible avec le chaos pré-collisionnel consiste à remarquer que si l'on a chaos post-collisionnel, le raisonnement menant à l'équation de Boltzmann peut être effectué à nouveau en exprimant les probabilités à deux particules en fonction des probabilités post-collisionnelles... et on obtient alors l'équation de Boltzmann renversée :

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f = -Q(f, f).$$

Comme on l'a mentionné, la preuve de Lanford ne s'applique qu'aux sphères dures; cependant l'équation de Boltzmann est utilisée pour une gamme bien plus large d'interactions. La forme générale de l'équation, disons en dimension d , est la même qu'en (11) :

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f = Q(f, f), \quad (13)$$

mais maintenant

$$Q(f, f) = \int_{\mathbb{R}^d} \int_{S^{d-1}} (f' f'_* - f f_*) \tilde{B}(v - v_*, \omega) dv_* d\omega \quad (14)$$

où $\tilde{B}(v - v_*, \omega)$ ne dépend que de $|v - v_*|$ et $|(v - v_*) \cdot \omega|$. Il existe plusieurs représentations de cet opérateur intégral (voir [103]); il est souvent commode de

changer de variables en introduisant un autre angle, $\sigma = (v' - v'_*)/|v - v_*|$, de sorte que les formules (10) doivent être remplacées par

$$v' = \frac{v + v_*}{2} + \frac{|v - v_*|}{2} \sigma, \quad v'_* = \frac{v + v_*}{2} - \frac{|v - v_*|}{2} \sigma. \quad (15)$$

On doit alors changer le noyau de collision \tilde{B} par B tel que

$$B d\sigma = \tilde{B} d\omega.$$

Explicitement, on trouve

$$\frac{1}{2} \tilde{B}(z, \omega) = \left| 2 \frac{z}{|z|} \cdot \omega \right|^{d-2} B(z, \sigma).$$

La forme précise de B (ou, de manière équivalente, de \tilde{B}) est obtenue par un calcul de scattering classique qui remonte à Maxwell, et que l'on peut trouver dans [38] : pour un paramètre d'impact $p \geq 0$ et une vitesse relative $z \in \mathbb{R}^3$, l'angle de déviation θ vaut

$$\theta(p, z) = \pi - 2p \int_{s_0}^{+\infty} \frac{ds/s^2}{\sqrt{1 - \frac{p^2}{s^2} - 4 \frac{\phi(s)}{|z|^2}}} = \pi - 2 \int_0^{\frac{p}{s_0}} \frac{du}{\sqrt{1 - u^2 - \frac{4}{|z|^2} \phi\left(\frac{p}{u}\right)}},$$

où s_0 est la racine positive de

$$1 - \frac{p^2}{s_0^2} - 4 \frac{\phi(s_0)}{|z|^2} = 0.$$

Alors B est défini implicitement par

$$B(|z|, \cos \theta) = \frac{p}{\sin \theta} \frac{dp}{d\theta} |z|. \quad (16)$$

Quand $\phi(r) = 1/r$, on retrouve la formule de Rutherford pour la déviation coulombienne :

$$B(|v - v_*|, \cos \theta) = \frac{1}{|v - v_*|^3 \sin^4(\theta/2)}.$$

Quand $\phi(r) = 1/r^{s-1}$, $s > 2$, le noyau de collision ne se calcule pas explicitement, mais on peut montrer que (toujours en dimension 3)

$$B(|v - v_*|, \cos \theta) = b(\cos \theta) |v - v_*|^\gamma, \quad \gamma = \frac{s-5}{s-1}. \quad (17)$$

En outre, la fonction b , définie implicitement, est localement lisse avec une *singularité angulaire non intégrable* quand $\theta \rightarrow 0$:

$$\sin \theta b(\cos \theta) \sim K \theta^{-1-\nu}, \quad \nu = \frac{2}{s-1}. \quad (18)$$

Cette singularité correspond aux collisions de grand paramètre d'impact p , qui ne sont presque pas déviées. Elle est inévitable dès lors que les forces sont de rayon infini : en effet

$$\int_0^\pi B(|z|, \cos \theta) \sin \theta d\theta = |z| \int_0^\pi p \frac{dp}{d\theta} d\theta = |z| \int_0^{p_{\max}} p dp = \frac{|z| p_{\max}^2}{2}. \quad (19)$$

Dans le cas particulier $s = 5$, le noyau de collision ne dépend plus de la vitesse relative, mais seulement de l'*angle de déviation* : on parle de molécules maxwelliennes. Par extension, on dit que $B(v-v_*, \sigma)$ est un noyau de collision maxwellien s'il ne dépend que de l'angle entre $v - v_*$ et σ . Les molécules maxwelliennes sont avant tout un modèle phénoménologique, même si l'interaction entre un ion chargé et une particule neutre dans un plasma est régie par une telle loi [42, Vol.1, p. 149]. Les potentiels en $1/r^{s-1}$ pour $s > 5$ sont appelés potentiels durs, pour $s < 5$ potentiels mous. On tronque souvent la singularité angulaire $b(\cos \theta)$ aux petites valeurs de θ .

L'équation de Boltzmann est importante dans la modélisation des gaz raréfiés, comme expliqué dans [39]. Cependant, du fait de son histoire mouvementée et de son contenu conceptuel, ainsi que de l'impact du traité de Boltzmann [22], cette équation a exercé une fascination qui va bien au-delà de son utilité. Les premiers travaux mathématiques qui lui étaient consacrés ont été ceux de Carleman [26] [27], suivis de Grad [57]. Outre l'article de Lanford [69] déjà cité, un résultat qui a eu un grand retentissement est le théorème de stabilité faible de DiPerna–Lions [47]. L'équation est bien comprise dans le régime spatialement homogène pour les potentiels durs avec troncature angulaire, voir par exemple [84]; et dans le régime proche de l'équilibre, voir par exemple [60]. On renvoie aux traités de référence [38, 40, 103] pour quantité d'autres résultats.

4.3 Équation de Landau

L'intégrale collisionnelle de Boltzmann perd son sens pour des interactions coulombiennes, du fait de la décroissance extrêmement lente du potentiel coulombien. Les collisions rasantes, de grand paramètre d'impact, deviennent alors dominantes.

En 1936, Landau [67] établissait par des arguments formels une asymptotique du noyau de Boltzmann dans ce régime. Soient λ_D la distance d'écrantage (en-dessous de laquelle le potentiel coulombien n'est plus visible du fait de la neutralité globale du plasma), et r_0 la distance typique des collisions (distance de deux particules dont l'énergie d'interaction est comparable à l'énergie d'agitation moléculaire); le paramètre $\Lambda = 2\lambda_D/r_0$ est le **paramètre plasma**, et dans la limite $\Lambda \rightarrow \infty$ (justifiée pour les plasmas dits classiques), on peut formellement remplacer l'opérateur de Boltzmann avec potentiel coulombien écranté par un opérateur diffusif appelé opérateur de Landau :

$$Q_B(f, f) \simeq \frac{\log \Lambda}{2\pi\Lambda} Q_L(f, f), \quad (20)$$

$$Q_L(f, f) = \nabla_v \cdot \left(\int_{\mathbb{R}^3} a(v - v_*) [f(v_*) \nabla_v f(v) - f(v) \nabla_v f(v_*)] dv_* \right); \quad (21)$$

$$a(v - v_*) = \frac{L}{|v - v_*|} \Pi_{(v-v_*)^\perp} \quad (22)$$

où L est une constante et Π_{z^\perp} est le projecteur orthogonal sur z^\perp .

L'approximation de Landau est maintenant bien comprise mathématiquement, dans le cadre d'une limite appelée **asymptotique des collisions rasantes**; on pourra se reporter à [3] pour une discussion détaillée de ce problème.

L'opérateur de Landau (21), à la fois diffusif et intégral, présente une structure remarquable. On le généralise facilement à des dimensions $d \geq 2$ arbitraires, et on peut changer le coefficient $L/|z|$ en $L|z|^{\gamma+2}$, où γ est la puissance apparaissant

dans (17). Les modèles de type potentiel durs avec $\gamma > 0$ ont été complètement étudiés dans le cas spatialement homogène [45]; mais c'est bien le cas $\gamma = -3$ en dimension 3 qui est physiquement intéressant. Dans ce cas on sait seulement démontrer l'existence de solutions faibles dans le cas spatialement homogène (en adaptant [1, Section 7]) et l'existence de solutions fortes pour des perturbations de l'équilibre [59]. Cette situation est tout à fait insatisfaisante.

4.4 Équation de Balescu-Lenard

En 1960, Balescu [7] établit directement une équation cinétique qui décrit les interactions coulombiennes dans un plasma; il retrouve ainsi une équation publiée sous une autre forme par Bogoljubov [19] et simplifiée par Lenard. On pourra consulter [96] pour plus d'informations sur la genèse de l'équation, et [8] pour une présentation synthétique. Le noyau de collision dans cette équation prend la même forme que (21), la différence est dans l'expression de la matrice $a(v - v_*)$, qui maintenant dépend aussi de v et ∇f :

$$a_{BL}(v, v - v_*, \nabla f) = B \int_{|k| \leq K_0} \frac{k \otimes k}{|k|^4} \frac{\delta_{k \cdot (v - v_*)}}{|\epsilon(k, k \cdot v, \nabla f)|^2} dk, \quad (23)$$

$$\epsilon(k, k \cdot v, \nabla f) = 1 + \frac{b}{|k|^2} \int_{\mathbb{R}^3} \frac{k \cdot \nabla f(u)}{k \cdot (v - u) - i0} du.$$

Cette équation peut aussi être obtenue à partir de l'étude des fluctuations de l'équation de Vlasov en temps grand [71, Section 51].

L'équation de Balescu-Lenard n'est presque pas employée du fait de sa complexité. Sous des hypothèses raisonnables, l'équation de Landau en constitue une bonne approximation [8, 70]. La procédure s'adapte à d'autres interactions que l'interaction coulombienne, mais en contraste avec la limite des collisions rasantes, elle fournit toujours l'expression (21), le seul changement étant dans le coefficient L de (22) qui est proportionnel à

$$\int_{\mathbb{R}^3} |k| |\widehat{W}(k)|^2 dk,$$

où W est le potentiel d'interaction. Cette équation est brièvement passée en revue dans [95, Chapitre 6].

La théorie mathématique de l'équation de Balescu-Lenard est grande ouverte, aussi bien pour ce qui est de l'établir que d'étudier ses propriétés qualitatives; un des rares travaux rigoureux sur le sujet est l'étude linéarisée de R. Strain [96]. Bien que peu utilisée, l'équation de Balescu-Lenard n'en est pas moins le plus respecté des modèles collisionnels dans les plasmas, et c'est un intermédiaire qui permet de justifier l'usage de l'opérateur de collision de Landau pour représenter les fluctuations en temps grand dans les systèmes de particules; sa théorie représente un formidable défi.

5 Théorème H de Boltzmann

Dans cette section nous allons partir de l'équation de Boltzmann et examiner quelques-unes de ses propriétés les plus marquantes. On trouvera beaucoup plus d'informations dans mon long article de revue [103].

5.1 Modification des observables par les collisions

Les propriétés statistiques du gaz se manifestent, dans le modèle cinétique, par l'évolution des **observables** $\iint f(t, x, v) \varphi(x, v) dx dv$. Supposant toujours des conditions aux limites périodiques, et toute la régularité nécessaire, on peut écrire

$$\begin{aligned} \frac{d}{dt} \iint f \varphi dx dv &= \iint (\partial_t f) \varphi dx dv \\ &= - \iint v \cdot \nabla_x f \varphi dx dv + \iint Q(f, f) \varphi dx dv \\ &= \iint v \cdot \nabla_x \varphi f dx dv + \iiint \tilde{B} (f' f'_* - f f_*) \varphi dx dv dv_* d\omega, \end{aligned} \quad (24)$$

où l'on note toujours $f' = f(t, x, v')$, etc.

On effectue maintenant dans le terme intégral en $f' f'_*$ le changement de variables pré-postcollisionnel $(v, v_*) \rightarrow (v', v'_*)$, pour tout $\omega \in S^{d-1}$. Ce changement de variable est unitaire (déterminant jacobien égal à 1), et préserve \tilde{B} (ces propriétés sont des traces de la microréversibilité). Après avoir renommé les variables, on obtient

$$\frac{d}{dt} \iint f \varphi dx dv = \iint v \cdot \nabla_x \varphi f dx dv + \iiint \tilde{B} f f_* (\varphi' - \varphi) dv dv_* d\omega dx. \quad (25)$$

C'est d'ailleurs sous cette forme que Maxwell avait écrit l'équation de Boltzmann dès 1867... On déduit de (25) que $\iint f dx dv$ est constante (heureusement !!), et on en tire aussi une quantité importante, la *section efficace de transfert de quantité de mouvement*, $M(v - v_*)$, définie par

$$M(v - v_*) (v - v_*) = \int \tilde{B}(v - v_*, \omega) (v' - v) d\omega.$$

Même quand \tilde{B} est d'intégrale divergente, la quantité M doit être finie, exprimant le fait que les collisions modifient les vitesses de manière statistiquement raisonnable. On pourra se reporter à [103] [2, 3] pour plus de détails sur le traitement des singularités rasantes de \tilde{B} .

Boltzmann va améliorer le procédé de Maxwell en tirant mieux partie des symétries de l'équation. D'abord, en effectuant le changement de variable pré-postcollisionnel dans tout le second membre de (24), on obtient

$$\iiint \tilde{B} (f' f'_* - f f_*) \varphi dv dv_* d\omega dx = - \iiint \tilde{B} (f' f'_* - f f_*) \varphi' dv dv_* d\omega dx. \quad (26)$$

Au lieu d'échanger les configurations précollisionnelles et postcollisionnelles, on peut maintenant échanger les particules entre elles : $(v, v_*) \mapsto (v_*, v)$, ce qui est aussi bien évidemment de déterminant jacobien unité. Ceci nous donne deux nouvelles formes de (26) :

$$\iiint \tilde{B} (f' f'_* - f f_*) \varphi_* dv dv_* d\omega dx = - \iiint \tilde{B} (f' f'_* - f f_*) \varphi'_* dv dv_* d\omega dx. \quad (27)$$

En combinant les quatre formes apparaissant dans (26) et (27), on obtient

$$\begin{aligned} \frac{d}{dt} \iint f \varphi \, dv \, dx &= \iint f (v \cdot \nabla_x \varphi) \, dx \, dv \\ &\quad - \frac{1}{4} \iiint \tilde{B} (f' f'_* - f f_*) (\varphi' + \varphi'_* - \varphi - \varphi_*) \, dx \, dv \, dv_* \, d\omega. \end{aligned} \quad (28)$$

Comme conséquence de (28), on note en premier lieu que $\iint f \varphi$ est conservé si φ vérifie l'équation fonctionnelle

$$\varphi(v') + \varphi(v'_*) = \varphi(v) + \varphi(v_*) \quad (29)$$

pour tous choix des vitesses v , v_* et du paramètre ω . De telles fonctions sont appelées **invariants de collision**, et se réduisent, sous des hypothèses extrêmement faibles, aux seules combinaisons linéaires des fonctions

$$1, \quad v_j \quad (1 \leq j \leq d), \quad \frac{|v|^2}{2}.$$

On pourra consulter [40] à ce sujet. Ceci est encore une fois naturel : il s'agit du reflet macroscopique de la conservation de masse, quantité de mouvement et énergie cinétique au cours des interactions microscopiques.

5.2 Théorème H

Nous arrivons maintenant à une découverte qui installera Boltzmann parmi les plus grands noms de la physique. Choisissons $\varphi = \log f$, et supposons toute la régularité souhaitée pour mener les calculs; en particulier,

$$\iint f v \cdot \nabla_x (\log f) \, dv \, dx = \iint v \cdot \nabla_x (f \log f - f) \, dv \, dx = 0.$$

L'identité (28) devient ainsi, en tenant compte des propriétés additives du logarithme,

$$\frac{d}{dt} \iint f \log f \, dx \, dv = -\frac{1}{4} \iiint \tilde{B} (f' f'_* - f f_*) (\log f' f'_* - \log f f_*). \quad (30)$$

La fonction logarithme étant croissante, l'expression ci-dessus est toujours négative ! En outre, sachant que B ne s'annule que sur un ensemble de mesure nulle, on voit que l'expression (30) est strictement négative dès que $f' f'_*$ n'est pas égal à $f f_*$ presque partout, ce qui est vrai pour des distributions génériques. Ainsi, modulo la justification rigoureuse des intégrations par parties et changement de variables, nous avons prouvé que *dans le modèle de Boltzmann, l'entropie augmente avec le temps.*

L'impact de ce résultat est capital. D'abord, le raisonnement microscopique heuristique de la section 3.4 a été remplacé par un argument simple qui porte directement sur l'équation limite. Ensuite, même s'il s'agit d'une manifestation de la **seconde loi de la thermodynamique**, la croissance de l'entropie dans le modèle de Boltzmann est déduite d'un *raisonnement logique* et non d'un postulat (une loi) qui ne se discute pas. Enfin, bien sûr, ce faisant, Boltzmann a mis en évidence une **flèche du temps** associée à son équation.

Non seulement cette irréversibilité macroscopique n'est pas contradictoire avec la réversibilité microscopique, mais elle lui est en fait intimement liée : comme on l'a déjà expliqué, c'est la préservation du volume microscopique dans l'espace des phases qui garantit la non-décroissance de l'entropie. Au reste, comme s'en étonnait déjà L. Carleson en 1979 en examinant des modèles simplifiés de l'équation de Boltzmann [35], c'est bien quand les paramètres sont ajustés de manière à avoir la réversibilité microscopique, que l'on obtient du même coup le théorème H . Le phénomène est bien connu dans le contexte de la physique des milieux granulaires [105] : la dynamique microscopique y est dissipative (non réversible), incluant une déperdition d'énergie due aux frottements, et la dynamique macroscopique ne vérifie pas le Théorème H !

Du point de vue informationnel, l'augmentation de l'entropie signifie que le système se dirige toujours vers des états macroscopiques de plus en plus probables. Cette notion de probabilité est exacerbée par la formidable puissance de la combinatoire : supposons par exemple que l'on considère un gaz de $N \simeq 10^{16}$ particules (c'est en gros ce que l'on trouve dans 1mm^3 de gaz aux conditions habituelles !), et qu'entre le temps $t = t_1$ et le temps $t = t_2$ l'entropie augmente de seulement 10^{-5} . Le volume des possibilités microscopiques est alors multiplié par $e^{N[S(t_2)-S(t_1)]} = e^{10^{11}} \gg 10^{10^{10}}$. Ce facteur phénoménal dépasse de très loin le nombre de protons dans l'univers (10^{100} ?) ou le nombre de livres de 1000 pages que l'on peut écrire en combinant tous les caractères alphabétiques de toutes les langues du monde...

L'interprétation intuitive du Théorème H est donc assez parlante : les états de haute entropie occupent au niveau microscopique une place tellement monstrueusement plus grande que les états de basse entropie, que le système microscopique va automatiquement les visiter. Comme on l'a vu, le raisonnement logique justifiant ce scénario est complexe et indirect, passant par la propagation du chaos et le déterminisme macroscopique — et à ce jour seule une petite partie du programme a été réalisée rigoureusement.

5.3 Annulation de la production d'entropie

L'augmentation de l'entropie admet un complément non moins profond, que l'on énonce souvent en deuxième partie du Théorème H : la *caractérisation des cas d'égalité*, c'est à dire des états pour lesquels la production d'entropie s'annule.

On a vu dans (30) que la production d'entropie est égale à

$$\int \text{PE}(f(x, \cdot)) dx, \quad (31)$$

où PE est la fonctionnelle de "production locale d'entropie", agissant sur des distributions cinétiques $f = f(v)$:

$$\text{PE}(f) = \iiint \tilde{B}(v - v_*, \omega) (f(v')f(v'_*) - f(v)f(v_*)) \log \frac{f(v')f(v'_*)}{f(v)f(v_*)} dv dv_* d\omega. \quad (32)$$

Pour tous les modèles raisonnables, on a $\tilde{B}(z, \omega) > 0$ presque partout, et il s'ensuit que la production d'entropie ne s'annule que pour une distribution vérifiant l'équation fonctionnelle

$$f(v')f(v'_*) = f(v)f(v_*) \quad (33)$$

pour (presque) tous v, v_*, ω . En prenant le logarithme dans (33) on retrouve l'équation (29), ce qui montre que f doit être l'exponentielle d'un invariant de collision. Au vu de la forme de ces derniers, et tenant compte de la contrainte d'intégrabilité de f , on obtient $f(v) = e^{a+b\cdot v+c|v|^2/2}$, que l'on peut réécrire

$$f(v) = \rho \frac{e^{-\frac{|v-u|^2}{2T}}}{(2\pi T)^{d/2}}, \quad (34)$$

où $\rho \geq 0$, $T > 0$ et $u \in \mathbb{R}^d$ sont des constantes. Il s'agit donc d'une gaussienne particulière, avec matrice de covariance proportionnelle à l'identité.

Maxwell avait déjà remarqué que (34) annule l'opérateur de collision de Boltzmann : $Q(f, f) = 0$. En son honneur, une telle distribution est appelée **maxwellienne**. Cependant, c'est Boltzmann qui donna le premier un argument convainquant selon lequel les distributions (34) sont *les seules* solutions de l'équation $PE(f) = 0$, et en conséquence les seules solutions de $Q(f, f) = 0$. Rendons-lui hommage en esquissant une variante de sa preuve originale.

Commençons par moyenner (33) sur tous les angles $\sigma = (v' - v'_*)/|v - v_*|$; le membre de gauche $|S^{d-1}|^{-1} \int f' f'_* d\sigma$ est alors la moyenne de la fonction $\sigma \rightarrow f(c + r\sigma) f(c - r\sigma)$ où $c = (v + v_*)/2$ et $r = |v - v_*|$. Cette quantité ne dépend donc que de c et r , ou de manière équivalente de $m = v + v_*$ et $e = (|v|^2 + |v_*|^2)/2$, respectivement la quantité de mouvement et l'énergie totales impliquées dans une collision. Après passage au logarithme, on trouve pour $\varphi = \log f$ l'identité

$$\varphi(v) + \varphi(v_*) = G(m, e). \quad (35)$$

L'opérateur $\nabla_v - \nabla_{v_*}$, appliqué au membre de gauche de (35), donne $\nabla\varphi(v) - \nabla\varphi(v_*)$. Quand on applique le même opérateur au membre de droite, la contribution de m disparaît, et la contribution de e est colinéaire à $v - v_*$, on conclut donc que $F = \nabla\varphi$ vérifie

$$F(v) - F(v_*) \text{ est colinéaire à } v - v_* \text{ pour tous } v, v_*$$

et il est facile d'en déduire que $F(v)$ est une application affine, d'où la conclusion. (Une méthode brutale pour montrer le caractère affine de F , supposant la régularité : on commence par noter, en écrivant une formule de Taylor, que la matrice jacobienne de F est en tout point un multiple de l'identité, soit $\partial_i F^j(v) z_i = \lambda(v) z_j$; ensuite en dérivant par rapport à v_k on déduit que $\partial_{ik} F^j = 0$ si $i \neq j$, et il s'ensuit que tous les coefficients $\partial_i F^j$ s'annulent, après quoi on voit facilement que DF est un multiple de l'identité.)

En conséquence de (31) et (34), les distributions $f(x, v)$ qui annulent la production d'entropie de Boltzmann sont exactement les distributions de la forme

$$f(x, v) = \rho(x) \frac{e^{-\frac{|v-u(x)|^2}{2T(x)}}}{(2\pi T(x))^{d/2}}. \quad (36)$$

On les appelle **maxwelliennes locales** ou encore **états hydrodynamiques**. Par rapport à la description cinétique, ces états sont caractérisés par une réduction considérable de la complexité, puisqu'ils ne dépendent que de trois champs : le champ de densité ρ , le champ de vitesses macroscopiques u , et le champ de température T . Ce sont ces champs qui interviennent dans les équations hydrodynamiques, d'où la terminologie ci-dessus.

Cette découverte jette un pont entre la description cinétique et la description hydrodynamique : dans un régime où les collisions sont très nombreuses (faible nombre de Knudsen), la finitude de la production d'entropie force la dynamique à se concentrer près des distributions qui annulent cette production. Cette remarque donne lieu à un vaste programme d'approximation hydrodynamique de l'équation de Boltzmann, auquel Hilbert fait allusion dans son Sixième Problème. Sur ce sujet on pourra consulter [54, 55, 93]. Si l'équation de Boltzmann peut être approchée aussi bien par des équations compressibles qu'incompressibles, il convient de noter qu'elle ne mène pas à toute la gamme des équations hydrodynamiques, mais seulement à celles des gaz parfaits, c'est à dire suivant une loi de pression proportionnelle à ρT .

6 Convergence entropique : oubli à marche forcée

Si Maxwell a découvert l'importance des profils de vitesses gaussiens en théorie cinétique, il n'a pas, comme le remarque Boltzmann, démontré que ces profils sont effectivement induits par la dynamique. Boltzmann souhaite achever ce programme, et pour cela retrouver les profils maxwelliens non seulement comme distributions d'équilibre, mais aussi comme limites de l'équation cinétique dans l'asymptotique où le temps devient grand ($t \rightarrow +\infty$). Ce saut conceptuel visant à fonder la mécanique statistique d'équilibre sur sa contrepartie hors équilibre, d'habitude beaucoup plus délicate, est toujours d'actualité dans d'innombrables contextes.

J'ai beaucoup écrit sur ce thème et on pourra consulter l'article de revue [103, Chapitre 3], le cours [107], l'article de recherche [46] ou le mémoire de recherche [108]. Dans la suite, pour fixer les idées, je supposerai que la variable de position vit dans le tore \mathbb{T}^d .

6.1 Maxwellienne globale

Nous avons déjà rencontré les maxwelliennes locales qui annulent l'opérateur de collision. Pour annuler également l'opérateur de transport $v \cdot \nabla_x$, il est naturel de chercher des maxwelliennes dont les paramètres ρ, u, T sont homogènes, des constantes indépendantes de la position. Un seul jeu de ces paramètres est compatible avec les lois de conservation de masse, de quantité de mouvement et d'énergie. La distribution ainsi obtenue est appelée **maxwellienne globale**

$$M_{\rho u T} = \rho \frac{e^{-\frac{|v-u|^2}{2T}}}{(2\pi T)^{d/2}}.$$

Sans perte de généralité, quitte à changer de repère galiléen ou d'échelle physique, on pourra supposer que $\rho = 1$, $u = 0$ et $T = 1$ et on notera M la distribution correspondante.

Cette distribution est donc un équilibre pour l'équation de Boltzmann. En outre, il est facile de vérifier que c'est la distribution qui maximise l'entropie sous les contraintes de masse, quantité de mouvement et énergie totales fixées. Ce critère de sélection préfigure la théorie classique de la mécanique statistique d'équilibre et les fameux ensembles canoniques de Gibbs.

6.2 Argument entropique

Boltzmann utilise maintenant le Théorème H pour donner une justification plus solide à la maxwellienne globale : il note que

- l'entropie augmente strictement, sauf si elle est dans un état hydrodynamique
- la maxwellienne globale, stationnaire, est la seule solution hydrodynamique de l'équation de Boltzmann.

L'image qui se dessine est que l'entropie va continuer à augmenter autant que possible, puisque la distribution ne restera jamais "bloquée" sur une solution hydrodynamique; l'entropie finira par approcher l'entropie maximale de la maxwellienne globale, et la convergence en résultera.

À ce sujet, on peut faire deux remarques : la première est que la mesure de Lebesgue, que nous avons prise comme mesure de référence dans l'entropie de Boltzmann, peut être remplacée par la maxwellienne : en effet,

$$H(f) - H(M) = \iint f \log \frac{f}{M} dv dx = H_M(f),$$

où l'on a utilisé le fait que $\log M$ est un invariant de collision. La seconde remarque est que la différence d'entropies permet de quantifier l'écart à la gaussienne d'équilibre, en vertu par exemple de l'inégalité de Csiszàr–Kullback–Pinsker : $H_M(f) \geq \|f - M\|_{L^1}^2 / 2 \|M\|_{L^1}$.

Le raisonnement de Boltzmann est essentiellement correct et il n'est pas difficile de le transformer en un argument rigoureux, montrant que les solutions suffisamment régulières de l'équation de Boltzmann s'approchent de l'équilibre maxwellien. Dans le cadre des solutions spatialement homogènes, T. Carleman mettait en forme ce raisonnement dès 1932 [26].

Cependant, Boltzmann n'avait aucun moyen de rendre l'argument qualitatif; il faudra attendre près d'un siècle avant que l'on ose se poser la question de la vitesse de convergence vers l'équilibre gaussien, particulièrement pertinente car la validité de l'équation de Boltzmann n'est pas éternelle, et limitée dans le temps par des phénomènes tels que le théorème de récurrence de Poincaré.

6.3 L'approche probabiliste de Mark Kac

Au début des années 1950, Kac [66] cherche à comprendre la convergence vers l'équilibre pour l'équation de Boltzmann, et commence par simplifier le modèle. Kac oublie les positions, simplifie à outrance la géométrie des collisions, et invente un modèle stochastique où *l'aléatoire est présent dans l'interaction* : quand deux particules interagissent, on tire au hasard les paramètres qui décrivent la collision. Les positions étant absentes, les particules interagissent toutes les unes avec les autres, et il s'agit donc d'un modèle "de champ moyen". Ce modèle probabiliste simplifié est pour Kac l'occasion de formaliser mathématiquement la notion de propagation du chaos dans les équations de champ moyen, qui s'avèrera si féconde, et sera reprise plus tard par Sznitman [98] et bien d'autres.

Méfiant envers l'équation de Boltzmann, Kac souhaite expliquer la convergence par un raisonnement probabiliste microscopique, au niveau du système à N particules; il cherche à obtenir des estimations de trou spectral uniformes en N . Son approche paraît aujourd'hui naïve dans ce qu'elle sous-estime la difficulté du traitement de la dimension N ; cependant le problème de la détermination du trou spectral

optimal, résolu un demi-siècle plus tard, s'est avéré très intéressant [31, 65, 79]. À ce sujet on pourra consulter également [104, Section 6] et [32] où l'on s'intéresse à la version entropique de ce programme "microscopique".

En 1966 McKean [81] reprend le travail de Kac et dresse un parallèle avec la problématique du théorème central limite. Il introduit dans le sujet des outils de théorie de l'information, en particulier l'information de Fisher [41] qui mesure la difficulté à estimer un paramètre tel que la vitesse des particules. Le programme sera complété par Tanaka [100] qui découvrira de nouvelles distances contractantes, et culminera avec les travaux de Carlen, Carvalho, Gabetta, Lu, Toscani sur le "théorème central limite pour l'équation de Boltzmann" [29, 30, 33, 34]. Cette théorie comprend aussi bien des théorèmes de convergence basés sur la combinatoire des interactions entre particules et des outils communs avec le théorème central limite (distances faibles...), que des contre-exemples présentant des convergences extrêmement lentes vers l'équilibre.

Ce programme stochastique permet donc de se dispenser du Théorème H ; en fait il a aussi permis de mettre en jour plusieurs autres fonctionnelles de Lyapunov : distance contractante (de transport optimal) de Tanaka, information de Fisher. Cependant du point de vue technique toute la théorie reste essentiellement confiné aux interactions maxwelliennes (dans lesquelles $\tilde{B}(v - v_*, \omega)$ dépend uniquement de l'angle entre $v - v_*$ et ω), et aux gaz spatialement homogènes. Le chapitre 4 de [103] est consacré aux propriétés particulières, fort élégantes au demeurant, de ces interactions.

Pour gagner en généralité et étudier des situations inhomogènes ou des interactions non maxwelliennes, la seule approche robuste connue à ce jour est basée sur le Théorème H .

6.4 Conjecture de Cercignani

Le Théorème H de Boltzmann est général et pertinent, il est donc naturel d'en chercher des raffinements quantitatifs. Au début des années 1980, C. Cercignani se demandait si l'on pouvait *minorer* la production d'entropie locale en fonction de la "non-gaussianité" de la distribution cinétique; idéalement par un multiple de l'information $H_M(f)$. Il fallut attendre une dizaine d'années pour que Carlen et Carvalho [28], Desvillettes [44], sans répondre à la question de Cercignani, puissent cependant apporter des bornes inférieures quantitatives à la production d'entropie.

Une réponse plus précise a ce problème est obtenue dans mes articles [101] et [103] (le premier en collaboration avec G. Toscani). Sans perte de généralité on suppose que $\int f dv = 1$, $\int f v dv = 0$, $\int f |v|^2 dv = d$, le cas général pouvant se déduire par changement d'échelle ou de référentiel.

Commençons par mentionner un résultat étonnant de simplicité, tiré de [104], qui s'applique dans un cas non physique : si $B(v - v_*, \sigma) \geq K(1 + |v - v_*|^2)$, alors

$$\text{PE}(f) \geq \left(K_B \frac{|S^{d-1}|}{8} \frac{d-1}{1+2d} \right) T_f^* H_M(f), \quad (37)$$

où

$$T_f^* = \inf_{e \in S^{d-1}} \int f(v) (v \cdot e)^2 dv.$$

La quantité T_f^* quantifie la non-concentration de f près d'un hyperplan; elle se minore en fonction d'une information d'entropie, ou de régularité, ou bien automatiquement pour des distributions radialement symétriques. La concision du résultat masque une technique de preuve surprenante où l'on régularise f par un semigroupe auxiliaire de diffusion; sous l'effet de ce semigroupe la variation de la production d'entropie de Boltzmann est essentiellement minorée par la production d'entropie de Landau, que l'on estime en termes de l'information de Fisher avant d'intégrer le long du semigroupe. On renvoie à [104] ou [107] pour les détails.

L'hypothèse de croissance quadratique en la vitesse relative n'est pas physiquement réaliste; elle est cependant optimale au sens où il existe des contre-exemples [16] pour des noyaux à croissance $|v - v_*|^\gamma$, pour tout $\gamma < 2$. On peut ensuite travailler sur l'inégalité (37) pour en déduire une minoration plus faible s'appliquant à des sections efficaces réalistes, telles que le modèle des sphères dures; la principale difficulté consiste à contrôler la quantité de production d'entropie qui est induite par les petites vitesses relatives ($|v - v_*| \leq \delta$); les logarithmes rendent ce contrôle délicat. On renvoie à [104] pour les détails. À la fin, on obtient, pour tout $\varepsilon > 0$ arbitrairement petit, l'inégalité

$$\text{PE}(f) \geq K_\varepsilon(f) [H(f) - H(M^f)]^{1+\varepsilon}, \quad (38)$$

où M^f est la maxwellienne associée à f , i.e. celle dont les paramètres ρ, u, T correspondent à la densité, vitesse moyenne et température de f . La constante $K_\varepsilon(f)$ ne dépend que de ε , de la régularité C^r de f pour r assez grand, d'un moment $\int f|v|^s dv$ pour s assez grand, et d'une borne inférieure $f \geq K e^{-A|v|^q}$. Cela reste un problème ouvert de savoir si ces hypothèses peuvent être relaxées...

6.5 Convergence conditionnelle

L'inégalité (38) concerne une fonction $f = f(v)$ mais n'inclut pas de dépendance en espace; c'est inévitable, puisque la variable x n'intervient pas dans l'étude de la production globale d'entropie. Bien sûr, (38) implique immédiatement (modulo les bornes de régularité) une convergence en $O(t^{-\infty})$ pour l'équation spatialement homogène, c'est à dire que la distance entre la distribution et l'équilibre tend vers 0 plus vite que t^{-k} pour tout k ; cependant cette inégalité ne résout pas le problème inhomogène. L'obstacle à surmonter est la *dégénérescence de la production d'entropie sur les états hydrodynamiques*. Une clé de l'étude en temps grand de l'équation de Boltzmann consiste donc à montrer que l'on ne passe pas trop de temps dans un état hydrodynamique, ou approximativement hydrodynamique. Pour éviter ce piège, nous ne pouvons compter que sur le transport, représenté par l'opérateur $v \cdot \nabla_x$. Grad [56] l'avait compris dès 1965, dans un texte au demeurant passablement obscur : *“the question is whether the deviation from a local Maxwellian, which is fed by molecular streaming in the presence of spatial inhomogeneity, is sufficiently strong to ultimately wipe out the inhomogeneity” (...)* *“a valid proof of the approach to equilibrium in a spatially varying problem requires just the opposite of the procedure that is followed in a proof of the H-Theorem, viz., to show that the distribution function does not approach too closely to a local Maxwellian.”*

Dans les années 2000, Desvillettes et moi-même [46] redécouvrons ce principe formulé par Grad, et établissons une version de l'**instabilité de l'approximation**

hydrodynamique : si le système devient, à un moment donné, proche d'être hydrodynamique, sans être pour autant à l'équilibre, alors les phénomènes de transport le font sortir de cet état hydrodynamique. Ceci est quantifié, sous des hypothèses de forte régularité, par l'étude des variations secondes d'un carré de norme, $\|f - M^f\|^2$, entre $f = f(t, x, v)$ et la maxwellienne *locale* associée,

$$M^f(t, x, v) = \rho(t, x) \frac{e^{-\frac{|v-u(t,x)|^2}{2T(t,x)}}}{(2\pi T(t, x))^{d/2}},$$

$$\begin{aligned} \rho(t, x) &= \int f(t, x, v) dv, & u(t, x) &= \frac{1}{\rho(t, x)} \int f(t, x, v) v dv, \\ T(t, x) &= \frac{1}{d\rho(t, x)} \int f(t, x, v) |v - u(t, x)|^2 dv. \end{aligned}$$

En quelque sorte, M^f est la meilleure approximation possible de f par un état hydrodynamique, et l'étude des variations de $\|f - M^f\|$ permet de vérifier que f ne peut rester trop longtemps proche de l'état hydrodynamique.

En mettant bout à bout (de manière particulièrement technique, et à l'aide de nombreuses inégalités fonctionnelles) le Théorème H quantitatif et l'instabilité de l'approximation hydrodynamique, nous aboutissons à la **convergence conditionnelle** : une solution de l'équation de Boltzmann vérifiant des bornes de régularité uniforme converge vers l'équilibre en $O(t^{-\infty})$. Ce résultat est constructif au sens où les constantes de temps impliquées ne dépendent que des bornes de régularité, de la forme de l'interaction et des conditions aux limites. La convergence repose sur un système d'inégalités qui font intervenir simultanément l'entropie et la distance aux états hydrodynamiques : par exemple, l'une d'entre elles s'écrit

$$\frac{d^2}{dt^2} \|f - M^f\|_{L^2}^2 \geq K \int |\nabla T|^2 dx - \frac{C}{\delta^{1-\varepsilon}} (\|f - M^f\|_{L^2}^2)^{1-\varepsilon} - \delta[H(f) - H(M)]. \quad (39)$$

Pour comprendre l'apport d'une telle inégalité, supposons que f devienne hydrodynamique à un moment : alors $f = M^f$, et (38) ne sert plus à rien. Mais si la température est inhomogène, et si δ est très petit dans (39), alors on se retrouve avec $(d^2/dt^2)\|f - M^f\|_{L^2}^2 \geq \text{const.}$, ce qui bien sûr empêche f de rester proche de M^f pendant trop longtemps. Une fois que f est sortie de l'approximation hydrodynamique, on peut ré-appliquer (38), et ainsi de suite. Ce raisonnement ne fonctionne que quand la température est inhomogène, mais on peut trouver d'autres inégalités faisant intervenir les gradients de vitesse macroscopique et de densité. On est ainsi passé d'un argument "passif" à un argument "actif", où l'augmentation de l'entropie est forcée par des inégalités différentielles plutôt que par l'identification d'une limite.

Finissons cette section par quelques commentaires sur les hypothèses. La théorie de la régularité de l'équation de Boltzmann permet de ramener les bornes générales à des bornes bien particulières; par exemple, on sait que la distribution cinétique est automatiquement minorée par un multiple de $e^{-|v|^q}$ si par exemple l'équation est posée dans le tore et que la solution est régulière. On sait aussi que des bornes sur les moments d'ordre bas permettent d'avoir des bornes sur les moments arbitrairement élevés, etc. Mais la régularité dans le cadre général reste un célèbre problème ouvert. Le résultat de convergence conditionnelle montre que c'est le dernier obstacle qui

nous sépare des estimations quantitatives de convergence vers l'équilibre; il unifie également les résultats déjà connus sur la convergence : aussi bien le cas de distributions spatialement homogènes, que celui de distributions proches de l'équilibre, sont des situations dans lesquelles on dispose d'une théorie de la régularité à peu près complète.

6.6 Hypocoercivité

Dans l'étude de la convergence vers l'équilibre pour l'équation de Boltzmann, nous assistons à une interaction subtile entre l'opérateur de collision (non linéaire, dissipatif dégénéré) et l'opérateur de transport (linéaire, conservatif). Aucun des deux, pris séparément, ne serait suffisant pour induire la convergence, mais la combinaison des deux y parvient. Cette situation est assez fréquente et rappelle la problématique de l'hypoellipticité en théorie de la régularité des équations aux dérivées partielles. Par analogie, le problème de l'*hypocoercivité* s'intéresse aux propriétés de convergence pour des équations éventuellement dégénérées.

Une étude quelque peu systématique de ces situations, aussi bien pour des équations linéaires que non linéaires, est effectuée dans mon mémoire [108]; on pourra aussi consulter l'article introductif [106]. La stratégie générale consiste à construire des fonctionnelles de Lyapunov adaptées à la dynamique, en ajoutant à la fonctionnelle naturelle (comme l'entropie) un terme d'ordre inférieur bien choisi. Un cas d'école est le "théorème A^*A+B ", inspiré du théorème des carrés de Hörmander, qui donne des conditions suffisantes sur les commutateurs entre les opérateurs A et B , avec B antisymétrique, pour que l'évolution $e^{-t(A^*A+B)}$ soit hypocoercive. Dans la variante la plus simple, l'une de ces conditions, réminiscente de la condition d'algèbre de Lie de Hörmander, est la coercivité de $A^*A + [A, B]^*[A, B]$.

La théorie de l'hypocoercivité a maintenant acquis sa vie propre avec déjà un certain nombre de résultats marquants, et continue son expansion, surtout dans un cadre linéaire. Ceci est vrai aussi bien dans en théorie cinétique, comme le travail [58] que l'on évoquera dans la section suivante, que en-dehors de la théorie cinétique, comme dans le travail de Liverani et Olla sur les limites hydrodynamiques de certains systèmes de particules [73].

Dans un cadre non linéaire, le principal résultat reste [108, Théorème 51]; cet énoncé général permet de simplifier la preuve du théorème de convergence conditionnelle pour l'équation de Boltzmann, et d'y inclure de nouvelles interactions et conditions aux limites. On renvoie à [108, Partie III] pour plus de détails.

6.7 Régime linéarisé

Le taux de convergence vers l'équilibre peut être précisé au moyen de l'étude du linéarisé. Commençons par débusquer une faute logique classique : l'étude du linéarisé ne peut en aucun cas se substituer à l'étude non linéaire, puisque la linéarisation n'est valable précisément qu'à partir du moment où la distribution est très proche de l'équilibre.

L'étude de la convergence linéarisée suppose de surmonter trois principales difficultés :

- estimer quantitativement le trou spectral de l'opérateur de collision linéarisé;
- effectuer une étude spectrale du linéarisé dans un espace qui convienne au problème non linéaire, afin d'effectuer un "raccord" entre l'étude non linéaire et

l'étude linéarisée;

- prendre en compte la dégénérescence hydrodynamique, dans une perspective hypocoercive : en effet, l'équation linéarisée est tout autant dégénérée que l'équation non linéaire.

Toutes ces difficultés ont été résolues dans la dernière décennie par C. Mouhot et ses collaborateurs Baranger, Gualdani et Mischler [11, 58, 82], au moins dans le cas emblématique des sphères dures. C'est ainsi que l'article récent [58] établit un résultat de convergence conditionnelle avec taux exponentiel $O(e^{-\lambda t})$ au lieu de $O(t^{-\infty})$, et le taux λ est estimé de manière constructive.

La convergence exponentielle n'est pas une caractéristique universelle de l'équation de Boltzmann : on ne l'attend que pour des potentiels durs ou modérément mous. Supposons pour fixer les idées que le noyau de collision se comporte comme $|v - v_*|^\gamma b(\cos \theta)$. Dans le cas où $b(\cos \theta) \sin^{d-2} \theta$ est intégrable (souvent obtenu par troncature angulaire aux collisions rasantes), l'opérateur de collision linéarisé n'admet un trou spectral que pour $\gamma \geq 0$. Une abondance de collisions rasantes permet d'étendre cette condition, comme l'ont montré Mouhot et Strain [83] : si $b(\cos \theta) \sin^{d-2} \theta \simeq \theta^{-(1+\nu)}$ pour $\theta \rightarrow 0$ (collisions rasantes importantes), alors l'opérateur de collision linéarisé n'admet un trou spectral que pour $\gamma + \nu \geq 0$. La théorie de la régularité est actuellement en cours de développement pour de telles équations (travaux de Gressman–Strain, Alexandre–Morimoto–Ukai–Xu–Yang), et on peut parier que d'ici quelques années la théorie linéarisée couvrira tous ces cas.

Pour des potentiels trop mous (ou pour le modèle de collisions coulombiennes de Landau), on n'a pas de trou spectral et le meilleur résultat espérable est une convergence en exponentielle fractionnaire $O(e^{-\lambda t^\beta})$, $0 < \beta < 1$. On trouvera de telles estimées dans les travaux de Guo et Strain [60].

6.8 Évolution qualitative de l'entropie

Un thème récurrent dans toute cette partie est la dégénérescence liée aux états hydrodynamiques, qui gêne la convergence vers l'équilibre. Au début des années 2000, Desvillettes et moi-même suggérions que cette dégénérescence se reflétait dans des oscillations de la production d'entropie. Jamais observées auparavant, ces oscillations ont été mises en évidence dans les simulations numériques très précises de F. Filbet. J'ai reproduit ci-dessous une courbe frappante, obtenue avec l'équation de Boltzmann dans une géométrie monodimensionnelle périodique.

Sur ce schéma, on a tracé le logarithme de la fonction H y en fonction du temps; la décroissance globalement rectiligne correspond donc à une convergence exponentielle vers l'état d'équilibre. On a également séparé l'information cinétique en information hydrodynamique et information "purement cinétique" :

$$\int f \log \frac{f}{M} = \left(\int \rho \log \frac{\rho}{T^{d/2}} \right) + \int f \log \frac{f}{Mf};$$

la deuxième quantité (information purement cinétique) est la courbe que l'on voit en-dessous de la courbe de la fonction H . Quand les deux courbes sont éloignées, la distribution est presque hydrodynamique, quand elles sont proches la distribution est presque homogène. Partant d'une distribution hydrodynamique, on s'en écarte tout de suite, conformément au principe d'instabilité de l'approximation hydrodynamique. Ensuite on distingue nettement des oscillations entre des états

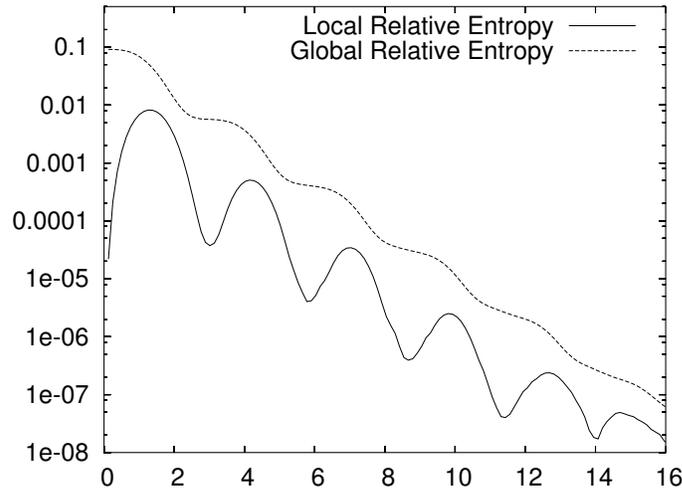


Figure 1: Évolution logarithmique de la fonction H cinétique et de la fonction H hydrodynamique, pour l'équation de Boltzmann dans une boîte périodique.

plutôt hydrodynamiques, associés à un ralentissement de la production d'entropie, et des états plutôt homogènes; ces oscillations sont importantes compte tenu de la nature logarithmique du diagramme. Filbet, Mouhot et Pareschi [50] présentent d'autres courbes et tentent d'expliquer la fréquence d'oscillation dans un certain régime asymptotique.

Ici l'équation de Boltzmann révèle bien sa double nature, relevant aussi bien des transferts d'information via les collisions, que de la mécanique des fluides via l'opérateur de transport. C'est souvent le mariage entre les deux aspects qui s'avère délicat.

L'importance relative du transport et des collisions peut être modulée par les conditions aux limites; dans un cadre périodique, il s'agit de faire varier la taille de la boîte. Une boîte large permettra des variations spatiales importantes, laissant libre cours aux effets hydrodynamiques, comme dans la simulation ci-dessus. Cependant, on voit bien que même dans ce cas, et contrairement à une idée bien enracinée même chez les spécialistes, *le régime asymptotique n'est pas hydrodynamique*, au sens où le rapport entre l'entropie hydrodynamique et l'entropie cinétique totale n'augmente pas significativement au fur et à mesure que le temps passe, oscillant plutôt entre des valeurs minimum et maximum.

On peut se demander ce qui se passe dans une boîte plutôt petite. Une telle simulation est présentée ci-après.

La conclusion de cette figure est précisément à l'opposé de l'intuition selon laquelle les effets hydrodynamiques dominent en temps grand : bien au contraire, partant d'une situation hydrodynamique, on arrive rapidement à une situation presque homogène (à vue on a l'impression que au temps $t \simeq 0.7$ l'information hydrodynamique ne représente guère plus de 1% de l'information totale !). Les effets inhomogènes reprennent ensuite leurs droits (au temps $t = 1$ l'information est partagée en parts du même ordre), après quoi on redevient résolument homogène. Dans cet exemple, l'homogénéisation a agi plus vite que la convergence vers l'équilibre. On reviendra dans la section 8 sur ce graphique qui a suscité une certaine perplexité.

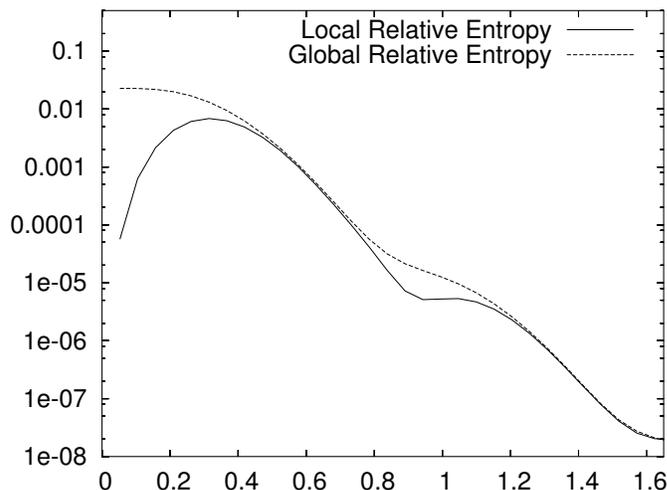


Figure 2: Même chose, dans une boîte plus petite.

6.9 Deux problèmes non conventionnels

Je terminerai cette section en mentionnant deux curieux problèmes liés à la flèche du temps dans l'équation de Boltzmann, qui sont peut-être simplement des curiosités. Le premier est la classification de solutions éternelles de l'équation de Boltzmann : j'avais tenté de montrer, au cours de ma thèse, que, au moins pour l'équation de Boltzmann spatialement homogène avec molécules maxwelliennes, il n'existe pas de solution éternelle d'énergie finie. La seconde consistait au contraire à chercher des solutions autosimilaires d'énergie infinie, ne convergeant pas vers l'équilibre maxwellien. Sur le premier problème, on pourra consulter [110] pour des résultats partiels; la conjecture tient toujours, et Bobylev et Cercignani [18] ont pu montrer qu'il n'existe pas de solution éternelle ayant des moments finis à tous ordres. Quant au second problème, il a été résolu par les mêmes auteurs [17], en utilisant des techniques de transformée de Fourier.

7 Relaxation isentropique : s'accommoder de ses souvenirs

On considère maintenant l'équation de Vlasov avec potentiel d'interaction W :

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f - \left(\nabla W * \int f dv \right) \cdot \nabla_v f = 0. \quad (40)$$

Contrairement à l'équation de Boltzmann, l'équation (40) n'impose pas de flèche du temps, et reste inchangée sous l'action d'un renversement temporel. La constance de l'entropie correspond à une préservation d'information microscopique. L'équation de Vlasov au temps t permet théoriquement de reconstituer la condition initiale sans perte de précision, en résolvant simplement l'équation de Vlasov après avoir renversé les vitesses.

En outre, alors que l'équation de Boltzmann n'admet qu'un très petit nombre d'équilibres (les maxwelliennes déterminées par les lois de conservation), l'équation

de Vlasov en admet une quantité considérable. Par exemple, *toutes* les distributions homogènes $f^0 = f^0(v)$ sont stationnaires. Il existe encore bien d'autres distributions stationnaires, par exemple la famille des ondes de Bernstein–Greene–Kruskal [14].

Pour toutes ces raisons, on ne voit rien a priori qui laisse supposer un comportement bien déterminé en temps grand; aucune indication de flèche du temps. Pourtant, en 1946, L. Landau, sorti quelques années auparavant des geôles du régime communiste soviétique où son franc-parler l'avait mené, suggéra un comportement bien spécifique pour l'équation de Vlasov en temps grand. Il se basait sur une analyse de l'équation linéarisée autour d'un équilibre homogène. La prédiction de Landau a provoqué une secousse et un changement conceptuel qui suscitent encore aujourd'hui de vives discussions [92]; à sa suite on se met à soupçonner que la convergence vers l'équilibre n'est pas forcément liée à une augmentation d'entropie.

Cette section est consacrée à un survol de la question de la convergence isentropique, en insistant sur le régime perturbatif qui est le seul sur lequel nous ayons des éléments solides. On trouvera plus de détails dans mon cours [109].

7.1 Analyse linéarisée

Étudions l'équation de Vlasov près d'un équilibre homogène $f^0(v)$. Si l'on pose $f(t, x, v) = f^0(v) + h(t, x, v)$, l'équation devient

$$\frac{\partial h}{\partial t} + v \cdot \nabla_x h + F[h] \cdot \nabla_v f^0 + F[h] \cdot \nabla_v h = 0, \quad (41)$$

où

$$F[h](t, x, v) = - \iint \nabla W(x - y) h(t, y, v) dv$$

est la force induite par la distribution h .

En négligeant le terme quadratique $F[h] \cdot \nabla_v h$ dans (41), on obtient l'**équation de Vlasov linéarisée** près d'un équilibre homogène :

$$\frac{\partial h}{\partial t} + v \cdot \nabla_x h + F[h] \cdot \nabla_v f^0 = 0. \quad (42)$$

Avant d'examiner (42), considérons le cas sans interaction ($W = 0$), soit le **transport libre** $\partial_t f + v \cdot \nabla_x f = 0$. Cette équation se résout dans $\mathbb{T}_x^d \times \mathbb{R}_v^d$ en $f(t, x, v) = f_i(x - vt, v)$, où f_i est la distribution initiale. Passons en variables de Fourier, en posant

$$\tilde{g}(k, \eta) = \iint g(x, v) e^{-2i\pi k \cdot x} e^{-2i\pi \eta \cdot v} dx dv;$$

la solution du transport libre s'écrit alors

$$\tilde{f}(t, k, \eta) = \tilde{f}_i(k, \eta + kt). \quad (43)$$

Dès que $k \neq 0$, cette expression tend vers 0 quand $t \rightarrow \infty$, avec une vitesse déterminée par la régularité de f_i dans la variable de vitesse (principe de Riemann–Lebesgue). Tous les modes spatiaux non nuls relaxent donc vers 0; c'est l'action **homogénéisante** du transport libre.

L'équation (42) ne se résout pas si aisément; cependant, si l'on pose $\rho(t, x) = \int h(t, x, v) dv$, alors on constate que les modes $\hat{\rho}(t, k)$ vérifient tous des équations

indépendantes pour des valeurs distinctes de k . Cette remarquable propriété de *découplage* des modes est à la base de l'analyse de Landau. Pour chaque k on dispose d'une **équation de Volterra** sur le mode k :

$$\widehat{\rho}(t, k) = \widetilde{f}_i(k, kt) + \int_0^t K^0(k, t - \tau) \widehat{\rho}(\tau, k) d\tau,$$

où

$$K^0(t, k) = -4\pi^2 \widehat{W}(k) \widetilde{f}^0(kt) |k|^2 t.$$

La stabilité des équations de Volterra est un problème classique. Si u vérifie $u(t) = S(t) + \int_0^t K(t - \tau) u(\tau) d\tau$, alors le taux de décroissance de u est dicté par le pire de deux taux : le taux de décroissance de S bien sûr, et d'autre part la largeur de la plus grande bande $\{0 \leq \operatorname{Re} \xi \leq \Lambda\}$ qui ne rencontre pas de solution de l'équation $K^L = 1$, où K^L est la transformée de Laplace de K . Si $\Lambda > 0$, on a donc stabilité exponentielle pour le linéarisé.

Adapté à notre contexte, ce résultat mène au **critère de stabilité de Penrose**, dont on va énoncer la version multidimensionnelle. Pour tout $k \in \mathbb{Z}^d$, on définit $f_k^0 : \mathbb{R} \rightarrow \mathbb{R}_+$ par

$$f_k^0(r) = \int_{k^\perp} f^0\left(\frac{k}{|k|} r + z\right) dz;$$

en clair, f_k^0 est la marginale de f^0 dans la direction k . Le critère de Penrose [88] demande que pour tout $k \in \mathbb{Z}^d$,

$$\forall \omega \in \mathbb{R}, \quad (f_k^0)'(\omega) = 0 \implies \widehat{W}(k) \int \frac{(f_k^0)'(v)}{v - \omega} dv < 1.$$

Si ce critère (essentiellement optimal) est satisfait, alors il y a stabilité exponentielle pour le linéarisé : la force décroît exponentiellement vite, de même que toutes les inhomogénéités de la densité spatiale $\int h dv$.

Le critère de stabilité de Penrose est vérifié dans de nombreuses situations : dans le cas d'une interaction coulombienne, dès que les marginales de f^0 sont croissantes à gauche de 0, décroissantes à droite (autrement dit, si $(f_k^0)'(z)/z < 0$ pour $z \neq 0$); en particulier si f^0 est une fonction décroissante de $|v|$, une gaussienne par exemple. Toujours dans le cas coulombien, en dimension 3 ou plus, le critère est vérifié si f^0 est isotrope. Dans le cas de l'attraction newtonienne, les choses sont plus complexes : par exemple, pour une distribution gaussienne, la stabilité dépend de la masse et de la température de la distribution. Ceci reflète la célèbre **instabilité de Jeans**, selon laquelle l'équation de Vlasov est linéairement instable aux longueurs plus grandes que

$$L_J = \sqrt{\frac{\pi T}{\mathcal{G} \rho^0}},$$

avec \mathcal{G} la constante de gravitation universelle, ρ^0 la masse de la distribution f^0 et T sa température. C'est cette instabilité qui est responsable de la tendance des corps massifs à se regrouper en "clusters" (galaxies, amas de galaxies, etc.).

En résumé, l'équation de Vlasov linéarisée autour d'un équilibre homogène stable (au sens de Penrose) prédit un amortissement exponentiel de la force, de manière apparemment irréversible. Cette découverte faisait revenir la problématique de la flèche du temps dans la théorie de l'équation de Vlasov.

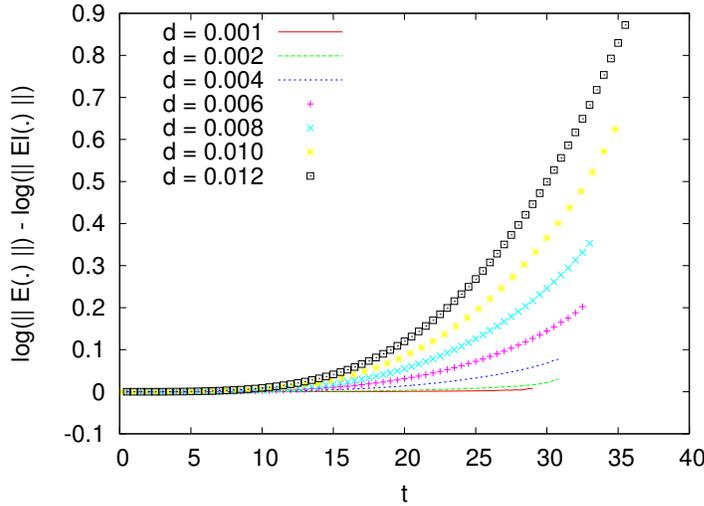


Figure 3: Pour une évolution de Vlasov, rapport logarithmique entre les normes de l'énergie suivant l'équation non linéaire, et suivant l'équation linéaire, pour différentes amplitudes de perturbation.

Le traitement linéarisé de l'équation de Vlasov se trouve dans tous les traités avancés de physique des plasmas, comme [71]; cependant, le traitement y est systématiquement obscurci par l'usage d'intégrales de contour dans le plan complexe, qui proviennent de l'inversion de la transformée de Laplace. Ceci est évité dans la présentation de [85, Section 3], basée sur la simple transformée de Fourier; ou dans la version courte [109].

7.2 Amortissement Landau non linéaire

La linéarisation effectuée par Landau n'est peut-être pas une opération innocente, et depuis un demi-siècle des doutes ont été émis sur sa validité. En 1960, Backus [6] faisait remarquer que remplacer dans le terme de force $\nabla_v(f^0 + h)$ par $\nabla_v f^0$ serait conceptuellement simple si $\nabla_v h$ restait petit pour tous temps; mais si l'on remplace h par la solution de l'équation linéarisée, on constate que son gradient en vitesses croît linéairement en temps, devenant arbitrairement grand. Ceci, suggère Backus, "détruit la validité de la théorie linéaire". L'argument de Backus est contestable car $\nabla_v h$ est multiplié par $F[h]$ que l'on s'attend à voir décroître exponentiellement; il n'empêche que des considérations heuristiques [86] suggèrent l'échec de l'approximation linéaire au bout d'un temps $O(1/\sqrt{\delta})$ où δ est la taille de la perturbation initiale. La courbe ci-après (tracée par F. Filbet) représente le logarithme du quotient entre l'énergie calculée selon l'équation non linéaire, et celle qui est obtenue selon l'équation linéaire, pour différentes valeurs de l'amplitude δ de perturbation; on voit clairement que même si δ est petit, on finit par arriver dans un régime où les effets non linéaires ne peuvent être négligés.

Il y a d'autres raisons pour se méfier de la linéarisation. D'abord, le terme éliminé, $F[h] \cdot \nabla_v h$, est de plus haut degré en termes de dérivées de h en vitesse. Ensuite, la linéarisation élimine la conservation de l'entropie, et privilégie l'état particulier f^0 , ce qui rend caduque la discussion sur la réversibilité.

En 1997, Isichenko [63] jette un pavé dans la mare en argumentant que la

convergence vers l'équilibre ne peut être en général plus rapide que $O(1/t)$ pour l'équation non linéaire. Il est aussitôt contredit par Caglioti et Maffei [25] qui construisent des solutions exponentiellement amorties de l'équation non linéaire. Les simulations numériques (voir ci-dessous) ne sont pas très fiables en temps très grand, et le besoin d'un théorème se fait sentir.

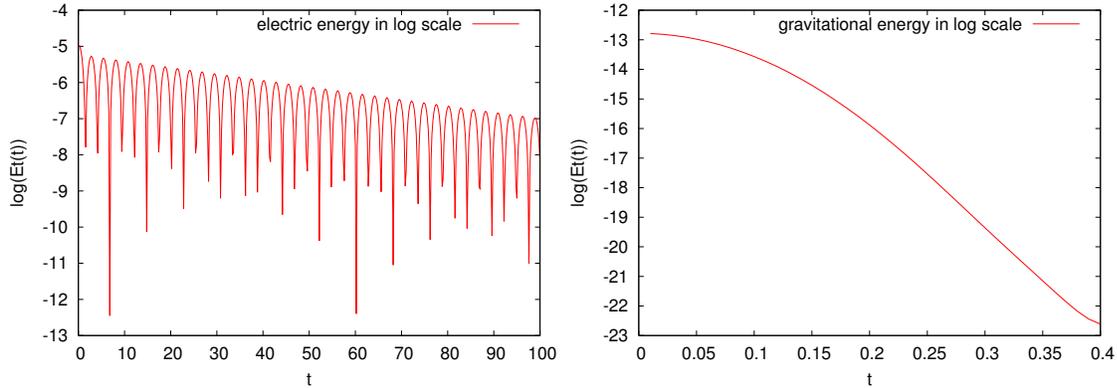


Figure 4: Évolution de la norme du champ de force, pour des interactions électrostatique (à gauche) et gravitationnelle (à droite). Dans le cas électrostatique, les oscillations rapides sont appelées ondes de Langmuir.

En 2009, Mouhot et moi-même établissons un tel résultat [85]. Si le potentiel d'interaction W n'est pas trop singulier, au sens où $\widehat{W}(k) = O(1/|k|^2)$ (cette hypothèse autorise tout juste les interactions de Coulomb et Newton !), et si f^0 est un équilibre homogène analytique vérifiant la condition de stabilité de Penrose, alors il y a stabilité non linéaire dynamique : partant d'une donnée initiale f_i analytique telle que $\|f_i - f^0\| = O(\delta)$, avec δ très petit, on a décroissance de la force en $O(e^{-2\pi\lambda|t|})$, pour tout $\lambda < \min(\lambda_0, \lambda_i, \lambda_L)$, où λ_0 est la largeur de la bande d'analyticité complexe de f^0 autour de \mathbb{R}_v^d , λ_i est la largeur de la bande d'analyticité complexe de f_i dans la variable v , et λ_L est le taux de convergence de Landau. En somme, l'amortissement linéaire implique l'amortissement non linéaire, avec une perte arbitrairement petite sur le taux de convergence.

Le théorème établit aussi la convergence faible de $f(t, \cdot)$ vers un état asymptotique homogène $f_\infty(v)$. Plus précisément, l'équation étant invariante par renversement du temps, il y a un profil asymptotique $f_{+\infty}$ pour $t \rightarrow +\infty$, et un autre profil $f_{-\infty}$ pour $t \rightarrow -\infty$. Si l'on voit l'équation de Vlasov comme un système dynamique, on a donc un comportement remarquable : les trajectoires homoclines/hétéroclines sont si nombreuses qu'elles emplissent un plein voisinage de f^0 en topologie analytique.

L'amortissement non linéaire de l'équation de Vlasov repose sur le **confinement** et le **mélange**. Le confinement est indispensable : on sait que l'amortissement Landau n'a pas lieu dans tout l'espace, même pour l'équation linéarisée [52, 53]; dans notre cas il est automatique car l'espace des phases est le tore. Le mélange a lieu à cause du phénomène de vitesse différentielle : des particules qui ont des vitesses différentes se déplacent à des vitesses différentes dans l'espace des phases; ici aussi c'est presque une tautologie. Un exemple de système non mélangeant est l'oscillateur harmonique, où les trajectoires portées par des variables d'action différentes se déplacent à une vitesse angulaire constante. Certains des autres ingrédients sous-

jacents à l'étude non linéaire sont

- une ré-interprétation du problème en termes de régularité : au lieu de montrer qu'il y a amortissement, on montre que $f(t, x, v)$ est "aussi régulière" que la solution du transport libre, uniformément en temps;
- des estimations de "scattering" : une particule placée dans un champ de forces exponentiellement décroissant suit une trajectoire asymptotique au transport libre, en un sens que l'on peut quantifier précisément;
- le rôle stabilisant de la réponse avec retard, *en échos*, du plasma : quand un des modes du plasma est perturbé, la réaction des autres modes n'est pas instantanée, mais survient avec un léger retard, car l'effet des modes se compense en-dehors de certains instants de résonance;
- un schéma de Newton, qui tire parti de ce que le système de Vlasov linéarisé est en quelque sorte complètement intégrable; la vitesse de convergence de ce schéma permet de compenser la perte de décroissance qui vient avec la résolution du linéarisé.

Tous ces ingrédients sont décrits plus en détail dans [109]. La place particulière du schéma de Newton et de la complète intégrabilité forment un pont inattendu avec la théorie KAM (Kolmogorov–Arnold–Moser). En quelque sorte, l'équation de Vlasov non linéaire, dans le régime perturbatif, hérite certaines des bonnes propriétés de l'équation complètement intégrable de Vlasov linéarisé.

Du point de vue physique, l'information va vers les **petites échelles cinétiques** : les oscillations de la fonction de distribution s'amplifient quand le temps devient grand, et deviennent *invisibles*. Lynden-Bell [75, 76] l'avait bien compris, et utilise une formule frappante pour l'expliquer : "A [galactic] system whose density has achieved a steady state will have information about its birth still stored in the peculiar velocities of its stars".

Ces oscillations, bien visibles sur les figures ci-dessous, sont à la fois une nuisance du point de vue technique, et le mécanisme physique fondamental qui produit l'impression d'irréversibilité. On note la différence avec le mécanisme dit de *radiation*, dans lequel l'énergie est émise à échelle macroscopique et s'en va à l'infini : ici au contraire l'énergie disparaît littéralement dans le néant...

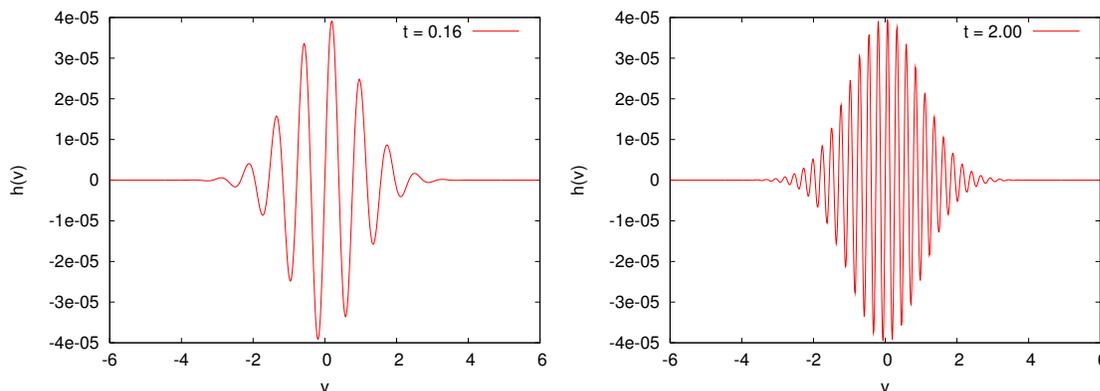


Figure 5: Une tranche de la fonction de distribution (par rapport à un équilibre homogène) pour l'amortissement Landau gravitationnel, à deux temps différents.

7.3 Régularité glissante

Le théorème d'amortissement non linéaire repose sur une ré-interprétation nouvelle en termes de régularité, qui mérite quelques commentaires. Commençons par parler de la **cascade** associée au transport libre, représentée sur le diagramme ci-dessous :

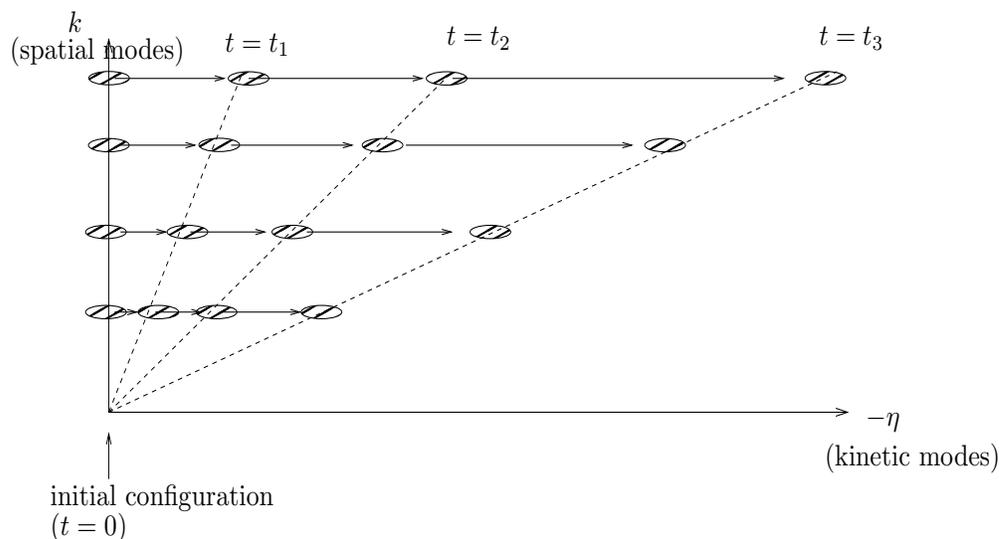


Figure 6: Évolution de l'énergie dans l'espace des fréquences le long du transport libre ou d'une perturbation de ce dernier; les marques indiquent la localisation de l'énergie dans l'espace des phases.

Cette image, qui se déduit de la formule (43), montre que les fréquences qui comptent varient au cours du temps : il y a un mouvement d'ensemble vers les hautes fréquences cinétiques, et ce mouvement est d'autant plus rapide que la fréquence spatiale est élevée. Plus précisément, le mode spatial de fréquence k oscille en vitesse avec une période en $O(1/|k|t)$. L'enjeu de l'amortissement Landau consiste à montrer que cette cascade, bien que déformée, est globalement préservée par l'effet des interactions, qui *couplent* les différents modes.

Ces fortes oscillations empêchent tout espoir d'obtenir des bornes uniformes en temps, par exemple en régularité analytique au sens usuel. Une idée maîtresse dans [85] consiste à se concentrer sur les modes de Fourier qui comptent dans la solution du transport libre, et donc à suivre la cascade au fur et à mesure que le temps passe. On appelle ce concept **régularité glissante**. La régularité glissante vient avec une dégradation des bornes de régularité en vitesse, mais en même temps avec une *amélioration de la régularité en position*, dès que l'on fait des moyennes en vitesse. Notre interprétation de l'amortissement Landau est donc un **transfert de régularité** de la variable v vers la variable x , la régularité de la force allant en s'améliorant, ce qui implique que son amplitude s'éteint.

La norme analytique utilisée dans [85] est assez complexe : elle a de bonnes propriétés algébriques qui permettent de suivre les erreurs obtenues par composition, elle s'adapte bien à la géométrie du problème, et suit le transport libre pour mesurer

la régularité glissante :

$$\|f\|_{\mathcal{Z}_\tau^{\lambda,(\mu,\gamma);p}} = \sum_{k \in \mathbb{Z}^d} \sum_{n \in \mathbb{N}^d} \frac{\lambda^n}{n!} e^{2\pi\mu|k|} (1 + |k|)^\gamma \left\| (\nabla_v + 2i\pi\tau k)^n \widehat{f}(k, v) \right\|_{L^p(dv)} \quad (44)$$

(Ici \widehat{f} désigne la transformée de Fourier dans la variable de position, pas de vitesse.) L'exposant λ quantifie la régularité analytique en vitesse, les exposants μ et γ (par défaut $\gamma = 0$) quantifient la régularité en position, et le paramètre τ est à prendre comme un décalage en temps. On renvoie à [85] pour une étude des propriétés remarquables de ce type de normes, et aussi pour des résultats de comparaison aux normes plus naïves dans lesquelles le théorème d'amortissement non linéaire est énoncé.

Le résultat principal de [85] consiste à prouver une borne uniforme du type

$$\|f(t, \cdot) - f^0\|_{\mathcal{Z}_t^{\lambda,\mu;1}} = O(\delta).$$

Cette borne implique l'amortissement Landau, mais contient aussi beaucoup plus d'informations : par exemple, elle montre que les fréquences spatiales élevées relaxent plus vite; elle entraîne aussi la stabilité orbitale non linéaire sous condition de Penrose, un problème qui avait jusqu'ici résisté à toutes les méthodes classiques.

7.4 Échos non linéaires et régularité critique

La célèbre expérience de l'écho plasma [77, 78] décrit l'interaction de deux ondes engendrées par des perturbations spatiales distinctes. Si l'on envoie une première perturbation au temps initial avec une fréquence k , il s'ensuit des oscillations à fréquence cinétique $|k|t$, oscillations qui ne s'atténuent pas au cours du temps et au contraire deviennent de plus en plus frénétique. Si maintenant au temps τ on fait intervenir une seconde perturbation avec fréquence ℓ , alors on engendrera des oscillations à la fréquence cinétique $|\ell|(t - \tau)$. Les deux trains d'oscillations seront invisibles l'un à l'autre, par effet de moyenne, sauf quand ils auront la même fréquence cinétique; ceci se produit à un temps t tel que $kt + \ell(t - \tau) = 0$, soit

$$t = \frac{\ell\tau}{k + \ell}; \quad (45)$$

il est sous-entendu ici que k et ℓ sont colinéaires et de direction opposée, avec $|\ell| > |k|$. En un certain sens, dans l'asymptotique en temps grand, la réaction à la seconde perturbation τ s'effectue à un temps t qui est strictement plus grand que t . Ce *retard* est capital pour expliquer la stabilité non linéaire de l'équation de Vlasov. Pour se donner une idée de ce gain, comparer l'inégalité $u(t) \leq A + \int_0^t \tau u(\tau) d\tau$, qui implique pour u une croissance essentiellement en $O(e^{t^2})$, à l'inégalité $u(t) \leq A + t u(t/2)$, qui implique une croissance très lente en $O(t^{\log t})$.

On peut proposer comme caricature des estimées pour l'équation de Vlasov-Poisson la famille d'inégalités

$$\varphi_k(t) \leq a(kt) + \frac{ct}{k^2} \varphi_{k+1} \left(\frac{kt}{k+1} \right).$$

Ici $\varphi_k(t)$ représente en gros la norme du mode k de la densité spatiale au temps t ; $a(kt)$ représente l'effet de la source (oublions le terme linéaire représenté par

une équation de Volterra), le coefficient t traduit le fait que le couplage se fait par l'intermédiaire du gradient de f en vitesses, et que ce gradient croît linéairement avec le temps; $1/k^2$ est la transformée de Fourier du potentiel d'interaction, on note à ce sujet que l'interaction entre modes est d'autant plus dangereuse que le potentiel est singulier; on n'a gardé que l'interaction entre le mode k et le mode $k + 1$; enfin l'argument du mode $k + 1$ n'est pas t mais $kt/(k + 1)$, ce qui représente un léger retard par rapport à t , comme dans la formule des échos. Une résolution explicite montre que

$$\varphi_k(t) \lesssim a(kt) \exp((ckt)^{1/3}).$$

Ces estimations peuvent être adaptées à l'équation originelle de Vlasov–Poisson; on trouve ainsi, dans la résolution de l'équation linéarisée autour d'une solution non stationnaire, une *perte de régularité/décroissance* qui est en exponentielle fractionnaire. Sous de bonnes hypothèses (aussi fortes que la condition de Penrose dans le cas gravitationnel, plus fortes dans le cas coulombien) on trouve effectivement $\exp((kt)^{1/3})$; dans des cas plus généraux la croissance reste comme une exponentielle fractionnaire en kt . Comme elle reste sous-exponentielle, cette perte de régularité peut être compensée par la décroissance exponentielle issue du problème linéaire.

La perte de régularité dépend essentiellement de l'interaction, alors que le gain linéaire dépend surtout de la régularité des données : exponentiel pour des données analytiques, polynomial pour des données C^r , exponentiel fractionnaire pour des données Gevrey. La discussion précédente suggère donc qu'il est possible d'étendre le théorème d'amortissement non linéaire à des données Gevrey. Par exemple, dans le cas gravitationnel, l'exposant critique $1/3$ correspond à une régularité critique Gevrey-3. On rappelle qu'une fonction est dite Gevrey- ν si ses dérivées successives ne croissent pas plus vite que $O(\nu!^\nu)$. Quitte à perdre arbitrairement peu sur ν , il est équivalent de demander que sa transformée de Fourier décroisse en exponentielle fractionnaire $\exp(c|\xi|^{1/\nu})$.

7.5 Spéculations

Le théorème d'amortissement Landau non linéaire ouvre un grand nombre de questions. D'abord son extension à d'autres géométries que \mathbb{T}^d est un vrai défi, car on perd alors la magique transformation de Fourier. L'extension à des équilibres inhomogènes est encore un rêve lointain; en fait la stabilité linéaire des ondes de Bernstein–Greene–Kruskal n'est toujours pas connue !

Ensuite, on a vu que l'on sait le traiter en régularité Gevrey, en revanche l'extension à des régularités plus basses telle que la régularité C^r est un problème ouvert. On a déjà souligné le parallèle avec la théorie KAM, dans laquelle on sait traiter des problèmes en classe C^r ; mais dans KAM, la perte de régularité est seulement polynomiale, ici elle est bien plus sévère. Certaines variantes du problème KAM mènent à une perte de régularité en exponentielle fractionnaire, et alors c'est également un problème ouvert de travailler en régularité plus basse que Gevrey. Dans l'immédiat, la seule avancée en régularité C^r suggérée par [85] est la possibilité de prouver l'amortissement sur des échelles de temps bien plus grandes que l'échelle non linéaire ($O(1/\delta)$ au lieu de $O(1/\sqrt{\delta})$), voir [85, Section 13]; ce développement semble dépendre d'une conjecture originale concernant les constantes optimales survenant dans certaines inégalités d'interpolation). On reparlera dans la section 8 de

stratégie permettant de contourner conceptuellement cette limitation de très haute régularité.

Quelle que soit la régularité optimale, il est exclu d'obtenir un amortissement Landau dans l'espace d'énergie naturel associé aux lois de conservations physiques. En effet, Lin et Zeng [72] montrent que l'amortissement Landau non linéaire est faux si l'on a strictement moins de deux dérivées, dans un sens approprié.

Enfin, même si l'amortissement Landau n'est qu'un phénomène perturbatif, il faut noter que son importance conceptuelle reste considérable parce que c'est à l'heure actuelle le seul îlot que nous parvenons à explorer dans l'océan des questions ouvertes ayant trait à la relaxation isentropique. Par sa découverte, Landau a fait prendre conscience que des systèmes physiques pouvaient relaxer sans qu'il y ait pour autant d'irréversibilité et d'augmentation d'entropie. Dans les années 1960, Lynden-Bell [75, 76] invoquait cette avancée conceptuelle pour résoudre le paradoxe de la relaxation des galaxies, qui apparaissent dans un état approximativement quasi-stationnaire alors que les temps de relaxation associés à l'équation de Vlasov galactique sont largement supérieurs à l'âge de l'univers. Depuis, le principe de la **relaxation violente**, relaxation du champ de forces sur quelques temps caractéristiques de la dynamique, est bien accepté par les astrophysiciens, sans que l'on ait aucune explication théorique à lui avancer. Il y a là un défi scientifique majeur.

8 Faible dissipation

Entre le modèle de Boltzmann qui fait la part belle aux collisions et celui de Vlasov qui les néglige complètement, nous trouvons un compromis particulièrement intéressant dans le modèle de Landau (ou Fokker–Planck–Landau), à faible dissipation :

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f + F[f] \cdot \nabla_v f = \varepsilon Q_L(f, f), \quad (46)$$

où Q_L est l'opérateur de Landau (21).

En physique des plasmas classique, le coefficient ε vaut $(\log \Lambda)/(2\pi\Lambda)$, où Λ est le paramètre plasma, d'ordinaire très grand (entre 10^2 et 10^{30}). Dans une approche particulière, le coefficient ε est une fluctuation par rapport à la limite de champ moyen, proportionnelle à $\log N/N$. Les effets entropiques irréversibles modélisés par l'opérateur de collision ne se font donc sentir notablement que sur de grands temps $O(1/\varepsilon)$. Par ailleurs, des effets régularisants se font sentir instantanément, même s'ils sont faibles. L'intérêt de l'étude est donc multiple :

- c'est un modèle physique plus réaliste que l'équation de Vlasov "pure", sans collisions;
- il permet de quantifier, en fonction du petit paramètre ε , les vitesses relatives des phénomènes d'homogénéisation (amortissement Landau) et de convergence entropique;
- il permet de contourner l'obstacle de la régularité Gevrey auquel se heurte l'étude du modèle sans collision.

Tout reste à faire en la matière et je vais me contenter d'esquisser un programme de longue haleine.

8.1 Scénario plausible

Partant d'une perturbation d'équilibre homogène à décroissance en vitesses très rapide, on doit, au cours de l'évolution temporelle par (46), rester proche du régime homogène; ceci est dans l'esprit de résultats d'Arkeryd, Esposito et Pulvirenti [5] sur l'équation de Boltzmann faiblement inhomogène. Dans le cadre homogène, l'opérateur du membre de droite a sans doute les mêmes propriétés régularisantes qu'un laplacien en vitesses, du moins localement (les propriétés de régularisation deviennent très faibles aux grandes vitesses, mais on a imposé une décroissance très forte en vitesses). Admettant que cela reste vrai dans un cadre faiblement inhomogène, on se retrouve avec une équation hypoelliptique, qui va régulariser dans toutes les variables, plus vite bien sûr dans la variable de vitesse que dans la variable de position.

La **régularisation hypoelliptique en classe Gevrey** n'a été que très peu étudiée, mais par des arguments dimensionnels on peut penser que dans ce contexte il y a régularisation en classe Gevrey- $1/\alpha$, avec vitesse $O(\exp((\varepsilon t)^{-\alpha/(2-\alpha)}))$ en v , et $O(\exp((\varepsilon t)^{-3\alpha/(2-3\alpha)}))$ en x .

D'un autre côté, dans la classe Gevrey- $1/\alpha$, pour $\alpha > 1/3$ on doit avoir décroissance vers le régime homogène en $O(\exp -t^\alpha)$.

En combinant les deux effets, on obtiendra l'homogénéisation sur une échelle de temps $O(\varepsilon^{-\zeta})$, avec $\zeta < 1$, ce qui est un taux plus rapide que le taux d'augmentation de l'entropie en $O(\varepsilon^{-1})$.

Quand on fait les comptes, le coefficient ζ que l'on peut espérer est décevant, de l'ordre de $8/9$. Parmi les étapes utilisées, le maillon le plus faible semble être la régularisation Gevrey en x , qui est extrêmement coûteuse et peut-être pas optimale puisque cette régularité n'est pas nécessaire dans l'analyse linéaire. Ceci motive le développement d'une version du théorème d'amortissement non linéaire en basse régularité en x . Si l'on se passe de cette régularité, le coefficient devient bien meilleur, de l'ordre de $1/6$...

8.2 Réexamen des simulations

Avec cette interprétation, nous pouvons maintenant revenir sur la figure 6.8 : l'usage d'une petite boîte spatiale renforce l'effet de l'opérateur $v \cdot \nabla_x$ au détriment de l'opérateur de collision, de sorte que l'on est dans un régime faiblement dissipatif. (Le champ de force est nul.) Alors en temps grand l'homogénéisation intervient plus rapidement que la relaxation entropique. Ceci n'explique pas tout, pour deux raisons : d'abord, dans cette figure la condition initiale est fortement (et non faiblement) inhomogène; ensuite l'opérateur de Boltzmann ne régularise pas. Cependant on veut bien croire que c'est l'homogénéisation par amortissement Landau qui se manifeste en premier dans cette figure, avant que les collisions ne fassent leur travail d'augmentation de l'entropie. (Comment décrire la sortie temporaire du régime homogène, semble un mystère.)

9 Métastatistiques

J'utilise ici le mot "métastatistiques" pour parler de statistiques sur la fonction de distribution, qui elle-même a un contenu statistique. Cette section sera courte car nous n'avons guère que des spéculations en la matière.

Le théorème de Hewitt–Savage, un avatar du théorème de Krein–Milman, décrit les mesures de probabilité symétriques en un grand nombre de variables comme des combinaisons convexes de mesures chaotiques :

$$\mu^N = \int_{P(\mathcal{Y})} \mu^{\otimes N} \Pi(d\mu),$$

où Π est une mesure de probabilité sur $P(\mathcal{Y})$, l'espace des mesures de probabilité sur l'espace macroscopique. En somme, une incertitude microscopique peut se décomposer en deux niveaux : outre le chaos à profil macroscopique fixé, il y a l'incertitude sur le profil macroscopique, c'est-à-dire le choix du profil μ , qui s'effectue avec la mesure de probabilité Π .

Maintenant y a-t-il une mesure de probabilité Π naturelle sur l'espace des profils admissibles ? Idéalement, une telle mesure serait **invariante par la dynamique**. Dans le cadre de l'équation de Boltzmann, la question ne se pose pas vraiment : seules restent en lice des mesures triviales portées par des équilibres maxwelliens. En revanche, dans le cadre de l'équation de Vlasov, la construction de mesures invariantes non triviales est un problème passionnant. De telles mesures refléteraient la nature hamiltonienne de l'équation de Vlasov, étudiée pour des interactions simplifiées par Ambrosio et Gangbo [4].

Un candidat assez sérieux au statut de mesure invariante est la **mesure entropique** de Sturm [97], issue de la théorie du transport optimal, formellement de la forme $\mathbb{P} = e^{-\beta H_\nu}$. Sa définition complexe a empêché jusqu'ici qu'on réussisse à prouver son invariance. Il ne doit pas être très difficile de modifier la construction pour ajouter un terme d'énergie. La mesure de Sturm est définie sur un espace compact, et il y a peut-être des subtilités à l'étendre dans un contexte cinétique où l'espace des vitesses est non borné. Mais la pire difficulté vient sans doute de la singularité des mesures typiques : on s'attend à ce que \mathbb{P} -presque toute mesure soit totalement étrangère à la mesure de Lebesgue, et portée par un ensemble de codimension 1. Ceci semble fermer la porte à toute étude statistique d'amortissement basée sur la régularité, et accroît le mystère.

Robert [91] et d'autres ont tenté de faire une théorie statistique de l'équation de Vlasov, en partant de la notion d'entropie, essayant de prédire l'état asymptotique *probable* de l'évolution dynamique. La théorie a remporté quelques succès, cependant elle reste controversée. En outre, comme l'état asymptotique est obtenu par une limite faible, la question se pose de savoir s'il faut imposer une égalité ou une inégalité sur les contraintes faisant intervenir des fonctionnelles non linéaires de la densité. À ce sujet on pourra consulter [102].

Ensuite, cette théorie ne tient presque pas compte de l'équation d'évolution sous-jacente, postulant une certaine universalité par rapport à l'interaction. Isichenko [63] a fait remarquer que l'état asymptotique en temps grand, s'il existe, doit dépendre de détails fins de la distribution initiale et de l'interaction, alors que les mesures construites par la théorie statistique ne dépendent que d'invariants : énergie, entropie, voire d'autres fonctionnelles de la forme $\iint A(f) dx dv$. Cette objection a trouvé de la substance avec les contre-exemples construits dans [85, Section 14], qui montrent que la transformation $f(x, v) \rightarrow f(x, -v)$ peut modifier l'état asymptotique final, alors qu'elle préserve tous les invariants connus de la dynamique. L'objection est peut-être surmontable, car ces contre-exemples sont construits en régularité analytique, c'est-à-dire dans une classe qui doit être invisible à un traite-

ment statistique; mais ces contre-exemples montrent la subtilité du problème, et renforcent le sentiment de difficulté de construction de mesures invariantes.

10 Paradoxes perdus

Dans cette dernière section je vais passer en revue une série de paradoxes plus ou moins célèbres associés à la flèche du temps et aux équations cinétiques, et présenter la résolution communément admise de ces paradoxes. Un certain nombre d'entre eux font intervenir l'infini, source classique de paradoxes comme l'"hôtel Hilbert" disposant d'un nombre infini de chambres, où l'on peut toujours loger un nouvel arrivant même s'il est déjà complet. À notre échelle, ce paradoxe reflète notre incapacité de nous rendre compte de l'apparition ou disparition d'une particule par rapport à la quantité gigantesque qui composent notre univers. La limite $N \rightarrow \infty$ (ou l'asymptotique $N \gg 1$, si comme Boltzmann on préfère éviter de manipuler les infinis) étant le fondement de la mécanique statistique, il n'est pas étonnant que ce paradoxe se manifeste.

Dans toute la suite, quand on parle de temps positif ou négatif, ou de configuration pré-collisionnelle ou post-collisionnelle, on fait référence au temps microscopique absolu des équations de Newton.

10.1 Paradoxe de Poincaré–Zermelo

Poincaré [89] en 1895 met en doute la théorie de Boltzmann au motif qu'elle semble contredire les propriétés fondamentales des systèmes dynamiques. Zermelo [112] peu après développe ce point et note que l'augmentation inexorable de l'entropie interdit le retour du système à l'état initial, qui est pourtant prédit par le théorème de récurrence (à une erreur arbitrairement petite près).

La même objection peut s'appliquer au problème de l'amortissement Landau : si la distribution tend vers un équilibre homogène, elle ne reviendra jamais près de son état initial.

Du point de vue mathématique, ce raisonnement ne s'applique évidemment pas, puisque l'équation de Boltzmann comme celle de Landau font intervenir un nombre infini de degrés de liberté; ce n'est que pour un nombre de particules fixé que le théorème de récurrence s'applique. Du point de vue physique, la réponse est un peu plus subtile. D'une part, le temps de récurrence diverge quand le nombre N de particules tend vers l'infini, et cette divergence est probablement monstrueusement rapide ! Pour un système de taille macroscopique, même petit, le théorème de récurrence ne s'applique simplement jamais, il met en oeuvre des temps bien plus grands que l'âge de l'univers. D'autre part, la validité des équations de Boltzmann ou de Landau n'est pas éternelle : à N fixé, la qualité de l'approximation va se dégrader avec le temps, car le chaos (simple ou pré-collisionnel) n'est préservé qu'approximativement. Quand le temps de récurrence de Poincaré aura lieu, l'équation de Boltzmann aura cessé d'être valable depuis bien longtemps !!

10.2 Conservation microscopique du volume

Le théorème de récurrence de Poincaré est basé sur la conservation du volume dans l'espace des phases microscopiques (préservation de la mesure de Liouville).

L'entropie est directement fonction du volume des états microscopiques admissibles, comment peut-elle augmenter si le volume des états microscopiques est constant ?

La réponse à cette question peut sembler surprenante : on peut argumenter que l'augmentation de l'entropie n'a pas lieu *malgré* la conservation du volume microscopique, mais *à cause* de cette conservation; plus précisément, c'est elle qui empêche l'entropie de diminuer. En effet, partons au temps initial de toutes les configurations typiques associées à une distribution f_i . Après un temps t , ces configurations typiques ont évolué et sont maintenant associées à une distribution f_t , la transition entre f_i et f_t étant régie par l'équation de Boltzmann. Les configurations typiques associées à f_t sont donc *au moins aussi nombreuses* que les configurations typiques associées à f_i , ce qui veut bien dire que l'entropie ne peut diminuer.

Dans un modèle microscopique irréversible, on aura typiquement une contraction de l'espace des phases microscopique, lié à un phénomène dissipatif. L'argument précédent ne s'applique alors plus, et on peut imaginer que l'entropie diminue, au moins pour certaines données initiales. C'est de fait ce qui se produit par exemple dans les modèles de gaz granulaires subissant des collisions inélastiques.

10.3 Apparition spontanée de la flèche du temps

Comment, partant d'une équation microscopique qui ne privilégie aucune direction du temps, l'équation de Boltzmann peut-elle prédire une évolution inexorable vers les temps positifs ?

La réponse est simple : il n'y a pas d'évolution inexorable vers les temps positifs, et la double direction du temps est préservée. Simplement, il y a eu un choix particulier de la donnée initiale (instant de préparation de l'expérience), qui a fixé un temps particulier, disons $t = 0$. À partir de là, on a une double flèche du temps, l'entropie augmente pour les temps positifs, et diminue pour les temps négatifs.

10.4 Paradoxe de Loschmidt

Le paradoxe de Loschmidt [74] formalise l'apparente contradiction qu'il y a dans la coexistence d'une dynamique microscopique réversible et d'une évolution irréversible de l'entropie. Supposons que l'on parte d'une configuration initiale donnée, et qu'au temps t on arrête le gaz, et on renverse les vitesses de toutes les particules. Cette opération ne change pas l'entropie, et partant de cette nouvelle donnée initiale on peut laisser la dynamique agir à nouveau. Par réversibilité microscopique, au bout d'un temps $2t$ on sera revenu au point de départ; mais l'entropie n'aura cessé d'augmenter, d'où contradiction.

On peut résoudre ce paradoxe de plusieurs manières, qui reviennent toutes à la même constatation : la *dégradation de la notion de chaos* entre le temps initial et le temps $t > 0$. Au niveau mathématique, on ne sait prouver que la convergence faible de $\mu_t^{1:N}$ vers $f(t, \cdot)$ quand $N \rightarrow \infty$, alors que la convergence est supposée uniforme au temps initial. En fait, on conjecture que la donnée (μ_t^N) vérifie la propriété (encore à définir...) de *chaos pré-collisionnel*, alors que la donnée initiale était supposée vérifier une propriété de chaos complète. Quand on inverse les vitesses, on transforme l'hypothèse de chaos pré-collisionnel en chaos post-collisionnel, et l'équation pertinente n'est plus l'équation de Boltzmann, mais l'équation de Boltzmann "inverse", dans laquelle on a mis un signe négatif devant l'opérateur de colli-

sion. L'entropie croît alors vers les temps négatifs et non plus vers les temps positifs, et toute contradiction disparaît.

Pour dire les choses de manière plus informelle : au temps initial, les particules sont toutes étrangères les unes aux autres. Après un temps t , les particules qui viennent de se rencontrer se connaissent encore, celles qui vont se rencontrer ne se connaissent pas : les particules ont une mémoire du passé et pas du futur. Quand on inverse les vitesses, les particules ont une mémoire du futur et pas du passé, et le temps se met à s'écouler à l'envers !

La légende dit que Boltzmann, confronté au paradoxe du renversement des vitesses, a répondu "Vas-y, renverse-les !" Derrière la boutade, il y a une observation profonde : le renversement des vitesses est une opération qui nous est inaccessible car elle nécessite une connaissance microscopique du système; et précisément la notion d'entropie émerge de ce que nous ne pouvons agir que macroscopiquement dessus. À partir des années 1950, les expériences d'écho de spin permettaient de voir le paradoxe sous un autre angle [10].

10.5 Non-validité universelle de l'équation de Boltzmann

Ce paradoxe est une variante du précédent. Ayant compris que l'équation de Boltzmann ne s'applique pas après renversement des vitesses, on va exploiter ce fait pour mettre Monsieur Boltzmann en défaut. On refait l'expérience précédente et on choisit comme donnée initiale la distribution obtenue après renversement des vitesses au temps t . On laisse alors le temps agir, et l'équation pertinente n'est bien sûr pas l'équation de Boltzmann.

Ce paradoxe montre effectivement qu'il y a des configurations microscopiques qui *ne mènent pas* à l'équation de Boltzmann. Cependant, et c'est ainsi que Boltzmann argumenta, ces configurations sont rares : précisément, elles font apparaître des corrélations entre vitesses pré-collisionnelles. Ceci n'est pas plus rare que des corrélations entre vitesses post-collisionnelles, mais c'est plus rare que de n'avoir pas de corrélations du tout ! L'équation de Boltzmann est approximativement vraie si l'on part d'une configuration *typique*, c'est-à-dire tirée selon une loi microscopique "fortement chaotique", mais elle n'est pas vérifiée pour *toutes* les configurations initiales. Une fois ces grands principes énoncés, le travail quantitatif reste à faire.

10.6 Arbitraire de la procédure de Boltzmann

Pour établir l'équation de Boltzmann, on exprime les probabilités de rencontre de particules en fonction des probabilités pré-collisionnelles, ce qui est arbitraire. Si l'on avait utilisé à la place les probabilités post-conditionnelles, on aurait obtenu une équation différente, avec un signe négatif devant l'opérateur de collision ! Pourquoi alors faire confiance à Boltzmann ?

La réponse est encore la même, bien sûr, et dépend du côté de l'origine où l'on se place : pour les temps positifs, ce sont les probabilités pré-collisionnelles qui sont presque factorisées, alors que pour les temps positifs, ce sont les probabilités post-collisionnelles.

10.7 Démon de Maxwell

Maxwell imagina une expérience de pensée dans laquelle un démon malicieux s'installe dans une boîte à deux compartiments, et manie adroitement un petit clapet de manière à ce qu'il y ait un flux de boules allant du compartiment de droite vers le compartiment de gauche, et pas l'inverse. Ainsi le système évolue vers plus d'ordre, et l'entropie diminue.

Bien sûr, ceci ne peut être considéré comme une objection à la loi de croissance de l'entropie, et l'expérience est destinée à nous faire réfléchir : d'abord le démon devrait faire partie du modèle, et être lui-même soumis à des lois mécaniques réversibles, tenant compte de l'énergie qu'il faut mettre en jeu pour reconnaître qu'une particule s'approche et évaluer sa vitesse, du travail cérébral qu'il effectue, etc. Si l'on fait le bilan complet, on retrouvera, soyons-en sûrs, la seconde loi de la thermodynamique.

Notons à ce sujet que récemment, des expériences de démon de Maxwell ont pu être réalisés avec des *gaz granulaires* : comme je l'ai moi-même vu avec stupéfaction sur un film expérimental, on part d'un récipient avec deux compartiments séparés verticalement, et une ouverture en haut permettant la communication, on remplit les deux compartiments de particules inélastiques en nombre à peu près égal, on agite le tout automatiquement, et peu à peu l'un des compartiments se vide au profit de l'autre. Un principe sous-jacent est que dans le compartiment le plus plein, l'abondance des collisions entraîne par dissipation d'énergie un refroidissement, et les particules sautent moins haut, ce qui rend de plus en plus difficile leur évation du compartiment plein. On retrouve à cette occasion le principe déjà mentionné selon lequel une dynamique microscopique dissipative (irréversible) ne mène pas forcément à une augmentation de l'entropie, bien au contraire.

10.8 Convergence et réversibilité

Ce paradoxe est une variante du paradoxe de Loschmidt; il s'applique aussi bien à la thématique de la relaxation entropique de Boltzmann, qu'à l'amortissement de Landau non linéaire : comment peut-on avoir la convergence quand $t \rightarrow +\infty$ si on a réversibilité de la dynamique ? ? La réponse est d'une simplicité enfantine : il y a aussi convergence quand $t \rightarrow -\infty$. Pour Vlasov, ceci se fait avec la même équation, et on a donc un phénomène d'homoclinie/hétéroclinie généralisé. Pour Boltzmann, l'équation change selon que l'on considère des temps qui sont antérieurs ou postérieurs à la donnée chaotique.

10.9 Stabilité et réversibilité

Ce paradoxe est plus subtil et s'applique à l'amortissement Landau non linéaire : stabilité asymptotique et réversibilité de la dynamique entraînent automatiquement une instabilité, ce qui semble contradictoire.

Détaillons l'argument. Si l'on a stabilité en temps $t \rightarrow +\infty$, soit $\bar{f}_\infty(v)$ un profil asymptotique stable, que l'on supposera pair. Prenons une solution $\bar{f}(t, x, v)$, inhomogène, qui converge vers $f_\infty(v)$. Choisissons alors pour donnée initiale $f(T, x, -v)$ avec T très grand, on aura ainsi une donnée arbitrairement proche de $f_\infty(-v) = f_\infty(v)$, et qui nous ramène au bout d'un temps T à la donnée $\bar{f}(0, x, v)$, assez éloignée

de $f_\infty(v)$. Autrement dit, la distribution f_∞ est *instable*. Comment cela est-il compatible avec la stabilité ? ?

La réponse, comme expliquée par exemple dans [25], tient dans la topologie : dans le théorème de stabilité asymptotique (amortissement Landau non linéaire), la convergence en temps grand a lieu *au sens de la topologie faible*, avec des oscillations frénétiques dans la distribution de vitesses, qui se compensent localement. Quand on dit qu'une distribution f^0 est stable, cela veut dire que si l'on part près de f^0 au sens de la topologie forte (par exemple analytique ou Gevrey), alors on restera près de f^0 au sens de la topologie faible. La stabilité asymptotique combinée avec la réversibilité impliquent donc l'*instabilité au sens de la topologie faible*, ce qui est parfaitement compatible avec la stabilité au sens de la topologie forte.

10.10 Relaxation conservative

Ce problème est de nature assez générale. L'équation de Vlasov vient avec une préservation de la quantité d'incertitude microscopique (conservation de l'entropie). En outre la distribution au temps $t > 0$ permet de reconstruire exactement la distribution au temps $t = 0$: il suffit de résoudre l'équation de Vlasov après renversement des vitesses. On peut dire que l'équation de Vlasov n'oublie rien; or la convergence consiste précisément à oublier les périétés de l'évolution dynamique !

La réponse tient encore dans la convergence faible et les oscillations. Une information va se loger dans ces oscillations, information qui est *invisible* car en pratique nous ne mesurons jamais la fonction de distribution complète, mais des moyennes de cette fonction de distribution (rappelons-nous la citation de Lynden-Bell reproduite dans la fin de la section 7.2). Chaque observable va converger vers sa valeur limite, et il y aura un "oubli". Le champ de force, obtenu comme moyenne de la distribution cinétique, converge vers 0 sans que cela soit contradictoire avec la préservation de l'information : l'information quitte les variables spatiales pour aller dans les variables cinétiques. En particulier, l'entropie spatiale $\int \rho \log \rho$ (où $\rho = \int f dv$) tend vers 0, alors que l'entropie cinétique totale $\int f \log f$ est conservée (mais ne converge pas ! l'information est conservée pour tout temps, mais à cause de la convergence faible il y a une perte d'information dans le passage à la limite $t \rightarrow \infty$).

De même, dans l'amortissement Landau non linéaire, l'énergie d'interaction $\int W(x-y) \rho(x) \rho(y) dx dy$ tend vers 0, et elle est convertie en énergie cinétique (qui peut croître ou décroître en fonction de l'interaction).

10.11 Expérience des échos

Dans cette célèbre expérience [77, 78], on prépare un plasma en état d'équilibre, et on l'excite au temps initial par une impulsion de fréquence spatiale k . Au bout d'un temps τ , après relaxation du plasma, on l'excite à nouveau avec une fréquence spatiale ℓ , colinéaire et de direction opposée à k , d'amplitude plus grande. On attend ensuite et on observe une réponse spontanée du champ électrique du plasma, appelée **écho**, qui se produit à la fréquence spatiale $k+\ell$ et autour du temps $t_e = (|\ell|/|k+\ell|)\tau$.

Cette expérience montre que la distribution cinétique du plasma a gardé trace des impulsions passées : même si le champ de force s'est amorti jusqu'à devenir négligeable, les oscillations cinétiques de la distribution restent présentes, et évoluent au cours du temps. La première impulsion subsiste sous la forme d'oscillations très rapides de période $(|k|t)^{-1}$, la seconde sous forme d'oscillations de période $(|\ell|(t -$

$\tau))^{-1}$. Un calcul que l'on trouvera par exemple dans [109, Section 7.3] montre que la distribution continue à osciller rapidement en vitesse, et la force associée reste négligeable, jusqu'à ce que les deux trains d'oscillations se compensent presque exactement, ce qui se manifeste par l'écho.

Références

- [1] Alexandre, R., Desvillettes, L., Villani, C., et Wennberg, B. : Entropy dissipation and long-range interactions. *Arch. Rat. Mech. Anal.* **152**, 327–355 (2000).
- [2] Alexandre, R., et Villani, C. : On the Boltzmann equation for long-range interaction and the Landau approximation in plasma physics. *Comm. Pure Appl. Math.* **55**, 1, 30–70 (2002).
- [3] Alexandre, R., et Villani, C. : On the Landau approximation in plasma physics. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **21**, 1, 61–95 (2004).
- [4] Ambrosio, L., et Gangbo, W. : Hamiltonian ODE's in the Wasserstein space of probability measures. *Comm. Pure Appl. Math.* **51**, 18–53 (2007).
- [5] Arkeryd, L., Esposito, R., et Pulvirenti, M. : The Boltzmann equation for weakly inhomogeneous data. *Comm. Math. Phys.* **111**, 3, 393–407 (1987).
- [6] Backus, G. : Linearized plasma oscillations in arbitrary electron distributions. *J. Math. Phys.* **1**, 178–191, 559 (1960).
- [7] Balescu, R. : Irreversible processes in ionized gases. *Phys. Fluids* **3**, 52–63 (1960).
- [8] Balescu, R. *Statistical Mechanics of Charged Particles*. Wiley-Interscience, New York, 1963.
- [9] Balian, R. : Entropy, a protean concept. *Poincaré Seminar* **2**, 119–145 (2003). Birkhäuser Verlag, Basel.
- [10] Balian, R. : Information in statistical physics. *Studies in Hist. and Phil. of Mod. Phys.* **36**, 323–353 (2005).
- [11] Baranger, C., et Mouhot, C. : Explicit spectral gap estimates for the linearized Boltzmann and Landau operators with hard potentials. *Rev. Matem. Iberoam.* **21**, 819–841 (2005).
- [12] Batt, J., et Rein, G. : Global classical solutions of the periodic Vlasov-Poisson system in three dimensions. *C. R. Acad. Sci. Paris Sér. I Math.* **313**, 6, 411–416 (1991).
- [13] Ben Arous, G., et Zeitouni, O. : Increasing propagation of chaos for mean field models. *Ann. Inst. H. Poincaré Probab. Statist.* **35**, 1, 85–102 (1999).
- [14] Bernstein, I.B., Greene, J.M., et Kruskal, M.D. : Exact nonlinear plasma oscillations. *Phys. Rev.* **108**, 3, 546–550 (1957).
- [15] Binney, J., et Tremaine, S. : *Galactic Dynamics*, second ed. Princeton Series in Astrophysics. Princeton University Press, 2008.

- [16] Bobylev, A.V., et Cercignani, C. : On the rate of entropy production for the Boltzmann equation. *J. Statist. Phys.* **94**, 3-4, 603–618 (1999).
- [17] Bobylev, A.V., et Cercignani, C. : Exact eternal solutions of the Boltzmann equation. *J. Statist. Phys.* **106**, 5–6, 1019–1038 (2002).
- [18] Bobylev, A.V., et Cercignani, C. : Self-similar solutions of the Boltzmann equation and their applications. *J. Statist. Phys.* **106**, 5–6, 1039–1071 (2002).
- [19] Bogoliubov, N.N. : Problems of a dynamical theory in statistical physics. *Studies in Statistical mechanics* **1**, 1–118 (1962). Traduction de l’original russe de 1946.
- [20] Bolley, F., Guillin, A., et Villani, C. : Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab. Theory Related Fields* **137**, 3-4, 541–593 (2007).
- [21] Boltzmann, L. : Weitere Studien über das Wärme gleichgewicht unter Gas-moläkulen. *Sitzungsberichte Akad. Wiss.* **66** (1872), 275–370. Traduction : Further studies on the thermal equilibrium of gas molecules, in *Kinetic Theory* **2**, 88–174, Éd. S.G. Brush, Pergamon, Oxford (1966).
- [22] Boltzmann, L. : *Lectures on gas theory*. University of California Press, Berkeley, 1964. Traduit en anglais par Stephen G. Brush. Réimpression de l’édition de 1896–1898. Dover Publications, 1995.
- [23] Bouchut, F. : Introduction à la théorie mathématique des équations cinétiques. In *Session “L’Etat de la Recherche” de la SMF* (1998). “Kinetic equations and asymptotic theory”, F. Bouchut, F. Golse, M. Pulvirenti, coll. “Series in Appl. Math.”, Gauthier-Villars, 2000.
- [24] Braun, W., et Hepp, K. : The Vlasov dynamics and its fluctuations in the $1/N$ limit of interacting classical particles. *Commun. Math. Phys.* **56**, 125–146 (1977).
- [25] Caglioti, E., et Maffei, C. : Time asymptotics for solutions of Vlasov–Poisson equation in a circle. *J. Statist. Phys.* **92**, 1–2, 301–323 (1998).
- [26] Carleman, T. : Sur la théorie de l’équation intégrodifférentielle de Boltzmann. *Acta Math.* **60**, 369–424 (1932).
- [27] Carleman, T. : *Problèmes Mathématiques dans la Théorie Cinétique des Gaz*. Almqvist & Wiksell, 1957.
- [28] Carlen, E.A., et Carvalho, M.C. : Strict entropy production bounds and stability of the rate of convergence to equilibrium for the Boltzmann equation. *J. Statist. Phys.* **67**, 3–4, 575–608 (1992).
- [29] Carlen, E.A., Carvalho, M. C. et Gabetta, E. : Central limit theorem for Maxwellian molecules and truncation of the Wild expansion. *Comm. Pure Appl. Math.* **53**, 370–397 (2000).
- [30] Carlen, E.A., Carvalho, M. C. et Gabetta, E. : On the relation between rates of relaxation and convergence of Wild sums for solutions of the Kac equation. *J. Funct. Anal.* **220**, 362–387 (2005).

- [31] Carlen, E.A., Carvalho, M.C. et Loss, M. : Determination of the spectral gap for Kac's master equation and related stochastic evolution. *Acta Math.* **191**, 1–54 (2003).
- [32] Carlen, E.A., Carvalho, M.C., Le Roux, J., Loss, M., et Villani, C. Entropy and chaos in the Kac model. *Kinet. Relat. Models* **3**, 1, 85–122 (2010).
- [33] Carlen, E.A., Gabetta, E., et Toscani, G. : Propagation of smoothness and the rate of exponential convergence to equilibrium for a spatially homogeneous Maxwellian gas. *Comm. Math. Phys.* **199**, 3, 521–546 (1999).
- [34] Carlen, E.A., et Lu, M. : Fast and slow convergence to equilibrium Maxwellian molecules via Wild sums *J. Stat. Phys.* **112**, 1–2, 59–134 (2003).
- [35] Carleson, L. : Some analytic problems related to statistical mechanics. In *Euclidean Harmonic Analysis*, J.J. Benedetto, Éd. (Univ. of Maryland, 1979), *Lect. Notes Math.* **779**, 5–45.
- [36] Cercignani, C. : On the Boltzmann equation with cutoff potentials. *Phys. Fluids* **10**, 2097 (1967).
- [37] Cercignani, C. : On the Boltzmann equation for rigid spheres. *Transport Theory Statist. Phys.* **2**, 3, 211–225 (1972).
- [38] Cercignani, C. : *The Boltzmann equation and its applications*. Springer-Verlag, New York, 1988.
- [39] Cercignani, C. : *Rarefied gas dynamics*. Cambridge University Press, Cambridge, 2000. From basic concepts to actual calculations.
- [40] Cercignani, C., Illner, R., et Pulvirenti, M. : *The Mathematical Theory of Dilute Gases*. Springer, New York, 1994.
- [41] Cover, T. M., et Thomas, J. A. : *Elements of information theory*. John Wiley & Sons Inc., New York, 1991. A Wiley-Interscience Publication.
- [42] Delcroix, J.-L., et Bers, A. : *Physique des plasmas* (2 Vol.). InterEditions/CNRS Éditions, 1994.
- [43] Dembo, A., et Zeitouni, O. : *Large deviations techniques and applications*, 2nde éd. Springer-Verlag, New York, 1998.
- [44] Desvillettes, L. : Entropy dissipation rate and convergence in kinetic equations. *Comm. Math. Phys.* **123**, 4, 687–702 (1989).
- [45] Desvillettes, L., et Villani, C. : On the spatially homogeneous Landau equation for hard potentials. Part I : existence, uniqueness and smoothness. *Comm. Partial Differential Equations* **25**, 1-2, 179–259 (2000).
- [46] Desvillettes, L., et Villani, C. : On the trend to global equilibrium for spatially inhomogeneous kinetic systems : the Boltzmann equation. *Invent. Math.* **159**, 2, 245–316 (2005).
- [47] DiPerna, R., et Lions, P.-L. : On the Cauchy problem for the Boltzmann equation : Global existence and weak stability. *Ann. of Math.* (2) **130**, 312–366 (1989).

- [48] Dobrušin, R.L. : Vlasov equations. *Funktsional. Anal. i Prilozhen.* **13**, 2, 48–58, 96 (1979).
- [49] Dudley, R. M. : *Real analysis and probability*, vol. 74 de *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. Réédition révisée de l’original de 1989.
- [50] Filbet, F., Mouhot, C., et Pareschi, L. : Solving the Boltzmann equation in $N \log_2 N$. *SIAM J. Sci. Comput.* **28**, 3, 1029–1053 (2006).
- [51] Glassey, R.T. : *The Cauchy problem in kinetic theory*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996.
- [52] Glassey, R., et Schaeffer, J. : Time decay for solutions to the linearized Vlasov equation. *Transport Theory Statist. Phys.* **23**, 4, 411–453 (1994).
- [53] Glassey, R., et Schaeffer, J. : On time decay rates in Landau damping. *Comm. Partial Differential Equations* **20**, 3–4, 647–676 (1995).
- [54] Golse, F. : From kinetic to macroscopic models. In *Session “L’Etat de la Recherche” de la SMF* (1998), “Kinetic equations and asymptotic theory”, F. Bouchut, F. Golse, M. Pulvirenti, coll. “Series in Appl. Math.”, Gauthier-Villars, 2000.
- [55] Golse, F., et Saint-Raymond, L. : The Navier–Stokes limit of the Boltzmann equation for bounded collision kernels. *Invent. Math.* **155**, 1, 81–161 (2004).
- [56] Grad, H. : On Boltzmann’s H -theorem. *J. Soc. Indust. Appl. Math.* **13**, 1, 259–277 (1965).
- [57] Grad, H. : Principles of the kinetic theory of gases. In *Flügge’s Handbuch des Physik*, vol. XII. Springer-Verlag, 1958, pp. 205–294.
- [58] Gualdani, M.P., Mischler, S. et Mouhot, C. : Factorization for non-symmetric operators and exponential H -theorem. Preprint, available online at <http://hal.archives-ouvertes.fr/ccsd-00495786>
- [59] Guo, Y. : The Landau equation in a periodic box. *Commun. Math. Phys.* **231**, 391–434 (2002).
- [60] Guo, Y., et Strain, R. M. : Exponential decay for soft potentials near Maxwellian. *Arch. Ration. Mech. Anal.* **187**, 2, 287–339 (2008).
- [61] Illner, R., et Pulvirenti, M. : Global validity of the Boltzmann equation for a two-dimensional rare gas in vacuum. *Comm. Math. Phys.* **105**, 2, 189–203 (1986). “Erratum and improved result”, *Comm. Math. Phys.* **121**, 1, 143–146 (1989).
- [62] Illner, R., et Pulvirenti, M. : A derivation of the BBGKY-hierarchy for hard sphere particle systems. *Transport Theory Statist. Phys.* **16**, 7, 997–1012 (1987).
- [63] Isichenko, M. : Nonlinear Landau damping in collisionless plasma and inviscid fluid. *Phys. Rev. Lett.* **78**, 12, 2369–2372 (1997).

- [64] Jabin, P.-E., et Hauray, M. : N -particles approximation of the Vlasov equations with singular potential. *Arch. Ration. Mech. Anal.* **183**, 3, 489–524 (2007).
- [65] Janvresse, E. : Spectral gap for Kac’s model of Boltzmann equation. *Ann. Probab.* **29**, 1, 288–304 (2001).
- [66] Kac, M. : Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III* (Berkeley and Los Angeles, 1956), University of California Press, pp. 171–197.
- [67] Landau, L.D. : Die kinetische Gleichung für den Fall Coulombscher Wechselwirkung. *Phys. Z. Sowjet* **10** (1936), 154. English translation : The transport equation in the case of Coulomb interactions, in *Collected papers of L.D. Landau*, edited and with an introduction by D. ter Haar, Pergamon Press, 1965, pp.163–170.
- [68] Landau, L.D. : On the vibration of the electronic plasma. *J. Phys. USSR* **10**, 25 (1946). Traduction en anglais dans *JETP* **16**, 574. Reproduit dans *Collected papers of L.D. Landau*, édité par et avec une introduction de D. ter Haar, Pergamon Press, 1965, pp. 445–460 ; et dans *Men of Physics : L.D. Landau*, Vol. 2, Pergamon Press, D. ter Haar, ed. (1965).
- [69] Lanford, O.E. : Time evolution of large classical systems. In *Dynamical systems, theory and applications (Recontres, Battelle Res. Inst., Seattle, Wash., 1974)*. Springer, Berlin, 1975, pp. 1–111. *Lecture Notes in Phys.*, Vol. 38.
- [70] Lenard, A. : On Bogoliubov’s kinetic equation for a spatially homogeneous plasma. *Ann. Physics* **10**, 390–400 (1960).
- [71] Lifshitz, E. M., et Pitaevskiĭ, L. P. : *Course of theoretical physics [“Landau-Lifshits”]. Vol. 10*. Pergamon Press, Oxford, 1981. Translated from the Russian by J. B. Sykes and R. N. Franklin.
- [72] Lin, Z., et Zeng, C. : BGK waves and nonlinear Landau damping. Preprint, 2010.
- [73] Liverani, C., et Olla, S. : Toward the Fourier law for a weakly interacting anharmonic crystal. Preprint, 2010.
- [74] Loschmidt, J. : Über den Zustand des Wärmegleichgewichtes einse Systems von Körpern mit Rücksicht auf die Schwerkraft. *Wien. Ber.* **73**, 128 (1876).
- [75] Lynden-Bell, D. : The stability and vibrations of a gas of stars. *Mon. Not. R. astr. Soc.* **124**, 4, 279–296 (1962).
- [76] Lynden-Bell, D. : Statistical mechanics of violent relaxation in stellar systems. *Mon. Not. R. astr. Soc.* **136**, 101–121 (1967).
- [77] Malmberg, J., et Wharton, C. : Collisionless damping of electrostatic plasma waves. *Phys. Rev. Lett.* **13**, 6, 184–186 (1964).
- [78] Malmberg, J., Wharton, C., Gould, R., et O’Neil, T. : Plasma wave echo experiment. *Phys. Rev. Letters* **20**, 3, 95–97 (1968).

- [79] Maslen, D. : The eigenvalues of Kac's master equation. *Math. Zeit.* **243**, 291–331 (2003).
- [80] Maxwell, J.C. : On the dynamical theory of gases. *Philos. Trans. Roy. Soc. London Ser. A* **157**, 49–88 (1867).
- [81] McKean, H.J. : Speed of approach to equilibrium for Kac's caricature of a Maxwellian gas. *Arch. Rational Mech. Anal.* **21**, 343–367 (1966).
- [82] Mouhot, C. : Rate of convergence to equilibrium for the spatially homogeneous Boltzmann equation with hard potentials. *Comm. Math. Phys.* **261**, 629–672 (2006).
- [83] Mouhot, C., et Strain, R.M. : Spectral gap and coercivity estimates for linearized Boltzmann collision operators without angular cutoff. *J. Math. Pures Appl.* **87**, 5, 515–535 (2007).
- [84] Mouhot, C., et Villani, C. : Regularity theory for the spatially homogeneous Boltzmann equation with cut-off. *Arch. Rational Mech. Anal.* **173**, 2, 169–212 (2004).
- [85] Mouhot, C., et Villani, C. : On Landau damping. À paraître dans *Acta Math.* Disponible en ligne à [arXiv :0904.2760](https://arxiv.org/abs/0904.2760)
- [86] O'Neil, T. : Collisionless damping of nonlinear plasma oscillations. *Phys. Fluids* **8**, 12, 2255–2262 (1965).
- [87] Neunzert, H. : An introduction to the nonlinear Boltzmann–Vlasov equation. In *Kinetic theories and the Boltzmann equation*, C. Cercignani, Ed., vol. 1048 of *Lecture Notes in Math.*, Springer, Berlin, Heidelberg, 1984, pp. 60–110.
- [88] Penrose, O. : Electrostatic instability of a non-Maxwellian plasma. *Phys. Fluids* **3**, 258–265 (1960).
- [89] Poincaré, H. : Le mécanisme et l'expérience. *Revue de Métaphysique et de Morale* **I**, 534–537 (1893).
- [90] Pulvirenti, M. : From particle to transport equations. In *Session "L'Etat de la Recherche" de la SMF* (1998). "Kinetic equations and asymptotic theory", F. Bouchut, F. Golse, M. Pulvirenti, coll. "Series in Appl. Math.", Gauthier-Villars, 2000.
- [91] Robert, R. : Statistical mechanics and hydrodynamical turbulence. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)* (Basel, 1995), Birkhäuser, pp. 1523–1531.
- [92] Ryutov, D. D. : Landau damping : half a century with the great discovery. *Plasma Phys. Control. Fusion* **41**, A1–A12 (1999).
- [93] Saint-Raymond, L. : *Hydrodynamic limits of the Boltzmann equation*. Lectures at SISSA, Trieste (2006), Springer Monograph, à paraître.
- [94] Shannon, C.E., et Weaver, W. : *The mathematical theory of communication*. University of Illinois Press, Urbana, Illinois, 1949.

- [95] Spohn, H. : *Large scale dynamics of interacting particles*. Texts and Monographs in Physics. Springer-Verlag, Berlin, 1991.
- [96] Strain, R.M. : On the linearized Balescu–Lenard equation. *Comm. Partial Differential Equations* **32**, 1551–1586 (2007).
- [97] Sturm, K.-Th. : Entropic measure on multidimensional spaces. In *Stochastic analysis, random fields and applications VI*, Dalang, R., Dozzi, M., and Russo, F., Eds. Progress in Probability, Birkhuser.
- [98] Sznitman, A.-S. : Equations de type de Boltzmann, spatialement homogènes. *Z. Wahrsch. Verw. Gebiete* **66**, 559–562 (1984).
- [99] Sznitman, A.-S. : *Topics in propagation of chaos*. In *École d’Été de Probabilités de Saint-Flour XIX (1989)*, Lect. Notes in Math. **1464** (1991), Springer, Berlin, pp. 165–251.
- [100] Tanaka, H. : An inequality for a functional of probability distributions and its application to Kac’s one-dimensional model of a Maxwellian gas. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **27**, 47–52 (1973).
- [101] Toscani, G., et Villani, C. : Sharp entropy dissipation bounds and explicit rate of trend to equilibrium for the spatially homogeneous Boltzmann equation. *Comm. Math. Phys.* **203**, 3, 667–706 (1999).
- [102] Turkington, B. : Statistical equilibrium measures and coherent states in two-dimensional turbulence. *Comm. Pure Appl. Math.* **52**, 7, 781–809 (1999).
- [103] Villani, C. : *A review of mathematical topics in collisional kinetic theory*. In *Handbook of Mathematical Fluid Dynamics I*, 71–305, S. Friedlander et D. Serre, Édts, North-Holland, Amsterdam, 2002.
- [104] Villani, C. Cercignani’s conjecture is sometimes true and always almost true. *Commun. Math. Phys.* **234**, 455–490 (2003).
- [105] Villani, C. : Mathematics of granular materials. *J. Stat. Phys.* **124**, no. 2-4, 781–822 (2006).
- [106] Villani, C. : Hypocoercive diffusion operators. Texte de ma conférence au Congrès International des Mathématiciens (Madrid, août 2006).
- [107] Villani, C. : Entropy production and convergence to equilibrium. Notes pour une série de cours à l’Institut Henri Poincaré, Paris (automne 2001). In *Entropy methods for the Boltzmann equation*, 1–70, Lecture Notes in Math., Vol. 1916, Springer, Berlin, 2008.
- [108] Villani, C. : *Hypocoercivity*. *Mem. Amer. Math. Soc.* **202**, no. 950 (2009).
- [109] Villani, C. : *Landau damping*. Notes de cours, CEMRACS 2010. Disponible sur www.math.univ-lyon1.fr/~villani
- [110] Villani, C. : Is there any backward solution of the Boltzmann equation ? Travail non publié, disponible sur www.math.univ-lyon1.fr/~villani

- [111] Vlasov, A.A. : On the oscillation properties of an electron gas. Zh. Èksper. Teoret. Fiz. **8**, 291–318 (1938).
- [112] Zermelo, E. : Über einen Satz der Dynamik und die mechanische Wärmetheorie. Annalen der Physik **54**, 485–494 (1896).