Landau damping

Cédric Villani

École Normale Supérieure de Lyon & Institut Henri Poincaré, 11 rue Pierre et Marie Curie, F-75231 Paris Cedex 05, FRANCE

E-mail address: `cvillani@umpa.ens-lyon.fr`

Contents

Abstract. These notes constitute an introduction to the Landau damping phenomenon in the linearized and perturbative nonlinear regimes, following the recent work [**74**] by Mouhot & Villani.

Foreword

In 1936, Lev Landau devised the basic collisional kinetic model for plasma physics, now commonly called the Landau–Fokker–Planck equation. With this model he was introducing the notion of relaxation in plasma physics: relaxation *à la* Boltzmann, by increase of entropy, or equivalently loss of information.

In 1946, Landau came back to this field with an astonishing concept: relaxation without entropy increase, with preservation of information. The revolutionary idea that conservative phenomena may exhibit irreversible features has been extremely influential, and later led to the concept of violent relaxation.

This idea has also been controversial and intriguing, triggering hundreds of papers and many discussions. The basic model used by Landau was the linearized Vlasov–Poisson equation, which is only a formal approximation of the Vlasov–Poisson equation. In the present notes I shall present the recent work by Clément Mouhot and myself, extending Landau's results to the nonlinear Vlasov–Poisson equation in the perturbative regime. Although this extension is still far from handling the mysterious fully nonlinear regime, it already turned out to be rich and tricky, from both the mathematical and the physical points of view.

These notes start with basic reminders about classical particle systems and Vlasov equations, assuming no prerequisite from modeling nor physics. Standard notation is used throughout the text, except maybe for the Fourier transforms: if $h = h(x, v)$ is a function on the position-velocity phase space, then $\widehat{h}$ stands for the Fourier transform in the x variable only, while $\widetilde{h}$ stands for the Fourier transform in both x and v variables. Precise conventions will be given later on.

This course was first taught in the summer of 2010 in Cotonou, Benin, on the invitation of Wilfrid Gangbo; and in Luminy, France, on the invitation of Éric Sonnendrücker. It is a pleasure to thank the audience for their interest and enthusiasm. Typing was mostly performed in the welcoming library of the gorgeous Domaine des Treilles

of the Fondation Schlumberger, whose hospitality is gratefully acknowledged, during a meeting on wave turbulence organized by Christophe Josserand.

CHAPTER 1

Mean field approximation

The two main classes of kinetic equations are the collisional equations of Boltzmann type, modeling short-range interactions, and the mean field equations of Vlasov type, modeling long-range interactions. The disinction between short-range and long-range does not refer to the decay of the microscopic interaction, but to the fact that the relevant interaction takes place at distances which are much smaller than, or comparable to, the macroscopic scale; in fact both types of interaction may occur simultaneously. Collisional equations are discussed in my survey [**99**]. In this chapter I will concisely present the archetypal mean field equations.

1. The Newton equations

The collective interaction of a large population of "particles" arises in a number of physical situations. The basic model consists in the system of Newton equations in $\mathbb{R}^d$ (typically $d = 3$):

$$m_i \, \ddot{x}_i(t) = \sum_j F_{j\to i}(t), \tag{1.1}$$

where m_i is the mass of particle i, $x_i(t) \in \mathbb{R}^d$ its position at time t, $\ddot{x}_i(t)$ its acceleration, and $F_{j\to i}$ is the force exerted by particle j on particle i. Even if this model does not take into account quantum or relativistic effects, huge theoretical and practical problems remain dependent on our understanding of (1.1).

The masses in (1.1) may differ my many orders of magnitude; actually this disparity of masses plays a key role in the study of the solar system, or the Kolmogorov–Arnold–Moser theory [**27**], among other things. But it also often happens that the situation where all masses m_i are equal is relevant, at least qualitatively. In the sequel, I shall only consider this situation, so $m_i = m$ for all i.

If the interaction is translation invariant, it is often the case that the the force derives from an **interaction potential**: there is $W : \mathbb{R}^d \to \mathbb{R}$ such that

$$F = -\nabla W(x - y)$$

is the force exterted at position x by a particle located at position y. This formalism misses important classes of interaction such as magnetic forces, but it will be sufficient for our purposes.

EXAMPLES 1.1. (a) $W(x-y) = \text{const.}\, \rho\, \rho'/|x-y|$ is the electrostatic interaction potential between particles with respective electric charges ρ and ρ', where $|x-y|$ is the Euclidean distance in $\mathbb{R}^3$; (b) $W(x-y) = -\text{const.}\, m\, m'/|x-y|$ is the gravitational interaction potential between particles with respective masses m and m', also in $\mathbb{R}^3$; (c) Essentially any potential W arises in some physical problem or the other, and even a smooth (or analytic!) interaction potential W leads to relevant and difficult problems.

As an example, let us write the basic equation governing the positions of stars in a galaxy:

$$\ddot{x}_i(t) = \mathcal{G} \sum_{j \neq i} m_j \, \frac{x_j - x_i}{|x_j - x_i|^3},$$

where $\mathcal{G}$ is the gravitational constant. Note that in this example, a star is considered as a "particle"! There are similar equations describing the behavior of ions and electrons in a plasma, involving the dielectric constant, mass and electric charges.

In the sequel, I will assume that all masses are equal and work in adimensional units, so masses will not explicitly appear in the equations.

But now there are as many equations as there are particles, and this means a lot. A galaxy may be made of $N \simeq 10^{13}$ stars, a plasma of $N \simeq 10^{20}$ particles... thus the equations are untractable in practice. Computer simulations, available on Internet, give a flavor of the rich and complex behavior displayed by large particle systems interacting through gravity. It is very difficult to say anything intelligent in front of these complex pictures. This complex behavior is partly due to the fact that the gravitational potential is attractive and singular at the origin; but even for a smooth interaction W would the large value of N cause much trouble in the quantitative analysis.

The **mean field limit** will lead to another model, more amenable to mathematical treatment.

2. Mean field limit

The limit $N \to \infty$ allows to replace a very large number of simple equations by just one complicated equation. Although we are trading reassuring ordinary differential equations for dreaded partial differential equations, the result will be more tractable.

From the theoretical point of view, the mean field approximation is fundamental: not only because it establishes the basic limit equation, but also because it shows that the qualitative behavior of the system does not depend much on the exact value of the number of particles, and then in numerical simulations for instance we can replace trillions of particles by, say, millions or even thousands.

It is not a priori obvious how one can let the dimension of the phase space go to infinity. As a first step, let us double variables to convert the second-order Newton equations into a first-order system. So for each position variable x_i we introduce the velocity variable $v_i = \dot{x}_i$ (time-derivative of the position), so that the whole state of the system at time t is described by $(x_1, v_1), \ldots, (x_N, v_N)$. Let us write X^d for the d-dimensional space of positions, which may be $\mathbb{R}^d$, or a subset of $\mathbb{R}^d$, or the d-dimensional torus $\mathbb{T}^d$ if we are considering periodic data; then the space of velocities will be $\mathbb{R}^d$.

Since all particles are identical, we do not really care about the state of each particle individually: it is sufficient to know the state of the system *up to permutation of particles.* In slightly pedantic terms, we are taking the quotient of the phase space $(X^d \times \mathbb{R}^d)^N$ by the permutation group $\mathcal{S}_N$, thus obtaining a cloud of undistinguishable points.

There is a one-to-one correspondence between such a cloud $\mathcal{C} = \{(x_1, v_1), \ldots, (x_N, v_N)\}$ and the associated *empirical measure*

$$\widehat{\mu}^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{(x_i, v_i)},$$

where $\delta_{(\overline{x}, \overline{v})}$ is the Dirac mass in phase space at $(\overline{x}, \overline{v})$. From the physical point of view, the empirical measure counts particles in phase space.

Now the empirical measure $\widehat{\mu}^N$ belongs to the space $P(X^d \times \mathbb{R}^d)$, the space of probability measures on the single-particle phase space. This space is infinite-dimensional, but it is *independent of the number of particles.* So the plan is to re-express the Newton equations in terms of the empirical measure, and then pass to the limit as $N \to \infty$.

For simplicity I shall assume that X^d is either $\mathbb{R}^d$ or $\mathbb{T}^d$, and that the force derives from an interaction potential W. The following proposition, slightly informal, establishes the link between the Newton equations and the empirical measure equation.

Proposition 1.2. *(i) Let $W \in C^1(X^d; \mathbb{R})$, and for each i let $x_i = x_i(t)$; then with the notation $\widehat{\mu}^N = N^{-1} \sum \delta_{(x_i, \dot{x}_i)}$ the following two* *statements are equivalent:*

$$\forall i, \quad \ddot{x}_i = -c \sum_j \nabla W(x_i - x_j) \tag{1.2}$$

$$\frac{\partial \widehat{\mu}^N}{\partial t} + v \cdot \nabla_x \widehat{\mu}^N + F^N(t,x) \cdot \nabla_v \widehat{\mu}^N = 0, \tag{1.3}$$

where

$$F^N(t,x) = -c \sum_j \nabla W(x - x_j) = -c\, N \left(\nabla W *_{x,v} \widehat{\mu}^N\right).$$

(ii) If ∇W is uniformly continuous and $\widehat{\mu}_0^N$ converges weakly to some measure μ_0 and $c = c(N)$ satisfies $cN \to \gamma \geq 0$ as $N \to \infty$, then up to extraction of a subsequence, $\widehat{\mu}^N$ converges as $t \to \infty$ to a time-dependent measure $\mu = \mu_t(dx\,dv)$ solving the system

$$\begin{cases} \dfrac{\partial \mu}{\partial t} + v \cdot \nabla_x \mu + F(t,x) \cdot \nabla_v \mu = 0 \\ F = -\gamma\, \nabla W *_{x,v} \mu \end{cases} \tag{1.4}$$

Remark 1.3. Equations (1.3) and (1.4) are to be understood in distributional sense, that is, after integrating on the phase space against a nice test function $\varphi(x,v)$, say smooth and compactly supported. To rewrite these equations in distributional form, note that

$$v \cdot \nabla_x \mu = \nabla_x \cdot (v\mu), \qquad F(t,x) \cdot \nabla_v \mu = \nabla_v \cdot \big(F(t,x)\,\mu\big).$$

(To be rigorous one should also use a test function in time, but this is not a serious issue and I shall leave it aside.)

Remark 1.4. The second formula in (1.4) can be made more explicit as

$$F(t,x) = -\iint W(x-y)\,\mu_t(dy\,dw);$$

of course the convolution in the velocity variable is trivial since ∇W does not depend on it; so this is just an integration in velocity space.

Remark 1.5. By definition, a sequence of measures μ^N converges to a measure μ in the weak sense if, for any bounded continuous function $\varphi(x,v)$,

$$\iint \varphi(x,v)\,\mu^N(dx\,dv) \xrightarrow[N\to\infty]{} \iint \varphi(x,v)\,\mu(dx\,dv).$$

If μ^N and μ are probability measures, then weak convergence is equivalent to convergence in the sense of distributions.

Sketch of proof of Proposition 1.2. Let us forget about issues of regularity and well-posedness, and focus on the core computations, assuming that $x_i(t)$ is a smooth function of t. When we test equation (1.3) against an arbitrary function $\varphi = \varphi(x,v)$ we obtain

$$\frac{d}{dt}\left[\frac{1}{N}\sum_i \varphi(x_i,v_i)\right] - \frac{1}{N}\sum_i \left(v\cdot\nabla_x\varphi\right)\big|_{(x_i,v_i)} - \frac{1}{N}\sum_i \left(F^N\cdot\nabla_v\varphi\right)\big|_{(x_i,v_i)} = 0,$$

where the time-dependence is implicit; by chain-rule this means

$$\frac{1}{N}\sum_i \Big(\nabla_x\varphi\cdot\dot{x}_i + \nabla_v\varphi\cdot\dot{v}_i - \nabla_x\varphi\cdot v_i - \nabla_v\varphi\cdot F^N(x_i)\Big) = 0,$$

where φ inside the summation is evaluated at (x_i,v_i). Since $v_i = \dot{x}_i$, this equation reduces to

$$\frac{1}{N}\sum_i \big[\dot{v}_i - F^N(t,x_i)\big]\cdot\nabla_v\varphi(x_i,v_i) = 0. \tag{1.5}$$

Now this should hold true for *any* test function $\varphi(x,v)$. Choosing one which takes the form $e\cdot v$ near (x_i,v_i) (with e an arbitrary vector) and which vanishes near (x_j,v_j) for all $j\neq i$, we deduce that $\dot{v}_i = F^N(t,x_i)$. (This argument is not fully rigorous since it may happen that two distinct particles occupy similar positions in phase space, but that is not a big deal to fix.) Now (1.5) is just a way to rewrite (1.2); the equivalence between (1.2) and (1.3) follows easily.

Next we note that $\sum \nabla W(x-x_j) = N\nabla W * \widehat{\mu}$, where the convolution is in both variables x and v. In retrospect, it is normal that the force should be expressed in terms of the empirical measure, since this is a symmetric expression of the positions of particles.

Now let us consider the limit $N\to\infty$. Let us fix a finite time-horizon $T>0$ and work on the time-interval $[-T,T]$. By assumption the initial data $\widehat{\mu}^N(0,\cdot)$ form a tight family; then from the differential equation satisfied by the measures $\widehat{\mu}^N(t,\,\cdot\,)$ it is not difficult to show that $\widehat{\mu}^N(t,\,\cdot\,)$ is also tight, uniformly in $t\in[-T,T]$. Then, up to extraction of a subsequence, $\widehat{\mu}^N(t,\cdot)$ will converge in $C([-T,T];\mathcal{D}'(X^d\times\mathbb{R}^d))$ for any $T>0$, to some limit measure $\mu(t,dx\,dv)$. It only remains to pass to the limit in the equation.

Being the convolution of a uniformly continuous function with a probability measure, hte force field $F^N = -cN\,\nabla W * \widehat{\mu}^N$ is uniformly continuous on $[-T,T]\times X^d$, and will converge uniformly as $N\to\infty$ to $-\gamma\nabla W * \mu$. This easily implies that

$$F^N\,\widehat{\mu}^N \longrightarrow F\,\mu$$

in distributional sense, whence $\nabla_v \cdot (F^N \widehat{\mu}^N) \longrightarrow \nabla_v \cdot (F\,\mu)$. Similarly, $\nabla_x \cdot (v\widehat{\mu}^N)$ converges to $\nabla_x \cdot (v\mu)$, and the proof is complete. $\square$

The limit equation (1.4) is called the **nonlinear Vlasov equation** associated with the interaction potential W. It makes sense just as well for $\mu_t(dx\,dv) = N^{-1}\sum \delta_{(x_i(t),v_i(t))}$ (in which case it reduces to the Newton dynamics) as for $\mu_t(dx\,dv) = f(t,x,v)\,dx\,dv$, that is, for a continuous distribution of matter. In fact the nonlinear Vlasov equation is the *completion*, in the space of measures, of the system of Newton equations.

It is customary and physically relevant to restrict to the case of a continuous distribution function, and then focus on the equation satisfied by $f(t,x,v)$. Since the Lebesgue measure $dx\,dv$ is transparent to the differential operators ∇_x and ∇_v, one easily obtains the **nonlinear Vlasov equation for the density function** $f = f(t,x,v)$:

$$\begin{cases} \dfrac{\partial f}{\partial t} + v \cdot \nabla_x f + F(t,x) \cdot \nabla_v f = 0 \\ F = -\nabla W *_x \rho, \qquad \rho(t,x) = \displaystyle\int f(t,x,v)\,dv, \end{cases} \tag{1.6}$$

where the (x,v)-convolution has been explicitly replaced by a convolution in x and an integration in v.

Equation (1.6) is the single most important partial differential equations of mean field systems, and will be the object of study of this course.

3. Precised results

In Proposition 1.2 it was assumed that W is continuously differentiable. If W is smoother then one can prove more precise results of quantitative convergence, involving distances on probability measures, for instance the Wasserstein distances W_p. For the present section, it will be sufficient to know the 1-Wasserstein distance, defined by the formula

$$W_1(\mu,\nu) := \sup\ \left\{ \int \psi\,d\mu - \int \psi\,d\nu; \quad \|\psi\|_{\mathrm{Lip}} \le 1 \right\},$$

where the supremum is over all 1-Lipschitz functions ψ of (x,v), and it is assumed that μ and ν possess a finite moment of order 1. (If one imposes that ψ is also bounded in supremum norm, one obtains the closely related "bounded Lipschitz" distance, which does not need any moment assumption.)

Here is a typical estimate of convergence for the mean-field limit, stated here without proof, going back to Dobrushin:

Proposition 1.6. *If $\mu_t(dx\,dv)$ and $\nu_t(dx\,dv)$ are two solutions of the nonlinear Vlasov equation with interaction potential W, then for any $t \in \mathbb{R}$*

$$W_1(\mu_t, \nu_t) \le e^{2C|t|}\, W_1(\mu_0, \nu_0), \qquad C = \max\big(\|\nabla^2 W\|_{L^\infty}, 1\big). \tag{1.7}$$

It might not be obvious why this provides a convergence estimate in the mean-field limit. To see this, choose $\mu_t(dx\,dv) = f(t,x,v)\,dx\,dv$ and $\nu_t = \widehat{\mu}_t^N$; then (1.7) controls at time t the distance between the limit mean-field behavior and the Newton equation for N particles, in terms of how small this distance is at initial time $t = 0$. If the particles at $t = 0$ are chosen randomly, then typically the W_1 distance at $t = 0$ is $O(1/\sqrt{N})$, so $W_1(\mu_t, \nu_t) = O(e^{2C|t|}/\sqrt{N})$, which solves the problem. (Note that this estimate requires crazy amounts of particles to get a good precision in large time.)

Another type of estimates are large deviation bounds:

Proposition 1.7. *If $\nabla^2 W$ is bounded, $f_0 = f_0(x,v)$ is given with $\iint f_0(x,v)\, e^{\beta(|x|^2+|v|^2)}\,dx\,dv \le C_0$, $(x_i(0), \dot{x}_i(0))$, $1 \le i \le N$, are chosen randomly and independently according to $f_0(x,v)\,dx\,dv$, $(x_i(t))$ solve the Newton equations* (1.2) *with $c = 1/N$, and $f(t,x,v)$ solves the nonlinear Vlasov equation* (1.6)*, then there is $K > 0$ such that for any $T \ge 0$ there is $C = C(T)$ such that*

$$N \ge N_0 \max\big(\varepsilon^{-(2d+3)}, 1\big) \implies \mathbb{P}\left[\sup_{0\le t\le T} W_1\Big(\widehat{\mu}_t^N, f(t,x,v)\,dx\,dv\Big) > \varepsilon\right] \le C\left(1+\varepsilon^{-2}\right) e^{-KN\varepsilon^2}, \tag{1.8}$$

where $\mathbb{P}$ stands for probability.

Many refinements are possible: for instance, one can estimate the density error between $f(t,x,v)$ and the empirical measure, after smoothing by a peaked convolution kernel; study the evolution of (de)correlations between particles which are initially randomly distributed; show that trajectories of particles in the system of size N are well approximated by trajectories of particles evolving in the limit mean-field force, etc.

4. Singular potentials

Fine. But eventually, more often than not, the interaction potential is not smooth at all, instead it is rather singular. Then nobody has a clue of why the mean-field limit should be true. The problem might be just technical, but on the contrary it seems very deep.

Such is the case in particular for the most important nonlinear Vlasov equations, namely the **Vlasov–Poisson equations**, where W

is the fundamental solution of $\pm\Delta$. In dimension $d = 3$, writing $r = |x - y|$, we have

- the Coulomb interaction (repulsive) $W = \dfrac{1}{4\pi r}$;
- the Newton interaction (attractive) $W = -\dfrac{1}{4\pi r}$.

Then the equation $F = -\nabla W * \rho$ becomes $F = \pm\nabla\Delta^{-1}\rho$.

It is remarkable that, up to a change of sign in the interaction, the very same equation describes systems of such various scales as a plasma and a galaxy, in which each star counts as one particle! In fact to be more precise, we should slightly change the equation for plasmas, by taking into account the contribution of heavy ions, which is usually considered in the form of a fixed density of positive charges, say $\rho_I(x)$, and by considering magnetic effects, which in some situations play an important role. Things become much more messy when irreversible phenomena are taken into account, but these phenomena occur only as corrections to the mean-field limit, due to the fact that N is finite.

While the mean-field limit for smooth potential has been well-understood for more than three decades, in the case of singular potentials the only available results are those obtained a few years ago by Hauray and Jabin: they assume that (a) the interaction is no too singular: essentially $|\nabla W| = O(r^{-s})$ with $0 < s < 1$ (independently of the dimension d); and (b) particles are well-separated in phase space initially, so

$$\inf_{j\neq i}\Big(|x_i - x_j| + |v_i - v_j|\Big) \geq \frac{c}{N^{\frac{1}{2d}}}, \tag{1.9}$$

where c is of course independent of N.

Both conditions are not so satisfactory: assumption (a) misses the Coulomb/Newton singularity by an order $1 + 0$, while assumption (b) cannot be true in the simplest case when particles are chosen randomly and idependently of each other. It might be that assumption (b) can be given a physical justification, though, based on the ionization process for instance; but that remains to be done. For numerical purpose, assumption (b) is more satisfactory since we can choose the discretization as we wish.

In any case, a key ingredient in the proof of the Hauray–Jabin theorem consists in showing that the separation (1.9) property is propagated in time: if true at $t = 0$, it remains true for later times, up to a deterioration of constants. This implies that the proportion of particles located in a box of side ε in phase space remains bounded like $O(\varepsilon^{2d})$ as time goes by, uniformly in N. (This is a discrete analogue of the property of propagation of L^∞ bounds for the nonlinear Vlasov equation, which will be examined in the next chapter)

What about the theory of the nonlinear Vlasov equation? Is the system well-posed for a given initial datum? For smooth interactions this does not pose any problem, but when the interaction potential is singular, this becomes highly nontrivial. Most efforts have been focused on the Poisson coupling in dimension 3. Although this may not have been considered carefully, the theory would probably work just the same in arbitrary dimensions and with a coupling that is no more singular than Poisson. There are two famous theories for the Vlasov–Poisson equation with large data:

• The Pfaffelmoser theory, developed and simplified in particular by Batt, Rein, Glassey, Scheffer, construct smooth solutions assuming essentially that f_i is C^1 and compactly supported in (x, v).

• The Lions–Perthame theory constructs a unique solution for an initial datum f_i on $\mathbb{R}^3_x \times \mathbb{R}^3_v$ which satisfies, say,

$$|f_i(x,v)| + |\nabla f(x,v)| \leq \frac{C}{(1+|x|+|v|)^{10}}. \tag{1.10}$$

(The exponent 10 depends on the fact that dimension is 3, and anyway should not be taken seriously.) Besides velocity averaging phenomena, the key insight of the analysis is the propagation of bounds on velocity moments of order greater than 3. Then one can show that the spatial density is uniformly bounded, and the smoothness is propagated too.

Both theories are still incomplete. The Lions–Perthame theory takes advantage of the dispersion at large positions to control velocity-moments; it has never been checked that it can be adapted in bounded geometries, like the torus $\mathbb{T}^3$. As for the Pfaffelmoser theory, it does adapt to bounded geometries, but the assumption of compact support in the velocity space, is a heresy, since it does not include even the single most important distribution in kinetic theory, namely the Gaussian distribution.

Perturbative theories of the nonlinear Vlasov equation near an equilibrium are in better shape. We shall see an example in this course. This suggests that the problem of the mean-field limit in a perturbative setting could be attacked.

Bibliographical notes

Impressive particle simulations of large systems, performed by John Dubinski, can be found online at `www.galaxydynamics.org`

The kinetic theory of plasmas was born in Soviet Union in the thirties, when Landau adapted the Boltzmann collision operator to the Coulomb interaction [**55**] and Vlasov argued that long-range interactions should be taken into account by a conceptually simpler mean-field term [**103**]. The collisional kinetic theory of plasmas is described in a number of physics textbooks [**1, 54, 61**] and in the mathematical review [**99**]; see also [**2**, Sections 1 and 2].

The mean field limit however did not become a mathematical subject until the classical works by Dobrushin [**31**], Braun & Hepp [**20**], and Neunzert [**77**]. Braun & Hepp were also interested in the propagation of chaos and the study of fluctuations; these topics are addressed again in Sznitman's Saint-Flour lecture notes [**95**]. Other synthetic sources are the book by Spohn [**92**] and my incomplete lecture notes on the mean field limit [**102**], which both contain a recast of the proof of Proposition 1.6. Quantitative estimates of the mean field limit for simple (stochastic) models and smooth interaction are found in my work [**15**] joint with Bolley and Guillin; the proof of Proposition 1.7 can be obtained by adapting the estimates therein.

The mean-field limit for mildly singular interactions was considered by Hauray and Jabin [**42**] in a pioneering work that still needs to be digested and simplified by the mathematical community.

Early contributions to the Cauchy problem for the Vlasov–Poisson equation, working either in short time, or with weak solutions, or in small dimension, are due to Arsen'ev, Horst, Bardos, Degond, Benachour, DiPerna & Lions in the seventies and eighties [**6, 7, 9, 12, 28, 46, 47**]. The theory reached a more mature stage with the groundbreaking works by Pfaffelmoser [**83**] and Lions & Perthame [**60**] at the dawn of the nineties. Pfaffelmoser's approach was simplified by Schaeffer [**90**] and Horst [**48**], and is well exposed by Glassey [**35**]; the adaptation to periodic data was performed by Batt & Rein [**10**]. The alternative Lions–Perthame approach is presented by Bouchut [**17**].

CHAPTER 2

Qualitative behavior of the Vlasov equation

In the previous chapter we were interested in the derivation and well-posedness of the Vlasov equation

$$
\begin{cases}
\dfrac{\partial f}{\partial t} + v \cdot \nabla_x f + F(t,x) \cdot \nabla_v f = 0 \\
F = -\nabla W *_x \rho \qquad \rho(t,x) = \displaystyle\int f(t,x,v)\, dv.
\end{cases}
\tag{2.1}
$$

But now the emphasis will be different: starting from the Vlasov equation, we shall enquire about its *qualitative behavior*. This problem fills up textbooks in physics, and has been the subject of an enormous amount of literature.

1. Boundary conditions

There is a zoology of boundary conditions for the Vlasov equation. To avoid discussing them, I shall continue to assume that the position space is either $X^d = \mathbb{R}^d$, the whole space, or $X^d = \mathbb{T}^d/\mathbb{Z}^d$, the d-dimensional torus. The latter case deserves some comments.

If W is a given potential in $\mathbb{R}^d$, then in the periodic setting, formally W should be replaced by its periodic version W^{per}:

$$W^{\mathrm{per}}(x) = \sum_{k \in \mathbb{Z}^d} W(x-k).$$

If W decays fast enough, this is well-defined, but if W has slow decay, like in the case of Poisson interaction, this will not converge! Then the justification requires some argument. In fact, it is clear that for Poisson coupling the potential cannot converge: in the case of the Poisson coupling, the total potential $W * \rho$ should formally be equal to $\pm\Delta^{-1}\rho$, which does not make sense since ρ does not have zero mean... To get around this problem, we would like to take out the mean of ρ.

In the plasma case, one can justify this by going back to the model: indeed, one may argue that the density of ions should be taken into account, that it can be modelled as a uniform background because ions are much heavier and move on longer time scales than electrons, and that the density of ion charges should be equal to the mean density of electrons because the plasma should be globally neutral. This amounts to replace the potential $W * \rho$ by $W * (\rho - \langle \rho \rangle)$, where $\langle \rho \rangle = \int \rho \, dx$.

The preceding reasoning is based on the existence of two different species of particles. But even if there is just one species of particles, as is the case for gravitational interaction, it is still possible to argue that the mean should be removed. Indeed, in (2.1) W only appears through its gradient, and, whenever c is a constant,

$$\nabla W * (\rho - c) = \nabla W * \rho - \nabla W * c = \nabla W * \rho.$$

Thus, if W decays fast enough at infinity and ρ is periodic,

$$\nabla W * \rho = \nabla W^{\mathrm{per}} * \rho = \nabla W^{\mathrm{per}} * (\rho - \langle \rho \rangle).$$

If W does not decay fast enough at infinity, then at least we can write $W = \lim_{\varepsilon \to 0} W_\varepsilon$, where W_ε is an approximation decaying fast at infinity (say $\pm e^{-r/\varepsilon}/(4\pi r)$), then $\nabla W_\varepsilon * \rho = \nabla W_\varepsilon^{\mathrm{per}} * (\rho - \langle \rho \rangle)$, which in the limit $\varepsilon \to 0$ converges to $\nabla W^{\mathrm{per}} * (\rho - \langle \rho \rangle)$. Of course this might not be so convincing in the absence of a clear discussion of the meaning of the parameter ε, but at least makes sense in some regime and allows to take out the mean $\langle \rho \rangle$ from the density in (2.1). This operation is similar to the so-called **Jeans swindle** in astrophysics.

Having warned the reader that there is a subtle point here, from now on in the periodic setting I shall always write $\nabla W * \rho$ for $\nabla W * (\rho - \langle \rho \rangle)$. As a final comment, one may argue against the relevance of periodic boundary conditions, especially in view of the above discussion; but this is still by far the simplest way to have access to a confined geometry, avoiding effects such as dispersion at infinity which completely change the qualitative behavior of the nonlinear Vlasov equation.

2. Structure

The nonlinear Vlasov equation is a transport equation, and can therefore be solved by the well-known **method of characteristics**: if f solves the equation, then the measure $f(t, x, v)\, dx\, dv$ is the **push-forward** of the initial measure $f_i(x, v)\, dx\, dv$ by the flow $S_{0,t} = (X_t, V_t)$ in phase space, solving the characteristic equations

$$\begin{cases} \dot{X}_t = V_t, \qquad \dot{V}_t = F(t, X_t), \qquad F = -\nabla W * \rho, \\[2ex] (X_0, V_0) = (x, v). \end{cases}$$

Of course this does not solve the problem "explicitly", since the force F at time t depends on the whole distribution of particles via the formula $F = -\nabla W * (\int f \, dv)$.

Recall that the push-forward of a measure μ_0 by a map S is defined by $S_\#\mu_0[A] = \mu_0[S^{-1}(A)]$. The resulting equation on the densities generally involves the Jacobian determinant of the flow at time t. However in the present case, the flow S_t induced by $F(t,x)$ preserves the Liouville measure $dx\,dv$ (that is a consequence from its Hamiltonian nature), so the push-forward equation can be simplified in a pull-back equation for densities. In other words, the solution $f(t,x,v)$ will satisfy

$$f(t, S_{0,t}(x,v)) = f(0,x,v). \tag{2.2}$$

Thus, to get the distribution function at time t we should invert the map S_t, in other words *solve the characteristics backwards.* If $S_{t,0}$ stands for the inverse of $S_{0,t}$, then (2.2) becomes

$$f(t,x,v,) = f(0, S_{t,0}(x,v)). \tag{2.3}$$

Depending on situation, taste and theory, one considers the nonlinear Vlasov equation either from the Eulerian point of view (focus on $f(t,x,v)$), or from the Lagrangian point of view (focus on particle trajectories in a force field reconstructed from the particle distribution). This affects not only the theory, but also the numerics, since numerical methods may be Eulerian (look at values of f on a grid, say), or Lagrangian (consider particles moving), or semi-Lagrangian (make particles move and interpolate at each step to reconstruct values of f on a grid).

Apart from that, equation (2.1) is a limit of Hamiltonian equations (the Newton equations), and actually has a Hamiltonian structure in a certain sense, in relation with optimal transport theory; this link was explored in particular by Ambrosio, Gangbo and Lott. For the moment it is not clear whether this striking structure has physically relevant implications beyond what is already known.

3. Invariants and identities

In this section I shall review the four main invariances and identities associated with the nonlinear Vlasov equation, assuming that everything is well-defined and being content with formal identities.

- The nonlinear Vlasov equation preserves the **total energy**

$$\iint f(x,v)\frac{|v|^2}{2}\,dx\,dv + \frac{1}{2}\iint W(x-y)\,\rho(x)\,\rho(y)\,dx\,dy \;=: T + U$$

is constant in time along solutions. The total energy is the sum of the kinetic energy T and the potential energy U. (The factor $1/2$ in the definition of U comes from the fact that we should count unordered pairs of particles.)

• The nonlinear Vlasov equation preserves all the **nonlinear integrals of the density**: often called the Casimirs of the equation, they take the form

$$\iint A(f(x,v))\,dx\,dv,$$

where A is arbitrary. These millions of conservation laws are immediately deduced from (2.3); in other words, they express the fact that the Vlasov equation induces a transport by a measure-preserving (in fact Hamiltonian) flow. In particular, all L^p norms are preserved, the supremum is preserved... and so is the **entropy**:

$$S = -\iint f \log f\,dx\,dv.$$

The latter property is in sharp contrast with the Boltzmann equation, for which the entropy can only increase in time, unless it is at equilibrium. Physically speaking, it reflects the *preservation of information*: whatever information we have about the distribution of particles at initial time, is preserved at later times.

• The equation is time-reversible: choose an initial datum f_i, let it evolve by the nonlinear Vlasov equation from time 0 to time T, then reverse velocities (that is replace $f(T,x,v)$ by $f(T,x,-v)$) let it evolve again for an additional time T, reverse velocities again, and you are back to the initial datum f_i. This again is in contrast with the time-irreversibility of the Boltzmann equation. As a consequence, the nonlinear Vlasov equation does not have any regularizing effect, at least in the usual sense.

• The last identity is called the **virial theorem**; it only holds in the whole space and for specific classes of interaction. The virial[1] is defined as

$$\mathcal{V} = \iint f(x,v)\,x\cdot v\,dx\,dv$$

is the time-derivative of the inertia

$$I = \iint f(x,v)\,\frac{|x|^2}{2}\,dx\,dv.$$

If the potential W is even and λ-homogeneous, that is, for any $z \in \mathbb{R}^d$ and $\alpha \neq 0$,

$$W(-z) = W(z), \qquad W(\alpha z) = |\alpha|^\lambda\,W(z),$$

[1]This word was made up by Clausius using the latine root for "force".

then one has the virial identity

$$\frac{d\mathcal{V}}{dt} = 2T - \lambda U.$$

The most famous case of application is of course the case of Coulomb/Newton equation, for which $\lambda = -1$, which yields

$$\frac{d\mathcal{V}}{dt} = 2T + U.$$

When one takes a time-average and looks over large times, the contribution of the time-derivative is likely to disappear, and we are left with the plausible guess

$$2\langle T\rangle + \langle U\rangle = 0, \tag{2.4}$$

where $\langle u\rangle = \lim_{T\to\infty} T^{-1}\int_0^T u(t)\,dt$. Identity (2.4) suggests some kind of biased, but universal partition between the kinetic and potential energies.

4. Equilibria

A famous property of the Boltzmann equation is that it only has Gaussian equilibria. In contrast, the Vlasov equation has infinitely many shapes of equilibria.

First of all, any distribution $f(x,v) = f^0(v)$ defines a **spatially homogeneous equilibrium**. Indeed, $v\cdot\nabla_x f^0 = 0$, and the density ρ^0 associated to f^0 is constant, so the corresponding force vanishes ($\nabla W * \rho^0 = W * (\nabla\rho^0) = 0$).

The construction of other classes of equilibria is easy by means of the so-called Jeans theorem: *any function of the invariants of the flow is an equilibrium.* As the most basic example, let us search for a stationary f in the form of a function of the microscopic energy

$$E(x,v) = \frac{|v|^2}{2} + \Phi(x), \qquad \Phi = W * \rho,$$

where $\rho = \int f\,dv$. Using the ansatz $f(x,v) = \overline{f}(E)$, where $\overline{f}$ is an arbitrary function $\mathbb{R}\to\mathbb{R}_+$, we get by chain-rule

$$v\cdot\nabla_x f - \nabla\Phi\cdot\nabla_v f = (\overline{f})'(E)\,\big[v\cdot\nabla\Phi - \nabla\Phi\cdot v\big] = 0,$$

so f is an equilibrium.

Of course this works only if the potential Φ is indeed induced by f, which leads to the compatibility condition

$$\int \overline{f}\left(\frac{|v|^2}{2} + \Phi(y)\right) W(x-y)\,dy\,dv = \Phi(x).$$

For a given $\overline{f}$ this is a nonlinear integral equation on the unknown Φ; in the general case it is certainly too hard to solve, but if we are looking for solutions with symmetries, depending on just one parameter, this can often be done in practice.

If W is the Coulomb or Newton potential, the integral equation transforms into a differential equation; as a typical situation, consider the three-dimensional gravitational case with radial symmetry, then ρ and Φ are functions of r, and we have after a few computations

$$\rho(r) = 4\pi \int_{\Phi(r)}^{0} \sqrt{2(E-\Phi(r))}\,\overline{f}(E)\,dE.$$

This gives ρ as a function of Φ, and then the formulas for spherical Laplace operator applied to radial functions yield

$$\frac{1}{r^2}\frac{d}{dr}(r^2\,\Phi'(r)) = 4\pi\,\rho(\Phi),$$

whence $f(x,v) = \overline{f}(|v|^2/2 + \Phi(r))$ can be reconstructed.

Another typical situation is the one-dimensional Coulomb interaction with periodic data: then the equation is

$$-\Phi''(x) = \int \overline{f}\left(\frac{v^2}{2} + \Phi(x)\right) - 1,$$

subject to the condition $\int \overline{f}(v^2/2 + \Phi(x))\,dv = 1$. Such a solution is called a **BGK equilibrium**, after Bernstein, Greene and Kruzkal; or BGK wave, to emphasize the periodic nature of the solution. Such waves exist as soon as $\overline{f}$ is smooth and decays fast enough at infinity, and satisfies $\int \overline{f}(v^2/2)\,dv = 1$.

5. Speculations

The general concern by physicists is about the large time asymptotics, $t \to \infty$. Can one somehow draw a picture of the possible qualitative behavior of solutions to the nonlinear Vlasov equations?

Usually a first step in the understanding of the large-time behavior is the identification of stable structures such as equilibria. In the present case, the abundance of equilibria is a bit disorienting, and we would like to find selection criteria allowing to make predictions in large time.

Are equilibria stable? There is a convincing stability criterion for homogeneous equilibria, due to Penrose, which will be studied in Chapter 3. But no such thing exists for BGK waves, and nobody has a clue whether these equilibria should be stable or unstable.

Having no convincing answer to the previous question, we may turn to an even more difficult question, that is, which equilibria are attractive? Can one witness convergence to equilibrium even in the absence of dissipative features in the equation? Does the Vlasov equation exhibit **non-entropic relaxation**, that is, relaxation without increase of entropy? This has been the object of considerable debate, and suggested by numerical experiments on the one hand, observation on the other hand: as pointed out by the astrophysicist Lynden-Bell in the sixties, galaxies, roughly speaking, seem to be in equilibrium at relevant scales, although the relaxation times associated with entropy production in galaxies exceed by far the age of the universe. Lynden-Bell argued that there should be a mechanism of **violent relaxation**, of which nobody has a decent understanding.

If the final state is impossible to predict, maybe this problem can be attacked in a statistical way: Lynden-Bell and followers argued that some equilibria, in particular those having high entropy, may be favored by statistical considerations. Maybe there are invariant measures on the space of solutions of the nonlinear Vlasov equation, which can be used to statistically predict the large-time behavior of solutions??

In all this maze of speculations, questions and religions, the only tiny island on which we can stand on our feet is the **Landau damping** phenomenon: a relaxation property near stable equilibria, which is driven by conservative phenomena. In the sequel I shall describe this phenomenon in great detail; for the moment let me emphasize that besides its theoretical and practical importance by itself, it is the only serious theoretical hint of the possibility of dissipation-free relaxation in confined systems, without appealing to an extra randomness assumption.

Bibliographical notes

I am not aware of any good synthetic introductory source for boundary conditions of the nonlinear Vlasov equations; but this topic is discussed for instance in the research paper [**40**]. Boundary conditions for kinetic equations are also evoked in [**23**, Chapter 8] or [**99**, Section 1.5]. The Cauchy problem for Vlasov–Poisson in a bounded convex domain is studied in [**49**]; for nonconvex domain it is expected that serious issues arise about the regularity.

The Jeans swindle appears in many textbooks in astrophysics to justify asymptotic expansions when the density is a perturbation of a uniform constant in the whole space, see e.g. [**14**]. The underlying mathematical meaning of the procedure is neatly explained by Kiessling

[**52**]. The explanations given in Section 1 are just an adaptation of the argument to the periodic situation.

The Hamiltonian nature of the nonlinear Vlasov equation, in relation with optimal transport theory, is discussed informally in my introductory book on optimal transport [**100**, Section 8.3.2], and more rigorously by Ambrosio & Gangbo [**4**], and Lott [**62**, Section 6]. Some of these features are shared by other partial differential equations, in particular the two-dimensional incompressible Euler equation, for which a good concise source is [**68**]. The similarity between the one-dimensional Vlasov equation and the two-dimensional Euler equation with nonnegative vorticity is well-known; physicists have systematically tried to adapt tools and theories from one equation to the other.

The statistical meaning of the entropy, and its relation to the Boltzmann formula $S = k \log W$ is discussed in many sources; a concise account can be found in my tribute to Boltzmann [**101**].

Formal properties of the Vlasov equation, including the virial theorem, are covered in many textbooks such as Binney & Tremaine [**14**]. This reference also discusses the procedure for constructing inhomogeneous equilibria.

BGK waves were introduced in the seminal paper [**13**] and have been the object of many speculations in the literature; see [**57, 58**] for a recent treatment. No BGK wave has been proven to be stable with respect to periodic perturbations (that is, whose period is equal to the period of the wave). The only known related statement is the instability against perturbations with period twice as long [**57, 58**]. (This holds in dimension 1, but can probably be translated into a multidimensional result.) At least this means that a BGK wave f on $\mathbb{T} \times \mathbb{R}$ cannot be hoped to be stable if f is $1/2$-periodic in x.

The idea of violent relaxation was introduced in the sixties by Lynden-Bell [**63, 64**], who at the same time founded the statistical theory of the Vlasov equation. The theory has been pushed by several authors, and also adapted to the two-dimensional incompressible Euler equation [**24, 70, 87, 96, 97, 104**]. Since it is based on purely heuristic grounds and on just the conservation laws satisfied by the Vlasov equation (not on the equation itself), the statistical theory has been the object of criticism, see e.g. [**51**].

The construction of invariant measures on infinite-dimensional Hamiltonian systems has failed for classical equations such as the Vlasov or (two-dimensional, positive vorticity, incompressible) Euler equations [**85**]; but it was solved for certain dispersive equations, such as the cubic nonlinear Schrödinger equations, treated by Bourgain [**19**]. As far as the Vlasov or Euler equations are concerned, there is no canonical choice of what could be a Gibbs measure, but now there might be hope with Sturm's construction of a fascinating canonical "entropic measure" on the space of probability measures [**94**], coming from the theory of optimal transport. But for the moment very little is known about Sturm's measure, and measures drawn according to this measure are not even absolutely continuous.

CHAPTER 3

Linearized Vlasov equation near homogeneity

Vlasov, Landau and other pioneers of kinetic theory of plasmas discovered a fundamental property: when one linearizes the Vlasov equation around a homogeneous equilibrium, the resulting linear equation is "explicitly" solvable; in a way this is a **completely integrable** system. This allowed Landau to solve the stability and asymptotic behavior for the linearized equation — two problems which seem out of reach now for inhomogeneous equilibria.

0. Free transport

As a preliminary, let us study the properties of free transport, that is, when there is no interaction ($W = 0$):

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f = 0. \tag{3.1}$$

The properties of this equation differ much in the whole space $\mathbb{R}^d$ and in the confined periodic space $\mathbb{T}^d$. In the former case, dispersion at infinity dominates the large-time behavior, while in the latter case one observes **homogenization** phenomena due to **phase mixing** as illustrated in Fig. 3.1.

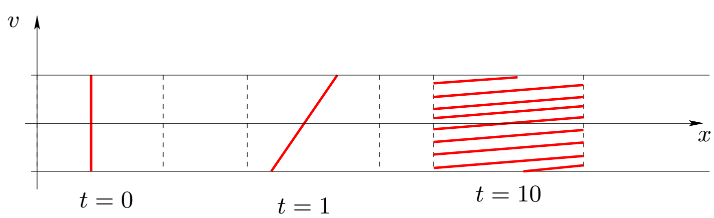

FIGURE 3.1. Put an initial disturbance along a line at $t = 0$. As time goes by, the free transport evolution makes this line twist and homogenize very fast.

Phase mixing occurs for mechanical systems expressed in action-angle variables when the angular velocity genuinely changes with the action variable. In the present case, the angular variable is the position, so the angular velocity is the plain velocity, which coincides precisely with the action variable.

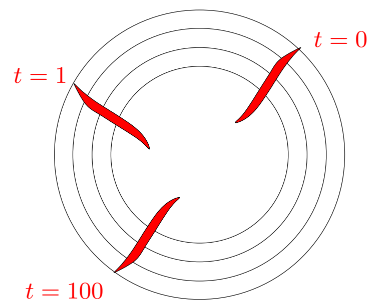

FIGURE 3.2. An example of a system which is not mixing: for the harmonic oscillator (linearized pendulum) the angular velocity is independent of the action variable, so a disturbance in phase space keeps the same shape as time goes by.

The free transport equation can be solved explicitly (which should not prevent us from keeping the qualitative picture in mind): if f_i is the datum at $t = 0$, then

$$f(t, x, v) = f_i(x - vt, v) \tag{3.2}$$

To study fine properties of this solution, it is most convenient to use the **Fourier transform**. Let us introduce the position-velocity Fourier transform

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

where $k \in \mathbb{Z}^d$ is dual to $x \in \mathbb{T}^d$, and $\eta \in \mathbb{R}^d$ is dual to $v \in \mathbb{R}^d$. Then (3.2) implies

$$\widetilde{f}(t, k, \eta) = \iint f_i(x - vt, v)\, e^{-2i\pi k\cdot x}\, e^{-2i\pi\eta\cdot v}\, dx\, dv \tag{3.3}$$

$$= \iint f_i(x, v)\, e^{-2i\pi k\cdot(x+vt)}\, e^{-2i\pi\eta\cdot v}\, dx\, dv \tag{3.4}$$

$$= \widetilde{f}_i(k, \eta + kt). \tag{3.5}$$

We deduce that

- $\widetilde{f}(t, 0, \eta) = \widetilde{f}_i(0, \eta)$: the zero spatial mode of f is preserved;

- for fixed η and $k \neq 0$, $\widetilde{f}_i(k, \eta + kt) \longrightarrow 0$ as $t \to \infty$, at a rate which is (a) determined by the smoothness of f_i in v(Riemann–Lebesgue lemma), (b) faster when k is large. In fact, the relevant time scale for the mode k is $|k|t$.

In particular, if f_i is analytic in v then $\widetilde{f}_i$ decays exponentially fast in η, so the mode k of the solution of the free transport equation will decay like $O(e^{-2\pi\lambda|k|t})$. Also, if f is only assumed to be Sobolev regular, say $W^{s,1}$ in the velocity variable for some $s > 0$, then the Fourier transform will decay like $O(|\eta|^{-s})$ at large values of $|\eta|$, so the mode of order k will decay like $O((|k|t)^{-s})$.

We can represent this behavior of the free transport equation, in Fourier space, as a *cascade* from low to high velocity modes, the cascade being faster for higher spatial modes.

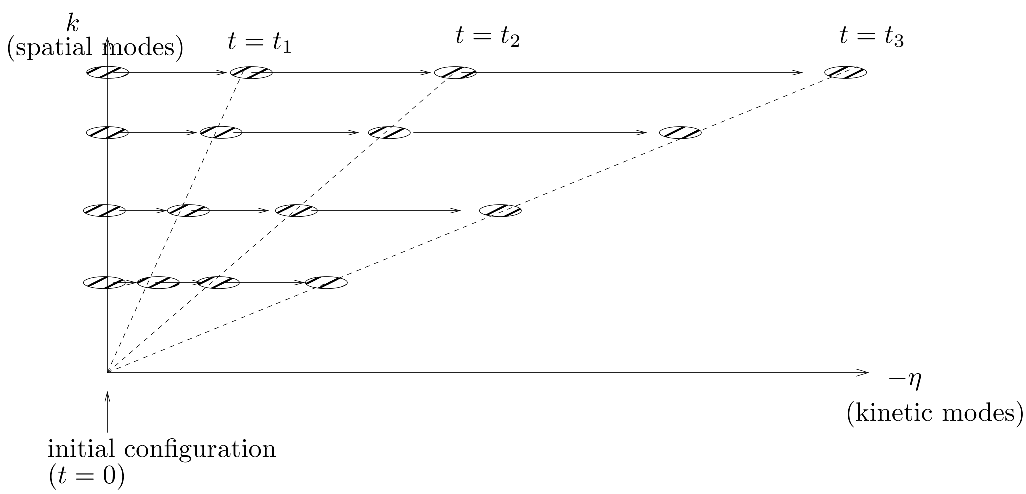

Figure 3.3. Schematic picture of the evolution of energy by free transport, or perturbation thereof; marks indicate localization of energy in phase space.

Remark 3.1. In view of this discussion, the free transport equation appears to be a natural way to convert regularity into a time decay, which can in principle be measured from a physical experiment!

Remark 3.2. Even though the free transport equation is reversible, there is a definite behavior as $t \to \pm\infty$: $\widetilde{f}(t, k, \eta)$ converges to 0 for all $k \neq 0$, and to $\widetilde{f}_i(0, \eta)$ for $k = 0$. This means that

$$f(t, x, v) \xrightarrow[t\to\infty]{\text{weakly}} \langle f_i \rangle,$$

where the brackets stand for spatial average:

$$\langle h \rangle(v) = \int h(x,v)\,dx.$$

The convergence holds as long as the initial measure does have a density, that is, f_i is well-defined as an integrable function; and it is faster if f_i is smooth.

Remark 3.3. Why don't we see such phenomena as recurrence, which are associated with confined mechanical systems? The answer is that as soon as the distribution is spread out and has a density, we do not expect such phenomena because the system truly is infinite-dimensional. Recurrence would occur with a singular distribution function, say Dirac masses, but we ruled out this situation.

1. Linearization

Now let us go back to the Vlasov equation. Let $f^0 = f^0(v)$ be a homogeneous equilibrium. We write $f(t,x,v) = f^0(v) + h(t,x,v)$, where $\|h\| \ll 1$ in some sense. Since f^0 does not contribute to the force field, the nonlinear Vlasov equation becomes

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

where

$$F[h](t,x) = -\iint \nabla W(x-y)\,h(t,y,w)\,dy\,dw = -\nabla_x W *_{x,v} h.$$

When h is very small we expect the quadratic term $F[h] \cdot \nabla_v h$ to be negligible in front of the linear terms, and obtain

$$\frac{\partial h}{\partial t} + v \cdot \nabla_x h + F[h] \cdot \nabla_v f^0 = 0. \tag{3.6}$$

The physical interpretation of (3.6) is not so obvious. Assume that we have two species of particles, one that has distribution h and the other one that has distribution f^0, and that the h-particles act on the f^0-particles by forcing, still they are unable to change the distribution f^0 (like you are pushing a wall, to no effect). In this case, we can imagine that the changes in the f^0 density would be compensated by the transmutation of h-particles into f^0-particles, or the reverse. Then the equation for f^0 will be

$$F[h] \cdot \nabla_v f^0 = S, \tag{3.7}$$

where S is the source of f^0 particles, and thus the equation for h would be

$$\frac{\partial h}{\partial t} + v \cdot \nabla_x h = -S. \tag{3.8}$$

The combination of (3.7) and (3.8) implies (3.6). Thus, in some sense, equation (3.6) can be interpreted as expressing the **reaction** exerted by the "wall" f^0 on the particle density.

We note that the last term on the right-hand side of (3.6) has the form $F[h] \cdot \nabla_v f^0$, where $F[h]$ is a function of t and x, and $f^0(v)$ is a function of v. This property of separation of variables will be crucial. As a start, it implies the statement below.

PROPOSITION 3.4. *If $h = h(t, x, v)$ evolves according to the linearized Vlasov equation* (3.6)*, then the function $\langle h \rangle = \int h(t, x, v)\, dx$ depends only on v and not on t.*

An equivalent statement is that the linearized Vlasov equation has an infinite number of conservation laws: for any function $\psi(v)$, $\int h\, \psi\, dv\, dx$ is a conserved quantity.

PROOF OF PROPOSITION 3.4. First note that $\langle \nabla_x h \rangle = 0$ and $\langle F[h] \rangle = 0$, since $F[h]$ is a gradient. So (3.6) implies

$$\begin{aligned} \partial_t \langle h \rangle = -\langle v \cdot \nabla_x h \rangle - \big\langle F[h] \cdot \nabla_v f^0 \big\rangle \\ = -v \cdot \nabla_x \langle h \rangle - \langle F[h] \rangle \cdot \nabla_v f^0 = 0. \end{aligned}$$

□

2. Separation of modes

Let us now work on the linearized equation, in the form

$$\frac{\partial h}{\partial t} + v \cdot \nabla_x h + F[h] \cdot \nabla_v f^0 = 0. \tag{3.9}$$

Solving this equation is a beautiful exercice in linear partial differential equations, involving three ingredients (whose order does not matter much): the method of characteristics, the integration in v, and the Fourier transform in x.

- **First step: the method of characteristics.** We apply the Duhamel principle to (3.9), treating it as a perturbation of free transport. It is easily checked that the solution of $\partial_t h + v \cdot \nabla_x h = -S$ takes the form

$$h(t, x, v) = h_i(x - vt, v) - \int_0^t S(\tau, x - v(t - \tau), v)\, d\tau,$$

where $h_i(x,v) = h(0,x,v)$.

• **Second step: Fourier transform.** Taking Fourier transform in both x and v yields

$$\begin{aligned}
\widetilde{h}(t,k,\eta) &= \iint h_i(x-vt,v)\, e^{-2i\pi k\cdot x}\, e^{-2i\pi\eta\cdot v}\, dx\, dv \\
&\qquad - \iint \int_0^t S(\tau, x - v(t-\tau), v)\, e^{-2i\pi k\cdot x}\, e^{-2i\pi\eta\cdot v}\, dx\, dv\, d\tau \qquad (3.10)\\
&= \iint h_i(x,v)\, e^{-2i\pi k\cdot(x+vt)}\, e^{-2i\pi\eta\cdot v}\, dx\, dv \qquad (3.11)\\
&\qquad - \iiint S(\tau,x,v)\, e^{-2i\pi k\cdot x}\, e^{-2i\pi k\cdot v(t-\tau)}\, e^{-2i\pi\eta\cdot v}\, dx\, dv\, d\tau \\
&= \widetilde{h}_i(k,\eta+kt) - \int_0^t \widetilde{S}\big(\tau,k,\eta+k(t-\tau)\big)\, d\tau,
\end{aligned}$$

where I used the measure-preserving change of variables $(x-vt,v)\to(x,v)$, and the obvious identity $k\cdot(vs) = v\cdot(ks)$ to absorb the time-shift into a change of arguments in the Fourier variables.

Now we note that the structure of separated variables in the term S and the properties of Fourier transform imply

$$\begin{aligned}
\widetilde{S}(\tau,k,\eta) &= \widehat{F}(\tau,k)\cdot \widetilde{\nabla_v f^0}(\eta) \\
&= (-\nabla W * \rho)\widehat{\ }(\tau,k)\cdot \widetilde{\nabla_v f^0}(\eta) \\
&= \big(-2i\pi k \widehat{W}(k)\, \widehat{\rho}(\tau,k)\big)\cdot\big(2i\pi\eta \widetilde{f}^0(\eta)\big) \\
&= 4\pi^2\, k\cdot\eta\, \widehat{W}(k)\, \widehat{\rho}^1(\tau,k)\, \widetilde{f}^0(\eta),
\end{aligned}$$

where $\rho^1(t,x) = \int h(t,x,v)\, dv$ is the first-order correction to the spatial density. Combining this with (3.10) we end up with

$$\begin{aligned}
\widetilde{h}(t,k,\eta) &= \widetilde{h}_i(k,\eta+kt) \qquad (3.12)\\
&\quad - 4\pi^2\, \widehat{W}(k) \int_0^t \widehat{\rho}^1(\tau,k)\, \widetilde{f}^0(\eta+k(t-\tau))\, k\cdot[\eta+k(t-\tau)]\, d\tau.
\end{aligned}$$

Third step: Integrate in v. This amounts to consider the Fourier mode $\eta = 0$ in (3.12):

$$\widehat{\rho}^1(t,k) = \widetilde{h}_i(k,kt) - 4\pi^2 \widehat{W}(k) \int_0^t \widehat{\rho}^1(\tau,k)\, \widetilde{f}^0(k(t-\tau))\, |k|^2\, (t-\tau)\, d\tau.$$

To recast it more synthetically:

$$\widehat{\rho}^1(t,k) = \widetilde{h}_i(k,kt) + \int_0^t K^0(t-\tau,k)\,\widehat{\rho}^1(\tau,k)\,d\tau, \tag{3.13}$$

where

$$K^0(t,k) = -4\pi^2\,\widehat{W}(k)\,\widetilde{f}^0(kt)\,|k|^2 t. \tag{3.14}$$

Now appreciate the sheer miracle: the Fourier modes $\widehat{\rho}^1(k)$, $k \in \mathbb{Z}$, evolve in time independently of each other! In a way this expresses a property of complete integrability, which can actually be made more formal.

Of course identity (3.14) is interesting only for $k \neq 0$; we already know that $\widehat{\rho}^1(t,0) = \widetilde{h}_i(0,0)$ is preserved in time.

3. Mode-by-mode study

If k is given, equation (3.13) is a **Volterra equation**, which in principle can be solved by **Laplace transform**. Generally speaking, if we have an equation of the form $\varphi = a + K * \varphi$, that is

$$\varphi(t) = a(t) + \int_0^t K(t-\tau)\,\varphi(\tau)\,d\tau,$$

then it can be changed, via the Laplace transform

$$\varphi^L(\lambda) = \int_0^\infty e^{2\pi\lambda t}\,\varphi(t)\,dt, \tag{3.15}$$

into the simple equation

$$\varphi^L = a^L + K^L\,\varphi^L,$$

whence

$$\varphi^L = \frac{a^L}{1-K^L}, \tag{3.16}$$

which is well-defined at $\lambda \in \mathbb{R}$ if $A^L(\lambda)$ and $K^L(\lambda)$ are well-defined (for instance if a and K decay exponentially fast and λ is small enough), and (careful!) if $K^L(\lambda) \neq 1$.

At this point it is useful to define the complex Laplace transform: for $\xi \in \mathbb{C}$,

$$\varphi^L(\xi) = \int_0^\infty e^{2\pi\xi^* t}\,\varphi(t)\,dt. \tag{3.17}$$

It is well-known that the reconstruction of φ from its Laplace transform involves integrating φ^L on a well-chosen contour in the complex plane, *which has to go out of the real line* and should be chosen appropriately.

Since Landau, many authors have discussed this tricky issue, by now very classical in plasma physics.

However, the reconstruction gives more information than we need: what we want is not the complete description of h, but its time-asymptotics. The following lemma will be enough to achieve this goal:

LEMMA 3.5. *Let $K = K(t)$ be a kernel defined for $t \geq 0$, such that*
(i) $|K(t)| \leq C_0\, e^{-2\pi\lambda_0 t}$;
(ii) $|K^L(\xi) - 1| \geq \kappa > 0$ for $0 \leq \mathcal{R}e\,\xi \leq \Lambda$.
Let further $a = a(t)$ satisfy $|a(t)| \leq \alpha\, e^{-2\pi\lambda t}$, and let φ solve the equation $\varphi = a + K * \varphi$. Then for any $\lambda' < \min(\lambda, \lambda_0, \Lambda)$,

$$|\varphi(t)| \leq C\,\alpha\, e^{-2\pi\lambda' t},$$

where $C = C(\lambda, \lambda', \Lambda, \lambda_0, \kappa, C_0)$.

Let us express this lemma in words: If the kernel K decays exponentially fast and satisfies the stability condition $K^L \neq 1$ on a strip of width $\Lambda > 0$ (see Fig. 3.4), then the solution φ decays in time at a rate which is limited only by the time-decay of the source, the time-decay of the kernel, and the width of the strip.

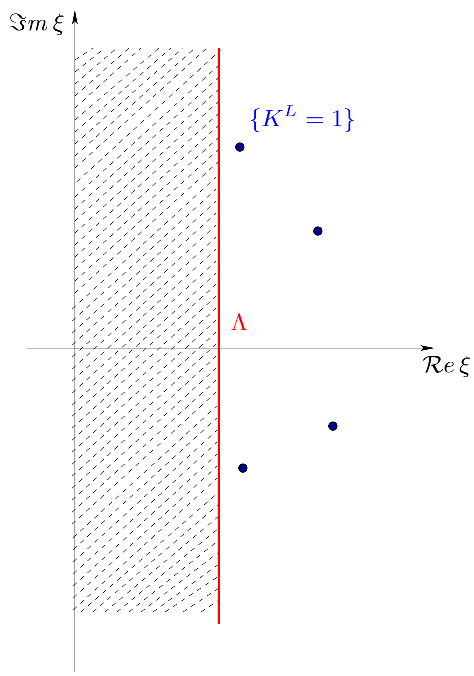

FIGURE 3.4. Λ is the width of a strip starting from the imaginary axis, containing no complex root of $\{K^L = 1\}$.

Proof of Lemma 3.5. Let us write $\Phi(t) = e^{2\pi\lambda' t}\,\varphi(t)$, $A(t) = e^{2\pi\lambda' t}\,a(t)$. The equation becomes

$$\Phi(t) = A(t) + \int_0^t K(t-\tau)\, e^{2\pi\lambda'(t-\tau)}\,\Phi(\tau)\,d\tau. \tag{3.18}$$

Extend Φ, A and K by 0 for $t \le 0$, then take Fourier transforms in the time variable: recalling the definition of the complex Laplace transform (3.17), this gives, for any $\omega \in \mathbb{R}$,

$$\widehat{\Phi}(\omega) = \widehat{A}(\omega) + K^L(\lambda' + i\omega)\,\widehat{\Phi}(\omega),$$

whence

$$\widehat{\Phi}(\omega) = \frac{\widehat{A}(\omega)}{1 - K^L(\lambda' + i\omega)}.$$

By assumption $|1 - K^L(\lambda' + i\omega)| \ge \kappa$, whence

$$\|\widehat{\Phi}\|_{L^2(d\omega)} \le \frac{\|\widehat{A}\|_{L^2(d\omega)}}{\kappa};$$

therefore, by Plancherel's identity and the decay assumption on A,

$$\|\Phi\|_{L^2(dt)} \le \frac{\|A\|_{L^2(dt)}}{\kappa} \le \frac{\alpha}{\kappa\,\sqrt{4\pi(\lambda - \lambda')}}.$$

Now plug this back in the equation (3.18), to get

$$\begin{aligned}\|\Phi\|_{L^\infty(dt)} &\le \|A\|_{L^\infty(dt)} + \left\|(K\,e^{2\pi\lambda' t}) * \Phi\right\|_{L^\infty(dt)}\\ &\le \|A\|_{L^\infty(dt)} + \|K\,e^{2\pi\lambda' t}\|_{L^2(dt)}\,\|\Phi\|_{L^2(dt)}\\ &\le \alpha + \frac{C_0}{\sqrt{4\pi(\lambda_0 - \lambda')}}\;\frac{\alpha}{\kappa\,\sqrt{4\pi(\lambda - \lambda')}},\end{aligned}$$

whence the desired result. □

Remark 3.6. It seems that I did not use the stability assumption in the whole strip $0 \le \xi \le \Lambda$, but only in a small strip near $\mathcal{R}e\,\xi = \Lambda$. But in fact I have *cheated* in the above proof, because I did not check that $\widehat{\Phi}(\omega)$ is well-defined. Making the reasoning rigorous will make the condition come back via a continuity argument. Further note that under appropriate decay conditions at infinity, $(1 - K^L)^{-1}$, if well-defined as a holomorphic function on the strip of width Λ, has maximum modulus near the axes $\mathcal{R}e\,\xi = 0$ and $\mathcal{R}e\,\xi = \Lambda$.

Let us apply Lemma 3.5 to (3.13). The kernel $K^0(t,k)$ decays as a function of t, exponentially fast if f^0 is analytic, more precisely like $O(e^{-2\pi\lambda_0|k|t})$. (The important remark is that time appears through $|k|t$.) Similarly, the source term $\widetilde{h}_i(k, kt)$ is $O(e^{-2\pi\lambda_i|k|t})$ if h_i is analytic.

So in order to ensure the exponential decay of $\widehat{\rho}^1(t,k)$ like $O(e^{-2\pi\lambda'|k|t})$, it only remains to check that

$$0 \le \mathcal{R}e\,\xi \le \lambda_L|k| \Longrightarrow |(K^0)^L(\xi) - 1| \ge \kappa > 0. \tag{3.19}$$

When that condition is satisfied, $\widehat{\rho}^1(t,k)$ converges to 0 at a rate which is exponential, *uniformly for* $|k| \ge 1$, so $\rho^1(t,\cdot)$ converges exponentially fast to its mean, and the associated force $F[h]$ converges exponentially fast to 0; this phenomenon is called **Landau damping**. For mnemonic means, you can figure it in the following way: if you keep pushing on a wall, the wall will not move and you will exhaust itself.

Before going on, note that the conclusion would be different if the position space X^d was the whole space $\mathbb{R}^d$ rather than $\mathbb{T}^d$: then the spatial mode k would live in $\mathbb{R}^d$ rather than $\mathbb{Z}^d$ (no "infrared cutoff"), and there would be no uniform lower bound for the convergence rate when k becomes small. As a matter of fact, counterexamples by Glassey and Scheffer show that the exponential damping of the force does not hold true in natural norms if $X^d = \mathbb{R}^d$, f^0 is a Gaussian and the interaction is Coulomb. Numerical computations by Landau suggest that the Landau damping rate in a periodic box of length ℓ decays extremely fast with ℓ, like $\exp(-c/\ell^2)$.

In the sequel, I shall continue to stick to the case when the position space is $\mathbb{T}^d$.

4. The Landau–Penrose stability criterion

Of course, the previous computation is hardly a solution of the problem, because the stability criterion (3.19) is only indirectly linked to the form of the distribution function f^0. Now the problem is to find more explicit stability conditions expressed in terms of f^0.

As a start, let us assume $d = 1$. Let us also rescale time by a factor $|k|$; the Laplace transform of $K^0(t,k)$, evaluated at $(\lambda - i\omega)|k|$, is, by integration by parts in the v variable,

$$
\begin{aligned}
&\int_0^\infty e^{2\pi(\lambda+i\omega)|k|t}\, K^0(t,k)\,dt \\
&\qquad = -4\pi^2 \widehat{W}(k) \int_0^\infty \int_{\mathbb{R}} f^0(v)\, e^{-2i\pi|k|tv}\, e^{2\pi(\lambda+i\omega)|k|t}\, |k|^2\, t\, dv\, dt \\
&\qquad = 2i\pi\, \widehat{W}(k) \int_0^\infty \int_{\mathbb{R}} (f^0)'(v)\, e^{-2i\pi|k|tv}\, e^{2\pi(\lambda+i\omega)|k|t} |k|\, dv\, dt \\
&\qquad = 2i\pi\, \widehat{W}(k) \int_{\mathbb{R}} (f^0)'(v) \left(\int_0^\infty e^{-2i\pi|k|tv}\, e^{2\pi(\lambda+i\omega)|k|t}\, |k|\, dt \right) dv \\
&\qquad = \widehat{W}(k) \int_{\mathbb{R}} \frac{(f^0)'(v)}{v-\omega+i\lambda}\, dv, \qquad (3.20)
\end{aligned}
$$

where I have used the formula for the generalized Laplace transform of a complex exponential. (This is justified if $(f^0)'$ decays fast enough at infinity.) The final result is an integral transform of $(f^0)'$, sometimes called Cauchy transform.

As soon as $(f^0)'(v) = O(1/|v|)$, the expression in (3.20) decays like $O(1/|\omega|)$ as $|\omega| \to \infty$, uniformly for $\lambda \in [0, \lambda_0]$; so if we wish to check that this expression does not approach 1, we can restrict ω to a bounded interval $|\omega| \le \Omega$. If in the limit $\lambda \to 0^+$ (3.20) does not approach 1, then by uniform continuity we can find $\Lambda > 0$ such that (3.20) does not approach 1 throughout the domain $\{|\omega| \le \Omega, 0 \le \lambda \le \Lambda\}$, and thus throughout the strip $0 \le \lambda \le \Lambda$. So let us focus on the limit $\lambda \to 0^+$.

From (3.20) we deduce

$$
\begin{aligned}
(K^0)^L((\lambda+i\omega)|k|) \xrightarrow[\lambda\to 0^+]{} \; & \widehat{W}(k) \int \frac{(f^0)'(v)}{v-\omega+i0}\, dv \\
&= \widehat{W}(k) \left[\text{p.v.} \int \frac{(f^0)'(v)}{v-\omega}\, dv \;-\; i\pi (f^0)'(\omega) \right] \\
&=: Z(k,\omega).
\end{aligned}
$$

Here I have used the so-called **Plemelj formula**,

$$
\frac{1}{z+i0} = \text{p.v.} \left(\frac{1}{z}\right) - i\pi\, \delta_0,
$$

which has become a standard in plasma physics. The abbreviation p.v. stands for principal value, that is, simplifying the possibly divergent part by symmetry around the vanishing of the numerator; in simple-minded terms,

$$
\text{p.v.} \int \frac{(f^0)'(v)}{v-\omega}\, dv = \int \frac{(f^0)'(v) - (f^0)'(\omega)}{v-\omega}\, dv.
$$

Now the goal is to find conditions so that Z does not approach 1. If the imaginary part of $Z(k, \omega)$ stays away from 0, then of course Z does not approach 1. But the imaginary part can approach 0 only if $\widehat{W}(k)$ approaches 0 (as $k \to \infty$), in which case the real part will also approach 0; or if $\omega \to \infty$, in which case the real part will also approach 0; or if ω approaches a zero of $(f^0)'$. So we only need to worry about zeros of $(f^0)'$, the problem becomes compact, and we have obtained a simple criterion for stability:

$$\forall \omega \in \mathbb{R}, \qquad (f^0)'(\omega) = 0 \implies \widehat{W}(k) \int \frac{(f^0)'(v)}{v - \omega}\, dv < 1. \tag{3.21}$$

This is the **Penrose stability condition**.

Example 3.7. Consider the Newton interaction, $\widehat{W}(k) = -1/|k|^2$, with a Gaussian distribution

$$f^0(v) = \rho^0 \sqrt{\frac{\beta}{2\pi}}\, e^{-\beta v^2/2}.$$

Then (3.21) is satisfied if $\rho^0 \beta < |k|^2$ for all $k \neq 0$, that is if $\beta < 1/\rho^0$: the Gaussian should be *spread enough* to be stable. In physics, there is a multiplicative factor $\mathcal{G}$ in front of the potential, the temperature $T = \beta^{-1}$ is typically given and determines the spreading of the distribution, the density is given, but one can change the size of the periodic box by performing a rescaling in space: the result is that the stability condition is satisfied if and only if $L < L_J$, where L_J is the so-called **Jeans length**,

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

It is widely accepted that this is a typical instability length for the Newtonian Vlasov–Poisson equation, which determines the typical length scale for the inter-galactic separation distance, and thus provides a qualitative answer to the basic question "Why are stars forming clusters (galaxies) rather than a uniform background?"

Example 3.8. Consider the Coulomb interaction, $\widehat{W}(k) = 1/|k|^2$. If f^0 has only one maximum at the origin, and is nondecreasing for $v < 0$, nonincreasing for $v > 0$ (for brevity we say that f^0 is increasing/decreasing), then obviously

$$\int \frac{(f^0)'(v)}{v}\, dv < 0,$$

and (3.21) trivially holds true, *independently of the length scale*. This is the **Landau stability criterion**.

EXAMPLE 3.9. If f^0 is a small perturbation of an increasing/decreasing distribution, so that it has a slight secondary bump, then the Landau criterion will no longer hold, but the Penrose criterion will still be satisfied, and linear stability will follow. If the bump becomes larger, there will be linear instability (**bump-on-tail instability**, or **two-stream instability**).

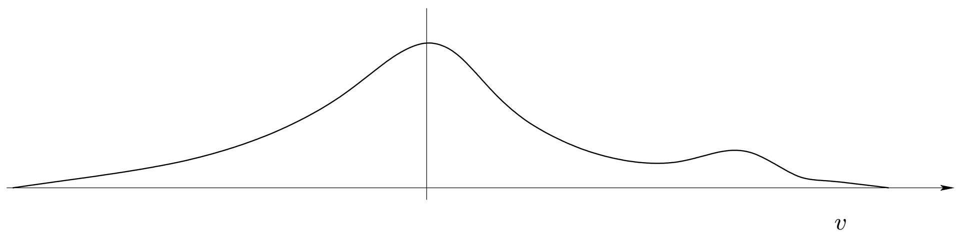

FIGURE 3.5. Bump-on-tail instability: For a given length of the box, a large enough secondary bump in the distribution function implies a linear instability. Conversely, if a non-monotone velocity distribution is given, there will be instability when the size of the box is large enough.

Now let us turn to the multidimensional setting. If $k \in \mathbb{Z}^d \setminus \{0\}$ and $\xi \in \mathbb{C}$, we use the splitting

$$v = \frac{k}{|k|}\, r + w, \qquad w \perp k, \qquad r = \frac{k}{|k|} \cdot v$$

to rewrite

$$(K^0)^L(\xi, k) = -4\pi^2\, \widehat{W}(k)\, |k|^2 \int_0^\infty \int_{\mathbb{R}^d} f^0(v)\, e^{-2i\pi k t \cdot v}\, t\, e^{2\pi \xi^* t}\, dv\, dt$$

$$= -4\pi^2 \widehat{W}(k) \int_0^\infty \int_{\mathbb{R}} \left(\int_{\frac{k}{|k|} r + k^\perp} f^0\left(\frac{k}{|k|} r + w \right) dw \right) e^{-2i\pi |k| r t} t\, e^{2\pi \xi^* t}\, dr\, dt,$$

where $k^\perp$ is the hyperplane orthogonal to k. So everything is expressed in terms of the one-dimensional marginals of f^0. If f is a given function of $v \in \mathbb{R}^d$, and σ is a unit vector, let us write $\sigma^\perp$ for the hyperplane orthogonal to σ, and

$$\forall v \in \mathbb{R} \qquad f_\sigma(v) = \int_{v\sigma + \sigma^\perp} f(w)\, dw. \tag{3.22}$$

Then the computation above shows that the multidimensional stability criterion reduces to the one-dimensional criterion in each direction $k/|k|$. Let us formalize this:

DEFINITION 3.10 (Penrose's stability criterion). We say that $f^0 = f^0(v)$ satisfies the (generalized) Penrose stability criterion for the interaction potential W if for any $k \in \mathbb{Z}^d$, and any $\omega \in \mathbb{R}$,

$$(f^0_\sigma)'(\omega) = 0 \implies \widehat{W}(k) \int \frac{(f^0_\sigma)'(v)}{v-\omega}\, dv < 1, \qquad \sigma = \frac{k}{|k|}.$$

EXAMPLE 3.11. The multidimensional generalization of Landau's stability criterion is that all marginals of f^0 are increasing/decreasing.

EXAMPLE 3.12. If f^0 is radially symmetric and positive, and $d \geq 3$, then all marginals of f^0 are decreasing functions of $|v|$. Indeed, if $\varphi(v) = \int_{\mathbb{R}^{d-1}} f(\sqrt{v^2+|w|^2})\, dw$, then after differentiation and integration by parts we find

$$\begin{cases} \varphi'(v) = -(d-3)\, v \displaystyle\int_{\mathbb{R}^{d-1}} f\big(\sqrt{v^2+|w|^2}\big)\, \frac{dw}{|w|^2} & (d \geq 4) \\ \varphi'(v) = -2\pi\, v\, f(|v|) \qquad (d=3). \end{cases}$$

5. Asymptotic behavior of the kinetic distribution

Let us assume stability, so that the force $F[h]$ converges to 0 as $t \to \infty$, exponentially fast in an analytic setting. What happens to h itself?

Starting again from

$$(3.23)\quad \widetilde{h}(t,k,\eta) = \widetilde{h}_i(k, \eta + kt) - 4\pi^2\, \widehat{W}(k) \int_0^t \widehat{\rho}^1(\tau,k)\, \widetilde{f}^0(\eta + k(t-\tau))\, k \cdot \big[\eta + k(t-\tau)\big]\, (t-\tau)\, d\tau$$

we can control the integrand on the right-hand side by the bounds

$$|\widehat{\rho}(\tau,k)| = O(e^{-2\pi\lambda'|k|\tau}), \qquad |\widetilde{f}^0(\eta)| = O(e^{-2\pi\lambda'|\eta+k(t-\tau)|}).$$

Sacrificing a little bit of the τ-decay of $|\widehat{\rho}|$ to ensure the convergence of the τ-integral, using $|\eta+kt| \leq |\eta+k(t-\tau)| + |k\tau|$, and assuming $|k|\widehat{W}(k) = O(1)$ (which is true if $\nabla W \in L^1$), we end up with

$$\left| 4\pi^2\, \widehat{W}(k) \int_0^t \widehat{\rho}^1(\tau,k)\, \widetilde{f}^0(\eta+k(t-\tau))\, |k|^2\, (t-\tau)\, d\tau \right| = O\left(e^{-2\pi\lambda''|\eta+kt|}\right),$$

where λ'' is arbitrarily close to λ'. Plugging this back in (3.23) implies

$$(3.24)\qquad \big|\widetilde{h}(t,k,\eta) - \widetilde{h}_i(k,\eta+kt)\big| \leq C\, e^{-2\pi\lambda''|\eta+kt|}.$$

It is not difficult to show that the bound (3.24) is qualitatively optimal; it is interesting only for $k \neq 0$, since we already know $\widetilde{h}(t,0,\eta) = \widetilde{h}_i(0,\eta)$.

Let us analyze (3.24) as time becomes large. First, for each fixed (k, η) we have $\widetilde{h}(t, k, \eta) \longrightarrow 0$ exponentially fast, in particular

$$h(t, \cdot) \xrightarrow[\text{weakly}]{t \to \infty} \langle h_i \rangle,$$

and the speed of convergence is determined by the regularity in velocity space: exponential convergence for analytic data, inverse polynomial for Sobolev data, etc.

However, for each t one can find (k, η) such that $|\widetilde{h}(t, k, \eta)|$ is $O(1)$ (not small!). In other words, *the decay of Fourier modes is not uniform*, and the convergence is not strong. In fact, the spatial mode k of $h(t, \cdot)$ undergoes oscillations along the kinetic frequency $\eta \simeq -kt$ in the velocity variable as $t \to \infty$; so at time t, the typical oscillation scale in the velocity variable is $O(1/|k|t)$ for the mode k. How much large $|k|$ modes affect the whole distribution h depends on the respective strength of the modes, that is, on the regularity in the x variable; but in any case the kinetic distribution h will exhibit fast velocity oscillations at scales $O(1/t)$ as t goes by. The problem only arises in the velocity variable: it is an easy exercice to check that the smoothness in the position variable is essentially preserved.

However, if one considers h *along trajectories of free transport*, the smoothness is restored: (3.24) shows that $\widetilde{h}(t, k, \eta - kt)$ is bounded like $O(e^{-2\pi\lambda''|\eta|})$, so we do not see oscillations in the velocity variable any longer. Let us call this the **gliding regularity**: if we change the focus in time to concentrate on modes $\eta \simeq -kt$ in Fourier space, we do see a good decay. Equivalently, if we look at $h(x + vt, v)$, what we see is uniformly smooth as $t \to \infty$.

This point of view can also be given an appealing interpretation as a finite-time **scattering procedure**. As $t \to \infty$, the force field vanishes, so the linearized Vlasov equation is asymptotic to the free transport evolution. Now the idea is to let the distribution evolve according to the linearized Vlasov for time t, then apply the free evolution backwards from time t to initial time, and study the result. This is the same as one does in classical scattering theory, except that in scattering theory one would take the limit $t \to \infty$, and we prefer to have estimates that are uniform in time, not just in the limit.

6. Qualitative recap

Let me reformulate and summarize what we learnt in this section. I shall start with a precise mathematical statement.

THEOREM 3.13. *Let $f^0 = f^0(v)$ be an analytic homogeneous equilibrium, with $|\widetilde{f^0(\eta)}| = O(e^{-2\pi\lambda_0|\eta|})$, and let W be an interaction potential such that $\nabla W \in L^1(\mathbb{T}^d)$. Let K^0 be defined in* (3.14)*; assume that there is $\lambda_L > 0$ such that the Laplace transform $(K^0)^L(\xi, k)$ of $K^0(t, k)$ stays away from the value 1 when $0 \le \mathcal{R}e\, \xi < \lambda_L|k|$. Let $h_i = h_i(x, v)$ be an analytic initial perturbation such that $\widetilde{h}_i(k, \eta) = O(e^{-2\pi\lambda|\eta|})$. Then if h solves the linearized Vlasov equation* (3.6) *with initial datum h_i, one has exponential decay of the force field: for any $k \neq 0$,*

$$\widehat{F[h]}(t, k) = O(e^{-2\pi\lambda'|k|t}),$$

for any $\lambda' < \min(\lambda_0, \lambda_L, \lambda)$. Moreover, Penrose's stability condition (Definition 3.10) guarantees the existence of $\lambda_L > 0$.

REMARK 3.14. Following Landau, physics textbooks usually care only on λ_L and forget about λ_0, λ, assuming that f^0 and h_i are entire functions (so one can choose λ_0 and λ_L arbitrarily large). But in general one should not forget that the damping rate does depend on the analytic regularity of f^0 and h_i.

Beyond Theorem 3.13, one can argue that the three key ingredients leading to the decay of the force field are

- the confinement ensured by the torus;
- the mixing property of the geodesic flow $(x, v) \to (x + vt, v)$;
- the Riemann–Lebesgue principle converting smoothness into decay in Fourier space.

The first two ingredients are important: as I already mentioned, there are counterexamples showing that decay does not hold in the whole space, and it is rather well-known from experiments that damping may cease when the flow ceases to be mixing, so that for instance trapped trajectories appear. As for the third ingredient, it is subject to debate, since there are many points of view around as to *why* damping holds (wave-particle interaction, etc.), but in these notes I will advocate the Riemann–Lebesgue point of view as natural and robust.

Now as far as the regularity of h is concerned, one should keep in mind that

- the regularity of h deteriorates in the velocity variable, as it oscillates faster and faster in v as time increases;
- there is a cascade in Fourier space from low to high kinetic modes, which on the mean is faster for higher position modes — it is like a shear flow in Fourier space;

• the distribution function evaluated along trajectories of the free flow, $h(t, x + vt, v)$, remains very smooth, uniformly in time (gliding regularity).

While the regularity of h deteriorates in the kinetic variable, on the contrary, the regularity of the force field *increases with time*, since (in analytic regularity)

$$\widehat{F}(t,0) = 0, \qquad |\widehat{F}(t,k)| = O(e^{-2\pi\lambda|k|t}). \tag{3.25}$$

Of course this implies the time decay of F like $O(e^{-2\pi\lambda t})$, but (3.25) is much more precise by keeping track of the respective size of the various modes. A simplistic way to summarize these apparently conflicting behaviors is that there is *deterioration of the regularity in v, improvement of the regularity in x.*

In the study of the linearized equation, we can live without knowing all this qualitative information, and it is not surprising that it has apparently never been recorded in a fully explicit way. But this will become crucial in our analysis of the nonlinear equation.

Bibliographical notes

Landau [**56**] solved the linearized Landau equation by using the separation of modes and the Fourier–Laplace transform. His treatment, based on the inversion of the Laplace transform, has been reproduced in countably many sources [**1, 8, 14, 29, 43, 54, 61, 69, 89**]. The rigorous justification is somewhat tricky because inverting a Laplace transform is not such a simple matter and involves integration over a complex contour, which has to be chosen properly; in fact Vlasov had it wrong in this respect, and Landau was the first to identify this subtlety. The first completely rigorous treatment is due to Backus [**8**]. Morrison [**71**] formalized the complete integrability property of the system, thanks to the so-called R-transform, which is related to the Laplace inversion.

An alternative solution consists in expressing the solution as a combination of generalized eigenfunctions, called **Van Kampen modes** [**22, 61, 98**]. This reduces the stability analysis to the study of a dispersion equation, but this is even more complicated to justify.

The presentation adopted in this chapter, taken from [**74**], is much more elementary since it replaces the tricky Laplace inversion formula by the standard Fourier inversion formula. This simple but downright unorthodox treatment was suggested by a conversation with Sigal. The price to pay (but we don't really care) for this simple approach is that instead of having an exact representation of the solution, we just have an estimate on it. On the other hand, a considerable reward is that this estimate easily goes through the nonlinear study.

Counterexamples by Glassey and Scheffer, showing that there is no Landau damping in the whole space for the linearized Vlasov–Poisson equation, can be found in [**36, 37**]. Estimates of the decay rate at small wavelengths (large distances) are performed in [**56**] or [**61**, Section 32].

In 1960 Penrose [**82**] suggested that the violation of the criterion (3.21) would lead to instability. In particular, he argued that if the distribution function has a secondary bump (is nonmonotone) then the distribution is linearly unstable at large enough scales. Conversely, the Penrose criterion implies stability under small-scale perturbations. Lin and Zeng [**59**] have shown that the Penrose criterion is close to be a necessary and sufficient condition. Example 3.12 is taken from [**61**, Problem, Section 30] (in dimension $d = 3$).

The interpretation of the Jeans instability can be found in [**14**]; it does not work quantitatively so well to predict the typical galaxy diameter, because galaxies are not really a continuum of stars. In a phase diagram for galaxies, the Jeans length is a "spinodal" (metastability) point, which only gives an upper bound for the phase transition regime [**93**].

The scattering approach to Landau damping was considered by Caglioti and Maffei [**21**], and amplified in my work with Mouhot [**74**]. A self-contained treatment of the linearized Vlasov–Poisson equation can be found in Section 3 of the latter work.

Ryutov [**88**] mentions that interpretations of the Landau damping phenomenon were still regularly appearing fifty years after the discovery of this effect. The wave-particle interpretation is surveyed, sometimes critically, in the works of Elkens and Escande [**32, 33, 34**]

Belmont [**11**] noticed that the damping rate in the linearized equation depends not only on the Landau stability condition, but also on the regularity of f^0, so that "for special distribution functions" the Landau damping rate does not govern the damping of the force.

CHAPTER 4

Nonlinear Landau damping

The damping phenomenon discovered by Landau, and considered in the previous chapter, is based on the study of the linearized Vlasov equation. But the physical model, of course, is the nonlinear equation, so the question naturally arises whether damping still holds for that model, at least in the perturbative regime, that is, near a spatially homogeneous equilibrium.

1. Nonlinear stability?

Linear stability is often a necessary condition for nonlinear stability, but is it sufficient? Starting from the nonlinear Vlasov equation, we have implicitly considered two distinct asymptotic regimes: $t \to \infty$ and $\varepsilon \to 0$, where $\varepsilon = \|f_i - f^0\|$; and these two limits a priori do not commute! In large time, small cumulated nonlinear effects might lead to a significant departure from the linearized equation.

To estimate the time scale on which this may occur, let us look for a scale invariance of the Vlasov equation. Let us assume that $f = 1+h$ solves the Vlasov equation (forget the fact that f has infinite mass), and set

$$f_\varepsilon(t,x,v) = 1 + \varepsilon^{1+d\nu}\, h\big(\varepsilon^\theta t, x, \varepsilon^\nu v\big),$$

where ν and θ are unknown parameters. (We cannot rescale in x since we work with periodic boundary conditions.) Note that $f_\varepsilon - 1$ is of size ε in L^1 norm, and $\int (f_\varepsilon - 1)\, dv = O(\varepsilon)$. Then

$$\begin{aligned}\partial_t f_\varepsilon &+ v\cdot\nabla_x f_\varepsilon + F[f_\varepsilon]\cdot\nabla_v f_\varepsilon \\ &= \varepsilon^{1+d\nu}\Big[\varepsilon^\theta\,\partial_t h + \varepsilon^{-\nu}\,(v\cdot\nabla_x h) + \varepsilon^{1+\nu}\,(F\cdot\nabla_v h)\Big](\varepsilon^\theta\, t, x, \varepsilon^\nu\, v),\end{aligned}$$

so f_ε solves the Vlasov equation if $\theta = -\nu = 1+\nu$, i.e. $\theta = 1/2 = -\nu$. In other words,

$$f_\varepsilon(t,x,v) = 1 + \varepsilon^{1-\frac{d}{2}}\, h\left(\sqrt{\varepsilon}\, t, x, \frac{v}{\sqrt{\varepsilon}}\right) \tag{4.1}$$

also solves the Vlasov equation. This suggests the typical nonlinear time scale $O(1/\sqrt{\varepsilon})$, where ε is the size of the perturbation. This is the **O'Neil time scale** and it is indeed well satisfied in numerical experiments.

To summarize: after a time scale $O(1/\sqrt{\varepsilon})$ we expect the solution of the nonlinear Vlasov equation to be significantly different from the solution of the linearized approximation; in particular, it is not clear that the large-time limit of the linearization coincides with the linearization of the large-time limit.

The problem arises not only for the study of the damping, but also already for the a priori simpler stability problem: for many years it has been an open problem to show that Penrose's linear stability condition guarantees nonlinear stability. Only the simpler stability condition by Landau (monotone profile) could be treated by Lyapunov functional techniques.

2. Elusive bounds

The study of the linearized Vlasov equation $\partial_t h + v \cdot \nabla_x h + F[h] \cdot \nabla_v f^0 = 0$ showed that the expected decay rate depends (among other things) on the smoothness of f^0. In the nonlinear case we have $\partial_t f + v \cdot \nabla_x f + F[f] \cdot \nabla_v f = 0$, so the uniformly smooth background $f^0(v)$ is replaced by the time-dependent distribution $f(t,x,v)$, which may be analytic, but *not uniformly in time*: fast oscillations in the velocity variable, a phenomenon which is also known as **filamentation** in phase space, will imply the blow-up of all regularity bounds of f in large time. Then how can one hope to adapt the tools on which the linearized study was based??

3. Backus's objection

When one linearizes the Vlasov equation as in the beginning of the previous chapter, one neglects the quadratic term $F[h] \cdot \nabla_v h$ in front of the others, assuming implicitly that $\|h\|$ is much smaller than $\|f^0\|$ in some sense. However, $\nabla_v h$ *cannot stay small* in the usual sense: its norm will typically grow in time — unless we use a weak norm, but we are in a context where smoothness matters much. To see this growth, let us consider just the simpler free transport: if $\partial_t h + v \cdot \nabla_x h = 0$, then

$$\widetilde{\nabla_v h}(t,k,\eta) = 2i\pi\eta\, \widetilde{h}(t,k,\eta) = 2i\pi\eta\, \widetilde{h}_i(k, \eta + kt).$$

Then even if h_i is of size $\varepsilon \ll 1$, the choice $\eta \simeq -kt$ shows that

$$\sup_\eta \left| \widetilde{\nabla_v h}(t,k,\eta) \right| \geq \text{const.}\, \varepsilon |k| t$$

as $t \to +\infty$, so $\|\nabla_v h\|_{L^1(dx\,dv)}$ grows at least linearly in time. As a consequence, if we wait long enough, there will necessarily come a time when the linearization postulate is no longer satisfied! This objection was raised by Backus in 1960.

Of course, $F[h] \cdot \nabla_v h$ might still decay in time, since we expect $F[h]$ to decay exponentially fast, and $\nabla_v h$ to grow at most linearly. But right now we cannot understand why the effect of this term would be negligible compared to the effect of other terms like $F[h] \cdot \nabla_v f^0$, which we also expect to be decaying exponentially fast!

4. Numerical simulations

Numerical simulations about the large-time behavior of the nonlinear Vlasov–Poisson equation are nonconclusive because of the difficulties in getting reliable simulations on very large times. If ε is the size of the perturbation, nonlinear effects start to appear at time scale $O(1/\sqrt{\varepsilon})$, and then tiny numerical errors cumulated over very large times can be dreadful.

To summarize the situation, one can say that

- for a slight perturbation of the equilibrium, numerical schemes do display the Landau damping phenomenon for large times, and some of them continue to display damping at very large times, while other ones present tiny bumps of the electric field, which sometimes do not vanish as $t \to \infty$;
- for a larger perturbation of the equilibrium, numerical schemes agree that damping may be replaced by a much more complicated behavior, leading to a persistent electric field. Some authors claim to observe BGK waves in very large times, while others remain more cautious.

Francis Filbet kindly accepted to do a few precise simulations for me in the perturbative regime, with different methods; they led to different results, but the one that was supposed to be the most precise displayed damping at spectacular precision (more than 20 orders of magnitude for the amplitude of the electric field, and at times so large that the nonlinear effects can definitely not be neglected; see Figures 4.1 and 4.2).

5. Theorem

Some of the previous questions are solved by the following theorem by Mouhot and myself. If n is a multi-integer and f a function I shall write $f^{(n)} = \nabla^n f = \partial_1^{n_1} \dots \partial_d^{n_d} f$.

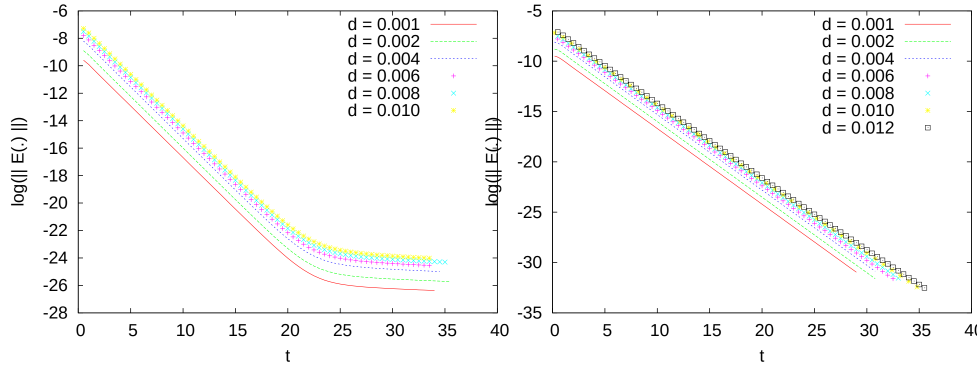

FIGURE 4.1. Large time behavior of the logarithm of the norm of the electric field, with two different numerical methods — the second one is supposed to be more precise. The interaction is gravitational, the initial datum is a Gaussian multiplied by $1 + \varepsilon \cos(2\pi kx)$.

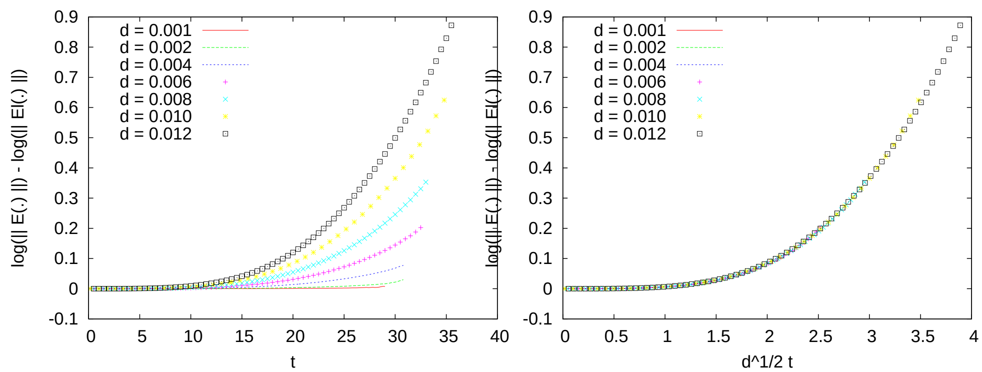

FIGURE 4.2. With the more precise method from Figure 4.1, a plot of $\log(\|E_{\rm NL}\|/\|E_L\|)$, the logarithmic ratio of the norm of the nonlinear electric field to the norm of the linearized electric field. On the left the time-scale is 1, on the right the time scale is $1/\sqrt{\varepsilon}$. Here we see that we arrive in a time regime where the nonlinearity can definitely not be neglected.

THEOREM 4.1. *Let $f^0 = f^0(v)$ be an analytic profile satisfying the Penrose linear stability condition. Further assume that the interaction potential W satisfies*

$$\widehat{W}(k) = O\left(\frac{1}{|k|^2}\right). \tag{4.2}$$

Then one has nonlinear stability and nonlinear damping close to f^0.

More precisely, assume that (a) f^0 *is analytic in a strip of width* $\lambda_0 > 0$, *in the sense that*

$$|\widetilde{f}^0(\eta)| \leq C_0\, e^{-2\pi\lambda_0|\eta|}, \qquad \sum_{n\in\mathbb{N}^d} \frac{\lambda_0^n}{n!} \|\nabla_v^n f^0\|_{L^1(dv)} \leq C_0;$$

(b) the Penrose linear stability condition is satisfied in a strip of width λ_L, *in the sense that, if* $K^0(t,k) = -4\pi^2\, \widehat{W}(k)\, \widetilde{f}^0(kt)\,|k|^2 t$, *then*

$$0 \leq \mathcal{R}e\,\xi \leq \lambda_L|k| \Longrightarrow |(K^0)^L - 1| \geq \kappa > 0. \tag{4.3}$$

(c) the initial condition f_i *is a perturbation of* f^0 *in a strip of width* $\lambda > 0$, *in the sense that*

$$|\widetilde{f}_i - \widetilde{f}_0|(k,\eta) \leq \varepsilon\, e^{-2\pi\mu|k|}\, e^{-2\pi\lambda|\eta|}, \qquad \int |f_i(x,v) - f^0(v)|\, e^{2\pi\beta|v|}\, dx\, dv \leq \varepsilon$$

for some $\mu > 0$, $\beta > 0$. *Then if* $\varepsilon \leq \varepsilon_* = \varepsilon_*(\lambda_0, \lambda_L, \lambda, \mu, \kappa, C_0, \beta)$, *for any* $\lambda' < \min(\lambda_0, \lambda_L, \lambda)$, *if* $f(t,\cdot\,)$ *is the solution of the nonlinear Vlasov equation with interaction* W *and initial datum* f_i, *and* $F = F[f]$ *is the associated force field, one has*

$$\|F(t,\,\cdot\,)\| = O(\varepsilon\, e^{-2\pi\lambda' t}) \qquad \text{as } t \to +\infty;$$

also $\rho(t,x)$ *converges (strongly and exponentially fast) to the average* $\rho_\infty = \iint f_i(x,v)\, dx\, dv$.
Moreover,

$$f(t,\,\cdot\,) \xrightarrow[weakly]{t\to\pm\infty} f_{\pm\infty},$$

where $f_{\pm\infty} = f_{\pm\infty}(v)$ *is an analytic homogeneous equilibrium; and*

$$\langle f(t,\,\cdot\,)\rangle \xrightarrow[strongly]{t\to\pm\infty} f_{\pm\infty},$$

where $\langle\,\cdot\,\rangle$ *stands for the* x*-average.*

Remarks 4.2. Here is a long series of remarks:

1. The condition $\widehat{W}(k) = O(1/|k|^2)$ seems critical, and only appears in the study of the nonlinear problem. It would somehow be easier to handle a decrease like $O(1/|k|^{2+\delta})$; this may be a coincidence or some deep thing.

2. The analytic norm used for $f_i - f^0$ is the most naive norm controlling exponential localization in Fourier space and in physical space. Also the smallness restriction on the size of ε is natural.

3. The decay of the force field is the damping phenomenon; the existence of limit distributions $f_{+\infty}$ and $f_{-\infty}$ is a bonus. These distributions are not determined by variational principles (at least, not that

I know) and may differ for $t \to +\infty$ and $t \to -\infty$. When $f_{+\infty} \neq f_{-\infty}$, such trajectories may be called **heteroclinic**, as opposed to homoclinic trajectories for which the behavior for $t \to +\infty$ is the same as for $t \to -\infty$ (as in the case of the linearized Vlasov equation). It is striking to see that *a whole neighborhood of* f^0 (stable equilibrium) is filled by homoclinic/heteroclinic trajectories. This is not predicted by the (random) quasilinear theory of the Vlasov–Poisson equation, neither by the statistical theory: there is in fact no randomness in Theorem 4.1. This abundance of homoclinic/heteroclinic orbits is possible only because, thanks to the infinite dimension, one can play with the various topologies. The fact that the conditions for damping are expressed in terms of the initial condition alone is a considerable improvement over previous works in the field.

4. The proof provides a constructive approach of the long-time behavior of f, which makes it possible to exchange the limits $\varepsilon \to 0$ and $t \to \infty$, perform asymptotic expansions, etc. In particular, one can check that the asymptotic state $f_{+\infty}$ "keeps the memory" of the initial datum, in a way that goes beyond the invariants of motion. The mere existence of heteroclinic trajectories demonstrates this: reversing time does not change the energy neither any of the constants of motion which were mentioned in Chapter 2, but will change the asymptotic behavior. This seems to confirm an objection raised by Isichenko in 1966 against the statistical theory of the nonlinear Vlasov equation. However, on second thoughts, the statistical theory can counterattack, because of the high *regularity* which is involved in the result. Theorem 4.1 is based on an analytic regularity; even if this is later relaxed in a Sobolev or even C^r regularity, these classes of regularity are probably negligible in a statistical context, where typical distributions are probably not smooth. (For instance, typical distributions for Sturm's entropic measure are not even absolutely continuous!)

5. The theorem in this section provides a perturbative large-time existence result which is covered neither by the Lions–Perthame theory (because this is a periodic setting) nor by the Pfaffelmoser–Batt–Rein theory (because the analyticity assumption is not compatible with a compact support).

6. An important corollary of the theorem is the *orbital stability* in a strong sense: if $H^s_{x,v}$ stands for the L^2-Sobolev space of order s on $\mathbb{T}^d_x \times \mathbb{R}^d_v$, then under the assumptions of Theorem 4.1, for any $s > 0$ we have

$$\left\| f(t, x + vt, v) - f^0(v) \right\|_{H^s_{x,v}} = O(\varepsilon), \tag{4.4}$$

uniformly as $t \to +\infty$. Since $L^2(dx\,dv)$ is invariant under the action of free transport, the norms in (4.4) control $\|f(t,\cdot) - f^0\|_{L^2(\mathbb{R}^d\times\mathbb{T}^d)}$. Using moment bounds and Sobolev injections, one may also control $f - f^0$ in L^p norms, for all $p \geq 1$.

7. The critical regularity to which the proof applies is the Gevrey class $\mathcal{G}^\nu$, for any $\nu < 3$ in the favorable cases. By convention, a function f belongs to the Gevrey class $\mathcal{G}^\nu$ if $f^{(n)} = O(n!^\nu)$, say in the supremum norm. Modulo an arbitrarily small loss on the exponent ν, it is equivalent to say that the Fourier transform $\widehat{f}(\xi)$ of f decays at infinity like $O(e^{-c|\xi|^{1/\nu}})$. The favorable case alluded to above is the one in which $\int |K(t,k)|\,dt < 1$; this applies for instance for gravitational interaction below the Jeans length. More generally, if the interaction satisfies $|\widehat{W}(k)| = O(1/|k|^{1+\gamma})$ and $\int |K(t,k)|\,dt < 1$ then the critical exponent is $\nu = \gamma + 2$. Here is a precise statement where we do not care about the critical exponent:

Theorem 4.3 (Nonlinear Landau damping in Gevrey regularity). *Let $W : \mathbb{T}^d \to \mathbb{R}$ be an interaction potential such that $|\widehat{W}(k)| = O(1/|k|^2)$. Let $f^0 = f^0(v)$ be an analytic profile such that*

$$|\widetilde{f}^0(\eta)| \leq C_0\, e^{-2\pi\lambda_0|\eta|}, \qquad \sum_n \frac{\lambda_0^n}{n!}\, \|\nabla_v^n f^0\|_{L^1(dv)} \leq C_0,$$

for some $\lambda_0 > 0$, and satisfying the Penrose stability condition with Landau width $\lambda_L > 0$. Then there is $\theta > 0$ such that for any $\nu \in (1, 1+\theta)$, $\beta > 0$, $\alpha < 1/\nu$, there is $\varepsilon_* > 0$ such that if

$$\varepsilon := \sup_{k,\eta}\Big(\big|(\widetilde{f}_i - \widetilde{f}^0)(k,\eta)\big|\, e^{\lambda|\eta|^{1/\nu}}\, e^{\lambda|k|^{1/\nu}}\Big) + \iint |f_i(x,v) - f^0(v)|\, e^{\beta|v|}\, dv\, dx$$

satisfies $\varepsilon \leq \varepsilon_*$, then the solution $f = f(t,x,v)$ of the nonlinear Vlasov equation with interaction W and initial datum f_i satisfies

$$F[f](t,\,\cdot\,) = O\big(\varepsilon\, e^{-ct^\alpha}\big),$$

and $f(t,\,\cdot\,)$ converges weakly to some asymptotic Gevrey profile $f_{\pm\infty} = f_{\pm\infty}(v)$ as $t \to \pm\infty$, with convergence rate $O(\varepsilon\, e^{-ct^\alpha})$.

6. The information cascade

How can one reconciliate the reversibility of the nonlinear Vlasov equation and the convergence in large time? Convergence seems to be about loss of information, which should go with an increase of entropy; but we have seen that the Vlasov equation preserves the entropy. So what?

The solution to this apparent paradox (well understood by some physicists at least fifty years ago, but not by all of them) is that the information all goes out to high frequencies, where it is *inaccessible*, *hidden*. The following numerical simulations will illustrate this.

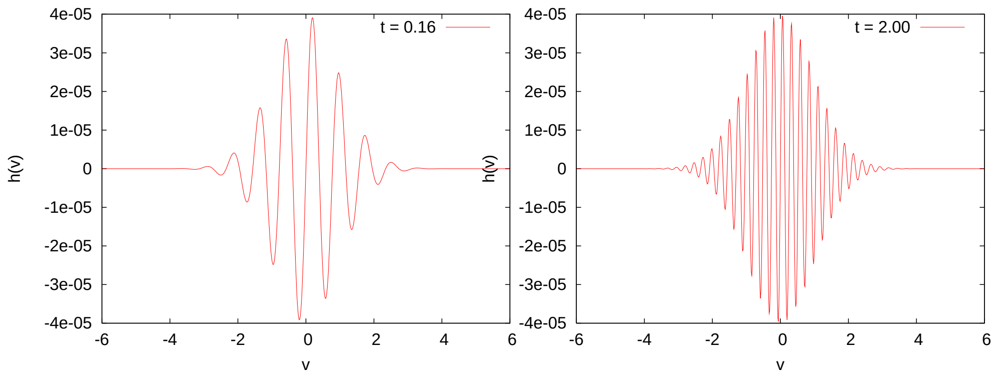

FIGURE 4.3. A slice of the distribution function (relative to a homogeneous equilibrium) for gravitational Landau damping, at two different times; notice the fast oscillations of the distribution function, which are very difficult to capture by an experiment.

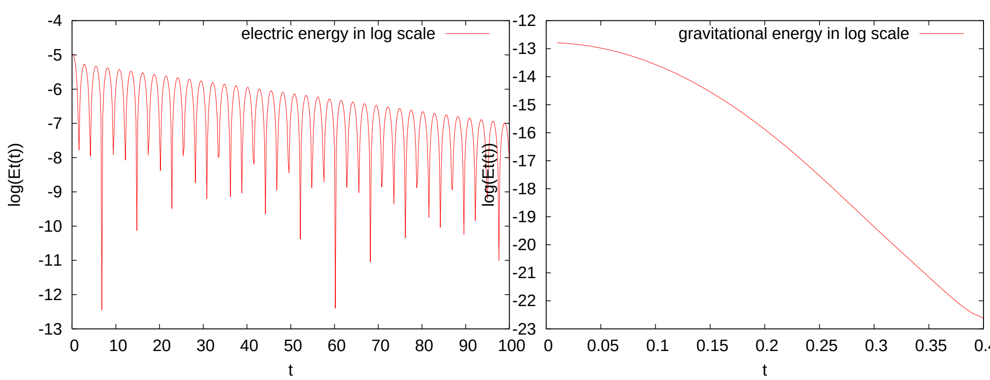

FIGURE 4.4. Time-evolution of the norm of the field, for electrostatic (on the left) and gravitational (on the right) interactions. In the electrostatic case, the fast time-oscillations are called Langmuir oscillations, and should not be mistaken with the velocity oscillations.

In fact, the cascade of energy, already present in the free transport evolution, still holds: energy (or information) flows from low to high frequencies. The complete integrability has been lost, but the energy transfer still holds. This is similar in spirit to the **KAM phenomenon** (KAM = Kolmogorov–Arnold–Moser), to which we shall come back later; for the moment I will just mention that there are common points and differences with the KAM theory.

To go back to the information, let us note that

$$\iint f \log f = \int \rho \log \rho + \iint f \log \frac{f}{\rho},$$

and the first term on the right-hand side converges to 0 because ρ converges strongly to a constant. So all the information becomes "kinetic" in the limit. (Due to the oscillations, a priori we cannot pass to the limit; but that becomes possible if we go along trajectories of the free transport and use the gliding regularity.)

A final remark is in order: because of the time-reversibility, any stability result, read backwards in time, should imply an *instability* result. This is true, however with a catch on the topologies involved. Landau damping asserts convergence in large time in the weak topology, when the initial datum is perturbed in the strong topology. Reading it backwards implies a instability in the strong topology, when the initial datum is perturbed in the weak topology. In particular, Landau damping is by no means in contradiction with the time-reversibility.

Bibliographical notes

Backus [**8**] was certainly the first one to point out the conceptual problem caused by the interversion of the asymptotic regimes $t \to \infty$ and $\|f_i - f^0\| \to 0$. Thus his paper contains both positive statements (the first mathematically rigorous treatment of the linear Landau damping) and expression of skepticism, which has not been answered until [**74**].

Nonlinear stability of monotone homogeneous profiles for the electrostatic Vlasov–Poisson equation was studied by Rein [**86**]. The two-stream instability has been established by Guo and Strauss [**41**].

The nonlinear time scale $O(1/\sqrt{\varepsilon})$ was predicted in 1965 by O'Neil [**81**] with a more sophisticated argument than the naive scaling approach presented in Section 1, and is anyway well checked in numerical simulations. This is the time scale where significant departure appears between the solutions of the linear and nonlinear equations; it does not mean that the qualitative behavior of these solutions differ much. Actually, O'Neil argues that damping still holds true for larger time scales, even though not infinite. He also stumbles on the problem pointed out by Backus.

Numerical investigation of the Landau damping, for small and larger perturbations, was performed by several authors at the end of the nineties, when precise methods started to be available [**67, 105**]. Since then, more efficient schemes have become available [**44**].

The mysterious but popular quasilinear theory is presented in a number of sources [**61**, Section 49] [**1**, Section 9.1.2] [**54**, Chapter 10], with more or less convincing arguments. Putting this theory on a decent level of rigor looks like a challenge.

Isichenko's criticism of the statistical theory of Vlasov equation appears in [**51**]. Using an analogy with a random potential problem, he argues that the convergence to equilibrium should be slow; however the analogy used by Isichenko is not so sound, and in fact very fast convergence can occur, as demonstrated in [**21**] and in the present paper. Isichenko's paper is still worth reading for his interesting insights, though.

Theorem 4.1 is taken from [**74**], as well as the comments made right after its statement.

Caglioti and Maffei [**21**] were the first to construct some exponentially damped solutions to the Vlasov–Poisson equation (in dimension 1); they also noted that this implies, by time-reversibility, the instability in the weak topology. Another construction to damped solutions was performed by Hwang and Velázquez [**50**].

The numerical simulations in this chapter were kindly provided by Filbet.

CHAPTER 5

Gliding analytic regularity

Analytic regularity may seem very specific, but it is for sure the first setting to understand in the study of Landau damping, for physical reasons (because it is associated to exponential damping) and for historical reasons as well (because this is the case that was treated by Landau). But the obstacles related to the study of the limit $t \to \infty$ require much care in the choice of functional space.

1. Preliminary analysis

There are many families of analytic norms. A particularly simple family is defined by the formula

$$\|f\|_{\lambda,\mu} = \sup_{k,\eta} \big|\widetilde{f}(k,\eta)\big|\, e^{2\pi\mu|k|}\, e^{2\pi\lambda|\eta|}. \tag{5.1}$$

One can interpret λ and μ as the width of the analyticity strip in the v and x variables, respectively.

To evaluate the relevance of these functional spaces, let us examine the equation and seek to approximate it by an iterative scheme. The natural quasilinear scheme is

$$\partial_t f^{n+1} + v \cdot \nabla_x f^{n+1} + F[f^n] \cdot \nabla_v f^{n+1} = 0, \tag{5.2}$$

which amounts to let particles evolve at stage $n+1$ in the force field created by the distribution at stage n. Equivalently, one solves the transport equation by the characteristic method:

$$f^{n+1}(t,x,v) = f_i\big(S^n_{t,0}(x,v)\big),$$

where $S^n_{t,0}(x,v)$ is the position (in phase space) at time 0 of particles evolving in the force field induced by f^n, which at time t will be at (x,v). Thus one is naturally led to study the behavior of norms with respect to composition. But these norms do behave very badly: to fix ideas, let us assume that $S_t(x,v)$ satisfies "perfect" estimates, as good as (say) $2\,\mathrm{Id}$; and observe that

$$\begin{aligned}\big\|f\circ(2\,\mathrm{Id})\big\|_{\lambda,\mu} &= 2^d \sup_{k,\eta}\left|\widetilde{f}\left(\frac{k}{2},\frac{\eta}{2}\right)\right| e^{2\pi\mu|k|}\,e^{2\pi\lambda|\eta|}\\ &= 2^d\sup_{k,\eta}\big|\widetilde{f}(k,\eta)\big|\,e^{2\pi(2\mu)|k|}\,e^{2\pi(2\lambda)|\eta|}\\ &= 2^d\,\|f\|_{2\lambda,2\mu}.\end{aligned}$$

In other words, the norms λ,μ are *not stable under composition*. This is in sharp contrast with Sobolev or C^r norms. Now imagine the disaster: as one iterates the estimates, one loses a factor 2 on the width of the analyticity strip, so that there is nothing left in the end...

What is more, (5.2) will not express a reaction: recall the heuristic image about pushing the wall and exhausting oneself. It would be more promising to use the alternative scheme

$$\partial_t f^{n+1} + v\cdot\nabla_x f^{n+1} + F[f^{n+1}]\cdot\nabla_v f^n = 0.$$

But this would mean treating a higher order degree, $\nabla_v f^n$, as a perturbation: thus there would be a *loss of derivative*.

A third problem is related to the *filamentation*: the best one can hope for f is estimates in the style of the solution of free transport $\widetilde{g}(k,\eta+kt)$; so in the best of worlds,

$$\big|\widetilde{f}(t,k,\eta)\big| \le C\,e^{-2\pi\lambda|\eta+kt|}\,e^{-2\pi\mu|k|}.$$

But then $\|f\|_{\lambda,\mu} \gtrsim e^{2\pi\lambda t}$ (choose $\eta=-kt$, $|k|=1$); and even worse, $\|f\|_{\lambda,\mu} \gtrsim e^{2\pi\lambda|k|t}$ for any k such that the mode k does not vanish. If all modes k are represented, one expects the norm of f to grow faster than any exponential! This of course is a disaster for the large-time analysis.

2. Algebra norms

Among all families of analytic norms, two deserve a special mention; let us present them in dimension 1:

$$\|f\|_{\mathcal{F}^\lambda} = \sum_{k\in\mathbb{Z}} e^{2\pi\lambda|k|}\,|\widehat{f}(k)| \qquad \|f\|_{\mathcal{C}^\lambda} = \sum_{n\in\mathbb{N}_0}\frac{\lambda^n}{n!}\|f^{(n)}\|_{L^\infty}, \tag{5.3}$$

where $f^{(n)}$ stands for the derivative of order n of f, and $\mathbb{N}_0=\{0,1,2,\ldots\}$. The first norm (as it is written) makes sense only for periodic functions, while the second one makes sense for any smooth function on $\mathbb{R}$.

PROPOSITION 5.1. *With $\|\cdot\|_\lambda$ standing either for the $\mathcal{F}^\lambda$ or the $\mathcal{C}^\lambda$ norm, one has*

$$\|fg\|_\lambda \le \|f\|_\lambda\,\|g\|_\lambda.$$

Sketch of proof. Let us prove the inequality for, say, the $\mathcal{C}^\lambda$ norm: the Leibniz differentiation formula implies

$$\begin{aligned}\sum_n \frac{\lambda^n}{n!}\left\|(fg)^{(n)}\right\|_{L^\infty} &\le \sum_n \sum_{m\le n} \frac{\lambda^n}{n!}\,\frac{n!}{m!\,(n-m)!}\,\|f^{(m)}\|_{L^\infty}\,\|g^{(n-m)}\|_{L^\infty}\\ &= \sum_{m,m'} \frac{\lambda^{m+m'}}{m!\,m'!}\,\|f^{(m)}\|_{L^\infty}\,\|g^{(m')}\|_{L^\infty}\\ &= \left(\sum_m \frac{\lambda^m}{m!}\,\|f^{(m)}\|_{L^\infty}\right)\left(\sum_{m'} \frac{\lambda^{m'}}{m'!}\,\|f^{(m')}\|_{L^\infty}\right).\end{aligned}$$

$\square$

As an immediate corollary, we have $\|f^n\|_\lambda \le \|f\|_\lambda^n$. This remarkable algebra property implies good properties with respect to composition as well: there will be a loss of exponent, that is unavoidable, but it will be controlled.

Proposition 5.2. *With $\|\cdot\|_\lambda$ standing either for the $\mathcal{F}^\lambda$ or the $\mathcal{C}^\lambda$ norm, one has*

$$\Big\| f\circ(\mathrm{Id}\, + G)\Big\|_\lambda \le \|f\|_\nu, \qquad \nu = \lambda + \|G\|_\lambda.$$

Note carefully: the constant in front of the right-hand side is equal to 1; and the size of G (in the λ-norm) is found in the *regularity index* on the right-hand side.

Sketch of proof of Proposition 5.2. Let us do it for the $\mathcal{C}^\lambda$ norm. Writing $h = f\circ(\mathrm{Id}\, + G)$ and using the Taylor formula (leaving aside the issue of convergence), we have

$$h(x) = \sum_n \frac{f^{(n)}(x)}{n!}\,G^n(x),$$

whence

$$h^{(m)}(x) = \sum_{k+\ell=m}\sum_n \frac{m!}{k!\,\ell!\,n!}\,f^{(n+k)}(x)\,(G^n)^{(\ell)}(x),$$

so

$$
\begin{aligned}
\sum_m \frac{\lambda^m}{m!}\|h^{(m)}\|_{L^\infty} &\leq \sum_m \sum_{k+\ell=m} \sum_n \frac{\lambda^{k+\ell}}{k!\,\ell!\,n!}\,\|f^{(n+k)}\|_{L^\infty}\,\|(G^n)^{(\ell)}\|_{L^\infty} \\
&= \sum_{k,\ell,n} \frac{\lambda^{k+\ell}}{k!\,\ell!\,n!}\,\|f^{(n+k)}\|_{L^\infty}\,\|(G^n)^{(\ell)}\|_{L^\infty} \\
&\leq \sum_{k,n} \frac{\lambda^k}{k!\,n!}\,\|G^n\|_\lambda\,\|f^{(n+k)}\|_{L^\infty} \\
&\leq \sum_r \left(\sum_{k+n=r} \frac{r!}{k!\,n!}\,\lambda^k\,\|G\|_\lambda^n\right) \frac{\|f^{(r)}\|_{L^\infty}}{r!} \\
&\leq \sum_r \frac{\big(\lambda+\|G\|_\lambda\big)^r}{r!}\,\|f^{(r)}\|_{L^\infty} \\
&= \|f\|_{\lambda+\|G\|_\lambda},
\end{aligned}
$$

where Newton's binomial formula was used. □

Proposition 5.2 admits some variants: in particular, it is possible to mix norms:

$$
\Big\| f\circ(\mathrm{Id}+G)\Big\|_{\mathcal{C}^\lambda} \leq \|f\|_{\mathcal{F}^\nu}, \qquad \nu=\lambda+\|G\|_{\dot{\mathcal{C}}^\lambda}, \tag{5.4}
$$

where the $\dot{\mathcal{C}}^\lambda$ seminorm is obtained from the $\mathcal{C}^\lambda$ norm by throwing away the zero mode. (The proof of (5.4) involves the dreaded Faà di Bruno formula.)

Working in kinetic theory, it is particularly convenient to *hybridize* the two spaces: apply the recipe $\mathcal{C}$ to the velocity space, and the recipe $\mathcal{F}$ to the position space. This will take advantage of the periodic geometry of the position space, and property (5.4) will guarantee that composition by the characteristics is still properly handled. Let us also generalize to d dimensions, add a parameter γ to count derivatives (for technical reasons, this is needed only in the x variable), and another parameter p to modulate the integrability; we are led to the formula

$$
\|f\|_{\mathcal{Z}^{\lambda,(\mu,\gamma);p}} = \sum_{k\in\mathbb{Z}^d}\sum_{n\in\mathbb{N}_0^d} e^{2\pi\mu|k|}\,(1+|k|)^\gamma\,\frac{\lambda^n}{n!}\,\big\|\nabla_v^n\widehat{f}(k,v)\big\|_{L^p(dv)}, \tag{5.5}
$$

where $\widehat{f}(k,v)=\int f(x,v)\,e^{-2i\pi k\cdot x}\,dx$ is the Fourier transform of f in the x variable only. Then we have good properties generalizing Property 5.1 such as

$$
\|fg\|_{\mathcal{Z}^{\lambda,(\mu,\gamma);p}} \leq \|f\|_{\mathcal{Z}^{\lambda,(\mu,\gamma);q}}\,\|g\|_{\mathcal{Z}^{\lambda,(\mu,\gamma);r}}, \qquad \frac{1}{r}=\frac{1}{p}+\frac{1}{q}.
$$

3. Gliding regularity

Now it remains to take into account filamentation, that is, the appearance of fast oscillations in the velocity variable. Since we cannot avoid it, let us accept it and *incorporate it* in the norm. This amounts to introducing a parameter τ (time-like) and to let the free transport evolution act backwards on the distribution function, for a time τ:

$$\|f\|_{\mathcal{Z}_\tau} = \|f \circ S^0_{-\tau}\|_{\mathcal{Z}},$$

where the regularity indices are implicit, and $S^0_t(x,v) = (x+vt, v)$.

This provides a family of functional spaces depending on a parameter τ, which can a priori be chosen as one wishes, the idea being that τ is equal to, or at least asymptotic to, the time of the equation. Thus we adapt our regularity scale to the filamentation; or, we focus on the relevant Fourier modes as time goes by.

All in all, we are led to the final definition of the $\mathcal{Z}$ norms: for a function $f = f(x,v)$,

$$\|f\|_{\mathcal{Z}^{\lambda,(\mu,\gamma);p}_\tau} = \sum_{k\in\mathbb{Z}^d} \sum_{n\in\mathbb{N}^d_0} e^{2\pi\mu|k|}\,(1+|k|)^\gamma\, \frac{\lambda^n}{n!}\, \left\| \big(\nabla_v + 2i\pi\tau k\big)^n \widehat{f}(k,v) \right\|_{L^p(dv)}. \tag{5.6}$$

By default, $\tau = 0$, $\gamma = 0$ and $p = \infty$.

Remarks 5.3. Here are some important remarks about the $\mathcal{Z}$ norms.

1. If $f = f(t,x,v)$ is a solution of the free transport equation, then

$$\|f(t,\cdot)\|_{\mathcal{Z}_t} = \|f(0,\cdot)\|_{\mathcal{Z}}$$

(the regularity indices being implicit).

2. If $f = f(v)$, then the $\|f\|_{\mathcal{Z}^{\lambda,(\mu,\gamma)}_\tau}$ norm reduces to the $\mathcal{C}^\lambda$ norm.

3. On the other hand, if $f = f(x)$, it reduces to the $\mathcal{F}^\nu$ norm with $\nu = (\lambda\tau+\mu, \gamma)$ (that is, the $\mathcal{F}^{\lambda\tau+\mu}$ space with γ additional derivatives). The crucial point is that the regularity in x improves with t, as it should be in view of our discussion at the end of Chapter 3!

4. As a final remark, we shall almost never try to compare norms $\mathcal{Z}_\tau$ for various values of τ, because this is very costly in the velocity regularity: we don't have anything better than

$$\|f\|_{\mathcal{Z}^{\lambda,\mu}_{\tau'}} \le \|f\|_{\mathcal{Z}^{\lambda,\mu+\lambda|\tau-\tau'|}_\tau},$$

and this becomes unaffordable as soon as $|\tau - \tau'|$ is large.

4. Functional analysis

Now one can study the main properties of the $\mathcal{Z}$ spaces, with respect to product, composition, differentiation (the analytic rigidity implies a control of the derivative in terms of the function itself), inversion, averaging... (The inversion is estimated by means of a fixed point theorem.) Here are some of the main results (I do not mention those indices which are similar on the left and right-hand sides of the inequalities):

$$\|fg\|_{\mathcal{Z};r} \le \|f\|_{\mathcal{Z};p}\, \|g\|_{\mathcal{Z};q}, \qquad \frac{1}{r} = \frac{1}{p} + \frac{1}{q}; \tag{5.7}$$

$$\Big\| f\big(x + X(x,v), v + V(x,v)\big)\Big\|_{\mathcal{Z}_\tau^{\lambda,\mu;p}} \le \|f\|_{\mathcal{Z}_\sigma^{\alpha,\beta;p}}, \tag{5.8}$$

where $\alpha = \lambda + \|V\|_{\mathcal{Z}_\tau^{\lambda,\mu}}$, $\beta = \mu + \lambda|\tau - \sigma| + \|X - \sigma V\|_{\mathcal{Z}_\tau^{\lambda,\mu}}$;

$$\|\nabla_x f\|_{\mathcal{Z}^{\lambda,\mu}} \le \frac{C}{\overline{\mu} - \mu}\, \|f\|_{\mathcal{Z}^{\lambda,\overline{\mu}}}; \tag{5.9}$$

(5.10)

$$1 < \overline{\lambda}/\lambda \le 2 \Longrightarrow \qquad \|\nabla_v f\|_{\mathcal{Z}_\tau^{\lambda,\mu}} \le C\left(\frac{1}{\overline{\lambda} - \lambda} + \frac{1+\tau}{\overline{\mu} - \mu}\right)\|f\|_{\mathcal{Z}_\tau^{\overline{\lambda},\overline{\mu}}};$$

$$\begin{aligned} \text{(5.11)} \quad \exists \alpha = \alpha(d) > 0; \quad & \|\nabla(F - \mathrm{Id}\,)\|_{\mathcal{Z}_\tau^{\lambda',\mu'}} \le \alpha \\ & \Longrightarrow \ \big\|F^{-1} \circ G - \mathrm{Id}\,\big\|_{\mathcal{Z}_\tau^{\lambda,\mu}} \le 2\,\|F - G\|_{\mathcal{Z}_\tau^{\lambda,\mu}}, \end{aligned}$$

where $\lambda' = \lambda + 2\,\|F - G\|_{\mathcal{Z}_\tau^{\lambda,\mu}}$, $\mu' = \mu + 2\,(1 + |\tau|)\,\|F - G\|_{\mathcal{Z}_\tau^{\lambda,\mu}}$;

$$\left\|\int f\, dv\right\|_{\mathcal{F}^{\lambda|\tau|+\mu}} \le \|f\|_{\mathcal{Z}_\tau^{\lambda,\mu;1}}; \tag{5.12}$$

$$\left\|\int f\, dx\right\|_{\mathcal{C}^\lambda} \le \|f\|_{\mathcal{Z}_\tau^\lambda}. \tag{5.13}$$

A more subtle inequality, which will allow us to "cheat" with the time index, is

$$\left\|\int f\big(x - v(t-\tau), v\big)\, dv\right\|_{\mathcal{F}^{\lambda t+\mu}} \le \|f\|_{\mathcal{Z}_{\tau - \frac{bt}{1+b}}^{\lambda(1+b),\mu;1}}, \tag{5.14}$$

for any $b > -1$ and $t > 0$. To prove (5.14), write

$$\begin{aligned} \left\|\int f \circ S^0_{\tau-t}\, dv\right\|_{\mathcal{F}^{\lambda t+\mu}} &\le \|f \circ S^0_{\tau-t}\|_{\mathcal{Z}_{t'}^{\lambda',\mu;1}} \\ &= \|f\|_{\mathcal{Z}_{t'+\tau-t}^{\lambda',\mu;1}}, \end{aligned}$$

where λ' and t' are subject to $\lambda' t' = \lambda t$. The choice $\lambda' = \lambda(1+b)$, $t' = t/(1+b)$ completes the proof of (5.14).

All these properties will be convenient to study the nonlinear Vlasov equation. One may complain about the complicated nature of the norms; but it is possible to inject these norms into more standard norms, up to an arbitrarily small loss on the regularity indices. Thus, even if we work out the estimates in the complicated $\mathcal{Z}$ norms, we will be able to state the result in the simple-minded $\mathcal{Y}$ norms defined by

$$\|f\|_{\mathcal{Y}^{\lambda,\mu}_\tau} := \sup_{k,\eta} \left|\widetilde{f}(k,\eta)\right| e^{2\pi\lambda|\eta+kt|}\, e^{2\pi\mu|k|}.$$

To estimate the $\mathcal{Y}$ norms by the $\mathcal{Z}$ norms, we have the simple inequality

$$\|f\|_{\mathcal{Y}^{\lambda,\mu}_\tau} \le \|f\|_{\mathcal{Z}^{\lambda,\mu;1}_\tau}. \tag{5.15}$$

Conversely, to estimate the $\mathcal{Z}$ norms by the $\mathcal{Y}$ norms, we have the more subtle inequalities

$$\|f\|_{\mathcal{Z}^{\lambda,\mu;\infty}_\tau} \le \frac{C(d,\overline{\mu})}{(\overline{\lambda}-\lambda)^d\,(\overline{\mu}-\mu)^d}\,\|f\|_{\mathcal{Y}^{\overline{\lambda},\overline{\mu}}\tau} \tag{5.16}$$

and

(5.17)
$$\|f\|_{\mathcal{Z}^{\lambda,\mu;1}_\tau} \le C^{\frac{1}{\min(\overline{\lambda}-\lambda,\overline{\mu}-\mu)}} \left(\|f\|_{\mathcal{Y}^{\overline{\lambda},\overline{\mu}}\tau} + \int |f(x,v)|\, e^{2\pi\beta|v|}\,dv\,dx\right).$$

The last inequality holds as soon as $\lambda < \overline{\lambda} \le \Lambda$, $\mu < \overline{\mu} < M$, $0 < b \le \beta \le B$, $\int |f| e^{2\pi\beta|v|}\,dv\,dx \le E$, and the constant C only depends on Λ, M, B, β, E. The mechanism of proof is similar to the **Sobolev injections**.

All inequalities in this section seem pretty much optimal, except for (5.17), for which one may conjecture that the best constants in the right-hand side of (5.17) are polynomial (instead of exponential) in $1/\min(\overline{\lambda}-\lambda,\overline{\mu}-\mu)$.

Bibliographical notes

Algebra norms similar to those in Section 2 are well-known in certain circles, and appear for instance in [**3**]. The idea to combine them, the resulting composition formulae, and the notion of gliding regularity were introduced in [**74**]. Detailed proofs can be found in Section 4 of that work.

Gliding regularity is somehow reminiscent of the philosophy used by Bourgain [**18**] in the definition of his $X^{s,b}$ spaces, which are defined by comparison with some unperturbed reversible dynamics; a difference is the role of the time variable, which in our treatment is just a parameter.

The conjecture according to which the constants in (5.17) might be polynomial is briefly discussed in [**74**], and an application to the study of the nonlinear stability in "low" regularity is sketched. The picture is far from clear.

CHAPTER 6

Characteristics in damped forcing

Before turning to the nonlinear problem where the distribution function determines the force, let us address the linear problem in which the force field is given and drives the force, and let us assume on the force field the desired qualitative features. Of course, the study of the transport equation can be reduced to the understanding of particle trajectories (characteristics), so we shall focus on those trajectories.

1. Damped forcing

Let be given a small gradient force field $F(t,\cdot)$ whose analytic regularity improves linearly in time: with the notation (5.3),

$$\big\|F(t,\,\cdot\,)\big\|_{\mathcal{F}^{\lambda t}} = O(\varepsilon).$$

The question is about the qualitative behavior of trajectories; in particular, are they transient like free transport trajectories, or can they be trapped and go along complicated trajectories?

To get a first feeling, let us proceed to a formal asymptotic analysis. As before, we write $X_{s,t}(x,v)$, $V_{s,t}(x,v)$ for respectively the position and velocity at time t, starting from time s at (x,v) Writing formally

$$V_{0,t}(x,v) = v + \varepsilon\, v^{(1)}(t,x,v) + \varepsilon^2\, v^{(2)}(t,x,v) + \dots,$$

$$X_{0,t}(x,v) = x + vt + \varepsilon \int_0^t v^{(1)}(s,x,v)\,ds + \varepsilon^2 \int_0^t v^{(2)}(s,x,v)\,ds + \dots,$$

we expect $F(t,X_{0,t}(x,v)) = F(t,x+vt) + O(\varepsilon^2)$, then the equation $\ddot{X} = F(t,X)$ leads to

$$V_{0,t}(x,v) = v + \int_0^t F(s,x+vs)\,ds + O(\varepsilon^2),$$

so we expect $V_{0,t}(x,v) = v + O(\varepsilon)$. If this is in an analytic norm taking derivatives into account, we expect in particular

$$|\nabla_v V_{0,t}(x,v) - I| = O(\varepsilon),$$

in particular the flow should be invertible if ε is small enough.

So our guess is that the trajectories remain perturbations of the free flow $(x+vt,v)$, uniformly in time.

2. Scattering

To compare the perturbed dynamics to the free dynamics, let us write $S_{t,\tau} = (X_{t,\tau}, V_{t,\tau})$, $S^0_{t,\tau} = (x - (t-\tau)v, v)$, and define the **finite-time scattering operator**: for $0 \le \tau \le t$,

$$\Omega_{t,\tau} = S_{t,\tau} \circ S^0_{\tau,t}. \tag{6.1}$$

That is, start from time τ, evolve by the free dynamics up to time t, and then evolve it backwards by the perturbed dynamics to time τ. As $t \to \infty$, $\Omega_{t,\tau}$ converges to what is usually called a scattering transform, whence the terminology.

Proposition 6.1. *With the above notation, if $\lambda' < \lambda$, $\mu' < \mu$ and*

$$\|\!| F |\!\| := \sup_{t \ge 0} \|F(t,\,\cdot\,)\|_{\mathcal{F}^{\lambda t + \mu}} \le \frac{(\mu - \mu')\,(\lambda - \lambda')^2}{C}, \tag{6.2}$$

where C is large enough, then

$$\Big\| \Omega_{t,\tau} - \mathrm{Id} \Big\|_{\mathcal{Z}^{\lambda',\mu'}_\tau} \le C\, \|\!| F |\!\|\, e^{-2\pi(\lambda - \lambda')\tau}\, \min\left(t - \tau, \frac{1}{\lambda - \lambda'} \right).$$

Proposition 6.1 provides an analytic scattering, modulo a small loss on the regularity index; it can be generalized (e.g. by changing τ on the left within some constraints), but for the moment this is quite sufficient to give a first idea. This estimate is

(a) uniform as $t \to \infty$;
(b) small as $\tau \to t$;
(c) exponentially small as $\tau \to \infty$.

Choosing $\lambda - \lambda' \simeq \mu - \mu'$, we see that the loss of regularity index is roughly of order $O(\varepsilon^{1/3})$. We shall see in Chapter 8 how to overcome this loss by changing the estimate.

Sketch of proof of Proposition 6.1. The principle of the proof is a standard fixed point reasoning. First let us make the ansatz

$$S_{t,\tau}(x,v) = \Big(x - v(t-\tau) + Z_{t,\tau}(x,v),\; v + \partial_\tau Z_{t,\tau}(x,v) \Big)$$

(the second component of $S_{t,\tau}$ is the τ-derivative of the first one). The equation on Z is

$$\begin{cases} \dfrac{\partial^2 Z}{\partial \tau^2} = F\Big(\tau, x - v(t-\tau) + Z_{t,\tau}\Big) \\ Z_{t,t} = 0, \qquad \partial_\tau\big|_{\tau = t} Z_{t,\tau} = 0. \end{cases} \tag{6.3}$$

So Z appears as a fixed point of $\Psi : W \longmapsto Z$, where Z is the solution of

$$(6.4)\qquad \begin{cases} \dfrac{\partial^2 Z}{\partial \tau^2} = F\Big(\tau, x - v(t-\tau) + W_{t,\tau}\Big) \\ Z_{t,t} = 0, \qquad \partial_\tau\big|_{\tau=t} Z_{t,\tau} = 0. \end{cases}$$

For given t, $Z_{t,\tau}$ is a function of $\tau \in [0,t]$; let us introduce the norm

$$\Big\|(Z_{t,\tau})_{0\le\tau\le t}\Big\|_{[0,t]} := \sup_{0\le\tau\le t} \left\{ \frac{\|Z_{t,\tau}\|_{\mathcal{Z}_\tau^{\lambda',\mu'}}}{R(\tau,t)} \right\}$$

where

$$R(\tau,t) = C\, e^{-2\pi(\lambda-\lambda')\tau}\, \min\left[\frac{(t-\tau)^2}{2},\ \frac{1}{(\lambda-\lambda')^2}\right].$$

(For the sake of pedagogy, I am slightly cheating, the right definition is slightly more complicated, see the original work for details.) Then the goal is to check that (a) $\|\Psi(0)\|_{[0,t]} \le 1$ and (b) Ψ is 1-Lipschitz in the norm $\|\cdot\|_{[0,t]}$, on the ball of radius 2 (for the same norm). If that is true, it follows by a classical fixed-point theorem that Ψ has a unique fixed point in the ball of radius 2, and this provides the desired estimate.

Let us give a hint of how to perform these estimates. For (a), we see that $\Psi(0) = Z^0$ such that

$$Z_{t,\tau} = \int_\tau^t (s-\tau)\, F\big(s, x - v(t-s)\big)\, ds.$$

We are estimating this in $\mathcal{Z}_\tau^{\lambda',\mu'}$ norm, so (recalling Remark 5.3(3)) this is trivially bounded above by

$$\int_\tau^t (s-\tau)\, \|F(s,\,\cdot\,)\|_{\mathcal{F}^{\lambda'\tau+\mu'}}\, ds.$$

Since F is a gradient, for $s \ge \tau$ we have the estimate

$$\|F(s,\,\cdot\,)\|_{\mathcal{F}^{\lambda'\tau+\mu'}} \le e^{2\pi(\lambda'\tau-\lambda s)}\, \|F(s,\,\cdot\,)\|_{\mathcal{F}^{\lambda s+\mu}} \le e^{-2\pi(\lambda-\lambda')s}\, |\!|\!|F|\!|\!|;$$

in other words, *thanks to the gradient structure of F, the gliding regularity has been converted into a time decay.* Now we obtain the bound on Z^0 by time-integration:

$$\int_\tau^t (s-\tau)\, e^{-2\pi(\lambda-\lambda')s}\, ds \le C\, e^{-2\pi(\lambda-\lambda')\tau}\, \min\left[(t-\tau)^2, \frac{1}{(\lambda-\lambda')^2}\right],$$

and the desired property follows easily.

To check the Lipschitz property (b) is hardly more tricky: if $Z = \Psi(W)$, $\widetilde{Z} = \Psi(\widetilde{W})$ then we have

$$Z_{t,\tau} - \widetilde{Z}_{t,\tau} = \int_\tau^t (s-\tau) \left[F\big(s, x-v(t-s)+W\big) - F\big(s, x-v(t-s)+\widetilde{W}\big)\right] ds.$$

To estimate this we write

$$\begin{aligned} & F\big(s, x-v(t-s)+W\big) - F\big(s, x-v(t-s)+\widetilde{W}\big) \\ & \qquad = \int_0^1 \nabla_x F\Big(s, x-v(t-s)+(1-\theta)W+\theta\widetilde{W}\Big)\, d\theta \cdot (W-\widetilde{W}), \end{aligned}$$

and then use the functional analysis of the $\mathcal{Z}$ spaces (with respect to product, composition, differentiation, and evolution by free transport) to bound this. A source of loss of regularity is the composition by something which has size $1 + O(\|W\|)$. Since $\|W\| = O(\varepsilon/(\lambda-\lambda')^2)$, we can absorb this loss of regularity (due to composition) into the loss of regularity in x, if $\varepsilon/(\lambda-\lambda')^2$ is significantly smaller than $\mu-\mu'$, and this explains where condition (6.2) comes from. □

Bibliographical notes

This chapter is entirely based on [**74**, Section 5].

CHAPTER 7

Reaction against an oscillating background

In the past chapter we were considering the time-evolution of an unknown distribution evolving in a given force field, now we shall consider the dual point of view: the force will be the unknown, and the forced distribution will be given. So the equation will be

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f + F[f](t,x) \cdot \nabla_v \overline{f}(t,x,v) = 0. \tag{7.1}$$

Formally, this equation describes the evolution of a gas of particles which tries to force the distribution $\overline{f}$, however there is a flux (or transmutation) of particles from distribution f to distribution $\overline{f}$, compensating exactly the effect of the force. Accordingly, I will informally call (7.1) the reaction equation. We shall assume on $\overline{f}$ the same estimates as on a typical solution of a transport equation, so in large time $\nabla_v \overline{f}$ will oscillate fast in the velocity variable. — as in Chapter 3.

1. Regularity extortion

For mnemonic purpose, one may interpret (7.1) saying that one is pushing against an oscillating wall, which at times takes energy and at times gives it back, so that it is not clear whether at the end of the day one gets exhausted or not. The goal of this chapter is to show that if $\overline{f}$ is quite smooth (in gliding regularity), then the force associated with f will gain regularity in time, eventually causing the exhaustion.

Proposition 7.1. *Let f^0 such that $|\widetilde{f}^0(\eta)| = O(e^{-2\pi\lambda_0|\eta|})$ and f^0 satisfies the Penrose stability condition with stability width λ_L. Assume that the interaction satisfies $\widehat{W}(k) = O(1/|k|^{1+\gamma})$, $\gamma \geq 1$. Let $f_i = f_i(x,v)$ such that*

$$\|f_i\|_{\mathcal{Z}^{\lambda,\mu;1}} \leq \varepsilon$$

and $\overline{f}(t,x,v) = f^0(v) + \overline{h}(t,x,v)$ with

$$\|\overline{h}(t,x,v)\|_{\mathcal{Z}_t^{\lambda,\mu;1}} \leq \delta,$$

where $\mu > 0$ and $0 < \lambda < \min(\lambda_0, \lambda_L)$. Then there are $C > 0$ and $r, s > 0$ such that the solution of (7.1) satisfies, for all λ', μ' with

$\mu'/\mu < 1$ and $1/2 < \lambda'/\lambda < 1$,

$$\left\| F[f](t,\,\cdot\,) \right\|_{\mathcal{F}^{\lambda' t+\mu'}} \le C\, \left(1 + \delta\, \frac{e^{\frac{1}{(\lambda-\lambda')^r}}}{(\mu-\mu')^s} \right) \varepsilon$$

for all times $t \ge 0$.

Remark 7.2. The assumption on the interaction potential is satisfied for the Coulomb or Newton interaction with $\gamma = 1$. The formulation above allows to discuss the influence of γ on the estimates.

The rest of this chapter is devoted to a presentation of the main ingredients behind Proposition 7.1.

2. Solving the reaction equation

Exactly as in Chapter 3, let us apply the Duhamel principle, the Fourier transform, and integrate over v: with $\rho = \int f\,dv$, we get

$$\begin{aligned}(7.2)\quad & \widehat{\rho}(t,k) = \widetilde{f}_i(k,kt) \\ & + \int_0^t \iint (\nabla W * \rho)\big(\tau, x - v(t-\tau)\big) \cdot (\nabla_v \overline{f^0})\big(\tau, x - v(t-\tau), v\big)\, e^{-2i\pi k\cdot x}\, dx\, dv\, d\tau \\ & + \int_0^t \iint (\nabla W * \rho)\big(\tau, x - v(t-\tau)\big) \cdot (\nabla_v \overline{h})\big(\tau, x - v(t-\tau), v\big)\, e^{-2i\pi k\cdot x}\, dx\, dv\, d\tau.\end{aligned}$$

The first and second terms on the right-hand side are the same as in the linearized study, and the novelty is in the last integral. Since both ρ and $\overline{h}$ depend on x, and the product becomes a convolution in Fourier space, this last integral can be rewritten as

$$\begin{aligned}(7.3)\quad & \\ & \int_0^t \int (\nabla W * \rho) \cdot \nabla_v \overline{h}(\tau,x,v)\, e^{-2i\pi k\cdot x}\, e^{-2i\pi k\cdot v(t-\tau)}\, dv\, d\tau \\ & \qquad \int_0^t \int \sum_{\ell\in\mathbb{Z}^d} \widehat{\nabla W}(k-\ell)\, \widehat{\rho}(\tau, k-\ell)\, (\nabla_v \overline{h})\widehat{\ }(\tau,\ell,v)\, e^{-2i\pi v\cdot k(t-\tau)}\, d\tau\, dv \\ & \qquad = \int_0^t \sum_{\ell} \widehat{\nabla W}(k-\ell)\, \widehat{\rho}(\tau,k-\ell)\, (\widetilde{\nabla_v \overline{h}})\big(\tau,\ell,k(t-\tau)\big)\, d\tau.\end{aligned}$$

The difference with the linearized situation is that now there are all the nonzero values of ℓ, so that the various Fourier modes of ρ are *coupled* to each other. To estimate the expression above, let us use the bound $\widehat{\nabla W}(k-\ell) = O(1/|k-\ell|^{\gamma})$ and the gliding regularity assumption on $\overline{h}$: using $\mathcal{Y}$ spaces for simplicity, we shall assume

$$\widetilde{\nabla_v \overline{h}}\big(\tau, \ell, k(t-\tau)\big) \leq C\,\delta\,|k|(t-\tau)\,e^{-2\pi\overline{\lambda}|k(t-\tau)+\ell\tau|}\,e^{-2\pi\overline{\mu}|k|}.$$

Multiplying this estimate by $e^{2\pi\lambda|k|t}$ and summing over k, we can bound the $\mathcal{F}^{(\lambda t+\mu)}$ norm of the last term in (7.2) by

$$C\int_0^t \sum_{k\neq\ell} \frac{1}{|k-\ell|^\gamma}\,\big|\widehat{\rho}(\tau, k-\ell)\big|\,e^{2\pi(\lambda\tau+\mu)|k-\ell|}\,|k|(t-\tau)$$
$$e^{2\pi\lambda|k(t-\tau)+\ell\tau|}\,e^{2\pi\mu|\ell|}\,e^{-2\pi\overline{\lambda}|k(t-\tau)+\ell\tau|}\,e^{-2\pi\overline{\mu}|\ell|}\,d\tau$$
$$\leq C\delta \int_0^t K(t,\tau)\,\|\rho(\tau)\|_{\mathcal{F}^{\lambda\tau+\mu}}\,d\tau,$$

where the kernel $K(t,\tau)$ is defined by

$$K(t,\tau) = \sup_{k,\ell}\left(\frac{|k|(t-\tau)\,e^{-2\pi(\overline{\lambda}-\lambda)|k(t-\tau)+\ell\tau|}\,e^{-2\pi(\overline{\mu}-\mu)|\ell|}}{1+|k-\ell|^\gamma}\right). \tag{7.4}$$

Notice, we took advantage of the fact that $\overline{\lambda} > \lambda$ to get some decay in the exponentials appearing in (7.4).

Plugging these bounds in (7.2), we conclude that

$$\sum_k e^{2\pi(\lambda t+\mu)|k|t}\left|\widehat{\rho}(t,k) - \int_0^t K^0(t-\tau,k)\,\widehat{\rho}(\tau,k)\,d\tau\right| \tag{7.5}$$
$$\leq A(t) + \delta\int_0^t \|\rho(\tau)\|_{\mathcal{F}^{\lambda\tau+\mu}}\,K(t,\tau)\,d\tau,$$

where $A(t) = \sum e^{2\pi\lambda|k|t}\,|\widetilde{f}_i(k,kt)|$ is the contribution of the initial datum.

To appreciate the effect of the new kernel $K(t,\tau)$ let us set $f^0 = 0$ in (7.5); then this inequality turns into a closed inequality on $\varphi(t) = \|\rho(t,\,\cdot\,)\|_{\mathcal{F}^{\lambda t+\mu}}$:

$$\|\rho(t)\|_{\mathcal{F}^{\lambda t+\mu}} \leq A(t) + C\int_0^t K(t,\tau)\,\|\rho(\tau)\|_{\mathcal{F}^{\lambda\tau+\mu}}\,d\tau. \tag{7.6}$$

Let us analyze the expression (7.4). The decay in ℓ is good, the decay in $|k-\ell|$ is not so good (depending on the smoothness of the interaction), and the decay in k is not good at all since $|k(t-\tau)+\ell\tau|$ can be small even though k is very large: just choose ℓ opposite to k and $\tau = (|k|/|k-\ell|)t$. Stated otherwise, in the time-integral there is a **resonance** phenomenon occurring for

$$k(t-\tau)+\ell\tau = 0.$$

3. Analysis of the kernel K

Let us analyze the kernel appearing in (7.4). Inside the supremum there is a good decay in $|k(t-\tau)+\ell\tau|$, so for practical purposes one may replace the factor $|k|(t-\tau)$ appearing in front of the exponential, by $1+|\ell|\tau$. (This is true also for $\ell=0$: if $f=O(1)$ in gliding analytic regularity, then in general $\nabla_v f = O(t)$, but $\langle \nabla_v f\rangle = \nabla_v \langle f\rangle = O(1)$, where $\langle f\rangle$ is the spatial average of f.) Then it is not difficult to see that $K(t,\tau)$ is not better than $O(\tau)$ and that its time-integral $\int K(t,\tau)\,d\tau$ is not better than $O(t)$. Now this is *bad news*, because it suggests the possibility of a very serious instability as $t\to\infty$, with a growth like, say, e^{t^2}.

But, let us analyze the kernel more precisely. To get an idea of its quantitative behavior, let us set $d=1$, replace the slower power decay $|k-\ell|^{-\gamma}$ by the faster decay $e^{-c|k-\ell|}$, keep only the dominant mode $\ell=-1$ and the modes $k\geq 0$, and replace $|k|(t-\tau)$ by $1+\tau$. The resulting approximation is, up to a multiplicative constant,

$$\overline{K}(t,\tau) = \sup_{k=1,2,\ldots} \Big((1+\tau)\, e^{-2\pi(\overline{\lambda}-\lambda)|k(t-\tau)-\tau|}\, e^{-c\,k}\Big). \tag{7.7}$$

Below is a numerical plot of $\overline{K}(t,\tau)$; that is, the function $\tau \longmapsto K(t,\tau)$ is plotted for various values of t.

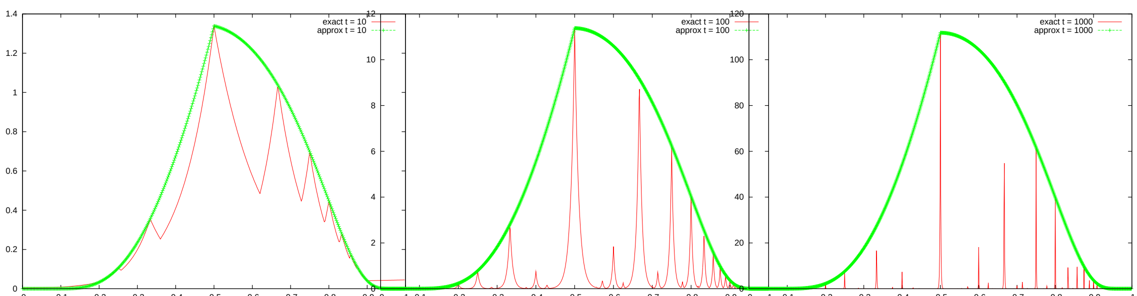

FIGURE 7.1. The kernel $\overline{K}(t,\tau)$ fort $=10$, $t=100$, $t=1000$. The curve above is an approximate envelope which can be computed analytically, but which provides disastrous estimates; observe indeed how most of the mass of the kernel K concentrates on discrete times as t becomes large.

A preliminary conclusion is that in the integral equation

$$\varphi(t) \leq A(t) + \int_0^t K(t,\tau)\,\varphi(\tau)\,d\tau, \tag{7.8}$$

for each t, only certain specific values of τ seem to count (like $t/2$, etc.). This may seem crazy at first, but it is exactly the same principle which underlies the **echo phenomenon**, a beautiful experiment made in the sixties by Malmberg and co-workers. Let me describe the experiment and its interpretation. Prepare your favorite plasma in your favorite lab, and at time 0 excite it by a small impulse of frequency $\ell \in \mathbb{Z}$, say $\ell < 0$. Wait until the electric field damps, and at time $\tau > 0$ excite the plasma again by a small impulse of frequency $k - \ell$, with $k \in \mathbb{Z}$, $k > 0$. Then sit and measure the electric field, analyzing the strength of the mode k. Around time

$$t_e = \frac{(k-\ell)\tau}{k}, \tag{7.9}$$

a spontaneous response, the *echo*, will be recorded.

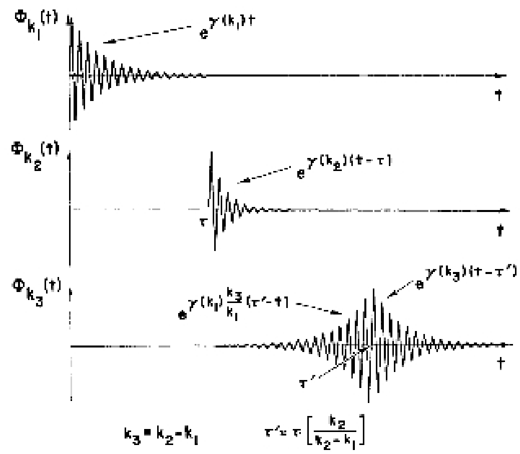

FIG. 1. Approximate variation of the principal Fourier coefficients of the self-consistent field for the case $k_3 \cong k_1 \cong \frac{1}{2}k_2$. Upper line: response to the first pulse; middle line: response to the second pulse; lower line: echo.

Figure 7.2. Representation of the plasma echo experiment, from the pioneering paper by Malmberg et al.

The interpretation is the following: initially disturbed, the electric field has damped, but the information is still there, *hidden* in the fast oscillations of the distribution function in the velocity variable. Start from a homogeneous background $f^0(v)$, apply an impulse in time, from the Vlasov equation the variation in the density is $-F(x) \cdot \nabla_v f^0(v)$, so right after the pulse the density will be roughly

$$f^0(v) - 2i\pi c\, e^{2i\pi\ell\cdot x}\, \ell \cdot \nabla_v f^0(v),$$

where c is a small constant and I use a complex notation for simplicity (only the real part makes sense). Then the distribution evolves by damping and oscillates more and more; if we retain only the contribution of free transport we obtain, at time τ,

$$f^0(v) - 2i\pi c\, e^{2i\pi\ell\cdot(x-v\tau)}\,\ell\cdot\nabla_v f^0(v).$$

At time τ we apply the new pulse; now the change in the density is proportional to $-F\cdot\nabla_v f$, where the force now oscillates at spatial frequency $k-\ell$: this gives an additional term proportional to

$$e^{2i\pi(k-\ell)\cdot x}\,\ell\tau\, e^{2i\pi\ell\cdot(x-v\tau)}\,\ell\cdot\nabla_v f^0(v) + e^{2i\pi\ell\cdot(x-v\tau)}\,\big(\ell\otimes(k-\ell)\big):\nabla_v^2 f^0(v).$$

The interesting term is the one involving both k and $k-\ell$ spatial frequencies; it evolves again at dominant order by free transport after time τ, with a contribution proportional to

$$e^{2i\pi(k-\ell)\cdot[x-v(t-\tau)]}\,e^{2i\pi\ell\cdot(x-vt)} = e^{2i\pi k\cdot x}\,e^{-2i\pi v\cdot[kt-(k-\ell)\tau]}.$$

As long as $kt-(k-\ell)\tau$ is large, the oscillations are very fast in the v variable; but when this becomes small, that is around time $t\simeq t_e$, then the oscillations in the velocity variable are not fast, and this gives a strong contribution to the electric field, located at spatial frequency k.

The beauty of the echo experiment is that it demonstrates that in the Landau damping phenomenon, the information has not disappeared: it is still there, but hidden in the high frequency oscillations in velocity. The interaction between the two spatial frequencies ℓ and $k-\ell$ has produced a response which can be measured: a visible manifestation of what was meant to remain hidden.

Back to the nonlinear damping problem, here is the picture which starts to emerge. While in the linearized Vlasov equation, each mode k was evolving independently of the other ones, in the nonlinear Vlasov equation that is not the case: there is a **coupling** of all modes by the interaction. Exciting one mode at some time has a nonnegligible conclusion on other modes at later times (by echoes), but this is controlled by the integral equation, mixing estimates on all modes together.

4. Analysis of the integral equation

Now let us come back to the analysis of the integral equation

$$\varphi(t) = A(t) + \int_0^t \overline{K}(t,\tau)\,\varphi(\tau)\,d\tau \tag{7.10}$$

appearing as a variant of (7.8).

From the past section, we have a bad news, namely that the kernel grows linearly in time; and a good news, namely that it concentrates on "resonant" times τ, *which are not too close to* t. The moral is the same as can be derived from the echo experiment: the Vlasov equation is an oscillatory system *which responds with time-delay.* To illustrate why this is a good news, let us examine a few examples of baby integral equations.

• a kernel that is uniformly $O(\tau)$:

$$\varphi(t) \leq A + c \int_0^t \tau\, \varphi(\tau)\, d\tau;$$

then this yields $\varphi(t) \leq A\, e^{ct^2/2}$, which is a disaster.

• a kernel whose integral is $O(t)$, and which is spread over times:

$$\varphi(t) \leq A + c \int_0^t \varphi(\tau)\, d\tau;$$

then $\varphi(t) \leq A\, e^{ct}$, which is better.

• a kernel whose integral is $O(t)$, which is concentrated at the final time:

$$\varphi(t) \leq A + c\, t\, \varphi(t) :$$

this is a complete disaster, the inequality does not even prevent φ from becoming infinite.

• a kernel whose integral is $O(t)$, which is $O(\tau)$, and concentrated near the final time:

$$\varphi(t) \leq A + c\, t \int_{t-1}^t \varphi(\tau)\, d\tau;$$

then this still allows for violent growth.

• a kernel whose integral is $O(t)$, but whose mass concentrates far away from t:

$$\varphi(t) \leq A + c\, \frac{t}{2}\, \varphi\left(\frac{t}{2}\right). \tag{7.11}$$

Then a power series expansion suggests

$$\varphi(t) \leq A \sum_n \frac{c^n\, t^n}{2^{n(n-1)/2}},$$

which is basically the same as $A^{c'(\log t)^2}$; one can also guess this behavior directly from (7.11). This is *very good* since this growth is subexponential (faster than any polynomial, slower than any fractional exponential). In particular, we can write $\varphi(t) \leq C_\varepsilon\, A\, e^{\varepsilon t}$, where ε is arbitrarily small and C_ε depends on ε.

Why is this good? Recall that in our case $\varphi(t)$ is an analytic norm whose regularity index increases *exponentially fast* with time, say $\varphi(t) = \|\rho(t)\|_{\mathcal{F}^{\lambda t+\mu}}$. Then for the force $F(t)$ we have, for $\lambda' < \lambda$, using as usual $\widehat{F}(t,0) = 0$,

$$\begin{aligned}\|F(t)\|_{\mathcal{F}^{\lambda' t+\mu}} &\le e^{-2\pi(\lambda-\lambda')t}\,\|F(t)\|_{\mathcal{F}^{\lambda t+\mu}}\\ &\le C\,e^{-2\pi(\lambda-\lambda')t}\,\|\rho(t)\|_{\mathcal{F}^{\lambda t+\mu}}\\ &\le C\,C_\varepsilon\,A\,e^{-2\pi(\lambda-\lambda')t}\,e^{\varepsilon t},\end{aligned}$$

and by choosing ε close enough we can make sure that the decay of the force is still exponential in time.

5. Effect of singular interactions

In the previous discussion and analysis of the kernel, we have replaced the power law decay $\widehat{W}(k) \simeq |k|^{-(1+\gamma)}$ of the interaction by the exponential decay $e^{-c|k|}$ which is typical of an analytic interaction. But of course, the most interesting cases occur when $\widehat{W}(k)$ only decay like a power law, corresponding to a singularity in physical space. The most important case of all is $\gamma = 1$ (Poisson coupling). How does this singularity modify the picture which we formed for an analytic interaction?

To get a feeling, and appreciate the influence of the strength of the singularity, let us consider the baby kernels

$$K_\gamma(t,\tau) = (1+\tau)\sup_{k=1,2,\ldots}\frac{e^{-\alpha|kt-(k+1)\tau|}}{(k+1)^\gamma}.$$

To appreciate the long-time behavior, let us perform a time-rescaling, setting $k_t(\theta) = t\,K(t,t\theta)$ (the t factor in front is there to keep the total mass of K invariant in the rescaling). As $t\to\infty$, the exponential $e^{-\alpha|kt-(k+1)\theta|}$ becomes localized in a neighborhood of size $O(1/kt)$ around $\theta = k/(k+1)$, and its mass becomes $2/(\alpha(k+1))$. Thus

$$\frac{k_t}{t} \longrightarrow \frac{2}{\alpha}\sum_k \frac{1}{(k+1)^\gamma}\,\frac{k}{(k+1)^2}\,\delta_{1-\frac{1}{k+1}},$$

where the convergence is in the weak sense on the time-interval $[0,t]$. This suggests the approximation

$$K_\gamma(t,\tau) \simeq c\,t\sum_{k\ge 1}\frac{1}{k^{1+\gamma}}\,\delta_{\left(1-\frac{1}{k}\right)t}. \tag{7.12}$$

Examination of (7.12) shows that the lower γ is, the more the echoes accumulate near t; then we are getting closer to the (very bad) regime of instantaneous response. For pedagogical purpose, one may keep in mind the image that if one sings in a very rough church (say), the abundance of small echoes may blur the sound in an uncontrollable way.

To evaluate the influence of the kernel (7.12), let us search once again for an exact power series solution $\varphi(t) = \sum a_n t^n$ to the integral equation

$$\varphi(t) = A + c\,t \sum_{k\geq 1} \frac{1}{k^{1+\gamma}}\,\varphi\left(\left(\frac{k-1}{k}\right)t\right).$$

This yields

$$a_0 = A, \qquad a_{n+1} = c\left[\sum_{k\geq 1} \frac{1}{k^{1+\gamma}}\left(1-\frac{1}{k}\right)^n\right] a_n.$$

The sum inside brackets is comparable to

$$\int_1^\infty t^{-(1+\gamma)}\left(1-\frac{1}{t}\right)^n dt = B(\gamma, n+1) \simeq n^{-\gamma},$$

where B is Euler's Beta function. So $a_{n+1} \simeq c\,n^{-\gamma}\,a_n$, thus $a_n \simeq A\,c^n/(n!)^\gamma$, and we expect

$$\varphi(t) \leq \text{const.}\, A \sum_{n\geq 1} \frac{c^n\,t^n}{n!^\gamma}. \tag{7.13}$$

This is subexponential for $\gamma > 1$ (which is good), but exponential for $\gamma = 1$!

Since $\gamma = 1$ is the most interesting case, it is tempting to believe that we stumbled on some deep difficulty. But this is a trap: a much more precise estimate can be obtained by *separating modes* and estimating them one by one, rather than seeking for an estimate on the whole norm. Namely, if we set

$$\varphi_k(t) = e^{2\pi(\lambda t+\mu)|k|}\,|\widehat{\rho}(t,k)|,$$

then we have a system of the form

$$\varphi_k(t) \leq a_k(t) + \frac{c\,t}{(k+1)^{\gamma+1}}\,\varphi_{k+1}\left(\frac{kt}{k+1}\right). \tag{7.14}$$

Let us assume that $a_k(t) = O(e^{-ak}\,e^{-2\pi\lambda|k|t})$. First we simplify the time-dependence by letting

$$A_k(t) = a_k(t)\,e^{2\pi\lambda|k|t}, \qquad \Phi_k(t) = \varphi_k(t)\,e^{2\pi\lambda|k|t}.$$

Then (7.14) becomes

$$\Phi_k(t) \le A_k(t) + \frac{c\,t}{(k+1)^{\gamma+1}}\,\Phi_{k+1}\left(\frac{kt}{k+1}\right). \tag{7.15}$$

(The exponential for the last term is right because $(k+1)(kt/(k+1)) = kt$.) Now if we get a subexponential estimate on $\Phi_k(t)$, this will imply an exponential decay for $\varphi_k(t)$.

Once again, we look for a power series, assuming that A_k is constant in time, decaying like e^{-ak} as $k \to \infty$; so we make the ansatz $\Phi_k(t) = \sum_m a_{k,m}\,t^m$ with $a_{k,0} = e^{-ak}$. As an exercice, the reader can work out the doubly recurrent estimate on the coefficients $a_{k,m}$ and deduce

$$a_{k,m} \le \text{const.}\, A\,(k\,e^{-ak})\,k^m\,c^m\,\frac{e^{-am}}{(m!)^{\gamma+2}},$$

whence

$$\Phi_k(t) \le \text{const.}\, A\,e^{(1-\alpha)(ckt)^\alpha}, \qquad \forall \alpha < \frac{1}{\gamma+2}. \tag{7.16}$$

This is subexponential even for $\gamma = 1$: in fact, we have taken advantage of the fact that *echoes at different values of k are asymptotically rather well separated in time.*

As a conclusion, as an effect of the singularity of the interaction, *we expect to lose a fractional exponential on the convergence rate*: if the mode k of the source decays like $e^{-2\pi\lambda|k|t}$, then φ_k, the mode k of the solution, should decay like $e^{-2\pi\lambda|k|t}\,e^{(c|k|t)^\alpha}$. More generally, if the mode k decays like $A(kt)$, one expects that $\varphi_k(t)$ decays like $A(kt)\,e^{(c|k|t)^\alpha}$. Then we conclude as before by absorbing the fractional exponential in a very slow exponential, at the price of a *very large* constant: say

$$e^{t^\alpha} \le \exp\!\left(c\,\varepsilon^{-\frac{\alpha}{1-\alpha}}\right) e^{\varepsilon\,t}.$$

6. Large time estimates via exponential moments

So far we have mainly done heuristics and power expansions, now arises the question of rigorously estimating solutions of integral equations. Let us leave apart the more tricky case when the modes are decoupled, and focus on the single case when there is just one equation, like (7.6).

So let $\varphi(t) \ge 0$ solve

$$\varphi(t) \le A + \int_0^t K(t,\tau)\,\varphi(\tau)\,d\tau,$$

where $K(t,\tau)$ is given by (7.4). To estimate φ in an exponential scale, we shall consider exponential moments of the kernel. The main idea is that

$$\int_0^t e^{-\varepsilon t} K_\gamma(t,\tau)\, e^{\varepsilon\tau}\, d\tau \tag{7.17}$$

will be smaller if K favors large values of $t-\tau$.

It can be shown by elementary means that for $t \geq 1$,

$$\int_0^t e^{-\varepsilon t} K(t,\tau)\, e^{\varepsilon\tau}\, d\tau \leq \frac{C}{\varepsilon^r\, t^{\gamma-1}}, \tag{7.18}$$

for some constants $C > 0$, $r > 0$. The important fact is that the bound on the right-hand side of (7.18) decays as $t \to \infty$.

Let us see how to exploit this information. Let $\psi(t) = B\, e^{\varepsilon t}$. If ψ satisfies

$$\begin{cases} \varphi(t) < \psi(t) & \text{for } 0 \leq t \leq T \\ \psi(t) \geq A + \int_0^t K(t,\tau)\, \psi(\tau)\, d\tau, \end{cases}$$

then $u := \psi - \varphi$ is positive for $t \geq T$, and satisfies the inequality $u(t) \geq \int_0^t K(t,\tau)\, u(\tau)\, d\tau$ for $t \geq T$, so u will never vanish and always remain positive — this is a maximum principle argument.

For small values of t, that is, $0 \leq t \leq T$, a crude bound, in Gronwall style, is easy: it may give a very bad constant like e^{T^2} or so, but that remains a finite constant, whatever the choice of T.

For large values of t, that is $t > T$, we just have to check that

$$A + B \int_0^t K(t,\tau)\, e^{\varepsilon\tau}\, d\tau \leq B\, e^{\varepsilon t}.$$

But from (7.18), the left-hand side is bounded above by

$$A + \frac{B\, C\, e^{\varepsilon t}}{\varepsilon^r\, t^{\gamma-1}},$$

and this is obviously bounded above by $B\, e^{\varepsilon t}$ as soon as $B \geq A/2$ and $t \geq (2BC/\varepsilon^r)^{1/(\gamma-1)}$, which in turn holds as soon as T is chosen large enough.

The estimate can be refined in many ways. Instead of exponential moments, one can consider fractional exponential moments

$$\int e^{-\varepsilon\, t^\alpha} K(t,\tau)\, e^{\varepsilon\, \tau^\alpha},$$

which gives much better results as far as the dependence on ε is concerned.

Also, there is a variant which covers the case when the kernel is used as a source term in the linearized Vlasov equation with a nontrivial f^0, as in (7.5). This argument is more tricky and goes not only through a maximum principle but also through L^2 estimates (as in Lemma 3.5) and an inequality of Young type. To work this out, one establishes decay estimates not only on L^1 exponential moments as (7.17), but also decay estimates on L^2 moments

$$\left(\int e^{-2\varepsilon t}\, K(t,\tau)^2\, e^{2\varepsilon\tau}\, d\tau\right)^{1/2},$$

and uniform bounds on dual L^1 moments,

$$\sup_{\tau\geq 0}\int_\tau^\infty e^{\varepsilon\tau}\, K(t,\tau)\, e^{-\varepsilon t}\, dt.$$

The elementary method from Lemma 3.5 can then be adapted to this tricky situation. (This is somewhat painful, but the use of the inversion of the Laplace transform would probably have been quite more painful.)

Finally, there is also a variant which allows to estimate the norms of all modes separately, and thus to treat the important case $\gamma = 1$.

Bibliographical notes

Basically all this chapter is taken from [**74**, Sections 6 and 7], where more precise computations and estimates are established. (But the efficiency of fractional exponential moments is not noticed in that reference.)

The echo experiment appears at the end of the sixties, in [**65**] (prediction) and [**66**] (report of experiment); I learnt it from Kiessling. In fact at first it was *spatial echoes* which were observed, and it is only later that temporal echoes could be produced. Nowadays they are used as an indirect way to measure the strength of irreversible phenomena going on in a plasma (defect of echo indicates dissipation!), see [**91**].

CHAPTER 8

Newton's scheme

In the past two chapters we have examined the two sides of the nonlinear Vlasov equation near equilibrium, first as a transport equation in a small force field whose regularity improves with time, secondly as the reaction for a gas forcing an oscillating background which is a perturbation of equililbrium. In both cases we obtained estimates in gliding regularity that are uniform in time, at the price of a *loss of regularity*, or consequently a loss on the time decay rate. (Recall that the gliding regularity automatically implies a time decay on velocity averages.)

Loss of regularity in the solution of the linearized problem is sometimes informally called the **Nash–Moser syndrome**. It was overcome in the fifties by Kolmogorov (in the proof of his celebrated 1954 theorem of the likely stability of trajectories of perturbed completely integrable Hamiltonian systems) and by Nash (in his celebrated 1956 construction of smooth isometric Riemannian embeddings). In both cases a key idea was to work out a perturbative analysis based on the **Newton iterative scheme** and use the supernatural speed of convergence of this scheme to overcome the loss of regularity. Nash also showed how to take advantage of this fast convergence to squeeze in a *regularization* at each stage, giving birth to what is now called the Nash–Moser method, arguably the most powerful perturbative technique known to this date. Moser used it to prove Kolmogorov's theorem in C^r regularity.

In the study of Landau damping, the mighty Newton scheme will save us again. It does not mean that this is the only solution: History has shown (for the Kolmogorov theorem and even more for the Nash theorem) that the Newton scheme can sometimes be replaced by a classical fixed point technique, applied to a clever reformulation of the problem.

0. The classical Newton scheme

The general formulation of the Newton scheme is as follows. Let be given an equation $\Phi(z) = 0$, where the unknown z lies in $\mathbb{R}$ or in some Banach space, and the equation should be solved in a neighborhood V

where the differential $D\Phi$ is invertible. Start from some guess z_0 and solve iteratively, approximating Φ at each step by its tangent to the previous approximation. So at step n, the equation to be solved is

$$\Phi(z_n) + D\Phi(z_n) \cdot (z_{n+1} - z_n) = 0, \tag{8.1}$$

or equivalently

$$z_{n+1} = z_n - [D\Phi(z_n)]^{-1} \cdot \Phi(z_n).$$

If z_{n+1} always remains in V then this procedure defines inductively a sequence $(z_n)_{n\in\mathbb{N}}$. Clearly if z_n converges to z, then from (8.1) we have $\Phi(z) = 0$. The claim is that if z_0 is close enough to the desired solution, and if Φ is twice continuously differentiable, then the convergence occurs *extremely fast*. To see this, use the Taylor expansion and (8.1) to deduce

$$\begin{aligned}\|\Phi(z_{n+1})\| &\le \Big\|\Phi(z_n) + D\Phi(z_n)\cdot(z_{n+1}-z_n)\Big\| + \frac{\|D^2\Phi\|_\infty}{2}\|z_{n+1}-z_n\|^2\\ &= \frac{\|D^2\Phi\|_\infty}{2}\|z_{n+1}-z_n\|^2,\end{aligned} \tag{8.2}$$

where $\|\cdot\|_\infty$ is the supremum norm over the domain V.

Plugging (8.2) in the identity $\Phi(z_{n+1})+D\Phi(z_{n+1})\cdot(z_{n+2}-z_{n+1}) = 0$ yields $\|D\Phi(z_{n+1})\cdot(z_{n+2}-z_{n+1})\| \le (\|D^2\Phi\|_\infty/2)\|z_{n+1}-z_n\|^2$, whence

$$\|z_{n+2}-z_{n+1}\| \le \left(\frac{\|(D\Phi)^{-1}\|_\infty\,\|D^2\Phi\|_\infty}{2}\right)\|z_{n+1}-z_n\|^2.$$

Iteration of the inequality $\|z_{n+2}-z_{n+1}\| \le C\,\|z_{n+1}-z_n\|^2$ yields

$$\|z_{n+1}-z_n\| \le C^n\,\|z_1-z_0\|^{2^n}. \tag{8.3}$$

Now if $\|\Phi(z_0)\|$ is small enough, then $\delta := \|\Phi(z_0)\|\,\|D\Phi(z_0)^{-1}\|$ is strictly less than 1, so $\|z_1-z_0\| < \delta$, and (8.3) implies inductively

$$\|z_{n+1}-z_n\| \le C^n\,\delta^{2^n}.$$

Then of course (z_n) converges to z, and by telescopic summation

$$\|z_n - z\| \le C^n\,\delta^{2^n}\Big[1 + C\,\delta^{2^n} + C^2\,\delta^{2^n\cdot 4^n} + C^3\,\delta^{2^n\cdot 4^n\cdot 8^n} + \dots\Big],$$

which is bounded above by $2\,C^n\delta^{2^n}$ if δ is small enough. Up to changing C, we conclude that

$$\|z_n - z\| \le C^n\,\delta^{2^n}. \tag{8.4}$$

That is, the Newton method converges like an **iterated exponential** (exponential of exponential). (One often says that the convergence is quadratic to express the fact that the number of significant digits doubles at each step.) All this is subject to the fact that (z_n) remains inside the neighborhood V where the root z belongs; but in view of the estimate (8.4), this is clearly the case if z is close enough to z_0.

1. Newton scheme for the nonlinear Vlasov equation

Consider an evolution partial differential equation $\partial_t f = Q(f)$, where the unknown is a solution $f = (f(t))_{t\geq 0}$ and the initial datum f_i is prescribed. To cast this equation in the setting of the Newton scheme, define

$$\Phi(f) = \left(\frac{\partial f}{\partial t} - Q(f),\ f(0) - f_i \right).$$

Then the equation $\Phi(f^n) + D\Phi(f^n)\cdot(f^{n+1} - f^n) = 0$ means

(8.5)
$$\begin{cases} \Big[\partial_t f^n - Q(f^n)\Big] + \partial_t (f^{n+1} - f^n) - Q'(f^n)\cdot(f^{n+1} - f^n) = 0 \\ f^n(0) = f_i \quad \text{for all } n. \end{cases}$$

The first equation is $\partial_t f^{n+1} = Q(f^n) - Q'(f^{n-1})\cdot(f^{n+1} - f^n)$, but this is not the most convenient form. It is best to see f^n as made of a series of successive layers: $f^n = f^0 + h^1 + \ldots + h^n$, where the unknowns h^n solve

$$\begin{aligned} \partial_t h^{n+1} &= Q'(f^n)\cdot h^{n+1} + Q(f^n) - \partial_t f^n \\ &= Q'(f^n)\cdot h^{n+1} + \Big[Q(f^{n-1} + h^n) - Q(f^{n-1}) - Q'(f^{n-1})\cdot h^n\Big], \end{aligned}$$

together with the initial conditions $h^1(0) = h_i$, $h^{n+1}(0) = 0$.

In the case of the nonlinear Vlasov equation, the nonlinearity is quadratic, so $Q(f^{n-1} + h^n) - Q(f^{n-1}) - Q'(f^{n-1})h^n = -F[h^n]\cdot\nabla_v h^n$. Then we arrive at the **Newton scheme for the nonlinear Vlasov equation** near a spatially homogeneous equilibrium f^0. First $f^0 = f^0(v)$ is given, then $f^n = f^0 + h^1 + \ldots + h^n$, where h^1 solves the linearized Vlasov equation

$$\begin{cases} \dfrac{\partial h^1}{\partial t} + v\cdot\nabla_x h^1 + F[h^1]\cdot\nabla_v f^0 = 0 \\ h^1(0,\,\cdot\,) = f_i - f^0, \end{cases} \tag{8.6}$$

and, for any $n \geq 1$,

(8.7)
$$\begin{cases} \dfrac{\partial h^{n+1}}{\partial t} + v\cdot\nabla_x h^{n+1} + F[f^n]\cdot\nabla_v h^{n+1} + F[h^{n+1}]\cdot\nabla_v f^n = -F[h^n]\cdot\nabla_v h^n \\ h^{n+1}(0,\,\cdot\,) = 0. \end{cases}$$

In this way the nonlinear Vlasov equation has been reduced to an infinite number of linear equations, each of which involves a source term which is quadratic in the solution of the previous equation.

The Newton scheme destroys many of the properties and invariances of the original equation, however note that it is still in divergence form, so

$$\forall n \geq 2, \quad \forall t \geq 0, \qquad \iint h^n(t,x,v)\,dx\,dv = 0. \tag{8.8}$$

For $n = 1$ we already know that

$$\forall t \geq 0, \qquad \iint h^1(t,x,v)\,dx\,dv = \iint (f_i - f^0)(x,v)\,dx\,dv.$$

Now the goal is to study the various layers h^n. We shall do this in two stages: short time, large time.

2. Short time estimates

Proposition 8.1. *Let $f^0 = f^0(v)$ be a spatially homogeneous profile satisfying $\|f^0\|_{\mathcal{Z}^{\lambda_0;1}} \leq C_0$ for some $\lambda_0 > 0$, and let W be an interaction potential with $\nabla W \in L^1(\mathbb{T}^d)$. If $\lambda' < \lambda < \lambda_0$ and $1 < a < 2$, then there are $\varepsilon_* > 0$ and $T_* > 0$ such that if f_i satisfies*

$$\|f_i - f^0\|_{\mathcal{Z}^{\lambda,\mu}} \leq \varepsilon,$$

then for $\varepsilon \leq \varepsilon_*$ one has

$$\forall n \in \mathbb{N}, \quad \forall t \in [0, T_*], \qquad \|h^n(t)\|_{\mathcal{Z}^{\lambda,\mu;1}} \leq C\,\varepsilon^{a^n}.$$

Remark 8.2. In this short-time estimate, it does not matter whether we use gliding regularity or not.

The Vlasov equation is a first-order nonlinear partial differential equation; for such equations there is a general method to establish short-time analytic estimates, known as the **Cauchy–Kowalevskaya theory**. For our current purposes, we do not need to explicitly appeal to that theory, and can give a self-contained treatment using the hybrid analytic norms, with a *regularity index which deteriorates in time.*

Lemma 8.3. *For t small enough,*

$$\frac{d^+}{dt}\|f\|_{\mathcal{Z}^{\lambda-Kt,\mu-Kt};1} \leq -c\,K\,\|\nabla f\|_{\mathcal{Z}^{\lambda-Kt,\mu-Kt};1}.$$

The proof of this lemma is easy and relies mainly on the identity $(d/dt)e^{2\pi\lambda|k|t} = 2\pi\lambda|k|\,e^{2\pi\lambda|k|t}$.

Sketch of proof of Proposition 8.1. Now let us estimate $h^n(t,\cdot)$ in a norm $\mathcal{Z}^{\lambda_n - Kt, \mu_n - Kt}$. We get rid of the linear transport term by using the free transport semigroup, and we are left with the contribution of the force term. So the estimate of the variation of the norm will involve several terms, one of which is nonlinear and involves derivatives of f^n and h^{n+1}, and one of which comes from the time-variation of the regularity index; that one is linear and proportional to $-\|\nabla h^{n+1}\|$. So

$$\frac{d^+}{dt}\|h^{n+1}\| \leq -K\|\nabla h^{n+1}\| + C\,\|\nabla(f^n - f^0)\|\,\|\nabla h^{n+1}\| + C\|\nabla h^n\|^2 + \dots \tag{8.9}$$

and all the norms are $\mathcal{Z}^{\lambda_{n+1}-Kt, \mu_{n+1}-Kt;1}$. The amount of regularity lost with time is the same for all indices n, and if λ_n, μ_n remain bounded from below and t is small enough, then these norms control a fixed norm $\mathcal{Z}^{\lambda,\mu;1}$.

Let us see how to estimate (8.9). We shall use the shorthand $\|h\|_{\lambda_n} = \|h\|_{\mathcal{Z}^{\lambda_n - Kt, \mu_n - Kt;1}}$ and assume that

$$\|h\|_{\lambda_n} \leq \delta_n,$$

then the goal is to get recursive estimates on δ_n. Using (5.9)–(5.10), we relate the norm of the gradient to the norm of the function, stratifying at the same time the estimate by estimating each layer h^n in a different norm (the regularity deteriorates with n, so we have more information for lower indices n). Thus we write

$$\begin{aligned}
\|\nabla(f^n - f^0)\|_{\lambda_{n+1}} &\leq \|\nabla h^1\|_{\lambda_{n+1}} + \|h^2\|_{\lambda_{n+1}} + \dots \\
&\leq \frac{\|h^1\|_{\lambda_1}}{\lambda_1 - \lambda_{n+1}} + \frac{\|\nabla h^2\|_{\lambda_2}}{\lambda_2 - \lambda_{n+1}} + \dots \\
&\leq \frac{\delta_1}{\lambda_1 - \lambda_{n+1}} + \frac{\delta_2}{\lambda_2 - \lambda_{n+1}} + \dots + \frac{\delta_n}{\lambda_n - \lambda_{n+1}}.
\end{aligned}$$

Assuming that the sequence $(\lambda_n - \lambda_{n+1})$ is decreasing, this is grossly bounded above by

$$\sum \frac{\delta_n}{\lambda_n - \lambda_{n+1}}, \tag{8.10}$$

and if that sum is small enough then the second term on the right-hand side of (8.9) is absorbed by the first one. After bounding in a similar way the last term on the right-hand side of (8.9), we end up with something like

$$\delta_{n+1} \leq \frac{C\,\delta_n^2}{\lambda_n - \lambda_{n+1}}. \tag{8.11}$$

At this stage we can choose, say, $\lambda_n - \lambda_{n+1} = \Lambda/n^2$, where Λ is very small: that is, we allow the regularity to decrease at each step in a controlled way. If $\sum n^2\,\delta_n$ is small enough, then the sum in (8.10) is very small, so (8.11) holds true, and the recursion relation $\delta_{n+1} \le C'\,n^2\,\delta_n^2$ implies that $\delta_n = O(\delta_1^{a^n})$ for $a < 2$, which a posteriori justifies the assumption of smallness of $\sum n^2\,\delta_n$. It is easy to conclude by propagating bounds inductively. $\square$

3. Large time estimates

Now the goal is to go for long-time estimates on the layers h^n, and establish

Proposition 8.4. *Let $f^0 = f^0(v)$ be a spatially homogeneous profile satisfying $|\widetilde{f^0}(\eta)| = O(e^{-2\pi\lambda_0|\eta|})$ and $\|f^0\|_{\mathcal{Z}^{\lambda_0;1}} < +\infty$, together with the Penrose stability condition with stability width $\lambda_L > 0$. Let W be an interaction potential with $\widehat{W} = O(1/|k|^2)$. Let $\lambda > 0$ and $\lambda' < \min(\lambda, \lambda_0, \lambda_L)$, let $\mu > 0$, let $a \in (1,2)$; then there is $\varepsilon_* > 0$ such that if f_i satisfies $\|f_i - f^0\|_{\mathcal{Z}^{\lambda,\mu;1}} \le \varepsilon \le \varepsilon_*$, then*

$$\forall n \in \mathbb{N}, \quad \forall t \ge 0 \qquad \|h^n(t)\|_{\mathcal{Z}_t^{\lambda,\mu;1}} \le C\,\varepsilon^{a^n}.$$

This is much more tricky than the short-time estimates. In particular, it will involve a Lagrangian point of view, where we will use trajectories induced by the force field rather than just free transport; and it will use the estimates from Chapters 6 and 7. The same global strategy of stratification of estimates will be useful, but a number of auxiliary estimates will be propagated. I shall only describe the main ideas in a sketchy way.

Sketch of proof of Proposition 8.4. First the general picture: Starting from

$$\partial_t h^{n+1} + v\cdot\nabla_x h^{n+1} + F[f^n]\cdot\nabla_v h^{n+1} + F[h^{n+1}]\cdot\nabla_v f^n = -F[h^n]\cdot\nabla_v h^n,$$

we formally get rid of the $F[f^n]\cdot\nabla_v h^{n+1}$ by using the characteristics S^n associated with the force field $F[h^n](t,x)$. This gives a kind of reaction equation in the style of (7.1), except that everything is composed with S^n. A notable unpleasant consequence is that we lose the gradient property: $(\nabla_v f^n)\circ S^n$ is not a gradient any longer; as a consequence, an additional zero mode will appear in the reaction estimates, associated with instantaneous response. Fortunately, this term will be uniformly bounded in time. The source term, quadratic in h^n, will not cause any serious problem.

From there, one sets up a constructive loop in the estimates: If the characteristics are close to the free transport trajectories, then the flow will have good mixing properties, and as a consequence the density will be uniformly smooth and the force will damp to 0. Conversely, if the force damps, then the characteristics remain close to free transport trajectories.

To quantify how close the characteristics are from free transport, one introduces the "finite time scattering operators"

$$\Omega_{t,\tau} = S^n_{t,\tau} \circ S^0_{\tau,t}, \tag{8.12}$$

where S^n is the characteristic flow generated by the force $F[h^n]$.

Now the core of the proof is to estimate simultaneously $h^n \circ \Omega^{n-1}$ (the density h^n along the characteristics generated by f^{n-1}) and $\rho^n = \int h^n\,dv$ (the spatial density). The density ρ^n is estimated in the natural gliding regularity: that is, in the space $\mathcal{F}^{\lambda_n\tau+\mu_n}$ at time τ. But the composed density $h^n \circ \Omega^{n-1}$ is estimated with a twist on the indices, depending on the final time t: using (5.14), the time velocity and regularity indices will be modulated by a function

$$b(t) = B/(1+t)$$

So the main estimates to propagate are

$$\sup_{\tau\geq 0} \|\rho^n(\tau)\|_{\mathcal{F}^{\lambda_n\tau+\mu_n}} \leq \delta_n, \qquad \sup_{t\geq\tau\geq 0} \left\| h^n_\tau \circ \Omega^{n-1}_{t,\tau} \right\|_{\mathcal{Z}^{\lambda_n(1+b),\mu_n;1}_{\tau-\frac{bt}{1+b}}} \leq \delta_n, \tag{8.13}$$

where δ_n is a sequence of positive numbers which should converge very fast to 0. I shall write $\delta = \delta_1$: this is an estimate of the final error.

A number of auxiliary estimates will be propagated. Writing $\Omega = \Omega_{t,\tau}$ and $h = h(\tau)$ for simplicity, the desired estimates are (in appropriate norms)

- $\Omega^n - \mathrm{Id}$ and $\nabla\Omega^n - I$ are $O(\delta/\tau^s)$, uniformly in n;
- $\Omega^n - \Omega^k$ and $(\Omega^k)^{-1} \circ \Omega^n - \mathrm{Id}$ are $O(\delta_k/\tau^s)$, for all $k \leq n$ (so these expressions are small as $k \to \infty$, uniformly in n);
- $h^k \circ \Omega^{n-1}$, $\nabla_x h^k \circ \Omega^{n-1}$, $(\nabla_v + \tau\nabla_x)h^k \circ \Omega^{n-1}$ are all $O(\delta_k)$, for all $k \leq n$;
- $\nabla^2 h^k \circ \Omega^{n-1}$ is $O(\tau^2\,\delta_k)$, for all $k \leq n$;
- $(\nabla h^n) \circ \Omega^{n-1} - \nabla(h^n \circ \Omega^{n-1})$ is $O(\delta_n/\tau^s)$, for all n.

In the above expressions $s > 0$ is as large as desired; for the sequel it would be sufficient to choose $s = 4$, but in these and related estimates I shall continue to write s, meaning an integer which can change from line to line and can be fixed arbitrarily large in advance. The possibility of choosing s large comes from the fact that the scattering estimates of Chapter 6 are exponentially small in τ.

Let us see in a sketchy way how this works. To simplify notation, I shall forget about the x-regularity and the parameter μ in the estimate of the kinetic distribution, and focus on the v-regularity parameter λ. At each stage of the iteration a bit or regularity is lost in v (when I say a bit, this is still an infinite number of derivatives, but in the parameter λ of analytic regularity this is just a bit): say $\lambda_n - \lambda_{n+1} = \Lambda/n^2$ with Λ very small. At each stage there are a number of steps, at each step a little bit is lost, so five or six intermediate indices are used between λ_n and λ_{n+1}, say $\lambda_n > \lambda'_n > \lambda''_n > \ldots > \lambda_{n+1}$; but $\lambda_n - \lambda'_n$, $\lambda''_n - \lambda'_n$, etc. are all of the order $\lambda_n - \lambda_{n+1}$, so I will just forget about the intermediate indices and use the fact that these differences are all of order $O(1/n^2)$. Also I shall assume δ_k decreases very fast to 0: in fact this can only be checked a posteriori once all estimates have been performed.

With these conventions in mind, let us sketch the steps of the iteration.

(a) In this step one shows that Ω^n is asymptotically close to Id. From $\|\rho^n\|_{\mathcal{F}^{\lambda_n \tau}} = O(\delta_n)$ we deduce, as in Chapter 6, a bound on $\Omega^n_{t,\tau} - \mathrm{Id}$. If we do it in gliding regularity about λ_{n+1}, then to apply Proposition 6.1, we need $\lambda_n - \lambda_{n+1}$ to be at least of order $\delta^{1/3}$, because the force field is of size δ (it depends not only on h^n but also on h^1, h^2, ... Of course we cannot afford to lose such a fixed amount of regularity as $n \to \infty$. So we modify the estimates in Proposition 6.1 tby letting the velocity regularity depend on τ and t: replace $\mathcal{Z}^{\lambda_n}_{\tau}$ by $\mathcal{Z}^{\lambda_n(1+b)}_{\tau - bt/(1+b)}$, where $b(t) = B/(1+t)$. This works because

- $\|F^n(t)\|$ decays faster than $\lambda_n\, b(t)$,
- $\lambda_n(1+b)\left(\tau - \frac{bt}{1+b}\right) < \lambda_n\, \tau$, at least for t positive enough. So

we can still estimate the $\mathcal{Z}^{\lambda_n(1+b)}_{\tau - bt/(1+b)}$ norm of the force field by its $\mathcal{F}^{\lambda_n \tau}$ norm (recall (5.14)).

This does not work for small values of t, but short times have already been treated separately, as explained in Section 2. In all the rest of the argument one should consider separately small and not so small times.

Then, by amplification of the fixed point technique of Chapter 6, we arrive at

$$\|\Omega^n_{t,\tau} - \mathrm{Id}\,\|_{\mathcal{Z}^{\lambda_n(1+b)}_{\tau - \frac{bt}{1+b}}} = \frac{O(\delta)}{\tau^s}. \tag{8.14}$$

Next, with another fixed point argument, one establishes

$$\|\nabla\Omega^n_{t,\tau} - I\|_{\mathcal{Z}^{\lambda_n(1+b)}_{\tau - \frac{bt}{1+b}}} = O\left(\frac{\delta}{\tau^s}\right). \tag{8.15}$$

(b) In this step one shows that Ω^k is close to Ω^n, uniformly in $n \geq k$. This is done again by fixed point, and one arrives at something like

$$\|\Omega^n_{t,\tau} - \Omega^k_{t,\tau}\|_{\mathcal{Z}^{\lambda_n(1+b)}_{\tau-\frac{bt}{1+b}}} = O\left(\frac{\delta_k}{\tau^s}\right). \tag{8.16}$$

(The regularity is the worst of λ_n and λ_k, that is λ_n; and the size is the worst of δ_n and δ_k, that is δ_k.) The important point is that this estimate goes to 0 as $k \to \infty$, uniformly in n. From (8.16) is deduced (by fixed point again...)

$$\left\|(\Omega^k_{t,\tau})^{-1} \circ \Omega^n_{t,\tau} - \mathrm{Id}\right\|_{\mathcal{Z}^{\lambda_n(1+b)}_{\tau-\frac{bt}{1+b}}} = O\left(\frac{\delta_k}{\tau^s}\right). \tag{8.17}$$

(Inversion with a norm of time index τ is possible only if $\Omega^n - \Omega^k$ is much smaller than $1/\tau$; but this is guaranteed by (8.16).)

In the sequel I shall not always write the indices of the norms.

(c) The next step is to update the controls on the previous layers h^k by taking into account the change of the characteristics; so one should estimate $h^k \circ \Omega^n$ for all k. This is done by composition:

$$h^k \circ \Omega^n = (h^k \circ \Omega^{k-1}) \circ \big((\Omega^{k-1})^{-1} \circ \Omega^n\big),$$

so

$$\begin{aligned}\left\|h^k \circ \Omega^n - h^k \circ \Omega^k\right\| &\leq \|\nabla(h^k \circ \Omega^{k-1})\| \left\|\big(\Omega^{k-1}_{t,\tau})^{-1} \circ \Omega^n_{t,\tau}\big) - I\right\| \\ &\leq C\, \delta_k \left(\frac{\delta_k}{\tau^s}\right),\end{aligned}$$

as a consequence of the induction assumption $\nabla(h^k \circ \Omega^{k-1}) = O(\delta_k)$ and (8.17). It follows

$$\|h^k_\tau \circ \Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda_n(1+b)}_{\tau-\frac{bt}{1+b}}} = O(\delta_k). \tag{8.18}$$

Similar bounds are established for $(\nabla_x h^k)\circ\Omega^n$ and $((\nabla_v + \tau\nabla_x)h^k)\circ\Omega^n$, with just a small loss on the regularity index. Consequently $(\nabla_v h^k)\circ\Omega^n$ is $O(\tau\,\delta_k)$. Similarly, $(\nabla^2 h^k)\circ\Omega^n = O(\tau^2\,\delta_k)$.

(d) Next is the key step where an estimate is obtained for ρ^{n+1}. First write the equation for h^{n+1}, then the method of characteristics in the force field $F[f^n]$ yields

$$\begin{aligned}\rho^{n+1}(t,x) = -\int_0^t \int \big(F[h^{n+1}] \cdot \nabla_v f^n\big)\big(\tau, S^n_{t,\tau}(x,v)\big)\, dv\, d\tau \\ + \text{quadratic contribution from } h^n.\end{aligned}$$

To take advantage of the mixing effect of the free transport semigroup, introduce the scattering Ω^n by force, rewriting the estimate above as

(8.19)
$$\rho^{n+1}(t,x) = -\int_0^t \int \Big[\big(F[h^{n+1}]\cdot\nabla_v f^n\big)\circ\Omega^n_{t,\tau}\Big]\big(x-v(t-\tau),v\big)\,dv\,d\tau + O(n^r\,\delta_n^2),$$

where the contribution from h^n has been estimated in a crude way.

If it was not for the composition by Ω^n, we would be in the same situation as in Chapter 7, and we could use the estimates on the Vlasov equation seen as a reaction equation. But $(F[h^{n+1}]\cdot\nabla_v f^n)\circ\Omega^n$ *does not* have the structure $G(t,x)\cdot\nabla_v g(t,x,v)$ which was crucial in Chapter 7. The problem is to show that composition by Ω^n does not change much in the long run.

So we decompose that reaction term as follows:

(8.20)
$$\begin{aligned}\big(F[h^{n+1}]\cdot\nabla_v f^n\big)\circ\Omega^n_{t,\tau} =& F[h^{n+1}]\cdot\nabla_v\Big(f^0+\sum_{k\le n} h^k\circ\Omega^{k-1}\Big)\\ &+\Big(F[h^{n+1}]\circ\Omega^n - F[h^{n+1}]\Big)\cdot(\nabla_v f^n\circ\Omega^n)\\ &+F[h^{n+1}]\cdot\sum_{k\le n}\Big[(\nabla_v h^k)\circ\Omega^{k-1}-\nabla_v\big(h^k\circ\Omega^{k-1}\big)\Big]\\ &+F[h^{n+1}]\cdot\sum_{k\le n}\Big[(\nabla_v h^k)\circ\Omega^{n}-(\nabla_v h^k)\circ\Omega^{k-1}\Big].\end{aligned}$$

The first term on the right-hand side is fine, and the other three terms will be treated as perturbations in large time.

- The second term in the right-hand side of (8.20) is estimated as follows:

$$\begin{aligned}\Big\|F[h^{n+1}]\circ\Omega^n - F[h^{n+1}]\Big\|\,\|\nabla_v f^n\circ\Omega^n\| &\le \big\|\nabla F[h^{n+1}]\big\|\,\|\Omega^n-\mathrm{Id}\,\|\,\|\nabla_v f^n\circ\Omega^n\|\\ &\le C\,\|\rho^{n+1}\|\,\frac{\delta\tau}{\tau^s},\end{aligned}$$

where I used $\|\nabla F[h^{n+1}]\|\le C\,\|\rho^{n+1}\|$: indeed, thanks to the assumption $|\widehat{W}(k)| = O(1/|k|^2)$, passing from the density to the force should gain at least one derivative. (Note carefully: Here we cannot afford to lose regularity on ρ^{n+1} because we are trying to get a Gronwall-type estimate on the unknown ρ^{n+1}, so it is crucial to use the very same norm on the left-hand side and the right-hand side, and the only thing we can use to regain the derivative is the smoothing induced by the convolution inside the force.)

- Next, by recursion hypothesis,

$$\left\|\nabla_v(h^k \circ \ Om^{k-1}) - (\nabla_v h^k)\circ\Omega^{k-1}\right\| = O\left(\frac{\delta_k}{\tau^s}\right),$$

which allows to control the third term in the right-hand side of (8.20) by $\|\rho^{n+1}\|\,\delta_k/\tau^s$.

- Finally, to handle the last term in the right-hand side of (8.20), one writes

$$\left\|(\nabla_v h^k)\circ\Omega^{k-1} - (\nabla_v h^k)\circ\Omega^n\right\| \\ \leq \sup_{0\leq\theta\leq 1}\left\|\nabla^2 h^k \circ \big((1-\theta)\,\Omega^{k-1} + \theta\,\Omega^n\big)\right\| \,\|\Omega^{k-1}-\Omega^n\|.$$

From (8.17) the argument of $\nabla^2 h^k$ is close to Ω^{k-1}, uniformly in θ, so up to a slight loss we end up with a bound like

$$\left\|\nabla^2 h^k\circ\Omega^{k-1}\right\|\,\|\Omega^{k-1}-\Omega^n\| \leq C\,\delta_k\,\tau^2\,\left(\frac{\delta_k}{\tau^s}\right),$$

after use of the induction hypothesis on $\nabla^2 h^k\circ\Omega^{k-1}$ and the bound (8.16).

Plugging all these controls in (8.20) shows

$$\rho^{n+1}(t,x) = -\int_0^t F[h^{n+1}_\tau]\cdot\nabla_v\Big(f^0 + \sum_{k\leq n} h^k_\tau\circ\Omega^{k-1}_{t,\tau}\Big)(x - v(t-\tau), v)\,d\tau\,dv \\ + O\left(\int_0^t \frac{\|\rho^{n+1}(\tau)\|}{1+\tau^s}\,d\tau\right) + O(n^r\delta_n^2).$$

Then one can operate as in Chapter 7 and get a Gronwall estimate on $\|\rho^{n+1}(\tau)\|_{\mathcal{F}^{\lambda_{n+1}\tau}}$. The difference is an additional term in the kernel $K(t,\tau)$, which is $O(\delta\,\tau^{-s})$, uniformly in t. But this is harmless: think indeed that a solution of

$$\varphi(t) \leq A + \delta\int_0^t \frac{\varphi(\tau)}{1+\tau^2}\,d\tau$$

satisfies $\varphi(t)\leq A + C\,\delta$. The robustness of the moment estimates from Chapter 7 and the L^2 method in Lemma 3.5 makes it possible to adapt all these estimates to the present complicated situation, yielding in the end

$$\|\rho^{n+1}\|_{\mathcal{F}^{\lambda_{n+1}\tau}} = O\Big(e^{n^K}\,n^r\,\delta_n^2\Big).$$

So this step gives

$$\delta_{n+1} = O\Big(e^{n^K}\, n^r\, \delta_n^2\Big).$$

This is not as good as $\delta_{n+1} = O(\delta_n^2)$, and does not imply $\delta_n = O(C^n\, \delta^{2^n})$ as in the classical Newton scheme; but this is still compatible with $\delta_n = O(\delta^{a^n})$, $a < 2$.

The final conclusion of this step is

$$\|\rho^{n+1}\|_{\mathcal{F}^{\lambda_{n+1}\tau}} = O(\delta_{n+1}). \tag{8.21}$$

(e) From the estimate on ρ^{n+1} we immediately deduce an estimate on the force: $\|F[h^{n+1}]\|_{\mathcal{F}^{\lambda_{n+1}\tau}} = O(\delta_{n+1})$.

(f) Then use the equation for h^{n+1} once more, but now compose it with $\Omega^n_{t,\tau}$ where τ is given, and estimate $h^{n+1}_\tau \circ \Omega^n_{t,\tau}$. This is not so difficult as Step (d) because now there is no need to use the same norm on both sides: we already have an estimate on the force, we don't need any Gronwall-type inequality, we can afford to lose a little bit on the regularity of h^{n+1} compared with the regularity of ρ^{n+1}. After some computations, one gets something like

$$\Big\|h^{n+1}_\tau \circ \Omega^n_{t,\tau}\Big\|_{\mathcal{Z}^{\lambda_n(1+b)}_{\tau-\frac{bt}{1+b}}} = O(\delta_{n+1}). \tag{8.22}$$

(g) Deduce (by gliding regularity) that

$$\nabla(h^{n+1}_\tau \circ \Omega^n_{t,\tau}) = O(\delta_{n+1}\,\tau), \qquad \nabla^2(h^{n+1}_\tau \circ \Omega^n_{t,\tau}) = O(\delta_{n+1}\,\tau^2). \tag{8.23}$$

(h) Deduce that

$$(\nabla h^{n+1}) \circ \Omega^n = (\nabla\Omega^n)^{-1}\,\nabla(h^{n+1} \circ \Omega^n) = O(\delta_{n+1}\,\tau),$$

and similarly $\nabla^2 h^{n+1} \circ \Omega^n = O(\delta_{n+1}\,\tau^2)$. Finally, note that

$$\begin{aligned}\Big\|\nabla(h^{n+1} \circ \Omega^n) - (\nabla h^{n+1}) \circ \Omega^n\Big\| &\le \big\|\nabla(\Omega^n - \mathrm{Id}\,)\big\|\,\big\|\nabla h^{n+1} \circ \Omega^n\big\| \\ &\le C\,\left(\frac{\delta}{\tau^s}\right)\,(\delta_{n+1}\,\tau),\end{aligned}$$

so at the same time this is small like $O(\delta_{n+1})$, and it decays fast in τ.

Once this is done, all the estimates have been propagated from stage n to stage $n+1$, and we can go on! $\square$

4. Main result

Once we have obtained the estimates on all h^n, it is easy to conclude the proof of Theorem 4.1. Let us sketch the argument. Summing the estimates on all h^n, one obtains the uniform bound

$$\sup_{t\geq 0} \|f(t,\,\cdot\,) - f^0\|_{\mathcal{Z}_t^{\lambda,\mu;1}} + \sup_{t\geq 0} \|\rho - \rho^0\|_{\mathcal{F}^{\lambda t+\mu}} = O(\delta). \tag{8.24}$$

This bound is the true main result: actually, it contains much more information than Theorem 4.1. It implies immediately that the force $F(t,\,\cdot\,)$ satisfies $\|F(t)\|_{\mathcal{F}^{\lambda t+\mu}} = O(\delta)$, and since F is a gradient, it immediately follows that $F(t)$ decays exponentially fast with t. On the other hand, $\nabla_v f$ grows at most linearly in t, so $F\cdot\nabla_v f$ decays exponentially fast in gliding regularity. This implies that $(d/dt)f(t, x+vt, v)$ also decays exponentially; in particular, $f(t, x+vt, v)$ has a large-time limit $g(x, v)$, which is analytic, and the convergence actually holds in an appropriate $\mathcal{Z}$ function space. As a consequence, $f(t, x, v)$ has the same asymptotic behavior as $g(x-vt, v)$, which converges weakly to $\langle g\rangle(v)$. The conclusion of Theorem 4.1 follows easily.

I shall conclude with a few indications on non-analytic data (Theorem 4.3). It was noted in Chapter 7 that the expected loss of regularity is like a fractional exponential, say $e^{|\xi|^\alpha}$. Then it is expected that all results hold true in a regularity which is better, that is, Gevrey-ν with $\nu > 1/\alpha$.

All the estimates can indeed be adapted to this setting, either by changing all our norms to handle Gevrey regularity, or by decomposing a Gevrey function in a sum of analytic contributions with analyticity width going to 0 in a controlled way. In practice, we decompose the initial datum $f_i - f^0$ in a sum of data h_i^n, such that h_i^n satisfies some analyticity condition in a strip of width λ'_n, and the norm of h_i^n decays in a controlled way as $n\to\infty$. Then we use h_i^n as an initial datum in the step n of the Newton scheme. Of course our long-time estimates on h^n only hold in regularity less than λ'_n, but the way λ'_n goes to 0 is controlled, so in the end we can reconstruct Gevrey regularity for $\sum h^n$, losing just a bit on the Gevrey exponent in the process. Then the iteration can be performed as in the analytic case.

Bibliographical notes

Newton presented his approximation scheme in a 1669 treatise [**78**] which was published only decades later. The presentation and the study of the scheme were revised and improved by a series of English mathematicians: Wallis, Raphson, Simpson, Cayley. An ancestor of the

Newton scheme is the so-called *Babylonian method* for the numerical solution of square roots, which some experts conjecture to have been known to Babylonian mathematicians as early as 1900 BC, and to Indian mathematicians before 800 BC: to compute e.g. $\sqrt{2}$ apply the Newton scheme to the function $\Phi(x) = x^2 - 2$.

Kolmogorov's perturbation theorem for Hamiltonian systems was announced in [**53**] in analytic regularity, and Nash's embedding theorem appeared in [**76**]. Kolmogorov's sketchy proof did not convince everybody at the time, which was very fortunate since it motivated Moser to devise his own proof [**72, 73**] in a differentiable setting, using Nash's work as an inspiration. Around the same time, Kolmogorov's analytic result was, after all, validated by Arnold [**5**] with an alternative proof. Much later, Chierchia [**26**] reconstructed the details of was is likely to have been Kolmogorov's original argument. Chierchia wrote a survey of KAM theory for the online encyclopedia Scholarpedia [**27**].

A fixed point approach to the KAM theory was proposed by Herman [**45**]; while it does not seem to apply in full generality, it does suffice to cover certain simple situations. A fixed point approach to Nash's embedding theorem was devised by Günther [**39**].

The Cauchy–Kowalevsakaya method is presented in a number of sources; Nirenberg's presentation [**79**] is based on a Newton scheme, and is close in spirit to the treatment sketched in these notes, whose details are provided in [**74**]. (I learnt about Nirenberg's work from Klainerman, Alinhac and Gérard after [**74**] was written.) Once again, after a few years, Nirenberg's use of a Newton scheme has been replaced by a fixed-point theorem [**80**], and maybe this will also happen some day for our theory of Landau damping.

Short-time analyticity estimates on the solutions of the Vlasov–Poisson equation go back to Benachour [**12**], with an alternative method.

A few remarks about Lemma 8.3 can be made. Differentiation of the norm with respect to time-dependent integrability index is classical in the field of hypercontractivity [**38**]. Differentiation with respect to a time-dependent regularity index is not so common, but appears in the work of Chemin [**25**] on the short-time regularity of the incompressible Navier–Stokes system.

The long-time analysis of the Newton scheme is performed in painful detail in [**74**]. The adaptation to Gevrey data is sketched in the same source. As I learnt later, Moser already used the idea to decompose a smooth, nonanalytic function h into a sum of analytic functions h^n whose norm and analyticity width decay in a controlled way as $n \to \infty$.

CHAPTER 9

Conclusions

The main result in this course is that *Landau damping survives nonlinearity*, and the long-time behavior of the linearized Vlasov equation is, after all, a good approximation of the long-time behavior of the nonlinear Vlasov equation. This ends up a controversy and provides a final answer to the objection formulated by Backus half a century ago. In the end Landau was right, although the proof involves many ingredients which were inaccessible at his time.

Remarkably, the range of interactions which are admissible in the main result includes the Poisson coupling (repulsive or attractive) as a limit case.

Moreover, the theory provides an interpretation of Landau damping: this is a relaxation by **mixing**, **confinement** and **smoothness**. The mixing transport equation converts smoothness into decay, in the spirit of Fourier transform (Riemann–Lebesgue lemma). Regularity goes away from the v variable to the x variable, so the force becomes very smooth, and because it has a gradient structure this implies time decay.

Even though the solution of the linearized problem involves a loss of gliding regularity (which implies relaxation), regularity estimates survive the nonlinear perturbation by a mathematical (rather than physical) phenomenon comparable to the KAM theory, which takes advantage of the complete integrability of the original system (in our case the linearized Vlasov equation) and a Newton scheme to overcome the loss of regularity.

In this sense the proof provides an unexpected bridge between three of the most famous paradoxical statements from classical mechanics in the twentieth century: Landau damping, KAM theory, and the echo experiment. This is all the more remarkable that this bridge only appears in the treatment of the nonlinear Vlasov equation, while Landau was dealing specifically with the linearized equation.

The fully constructive property of the Newton scheme allows to *construct* the asymptotic state, opening the door to asymptotic studies. For instance, one can construct in this way heteroclinic trajectories of the nonlinear Vlasov equation (solutions are automatically homoclinic at order 2 in the perturbation size ε; but heteroclinic corrections of order $O(\varepsilon^3)$ can appear). This shows that the asymptotic behavior cannot be predicted on the basis on invariants of motion alone: indeed, these invariants are all preserved by the reversal of velocities, which amounts to a change of the direction of time.

Another striking feature of our proof is that, compared to KAM theory, the loss of regularity is much more severe in the present case: infinitely many derivatives are lost, corresponding to a fractional exponential in Fourier space. Such high losses prevent the application of the classical Nash–Moser regularization scheme; for this reason in particular, we have not been able to establish nonlinear Landau damping below Gevrey regularity.

The issue of nonlinear Landau damping for less smooth data appears wide open. In a recent contribution, Lin proved that one cannot hope for Landau damping in low regularity, that is, with less than 2 derivatives in an appropriate Sobolev space. Indeed, in such a low regularity topology, BGK waves are *dense* around stable homogeneous equilibrium profiles; so damping to a homogeneous state might still be true for typical solutions, but cannot hold over a whole neighborhood of the equilibrium.

The benchmarking of reliable long-time numerical schemes, the study of the linear and nonlinear stability of BGK wave, the qualitative study of large perturbations of equilibrium, the statistical theory of the Vlasov–Poisson equation, remain wide open fascinating subjects. Another development of interest would be the adaptation of Landau damping theory to other models sharing some similar features; for the most natural candidate, the two-dimensional incompressible Euler equation, this turns out to be much more difficult than could be expected.

Bibliographical notes

KAM type problems with a loss of infinitely many derivatives (multiplication by a fractional exponential in Fourier space) have been considered by Popov [**84**]; in this case (as I learnt from Chierchia and Pöschel) nobody knows how to treat C^r regularity in the style of Moser [**72**].

Heteroclinic solutions of the nonlinear Vlasov equation are constructed in [**74**, Section 14].

Lin's negative results are presented in [**59**]. The precise statement is that there is density in $W^{1+1/p,p+0}$ topology for any $p \in (1,\infty)$.

A preliminary discussion of nonlinear Landau damping for two-dimensional incompressible Euler equation was performed by Bouchet and Morita [**16**]. Together with Mouhot, we tried to put this on rigorous footing by adapting the study of the nonlinear Vlasov equation, but stumbled upon formidable difficulties, even in the simple case of a perturbation of a linear shear flow [**75**].

Bibliography

[1] Akhiezer, A., Akhiezer, I., Polovin, R., Sitenko, A., and Stepanov, K. *Plasma electrodynamics. Vol. I: Linear theory, Vol. II: Non-linear theory and fluctuations.* Pergamon Press, 1975 (English Edition). Translated by D. ter Haar.

[2] Alexandre, R., and Villani, C. On the Landau approximation in plasma physics. *Ann. Inst. H. Poincaré Anal. Non Linéaire 21*, 1 (2004), 61–95.

[3] Alinhac, S., and Gérard, P. *Pseudo-differential operators and the Nash-Moser theorem*, vol. 82 of *Graduate Studies in Mathematics.* Amer. Math. Soc., Providence, RI, 2007. Translated from the 1991 French original by Stephen S. Wilson.

[4] Ambrosio, L., and Gangbo, W. Hamiltonian ODE's in the Wasserstein space of probability measures. *Comm. Pure Appl. Math. 51* (2007), 18–53.

[5] Arnold, V.I. Proof of a Theorem by A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian. *Russian Math. Survey 18* (1963), 13–40.

[6] Arsen'ev, A.A. Existence in the large of a weak solution of Vlasov's system of equations. *Ž. Vyčisl. Mat. i Mat. Fiz. 15* (1975), 136–147, 276.

[7] Arsen'ev, A.A. Existence and uniqueness of the classical solution of Vlasov's system of equations. *Ž. Vyčisl. Mat. i Mat. Fiz. 15* (1975), 1344–1349, 1363.

[8] Backus, G. Linearized plasma oscillations in arbitrary electron distributions. *J. Math. Phys. 1* (1960), 178–191, 559.

[9] Bardos, C. and Degond, P. Global existence for the Vlasov–Poisson equation in 3 space variables with small initial data. *Ann. Inst. H. Poincar Anal. Non Linaire 2*, 2 (1985), 101–118.

[10] Batt, J., and Rein, G. Global classical solutions of the periodic Vlasov-Poisson system in three dimensions. *C. R. Acad. Sci. Paris Sér. I Math. 313*, 6 (1991), 411–416.

[11] Belmont, G., Mottez, F., Chust, T., and Hess, S. Existence of non-Landau solutions for Langmuir waves. *Phys. of Plasmas 15* (2008), 052310, 1–14.

[12] Benachour, S. Analyticité des solutions des équations de Vlassov-Poisson. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 16*, 1 (1989), 83–104.

[13] Bernstein, I.B., Greene, J.M. and Kruskal, M.D. Exact nonlinear plasma oscillations. *Phys. Rev. 108*, 3 (1957), 546–550.

[14] Binney, J., and Tremaine, S. *Galactic Dynamics*, second ed. Princeton Series in Astrophysics. Princeton University Press, 2008.

[15] Bolley, F., Guillin, A., and Villani, C. Quantitative concentration inequalities for empirical measures on non-compact spaces. *Prob. Theory Related Fields 137*, 3–4 (2007), 541–593.

[16] Bouchet, F., and Morita, H. Large time behavior and asymptotic stability of the two-dimensional Euler and linearized Euler equations. Preprint Hal-00382670.

[17] Bouchut, F. Introduction à la théorie mathématique des équations cinétiques. In *Session "L'Etat de la Recherche" de la SMF* (1998). in "Kinetic equations and asymptotic theory", F. Bouchut, F. Golse, M. Pulvirenti, coll. "Series in Appl. Math.", Gauthier-Villars, 2000.

[18] Bourgain, J. Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations. *Geom. Funct. Anal. 3*, 2 (1993), 107–156.

[19] Bourgain, J. Periodic nonlinear Schrödinger equation and invariant measures. *Comm. Math. Phys. 166*, 1 (1994), 1–26.

[20] Braun, W. and Hepp, K. The Vlasov dynamics and its fluctuations in the $1/N$ limit of interacting classical particles. *Commun. Math. Phys. 56* (1977), 125–146.

[21] Caglioti, E., and Maffei, C. Time asymptotics for solutions of Vlasov–Poisson equation in a circle. *J. Statist. Phys. 92*, 1-2 (1998), 301–323.

[22] Case, K. Plasma oscillations. *Ann. Phys. 7* (1959), 349–364.

[23] Cercignani, C., Illner, R., and Pulvirenti, M. *The mathematical theory of rarefied gases.* Applied Mathematical Sciences, 106. Springer-Verlag, New York, 1994.

[24] Chavanis, P.H., Sommeria, J., and Robert, R. Statistical mechanics of two-dimensional vortices and collisionless stellar systems. *Astrophys. J. 471* (1996), 385–399.

[25] Chemin, J.-Y. Le système de Navier–Stokes incompressible soixante dix ans après Jean Leray. In *Actes des Journées Mathématiques à la Mémoire de Jean Leray*, vol. 9 of *Sémin. Congr.* Soc. Math. France, Paris, 2004, pp. 99–123.

[26] Chierchia, L. A. N. Kolmogorov's 1954 paper on nearly-integrable Hamiltonian systems. A comment on: "On conservation of conditionally periodic motions for a small change in Hamilton's function" [Dokl. Akad. Nauk SSSR (N.S.) **98** (1954), 527–530]. *Regul. Chaotic Dyn. 13*, 2 (2008), 130–139.

[27] Chierchia, L. *Kolmogorov–Arnold–Moser theory.* Online review available at `www.scholarpedia.org/article/KAM_theory_in_celestial_mechanics`

[28] Degond, P. Global existence of smooth solutions for the Vlasov–Fokker–Planck equation in 1 and 2 space dimensions. *Ann. Sci. École Norm. Sup. (4) 19*, 4 (1986), 519–542.

[29] Degond, P. Spectral theory of the linearized Vlasov–Poisson equation. *Trans. Amer. Math. Soc. 294*, 2 (1986), 435–453.

[30] DiPerna, R., and Lions, P.-L. Global weak solutions of Vlasov-Maxwell systems. *Comm. Pure Appl. Math. 42* (1989), 729–757.

[31] Dobrušin, R.L. Vlasov equations. *Funktsional. Anal. i Prilozhen. 13*, 2 (1979), 48–58, 96.

[32] Elskens, Y. Irreversible behaviours in Vlasov equation and many-body Hamiltonian dynamics: Landau damping, chaos and granularity in the kinetic limit. In *Topics in Kinetic Theory* (2005), T. Passot, C. Sulem, and P. L. Sulem, Eds., vol. 46 of *Fields Institute Communications*, Amer. Math. Soc., Providence, pp. 89–108.

[33] Elskens, Y., and Escande, D.F. *Microscopic dynamics of plasmas and chaos*. Institute of Physics, Bristol, 2003.

[34] Escande, D.F. *Wave–particle interaction in plasmas: a qualitative approach*, vol. 90 of *Lecture Notes of the Les Houches Summer School*. Oxford Univ. Press, 2009. Edited by Th. Dauxois, S. Ruffo and L. F. Cugliandolo.

[35] Glassey, R.T. *The Cauchy problem in kinetic theory*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996.

[36] Glassey, R., and Schaeffer, J. Time decay for solutions to the linearized Vlasov equation. *Transport Theory Statist. Phys. 23*, 4 (1994), 411–453.

[37] Glassey, R., and Schaeffer, J. On time decay rates in Landau damping. *Comm. Partial Differential Equations 20*, 3-4 (1995), 647–676.

[38] Gross, L. Logarithmic Sobolev inequalities. *Amer. J. Math. 97* (1975), 1061–1083.

[39] Günther, M. On the perturbation problem associated to isometric embeddings of Riemannian manifolds. *Ann. Global Anal. Geom. 7*, 1 (1989), 69–77.

[40] Guo, Y. Global weak solutions of the Vlasov–Maxwell system with boundary conditions. *Commun. Math. Phys. 154* (1993), 245–263.

[41] Guo, Y., and Strauss, W. A. Nonlinear instability of double-humped equilibria. *Ann. Inst. H. Poincaré Anal. Non Linéaire 12*, 3 (1995), 339–352.

[42] Hauray, M., and Jabin, P.-E. N-particles approximation of the Vlasov equations with singular potential. *Arch. Ration. Mech. Anal. 183*, 3 (2007), 489–524.

[43] Hayes, J. N. On non-Landau damped solutions to the linearized Vlasov equation. *Nuovo Cimento (10) 30* (1963), 1048–1063.

[44] Heath, R., Gamba, I., Morrison, P., and Michler, C. A discontinuous Galerkin method for the Vlasov–Poisson system. Work in progress, 2009.

[45] Herman, M. *Sur les courbes invariantes par les difféomorphismes de l'anneau. Vol. 1.* Astérisque, 103-104. Société Mathématique de France, Paris, 1983.

[46] Horst, E. On the classical solutions of the initial value problem for the unmodified nonlinear Vlasov equation. I. General theory. *Math. Methods Appl. Sci. 3*, 2 (1981), 229–248.

[47] Horst, E. On the classical solutions of the initial value problem for the unmodified nonlinear Vlasov equation. II. Special cases. *Math. Methods Appl. Sci. 4*, 1 (1982), 19–32.

[48] Horst, E. On the asymptotic growth of the solutions of the Vlasov–Poisson system. *Math. Methods Appl. Sci. 16*, 2 (1993), 75–86.

[49] Hwang, J.-H., and Velázquez, J. Global existence for the Vlasov-Poisson system in bounded domains. *Arch. Ration. Mech. Anal. 195*, 3 (2010), 763–796.

[50] Hwang, J.-H., and Velázquez, J. On the existence of exponentially decreasing solutions of the nonlinear Landau damping problem. *Indiana Univ. Math. J. 58*, 6 (2009), 2623–2660.

[51] Isichenko, M. Nonlinear Landau damping in collisionless plasma and inviscid fluid. *Phys. Rev. Lett. 78*, 12 (1997), 2369–2372.

[52] Kiessling, M. K.-H. The "Jeans swindle": a true story—mathematically speaking. *Adv. in Appl. Math. 31*, 1 (2003), 132–149.

[53] KOLMOGOROV, A.N. On conservation of conditionally periodic motions for a small change in Hamilton's function. *Dokl. Akad. Nauk. SSSR (N.S.) 98* (1954), 527–530.

[54] KRALL, N., AND TRIVELPIECE, A. *Principles of plasma physics*. San Francisco Press, 1986.

[55] 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.

[56] LANDAU, L.D. On the vibration of the electronic plasma. *J. Phys. USSR 10* (1946), 25. English translation in *JETP 16*, 574. Reproduced in *Collected papers of L.D. Landau*, edited and with an introduction by D. ter Haar, Pergamon Press, 1965, pp. 445–460; and in *Men of Physics: L.D. Landau*, Vol. 2, Pergamon Press, D. ter Haar, ed. (1965). [1]

[57] LIN, Z. Instability of periodic BGK waves. *Math. Res. Lett. 8* (2001), 521–534.

[58] LIN, Z. Nonlinear instability of periodic BGK waves for Vlasov–Poisson system. *Comm. Pure Appl. Math. 58* (2005), 505–528.

[59] LIN, Z., AND ZENG, C. BGK waves and nonlinear Landau damping. Work in progress, 2009.

[60] LIONS, P.-L., AND PERTHAME, B. Propagation of moments and regularity for the 3-dimensional Vlasov–Poisson system. *Invent. Math. 105*, 2 (1991), 415–430.

[61] LIFSHITZ, E. M., AND 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.

[62] LOTT, J. Some geometric calculations on Wasserstein space. *Comm. Math. Phys. 277*, 2 (2008), 423–437.

[63] LYNDEN-BELL, D. The stability and vibrations of a gas of stars. *Mon. Not. R. astr. Soc. 124*, 4 (1962), 279–296.

[64] LYNDEN-BELL, D. Statistical mechanics of violent relaxation in stellar systems. *Mon. Not. R. astr. Soc. 136* (1967), 101–121.

[65] MALMBERG, J., AND WHARTON, C. Collisionless damping of electrostatic plasma waves. *Phys. Rev. Lett. 13*, 6 (1964), 184–186.

[66] MALMBERG, J., WHARTON, C., GOULD, R., AND O'NEIL, T. Plasma wave echo experiment. *Phys. Rev. Letters 20*, 3 (1968), 95–97.

[67] MANFREDI, G. Long-time behavior of nonlinear Landau damping. *Phys. Rev. Lett. 79*, 15 (1997), 2815–2818.

[68] MARCHIORO, C. AND PULVIRENTI, M. *Mathematical theory of incompressible fluids*. Springer-Verlag, New York, 1994.

[69] MASLOV, V.P., AND FEDORYUK, M.V. The linear theory of Landau damping. *Mat. Sb. (N.S.) 127(169)*, 4 (1985), 445–475, 559.

[70] MILLER, J. Statistical mechanics of Euler equations in two dimensions. *Phys. Rev. Lett. 65*, 17 (1990), 2137–2140.

[1]There is a misprint in formula (17) of this reference (p. 104): replace $e^{-(ka)^2/2}$ by $e^{-1/(2(ka)^2)}$.

[71] Morrison, P.J. Hamiltonian description of Vlasov dynamics: Action-angle variables for the continuous spectrum. *Transp. Theory Statist. Phys. 29*, 3–5 (2000), 397–414.

[72] Moser, J. A rapidly convergent iteration method and non-linear differential equations. II. *Ann. Scuola Norm. Sup. Pisa (3) 20* (1966), 499–535.

[73] Moser, J. Recollections. In *The Arnoldfest (Toronto, ON, 1997)*, vol. 24 of *Fields Inst. Commun.* Amer. Math. Soc., Providence, RI, 1999, pp. 19–21. Concerning the early development of KAM theory.

[74] Mouhot, C. and Villani, C. On Landau damping. Preprint, 2009. Available online at `arXiv:0904.2760`

[75] Mouhot, C. and Villani, C. Landau damping of shear flow. Working manuscript, 2010.

[76] Nash, J. The imbedding problem for Riemannian manifolds. *Ann. of Math. (2) 63* (1956), 20–63.

[77] 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.

[78] Newton, I. *De analysi per aequationes numero terminorum infinitas.* Manuscript, 1669; published in 1711 by W. Jones.

[79] Nirenberg, L. An abstract form of the nonlinear Cauchy–Kowalewski theorem. *J. Differential Geometry 6* (1972), 561–576.

[80] Nishida, T. A note on a theorem of Nirenberg. *J. Differential Geometry 12* (1977), 629–633.

[81] O'Neil, T. Collisionless damping of nonlinear plasma oscillations. *Phys. Fluids 8*, 12 (1965), 2255–2262.

[82] Penrose, O. Electrostatic instability of a non-Maxwellian plasma. *Phys. Fluids 3* (1960), 258–265.

[83] Pfaffelmoser, K. Global classical solutions of the Vlasov-Poisson system in three dimensions for general initial data. *J. Differential Equations 95*, 2 (1992), 281–303.

[84] Popov, G. KAM theorem and quasimodes for Gevrey Hamiltonians. *Erg. Th. Dyn. Sys. 24*, 5 (2004), 1753–1786.

[85] Pulvirenti, M. On invariant measures for the 2-D Euler flow. In *Mathematical aspects of vortex dynamics* (Leesburg, VA, 1988). SIAM, Philadelphia, PA, 1989, pp. 88–96.

[86] Rein, G. Non-linear stability for the Vlasov–Poisson system — the energy-Casimir method. *Math. Meth. Appl. Sci. 17*, 14 (1994), 1129–1140

[87] 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.

[88] Ryutov, D. D. Landau damping: half a century with the great discovery. *Plasma Phys. Control. Fusion 41* (1999), A1–A12.

[89] Sáenz, A. W. Long-time behavior of the electric potential and stability in the linearized Vlasov theory. *J. Mathematical Phys. 6* (1965), 859–875.

[90] Schaeffer, J. Global existence of smooth solutions to the Vlasov-Poisson system in three dimensions. *Comm. Partial Differential Equations 16*, 8-9 (1991), 1313–1335.

[91] Spentzouris, L., Ostiguy, J., and Colestock, P. Direct measurement of diffusion rates in high energy synchrotrons using longitudinal beam echoes. *Phys. Rev. Lett. 76*, 4 (1996), 620–623.

[92] Spohn, H. *Large scale dynamics of interacting particles.* Texts and Monographs in Physics. Springer-Verlag, Berlin, 1991.

[93] Stahl, B., Kiessling, M.K.-H., and Schindler, K. Phase transitions in gravitating systems and the formation of condensed objects. *Planet. Space Sci. 43*, 3/4 (1995), 271–282.

[94] Sturm, K.-Th. Entropic measure on multidimensional spaces. Preprint (2009).

[95] Sznitman, A.-S. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989.* Springer, Berlin, 1991, pp. 165–251.

[96] Tremaine, S., Hénon, M., and Lynden-Bell, D. H-functions and mixing in violent relaxation. *Mon. Not. R. astr. Soc. 219* (1986), 285–297.

[97] Turkington, B. Statistical equilibrium measures and coherent states in two-dimensional turbulence. *Comm. Pure Appl. Math. 52*, 7 (1999), 781–809.

[98] van Kampen, N. On the theory of stationary waves in plasma. *Physica 21* (1955), 949–963.

[99] Villani, C. A review of mathematical topics in collisional kinetic theory. In *Handbook of mathematical fluid dynamics, Vol. I.* North-Holland, Amsterdam, 2002, pp. 71–305.

[100] Villani, C. *Topics in optimal transportation*, vol. 58 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2003.

[101] Villani, C. H-Theorem and beyond: Boltzmann's entropy in today's mathematics. Proceedings of the Boltzmann memorial conferences in Vienna and Munich (July & October 2006). In *Boltzmann's legacy, ESI Lect. Math. Phys., Eur. Math. Soc., Zürich*, 2008, pp.129-143.

[102] Villani, C. Limite de champ moyen. Incomplete lecture notes from a graduate course in École Normale Supérieure de Lyon, 2001-2002. Available online from `www.umpa.ens-lyon.fr`

[103] Vlasov, A.A. On the oscillation properties of an electron gas. *Zh. Èksper. Teoret. Fiz. 8* (1938), 291–318.

[104] Wiechen, H., Ziegler, H.J., and Schindler, K. Relaxation of collisionless self-gravitating matter – the lowest energy state. *Mon. Not. R. astr. Soc. 232* (1988), 623–646.

[105] Zhou, T., Guo, Y., and Shu, C.-W. Numerical study on Landau damping. *Physica D 157* (2001), 322–333.