Acta Math., 207 (2011), 29–201
DOI: 10.1007/s11511-011-0068-9
© 2011 by Institut Mittag-Leffler. All rights reserved

On Landau damping

by

Clément Mouhot
University of Cambridge
Cambridge, U.K.

Cédric Villani
Institut Henri Poincaré & Université de Lyon
Lyon, France

Dedicated to Vladimir Arnold and Carlo Cercignani.

Contents

Landau damping may be the single most famous mystery of classical plasma physics. For the past sixty years it has been treated in the linear setting at various degrees of rigor; but its non-linear version has remained elusive, since the only available results [13], [41] prove the existence of *some* damped solutions, without telling anything about their genericity.

In the present work we close this gap by treating the non-linear version of Landau damping in arbitrarily large times, under assumptions which cover both attractive and repulsive interactions, of any regularity down to Coulomb–Newton.

This will lead us to discover a distinctive mathematical theory of Landau damping, complete with its own functional spaces and functional inequalities. Let us make it clear that this study is not just for the sake of mathematical rigor: indeed, we shall get new insights into the physics of the problem, and identify new mathematical phenomena.

The plan of the paper is as follows.

In §1 we provide an introduction to Landau damping, including historical comments and a review of the existing literature. Then in §2 we state and comment on our main result about "non-linear Landau damping" (Theorem 2.6).

In §3 we provide a rather complete treatment of linear Landau damping, slightly improving on the existing results both in generality and simplicity. This section can be read independently of the rest.

In §4 we define the spaces of analytic functions which are used in the remainder of the paper. The careful choice of norms is one of the keys of our analysis; the complexity of the problem will naturally lead us to work with norms having up to five parameters. As a first application, we shall revisit linear Landau damping within this framework.

In §§5–7 we establish four types of new estimates (deflection estimates, short-term and long-term regularity extortion, echo control); these are the key sections containing in particular the physically relevant new material.

In §8 we adapt the Newton algorithm to the setting of the non-linear Vlasov equation. Then in §§9–11 we establish some iterative estimates along this scheme. (§11 is devoted specifically to a technical refinement allowing us to handle Coulomb–Newton interaction.)

From these estimates our main theorem is easily deduced in §12.

An extension to non-analytic perturbations is presented in §13.

Some counterexamples and asymptotic expansions are studied in §14.

Final comments about the scope and range of applicability of these results are provided in §15.

Even though it basically proves one main result, this paper is very long. This is due partly to the intrinsic complexity and richness of the problem, partly to the need to develop an adequate functional theory from scratch, and partly to the inclusion of remarks, explanations and comments intended to help the reader to understand the proof and the scope of the results. The whole process culminates in the extremely technical iteration performed in §10 and §11. A short summary of our results and methods of proofs can be found in the expository paper [69].

This project started from an unlikely conjunction of discussions of the authors with

various people, most notably Yan Guo, Dong Li, Freddy Bouchet and Étienne Ghys. We also got crucial inspiration from the books [9] and [10] by James Binney and Scott Tremaine; and [2] by Serge Alinhac and Patrick Gérard. Warm thanks to Julien Barré, Jean Dolbeault, Thierry Gallay, Stephen Gustafson, Gregory Hammett, Donald Lynden-Bell, Michael Sigal, Éric Séré and especially Michael Kiessling for useful exchanges and references; and to Francis Filbet and Irene Gamba for providing numerical simulations. We are also grateful to Patrick Bernard, Freddy Bouchet, Emanuele Caglioti, Yves Elskens, Yan Guo, Zhiwu Lin, Michael Loss, Peter Markowich, Govind Menon, Yann Ollivier, Mario Pulvirenti, Jeff Rauch, Igor Rodnianski, Peter Smereka, Yoshio Sone, Tom Spencer, and the team of the Princeton Plasma Physics Laboratory for further constructive discussions about our results. Finally, we acknowledge the generous hospitality of several institutions: Brown University, where the first author was introduced to Landau damping by Yan Guo in early 2005; the Institute for Advanced Study in Princeton, who offered the second author a serene atmosphere of work and concentration during the best part of the preparation of this work; Cambridge University, who provided repeated hospitality to the first author thanks to the Award No. KUK-I1-007-43, funded by the King Abdullah University of Science and Technology; and the University of Michigan, where conversations with Jeff Rauch and others triggered a significant improvement of our results.

Our deep thanks go to the referees for their careful examination of the manuscript. We dedicate this paper to two great scientists who passed away during the elaboration of our work. The first one is Carlo Cercignani, one of the leaders of kinetic theory, author of several masterful treatises on the Boltzmann equation, and a long-time personal friend of the second author. The other one is Vladimir Arnold, a mathematician of extraordinary insight and influence; in this paper we shall uncover a tight link between Landau damping and the theory of perturbation of completely integrable Hamiltonian systems, to which Arnold has made major contributions.

1. Introduction to Landau damping

1.1. Discovery

Under adequate assumptions (collisionless regime, non-relativistic motion, heavy ions, no magnetic field), a dilute plasma is well described by the non-linear Vlasov–Poisson equation

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f + \frac{F}{m} \cdot \nabla_v f = 0, \tag{1.1}$$

where $f=f(t,x,v)\geqslant 0$ is the density of electrons in phase space (x is the position and v the velocity), m is the mass of an electron, and $F=F(t,x)$ is the mean-field (self-consistent) electrostatic force:

$$F=-eE, \quad E=\nabla\Delta^{-1}(4\pi\varrho). \tag{1.2}$$

Here $e>0$ is the absolute electron charge, $E=E(t,x)$ is the electric field, and $\varrho=\varrho(t,x)$ is the density of charges

$$\varrho=\varrho_i-e\int_{\mathbb{R}^3} f\,dv, \tag{1.3}$$

ϱ_i being the density of charges due to ions. This model and its many variants are of tantamount importance in plasma physics [1], [5], [49], [54].

In contrast to models incorporating collisions [92], the Vlasov–Poisson equation is *time-reversible*. However, in 1946 Landau [52] stunned the physical community by predicting an irreversible behavior on the basis of this equation. This "astonishing result" (as it was called in [88]) relied on the solution of the Cauchy problem for the linearized Vlasov–Poisson equation around a spatially homogeneous Maxwellian (Gaussian) equilibrium. Landau formally solved the equation by means of Fourier and Laplace transforms, and after a study of singularities in the complex plane, concluded that the electric field decays exponentially fast; he further studied the rate of decay as a function of the wave vector k. Landau's computations are reproduced in [54, §34] and [1, §4.2].

An alternative argument appears in [54, §30]: there the thermodynamical formalism is used to compute the amount of heat Q which is dissipated when a (small) oscillating electric field $E(t,x)=Ee^{i(k\cdot x-\omega t)}$ (k is a wave vector and $\omega>0$ a frequency) is applied to a plasma whose distribution f^0 is homogeneous in space and isotropic in velocity space; the result is

$$Q=-|E|^2\frac{\pi me^2\omega}{|k|^2}\phi'\left(\frac{\omega}{|k|}\right), \tag{1.4}$$

where

$$\phi(v_1)=\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} f^0(v_1,v_2,v_3)\,dv_2\,dv_3.$$

In particular, (1.4) is always positive (see the last remark in [54, §30]), which means that the system reacts against the perturbation, and thus possesses some "active" stabilization mechanism.

A third argument [54, §32] consists of studying the dispersion relation, or equivalently searching for the (generalized) eigenmodes of the linearized Vlasov–Poisson equation, now with complex frequency ω. After appropriate selection, these eigenmodes are all *decaying* ($\operatorname{Im}\omega<0$) as $t\to\infty$. This again suggests stability, although in a somewhat weaker sense than the computation of heat release.

The first and third arguments also apply to the *gravitational* Vlasov–Poisson equation, which is the main model for non-relativistic galactic dynamics. This equation is similar to (1.1), but now m is the mass of a typical star (!), and f is the density of stars in phase space; moreover the first equation of (1.2) and the relation (1.3) should be replaced by

$$F = -\mathcal{G}mE \quad \text{and} \quad \varrho = m \int_{\mathbb{R}^3} f \, dv, \tag{1.5}$$

where $\mathcal{G}$ is the gravitational constant, E is the gravitational field and ϱ is the density of mass. The books [9] and [10] by Binney and Tremaine constitute excellent references about the use of the Vlasov–Poisson equation in stellar dynamics—where it is often called the "collisionless Boltzmann equation", see footnote on p. 276 in [10]. On "intermediate" time scales, the Vlasov–Poisson equation is thought to be an accurate description of very large star systems [28], which are now accessible to numerical simulations.

Since the work of Lynden-Bell [58] it has been recognized that Landau damping, and wilder collisionless relaxation processes generically dubbed "violent relaxation", constitute a fundamental stabilizing ingredient of galactic dynamics. Without these still poorly understood mechanisms, the surprisingly short time scales for relaxation of the galaxies would remain unexplained.

One main difference between the electrostatic and the gravitational interactions is that in the latter case Landau damping should occur only at wavelengths smaller than the *Jeans length* [10, §5.2]; beyond this scale, even for Maxwellian velocity profiles, the Jeans instability takes over and governs planet and galaxy aggregation.([1])

On the contrary, in (classical) plasma physics, Landau damping should hold at all scales under suitable assumptions on the velocity profile; and in fact one is in general not interested in scales smaller than the *Debye length*, which is roughly defined in the same way as the Jeans length.

Nowadays, not only has Landau damping become a cornerstone of plasma physics,([2]) but it has also made its way into other areas of physics (astrophysics, but also wind waves, fluids, superfluids, etc.) and even biophysics. One may consult the concise survey papers [76], [81] and [91] for a discussion of its influence and some applications.

([1]) Or at least would do, if galactic matter was smoothly distributed; in presence of "microscopic" heterogeneities, a phase transition for aggregation can occur far below this scale [48]. In the language of statistical mechanics, the Jeans length corresponds to a "spinodal point" rather than a phase transition [87].

([2]) Ryutov [81] estimated in 1998 that "approximately every third paper on plasma physics and its applications contains a direct reference to Landau damping".

1.2. Interpretation

True to his legend, Landau deduced the damping effect from a mathematical-style study,([3]) without bothering to give a physical explanation of the underlying mechanism. His arguments anyway yield exact formulae, which in principle can be checked experimentally, and indeed provide good qualitative agreement with observations [59].

A first set of problems in the interpretation is related to the arrow of time. In the thermodynamic argument, the exterior field is awkwardly imposed from time $-\infty$ onwards; moreover, reconciling a positive energy dissipation with the reversibility of the equation is not obvious. In the dispersion argument, one has to arbitrarily impose the location of the singularities taking into account the arrow of time (via the Plemelj formula); then the spectral study requires some thinking. All in all, the most convincing argument remains Landau's original one, since it is based only on the study of the Cauchy problem, which makes more physical sense than the study of the dispersion relation (see the remark in [9, p. 682]).

A more fundamental issue resides in the use of analytic function theory, with contour integration, singularities and residue computation, which has played a major role in the theory of the Vlasov–Poisson equation ever since Landau [54, Chapter 32], see also [10, §5.2.4], and helps little, if at all, to understand the underlying physical mechanism.([4])

The most popular interpretation of Landau damping considers the phenomenon from an energetic point of view, as the result of the interaction of a plasma wave with particles of nearby velocity [36, p. 18], [10, p. 412], [1, §4.2.3], [54, p. 127]. In a nutshell, the argument says that dominant exchanges occur with those particles which are "trapped" by the wave because their velocity is close to the wave velocity. If the distribution function is a decreasing function of $|v|$, among trapped particles more are accelerated than are decelerated, so the wave loses energy to the plasma—or the plasma surfs on the wave—and the wave is damped by the interaction.

Appealing as this image may seem, to a mathematically-oriented mind it will probably make little sense at first hearing.([5]) A more down-to-earth interpretation emerged in the fifties from the "wave packet" analysis of van Kampen [45] and Case [14]: Landau damping would result from *phase mixing*. This phenomenon, well known in galactic dynamics, describes the damping of oscillations occurring when a continuum is transported

([3]) Not completely rigorous from the mathematical point of view, but formally correct, in contrast to the previous studies by Landau's fellow physicists—as Landau himself pointed out without mercy [52].

([4]) van Kampen [45] summarizes the conceptual problems posed to his contemporaries by Landau's treatment, and comments on more or less clumsy attempts to resolve the apparent paradox caused by the singularities in the complex plane.

([5]) Escande [25, Chapter 4, footnote 6] points out some misconceptions associated with the surfer image.

in phase space along an anharmonic Hamiltonian flow [10, pp. 379–380]. The mixing results from the simple fact that particles following different orbits travel at different angular([6]) speeds, so perturbations start "spiraling" (see Figure 4.27 on p. 379 in [10]) and homogenize by fast spatial oscillation. From the mathematical point of view, phase mixing results in *weak convergence*; from the physical point of view, this is just the convergence of observables, defined as averages over the velocity space (this is sometimes called "convergence in the mean").

At first sight, both points of view seem hardly compatible: Landau's scenario suggests a very smooth process, while phase mixing involves tremendous oscillations. The coexistence of these two interpretations did generate some speculation on the nature of the damping, and on its relation to phase mixing, see e.g. [46] or [10, p. 413]. There is actually no contradiction between the two points of view: many physicists have rightly pointed out that Landau damping should come with filamentation and oscillations of the distribution function [45, p. 962], [54, p. 141], [1, Vol. 1, pp. 223–224], [57, pp. 294–295]. Nowadays these oscillations can be visualized spectacularly due to deterministic numerical schemes, see e.g. [95], [39, Figure 3] and [27]. In Figure 1.2 we reproduce some examples provided by Filbet.

In any case, there is still no definite interpretation of Landau damping: as noted by Ryutov [81, §9], papers devoted to the interpretation and teaching of Landau damping were still appearing regularly fifty years after its discovery; to quote just a couple of more recent examples let us mention works by Elskens and Escande [23], [24], [25]. The present paper will also contribute a new point of view.

1.3. Range of validity

The following issues are addressed in the literature [42], [46], [61], [95] and slightly controversial:

• Does Landau damping really hold for gravitational interaction? The case seems thinner in this situation than for plasma interaction, all the more as there are many instability results in the gravitational context; up to now there has been no consensus among mathematical physicists [79]. (Numerical evidence is not conclusive because of the difficulty of accurate simulations in very large time—even in one dimension of space.)

• Does the damping hold for unbounded systems? Counterexamples from [30] and [31] show that some kind of confinement is necessary, even in the electrostatic case. More precisely, Glassey and Schaeffer show that a solution of the linearized Vlasov–Poisson

([6]) "Angular" here refers to action-angle variables, and applies even for straight trajectories in a torus.

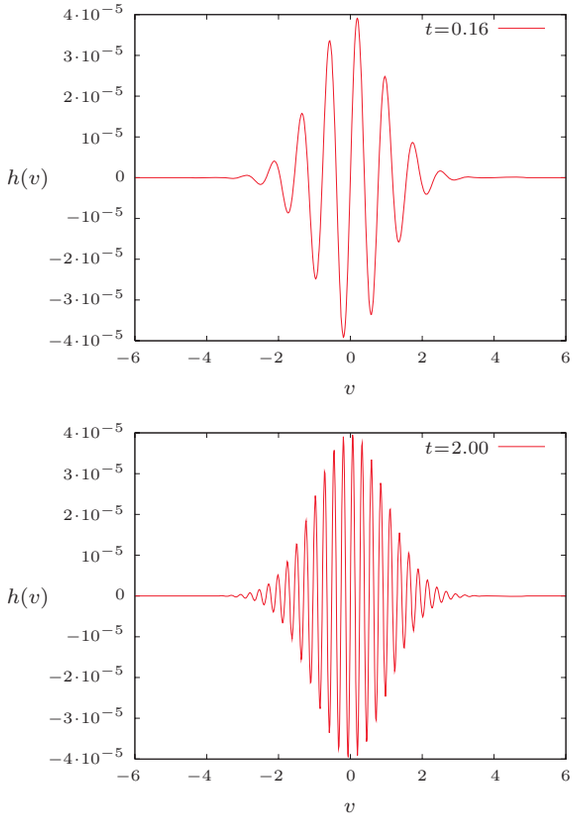

Figure 1. A slice of the distribution function (relative to a homogeneous equilibrium) for gravitational Landau damping, at two different times.

equation in the whole space (linearized around a homogeneous equilibrium f^0 of infinite mass) decays at best like $O(t^{-1})$, modulo logarithmic corrections, for $f^0(v)=c/(1+|v|^2)$; and like $O((\log t)^{-\alpha})$ if f^0 is a Gaussian. In fact, Landau's original calculations already indicated that the damping is *extremely* weak at large wavenumbers; see the discussion in [54, §32]. Of course, in the gravitational case, this is even more dramatic because of the Jeans instability.

• Does convergence hold in infinite time for the solution of the "full" non-linear equation? This is not clear at all since there is no mechanism that would keep the distribution close to the original equilibrium *for all times*. Some authors do not believe that there is convergence as $t\to\infty$; others believe that there is convergence but argue that it should be very slow [42], say $O(1/t)$. In the first mathematically rigorous study of the subject, Backus [4] notes that in general the linear and non-linear evolution break

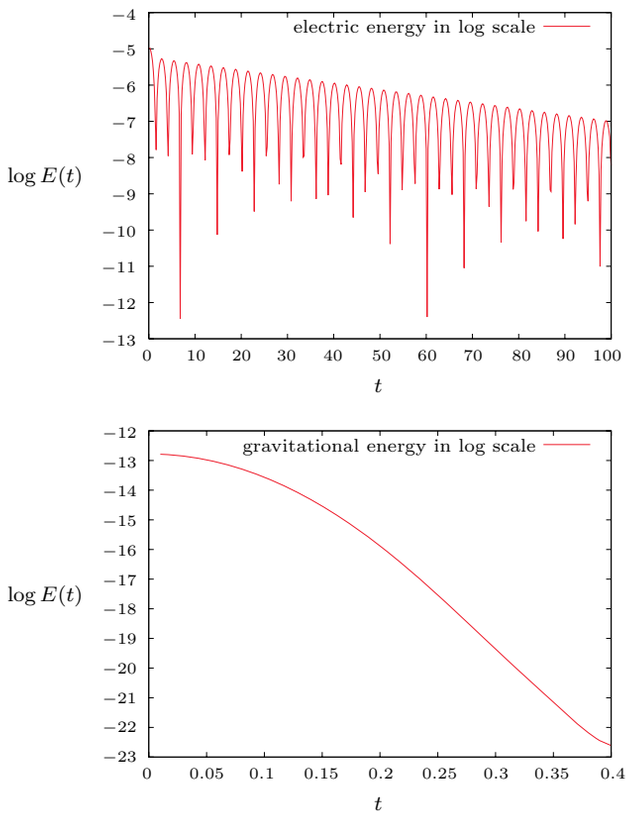

Figure 2. Time-evolution of the norm of the field, for electrostatic (left) and gravitational (right) interactions. Notice the fast Langmuir oscillations in the electrostatic case.

apart after some (not very large) time, and questions the validity of the linearization.([7]) O'Neil [75] argues that relaxation holds in the "quasilinear regime" on larger time scales, when the "trapping time" (roughly proportional the inverse square root of the size of the perturbation) is much smaller than the damping time. Other speculations and arguments related to trapping appear in many sources, e.g. [61] and [64]. Kaganovich [44] argues that non-linear effects may quantitatively affect Landau damping related phenomena by several orders of magnitude.

The so-called "quasilinear relaxation theory" [54, §49], [1, §9.1.2], [49, Chapter 10] uses second-order approximation of the Vlasov equation to predict the convergence of the spatial average of the distribution function. The procedure is most esoteric, involving av-

([7]) From the abstract: "The linear theory predicts that in stable plasmas the neglected term will grow linearly with time at a rate proportional to the initial disturbance amplitude, destroying the validity of the linear theory, and vitiating positive conclusions about stability based on it."

eraging over statistical ensembles, and diffusion equations with discontinuous coefficients, acting only near the resonance velocity for particle-wave exchanges. Because of these discontinuities, the predicted asymptotic state is discontinuous, and collisions are invoked to restore smoothness. Linear Fokker–Planck equations([8]) in velocity space have also been used in astrophysics [58, p. 111], but only on phenomenological grounds (the ad-hoc addition of a friction term leading to a Gaussian stationary state); and this procedure has been exported to the study of 2-dimensional incompressible fluids [15], [16].

Even if it were more rigorous, quasilinear theory only aims at second-order corrections, but the effect of higher-order perturbations might be even worse. Think of something like

$$e^{-t} \sum_{n \in \mathbb{N}_0} \frac{\varepsilon^n t^n}{\sqrt{n!}}$$

(where $\mathbb{N}_0 = \{0, 1, 2, \dots\}$), then truncation at any order in ε converges exponentially fast as $t \to \infty$, but the whole sum diverges to infinity.

Careful numerical simulation [95] seems to show that the solution of the non-linear Vlasov–Poisson equation does converge to a spatially homogeneous distribution, but only as long as the size of the perturbation is small enough. We shall call this phenomenon *non-linear Landau damping*. This terminology summarizes the problem well, still it is subject to criticism since (a) Landau himself sticked to the linear case and did not discuss the large-time convergence of the distribution function; (b) damping is expected to hold when the regime is close to linear, but not necessarily when the non-linear term dominates;([9]) and (c) this expression is also used to designate related but different phenomena [1, §10.1.3]. It should be kept in mind that in the present paper, non-linearity does manifest itself, not because there is a significant initial departure from equilibrium (our initial data will be very close to equilibrium), but because we are addressing very large times, and this is all the more tricky to handle, as the problem is highly oscillating.

• Is Landau damping related to the more classical notion of stability in orbital sense? Orbital stability means that the system, slightly perturbed at initial time from an equilibrium distribution, will always remain close to this equilibrium. Even in the favorable electrostatic case, stability is not granted; the most prominent phenomenon being the Penrose instability [77] according to which a distribution with two deep bumps may be unstable. In the more subtle gravitational case, various stability and instability criteria are associated with the names of Chandrasekhar, Antonov, Goodman, Doremus, Feix, Baumann, etc. [10, §7.4]. There is a widespread agreement (see e.g. the comments

([8]) These equations act on some ensemble average of the distribution; they are different from the Vlasov–Landau equation.

([9]) Although phase mixing might still play a crucial role in violent relaxation or other unclassified non-linear phenomena.

in [95]) that Landau damping and stability are related, and that Landau damping cannot be hoped for if there is no orbital stability.

1.4. Conceptual problems

Summarizing, we can identify three main conceptual obstacles which make Landau damping mysterious, even sixty years after its discovery:

(i) The equation is time-reversible, yet we are looking for an irreversible behavior as $t \to \infty$ (or $t \to -\infty$). The value of the entropy does not change in time, which physically speaking means that there is no loss of information in the distribution function. The spectacular experiment of the "plasma echo" illustrates this conservation of microscopic information [32], [60]: a plasma which is apparently back to equilibrium after an initial disturbance, will react to a second disturbance in a way that shows that it has not forgotten the first one.([10]) And at the linear level, if there are decaying modes, there also have to be growing modes!

(ii) When one perturbs an equilibrium, there is no mechanism forcing the system to go back to this equilibrium in large time; so there is no justification in the use of linearization to predict the large-time behavior.

(iii) At the technical level, Landau damping (in Landau's own treatment) rests on analyticity, and its most attractive interpretation is in terms of phase mixing. But both phenomena are *incompatible in the large-time limit*: phase mixing implies an irreversible deterioration of analyticity. For instance, it is easily checked that free transport induces an exponential growth of analytic norms as $t \to \infty$—except if the initial datum is spatially homogeneous. In particular, the Vlasov–Poisson equation is *unstable* (in large time) in any norm incorporating velocity regularity. (Space-averaging is one of the ingredients used in the quasilinear theory to formally get rid of this instability.)

How can we respond to these issues?

One way to solve the first problem (time-reversibility) is to appeal to van Kampen modes as in [10, p. 415]; however these are not so physical, as noticed in [9, p. 682]. A simpler conceptual solution is to invoke the notion of weak convergence: reversibility manifests itself in the conservation of the information contained in the density function; but information may be lost irreversibly in the limit when we consider weak convergence. Weak convergence only describes the long-time behavior of arbitrary observables, each of which does not contain as much information as the density function.([11]) As a very simple

([10]) Interestingly enough, this experiment was suggested as a way to evaluate the strength of irreversible phenomena going on inside a plasma, e.g. the collision frequency, by measuring attenuations with respect to the predicted echo. See [86] for an interesting application and striking pictures.

([11]) In Lynden-Bell's appealing words [57, p. 295], "a system whose density has achieved a steady

illustration, consider the time-reversible evolution defined by $u(t,x)=e^{itx}u_i(x)$, and notice that it does converge weakly to 0 as $t\to\pm\infty$; this convergence is even exponentially fast if the initial datum u_i is analytic. (Our example is not chosen at random: although it is extremely simple, it may be a good illustration of what happens in phase mixing.) In a way, microsocopic reversibility is compatible with macroscopic irreversibility, provided that the "microscopic regularity" is destroyed asymptotically.

Still in respect to this reversibility, it should be noted that the "dual" mechanism of *radiation*, according to which an infinite-dimensional system may lose energy towards very large scales, is relatively well understood and recognized as a crucial stability mechanism [3], [85].

The second problem (lack of justification of the linearization) only indicates that there is a wide gap between the understanding of linear Landau damping, and that of the non-linear phenomenon. Even if unbounded corrections appear in the linearization procedure, the effect of the large terms might be averaged over time or other variables.

The third problem, maybe the most troubling from an analyst's perspective, does not dismiss the phase mixing explanation, but suggests that we shall have to keep track of the initial time, in the sense that a rigorous proof cannot be based on the propagation of some phenomenon. This situation is of course in sharp contrast with the study of dissipative systems possessing a Lyapunov functional, as do many collisional kinetic equations [92], [93]; it will require completely different mathematical techniques.

1.5. Previous mathematical results

At the linear level, the first rigorous treatments of Landau damping were performed in the sixties; see Saenz [82] for rather complete results and a review of earlier works. The theory was rediscovered and renewed at the beginning of the eighties by Degond [20], and Maslov and Fedoryuk [63]. In all these works, analytic arguments play a crucial role (for instance for the analytic extension of resolvent operators), and asymptotic expansions for the electric field associated with the linearized Vlasov–Poisson equation are obtained.

Also at the linearized level, there are counterexamples by Glassey–Schaeffer [30], [31] showing that there is in general no exponential decay for the linearized Vlasov–Poisson equation without analyticity, or without confining.

In a *non-linear* setting, the only rigorous treatments so far are those by Caglioti–Maffei [13], and later Hwang–Vélazquez [41]. Both sets of authors work in the 1-dimensional torus and use fixed-point theorems and perturbative arguments to prove the *existence* of a class of analytic solutions behaving, asymptotically as $t\to\infty$, and in a

state will have information about its birth still stored in the peculiar velocities of its stars."

strong sense, like perturbed solutions of free transport. Since solutions of free transport weakly converge to spatially homogeneous distributions, the solutions constructed by this "scattering" approach are indeed damped. The weakness of these results is that they say nothing about the initial perturbations leading to such solutions, which could be very special. In other words: damped solutions do exist, but do we ever reach them?

Sparse as it may seem, this list is kind of exhaustive. On the other hand, there is a rather large mathematical literature on the orbital stability problem, due to Guo, Rein, Strauss, Wolansky and Lemou–Méhats–Raphaël. In this respect see for instance [35] for the plasma case, and [34] and [53] for the gravitational case; these sources contain many references on the subject. This body of works has confirmed the intuition of physicists, although with quite different methods. The gap between a formal, linear treatment and a rigorous, non-linear one is striking: compare the appendix of [34] to the rest of the paper. In the gravitational case, these works do not consider homogeneous equilibria, but only localized solutions.

Our treatment of Landau damping will be performed from scratch, and will not rely on any of these results.

2. Main result

2.1. Modeling

We shall work in adimensional units throughout the paper, in d dimensions of space and d dimensions of velocity ($d \in \mathbb{N} = \{1, 2, \dots\}$).

As should be clear from our presentation in §1, to observe Landau damping, we need to put a restriction on the length scale (anyway plasmas in experiments are usually confined). To achieve this we shall take the position space to be the d-dimensional torus of sidelength L, namely $\mathbb{T}^d_L = \mathbb{R}^d / L\mathbb{Z}^d$. This is admittedly a bit unrealistic, but it is commonly done in plasma physics (see e.g. [5]).

In a periodic setting the Poisson equation has to be reinterpreted, since $\Delta^{-1}\varrho$ is not well defined unless $\int_{\mathbb{T}^d_L} \varrho \, dx = 0$. The natural solution consists of removing the mean value of ϱ, independently of any "neutrality" assumption. Let us sketch a justification in the important case of Coulomb interaction: due to the screening phenomenon, we may replace the Coulomb potential V by a potential $V_\varkappa$ exhibiting a "cutoff" at large distances (typically $V_\varkappa$ could be of Debye type [5]; anyway the choice of approximation has no influence on the result). If $\nabla V_\varkappa \in L^1(\mathbb{R}^d)$, then $\nabla V_\varkappa * \varrho$ makes sense for a periodic ϱ, and moreover

$$(\nabla V_\varkappa * \varrho)(x) = \int_{\mathbb{R}^d} \nabla V_\varkappa(x-y)\varrho(y)\, dy = \int_{[0,L]^d} \nabla V^{(L)}_\varkappa(x-y)\varrho(y)\, dy,$$

where $V_\varkappa^{(L)}(z)=\sum_{l\in\mathbb{Z}^d} V_\varkappa(z+lL)$. Passing to the limit as $\varkappa\to 0$ yields

$$\int_{[0,L]^d} \nabla V^{(L)}(x-y)\varrho(y)\,dy = \int_{[0,L]^d} \nabla V^{(L)}(x-y)(\varrho-\langle\varrho\rangle)(y)\,dy = -\nabla\Delta_L^{-1}(\varrho-\langle\varrho\rangle),$$

where Δ_L^{-1} is the inverse Laplace operator on $\mathbb{T}_L^d$.

In the case of galactic dynamics there is no screening; however it is customary to remove the zeroth-order term of the density. This is known as the *Jeans swindle*, a trick considered as efficient but logically absurd. In 2003, Kiessling [47] reopened the case and acquitted Jeans, on the basis that his "swindle" can be justified by a simple limit procedure, similar to the one presented above; however, the physical basis for the limit is less transparent and subject to debate. For our purposes, it does not matter much: since anyway periodic boundary conditions are not realistic in a cosmological setting, we may just as well say that we adopt the Jeans swindle as a simple phenomenological model.

More generally, we may consider any interaction potential W on $\mathbb{T}_L^d$, satisfying the natural symmetry assumption $W(-z)=W(z)$ (that is, W is even), as well as certain regularity assumptions. Then the self-consistent field will be given by

$$F = -\nabla W * \varrho, \quad \varrho(x) = \int_{\mathbb{R}^d} f(x,v)\,dv,$$

where now $*$ denotes the convolution on $\mathbb{T}_L^d$.

In accordance with our conventions from Appendix A.3, we shall write

$$\widehat{W}^{(L)}(k) = \int_{\mathbb{T}_L^d} e^{-2i\pi k\cdot x/L} W(x)\,dx.$$

In particular, if W is the periodization of a potential $\mathbb{R}^d\to\mathbb{R}$ (still denoted W by abuse of notation), i.e.,

$$W(x) = W^{(L)}(x) = \sum_{l\in\mathbb{Z}^d} W(x+lL),$$

then

$$\widehat{W}^{(L)}(k) = \widehat{W}\left(\frac{k}{L}\right), \tag{2.1}$$

where

$$\widehat{W}(\xi) = \int_{\mathbb{R}^d} e^{-2i\pi\xi\cdot x} W(x)\,dx$$

is the original Fourier transform in the whole space.

2.2. Linear damping

It is well known that Landau damping requires some stability assumptions on the unperturbed homogeneous distribution function, say $f^0(v)$. In this paper we shall use a very general assumption, expressed in terms of the Fourier transform

$$\tilde{f}^0(\eta) = \int_{\mathbb{R}^d} e^{-2i\pi\eta\cdot v} f^0(v)\, dv, \tag{2.2}$$

the length L, and the interaction potential W. To state it, we define, for $t\geqslant 0$ and $k\in\mathbb{Z}^d$,

$$K^0(t,k) = -4\pi^2 \widehat{W}^{(L)}(k)\tilde{f}^0\left(\frac{kt}{L}\right)\frac{|k|^2 t}{L^2}; \tag{2.3}$$

and, for any $\xi\in\mathbb{C}$, we define a function $\mathcal{L}$ via the following Fourier–Laplace transform of K^0 in the time variable:

$$\mathcal{L}(\xi,k) = \int_0^\infty e^{2\pi\xi^*|k|t/L} K^0(t,k)\, dt, \tag{2.4}$$

where ξ^* is the complex conjugate to ξ. Our linear damping condition is expressed as follows:

(L) *There are constants $C_0,\lambda,\varkappa>0$ such that $|\tilde{f}^0(\eta)|\leqslant C_0 e^{-2\pi\lambda|\eta|}$ for any $\eta\in\mathbb{R}^d$; and for any $\xi\in\mathbb{C}$ with $0\leqslant \mathrm{Re}\,\xi<\lambda$,*

$$\inf_{k\in\mathbb{Z}^d} |\mathcal{L}(\xi,k)-1| \geqslant \varkappa.$$

We shall prove in §3 that **(L)** implies Landau damping. For the moment, let us give a few sufficient conditions for **(L)** to be satisfied. The first one can be thought of as a smallness assumption on either the length, or the potential, or the velocity distribution. The other conditions involve the marginals of f^0 along arbitrary wave vectors k:

$$\varphi_k(v) = \int_{kv/|k|+k^\perp} f^0(w)\, dw, \quad v\in\mathbb{R}. \tag{2.5}$$

All studies known to us are based on one of these assumptions, so **(L)** appears as a unifying condition for linear Landau damping around a homogeneous equilibrium.

PROPOSITION 2.1. *Let $f^0=f^0(v)$ be a velocity distribution such that $\tilde{f}^0$ decays exponentially fast at infinity, let $L>0$ and let W be an even interaction potential on $\mathbb{T}^d_L$, $W\in L^1(\mathbb{T}^d)$. If any of the following conditions is satisfied*:

(a) smallness:

$$4\pi^2\Big(\max_{k\in\mathbb{Z}^d_*}|\widehat{W}^{(L)}(k)|\Big)\sup_{|\sigma|=1}\int_0^\infty |\tilde{f}^0(r\sigma)|r\, dr < 1; \tag{2.6}$$

(b) repulsive interaction and decreasing marginals: *for all* $k\in\mathbb{Z}^d$ *and* $v\in\mathbb{R}$,

$$\widehat{W}^{(L)}(k)\geqslant 0 \quad \textit{and} \quad \varphi_k'(v)\begin{cases} >0, & \textit{if } v<0,\\ <0, & \textit{if } v>0,\end{cases} \tag{2.7}$$

(c) generalized Penrose condition on marginals: *for all* $k\in\mathbb{Z}^d$,

$$\varphi_k'(w)=0 \implies \widehat{W}^{(L)}(k)\Big(\text{p.v.}\int_{\mathbb{R}}\frac{\varphi_k'(v)}{v-w}\,dv\Big)<1 \quad \textit{for all } w\in\mathbb{R}; \tag{2.8}$$

then **(L)** *holds true for some* $C_0,\lambda,\varkappa>0$.

Remark 2.2. ([54, problem in §30]) If f^0 is radially symmetric and positive, and $d\geqslant 3$, then all marginals of f^0 are decreasing functions of $|v|$. Indeed, if

$$\varphi(v)=\int_{\mathbb{R}^{d-1}} f\big(\sqrt{v^2+|w|^2}\big)\,dw,$$

then after differentiation and integration by parts we find that

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

Example 2.3. Take a gravitational interaction and Mawellian background:

$$\widehat{W}(k)=-\frac{\mathcal{G}}{\pi|k|^2} \quad \text{and} \quad f^0(v)=\varrho^0\frac{e^{-|v|^2/2T}}{(2\pi T)^{d/2}}.$$

Recalling (2.1), we see that (2.6) becomes

$$L<\sqrt{\frac{\pi T}{\mathcal{G}\varrho^0}}=:L_J(T,\varrho^0). \tag{2.9}$$

The length L_J is the celebrated Jeans length [10], [47], so criterion (a) can be applied, all the way up to the onset of the Jeans instability.

Example 2.4. If we replace the gravitational interaction by the electrostatic interaction, the same computation yields

$$L<\sqrt{\frac{\pi T}{e^2\varrho^0}}=:L_D(T,\varrho^0), \tag{2.10}$$

and now L_D is essentially the Debye length. Then criterion (a) becomes quite restrictive, but because the interaction is repulsive we can use criterion (b) as soon as f^0 is a strictly monotone function of $|v|$; this covers in particular Maxwellian distributions, *independently of the size of the box*. Criterion (b) also applies if $d\geqslant 3$ and f^0 has radial symmetry. For a given $L>0$, the condition **(L)** being open, it will also be satisfied if f^0 is a small (analytic) perturbation of a profile satisfying (b); this includes the so-called "small bump on tail" stability. Then if the distribution presents two large bumps, the Penrose instability will take over.

Example 2.5. For the electrostatic interaction in dimension 1, (2.8) becomes

$$(f^0)'(w)=0 \quad \Longrightarrow \quad \int_{\mathbb{R}} \frac{(f^0)'(v)}{v-w}\,dv < \frac{\pi}{e^2 L^2}. \tag{2.11}$$

This is a variant of the Penrose stability condition [77]. This criterion is in general sharp for linear stability (up to the replacement of the strict inequality by the non-strict one, and assuming that the critical points of f^0 are non-degenerate); see [55, Appendix] for precise statements.

We shall show in §3 that **(L)** implies linear Landau damping (Theorem 3.1); then we shall prove Proposition 2.1 at the end of that section. The general ideas are close to those appearing in previous works, including Landau himself; the only novelties lie in the more general assumptions, the elementary nature of the arguments, and the slightly more precise quantitative results.

2.3. Non-linear damping

As others have done before in the study of the Vlasov–Poisson equation [13], we shall quantify the analyticity by means of natural norms involving Fourier transform in both variables (also denoted with a tilde in the sequel). So we define

$$\|f\|_{\lambda,\mu} = \sup_{\substack{k\in\mathbb{Z}^d \\ \eta\in\mathbb{R}^d}} |\tilde{f}^{(L)}(k,\eta)| e^{2\pi\lambda|\eta|} e^{2\pi\mu|k|/L}, \tag{2.12}$$

where k varies in $\mathbb{Z}^d$, $\eta\in\mathbb{R}^d$, λ and μ are positive parameters, and we recall the dependence of the Fourier transform on L (see Appendix A.3 for conventions). Now we can state our main result as follows.

THEOREM 2.6. (Non-linear Landau damping) *Let $f^0:\mathbb{R}^d\to\mathbb{R}_+$ be an analytic velocity profile. Let $L>0$ and let $W:\mathbb{T}^d_L\to\mathbb{R}$ be an even interaction potential satisfying*

$$|\widehat{W}^{(L)}(k)| \leqslant \frac{C_W}{|k|^{1+\gamma}} \quad \textit{for all } k\in\mathbb{Z}^d \tag{2.13}$$

for some constants $C_W>0$ and $\gamma\geqslant 1$. Assume that f^0 and W satisfy the stability condition **(L)** *from* §2.2, *with some constants $\lambda,\varkappa>0$; further assume that, for the same parameter λ,*

$$\sup_{\eta\in\mathbb{R}^d} |\tilde{f}^0(\eta)| e^{2\pi\lambda|\eta|} \leqslant C_0 \quad \textit{and} \quad \sum_{n\in\mathbb{N}_0^d} \frac{\lambda^n}{n!} \|\nabla_v^n f^0\|_{L^1(\mathbb{R}^d)} \leqslant C_0 < \infty. \tag{2.14}$$

Then for any $0<\lambda'<\lambda$, $\beta>0$ *and* $0<\mu'<\mu$, *there is*

$$\varepsilon=\varepsilon(d,L,C_W,C_0,\varkappa,\lambda,\lambda',\mu,\mu',\beta,\gamma)$$

with the following property: *if* $f_i=f_i(x,v)$ *is an initial datum satisfying*

$$\delta:=\|f_i-f^0\|_{\lambda,\mu}+\int_{\mathbb{T}^d_L}\int_{\mathbb{R}^d}|f_i-f^0|e^{\beta|v|}\,dv\,dx\leqslant\varepsilon, \tag{2.15}$$

then

- *the unique classical solution* f *to the non-linear Vlasov equation*

$$\frac{\partial f}{\partial t}+v\cdot\nabla_x f-(\nabla W*\varrho)\cdot\nabla_v f=0,\quad \varrho=\int_{\mathbb{R}^d}f\,dv, \tag{2.16}$$

with initial datum $f(0,\cdot)=f_i$, *converges in the weak topology as* $t\to\pm\infty$, *with rate* $O(e^{-2\pi\lambda'|t|})$, *to a spatially homogeneous equilibrium* $f_{\pm\infty}$ (*that is, it converges to* f_∞ *as* $t\to\infty$, *and to* $f_{-\infty}$ *as* $t\to-\infty$);

- *the distribution function composed with the backward free transport* $f(t,x+vt,v)$ *converges strongly to* $f_{\pm\infty}$ *as* $t\to\pm\infty$;

- *the density* $\varrho(t,x)=\int_{\mathbb{R}^d}f(t,x,v)\,dv$ *converges in the strong topology as* $t\to\pm\infty$, *with rate* $O(e^{-2\pi\lambda'|t|})$, *to the constant density*

$$\varrho_\infty=\frac{1}{L^d}\int_{\mathbb{T}^d_L}\int_{\mathbb{R}^d}f_i(x,v)\,dv\,dx;$$

in particular the force $F=-\nabla W*\varrho$ *converges exponentially fast to* 0;

- *the space average* $\langle f\rangle(t,v)=\int_{\mathbb{T}^d_L}f(t,x,v)\,dx$ *converges in the strong topology as* $t\to\pm\infty$, *with rate* $O(e^{-2\pi\lambda'|t|})$, *to* $f_{\pm\infty}$.

More precisely, there are $C>0$, *and spatially homogeneous distributions* $f_\infty(v)$ *and* $f_{-\infty}(v)$, *depending continuously on* f_i *and* W, *such that*

$$\begin{gathered}\sup_{t\in\mathbb{R}}\|f(t,x+vt,v)-f^0(v)\|_{\lambda',\mu'}\leqslant C\delta,\\ |\tilde f_{\pm\infty}(\eta)-\tilde f^0(\eta)|\leqslant C\delta e^{-2\pi\lambda'|\eta|}\quad \textit{for all } \eta\in\mathbb{R}^d\end{gathered} \tag{2.17}$$

and

$$|L^{-d}\tilde f^{(L)}(t,k,\eta)-\tilde f_{\pm\infty}(\eta)1_{k=0}|=O(e^{-2\pi\lambda'|t|/L})\quad \textit{as } t\to\pm\infty,\ \textit{for all } (k,\eta)\in\mathbb{Z}^d\times\mathbb{R}^d,$$

$$\|f(t,x+vt,v)-f_{\pm\infty}(v)\|_{\lambda',\mu'}=O(e^{-2\pi\lambda'|t|/L})\quad \textit{as } t\to\pm\infty,$$

$$\|\varrho(t,\cdot)-\varrho_\infty\|_{C^r(\mathbb{T}^d)}=O(e^{-2\pi\lambda'|t|/L})\quad \textit{as } |t|\to\infty,\ \textit{for all } r\in\mathbb{N}, \tag{2.18}$$

$$\|F(t,\cdot)\|_{C^r(\mathbb{T}^1d)}=O(e^{-2\pi\lambda'|t|/L})\quad \textit{as } |t|\to\infty,\ \textit{for all } r\in\mathbb{N}, \tag{2.19}$$

$$\|\langle f(t,\cdot,v)\rangle-f_{\pm\infty}\|_{C^r_\sigma(\mathbb{R}^d_v)}=O(e^{-2\pi\lambda'|t|/L})\quad \textit{as } t\to\pm\infty,\ \textit{for all } r\in\mathbb{N}\ \textit{and}\ \sigma>0. \tag{2.20}$$

In this statement C^r stands for the usual norm on r-times continuously differentiable functions, and C^r_σ involves in addition moments of order σ, namely

$$\|f\|_{C^r_\sigma} = \sup_{\substack{r' \leqslant r \\ v \in \mathbb{R}^d}} |f^{(r')}(v)(1+|v|^\sigma)|.$$

These results can be reformulated in a number of alternative norms, both for the strong and for the weak topology.

2.4. Comments

Let us start with a list of remarks about Theorem 2.6.

• The decay of the force field, statement (2.19), is the experimentally measurable phenomenon which may be called Landau damping.

• Since the energy

$$E = \frac{1}{2} \int_{\mathbb{T}^d_L} \int_{\mathbb{T}^d_L} \varrho(x)\varrho(y)\, W(x-y)\, dx\, dy + \int_{\mathbb{T}^d_L} \int_{\mathbb{R}^d} f(x,v) \frac{|v|^2}{2}\, dv\, dx$$

(= potential + kinetic energy) is conserved by the non-linear Vlasov evolution, there is a conversion of potential energy into kinetic energy as $t \to \infty$ (kinetic energy goes up for Coulomb interaction and goes down for Newton interaction). Similarly, the entropy

$$S = -\int_{\mathbb{T}^d_L} \int_{\mathbb{R}^d} f \log f\, dv\, dx = -\left(\int_{\mathbb{T}^d_L} \varrho \log \varrho\, dx + \int_{\mathbb{T}^d_L} \int_{\mathbb{R}^d} f \log \frac{f}{\varrho}\, dv\, dx \right)$$

(= spatial + kinetic entropy) is preserved, and there is a transfer of information from spatial to kinetic variables in large time.

• Our result covers both attractive and repulsive interactions, as long as the linear damping condition is satisfied; it covers the Newton–Coulomb potential as a limit case ($\gamma = 1$ in (2.13)). The proof breaks down for $\gamma < 1$; this is a non-linear effect, as any $\gamma > 0$ would work for the linearized equation. The singularity of the interaction at short scales will be the source of important technical problems.([12])

• Condition (2.14) could be replaced by

$$|\tilde{f}^0(\eta)| \leqslant C_0 e^{-2\pi\lambda|\eta|} \quad \text{and} \quad \int_{\mathbb{R}^d} f^0(v) e^{\beta|v|}\, dv \leqslant C_0. \tag{2.21}$$

([12]) In a related subject, this singularity is also the reason why the Vlasov–Poisson equation is still far from being established as a mean-field limit of particle dynamics (see [37] for partial results covering much less singular interactions).

But condition (2.14) is more general, in view of Theorem 4.20 below. For instance, $f^0(v)=1/(1+v^2)$ in dimension $d=1$ satisfies (2.14) but not (2.21); this distribution is commonly used in theoretical and numerical studies, see e.g. [39]. We shall also establish slightly more precise estimates under slightly more stringent conditions on f^0, see (12.1).

• Our conditions are expressed in terms of the initial datum, which is a considerable improvement over [13] and [41]. Still it is of interest to pursue the "scattering" program started in [13], e.g. in hope of better understanding of the non-perturbative regime.

• The smallness assumption on f_i-f^0 is expected, for instance in view of the work of O'Neil [75], or the numerical results of [95]. We also make the standard assumption that f_i-f^0 is well localized.

• No convergence can be hoped for if the initial datum is only close to f^0 in the weak topology: indeed there is *instability* in the weak topology, even around a Maxwellian [13].

• The well-posedness of the non-linear Vlasov–Poisson equation in dimension $d\leqslant 3$ was established by Pfaffelmoser [78] and Lions–Perthame [56] in the whole space. Pfaffelmoser's proof was adapted to the case of the torus by Batt and Rein [6]; however, like Pfaffelmoser, these authors imposed a stringent assumption of uniformly bounded velocities. Building on Schaeffer's simplification [84] of Pfaffelmoser's argument, Horst [40] proved well-posedness in the whole space assuming only inverse polynomial decay in the velocity variable. Although this has not been done explicitly, Horst's proof can easily be adapted to the case of the torus, and covers in particular the setting which we use in the present paper. (The adaptation of [56] seems more delicate.) Propagation of analytic regularity is not studied in these works. In any case, our proof will provide a new perturbative existence theorem, together with regularity estimates which are considerably stronger than what is needed to prove the uniqueness. We shall not come back to these issues which are rather irrelevant for our study: uniqueness only needs local-in-time regularity estimates, while all the difficulty in the study of Landau damping consists in handling (very) large time.

• We note in passing that while blow-up is known to occur for certain solutions of the Newtonian Vlasov–Poisson equation in dimension 4, blow-up does not occur in this perturbative regime, whatever the dimension. There is no contradiction since blow-up solutions are constructed with negative energy initial data, and a nearly homogeneous solution automatically has positive energy. (Also, blow-up solutions have been constructed only in the whole space, where the virial identity is available; but it is plausible, although not obvious, that blow-up is still possible in bounded geometry.)

• $f(t,\cdot)$ is *not* close to f^0 in analytic norm as $t\to\infty$, and does not converge to anything in the strong topology, so the conclusion cannot be improved much. Still we shall establish more precise quantitative results, and the limit profiles $f_{\pm\infty}$ are obtained

by a constructive argument.

• Estimate (2.17) expresses the orbital "traveling stability" around f^0; it is much stronger than the usual orbital stability in Lebesgue norms [34], [35]. An equivalent formulation is that if $(T_t)_{t\in\mathbb{R}}$ stands for the non-linear Vlasov evolution operator, and $(T_t^0)_{t\in\mathbb{R}}$ for the free transport operator, then in a neighborhood of a homogeneous equilibrium satisfying the stability criterion **(L)**, $T_{-t}^0 \circ T_t$ remains uniformly close to Id for all t. Note the important difference: unlike in the usual orbital stability theory, our conclusions are expressed in functional spaces involving smoothness, *which are not invariant under the free transport semigroup.* This a source of difficulty (our functional spaces are sensitive to the filamentation phenomenon), but it is also the reason for which this "analytic" orbital stability contains much more information, and in particular the damping of the density.

• Compared with known non-linear stability results, and even forgetting about the smoothness, estimate (2.17) is new in several respects. In the context of plasma physics, it is the first one to prove stability for a distribution which is not necessarily a decreasing function of $|v|$ ("small bump on tail"); while in the context of astrophysics, it is the first one to establish stability of a homogeneous equilibrium against periodic perturbations with wavelength smaller than the Jeans length.

• While analyticity is the usual setting for Landau damping, both in mathematical and physical studies, it is natural to ask whether this restriction can be dispensed with. (This can be done only at the price of losing the exponential decay.) In the linear case, this is easy, as we shall recall later in Remark 3.5; but in the non-linear setting, leaving the analytic world is much more tricky. In §13, we shall present the first results in this direction.

With respect to the questions raised above, our analysis brings the following answers:

(a) Convergence of the distribution f does hold for $t\to\infty$; it is indeed based on phase mixing, and therefore involves very fast oscillations. In this sense it is right to consider Landau damping as a "wild" process. But on the other hand, the spatial density (and therefore the force field) converges strongly and smoothly.

(b) The space average $\langle f\rangle$ does converge in large time. However the conclusions are quite different from those of quasilinear relaxation theory, since there is no need for extra randomness, and the limiting distribution is smooth, even without collisions.

(c) Landau damping is a linear phenomenon, which survives non-linear perturbation due to the structure of the Vlasov–Poisson equation. The non-linearity manifests itself by the presence of *echoes*. Echoes were well known to specialists of plasma physics [54, §35], [1, §12.7], but were not identified as a possible source of unstability. Controlling the echoes will be a main technical difficulty; but the fact that the response appears in

this form, with an associated *time-delay* and localized in time, will in the end explain the stability of Landau damping. These features can be expected in other equations exhibiting oscillatory behavior.

(d) The large-time limit is in general different from the limit predicted by the linearized equation, and depends on the interaction and initial datum (more precise statements will be given in §14); still the linearized equation, or higher-order expansions, do provide a good approximation. We shall also set up a systematic recipe for approximating the large-time limit with arbitrarily high precision as the strength of the perturbation becomes small. This justifies a posteriori many known computations.

(e) From the point of view of dynamical systems, the non-linear Vlasov equation exhibits a truly remarkable behavior. It is not uncommon for a Hamiltonian system to have many, or even countably many heteroclinic orbits (there are various theories for this, a popular one being the Melnikov method); but in the present case we see that heteroclinic/homoclinic orbits([13]) are so numerous as to fill up *a whole neighborhood* of the equilibrium. This is possible only because of the infinite-dimensional nature of the system, and the possibility to work with non-equivalent norms; such a behavior has already been reported for other systems [50], [51], in relation with infinite-dimensional KAM (Kolmogorov–Arnold–Moser) theory.

(f) As a matter of fact, non-linear Landau damping has strong similarities with the KAM theory. It has been known since the early days of the theory that the linearized Vlasov equation can be reduced to an infinite system of uncoupled Volterra equations, which makes this equation completely integrable in some sense. (Morrison [66] gave a more precise meaning to this property.) To see a parallel with classical KAM theory, one step of our result is to prove the preservation of the phase-mixing property under non-linear perturbation of the interaction. (Although there is no ergodicity in phase space, the mixing will imply an ergodic behavior for the spatial density.) The analogy is reinforced by the fact that the proof of Theorem 2.6 shares many features with the proof of the KAM theorem in the analytic (or Gevrey) setting. (Our proof is close to Kolmogorov's original argument, exposed in [18].) In particular, we shall invoke a Newton scheme to overcome a loss of "regularity" in analytic norms, only in a trickier sense than in KAM theory. If one wants to push the analogy further, one can argue that the resonances which cause the phenomenon of small divisors in KAM theory find an analogue in the time-resonances which cause the echo phenomenon in plasma physics. A notable difference is that in the present setting, time-resonances arise from the non-linearity at the level of the partial differential equation, whereas small divisors in KAM

([13]) Here we use these words just to designate solutions connecting two distinct/equal equilibria, without any mention of stable or unstable manifolds.

theory arise at the level of the ordinary differential equation. Another major difference is that in the present situation there is a time-averaging which is not present in KAM theory.

Thus we see that three of the most famous paradoxical phenomena from twentieth century classical physics: Landau damping, echoes and the KAM theorem are intimately related (only in the non-linear variant of Landau's linear argument!). This relation, which we did not expect, is one of the main discoveries of the present paper.

2.5. Interpretation

A successful point of view adopted in this paper is that Landau damping is a *relaxation by smoothness* and by *mixing*. In a way, phase mixing converts the smoothness into decay. Thus Landau damping emerges as a rare example of a physical phenomenon in which regularity is not only crucial from the mathematical point of view, but also can be "measured" by a physical experiment.

2.6. Main ingredients

Some of our ingredients are similar to those in [13]: in particular, the use of the Fourier transform to quantify analytic regularity and to implement phase mixing. New ingredients used in our work include the following.

• The introduction of a time-shift parameter to keep memory of the initial time (§4 and §5), thus getting uniform estimates in spite of the loss of regularity in large time. We call this the *gliding regularity*: it shifts in phase space from low to high modes. Gliding regularity automatically comes with an improvement of the regularity in x, and a deterioration of the regularity in v, as time passes by.

• The use of carefully designed flexible analytic norms behaving well with respect to composition (§4). This requires care, because analytic norms are very sensitive to composition, contrary to, say, Sobolev norms.

• A control of the deflection of *trajectories* induced by the force field, to reduce the problem to homogenization of *free flow* (§5) via composition. The physical meaning is the following: when a background with gliding regularity acts on (say) a plasma, the trajectories of plasma particles are asymptotic to free transport trajectories.

• New functional inequalities of bilinear type, involving analytic functional spaces, integration in time and velocity variables, and evolution by free transport (§6). These inequalities morally mean the following: when a plasma acts (by forcing) on a smooth background of particles, the background reacts by lending a bit of its (gliding) regularity

to the plasma, uniformly in time. Eventually the plasma will exhaust itself (the force will decay). This most subtle effect, which is at the heart of Landau's damping, will be mathematically expressed in the formalism of analytic norms with gliding regularity.

• A new analysis of the time response associated with the Vlasov–Poisson equation (§7), aimed ultimately at controlling the self-induced echoes of the plasma. For any interaction less singular than that of Coulomb–Newton, this will be done by analyzing time-integral equations involving a norm of the spatial density. To treat the Coulomb–Newton potential we shall refine the analysis, considering individual modes of the spatial density.

• A Newton iteration scheme, solving the non-linear evolution problem as a succession of linear ones (§10). Picard iteration schemes still play a role, since they are run at each step of the iteration process, to estimate the deflection.

It is only in the linear study of §3 that the length scale L will play a crucial role, via the stability condition **(L)**. In all the rest of the paper we shall normalize L to 1 for simplicity.

2.7. About phase mixing

A physical mechanism transferring energy from large scales to very fine scales, asymptotically in time, is sometimes called *weak turbulence*. Phase mixing provides such a mechanism, and in a way our study shows that the Vlasov–Poisson equation is subject to weak turbulence. But the phase mixing interpretation provides a more precise picture. While one often sees weak turbulence as a "cascade" from low to high Fourier modes, the relevant picture would rather be a 2-dimensional figure with an interplay between spatial Fourier modes and velocity Fourier modes. More precisely, phase mixing transfers the energy from each non-zero spatial frequency k, to large velocity frequences η, and this transfer occurs at a speed proportional to k. This picture is clear from the solution of free transport in Fourier space, and is illustrated in Figure 3. (Note the resemblance with a shear flow.) So there is transfer of energy from one variable (here x) to another (here v); homogenization in the first variable going together with filamentation in the second one. The same mechanism may also underlie other cases of weak turbulence.

The fact that the high modes are ultimately damped by some "random" microscopic process (collisions, diffusion, etc.) not described by the Vlasov–Poisson equation is certainly undisputed in plasma physics [54, §41],([14]) but is the object of debate in galactic dynamics; anyway this is a different story. Some mathematical statistical theo-

([14]) See [54, Problem 41]: due to Landau damping, collisions are expected to smooth the distribution quite efficiently; this is a hypoellipticity issue.

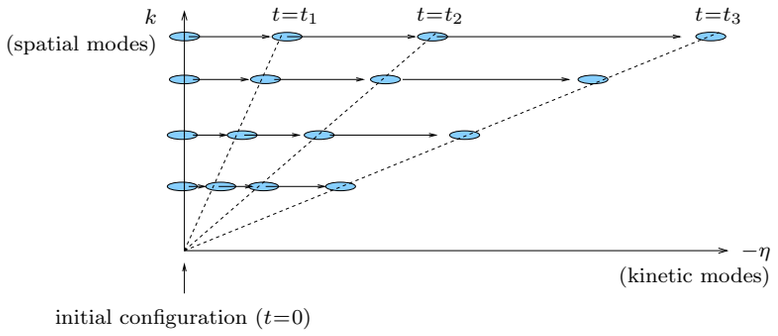

Figure 3. Schematic picture of the evolution of energy by free transport, or perturbation thereof; marks indicate localization of energy in phase space. The energy of the spatial mode k is concentrated in large time around $\eta \simeq -kt$.

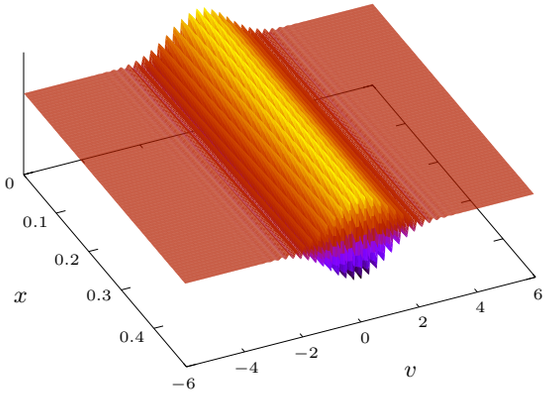

Figure 4. The distribution function in phase space (position, velocity) at a given time; notice how the fast oscillations in v contrast with the slower variations in x.

ries of Euler and Vlasov–Poisson equations do postulate the existence of some small-scale coarse graining mechanism, but resulting in mixing rather than dissipation [80], [90].

3. Linear damping

In this section we establish Landau damping for the linearized Vlasov equation. Beforehand, let us recall that the free transport equation

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

has a strong *mixing* property: any solution of (3.1) converges weakly in large time to a spatially homogeneous distribution equal to the space-averaging of the initial datum.

Let us sketch the proof.

If f solves (3.1) in $\mathbb{T}^d \times \mathbb{R}^d$, with initial datum $f_i = f(0, \cdot)$, then

$$f(t, x, v) = f_i(x - vt, v),$$

so the space-velocity Fourier transform of f is given by the formula

$$\tilde{f}(t, k, \eta) = \tilde{f}_i(k, \eta + kt). \tag{3.2}$$

On the other hand, if f_∞ is defined by

$$f_\infty(v) = \langle f_i(\cdot, v) \rangle = \int_{\mathbb{T}^d} f_i(x, v)\, dx,$$

then $\tilde{f}_\infty(k, \eta) = \tilde{f}_i(0, \eta) 1_{k=0}$. So, by the Riemann–Lebesgue lemma, for any fixed (k, η) we have

$$|\tilde{f}(t, k, \eta) - \tilde{f}_\infty(k, \eta)| \to 0 \quad \text{as } |t| \to \infty,$$

which shows that f converges weakly to f_∞. The convergence holds as soon as f_i is merely integrable; and by (3.2), the rate of convergence is determined by the decay of $\tilde{f}_i(k, \eta)$ as $|\eta| \to \infty$, or equivalently the smoothness in the velocity variable. In particular, the convergence is exponentially fast if (and only if) $f_i(x, v)$ is analytic in v.

This argument obviously works independently of the size of the box. But when we turn to the Vlasov equation, length scales will matter, so we shall introduce a length $L > 0$, and work in $\mathbb{T}^d_L = \mathbb{R}^d / L\mathbb{Z}^d$. Then the length scale will appear in the Fourier transform: see Appendix A.3. (This is the only section in this paper where the scale will play a non-trivial role, so in all the rest of the paper we shall take $L = 1$.)

Any velocity distribution $f^0 = f^0(v)$ defines a stationary state for the non-linear Vlasov equation with interaction potential W. Then the linearization of that equation around f^0 yields

$$\begin{cases} \dfrac{\partial f}{\partial t} + v \cdot \nabla_x f - (\nabla W * \varrho) \cdot \nabla_v f^0 = 0, \\ \varrho = \displaystyle\int_{\mathbb{R}^d} f\, dv. \end{cases} \tag{3.3}$$

Note that there is no force term in (3.3), due to the fact that f^0 does not depend on x. This equation describes what happens to a plasma density f which tries to force a stationary homogeneous background f^0; equivalently, it describes the *reaction* exerted by the background which is acted upon. (Imagine that there is an exchange of matter between the forcing gas and the forced gas, and that this exchange exactly compensates the effect of the force, so that the density of the forced gas does not change after all.)

THEOREM 3.1. (Linear Landau damping) *Let $f^0=f^0(v)$, $L>0$, $W:\mathbb{T}^d_L\to\mathbb{R}$ be such that $W(-z)=W(z)$ and $\|\nabla W\|_{L^1}\leqslant C_W<\infty$ and let $f_i=f_i(x,v)$ be such that*

(i) *condition* **(L)** *from* §2.2 *holds for some constants $\lambda,\varkappa>0$;*

(ii) *for all $\eta\in\mathbb{R}^d$, $|\tilde{f}^0(\eta)|\leqslant C_0e^{-2\pi\lambda|\eta|}$ for some constant $C_0>0$;*

(iii) *for all $k\in\mathbb{Z}^d$ and all $\eta\in\mathbb{R}^d$, $|\tilde{f}_i^{(L)}(k,\eta)|\leqslant C_ie^{-2\pi\alpha|\eta|}$ for some $\alpha,C_i>0$.*

Then, as $t\to\infty$, the solution $f(t,\cdot)$ to the linearized Vlasov equation (3.3) *with initial datum f_i converges weakly to $f_\infty=\langle f_i\rangle$ defined by*

$$f_\infty(v)=\frac{1}{L^d}\int_{\mathbb{T}^d_L} f_i(x,v)\,dx;$$

and $\varrho(x)=\int_{\mathbb{R}^d} f(x,v)\,dv$ converges strongly to the constant

$$\varrho_\infty=\frac{1}{L^d}\int_{\mathbb{T}^d_L}\int_{\mathbb{R}^d} f_i(x,v)\,dv\,dx.$$

More precisely, for any $\lambda'<\min\{\lambda,\alpha\}$,

$$\begin{cases}\|\varrho(t,\cdot)-\varrho_\infty\|_{C^r}=O(e^{-2\pi\lambda'|t|/L}) & \text{for all } r\in\mathbb{N},\\ |\tilde{f}^{(L)}(t,k,\eta)-\tilde{f}^{(L)}_\infty(k,\eta)|=O(e^{-2\pi\lambda'|kt|/L}) & \text{for all } (k,\eta)\in\mathbb{Z}^d\times\mathbb{Z}^d.\end{cases}$$

Remark 3.2. Even if the initial datum is more regular than analytic, the convergence will in general not be better than exponential (except in some exceptional cases [38]). See [10, pp. 414–416] for an illustration. Conversely, if the analyticity width α for the initial datum is smaller than the "Landau rate" λ, then the rate of decay will not be better than $O(e^{-\alpha t})$. See [7] and [19] for a discussion of this fact, often overlooked in the physical literature.

Remark 3.3. The fact that the convergence is to the average of the initial datum will not survive non-linear perturbation, as shown by the counterexamples in §14.

Remark 3.4. Dimension does not play any role in the linear analysis. This can be attributed to the fact that only longitudinal waves occur, so everything happens "in the direction of the wave vector". Transversal waves arise in plasma physics only when magnetic effects are taken into account [1, Chapter 5].

Remark 3.5. The proof can be adapted to the case when f^0 and f_i are only C^∞; then the convergence is not exponential, but still $O(t^{-\infty})$. The regularity can also be further decreased, down to $W^{s,1}$, at least for any $s>2$; more precisely, if $f^0\in W^{s_0,1}$ and $f_i\in W^{s_i,1}$, there will be damping with a rate $O(t^{-\varkappa})$ for any $\varkappa<\max\{s_0-2,s_i\}$. (Compare with [1, Volume 1, p. 189].) This is independent of the regularity of the interaction.

The proof of Theorem 3.1 relies on the following elementary estimate for Volterra equations. We use the notation of §2.2.

LEMMA 3.6. *Assume that* **(L)** *holds true for some constants* $C_0, \varkappa, \lambda > 0$. *Let*

$$C_W = \|W\|_{L^1(\mathbb{T}^d_L)}$$

and let K^0 *be defined by* (2.3). *Then any solution* $\varphi(t,k)$ *of*

$$\varphi(t,k) = a(t,k) + \int_0^t K^0(t-\tau,k)\varphi(\tau,k)\,d\tau \tag{3.4}$$

satisfies, for any $k \in \mathbb{Z}^d$ *and any* $\lambda' < \lambda$,

$$\sup_{t\geqslant 0} |\varphi(t,k)| e^{2\pi\lambda'|k|t/L} \leqslant (1 + C_0 C_W C(\lambda,\lambda',\varkappa)) \sup_{t\geqslant 0} |a(t,k)| e^{2\pi\lambda|k|t/L}.$$

Here $C(\lambda,\lambda',\varkappa) = C(1+\varkappa^{-1}(1+(\lambda-\lambda')^{-2}))$ *for some universal constant* C.

Remark 3.7. It is standard to solve these Volterra equations by Laplace transforms; but, with a view to the non-linear setting, we shall prefer a more flexible and quantitative approach.

Proof. If $k=0$ this is obvious since $K^0(t,0)=0$; so we assume $k \neq 0$. Consider $\lambda' < \lambda$, multiply (3.4) by $e^{2\pi\lambda'|k|t/L}$, and write

$$\Phi(t,k) = \varphi(t,k) e^{2\pi\lambda'|k|t/L} \quad \text{and} \quad A(t,k) = a(t,k) e^{2\pi\lambda'|k|t/L};$$

then (3.4) becomes

$$\Phi(t,k) = A(t,k) + \int_0^t K^0(t-\tau,k) e^{2\pi\lambda'|k|(t-\tau)/L} \Phi(\tau,k)\,d\tau. \tag{3.5}$$

A particular case. The proof is extremely simple if we make the stronger assumption

$$\int_0^\infty |K^0(\tau,k)| e^{2\pi\lambda'|k|\tau/L}\,d\tau \leqslant 1-\varkappa, \quad \varkappa \in (0,1).$$

Then from (3.5),

$$\sup_{0\leqslant t\leqslant T} |\Phi(t,k)| \leqslant \sup_{0\leqslant t\leqslant T} |A(t,k)| + \left(\sup_{0\leqslant t\leqslant T} \int_0^t |K^0(t-\tau,k)| e^{2\pi\lambda'|k|(t-\tau)/L}\,d\tau \right) \sup_{0\leqslant \tau\leqslant T} |\Phi(\tau,k)|,$$

whence

$$\sup_{0\leqslant\tau\leqslant t}|\Phi(\tau,k)|\leqslant\frac{\sup_{0\leqslant\tau\leqslant t}|A(\tau,k)|}{1-\int_0^\infty|K^0(\tau,k)|e^{2\pi\lambda'|k|\tau/L}\,d\tau}\leqslant\frac{\sup_{0\leqslant\tau\leqslant t}|A(\tau,k)|}{\varkappa},$$

and therefore

$$\sup_{t\geqslant0}|\varphi(t,k)|e^{2\pi\lambda'|k|t/L}\leqslant\frac{\sup_{t\geqslant0}|a(t,k)|e^{2\pi\lambda'|k|t/L}}{\varkappa}.$$

The general case. To treat the general case we take the Fourier transform in the time variable, after extending K, A and Φ by 0 at negative times. (This presentation was suggested to us by Sigal, and appears to be technically simpler than the use of the Laplace transform.) Denoting the Fourier transform with a hat and recalling (2.4), we have, for $\xi=\lambda'+i\omega L/|k|$,

$$\widehat{\Phi}(\omega,k)=\hat{A}(\omega,k)+\mathcal{L}(\xi,k)\,\widehat{\Phi}(\omega,k).$$

By assumption $\mathcal{L}(\xi,k)\neq1$, so

$$\widehat{\Phi}(\omega,k)=\frac{\hat{A}(\omega,k)}{1-\mathcal{L}(\xi,k)}.$$

From there, it is traditional to apply the Fourier (or Laplace) inversion transform. Instead, we apply Plancherel's identity to find (for each k)

$$\|\Phi\|_{L^2(dt)}\leqslant\frac{\|A\|_{L^2(dt)}}{\varkappa}.$$

We then plug this into the equation (3.5) to get

$$\begin{aligned}\|\Phi\|_{L^\infty(dt)}&\leqslant\|A\|_{L^\infty(dt)}+\|K^0e^{2\pi\lambda'|k|t/L}\|_{L^2(dt)}\,\|\Phi\|_{L^2(dt)}\\&\leqslant\|A\|_{L^\infty(dt)}+\frac{\|K^0e^{2\pi\lambda'|k|t/L}\|_{L^2(dt)}\,\|A\|_{L^2(dt)}}{\varkappa}.\end{aligned}\tag{3.6}$$

It remains to bound the second term. On the one hand,

$$\begin{aligned}\|A\|_{L^2(dt)}&=\left(\int_0^\infty|a(t,k)|^2e^{4\pi\lambda'|k|t/L}\,dt\right)^{1/2}\\&\leqslant\left(\int_0^\infty e^{-4\pi(\lambda-\lambda')|k|t/L}\,dt\right)^{1/2}\sup_{t\geqslant0}|a(t,k)|e^{2\pi\lambda|k|t/L}\\&=\left(\frac{L}{4\pi|k|(\lambda-\lambda')}\right)^{1/2}\sup_{t\geqslant0}|a(t,k)|e^{2\pi\lambda|k|t/L}.\end{aligned}\tag{3.7}$$

On the other hand,

$$\begin{aligned}\|K^0 e^{2\pi\lambda'|k|t/L}\|_{L^2(dt)} &= 4\pi^2|\widehat{W}^{(L)}(k)|\frac{|k|^2}{L^2}\left(\int_0^\infty e^{4\pi\lambda'|k|t/L}\left|\tilde{f}^0\left(\frac{kt}{L}\right)\right|^2 t^2\,dt\right)^{1/2}\\ &= 4\pi^2|\widehat{W}^{(L)}(k)|\frac{|k|^{1/2}}{L^{1/2}}\left(\int_0^\infty e^{4\pi\lambda' u}|\tilde{f}^0(\sigma u)|^2 u^2\,du\right)^{1/2},\end{aligned} \tag{3.8}$$

where $\sigma = k/|k|$ and $u = |k|t/L$. The estimate follows since

$$\int_0^\infty e^{-4\pi(\lambda-\lambda')u}u^2\,du = O((\lambda-\lambda')^{-3/2}).$$

(Note that the factor $|k|^{-1/2}$ in (3.7) cancels with $|k|^{1/2}$ in (3.8).)

It seems that we only used properties of the function $\mathcal{L}$ in a strip $\operatorname{Re}\xi \simeq \lambda$; but this is an illusion. Indeed, we have taken the Fourier transform of Φ without checking that it belongs to $(L^1+L^2)(dt)$, so what we have established is only an a-priori estimate. To convert it into a rigorous result, one can use a continuity argument after replacing λ' by a parameter α which varies from $-\varepsilon$ to λ'. (By the integrability of K^0 and Grönwall's lemma, φ is obviously bounded as a function of t; so $\varphi(k,t)e^{-\varepsilon|k|t/L}$ is integrable for any $\varepsilon>0$, and continuous as $\varepsilon\to 0$.) Then assumption **(L)** guarantees that our bounds are uniform in the strip $0\leqslant\operatorname{Re}\xi\leqslant\lambda'$, and the proof goes through. □

Proof of Theorem 3.1. Without loss of generality, we assume $t\geqslant 0$. Considering (3.3) as a perturbation of free transport, we apply Duhamel's formula to get

$$f(t,x,v) = f_i(x-vt,v) + \int_0^t [(\nabla W * \varrho)\cdot\nabla_v f^0](\tau, x-v(t-\tau), v)\,d\tau. \tag{3.9}$$

Integration in v yields

$$\varrho(t,x) = \int_{\mathbb{R}^d} f_i(x-vt,v)\,dv + \int_0^t\int_{\mathbb{R}^d}[(\nabla W*\varrho)\cdot\nabla_v f^0](\tau,x-v(t-\tau),v)\,dv\,d\tau. \tag{3.10}$$

Of course,

$$\int_{\mathbb{T}^d_L}\varrho(t,x)\,dx = \int_{\mathbb{T}^d_L}\int_{\mathbb{R}^d} f_i(x,v)\,dv\,dx.$$

For $k\neq 0$, taking the Fourier transform of (3.10), we obtain

$$\begin{aligned}
\hat\varrho^{(L)}(t,k) &= \int_{\mathbb{T}^d_L}\int_{\mathbb{R}^d} f_i(x-vt,v)e^{-2i\pi k\cdot x/L}\,dv\,dx \\
&\qquad + \int_0^t \int_{\mathbb{T}^d_L}\int_{\mathbb{R}^d} [(\nabla W*\varrho)\cdot\nabla_v f^0](\tau,x-v(t-\tau),v)e^{-2i\pi k\cdot x/L}\,dv\,dx\,d\tau \\
&= \int_{\mathbb{T}^d_L}\int_{\mathbb{R}^d} f_i(x,v)e^{-2i\pi k\cdot x/L}e^{-2i\pi k\cdot vt/L}\,dv\,dx \\
&\qquad + \int_0^t \int_{\mathbb{T}^d_L}\int_{\mathbb{R}^d} [(\nabla W*\varrho)\cdot\nabla_v f^0](\tau,x,v)e^{-2i\pi k\cdot x/L}e^{-2i\pi k\cdot v(t-\tau)/L}\,dv\,dx\,d\tau \\
&= \tilde f_i^{(L)}\left(k,\frac{kt}{L}\right) + \int_0^t (\widehat{\nabla W*\varrho})^{(L)}(\tau,k)\cdot\widetilde{\nabla_v f^0}\left(\frac{k(t-\tau)}{L}\right)d\tau \\
&= \tilde f_i^{(L)}\left(k,\frac{kt}{L}\right) + \int_0^t \left(2i\pi\frac{k}{L}\widehat{W}^{(L)}(k)\hat\varrho^{(L)}(\tau,k)\right)\cdot\left(2i\pi\frac{k(t-\tau)}{L}\tilde f^0\left(\frac{k(t-\tau)}{L}\right)\right)d\tau.
\end{aligned}$$

In conclusion, we have established the *closed equation* for $\hat\varrho^{(L)}$:

$$\hat\varrho^{(L)}(t,k) = \tilde f_i^{(L)}\left(k,\frac{kt}{L}\right) - 4\pi^2\widehat{W}^{(L)}(k)\int_0^t \hat\varrho^{(L)}(\tau,k)\tilde f^0\left(\frac{k(t-\tau)}{L}\right)\frac{|k|^2}{L^2}(t-\tau)\,d\tau. \quad (3.11)$$

Recalling (2.3), this is the same as

$$\hat\varrho^{(L)}(t,k) = \tilde f_i^{(L)}\left(k,\frac{kt}{L}\right) + \int_0^t K^0(t-\tau,k)\hat\varrho^{(L)}(\tau,k)\,d\tau.$$

Without loss of generality $\lambda\leqslant\alpha$, where α appears in Theorem 3.1. By assumption (**L**) and Lemma 3.6,

$$|\hat\varrho^{(L)}(t,k)| \leqslant C_0C_WC(\lambda,\lambda',\varkappa)C_i e^{-2\pi\lambda'|k|t/L}.$$

In particular, for $k\neq 0$ we have

$$|\hat\varrho^{(L)}(t,k)| = O(e^{-2\pi\lambda''t/L}e^{-2\pi(\lambda'-\lambda'')|k|/L}) \quad \text{for all } t\geqslant 1;$$

so any Sobolev norm of $\varrho-\varrho_\infty$ converges to zero like $O(e^{-2\pi\lambda''t/L})$, where λ'' is arbitrarily close to λ' and therefore also to λ. By Sobolev embedding, the same is true for any C^r norm.

Next, we go back to (3.9) and take the Fourier transform in both variables x and v,

to find

$$\begin{aligned}\tilde f^{(L)}(t,k,\eta) &= \int_{\mathbb{T}^d_L}\int_{\mathbb{R}^d} f_i(x-vt,v)e^{-2i\pi k\cdot x/L}e^{-2i\pi\eta\cdot v}\,dv\,dx\\ &\quad+\int_0^t\int_{\mathbb{T}^d_L}\int_{\mathbb{R}^d}(\nabla W*\varrho)(\tau,x-v(t-\tau))\cdot\nabla_v f^0(v)e^{-2i\pi k\cdot x/L}e^{-2i\pi\eta\cdot v}\,dv\,dx\,d\tau\\ &=\int_{\mathbb{T}^d_L}\int_{\mathbb{R}^d} f_i(x,v)e^{-2i\pi k\cdot x/L}e^{-2i\pi k\cdot vt/L}e^{-2i\pi\eta\cdot v}\,dv\,dx\\ &\quad+\int_0^t\int_{\mathbb{T}^d_L}\int_{\mathbb{R}^d}(\nabla W*\varrho)(\tau,x)\cdot\nabla_v f^0(v)\\ &\qquad\qquad\times e^{-2i\pi k\cdot x/L}e^{-2i\pi k\cdot v(t-\tau)/L}e^{-2i\pi\eta\cdot v}\,dv\,dx\,d\tau\\ &=\tilde f_i^{(L)}\biggl(k,\eta+\frac{kt}{L}\biggr)+\int_0^t\widehat{\nabla W}^{(L)}(k)\hat\varrho^{(L)}(\tau,k)\cdot\widetilde{\nabla_v f^0}\biggl(\eta+\frac{k(t-\tau)}{L}\biggr)\,d\tau.\end{aligned}$$

So

$$\tilde f^{(L)}\biggl(t,k,\eta-\frac{kt}{L}\biggr)=\tilde f_i^{(L)}(k,\eta)+\int_0^t\widehat{\nabla W}^{(L)}(k)\hat\varrho^{(L)}(\tau,k)\cdot\widetilde{\nabla_v f^0}\biggl(\eta-\frac{k\tau}{L}\biggr)\,d\tau. \tag{3.12}$$

In particular, for any $\eta\in\mathbb{R}^d$,

$$\tilde f^{(L)}(t,0,\eta)=\tilde f_i^{(L)}(0,\eta); \tag{3.13}$$

in other words, $\langle f\rangle=L^{-d}\int_{\mathbb{T}^d_L} f\,dx$ remains equal to $\langle f_i\rangle$ for all times.

On the other hand, if $k\neq0$, then

$$\begin{aligned}&\biggl|\tilde f^{(L)}\biggl(t,k,\eta-\frac{kt}{L}\biggr)\biggr|\\ &\leqslant|\tilde f_i^{(L)}(k,\eta)|+\int_0^t|\widehat{\nabla W}^{(L)}(k)|\,|\hat\varrho^{(L)}(\tau,k)|\,\biggl|\widetilde{\nabla_v f^0}\biggl(\eta-\frac{k\tau}{L}\biggr)\biggr|\,d\tau\\ &\leqslant C_ie^{-2\pi\alpha|\eta|}+\int_0^t C_W C(\lambda,\lambda',\varkappa)C_ie^{-2\pi\lambda'|k|\tau/L}\biggl(2\pi C_0\biggl|\eta-\frac{k\tau}{L}\biggr|e^{-2\pi\lambda|\eta-k\tau/L|}\biggr)\,d\tau\\ &\leqslant C\biggl(e^{-2\pi\alpha|\eta|}+\int_0^t e^{-2\pi\lambda'|k|\tau/L}e^{-\pi(\lambda'+\lambda)|\eta-k\tau/L|}\,d\tau\biggr),\end{aligned} \tag{3.14}$$

where we have used that $\lambda'<\frac12(\lambda'+\lambda)<\lambda$, and C only depends on C_W, C_i, λ, λ' and $\varkappa$.

In the end,

$$\begin{aligned}\int_0^t e^{-2\pi\lambda'|k|\tau/L}e^{-\pi(\lambda'+\lambda)|\eta-k\tau/L|}\,d\tau&\leqslant\int_0^t e^{-2\pi\lambda'|\eta|}e^{-\pi(\lambda-\lambda')|\eta-k\tau/L|}\,d\tau\\ &\leqslant\frac{L}{\pi(\lambda-\lambda')}e^{-2\pi[\lambda'-(\lambda-\lambda')/2]|\eta|}.\end{aligned}$$

Plugging this back into (3.14), we obtain, with $\lambda''=\lambda'-\frac{1}{2}(\lambda-\lambda')$,

$$\left|\tilde f^{(L)}\left(t,k,\eta-\frac{kt}{L}\right)\right|\leqslant Ce^{-2\pi\lambda''|\eta|}. \tag{3.15}$$

In particular, for any fixed η and $k\neq0$,

$$|\tilde f^{(L)}(t,k,\eta)|\leqslant Ce^{-2\pi\lambda''|\eta+kt/L|}=O(e^{-2\pi\lambda''|t|/L}).$$

We conclude that $\tilde f^{(L)}$ converges pointwise, exponentially fast, to the Fourier transform of $\langle f_i\rangle$. Since λ' and then λ'' can be taken as close to λ as wanted, this ends the proof. □

We close this section by proving Proposition 2.1.

Proof of Proposition 2.1. First assume (a). Since $\tilde f^0$ decreases exponentially fast, we can find $\lambda,\varkappa>0$ such that

$$4\pi^2\max|\widehat{W}^{(L)}(k)|\sup_{|\sigma|=1}\int_0^\infty|\tilde f^0(r\sigma)|re^{2\pi\lambda r}\,dr\leqslant1-\varkappa.$$

Performing the change of variables $kt/L=r\sigma$ inside the integral, we deduce that

$$\int_0^\infty 4\pi^2|\widehat{W}^{(L)}(k)|\left|\tilde f^0\left(\frac{kt}{L}\right)\right|\frac{|k|^2t}{L^2}e^{2\pi\lambda|k|t/L}\,dt\leqslant1-\varkappa,$$

and this obviously implies **(L)**.

The choice $w=0$ in (2.8) shows that condition (b) is a particular case of (c), so we only treat the latter assumption. The reasoning is more subtle than for case (a). Throughout the proof we shall abbreviate $\widehat{W}^{(L)}$ by $\widehat{W}$. As a start, let us assume $d=1$ and $k>0$ (so $k\in\mathbb{N}$). Then we compute: for any $\omega\in\mathbb{R}$,

$$\begin{aligned}
&\int_0^\infty e^{2i\pi\omega kt/L}K^0(t,k)\,dt\\
&\quad=\lim_{\lambda\to0^+}\int_0^\infty e^{-2\pi\lambda kt/L}e^{2i\pi\omega kt/L}\,K^0(t,k)\,dt\\
&\quad=-4\pi^2\widehat{W}(k)\lim_{\lambda\to0^+}\int_0^\infty\int_{\mathbb{R}}f^0(v)e^{-2i\pi kvt/L}e^{-2\pi\lambda kt/L}e^{2i\pi\omega kt/L}\frac{k^2}{L^2}t\,dv\,dt\\
&\quad=-4\pi^2\widehat{W}(k)\lim_{\lambda\to0^+}\int_0^\infty\int_{\mathbb{R}}f^0(v)e^{-2i\pi vt}e^{-2\pi\lambda t}e^{2i\pi\omega t}t\,dv\,dt.
\end{aligned}\tag{3.16}$$

Then by integration by parts, assuming that $(f^0)'$ is integrable,

$$\int_{\mathbb{R}}f^0(v)e^{-2i\pi vt}t\,dv=\frac{1}{2i\pi}\int_{\mathbb{R}}(f^0)'(v)e^{-2i\pi vt}\,dv.$$

Plugging this back into (3.16), we obtain the expression

$$2i\pi\widehat{W}(k)\lim_{\lambda\to 0^+}\int_{\mathbb{R}}(f^0)'(v)\int_0^\infty e^{-2\pi[\lambda+i(v-\omega)]t}\,dt\,dv.$$

Next, recall that for any $\lambda>0$,

$$\int_0^\infty e^{-2\pi[\lambda+i(v-\omega)]t}\,dt=\frac{1}{2\pi[\lambda+i(v-\omega)]};$$

indeed, both sides are holomorphic functions of $z=\lambda+i(v-\omega)$ in the half-plane $\{z\in\mathbb{C}:\operatorname{Re}z>0\}$, and they coincide on the real half-axis $\{z\in\mathbb{R}:z>0\}$, so they have to coincide everywhere. We conclude that

$$\int_0^\infty e^{2i\pi\omega kt/L}K^0(t,k)\,dt=\widehat{W}(k)\lim_{\lambda\to 0^+}\int_{\mathbb{R}}\frac{(f^0)'(v)}{v-\omega-i\lambda}\,dv. \tag{3.17}$$

The celebrated *Plemelj formula* states that

$$\frac{1}{z-i0}=\text{p.v.}\left(\frac{1}{z}\right)+i\pi\delta_0, \tag{3.18}$$

where the left-hand side should be understood as the limit, in weak sense, of $1/(z-i\lambda)$ as $\lambda\to 0^+$. The abbreviation p.v. stands for principal value, that is, simplifying the possibly divergent part by using compensations by symmetry when the denominator vanishes. Formula (3.18) is proven in Appendix A.5, where the notion of principal value is also precisely defined. Combining (3.17) and (3.18), we end up with the identity

$$\int_0^\infty e^{2i\pi\omega kt/L}K^0(t,k)\,dt=\widehat{W}(k)\left[\left(\text{p.v.}\int_{\mathbb{R}}\frac{(f^0)'(v)}{v-\omega}\,dv\right)+i\pi(f^0)'(\omega)\right]. \tag{3.19}$$

Since W is even, $\widehat{W}$ is real-valued, so the above formula yields the decomposition of the limit into real and imaginary parts. The problem is to check that the real part cannot approach 1 at the same time as the imaginary part approaches 0.

As soon as $(f^0)'(v)=O(1/|v|)$, we have

$$\int_{\mathbb{R}}\frac{(f^0)'(v)}{v-\omega}\,dv=O\left(\frac{1}{|\omega|}\right)\quad\text{as }|\omega|\to\infty,$$

so the real part in the right-hand side of (3.19) becomes small when $|\omega|$ is large, and we can restrict to a bounded interval $|\omega|\leqslant\Omega$.

Then the imaginary part, $\widehat{W}(k)\pi(f^0)'(\omega)$, can become small only in the limit $k\to\infty$ (but then also the real part becomes small) or if ω approaches one of the zeroes of $(f^0)'$. Since ω varies in a compact set, by continuity it will be sufficient to check the condition

only at the zeroes of $(f^0)'$. In the end, we have obtained the following stability criterion: for any $k\in\mathbb{N}$,

$$(f^0)'(\omega)=0 \implies \widehat{W}(k)\int_{\mathbb{R}}\frac{(f^0)'(v)}{v-\omega}\,dv\neq 1 \quad \text{for all } \omega\in\mathbb{R}. \tag{3.20}$$

Now, if $k<0$, we can restart the computation as follows:

$$\int_0^\infty e^{2i\pi\omega|k|t/L}K^0(t,k)\,dt$$
$$=-4\pi^2\widehat{W}(k)\lim_{\lambda\to 0^+}\int_0^\infty\int_{\mathbb{R}} f^0(v)e^{-2i\pi kvt/L}e^{-2\pi\lambda|k|t/L}e^{2i\pi\omega|k|t/L}\frac{|k|^2}{L^2}t\,dv\,dt.$$

Then the change of variable $v\mapsto -v$ brings us back to the previous computation with k replaced by $|k|$ (except in the argument of W) and $f^0(v)$ replaced by $f^0(-v)$. However, it is immediately checked that (3.20) is *invariant under reversal of velocities*, that is, if $f^0(v)$ is replaced by $f^0(-v)$.

Finally, let us generalize this to several dimensions. If $k\in\mathbb{Z}^d\setminus\{0\}$ and $\xi\in\mathbb{C}$, we can use the splitting

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

and Fubini's theorem to rewrite

$$\mathcal{L}(\xi,k)$$
$$=-4\pi^2\widehat{W}(k)\frac{|k|^2}{L^2}\int_0^\infty\int_{\mathbb{R}^d} f^0(v)e^{-2i\pi kt\cdot v/L}te^{2\pi|k|\xi^*t/L}\,dv\,dt$$
$$=-4\pi^2\widehat{W}(k)\frac{|k|^2}{L^2}\int_0^\infty\int_{\mathbb{R}}\left(\int_{kr/|k|+k^\perp} f^0\left(\frac{k}{|k|}r+w\right)dw\right)e^{-2i\pi|k|rt/L}te^{2\pi|k|\xi^*t/L}\,dr\,dt,$$

where $k^\perp$ is the hyperplane orthogonal to k. So everything is expressed in terms of the 1-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

$$f_\sigma(v)=\int_{v\sigma+\sigma^\perp} f(w)\,dw \quad \text{for all } v\in\mathbb{R}. \tag{3.21}$$

Then the computation above shows that the multi-dimensional stability criterion reduces to the 1-dimensional criterion in each direction $k/|k|$, and the claim is proven. □

4. Analytic norms

In this section we introduce some functional spaces of analytic functions on the spaces $\mathbb{R}^d$, $\mathbb{T}^d=\mathbb{R}^d/\mathbb{Z}^d$ and most importantly $\mathbb{T}^d\times\mathbb{R}^d$. (Changing the sidelength of the torus

only results in some changes in the constants.) Then we establish a number of functional inequalities which will be crucial in the subsequent analysis. At the end of this section we shall reformulate the linear study in this new setting.

Throughout the whole section, d is a positive integer. Working with analytic functions will force us to be careful with combinatorial issues, and proofs will at times involve summation over many indices.

4.1. Single-variable analytic norms

Here "single-variable" means that the variable lives in either $\mathbb{R}^d$ or $\mathbb{T}^d$, but d may be greater than 1.

Among many possible families of norms for analytic functions, two will be of particular interest for us; they will be denoted by $\mathcal{C}^{\lambda;p}$ and $\mathcal{F}^{\lambda;p}$. The $\mathcal{C}^{\lambda;p}$ norms are defined for functions on $\mathbb{R}^d$ or $\mathbb{T}^d$, while the $\mathcal{F}^{\lambda;p}$ norms are defined only for $\mathbb{T}^d$ (although we could easily cook up a variant in $\mathbb{R}^d$). We shall write $\mathbb{N}_0^d$ for the set of d-tuples of integers (the subscript being here to insist that 0 is allowed). If $n\in\mathbb{N}_0^d$ and $\lambda\geqslant 0$ we shall write $\lambda^n=\lambda^{|n|}$. Conventions about the Fourier transform and multi-dimensional differential calculus are gathered in the appendix.

Definition 4.1. (One-variable analytic norms) For any $p\in[1,\infty]$ and $\lambda\geqslant 0$, we define

$$\|f\|_{\mathcal{C}^{\lambda;p}} := \sum_{n\in\mathbb{N}_0^d} \frac{\lambda^n}{n!} \|f^{(n)}\|_{L^p} \quad \text{and} \quad \|f\|_{\mathcal{F}^{\lambda;p}} := \Big(\sum_{k\in\mathbb{Z}^d} e^{2\pi\lambda p|k|} |\hat{f}(k)|^p \Big)^{1/p}, \tag{4.1}$$

the latter expression standing for $\sup_{k\in\mathbb{Z}^d} |\hat{f}(k)| e^{2\pi\lambda|k|}$ if $p=\infty$. We further write

$$\mathcal{C}^\lambda := \mathcal{C}^{\lambda;\infty} \quad \text{and} \quad \mathcal{F}^\lambda := \mathcal{F}^{\lambda;1}. \tag{4.2}$$

Remark 4.2. The parameter λ can be interpreted as a radius of convergence.

Remark 4.3. The norms $\mathcal{C}^\lambda$ and $\mathcal{F}^\lambda$ are of particular interest because they are algebra norms.

We shall sometimes abbreviate $\|\cdot\|_{\mathcal{C}^{\lambda;p}}$ or $\|\cdot\|_{\mathcal{F}^{\lambda;p}}$ into $\|\cdot\|_{\lambda;p}$ when no confusion is possible, or when the statement works for either norm.

The norms in (4.1) extend to vector-valued functions in a natural way: if f is valued in $\mathbb{R}^d$ or $\mathbb{T}^d$ or $\mathbb{Z}^d$, define $f^{(n)}=(f_1^{(n)},...,f_d^{(n)})$ and $\hat{f}(k)=(\hat{f}_1(k),...,\hat{f}_d(k))$; then the formulae in (4.1) make sense provided that we choose a norm on $\mathbb{R}^d$ or $\mathbb{T}^d$ or $\mathbb{Z}^d$. Which norm we choose will depend on the context; the choice will always be done in such a way to get the duality right in the inequality $|a\cdot b|\leqslant\|a\|\,\|b\|_*$. For instance if f is valued

in $\mathbb{Z}^d$ and g in $\mathbb{T}^d$, and we have to estimate $f\cdot g$, we may norm $\mathbb{Z}^d$ by $|k|=\sum_{j=1}^d |k_j|$ and $\mathbb{T}^d$ by $|x|=\sup_j |x_j|$.([15]) This will not pose any problem, and the reader can forget about this issue; we shall just make remarks about it whenever needed. For the rest of this section, we shall focus on scalar-valued functions for simplicity of exposition.

Next, we define "homogeneous" analytic seminorms by removing the zeroth-order term. We write $\mathbb{N}_*^d=\mathbb{N}_0^d\setminus\{0\}$ and $\mathbb{Z}_*^d=\mathbb{Z}^d\setminus\{0\}$.

Definition 4.4. (One-variable homogeneous analytic seminorms) For $p\in[1,\infty]$ and $\lambda\geqslant 0$ we write

$$\|f\|_{\dot{\mathcal{C}}^{\lambda;p}} = \sum_{n\in\mathbb{N}_*^d} \frac{\lambda^n}{n!}\|f^{(n)}\|_{L^p} \quad \text{and} \quad \|f\|_{\dot{\mathcal{F}}^{\lambda;p}} = \Big(\sum_{k\in\mathbb{Z}_*^d} e^{2\pi\lambda p|k|}|\hat{f}(k)|^p\Big)^{1/p}.$$

It is interesting to note that affine functions $x\mapsto a\cdot x+b$ can be included in $\dot{\mathcal{C}}^\lambda=\dot{\mathcal{C}}^{\lambda;\infty}$, even though they are unbounded; in particular $\|a\cdot x+b\|_{\dot{\mathcal{C}}^\lambda}=\lambda|a|$. On the other hand, linear forms $x\mapsto a\cdot x$ do not naturally belong to $\dot{\mathcal{F}}^\lambda$, because their Fourier expansion is not even summable (it decays like $1/k$).

The spaces $\mathcal{C}^{\lambda;p}$ and $\mathcal{F}^{\lambda;p}$ enjoy remarkable properties, summarized in Propositions 4.5, 4.8 and 4.10 below. Some of these properties are well known, others not so.

Proposition 4.5. (Algebra property) (i) *For any $\lambda\geqslant 0$ and $p,q,r\in[1,\infty]$ such that $1/p+1/q=1/r$, we have*

$$\|fg\|_{\mathcal{C}^{\lambda;r}} \leqslant \|f\|_{\mathcal{C}^{\lambda;p}}\|g\|_{\mathcal{C}^{\lambda;q}}.$$

(ii) *For any $\lambda\geqslant 0$ and $p,q,r\in[1,\infty]$ such that $1/p+1/q=1/r+1$, we have*

$$\|fg\|_{\mathcal{F}^{\lambda;r}} \leqslant \|f\|_{\mathcal{F}^{\lambda;p}}\|g\|_{\mathcal{F}^{\lambda;q}}.$$

(iii) *As a consequence, for any $\lambda\geqslant 0$, $\mathcal{C}^\lambda=\mathcal{C}^{\lambda;\infty}$ and $\mathcal{F}^\lambda=\mathcal{F}^{\lambda;1}$ are normed algebras: for either space,*

$$\|fg\|_\lambda \leqslant \|f\|_\lambda\|g\|_\lambda.$$

In particular, $\|f^n\|_\lambda\leqslant\|f\|_\lambda^n$ for any $n\in\mathbb{N}_0$, and $\|e^f\|_\lambda\leqslant e^{\|f\|_\lambda}$.

Remark 4.6. Ultimately, property (iii) relies on the fact that L^∞ and L^1 are normed algebras for the multiplication and convolution, respectively.

Remark 4.7. It follows from the Fourier inversion formula and Proposition 4.5 that $\|f\|_{\mathcal{C}^\lambda}\leqslant\|f\|_{\mathcal{F}^\lambda}$ (and $\|f\|_{\dot{\mathcal{C}}^\lambda}\leqslant\|f\|_{\dot{\mathcal{F}}^\lambda}$); this is a special case of Proposition 4.8 (iv) below. The reverse inequality does not hold, because $\|f\|_\infty$ does not control $\|\hat{f}\|_{L^1}$.

([15]) Of course all norms are equivalent, still the choice is not innocent when the estimates are iterated infinitely many times; an advantage of the supremum norm on $\mathbb{R}^d$ is that it has the algebra property.

Analytic norms are very sensitive to composition; think that if $a>0$ then

$$\|f\circ(a\,\mathrm{Id})\|_{\mathcal{C}^{\lambda;p}} = a^{-d/p}\|f\|_{\mathcal{C}^{a\lambda;p}};$$

so we typically lose on the functional space. This is a major difference with more traditional norms used in partial differential equations theory, such as Hölder or Sobolev norms, for which composition may affect constants but not regularity indices. The next proposition controls the loss of regularity implied by composition.

PROPOSITION 4.8. (Composition inequality) (i) *For any $\lambda>0$ and any $p\in[1,\infty]$,*

$$\|f\circ H\|_{\mathcal{C}^{\lambda;p}} \leqslant \|(\det\nabla H)^{-1}\|_\infty^{1/p}\|f\|_{\mathcal{C}^{\nu;p}}, \quad \nu = \|H\|_{\dot{\mathcal{C}}^\lambda},$$

where H is possibly unbounded.

(ii) *For any $\lambda>0$, any $p\in[1,\infty]$ and any $a>0$,*

$$\|f\circ(a\,\mathrm{Id}+G)\|_{\mathcal{C}^{\lambda;p}} \leqslant a^{-d/p}\|f\|_{\mathcal{C}^{a\lambda+\nu;p}}, \quad \nu = \|G\|_{\mathcal{C}^\lambda}.$$

(iii) *For any $\lambda>0$,*

$$\|f\circ(\mathrm{Id}+G)\|_{\mathcal{F}^\lambda} \leqslant \|f\|_{\mathcal{F}^{\lambda+\nu}}, \quad \nu = \|G\|_{\dot{\mathcal{F}}^\lambda}.$$

(iv) *For any $\lambda>0$ and any $a>0$,*

$$\|f\circ(a\,\mathrm{Id}+G)\|_{\mathcal{C}^\lambda} \leqslant \|f\|_{\mathcal{F}^{a\lambda+\nu}}, \quad \nu = \|G\|_{\dot{\mathcal{C}}^\lambda}.$$

Remark 4.9. Inequality (iv), with $\mathcal{C}$ on the left and $\mathcal{F}$ on the right, will be most useful. The reverse inequality is not likely to hold, in view of Remark 4.7.

The last property of interest for us is the control of the loss of regularity involved by differentiation.

PROPOSITION 4.10. (Control of gradients) *For any $\bar\lambda>\lambda$ and any $p\in[1,\infty]$,*

$$\|\nabla f\|_{\mathcal{C}^{\lambda;p}} \leqslant \frac{1}{\lambda e\log(\bar\lambda/\lambda)}\|f\|_{\dot{\mathcal{C}}^{\bar\lambda;p}}, \tag{4.3}$$

$$\|\nabla f\|_{\mathcal{F}^{\lambda;p}} \leqslant \frac{1}{2\pi e(\bar\lambda-\lambda)}\|f\|_{\dot{\mathcal{F}}^{\bar\lambda;p}}. \tag{4.4}$$

The proofs of Propositions 4.5–4.10 will be preparations for the more complicated situations considered in the sequel.

Proof of Proposition 4.5. (i) Denoting the norm of $\mathcal{C}^{\lambda;p}$ by $\|\cdot\|_{\lambda;p}$ and using the multi-dimensional Leibniz formula from Appendix A.2, we have

$$\begin{aligned}
\|fg\|_{\lambda;r} &= \sum_{l\in\mathbb{N}_0^d} \|(fg)^{(l)}\|_{L^r} \frac{\lambda^l}{l!} \\
&\leqslant \sum_{\substack{l,m\in\mathbb{N}_0^d \\ m\leqslant l}} \binom{l}{m} \|f^{(m)} g^{(l-m)}\|_{L^r} \frac{\lambda^l}{l!} \\
&\leqslant \sum_{\substack{l,m\in\mathbb{N}_0^d \\ m\leqslant l}} \binom{l}{m} \|f^{(m)}\|_{L^p} \|g^{(l-m)}\|_{L^q} \frac{\lambda^l}{l!} \\
&= \sum_{\substack{l,m\in\mathbb{N}_0^d \\ m\leqslant l}} \frac{\|f^{(m)}\|_{L^p}\lambda^m}{m!} \frac{\|g^{(l-m)}\|_{L^q}\lambda^{l-m}}{(l-m)!} \\
&= \|f\|_{\lambda;p} \|g\|_{\lambda;q}.
\end{aligned}$$

(ii) Denoting now the norm of $\mathcal{F}^{\lambda;p}$ by $\|\cdot\|_{\lambda;p}$ and applying Young's convolution inequality, we get

$$\begin{aligned}
\|fg\|_{\lambda;r} &= \Big(\sum_{k\in\mathbb{Z}^d} |\widehat{fg}(k)|^r e^{2\pi\lambda r|k|} \Big)^{1/r} \\
&\leqslant \Big(\sum_{k\in\mathbb{Z}^d} \Big(\sum_{l\in\mathbb{Z}^d} |\hat{f}(l)|\, |\hat{g}(k-l)| e^{2\pi\lambda|k-l|} e^{2\pi\lambda|l|} \Big)^r \Big)^{1/r} \\
&\leqslant \Big(\sum_{k\in\mathbb{Z}^d} |\hat{f}(k)|^p e^{2\pi\lambda p|k-l|} \Big)^{1/p} \Big(\sum_{l\in\mathbb{Z}^d} |\hat{g}(l)|^q e^{2\pi\lambda q|l|} \Big)^{1/q}. \qquad \square
\end{aligned}$$

Proof of Proposition 4.8. (i) We use the (multi-dimensional) Faà di Bruno formula:

$$(f\circ H)^{(n)} = \sum_{\sum_{j=1}^n j m_j = n} \frac{n!}{m_1!\dots m_n!} (f^{(m_1+\dots+m_n)}\circ H) \prod_{j=1}^n \Big(\frac{H^{(j)}}{j!} \Big)^{m_j};$$

so

$$\|(f\circ H)^{(n)}\|_{L^p} \leqslant \sum_{\sum_{j=1}^n j m_j = n} \frac{n!}{m_1!\dots m_n!} \|f^{(m_1+\dots+m_n)}\circ H\|_{L^p} \prod_{j=1}^n \Big\| \frac{H^{(j)}}{j!} \Big\|_\infty^{m_j};$$

and thus

$$\begin{aligned}\sum_{n\geqslant 1}\frac{\lambda^n}{n!}&\|(f\circ H)^{(n)}\|_{L^p}\\ &\leqslant \|(\det\nabla H)^{-1}\|_\infty^{1/p}\sum_{k=1}^{\infty}\|f^{(k)}\|_{L^p}\sum_{\substack{\sum_{j=1}^n jm_j=n\\ \sum_{j=1}^n m_j=k}}\frac{\lambda^n}{m_1!\dots m_n!}\prod_{j=1}^n\left\|\frac{H^{(j)}}{j!}\right\|_\infty^{m_j}\\ &= \|(\det\nabla H)^{-1}\|_\infty^{1/p}\sum_{k\geqslant 1}\|f^{(k)}\|_{L^p}\frac{1}{k!}\Big(\sum_{|l|\geqslant 1}\frac{\lambda^l}{l!}\|H^{(l)}\|_\infty\Big)^k,\end{aligned}$$

where the last step follows from the multi-dimensional binomial formula.

(ii) We decompose $h(x):=f(ax+G(x))$ as

$$h(x)=\sum_{n\in\mathbb{N}_0^d}\frac{f^{(n)}(ax)}{n!}G(x)^n$$

and we apply ∇^k:

$$\nabla^k h(x)=\sum_{\substack{k_1,k_2\in\mathbb{N}_0^d\\ k_1+k_2=k}}\sum_{n\in\mathbb{N}_0^d}\frac{k!a^{k_1}}{k_1!k_2!n!}\nabla^{k_1+n}f(ax)\nabla^{k_2}G^n(x).$$

Then we take the L^p norm, multiply by $\lambda^k/k!$ and sum over k:

$$\begin{aligned}\|h\|_{\mathcal{C}^{\lambda;p}}&\leqslant|a|^{-d/p}\sum_{k_1,k_2,n\geqslant 0}\frac{\lambda^{k_1+k_2}|a|^{k_1}}{k_1!k_2!n!}\|\nabla^{k_1+n}f\|_{L^p}\|\nabla^{k_2}G^n\|_\infty\\ &=|a|^{-d/p}\sum_{k_1,n\geqslant 0}\frac{\lambda^{k_1}|a|^{k_1}}{k_1!n!}\|\nabla^{k_1+n}f\|_{L^p}\|G^n\|_{\mathcal{C}^\lambda}\\ &\leqslant|a|^{-d/p}\sum_{k_1,n\geqslant 0}\frac{\lambda^{k_1}|a|^{k_1}}{k_1!n!}\|\nabla^{k_1+n}f\|_{L^p}\|G\|_{\mathcal{C}^\lambda}^n\\ &=|a|^{-d/p}\sum_{m\geqslant 0}\frac{(a\lambda+\|G\|_{\mathcal{C}^\lambda})^m}{m!}\|\nabla^m f\|_{L^p},\end{aligned}$$

where Proposition 4.5 (iii) was used in the second-last step.

(iii) In this case we write, with $G_0=\widehat{G}(0)$,

$$h(x)=f(x+G(x))=\sum_{k\in\mathbb{Z}^d}\hat{f}(k)e^{2i\pi k\cdot x}e^{2i\pi k\cdot G_0}e^{2i\pi k\cdot(G(x)-G_0)};$$

so

$$\hat{h}(l)=\sum_{k\in\mathbb{Z}^d}\hat{f}(k)e^{2i\pi k\cdot G_0}[e^{2i\pi k\cdot(G-G_0)}]\hat{}\,(l-k).$$

Then, using again Proposition 4.5,

$$\begin{aligned}\sum_{l\in\mathbb{Z}^d}|\hat{h}(l)|e^{2\pi\lambda|l|} &\leqslant \sum_{l,k\in\mathbb{Z}_d}|\hat{f}(k)|e^{2\pi\lambda|k|}e^{2\pi\lambda|l-k|}|[e^{2i\pi k\cdot(G-G_0)}]\hat{}\,(l-k)|\\ &=\sum_{k\in\mathbb{Z}^d}|\hat{f}(k)|e^{2\pi\lambda|k|}\|e^{2i\pi k\cdot(G-G_0)}\|_\lambda\\ &\leqslant\sum_{k\in\mathbb{Z}_d}|\hat{f}(k)|e^{2\pi\lambda|k|}e^{\|2\pi k\cdot(G-G_0)\|_\lambda}\\ &\leqslant\sum_{k\in\mathbb{Z}^d}|\hat{f}(k)|e^{2\pi\lambda|k|}e^{2\pi|k|\|G-G_0\|_\lambda}\\ &=\|f\|_{\lambda+\|G-G_0\|_\lambda}\\ &=\|f\|_{\lambda+\nu},\end{aligned}$$

where $\nu=\|G\|_{\dot{\mathcal{F}}^\lambda}$.

(iv) We actually have the more precise result

$$\|f\circ H\|_{\mathcal{C}^\lambda}\leqslant\sum_{k\in\mathbb{Z}^d}|\hat{f}(k)|e^{2\pi|k|\,\|H\|_{\dot{\mathcal{C}}^\lambda}}. \tag{4.5}$$

Writing $f\circ H=\sum_{k\in\mathbb{Z}^d}\hat{f}(k)e^{2i\pi k\cdot H}$, we see that (4.5) follows from

$$\|e^{ih}\|_{\mathcal{C}^\lambda}\leqslant e^{\|h\|_{\dot{\mathcal{C}}^\lambda}}. \tag{4.6}$$

To prove (4.6), let P_n be the polynomial in the variables X_m, $m\leqslant n$, defined by the identity

$$(e^f)^{(n)}=P_n((f^{(m)})_{m\leqslant n})e^f;$$

this polynomial (which can be made more explicit from the Faà di Bruno formula) has non-negative coefficients, so $\|(e^{if})^{(n)}\|_\infty\leqslant P_n((\|f^{(m)}\|)_{m\leqslant n})$. The conclusion will follow from the identity (between formal series!)

$$1+\sum_{n\in\mathbb{N}_*^d}\frac{\lambda^n}{n!}P_n((X_m)_{m\leqslant n})=\exp\biggl(\sum_{k\in\mathbb{N}_*^d}\frac{\lambda^k}{k!}X_k\biggr). \tag{4.7}$$

To prove (4.7), it is sufficient to note that the left-hand side is the expansion of e^g in powers of λ at 0, where

$$g(\lambda)=\sum_{k\in\mathbb{N}_*^d}\frac{\lambda^k}{k!}X_k. \qquad \square$$

Proof of Proposition 4.10. (a) Writing $\|\cdot\|_{\lambda;p}=\|\cdot\|_{\mathcal{C}^{\lambda;p}}$, we have

$$\|\partial_j f\|_{\lambda;p}=\sum_{n\in\mathbb{N}_0^d}\frac{\lambda^n}{n!}\|\partial_x^n\partial_j f\|_{L^p},$$

where $\partial_j=\partial/\partial x_j$. If 1_j is the d-tuple of integers with 1 in position j and 0 elsewhere, then $(n+1_j)!\leqslant(|n|+1)n!$, so

$$\|\partial_j f\|_{\lambda;p}\leqslant\sup_{n\in\mathbb{N}_0^d}\frac{(|n|+1)\lambda^n}{\bar\lambda^{n+1}}\sum_{|m|\geqslant 1}\frac{\bar\lambda^m}{m!}\|\nabla^m f\|_{L^p},$$

and the proof of (4.3) follows easily.

(b) Writing $\|\cdot\|_{\lambda;p}=\|\cdot\|_{\mathcal{F}^{\lambda;p}}$, we have

$$\begin{aligned}\|\partial_j f\|_{\lambda;p}&=\Big(\sum_{k\in\mathbb{Z}^d}|k_j|^p|\hat f(k)|^p e^{2\pi\lambda p|k|}\Big)^{1/p}\\&\leqslant\Big(\sup_{k\in\mathbb{Z}^d}|k|e^{2\pi(\lambda-\bar\lambda)|k|}\Big)\Big(\sum_{k\in\mathbb{Z}^d}|\hat f(k)|^p e^{2\pi\bar\lambda p|k|}\Big)^{1/p},\end{aligned}$$

and (4.4) follows. □

4.2. Analytic norms in two variables

To estimate solutions and trajectories of kinetic equations we will work on the phase space $\mathbb{T}^d_x\times\mathbb{R}^d_v$, and use three parameters: λ (*gliding* analytic regularity), μ (analytic regularity in x) and τ (time-shift along the free transport semigroup). The regularity quantified by λ is said to be gliding because for $\tau=0$ this is an analytic regularity in v, but as τ grows the regularity is progressively transferred from velocity to spatial modes, according to the evolution by free transport. This catch is crucial to our analysis: indeed, the solution of a transport equation like free transport or Vlasov *cannot* be uniformly analytic([16]) in v as time goes by—except of course if it is spatially homogeneous. Instead, the best we can do is compare the solution at time τ to the solution of free transport at the same time.

The parameters λ and μ will be non-negative; τ will vary in $\mathbb{R}$, but often be restricted to $\mathbb{R}_+$, just because we shall work in positive time. When τ is not specified, this means $\tau=0$. Sometimes we shall abuse notation by writing $\|f(x,v)\|$ instead of $\|f\|$, to stress the dependence of f on the two variables.

Putting aside the time-shift for a moment, we may generalize the norms $\mathcal{C}^\lambda$ and $\mathcal{F}^\lambda$ in an obvious way.

([16]) By this we mean of course that some *norm* or seminorm quantifying the degree of analytic smoothness in v will remain uniformly bounded.

Definition 4.11. (Two-variables analytic norms) For any $\lambda, \mu \geqslant 0$, we define

$$\|f\|_{\mathcal{C}^{\lambda,\mu}} = \sum_{m\in\mathbb{N}_0^d} \sum_{n\in\mathbb{N}_0^d} \frac{\lambda^n}{n!}\frac{\mu^m}{m!} \|\nabla_x^m \nabla_v^n f\|_{L^\infty(\mathbb{T}^d_x\times\mathbb{R}^d_v)}, \tag{4.8}$$

$$\|f\|_{\mathcal{F}^{\lambda,\mu}} = \sum_{k\in\mathbb{Z}^d} \int_{\mathbb{R}^d} |\tilde{f}(k,\eta)| e^{2\pi\lambda|\eta|} e^{2\pi\mu|k|}\, d\eta. \tag{4.9}$$

Of course one might also introduce variants based on L^p or ℓ^p norms (with two additional parameters p and q, since one can make different choices for the space and velocity variables).

The norm (4.9) is better adapted to the periodic nature of the problem, and is very well suited to estimate solutions of kinetic equations (with fast decay as $|v|\to\infty$); but in the sequel we shall also have to estimate characteristics (trajectories) which are unbounded functions of v. We could hope to play with two different families of norms, but this would entail considerable technical difficulties. Instead, we shall mix the two recipes to get the following *hybrid norms*.

Definition 4.12. (Hybrid analytic norms) For any $\lambda, \mu \geqslant 0$, let

$$\|f\|_{\mathcal{Z}^{\lambda,\mu}} = \sum_{l\in\mathbb{Z}^d} \sum_{n\in\mathbb{N}_0^d} \frac{\lambda^n}{n!} e^{2\pi\mu|l|} \|\widehat{\nabla_v^n f}(l,v)\|_{L^\infty(\mathbb{R}^d_v)}. \tag{4.10}$$

More generally, for any $p\in[1,\infty]$, we define

$$\|f\|_{\mathcal{Z}^{\lambda,\mu;p}} = \sum_{l\in\mathbb{Z}^d} \sum_{n\in\mathbb{N}_0^d} \frac{\lambda^n}{n!} e^{2\pi\mu|l|} \|\widehat{\nabla_v^n f}(l,v)\|_{L^p(\mathbb{R}^d_v)}. \tag{4.11}$$

Now let us introduce the time-shift τ. We denote by $(S^0_\tau)_{\tau\geqslant 0}$ the geodesic semigroup: $(S^0_\tau)(x,v) = (x+v\tau, v)$. Recall that the backward free transport semigroup is defined by $(f\circ S^0_\tau)_{\tau\geqslant 0}$, and the forward semigroup by $(f\circ S^0_{-\tau})_{\tau\geqslant 0}$.

Definition 4.13. (Time-shift pure and hybrid analytic norms)

$$\|f\|_{\mathcal{C}^{\lambda,\mu}_\tau} = \|f\circ S^0_\tau\|_{\mathcal{C}^{\lambda,\mu}} = \sum_{m\in\mathbb{N}_0^d} \sum_{n\in\mathbb{N}_0^d} \frac{\lambda^n}{n!}\frac{\mu^m}{m!} \|\nabla_x^m (\nabla_v + \tau\nabla_x)^n f\|_{L^\infty(\mathbb{T}^d_x\times\mathbb{R}^d_v)}, \tag{4.12}$$

$$\|f\|_{\mathcal{F}^{\lambda,\mu}_\tau} = \|f\circ S^0_\tau\|_{\mathcal{F}^{\lambda,\mu}} = \sum_{k\in\mathbb{Z}^d} \int_{\mathbb{R}^d} |\tilde{f}(k,\eta)| e^{2\pi\lambda|k\tau+\eta|} e^{2\pi\mu|k|}\, d\eta, \tag{4.13}$$

$$\|f\|_{\mathcal{Z}^{\lambda,\mu}_\tau} = \|f\circ S^0_\tau\|_{\mathcal{Z}^{\lambda,\mu}} = \sum_{l\in\mathbb{Z}^d} \sum_{n\in\mathbb{N}_0^d} \frac{\lambda^n}{n!} e^{2\pi\mu|l|} \|(\nabla_v + 2i\pi\tau l)^n \hat{f}(l,v)\|_{L^\infty(\mathbb{R}^d_v)}, \tag{4.14}$$

$$\|f\|_{\mathcal{Z}^{\lambda,\mu;p}_\tau} = \sum_{l\in\mathbb{Z}^d} \sum_{n\in\mathbb{N}_0^d} \frac{\lambda^n}{n!} e^{2\pi\mu|l|} \|(\nabla_v + 2i\pi\tau l)^n \hat{f}(l,v)\|_{L^p(\mathbb{R}^d_v)}. \tag{4.15}$$

This choice of norms is one of the cornerstones of our analysis: first, because of their hybrid nature, they will connect well to both periodic (in x) estimates on the force field, and uniform (in v) estimates on the "deflection maps" studied in §5. Secondly, they are well behaved with respect to the properties of free transport, allowing us to keep track of the initial time. Thirdly, they will satisfy the algebra property (for $p=\infty$), the composition inequality and the gradient inequality (for any $p\in[1,\infty]$). Before going on with the proof of these properties, we note the following alternative representations.

PROPOSITION 4.14. *The norm $\mathcal{Z}^{\lambda,\mu;p}_\tau$ admits the following alternative representations*:

$$\|f\|_{\mathcal{Z}^{\lambda,\mu;p}_\tau}=\sum_{l\in\mathbb{Z}^d}\sum_{n\in\mathbb{N}_0^d}\frac{\lambda^n}{n!}e^{2\pi\mu|l|}\|\nabla_v^n(\hat f(l,v)e^{2i\pi\tau l\cdot v})\|_{L^p(\mathbb{R}^d_v)},\tag{4.16}$$

$$\|f\|_{\mathcal{Z}^{\lambda,\mu;p}_\tau}=\sum_{n\in\mathbb{N}_0^d}\frac{\lambda^n}{n!}|\!|\!|(\nabla_v+\tau\nabla_x)^n f|\!|\!|_{\mu;p},\tag{4.17}$$

where

$$|\!|\!|g|\!|\!|_{\mu;p}=\sum_{l\in\mathbb{Z}^d}e^{2\pi\mu|l|}\|\hat g(l,v)\|_{L^p(\mathbb{R}^d_v)}.\tag{4.18}$$

4.3. Relations between functional spaces

The next propositions are easily checked.

PROPOSITION 4.15. *With the notation from* §4.2, *for any* $\tau\in\mathbb{R}$,

(i) *if f is a function only of x then*

$$\|f\|_{\mathcal{C}^{\lambda,\mu}_\tau}=\|f\|_{\mathcal{C}^{\lambda|\tau|+\mu}}\quad\textit{and}\quad\|f\|_{\mathcal{F}^{\lambda,\mu}_\tau}=\|f\|_{\mathcal{Z}^{\lambda,\mu}_\tau}=\|f\|_{\mathcal{F}^{\lambda|\tau|+\mu}};$$

(ii) *if f is a function only of v then*

$$\|f\|_{\mathcal{C}^{\lambda,\mu;p}_\tau}=\|f\|_{\mathcal{Z}^{\lambda,\mu;p}_\tau}=\|f\|_{\mathcal{C}^{\lambda;p}}\quad\textit{and}\quad\|f\|_{\mathcal{F}^{\lambda,\mu}_\tau}=\|f\|_{\mathcal{F}^{\lambda}};$$

(iii) *for any function $f=f(x,v)$, if $\langle\cdot\rangle$ stands for spatial average then*

$$\|\langle f\rangle\|_{\mathcal{C}^{\lambda;p}}\leqslant\|f\|_{\mathcal{Z}^{\lambda,\mu;p}_\tau};$$

(iv) *for any function $f=f(x,v)$,*

$$\left\|\int_{\mathbb{R}^d}f\,dv\right\|_{\mathcal{F}^{\lambda|\tau|+\mu}}\leqslant\|f\|_{\mathcal{Z}^{\lambda,\mu;1}_\tau}.$$

Remark 4.16. Note, in Proposition 4.15 (i) and (iv), how the regularity in x is improved by the time-shift.

Proof. Only (iv) requires some explanations. Let

$$\varrho(x) = \int_{\mathbb{R}^d} f(x,v)\, dv.$$

Then for any $k \in \mathbb{Z}^d$,

$$\hat{\varrho}(k) = \int_{\mathbb{R}^d} \hat{f}(k,v)\, dv;$$

so for any $n \in \mathbb{N}_0^d$,

$$(2i\pi tk)^n \hat{\varrho}(k) = \int_{\mathbb{R}^d} (2i\pi tk)^n \hat{f}(k,v)\, dv = \int_{\mathbb{R}^d} (\nabla_v + 2i\pi tk)^n \hat{f}(k,v)\, dv.$$

Recalling the conventions from Appendix A.1 we deduce

$$\sum_{\substack{k\in\mathbb{Z}^d \\ n\in\mathbb{N}_0^d}} e^{2\pi\mu|k|} \frac{|2\pi\lambda tk|^n}{n!} |\hat{\varrho}(k)| \leqslant \sum_{\substack{k\in\mathbb{Z}^d \\ n\in\mathbb{N}_0^d}} e^{2\pi\mu|k|} \frac{\lambda^n}{n!} \int_{\mathbb{R}^d} |(\nabla_v + 2i\pi tk)^n \hat{f}(k,v)|\, dv = \|f\|_{\mathcal{Z}_t^{\lambda,\mu;1}}. \qquad \square$$

Proposition 4.17. *With the notation from* §4.2,

$$\lambda \leqslant \lambda' \text{ and } \mu \leqslant \mu' \quad \Longrightarrow \quad \|f\|_{\mathcal{Z}_\tau^{\lambda,\mu}} \leqslant \|f\|_{\mathcal{Z}_\tau^{\lambda',\mu'}}.$$

Moreover, for $\tau, \bar{\tau} \in \mathbb{R}$ *and any* $p \in [1,\infty]$,

$$\|f\|_{\mathcal{Z}_{\bar{\tau}}^{\lambda,\mu;p}} \leqslant \|f\|_{\mathcal{Z}_\tau^{\lambda,\mu+\lambda|\bar{\tau}-\tau|;p}}. \tag{4.19}$$

Remark 4.18. Note carefully that the spaces $\mathcal{Z}_\tau^{\lambda,\mu}$ are *not* ordered with respect to the parameter τ, which cannot be thought of as a regularity index. We could dispend with this parameter if we were working in time $O(1)$; but (4.19) is of course of absolutely no use. This means that errors on the exponent τ should remain somehow small, in order to be controllable by small losses on the exponent μ.

Finally we state an easy proposition which follows from the time invariance of the free transport equation.

Proposition 4.19. *For any* $X \in \{\mathcal{C}, \mathcal{F}, \mathcal{Z}\}$ *and any* $t, \tau \in \mathbb{R}$,

$$\|f \circ S_t^0\|_{X_\tau^{\lambda,\mu}} = \|f\|_{X_{t+\tau}^{\lambda,\mu}}.$$

Now we shall see that the hybrid norms, and certain variants thereof, enjoy properties rather similar to those of the single-variable analytic norms studied before. This will sometimes be technical, and the reader who would like to reconnect to physical problems is advised to go directly to §4.11.

4.4. Injections

In this section we relate $\mathcal{Z}_\tau^{\lambda,\mu;p}$ norms to more standard norms entirely based on Fourier space. In the next theorem we write

$$\|f\|_{\mathcal{Y}_\tau^{\lambda,\mu}} := \|f\|_{\mathcal{F}_\tau^{\lambda,\mu;\infty}} = \sup_{k\in\mathbb{Z}^d}\sup_{\eta\in\mathbb{R}^d} e^{2\pi\mu|k|}e^{2\pi\lambda|\eta+k\tau|}|\tilde{f}(k,\eta)|. \tag{4.20}$$

THEOREM 4.20. (Injections between analytic spaces) (i) *If* $\lambda,\mu\geqslant 0$ *and* $\tau\in\mathbb{R}$ *then*

$$\|f\|_{\mathcal{Y}_\tau^{\lambda,\mu}} \leqslant \|f\|_{\mathcal{Z}_\tau^{\lambda,\mu;1}}. \tag{4.21}$$

(ii) *If* $0<\lambda<\bar{\lambda}$, $0<\mu<\bar{\mu}\leqslant M$ *and* $\tau\in\mathbb{R}$, *then*

$$\|f\|_{\mathcal{Z}_\tau^{\lambda,\mu}} \leqslant \frac{C(d,\bar{\mu})}{(\bar{\lambda}-\lambda)^d(\bar{\mu}-\mu)^d}\|f\|_{\mathcal{Y}_\tau^{\bar{\lambda},\bar{\mu}}}. \tag{4.22}$$

(iii) *If* $0<\lambda<\bar{\lambda}\leqslant\Lambda$, $0<\mu<\bar{\mu}\leqslant M$ *and* $b\leqslant\beta\leqslant B$, *then there is* $C=C(\Lambda,M,b,B,d)$ *such that*

$$\|f\|_{\mathcal{Z}_\tau^{\lambda,\mu;1}} \leqslant C^{1/\min\{\bar{\lambda}-\lambda,\bar{\mu}-\mu\}} \times\left(\|f\|_{\mathcal{Y}_\tau^{\bar{\lambda},\bar{\mu}}} + \max\left\{\int_{\mathbb{T}^d}\int_{\mathbb{R}^d}|f(x,v)|e^{\beta|v|}\,dv\,dx, \left(\int_{\mathbb{T}^d}\int_{\mathbb{R}^d}|f(x,v)|e^{\beta|v|}\,dv\,dx\right)^2\right\}\right).$$

Remark 4.21. The combination of (ii) and (iii), plus elementary Lebesgue interpolation, enables us to control all norms $\mathcal{Z}_\tau^{\lambda,\mu;p}$, $1\leqslant p\leqslant\infty$.

Proof. By the invariance under the action of free transport, it is sufficient to do the proof for $\tau=0$.

By integration by parts in the Fourier transform formula, we have

$$\tilde{f}(k,\eta) = \int_{\mathbb{R}^d}\hat{f}(k,v)e^{-2i\pi\eta\cdot v}\,dv = \int_{\mathbb{R}^d}\nabla_v^m\hat{f}(k,v)\frac{e^{-2i\pi\eta\cdot v}}{(2i\pi\eta)^m}\,dv.$$

So

$$|\tilde{f}(k,\eta)| \leqslant \frac{1}{(2\pi|\eta|)^m}\int_{\mathbb{R}^d}|\nabla_v^m\hat{f}(k,v)|\,dv;$$

and therefore

$$\begin{aligned} e^{2\pi\mu|k|}e^{2\pi\lambda|\eta|}|\tilde{f}(k,\eta)| &\leqslant e^{2\pi\mu|k|}\sum_{n\in\mathbb{N}_0^d}\frac{(2\pi\lambda)^n}{n!}|\eta|^n|\tilde{f}(k,\eta)| \\ &\leqslant e^{2\pi\mu|k|}\sum_{n\in\mathbb{N}_0^d}\frac{\lambda^n}{n!}\int_{\mathbb{R}^d}|\nabla_v^n\tilde{f}(k,v)|\,dv. \end{aligned}$$

This establishes (i).

Next, by differentiating the identity

$$\hat{f}(k,v) = \int_{\mathbb{R}^d} \tilde{f}(k,\eta) e^{2i\pi\eta\cdot v}\, d\eta,$$

we get

$$\nabla_v^m \hat{f}(k,v) = \int_{\mathbb{R}^d} \tilde{f}(k,\eta)(2i\pi\eta)^m e^{2i\pi\eta\cdot v}\, d\eta. \tag{4.23}$$

Then we deduce (ii) by writing

$$\begin{aligned}
&\sum_{\substack{k\in\mathbb{Z}^d\\ m\in\mathbb{N}_0^d}} e^{2\pi\mu|k|}\frac{\lambda^m}{m!}\|\nabla_v^m \hat{f}(k,v)\|_{L^\infty(dv)}\\
&\qquad\leqslant \sum_{k\in\mathbb{Z}^d} e^{2\pi\mu|k|}\int_{\mathbb{R}^d} e^{2\pi\lambda|\eta|}|\tilde{f}(k,\eta)|\, d\eta\\
&\qquad\leqslant \left(\sum_{k\in\mathbb{Z}^d} e^{-2\pi(\bar{\mu}-\mu)|k|}\right)\left(\int_{\mathbb{R}^d} e^{-2\pi(\bar{\lambda}-\lambda)|\eta|}\, d\eta\right)\sup_{\substack{k\in\mathbb{Z}^d\\ \eta\in\mathbb{R}^d}} |\tilde{f}(k,\eta)|e^{2\pi\bar{\lambda}|\eta|}e^{2\pi\bar{\mu}|k|}.
\end{aligned}$$

The proof of (iii) is the most tricky. We start again from (4.23), but now we integrate by parts in the η variable:

$$\nabla_v^m \hat{f}(k,v) = (-1)^q \int_{\mathbb{R}^d} \nabla_\eta^q[\tilde{f}(k,\eta)(2i\pi\eta)^m]\frac{e^{2i\pi\eta\cdot v}}{(2i\pi v)^q}\, dv, \tag{4.24}$$

where $q{=}q(v)$ is a multi-index to be chosen.

We split $\mathbb{R}^d_v$ into 2^d disjoint regions $\Delta(i_1,...,i_n)$, where $i_1,...,i_n$ are distinct indices in $\{1,...,d\}$:

$$\Delta(I) = \{v\in\mathbb{R}^d : |v_j|\geqslant 1 \text{ for all } j\in I \text{ and } |v_j|<1 \text{ for all } j\notin I\}.$$

If $v{\in}\Delta(i_1,...,i_n)$ we apply (4.24) with the multi-index q defined by

$$q_j = \begin{cases} 2, & \text{if } j\in\{i_1,...,i_n\},\\ 0, & \text{otherwise.}\end{cases}$$

This gives

$$\begin{aligned}
&\int_{\Delta(i_1,...,i_n)} |\nabla_v^m \hat{f}(k,v)|\, dv\\
&\qquad\leqslant \left(\frac{1}{(2\pi)^{2n}}\int_{\Delta(i_1,...,i_n)}\frac{dv_{i_1}\,...\,dv_{i_n}}{|v_{i_1}|^2\,...\,|v_{i_n}|^2}\right)\sup_{\substack{k\in\mathbb{Z}^d\\ \eta\in\mathbb{R}^d}} |\nabla_\eta^q[\tilde{f}(k,\eta)(2i\pi\eta)^m]|.
\end{aligned}$$

Summing up all pieces and using the Leibniz formula, we get

$$\int_{\mathbb{R}^d} |\nabla_v^m \hat{f}(k,v)|\,dv \leqslant C(d)(1+m^{2d}) \sup_{\substack{k\in\mathbb{Z}^d\\ \eta\in\mathbb{R}^d}} \sup_{|q|\leqslant 2d} |\nabla_\eta^q \tilde{f}(k,\eta)|\,|2\pi\eta|^{m-q}.$$

At this point we apply Lemma 4.22 below with

$$\varepsilon = \frac{1}{4}\min\left\{\frac{\bar{\lambda}-\lambda}{\bar{\lambda}}, \frac{\bar{\mu}-\mu}{\bar{\mu}}\right\},$$

and we get, for $q\leqslant 2d$,

$$|\nabla_\eta^q \tilde{f}(k,\eta)| \leqslant C(d)^{\max\{\bar{\lambda}/(\bar{\lambda}-\lambda),\bar{\mu}/(\bar{\mu}-\mu)\}} K(b,B) e^{-\pi(\lambda+\bar{\lambda})|\eta|} \Big(\sup_{\eta\in\mathbb{R}^d} |\tilde{f}(k,\eta)| e^{2\pi\bar{\lambda}|\eta|}\Big)^{1-\varepsilon}$$
$$\times \max\left\{\left(\sup_{\substack{l\in\mathbb{N}_0^d\\ \eta\in\mathbb{R}^d}} \frac{\beta^l \|\nabla_\eta^l \tilde{f}\|_\infty}{l!}\right)^{\varepsilon}, \left(\sup_{\substack{l\in\mathbb{N}_0^d\\ \eta\in\mathbb{R}^d}} \frac{\beta^l \|\nabla_\eta^l \tilde{f}\|_\infty}{l!}\right)^{2\varepsilon}\right\}.$$

Of course,

$$\frac{\beta^l |\nabla_\eta^l \tilde{f}(k,\eta)|}{l!} \leqslant (2\pi\beta)^l \int_{\mathbb{R}^d} |\hat{f}(k,v)| \frac{|v|^l}{l!}\,dv$$
$$\leqslant \int_{\mathbb{R}^d} |f(x,v)| (2\pi\beta)^l \frac{|v|^l}{l!}\,dv \leqslant \int_{\mathbb{R}^d} |f(x,v)| e^{2\pi\beta|v|}\,dv.$$

So, all in all,

$$\sum_{\substack{k\in\mathbb{Z}^d\\ m\in\mathbb{N}_0^d}} e^{2\pi\mu|k|} \frac{\lambda^m}{m!} \int_{\mathbb{R}^d} |\nabla_v^m \hat{f}(k,v)|\,dv$$
$$\leqslant \sum_{|q|\leqslant 2d} C(d,\Lambda,M,b,B)^{1/\min\{\bar{\lambda}-\lambda,\bar{\mu}-\mu\}}$$
$$\times \left(\sup_{\eta\in\mathbb{R}^d} e^{-\pi(\lambda+\bar{\lambda})|\eta|} \sum_{m\in\mathbb{N}_0^d} \frac{\lambda^m (1+m)^{2d} |2\pi\eta|^{m-q}}{m!}\right)$$
$$\times \left(\sum_{k\in\mathbb{Z}^d} e^{-2\pi(\bar{\mu}(1-\varepsilon)-\mu)|k|}\right) \left(\sup_{\substack{k\in\mathbb{Z}^d\\ \eta\in\mathbb{R}^d}} |\tilde{f}(k,\eta)| e^{2\pi\bar{\mu}|k|} e^{2\pi\bar{\lambda}|\eta|}\right)^{1-\varepsilon}$$
$$\times \max\left\{\left(\int_{\mathbb{R}^d} |f(x,v)| e^{\beta|v|}\,dv\right)^{\varepsilon}, \left(\int_{\mathbb{R}^d} |f(x,v)| e^{\beta|v|}\,dv\right)^{2\varepsilon}\right\}.$$

Since

$$\sum_{m\in\mathbb{N}_0^d} \frac{\lambda^m (1+m)^{2d} |2\pi\eta|^{m-q}}{m!} \leqslant C(q,\Lambda) e^{\pi(\lambda+\bar{\lambda})|\eta|}$$

and

$$\sum_{k\in\mathbb{Z}^d} e^{-2\pi(\bar{\mu}(1-\varepsilon)-\mu)|k|} \leqslant \sum_{k\in\mathbb{Z}^d} e^{-\pi(\bar{\mu}-\mu)|k|} \leqslant \frac{C}{(\bar{\mu}-\mu)^d},$$

we easily end up with the desired result. □

LEMMA 4.22. *Let $f:\mathbb{R}^d\to\mathbb{C}$ and let $\alpha>0$, $A\geqslant 1$ and $q\in\mathbb{N}_0^d$. Let further β be such that $0<b\leqslant\beta\leqslant B$. If $|f(x)|\leqslant Ae^{-\alpha|x|}$ for all x, then for any $\varepsilon\in\left(0,\frac{1}{4}\right)$, one has*

$$|\nabla^q f(x)|\leqslant C(q,d)^{1/\varepsilon}K(b,B)A^{1-\varepsilon}e^{-(1-2\varepsilon)\alpha|x|} \times \sup_{r\in\mathbb{N}_0^d}\max\left\{\left(\beta^r\frac{\|\nabla^r f\|_\infty}{r!}\right)^{\varepsilon},\left(\beta^r\frac{\|\nabla^r f\|_\infty}{r!}\right)^{2\varepsilon}\right\}. \tag{4.25}$$

Remark 4.23. One may conjecture that the optimal constant in the right-hand side of (4.25) is in fact polynomial in $1/\varepsilon$; if this conjecture holds true, then the constants in Theorem 4.20 (iii) can be improved accordingly. Mironescu communicated to us a derivation of polynomial bounds for the optimal constant in the related inequality

$$\|f^{(k)}\|_{L^\infty(\mathbb{R})}\leqslant C(k)\|f\|_{L^1(\mathbb{R})}^{1/(k+2)}\|f^{(k+1)}\|_{L^\infty(\mathbb{R})}^{(k+1)/(k+2)},$$

based on a real interpolation method.

Proof. Let us first see f as a function of x_1, and treat $x'=(x_2,...,x_d)$ as a parameter. Thus the assumption is $|f(x_1,x')|\leqslant(Ae^{-\alpha|x'|})e^{-\alpha|x_1|}$. By a more or less standard interpolation inequality [22, Lemma A.1],

$$|\partial_1 f(x_1,x')|\leqslant 2\sqrt{Ae^{-\alpha|x'|}}\sqrt{e^{-\alpha|x_1|}}\|\partial_1^2 f(x_1,x')\|_\infty^{1/2}=2\sqrt{Ae^{-\alpha|x|}}\sqrt{\|\partial_1^2 f\|_\infty}. \tag{4.26}$$

Let C_{q_1,r_1} be the optimal constant (not smaller than 1) such that

$$|\partial_1^{q_1} f(x_1,x')|\leqslant C_{q_1,r_1}(Ae^{-\alpha|x|})^{1-q_1/r_1}\|\partial_r^{r_1} f(x_1,x')\|_\infty^{q_1/r_1}. \tag{4.27}$$

By iterating (4.26), we find $C_{q_1,r_1}\leqslant 2\sqrt{C_{q_1-1,r_1}C_{q_1+1,r_1}}$. It follows by induction that

$$C_{q,r}\leqslant 2^{q(r-q)}.$$

Next, using (4.27) and interpolating according to the second variable x_2 as in (4.26), we get

$$\begin{aligned}&|\partial_2^{q_2}\partial_1^{q_1} f(x)|\\ &\quad\leqslant C_{q_2,r_2}(C_{q_1,r_1}(Ae^{-\alpha|x|})^{1-q_1/r_1}\|\partial_1^{r_1} f\|_\infty^{q_1/r_1})^{1-q_2/r_2}\|\partial_2^{r_2}\partial_1^{q_1} f\|_\infty^{q_2/r_2}\\ &\quad\leqslant C_{q_1,r_1}C_{q_2,r_2}(Ae^{-\alpha|x|})^{(1-q_1/r_1)(1-q_2/r_2)}\|\partial_1^{r_1} f\|_\infty^{(q_1/r_1)(1-q_2/r_2)}\|\partial_2^{r_2}\partial_1^{q_1} f\|_\infty^{q_2/r_2}.\end{aligned}$$

We repeat this until we get

$$\begin{aligned}|\nabla^q f(x)|\leqslant{}& C_{q_1,r_1}\dots C_{q_d,r_d}(Ae^{-\alpha|x|})^{(1-q_1/r_1)\dots(1-q_d/r_d)}\\ &\times\|\partial_1^{r_1} f\|_\infty^{(q_1/r_1)(1-q_2/r_2)\dots(1-q_d/r_d)}\|\partial_1^{q_1}\partial_2^{r_2} f\|^{(q_2/r_2)(1-q_3/r_3)\dots(1-q_d/r_d)}\\ &\dots\|\partial_1^{q_1}\partial_2^{q_2}\dots\partial_{d-1}^{q_{d-1}}\partial_d^{r_d} f\|^{q_d/r_d}.\end{aligned} \tag{4.28}$$

Choose r_j, $1\leqslant j\leqslant d$, in such a way that

$$\frac{\varepsilon}{d}\leqslant\frac{q_j}{r_j}\leqslant\frac{2\varepsilon}{d};$$

this is always possible for $\varepsilon<\frac{1}{4}d$. Then $C_{q_j,r_j}\leqslant(2^{dq_j^2})^{1/\varepsilon}$, and (4.28) implies that

$$|\nabla^q f(x)|\leqslant(2^{d|q|^2})^{1/\varepsilon}(Ae^{-\alpha|x|})^{1-\varepsilon}\max_{s\leqslant r+q}\{\|\nabla^s f\|_\infty^\varepsilon,\|\nabla^s f\|_\infty^{2\varepsilon}\}.$$

Then, since $2(r+q)\varepsilon\leqslant 3dq$, we have, by a crude application of Stirling's formula (in quantitative form), for $s\leqslant r+q$,

$$\|\nabla^s f\|_\infty^\varepsilon\leqslant\left(\frac{\beta^s\|\nabla^s f\|_\infty}{s!}\right)^\varepsilon\left(\frac{s!}{\beta^s}\right)^\varepsilon\leqslant\left(\sup_{n\in\mathbb{N}_0^d}\frac{\beta^n\|\nabla^n f\|_\infty}{n!}\right)^\varepsilon C(\beta,q,d)\varepsilon^{-3dq},$$

and the result easily follows. □

4.5. Algebra property in two variables

In this section we only consider the norms $\mathcal{Z}_\tau^{\lambda,\mu;p}$; but similar results would hold true for the two-variables spaces $\mathcal{C}$ and $\mathcal{F}$, and could be proven with the same methods as those used for the one-variable spaces $\mathcal{F}^\lambda$ and $\mathcal{C}^\lambda$, respectively (note that the Leibniz formula still applies because ∇_x and $\nabla_v+\tau\nabla_x$ commute).

PROPOSITION 4.24. (i) *For any $\lambda,\mu\geqslant 0$, any $\tau\in\mathbb{R}$ and any $p,q,r\in[1,\infty]$ such that $1/p+1/q=1/r$, we have*

$$\|fg\|_{\mathcal{Z}_\tau^{\lambda,\mu;r}}\leqslant\|f\|_{\mathcal{Z}_\tau^{\lambda,\mu;p}}\|g\|_{\mathcal{Z}_\tau^{\lambda,\mu;q}}.$$

(ii) *As a consequence, $\mathcal{Z}_\tau^{\lambda,\mu}=\mathcal{Z}_\tau^{\lambda,\mu;\infty}$ is a normed algebra*:

$$\|fg\|_{\mathcal{Z}_\tau^{\lambda,\mu}}\leqslant\|f\|_{\mathcal{Z}_\tau^{\lambda,\mu}}\|g\|_{\mathcal{Z}_\tau^{\lambda,\mu}}.$$

In particular, $\|f^n\|_{\mathcal{Z}_\tau^{\lambda,\mu}}\leqslant\|f\|_{\mathcal{Z}_\tau^{\lambda,\mu}}^n$ for any $n\in\mathbb{N}_0$, and $\|e^f\|_{\mathcal{Z}_\tau^{\lambda,\mu}}\leqslant e^{\|f\|_{\mathcal{Z}_\tau^{\lambda,\mu}}}$.

Proof. First we note that (with the notation (4.18)) $\|\!|\cdot|\!\|_{\mu;r}$ satisfies the "(p,q,r) property": whenever $p,q,r\in[1,\infty]$ satisfy $1/p+1/q=1/r$, we have

$$\begin{aligned}
\|\!|fg|\!\|_{\mu;r}&=\sum_{l\in\mathbb{Z}^d}e^{2\pi\mu|l|}\|\hat{fg}(l,\cdot)\|_{L^r(\mathbb{R}^d_v)}\\
&=\sum_{l\in\mathbb{Z}^d}e^{2\pi\mu|l|}\left\|\sum_{k\in\mathbb{Z}^d}\hat f(k,\cdot)\hat g(l-k,\cdot)\right\|_{L^r(\mathbb{R}^d_v)}\\
&\leqslant\sum_{l\in\mathbb{Z}^d}\sum_{k\in\mathbb{Z}^d}e^{2\pi\mu|k|}e^{2\pi\mu|l-k|}\|\hat f(k,\cdot)\|_{L^p(\mathbb{R}^d_v)}\|\hat g(l-k,\cdot)\|_{L^q(\mathbb{R}^d_v)}\\
&=\|\!|f|\!\|_{\mu;p}\|\!|g|\!\|_{\mu;q}.
\end{aligned}$$

Next, we write

$$\begin{aligned}
\|\!| fg |\!\|_{\mathcal{Z}^{\lambda,\mu;r}_\tau} &= \sum_{n\in\mathbb{N}_0^d} \frac{\lambda^n}{n!} \|\!| (\nabla_v+\tau\nabla_x)^n (fg) |\!\|_{\mu;r} \\
&= \sum_{n\in\mathbb{N}_0^d} \frac{\lambda^n}{n!} \Big\|\!\Big| \sum_{m\leqslant n} \binom{n}{m} (\nabla_v+\tau\nabla_x)^m f (\nabla_v+\tau\nabla_x)^{n-m} g \Big|\!\Big\|_{\mu;r} \\
&\leqslant \sum_{n\in\mathbb{N}_0^d} \frac{\lambda^n}{n!} \sum_{m\leqslant n} \binom{n}{m} \|\!| (\nabla_v+\tau\nabla_x)^m f |\!\|_{\mu;p} \|\!| (\nabla_v+\tau\nabla_x)^{n-m} g |\!\|_{\mu;q} \\
&= \Big(\sum_{m\in\mathbb{N}_0^d} \frac{\lambda^m}{m!} \|\!| (\nabla_v+\tau\nabla_x)^m f |\!\|_{\mu;p} \Big) \sum_{l\in\mathbb{N}_0^d} \frac{\lambda^l}{l!} \|\!| (\nabla_v+\tau\nabla_x)^l f |\!\|_{\mu;q} \\
&= \|\!| f |\!\|_{\mathcal{Z}^{\lambda,\mu;p}_\tau} \|\!| g |\!\|_{\mathcal{Z}^{\lambda,\mu;q}_\tau}.
\end{aligned}$$

(We could also reduce to $\tau=0$ by means of Proposition 4.19.) □

4.6. Composition inequality

PROPOSITION 4.25. (Composition inequality in two variables) *For any $\lambda,\mu\geqslant 0$ and any $p\in[1,\infty]$, $\tau\in\mathbb{R}$, $\sigma\in\mathbb{R}$, $a\in\mathbb{R}\setminus\{0\}$ and $b\in\mathbb{R}$,*

$$\|f(x+bv+X(x,v),av+V(x,v))\|_{\mathcal{Z}^{\lambda,\mu;p}_\tau} \leqslant |a|^{-d/p} \|f\|_{\mathcal{Z}^{\alpha,\beta;p}_\sigma}, \tag{4.29}$$

where

$$\alpha = \lambda|a| + \|V\|_{\mathcal{Z}^{\lambda,\mu}_\tau} \quad \text{and} \quad \beta = \mu+\lambda|b+\tau-a\sigma| + \|X-\sigma V\|_{\mathcal{Z}^{\lambda,\mu}_\tau}. \tag{4.30}$$

Remark 4.26. The norms in (4.30) for X and V have to be based on L^∞, not just any L^p. Also note that the fact that the second argument of f has the form $av+V$ (and not $av+cx+V$) is related to Remark 4.7.

Proof. The proof is a combination of the arguments in Proposition 4.8. In a first step, we do it for the case $\tau=\sigma=0$, and we write $\|\cdot\|_{\lambda,\mu;p} = \|\cdot\|_{\mathcal{Z}^{\lambda,\mu;p}_0}$.

From the expansion $f(x,v) = \sum_{k\in\mathbb{Z}^d} \hat f(k,v) e^{2i\pi k\cdot x}$ we deduce that

$$\begin{aligned}
h(x,v) &:= f(x+bv+X(x,v), av+V(x,v)) \\
&= \sum_{k\in\mathbb{Z}^d} \hat f(k, av+V) e^{2i\pi k\cdot(x+bv+X)} \\
&= \sum_{\substack{k\in\mathbb{Z}^d \\ m\in\mathbb{N}_0^d}} \nabla_v^m \hat f(k,av) \cdot \frac{V^m}{m!} e^{2i\pi k\cdot x} e^{2i\pi k\cdot bv} e^{2i\pi k\cdot X}.
\end{aligned}$$

Taking the Fourier transform in x, we see that for any $l\in\mathbb{Z}^d$,

$$\hat{h}(l,v)=\sum_{\substack{k\in\mathbb{Z}^d\\ m\in\mathbb{N}_0^d}}\nabla_v^m\hat{f}(k,av)e^{2i\pi k\cdot bv}\sum_{j\in\mathbb{Z}^d}\frac{(\widehat{V^m})(j)}{m!}(e^{2i\pi k\cdot X})\hat{}\,(l-k-j).$$

Differentiating n times via the Leibniz formula (here applied to a product of four functions), we get

$$\begin{aligned}\nabla_v^n\hat{h}(l,v)=\sum_{\substack{k,j\in\mathbb{Z}^d\\ m,n_1,n_2,n_3,n_4\in\mathbb{N}_0^d\\ n_1+n_2+n_3+n_4=n}}&\frac{n!a^{n_1}}{n_1!n_2!n_3!n_4!}\nabla_v^{m+n_1}\hat{f}(k,av)\\ &\times\frac{\nabla_v^{n_2}(\widehat{V^m})(j)}{m!}\nabla_v^{n_3}(e^{2i\pi k\cdot X})\hat{}\,(l-k-j,v)(2i\pi bk)^{n_4}e^{2i\pi k\cdot bv}.\end{aligned}$$

Multiplying by $\lambda^n e^{2\pi\mu|l|}/n!$ and summing over n and l, taking L^p norms and using $\|fg\|_{L^p}\leqslant\|f\|_{L^p}\|g\|_{L^\infty}$, we finally obtain

$$\begin{aligned}&\|h\|_{\lambda,\mu}\\ &\leqslant|a|^{-d/p}\sum_{\substack{k,j,l\in\mathbb{Z}^d\\ m,n,n_1,n_2,n_3,n_4\in\mathbb{N}_0^d\\ n_1+n_2+n_3+n_4=n}}\frac{\lambda^n e^{2\pi\mu|l|}|a|^{n_1}}{n_1!n_2!n_3!n_4!}\|\nabla_v^{m+n_1}\hat{f}(k,\cdot)\|_{L^p}\left\|\frac{\nabla_v^{n_2}(\widehat{V^m})(j)}{m!}\right\|_\infty\\ &\qquad\qquad\times\|\nabla_v^{n_3}(e^{2i\pi k\cdot X})\hat{}\,(l-k-j)\|_\infty(2\pi|b|\,|k|)^{n_4}\\ &=|a|^{-d/p}\sum_{\substack{k,j,l\in\mathbb{Z}^d\\ m,n_1,n_2,n_3,n_4\in\mathbb{N}_0^d}}\frac{\lambda^{n_1+n_2+n_3+n_4}e^{2\pi\mu|k|}e^{2\pi\mu|j|}e^{2\pi\mu|l-k-j|}|a|^{n_1}}{n_1!n_2!n_3!n_4!}\|\nabla_v^{m+n_1}\hat{f}(k,\cdot)\|_{L^p}\\ &\qquad\qquad\times\left\|\frac{\nabla_v^{n_2}(\widehat{V^m})(j)}{m!}\right\|_\infty\|\nabla_v^{n_3}(e^{2i\pi k\cdot X})\hat{}\,(l-k-j)\|_\infty(2\pi|b|\,|k|)^{n_4}\\ &\leqslant|a|^{-d/p}\sum_{\substack{k\in\mathbb{Z}^d\\ n_1,m\in\mathbb{N}_0^d}}\frac{\lambda^{n_1}|a|^{n_1}}{n_1!}\|\nabla_v^{n_1+m}\hat{f}(k,\cdot)\|_{L^p}\,e^{2\pi\mu|k|}\\ &\qquad\qquad\times\left(\frac{1}{m!}\sum_{\substack{j\in\mathbb{Z}^d\\ n_2\in\mathbb{N}_0^d}}\frac{\lambda^{n_2}}{n_2!}e^{2\pi\mu|j|}\|\nabla_v^{n_2}(\widehat{V^m})(j)\|_\infty\right)\\ &\qquad\qquad\times\left(\sum_{\substack{h\in\mathbb{Z}^d\\ n_3\in\mathbb{N}_0^d}}\frac{\lambda^{n_3}}{n_3!}e^{2\pi\mu|h|}\|\nabla_v^{n_3}(e^{2i\pi k\cdot X})\hat{}\,(h)\|_\infty\right)\sum_{n_4\in\mathbb{N}_0^d}\frac{(2\pi\lambda|b|\,|k|)^{n_4}}{n_4!}\\ &=|a|^{-d/p}\sum_{\substack{k\in\mathbb{Z}^d\\ n_1,m\in\mathbb{N}_0^d}}\frac{(\lambda|a|)^{n_1}}{n_1!}e^{2\pi\mu|k|}\|\nabla_v^{n_1+m}\hat{f}(k,\cdot)\|_{L^p}\frac{\|V^m\|_{\lambda,\mu}}{m!}\|e^{2i\pi k\cdot X}\|_{\lambda,\mu}e^{2\pi\lambda|b|\,|k|}\end{aligned}$$

$$\leqslant |a|^{-d/p} \sum_{\substack{k\in\mathbb{Z}^d \\ n_1,m\in\mathbb{N}_0^d}} \frac{(\lambda|a|)^{n_1}}{n_1!} e^{2\pi(\mu+\lambda|b|)|k|} \|\nabla_v^{n_1+m} \hat{f}(k,\cdot)\|_{L^p} \frac{\|V\|_{\lambda,\mu}^m}{m!} e^{2\pi|k|\,\|X\|_{\lambda,\mu}}$$

$$= |a|^{-d/p} \sum_{\substack{k\in\mathbb{Z}^d \\ n\in\mathbb{N}_0^d}} \frac{1}{n!} (\lambda|a|+\|V\|_{\lambda,\mu})^n \|\nabla_v^n \hat{f}(k,\cdot)\|_{L^{\bar{p}}}\, e^{2\pi|k|(\mu+\lambda|b|+\|X\|_{\lambda,\mu})}$$

$$= |a|^{-d/p} \|f\|_{\lambda|a|+\|V\|_{\lambda,\mu},\mu+\lambda|b|+\|X\|_{\lambda,\mu}}.$$

Now we generalize this to arbitrary values of σ and τ. By Proposition 4.19,

$$\begin{aligned}
&\|f(x+bv+X(x,v), av+V(x,v))\|_{\mathcal{Z}_\tau^{\lambda,\mu;p}} \\
&\quad = \|f(x+v(b+\tau)+X(x+v\tau,v), av+V(x+v\tau,v))\|_{\mathcal{Z}^{\lambda,\mu;p}} \\
&\quad = \|f\circ S_\sigma^0 \circ S_{-\sigma}^0 (x+v(b+\tau)+X(x+v\tau,v), av+V(x+v\tau,v))\|_{\mathcal{Z}^{\lambda,\mu;p}} \\
&\quad = \|(f\circ S_\sigma^0)(x+v(b+\tau-a\sigma)+(X-\sigma V)(x+v\tau,v), av+V(x+v\tau,v))\|_{\mathcal{Z}^{\lambda,\mu;p}} \\
&\quad = \|(f\circ S_\sigma^0)(x+v(b+\tau-a\sigma)+Y(x,v), av+W(x,v))\|_{\mathcal{Z}^{\lambda,\mu;p}},
\end{aligned}$$

where

$$W(x,v) = V\circ S_\tau^0(x,v) \quad \text{and} \quad Y(x,v) = (X-\sigma V)\circ S_\tau^0(x,v).$$

Applying the result for $\tau=0$, we deduce that the norm of

$$h(x,v) = f(x+bv+X(x,v), av+V(x,v))$$

in $\mathcal{Z}_\tau^{\lambda,\mu}$ is bounded by

$$\|f\circ S_\sigma^0\|_{\mathcal{Z}^{\alpha,\beta;p}} = \|f\|_{\mathcal{Z}_\sigma^{\alpha,\beta;p}},$$

where

$$\alpha = \lambda|a|+\|V\circ S_\tau^0\|_{\mathcal{Z}^{\lambda,\mu}} = |a|\lambda+\|V\|_{\mathcal{Z}_\tau^{\lambda,\mu}}$$

and

$$\beta = \mu+\lambda|b+\tau-a\sigma|+\|(X-\sigma V)\circ S_\tau^0\|_{\mathcal{Z}^{\lambda,\mu}} = \mu+\lambda|b+\tau-a\sigma|+\|X-\sigma V\|_{\mathcal{Z}_\tau^{\lambda,\mu}}.$$

This establishes the desired bound. □

4.7. Gradient inequality

In the next proposition we shall write

$$\|f\|_{\dot{\mathcal{Z}}_\tau^{\lambda,\mu}} = \sum_{l\in\mathbb{Z}_*^d} \sum_{n\in\mathbb{N}_0^d} \frac{\lambda^n}{n!} \|(\nabla_v+2i\pi\tau l)^n \hat{f}(l,v)\|_{L^\infty(\mathbb{R}_v^d)}\, e^{2\pi\mu|l|}. \tag{4.31}$$

This is again a homogeneous (in the x variable) seminorm.

PROPOSITION 4.27. *For $\bar\lambda>\lambda\geqslant 0$ and $\bar\mu>\mu\geqslant 0$, we have the functional inequalities*

$$\|\nabla_x f\|_{\mathcal{Z}_\tau^{\lambda,\mu;p}} \leqslant \frac{C(d)}{\bar\mu-\mu}\|f\|_{\dot{\mathcal{Z}}_\tau^{\lambda,\bar\mu;p}},$$
$$\|(\nabla_v+\tau\nabla_x)f\|_{\mathcal{Z}_\tau^{\lambda,\mu;p}} \leqslant \frac{C(d)}{\lambda\log(\bar\lambda/\lambda)}\|f\|_{\mathcal{Z}_\tau^{\bar\lambda,\mu;p}}.$$

In particular, for $\tau\geqslant 0$ we have

$$\|\nabla_v f\|_{\mathcal{Z}_\tau^{\lambda,\mu;p}} \leqslant C(d)\bigg(\frac{1}{\lambda\log(\bar\lambda/\lambda)}\|f\|_{\mathcal{Z}_\tau^{\bar\lambda,\bar\mu;p}}+\frac{\tau}{\bar\mu-\mu}\|f\|_{\dot{\mathcal{Z}}_\tau^{\bar\lambda,\bar\mu;p}}\bigg).$$

The proof is similar to that of Proposition 4.10; the constant $C(d)$ arises in the choice of norm on $\mathbb{R}^d$. As a consequence, if $1<\bar\lambda/\lambda\leqslant 2$, we have e.g. the bound

$$\|\nabla f\|_{\mathcal{Z}_\tau^{\lambda,\mu;p}} \leqslant C(d)\bigg(\frac{1}{\bar\lambda-\lambda}+\frac{1+\tau}{\bar\mu-\mu}\bigg)\|f\|_{\mathcal{Z}_\tau^{\bar\lambda,\bar\mu;p}}.$$

4.8. Inversion

From the composition inequality follows an inversion estimate.

PROPOSITION 4.28. (Inversion inequality) (i) *Let $\lambda,\mu\geqslant 0$, $\tau\in\mathbb{R}$ and consider a function $F\colon \mathbb{T}^d\times\mathbb{R}^d\to\mathbb{T}^d\times\mathbb{R}^d$. Then there is $\varepsilon=\varepsilon(d)$ such that if F satisfies*

$$\|\nabla(F-\mathrm{Id})\|_{\mathcal{Z}_\tau^{\lambda',\mu'}} \leqslant \varepsilon(d),$$

where

$$\lambda'=\lambda+2\|F-\mathrm{Id}\,\|_{\mathcal{Z}_\tau^{\lambda,\mu}} \quad \text{and} \quad \mu'=\mu+2(1+|\tau|)\|F-\mathrm{Id}\,\|_{\mathcal{Z}_\tau^{\lambda,\mu}},$$

then F is invertible and

$$\|F^{-1}-\mathrm{Id}\,\|_{\mathcal{Z}_\tau^{\lambda,\mu}} \leqslant 2\|F-\mathrm{Id}\,\|_{\mathcal{Z}_\tau^{\lambda,\mu}}. \tag{4.32}$$

(ii) *More generally, if $F,G\colon \mathbb{T}^d\times\mathbb{R}^d\to\mathbb{T}^d\times\mathbb{R}^d$ are such that*

$$\|\nabla(F-\mathrm{Id})\|_{\mathcal{Z}_\tau^{\lambda',\mu'}} \leqslant \varepsilon(d), \tag{4.33}$$

where

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

then F is invertible and

$$\|F^{-1}\circ G-\mathrm{Id}\,\|_{\mathcal{Z}_\tau^{\lambda,\mu}} \leqslant 2\|F-G\|_{\mathcal{Z}_\tau^{\lambda,\mu}}. \tag{4.34}$$

Remark 4.29. The conditions become very stringent as τ becomes large: basically, $F-\mathrm{Id}$ (or $F-G$ in case (ii)) should be of order $o(1/\tau)$ for Proposition 4.28 to be applicable.

Remark 4.30. By Proposition 4.27, a sufficient condition for (4.33) to hold is that there be λ'' and μ'' such that $\lambda\leqslant\lambda''\leqslant 2\lambda$, $\mu\leqslant\mu''$ and

$$\|F-\mathrm{Id}\,\|_{\mathcal{Z}_\tau^{\lambda'',\mu''}}\leqslant\frac{\varepsilon'(d)}{1+\tau}\,\min\{\lambda''-\lambda',\mu''-\mu'\}.$$

However, this condition is in practice hard to fulfill.

Proof. We prove only (ii), being (i) a particular case. Let $f=F-\mathrm{Id}$, $h=F^{-1}\circ G-\mathrm{Id}$ and $g=G-\mathrm{Id}$, so that $\mathrm{Id}+g=(\mathrm{Id}+f)\circ(\mathrm{Id}+h)$, or equivalently

$$h=g-f\circ(\mathrm{Id}+h).$$

Thus h is a fixed point of

$$\Phi\colon Z\longmapsto g-f\circ(\mathrm{Id}+Z).$$

Note that $\Phi(0)=g-f$. If Φ is $\frac{1}{2}$-Lipschitz on the ball $B(0,2\|f-g\|)$ in $\mathcal{Z}_\tau^{\lambda,\mu}$, then (4.34) will follow by fixed-point iteration as in Theorem A.2. (Here $B(x,r):=\{y:\|x-y\|\leqslant r\}$.)

So let Z and $\widetilde{Z}$ be given with

$$\|Z\|_{\mathcal{Z}_\tau^{\lambda,\mu}},\|\widetilde{Z}\|_{\mathcal{Z}_\tau^{\lambda,\mu}}\leqslant 2\|f-g\|_{\mathcal{Z}_\tau^{\lambda,\mu}}.$$

We have

$$\Phi(Z)-\Phi(\widetilde{Z})=f(\mathrm{Id}+\widetilde{Z})-f(\mathrm{Id}+Z)=(\widetilde{Z}-Z)\cdot\int_0^1\nabla f(\mathrm{Id}+(1-\theta)Z+\theta\widetilde{Z})\,d\theta.$$

By Proposition 4.24,

$$\|\Phi(Z)-\Phi(\widetilde{Z})\|_{\mathcal{Z}_\tau^{\lambda,\mu}}\leqslant\|\widetilde{Z}-Z\|_{\mathcal{Z}_\tau^{\lambda,\mu}}\int_0^1\|\nabla f(\mathrm{Id}+(1-\theta)Z+\theta\widetilde{Z})\|_{\mathcal{Z}_\tau^{\lambda,\mu}}\,d\theta.$$

For any $\theta\in[0,1]$, by Proposition 4.25,

$$\|\nabla f(\mathrm{Id}+(1-\theta)Z+\theta\widetilde{Z})\|_{\mathcal{Z}_\tau^{\lambda,\mu}}\leqslant\|\nabla f\|_{\mathcal{Z}_\tau^{\hat{\lambda},\hat{\mu}}},$$

where

$$\hat{\lambda}=\lambda+\max\{\|Z\|,\|\widetilde{Z}\|\}\leqslant\lambda+2\|f-g\|_{\mathcal{Z}_\tau^{\lambda,\mu}}$$

and, writing $Z=(Z_x,Z_v)$ and $\widetilde{Z}=(\widetilde{Z}_x,\widetilde{Z}_v)$,

$$\hat{\mu}=\mu+\max\{\|Z_x-\tau Z_v\|,\|\widetilde{Z}_x-\tau\widetilde{Z}_v\|\}\leqslant\mu+2(1+|\tau|)\|f-g\|_{\mathcal{Z}_\tau^{\lambda,\mu}}.$$

If F and G satisfy the assumptions of Proposition 4.28, we deduce that

$$\|\Phi\|_{\mathrm{Lip}(B(0,2))}\leqslant C(d)\varepsilon(d),$$

and this is bounded above by $\frac{1}{2}$ if $\varepsilon(d)$ is small enough. □

4.9. Sobolev corrections

We shall need to quantify Sobolev regularity corrections to the analytic regularity, in the x variable.

Definition 4.31. (Hybrid analytic norms with Sobolev corrections) For $\lambda,\mu,\gamma\geqslant 0$, $\tau\in\mathbb{R}$ and $p\in[1,\infty]$ we define

$$\|f\|_{\mathcal{Z}_\tau^{\lambda,(\mu,\gamma);p}}=\sum_{l\in\mathbb{Z}^d}\sum_{n\in\mathbb{N}_0^d}\frac{\lambda^n}{n!}e^{2\pi\mu|l|}(1+|l|)^\gamma\|(\nabla_v+2i\pi\tau l)^n\hat{f}(l,v)\|_{L^p(\mathbb{R}^d_v)},$$
$$\|f\|_{\mathcal{F}^{\lambda,\gamma}}=\sum_{k\in\mathbb{Z}^d}e^{2\pi\lambda|k|}(1+|k|)^\gamma|\hat{f}(k)|.$$

Proposition 4.32. *Let $\lambda,\mu,\gamma\geqslant 0$, $\tau\in\mathbb{R}$ and $p\in[1,\infty]$. We have the following functional (in)equalities.*

(i) $\|f\|_{\mathcal{Z}_{t+\tau}^{\lambda,(\mu,\gamma);p}}=\|f\circ S_t^0\|_{\mathcal{Z}_\tau^{\lambda,(\mu,\gamma);p}}$;

(ii) *If $1/p+1/q=1/r$ then $\|fg\|_{\mathcal{Z}_\tau^{\lambda,(\mu,\gamma);r}}\leqslant\|f\|_{\mathcal{Z}_\tau^{\lambda,(\mu,\gamma);p}}\|g\|_{\mathcal{Z}_\tau^{\lambda,(\mu,\gamma);q}}$, and therefore in particular $\mathcal{Z}_\tau^{\lambda,(\mu,\gamma)}=\mathcal{Z}_\tau^{\lambda,(\mu,\gamma);\infty}$ is a normed algebra*;

(iii) *If f depends only on x then $\|f\|_{\mathcal{Z}_\tau^{\lambda,(\mu,\gamma)}}=\|f\|_{\mathcal{F}^{\lambda|\tau|+\mu,\gamma}}$*;

(iv) $\|f\|_{\mathcal{Z}_{\bar\tau}^{\lambda,(\mu,\gamma);p}}\leqslant\|f\|_{\mathcal{Z}_\tau^{\lambda,(\mu+\lambda|\tau-\bar\tau|,\gamma);p}}$;

(v) *For any $\sigma\in\mathbb{R}$, $a\in\mathbb{R}\setminus\{0\}$ and $b\in\mathbb{R}$,*

$$\|f(x+bv+X(x,v),av+V(x,v))\|_{\mathcal{Z}_\tau^{\lambda,(\mu,\gamma);p}}\leqslant|a|^{-d/p}\|f\|_{\mathcal{Z}_\sigma^{\alpha,(\beta,\gamma);p}},$$

where

$$\alpha=\lambda|a|+\|V\|_{\mathcal{Z}_\tau^{\lambda,(\mu,\gamma)}}\quad\text{and}\quad\beta=\mu+\lambda|b\tau-a\sigma|+\|X-\sigma V\|_{\mathcal{Z}_\tau^{\lambda,(\mu,\gamma)}};$$

(vi) Gradient inequality:

$$\|\nabla_x f\|_{\mathcal{Z}_\tau^{\lambda,(\mu,\gamma);p}}\leqslant\frac{C(d)}{\bar\mu-\mu}\|f\|_{\mathcal{Z}_\tau^{\lambda,(\bar\mu,\gamma);p}},$$
$$\|\nabla f\|_{\mathcal{Z}_\tau^{\lambda,(\mu,\gamma);p}}\leqslant C(d)\left(\frac{1}{\bar\lambda-\lambda}+\frac{1+\tau}{\bar\mu-\mu}\right)\|f\|_{\mathcal{Z}_\tau^{\bar\lambda,(\bar\mu,\gamma);p}};$$

(vii) Inversion: *if $F,G\colon\mathbb{T}^d\times\mathbb{R}^d\to\mathbb{T}^d\times\mathbb{R}^d$ are such that*

$$\|\nabla(F-\mathrm{Id})\|_{\mathcal{Z}_\tau^{\lambda',(\mu',\gamma)}}\leqslant\varepsilon(d),$$

where

$$\lambda'=\lambda+2\|F-G\|_{\mathcal{Z}_\tau^{\lambda,(\mu,\gamma)}}\quad\text{and}\quad\mu'=\mu+2(1+|\tau|)\|F-G\|_{\mathcal{Z}_\tau^{\lambda,(\mu,\gamma)}},$$

then

$$\|F^{-1}\circ G-\mathrm{Id}\,\|_{\mathcal{Z}_\tau^{\lambda,(\mu,\gamma)}}\leqslant 2\|F-G\|_{\mathcal{Z}_\tau^{\lambda,(\mu,\gamma)}}.\tag{4.35}$$

Proof. The proofs are the same as for the "plain" hybrid norms; the only notable point is that for the proof of (ii) we use, in addition to $e^{2\pi\lambda|k|}\leqslant e^{2\pi\lambda|k-l|}e^{2\pi\lambda|l|}$, the inequality

$$(1+|k|)^\gamma \leqslant (1+|k-l|)^\gamma (1+|l|)^\gamma. \qquad \square$$

Remark 4.33. Of course, some of the estimates in Proposition 4.32 can be "improved" by taking advantage of γ; e.g. for $\gamma\geqslant 1$ we have

$$\|\nabla_x f\|_{\mathcal{Z}^{\lambda,\mu;p}_\tau} \leqslant C(d)\|f\|_{\mathcal{Z}^{\lambda,(\mu,\gamma);p}_\tau}.$$

4.10. Individual mode estimates

To handle very singular cases, we shall at times need to estimate Fourier modes individually, rather than full norms. If $f=f(x,v)$, we write

$$(P_k f)(x,v) = \hat f(k,v) e^{2i\pi k\cdot x}. \tag{4.36}$$

In particular the following estimates will be useful.

PROPOSITION 4.34. *For any $\lambda,\mu\geqslant 0$, $\tau\in\mathbb{R}$, Lebesgue exponents $1/r=1/p+1/q$ and $k\in\mathbb{Z}^d$, we have the estimate*

$$\|P_k(fg)\|_{\mathcal{Z}^{\lambda,\mu;r}_\tau} \leqslant \sum_{l\in\mathbb{Z}^d} \|P_l f\|_{\mathcal{Z}^{\lambda,\mu;p}_\tau} \|P_{k-l} g\|_{\mathcal{Z}^{\lambda,\mu;q}_\tau}.$$

PROPOSITION 4.35. *For any $\lambda>0$, $\bar\mu\geqslant\mu\geqslant 0$, $\tau\in\mathbb{R}$, $p\in[1,\infty]$ and $k\in\mathbb{Z}^d$, we have the estimate*

$$\|P_k[f(x+X(x,v),v)]\|_{\mathcal{Z}^{\lambda,\mu;p}_\tau} \leqslant \sum_{l\in\mathbb{Z}^d} e^{-2\pi(\bar\mu-\mu)|k-l|} \|P_l f\|_{\mathcal{Z}^{\lambda,\nu;p}_\tau}, \quad \nu=\mu+\|X\|_{\mathcal{Z}^{\lambda,\bar\mu}_\tau}.$$

These estimates also have variants with Sobolev corrections. Note that when $\mu=\bar\mu$, Proposition 4.35 is a direct consequence of Proposition 4.25 with $V=0$, $b=0$ and $a=1$:

$$\|P_k[f(x+X(x,v),v)]\|_{\mathcal{Z}^{\lambda,\mu;p}_\tau} \leqslant \|f(x+X(x,v),v)\|_{\mathcal{Z}^{\lambda,\mu;p}_\tau} \leqslant \|f\|_{\mathcal{Z}^{\lambda,\nu;p}_\tau}, \quad \nu=\mu+\|X\|_{\mathcal{Z}^{\lambda,\mu}_\tau}.$$

Proof of Propositions 4.34 *and* 4.35. The proof of Proposition 4.34 is quite similar to the proof of Proposition 4.24. (It is no restriction to choose $\tau=0$ because P_k commutes with the free transport semigroup.) Proposition 4.35 needs a few words of explanation. As in the proof of Proposition 4.25 we let $h(x,v)=f(x+X(x,v),v)$, and readily obtain

$$\begin{aligned}\|P_k h\|_{\mathcal{Z}^{\lambda,\mu;p}_\tau} &= \sum_{n\in\mathbb{N}_0^d} \frac{\lambda^n e^{2\pi\mu|k|}}{n!} \|\nabla_v^n \hat h(k,v)\|_{L^p(dv)} \\ &\leqslant \sum_{\substack{n\in\mathbb{N}_0^d\\ l\in\mathbb{Z}^d}} \frac{\lambda^n e^{2\pi\mu|k|}}{n!} \|\nabla_v^n \hat f(l,v)\|_{L^p(dv)} \sum_{m\in\mathbb{N}_0^d} \frac{\lambda^m}{m!} \|\nabla_v^m (e^{2i\pi lX})\hat{}\,(k-l,v)\|_{L^\infty(dv)}.\end{aligned}$$

At this stage we write

$$e^{2\pi\mu|k|} \leqslant e^{2\pi\mu|l|} e^{-2\pi(\bar{\mu}-\mu)|k-l|} e^{2\pi\bar{\mu}|k-l|},$$

and use the following crude bound, which holds for all $l \in \mathbb{Z}^d$:

$$e^{2\pi\bar{\mu}|k-l|} \|\nabla_v^m (e^{2i\pi lX})\hat{}\,(k-l,v)\|_{L^\infty(dv)} \leqslant \sum_{j\in\mathbb{Z}^d} e^{2\pi\bar{\mu}|j|} \|\nabla_v^m (e^{2i\pi lX})\hat{}\,(j,v)\|_{L^\infty(dv)}.$$

The rest of the proof is as in Proposition 4.25. □

4.11. Measuring solutions of kinetic equations in large time

As we already discussed, even for the simplest kinetic equation, namely free transport, we cannot hope to have uniform-in-time regularity estimates in the velocity variable: rather, because of filamentation, we may have $\|\nabla_v f(t,\cdot)\| = O(t)$, $\|\nabla_v^2 f(t,\cdot)\| = O(t^2)$, etc. For analytic norms we may at best hope for exponential growth.

But the invariance of the "gliding" norms $\mathcal{Z}_\tau^{\lambda,\mu}$ under free transport (Proposition 4.19) makes it possible to look for uniform estimates such as

$$\|f(\tau,\cdot)\|_{\mathcal{Z}_\tau^{\lambda,\mu}} = O(1) \quad \text{as } \tau \to \infty. \tag{4.37}$$

Of course, by Proposition 4.27, (4.37) implies that

$$\|\nabla_v f(\tau,\cdot)\|_{\mathcal{Z}_\tau^{\lambda',\mu'}} = O(\tau) \quad \text{for } \lambda' < \lambda \text{ and } \mu' < \mu, \tag{4.38}$$

and nothing better as far as the asymptotic behavior of $\nabla_v f$ is concerned; but (4.37) is much more precise than (4.38). For instance it implies that

$$\|(\nabla_v + \tau\nabla_x) f(\tau,\cdot)\|_{\mathcal{Z}_\tau^{\lambda',\mu'}} = O(1) \quad \text{for } \lambda' < \lambda \text{ and } \mu' < \mu.$$

Another way to get rid of filamentation is to *average* over the spatial variable x, a common sense procedure which has already been used in physics [54, §49]. Think that, if f evolves according to free transport, or even according to the linearized Vlasov equation (3.3), then its space-average

$$\langle f \rangle(\tau, v) := \int_{\mathbb{T}^d} f(\tau, x, v)\, dx \tag{4.39}$$

is time-invariant. (We used these infinite number of conservation laws to determine the long-time behavior in Theorem 3.1.)

The bound (4.37) easily implies a bound on the space average: indeed,

$$\|\langle f\rangle(\tau,\cdot)\|_{\mathcal{C}^\lambda} = \|\langle f\rangle(\tau,\cdot)\|_{\mathcal{Z}^{\lambda,\mu}_\tau} \leqslant \|f(\tau,\cdot)\|_{\mathcal{Z}^{\lambda,\mu}_\tau} = O(1) \quad \text{as } \tau\to\infty; \tag{4.40}$$

and in particular, for $\lambda'<\lambda$,

$$\|\langle \nabla_v f\rangle(\tau,\cdot)\|_{\mathcal{C}^{\lambda'}} = O(1) \quad \text{as } \tau\to\infty. \tag{4.41}$$

Again, (4.37) contains a lot more information than (4.41).

Remark 4.36. The idea to estimate solutions of a non-linear equation by comparison with some unperturbed (reversible) linear dynamics is already present in the definition of Bourgain spaces $X^{s,b}$ [12]. The analogy stops here, since time is a dummy variable in $X^{s,b}$ spaces, while in $\mathcal{Z}^{\lambda,\mu}_t$ spaces it is frozen and appears as a parameter, on which we shall play later.

4.12. Linear damping revisited

As a simple illustration of the functional analysis introduced in this section, let us recast the linear damping (Theorem 3.1) in this language. This will be the first step for the study of the non-linear damping. For simplicity we set $L=1$.

THEOREM 4.37. (Linear Landau damping again) *Let $f^0=f^0(v)$, $W:\mathbb{T}^d\to\mathbb{R}$ with $\|\nabla W\|_{L^1}\leqslant C_W$ and $f_i(x,v)$ be such that*

(i) *condition* **(L)** *from §2.2 holds for some constants $C_0,\lambda,\varkappa>0$;*

(ii) $\|f^0\|_{\mathcal{C}^{\lambda;1}}\leqslant C_0$;

(iii) $\|f_i\|_{\mathcal{Z}^{\lambda,\mu;1}}\leqslant\delta$ *for some $\mu,\delta>0$;*

Then for any $\lambda'<\lambda$ and $\mu'<\mu$, the solution of the linearized Vlasov equation (3.3) *satisfies*

$$\sup_{t\in\mathbb{R}} \|f(t,\cdot)\|_{\mathcal{Z}^{\lambda',\mu';1}_t} \leqslant C\delta \tag{4.42}$$

for some constant $C=C(d,C_W,C_0,\lambda,\lambda',\mu,\mu',\varkappa)$. In particular, $\varrho=\int_{\mathbb{R}^d} f\,dv$ satisfies

$$\sup_{t\in\mathbb{R}} \|\varrho(t,\cdot)\|_{\mathcal{F}^{\lambda'|t|+\mu'}} \leqslant C\delta. \tag{4.43}$$

As a consequence, as $|t|\to\infty$, ϱ converges strongly to

$$\varrho_\infty = \int_{\mathbb{T}^d}\int_{\mathbb{R}^d} f_i(x,v)\,dv\,dx,$$

and f converges weakly to $\langle f_i\rangle=\int_{\mathbb{T}_d} f_i\,dx$, at rate $O(e^{-\lambda''|t|})$ for any $\lambda''<\lambda'$.

If moreover $\|f^0\|_{\mathcal{C}^{\lambda;p}}\leqslant C_0$ and $\|f_i\|_{\mathcal{Z}^{\lambda,\mu;p}}\leqslant\delta$ for all p in some interval $[1,\bar p]$, then (4.42) *can be reinforced into*

$$\sup_{t\in\mathbb{R}} \|f(t,\cdot)\|_{\mathcal{Z}^{\lambda',\mu';p}_t} \leqslant C\delta, \quad 1\leqslant p\leqslant\bar p. \tag{4.44}$$

Remark 4.38. The notions of weak and strong convergence are the same as those in Theorem 3.1. With respect to that statement, we have added an extra analyticity assumption in the x variable; in this linear context this is an overkill (as the proof will show), but later in the non-linear context this will be important.

Proof. Without loss of generality we restrict our attention to $t\geqslant 0$. Although (4.43) follows from (4.42) by Proposition 4.15, we shall establish (4.43) first, and deduce (4.42) due to the equation. We shall write C for various constants depending only on the parameters in the statement of the theorem.

As in the proof of Theorem 3.1, we have

$$\hat{\varrho}(t,k)=\tilde{f}_i(k,kt)+\int_0^t K^0(t-\tau,k)\hat{\varrho}(\tau,k)\,d\tau$$

for any $t\geqslant 0$ and $k\in\mathbb{Z}^d$. By Lemma 3.6, for any $\lambda'<\lambda$ and $\mu'<\mu$,

$$\begin{aligned}
&\sup_{t\geqslant 0}\biggl(\sum_{k\in\mathbb{Z}^d}|\hat{\varrho}(t,k)|e^{2\pi(\lambda' t+\mu')|k|}\biggr)\\
&\qquad\leqslant C(\lambda,\lambda',\varkappa)\biggl(\sum_{k\in\mathbb{Z}^d}e^{-2\pi(\mu-\mu')|k|}\biggr)\sup_{t\geqslant 0}\sup_{k\in\mathbb{Z}^d}|\tilde{f}_i(k,kt)|e^{2\pi(\lambda' t+\mu)|k|}\\
&\qquad\leqslant\frac{C(\lambda,\lambda',\varkappa)}{(\mu-\mu')^d}\sup_{t\geqslant 0}\sum_{k\in\mathbb{Z}^d}|\tilde{f}_i(k,kt)|e^{2\pi(\lambda t+\mu)|k|}.
\end{aligned}$$

Equivalently,

$$\sup_{t\geqslant 0}\|\varrho(t,\cdot)\|_{\mathcal{F}^{\lambda' t+\mu'}}\leqslant C\sup_{t\geqslant 0}\biggl\|\int_{\mathbb{R}^d}f_i\circ S^0_{-t}\,dv\biggr\|_{\mathcal{F}^{\lambda t+\mu}}.\tag{4.45}$$

By Propositions 4.15 and 4.19,

$$\biggl\|\int_{\mathbb{R}^d}f_i\circ S^0_{-t}\,dv\biggr\|_{\mathcal{F}^{\lambda t+\mu}}\leqslant\|f_i\circ S^0_{-t}\|_{\mathcal{Z}^{\lambda,\mu;1}_t}=\|f_i\|_{\mathcal{Z}^{\lambda,\mu;1}_0}\leqslant\delta.$$

This and (4.45) imply (4.43).

To deduce (4.42), we first write

$$f(t,\cdot)=f_i\circ S^0_{-t}+\int_0^t((\nabla W*\varrho_\tau)\circ S^0_{-(t-\tau)})\cdot\nabla_v f^0\,d\tau,$$

where $\varrho_\tau=\varrho(\tau,\cdot)$. Then for any $\lambda''<\lambda'$ we have, by Propositions 4.24 and 4.15, for all $t\geqslant 0$,

$$\begin{aligned}
\|f\|_{\mathcal{Z}^{\lambda'',\mu';1}_t}&\leqslant\|f_i\circ S^0_{-t}\|_{\mathcal{Z}^{\lambda'',\mu;1}_t}+\int_0^t\|(\nabla W*\varrho_\tau)\circ S^0_{-(t-\tau)}\|_{\mathcal{Z}^{\lambda'',\mu';\infty}_t}\|\nabla_v f^0\|_{\mathcal{Z}^{\lambda'',\mu;1}_t}\,d\tau\\
&=\|f_i\|_{\mathcal{Z}^{\lambda'',\mu;1}}+\|\nabla_v f^0\|_{\mathcal{C}^{\lambda'';1}}\int_0^t\|\nabla W*\varrho_\tau\|_{\mathcal{F}^{\lambda''\tau+\mu'}}\,d\tau.
\end{aligned}\tag{4.46}$$

Since $\widehat{\nabla W}(0)=0$, we have, for any $\tau\geqslant 0$,

$$\begin{aligned}\|\nabla W * \varrho_\tau\|_{\mathcal{F}^{\lambda''\tau+\mu}} &\leqslant e^{-2\pi(\lambda''-\lambda')\tau}\|\nabla W * \varrho_\tau\|_{\mathcal{F}^{\lambda'\tau+\mu'}}\\ &\leqslant \|\nabla W\|_{L^1}\, e^{-2\pi(\lambda''-\lambda')\tau}\|\varrho_\tau\|_{\mathcal{F}^{\lambda'\tau+\mu'}}\\ &\leqslant C_W\ C\delta e^{-2\pi(\lambda''-\lambda')\tau};\end{aligned}$$

in particular

$$\int_0^t \|\nabla W * \varrho_\tau\|_{\mathcal{F}^{\lambda''+\mu'}}\, d\tau \leqslant \frac{C\delta}{\lambda''-\lambda'}. \tag{4.47}$$

Also, by Proposition 4.10, for $1<\lambda'/\lambda''\leqslant 2$ we have

$$\|\nabla_v f^0\|_{\mathcal{C}^{\lambda'';1}} \leqslant \frac{C}{\lambda-\lambda''}\,\|f^0\|_{\mathcal{C}^{\lambda;1}} \leqslant \frac{C\,C_0}{\lambda-\lambda''}. \tag{4.48}$$

Plugging (4.47) and (4.48) into (4.46), we deduce (4.42). The end of the proof is an easy exercise if one recalls that $\langle f(t,\cdot)\rangle = \langle f_i\rangle$ for all t. □

5. Deflection estimates

Let a small time-dependent force field, denoted by $\varepsilon F(t,x)$, be given on $\mathbb{T}^d\times\mathbb{R}^d$, whose analytic regularity *improves* linearly in time. (Think of εF as the force created by a damped density.) This force field perturbs the trajectories $S^0_{\tau,t}$ of the free transport (τ is the initial time and t the current time) into trajectories $S_{\tau,t}$. The goal of this section is to get an estimate on the maps $\Omega_{t,\tau}=S_{t,\tau}\circ S^0_{\tau,t}$ (so that $S_{t,\tau}=\Omega_{t,\tau}\circ S^0_{t,\tau}$). These bounds should be in an analytic class about as good as F, with a loss of analyticity depending on ε; they should also be, for $0\leqslant\tau\leqslant t$,

- uniform in $t\geqslant\tau$,
- small as $\tau\to\infty$,
- small as $\tau\to t$.

We shall call $\Omega_{t,\tau}$ the *deflection map* (from time τ to time t). The idea is to compare the free (= without interaction) evolution to the true evolution is at the basis of the interaction representation, wave operators, and most famously the *scattering transforms* (in which the whole evolution from $t\to-\infty$ to $t\to\infty$ is replaced by well-chosen asymptotics). For all these methods which are of constant use in classical and quantum physics one can consult [21].

Remark 5.1. The order of composition in $\Omega_{t,\tau}=S_{t,\tau}\circ S^0_{\tau,t}$ is the only one which leads to useful asymptotics: it is easy to check that in general $S^0_{\tau,t}\circ S_{t,\tau}$ does not converge to anything as $t\to\infty$, even if the force is compactly supported in time.

5.1. Formal expansion

Before stating the main result, we sketch a heuristic perturbation study. Let us write a formal expansion of $V_{0,t}(x,v)$ as a perturbation series:

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

Then we deduce that

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

with $v^{(i)}(t{=}0){=}0$.

So

$$\frac{\partial^2 X_{0,t}}{\partial t^2} = \varepsilon\frac{\partial v^{(1)}}{\partial t}+\varepsilon^2\frac{\partial v^{(2)}}{\partial t}+\dots .$$

On the other hand,

$$\begin{aligned}\varepsilon F(t,X_{0,t}) &= \varepsilon\sum_{k\in\mathbb{Z}^d}\widehat{F}(t,k)e^{2i\pi k\cdot x}e^{2i\pi k\cdot vt}e^{2i\pi k\cdot(\varepsilon\int_0^t v^{(1)}\,ds+\varepsilon^2\int_0^t v^{(2)}\,ds+\dots)}\\ &= \varepsilon\sum_{k\in\mathbb{Z}^d}\widehat{F}(t,k)e^{2i\pi k\cdot x}e^{2i\pi k\cdot vt}\\ &\quad\times\left(1+2i\pi\varepsilon k\cdot\int_0^t v^{(1)}\,ds+2i\pi\varepsilon^2 k\cdot\int_0^t v^{(2)}\,ds-(2\pi)^2\varepsilon^2\left(k\cdot\int_0^t v^{(1)}\,ds\right)^2+\dots\right).\end{aligned}$$

By successive identification,

$$\begin{aligned}\frac{\partial v^{(1)}}{\partial t} &= \sum_{k\in\mathbb{Z}^d}\widehat{F}(t,k)e^{2i\pi k\cdot x}e^{2i\pi k\cdot vt},\\ \frac{\partial v^{(2)}}{\partial t} &= \sum_{k\in\mathbb{Z}^d}\widehat{F}(t,k)e^{2i\pi k\cdot x}e^{2i\pi k\cdot vt}2i\pi k\cdot\int_0^t v^{(1)}\,ds,\\ \frac{\partial v^{(3)}}{\partial t} &= \sum_{k\in\mathbb{Z}^d}\widehat{F}(t,k)e^{2i\pi k\cdot x}e^{2i\pi k\cdot vt}\left(2i\pi k\cdot\int_0^t v^{(2)}\,ds-(2\pi)^2\varepsilon^2\left(k\cdot\int_0^t v^{(1)}\,ds\right)^2\right),\end{aligned}$$

etc.

In particular notice that

$$\left|\frac{\partial v^{(1)}}{\partial t}\right| \leqslant \sum_{k\in\mathbb{Z}^d}|\widehat{F}(t,k)|,$$

so

$$\begin{aligned}\int_0^\infty\left|\frac{\partial v^{(1)}}{\partial t}\right|dt &\leqslant \int_0^\infty\sum_{k\in\mathbb{Z}^d}|\widehat{F}(t,k)|\,dt\\ &\leqslant \int_0^\infty\sum_{k\in\mathbb{Z}^d}|\widehat{F}(t,k)|e^{2\pi\mu t}e^{-2\pi\mu t}\,dt \leqslant C_F\int_0^\infty e^{-2\pi\mu t}\,dt = \frac{C_F}{2\pi\mu}.\end{aligned}$$

Thus, under our uniform analyticity assumptions we expect $V_{0,t}(x,v)$ to be a *uniformly bounded* analytic perturbation of v.

5.2. Main result

On $\mathbb{T}^d_x$ we consider the dynamical system

$$\frac{d^2X}{dt^2} = \varepsilon F(t, X);$$

its phase space is $\mathbb{T}^d \times \mathbb{R}^d$. Although this system is reversible, we shall only consider $t \geqslant 0$. The parameter ε is here only to recall the perturbative nature of the estimate.

For any $(x, v) \in \mathbb{T}^d \times \mathbb{R}^d$ and any two times $\tau, t \in \mathbb{R}_+$, let $S_{\tau,t}$ be the transform mapping the state of the system at time τ to the state of the system at time t. In more precise terms, $S_{\tau,t}$ is described by the equations

$$S_{\tau,t}(x,v) = (X_{\tau,t}(x,v), V_{\tau,t}(x,v)), \quad X_{\tau,\tau}(x,v) = x, \quad V_{\tau,\tau}(x,v) = v,$$

$$\frac{d}{dt} X_{\tau,t}(x,v) = V_{\tau,t}(x,v) \quad \text{and} \quad \frac{d}{dt} V_{\tau,t}(x,v) = \varepsilon\, F(t, X_{\tau,t}(x,v)). \tag{5.1}$$

From the definition we have the composition identity

$$S_{t_2,t_3} \circ S_{t_1,t_2} = S_{t_1,t_3}; \tag{5.2}$$

in particular $S_{t,\tau}$ is the inverse of $S_{\tau,t}$.

We also write $S^0_{\tau,t}$ for the same transform in the case of the free dynamics ($\varepsilon = 0$); in this case there is an explicit expression:

$$S^0_{\tau,t}(x,v) = (x + v(t-\tau), v), \tag{5.3}$$

where $x + v(t-\tau)$ is evaluated modulo $\mathbb{Z}^d$. Finally, we define the deflection map associated with εF:

$$\Omega_{t,\tau} = S_{t,\tau} \circ S^0_{\tau,t}. \tag{5.4}$$

(There is no simple semigroup property for the transforms $\Omega_{t,\tau}$.)

In this section we establish the following estimates.

THEOREM 5.2. (Analytic estimates on deflection in hybrid norms) *Let $\varepsilon > 0$ and let $F = F(t,x)$ on $\mathbb{R}_+ \times \mathbb{T}^d$ satisfy*

$$\widehat{F}(t,0) = 0 \quad \textit{and} \quad \sup_{t \geqslant 0} (\|F(t,\cdot)\|_{\mathcal{F}^{\lambda t+\mu}} + \|\nabla_x F(t,\cdot)\|_{\mathcal{F}^{\lambda t+\mu}}) \leqslant C_F \tag{5.5}$$

for some parameters $\lambda, \mu > 0$ and $C_F > 0$. Let $t \geqslant \tau \geqslant 0$ and let

$$\Omega_{t,\tau} = (\Omega X_{t,\tau}, \Omega V_{t,\tau})$$

be the deflection associated with εF. Let $0\leqslant\lambda'<\lambda$, $0\leqslant\mu'<\mu$ and $\tau'\geqslant 0$ be such that

$$\lambda'(\tau'-\tau)\leqslant\tfrac{1}{2}(\mu-\mu'). \tag{5.6}$$

Let

$$\begin{cases} R_1(\tau,t)=C_F e^{-2\pi(\lambda-\lambda')\tau}\min\{t-\tau,(2\pi(\lambda-\lambda'))^{-1}\},\\ R_2(\tau,t)=C_F e^{-2\pi(\lambda-\lambda')\tau}\min\{\frac{1}{2}(t-\tau)^2,(2\pi(\lambda-\lambda'))^{-2}\}.\end{cases}$$

Assume that

$$\varepsilon R_2(\tau,t)\leqslant\tfrac{1}{4}(\mu-\mu')\quad \textit{for all } 0\leqslant\tau\leqslant t, \tag{5.7}$$

and

$$\varepsilon C_F\leqslant 2\pi^2(\lambda-\lambda')^2. \tag{5.8}$$

Then

$$\|\Omega X_{t,\tau}-\mathrm{Id}\,\|_{\mathcal{Z}^{\lambda',\mu'}_{\tau'}}\leqslant 2\varepsilon R_2(\tau,t)\quad \textit{for all } 0\leqslant\tau\leqslant t, \tag{5.9}$$

and

$$\|\Omega V_{t,\tau}-\mathrm{Id}\,\|_{\mathcal{Z}^{\lambda',\mu'}_{\tau'}}\leqslant\varepsilon R_1(\tau,t)\quad \textit{for all } 0\leqslant\tau\leqslant t. \tag{5.10}$$

Remark 5.3. The proof of Theorem 5.2 is easily adapted to include Sobolev corrections. It is important to note that the deflection map is smooth uniformly in time, not just in gliding regularity ($\tau'=0$ is admissible in (5.6)).

Proof. For a start, let us make the ansatz

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

with

$$Z_{t,t}(x,v)=0\quad\text{and}\quad\partial_\tau Z_{t,\tau}|_{\tau=t}(x,v)=0.$$

Then it is easily checked that

$$\Omega_{t,\tau}-\mathrm{Id}=\varepsilon\,(Z,\partial_\tau Z)\circ S^0_{t-\tau};$$

in particular

$$\|\Omega_{t,\tau}-\mathrm{Id}\,\|_{\mathcal{Z}^{\lambda',\mu'}_{\tau'}}=\varepsilon\,\|(Z,\partial_\tau Z)\|_{\mathcal{Z}^{\lambda',\mu'}_{t+\tau'-\tau}}.$$

To estimate this we shall use a fixed-point argument based on the equation for $S_{t,\tau}$, namely

$$\frac{d^2X_{t,\tau}}{d\tau^2}=\varepsilon F(\tau,X_{t,\tau}),$$

or equivalently

$$\frac{d^2Z_{t,\tau}}{d\tau^2}=F(\tau,x-v(t-\tau)+\varepsilon Z_{t,\tau}).$$

So let us fix t and define

$$\Psi\colon (W_{t,\tau})_{0\leqslant\tau\leqslant t} \longmapsto (Z_{t,\tau})_{0\leqslant\tau\leqslant t}$$

such that $(Z_{t,\tau})_{0\leqslant\tau\leqslant t}$ is the solution of

$$\begin{cases} \dfrac{\partial^2 Z_{t,\tau}}{\partial\tau^2} = F(\tau, x-v(t-\tau)+\varepsilon W_{t,\tau}), \\ Z_{t,t} = 0, \\ (\partial_\tau Z_{t,\tau})|_{\tau=t} = 0. \end{cases} \tag{5.11}$$

What we are after is an estimate of the fixed point of Ψ. We do this in two steps.

Step 1. *Estimate of* $\Psi(0)$. Let $Z^0=\Psi(0)$. By integration of (5.11) (for $W=0$) we have

$$Z^0_{t,\tau} = \int_\tau^t (s-\tau)F(s, x-v(t-s))\,ds.$$

Let σ be such that $\lambda'\sigma\leqslant\frac{1}{2}(\mu-\mu')$. Applying the $\mathcal{Z}^{\lambda',\mu'}_{t+\sigma}$ norm and using Proposition 4.19, we get

$$\|Z^0_{t,\tau}\|_{\mathcal{Z}^{\lambda',\mu'}_{t+\sigma}} \leqslant \int_\tau^t (s-\tau)\|F(s,\cdot)\|_{\mathcal{Z}^{\lambda',\mu'}_{s+\sigma}}\,ds = \int_\tau^t (s-\tau)\|F(s,\cdot)\|_{\mathcal{F}^{\lambda' s+\lambda'\sigma+\mu'}}\,ds.$$

Of course $\lambda'\sigma+\mu'\leqslant\mu$, so in particular

$$\lambda' s+\lambda'\sigma+\mu' \leqslant -(\lambda-\lambda')s+\lambda s+\mu.$$

Combining this with the assumption $\widehat{F}(s,0)=0$ yields

$$\|F(s,\cdot)\|_{\mathcal{F}^{\lambda' s+\lambda'\sigma+\mu'}} \leqslant \|F(s,\cdot)\|_{\mathcal{F}^{\lambda s+\mu}} e^{-2\pi(\lambda-\lambda')s} \leqslant C_F e^{-2\pi(\lambda-\lambda')s}.$$

So

$$\begin{aligned} \|Z^0_{t,\tau}\|_{\mathcal{Z}^{\lambda',\mu'}_{t+\sigma}} &\leqslant C_F \int_\tau^t (s-\tau)e^{-2\pi(\lambda-\lambda')s}\,ds \\ &\leqslant C_F e^{-2\pi(\lambda-\lambda')\tau} \min\left\{\frac{(t-\tau)^2}{2}, \frac{1}{(2\pi(\lambda-\lambda'))^2}\right\} \leqslant R_2(\tau,t). \end{aligned}$$

With t still fixed, we define the norm

$$|\!|\!|(Z_{t,\tau})_{0\leqslant\tau\leqslant t}|\!|\!| := \sup\left\{\frac{\|Z_{t,\tau}\|_{\mathcal{Z}^{\lambda',\mu'}_{t+\sigma}}}{R_2(\tau,t)} : 0\leqslant\tau\leqslant t, \sigma+t\geqslant 0 \text{ and } \lambda'\sigma\leqslant\frac{\mu-\mu'}{2}\right\}. \tag{5.12}$$

The above estimates show that $|||\Psi(0)|||\leqslant 1$. (Since $t+(\tau'-\tau)\geqslant t-\tau\geqslant 0$, we may assume that $t+\sigma\geqslant 0$, and we aim at finally choosing $\sigma=\tau'-\tau$.)

Step 2. *Lipschitz constant of* Ψ. We shall prove that under our assumptions, Ψ is $\frac{1}{2}$-Lipschitz on the ball $B(0,2)$ in the norm $|||\cdot|||$. Let $W,\widetilde{W}\in B(0,2)$, $Z=\Psi(W)$ and $\widetilde{Z}=\Psi(\widetilde{W})$. By solving the differential inequality for $Z-\widetilde{Z}$, we get

$$Z_{t,\tau}-\widetilde{Z}_{t,\tau}=\varepsilon(W_{t,s}-\widetilde{W}_{t,s})\cdot\int_0^1\int_\tau^t(s-\tau)\nabla_x F(s,x-v(t-s)+\varepsilon(\theta W_{t,s}+(1-\theta)\widetilde{W}_{t,s}))\,ds\,d\theta.$$

We divide by $R_2(\tau,t)$, take the $\mathcal{Z}$ norm and note that $R_2(s,t)\leqslant R_2(\tau,t)$ to obtain

$$|||(Z_{t,\tau}-\widetilde{Z}_{t,\tau})_{0\leqslant\tau\leqslant t}|||\leqslant\varepsilon|||(W_{t,s}-\widetilde{W}_{t,s})_{0\leqslant s\leqslant t}|||A(t)$$

with

$$A(t)=\sup_{\sigma,\tau}\int_0^1\int_\tau^t(s-\tau)\|\nabla_x F(s,x-v(t-s)+\varepsilon(\theta W_{t,s}+(1-\theta)\widetilde{W}_{t,s}))\|_{\mathcal{Z}^{\lambda',\mu'}_{t+\sigma}}\,ds\,d\theta.$$

By Proposition 4.25 (composition inequality),

$$A(t)\leqslant\int_\tau^t(s-\tau)\|\nabla_x F(s,\cdot)\|_{\mathcal{Z}^{\lambda',\mu'+e(t,s,\sigma)}_{s+\sigma}}\,ds=\int_\tau^t(s-\tau)\|\nabla_x F(s,\cdot)\|_{\mathcal{F}^{\lambda' s+\lambda'\sigma+\mu'+e(t,s,\sigma)}}\,ds,$$

with

$$e(t,s,\sigma):=\varepsilon\|\theta W_{t,s}+(1-\theta)\widetilde{W}_{t,s}\|_{\mathcal{Z}^{\lambda',\mu'}_{t+\sigma}}\leqslant 2\varepsilon R_2(s,t)\leqslant 2\varepsilon R_2(\tau,t).$$

Using (5.7), we get

$$\lambda' s+\lambda'\sigma+\mu'+e(s,t,\sigma)\leqslant\lambda' s+\lambda'\sigma+\mu'+2\varepsilon R_2(\tau,t)\leqslant\lambda' s+\mu=(\lambda s+\mu)-(\lambda-\lambda')s.$$

By again using the bound on $\nabla_x F$ and the assumption $\widehat{F}(s,0)=0$, we deduce that

$$A(t)\leqslant\sup_\tau\int_\tau^t(s-\tau)C_F e^{-2\pi(\lambda-\lambda')s}\,ds\leqslant R_2(0,t)\leqslant\frac{C_F}{4\pi^2(\lambda-\lambda')^2}.$$

Using (5.8), we conclude that

$$|||(Z_{t,\tau}-\widetilde{Z}_{t,\tau})_{0\leqslant\tau\leqslant t}|||\leqslant\tfrac{1}{2}|||(W_{t,s}-\widetilde{W}_{t,s})_{0\leqslant s\leqslant t}|||.$$

So Ψ is $\frac{1}{2}$-Lipschitz on $B(0,2)$, and we can conclude the proof of (5.9) by applying Theorem A.2 and choosing $\sigma=\tau'-\tau$.

It remains to control the velocity component of Ω, i.e., establish (5.10); this will follow from the control of the position component. Indeed, if we write

$$Q_{t,\tau}=\varepsilon^{-1}(\Omega V_{t,\tau}-\mathrm{Id})(x,v),$$

we have

$$Q_{t,\tau} = \int_\tau^t F(s, x - v(t-s) + \varepsilon W_{t,s})\, ds,$$

so we can estimate as before

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

to get

$$\|Q_{t,\tau}\|_{\mathcal{Z}^{\lambda',\mu'}_{t+(\tau'-\tau)}} \leqslant \int_\tau^t C_F e^{-2\pi(\lambda-\lambda')s}\, ds \leqslant R_1(\tau,t).$$

Thus the proof is complete. □

Remark 5.4. Instead of directly studying $S_{t,\tau} \circ S^0_{\tau,t}$, one can study $S^0_{t,\tau} \circ S_{\tau,t}$ and then invert with the help of Proposition 4.28 (inversion inequality); in the end the results are similar.

Remark 5.5. Loss and Bernard independently suggested to compare the estimates in the present section with the Nekhoroshev theorem in dynamical systems theory [71], [72]. The latter theorem roughly states that for a perturbation of a completely integrable system, trajectories remain close to those of the unperturbed system for a time growing exponentially in the inverse of the size of the perturbation (unlike KAM theory, this result is not global in time; but it is more general in the sense that it also applies outside invariant tori). In the present setting the situation is better since the perturbation decays.

6. Bilinear regularity and decay estimates

To introduce this crucial section, let us reproduce and improve a key computation from §3. Let G be a function of v, and R a time-dependent function of x with $\widehat{R}(0)=0$; both G and R are vector-valued in $\mathbb{R}^d$. (Think of $G(v)$ as $\nabla_v f(v)$ and of $R(\tau,x)$ as $\nabla W * \varrho(\tau,x)$.) From now on we shall always use the supremum norm over the coordinates in the following sense

$$\|F\|_{\mathcal{Z}} := \max_{1\leqslant j\leqslant d} \|F_j\|_{\mathcal{Z}} \quad \text{and} \quad \|F\|_{\mathcal{F}} := \max_{1\leqslant j\leqslant d} \|F_j\|_{\mathcal{F}}.$$

Let further

$$\sigma(t,x) = \int_0^t \int_{\mathbb{R}^d} G(v)\cdot R(\tau, x - v(t-\tau))\, dv\, d\tau.$$

Then

$$\begin{aligned}\hat\sigma(t,k) &= \int_0^t \int_{\mathbb{T}^d}\int_{\mathbb{R}^d} G(v)\cdot R(\tau, x-v(t-\tau)) e^{-2i\pi k\cdot x}\,dv\,dx\,d\tau \\ &= \int_0^t \int_{\mathbb{T}^d}\int_{\mathbb{R}^d} G(v)\cdot R(\tau,x) e^{-2i\pi k\cdot x} e^{-2i\pi k\cdot v(t-\tau)}\,dv\,dx\,d\tau \\ &= \int_0^t \widetilde{G}(k(t-\tau))\cdot \widehat{R}(\tau,k)\,d\tau.\end{aligned}$$

Let us assume that G has a "high" gliding analytic regularity $\bar\lambda$, and estimate σ in regularity λt, with $\lambda<\bar\lambda$. Let $\alpha=\alpha(t,\tau)$ satisfy

$$0\leqslant \alpha(t,\tau)\leqslant (\bar\lambda-\lambda)\,(t-\tau).$$

Then, with $\mathbb{Z}^d_*=\mathbb{Z}^d\setminus\{0\}$,

$$\begin{aligned}\|\sigma(t)\|_{\mathcal{F}^{\lambda t}} &\leqslant d\sum_{k\in Z^d_*}\int_0^t e^{2\pi\lambda t|k|}|\widetilde{G}(k(t-\tau))|\,|\widehat{R}(\tau,k)|\,d\tau \\ &\leqslant d\int_0^t \Big(\sup_{k\in\mathbb{Z}^d_*}|\widetilde{G}(k(t-\tau))|e^{2\pi(\lambda(t-\tau)+\alpha)|k|}\Big)\Big(\sum_{k\in\mathbb{Z}^d} e^{2\pi(\lambda\tau-\alpha)|k|}|\widehat{R}(\tau,k)|\Big)\,d\tau \\ &\leqslant d\Big(\sup_{\eta\in\mathbb{Z}^d} e^{2\pi\bar\lambda|\eta|}|\widetilde{G}(\eta)|\Big)\Big(\sup_{0\leqslant\tau\leqslant t}\|R(\tau,\cdot)\|_{\mathcal{F}^{\lambda\tau-\alpha}}\Big)\int_0^t e^{-2\pi((\bar\lambda-\lambda)(t-\tau)-\alpha)}\,d\tau,\end{aligned}$$

where we have used that

$$k\in\mathbb{Z}^d_* \quad\Longrightarrow\quad 2\pi(\lambda(t-\tau)+\alpha)|k|\leqslant 2\pi\bar\lambda|k|(t-\tau)-((\bar\lambda-\lambda)(t-\tau)-\alpha).$$

Let us choose

$$\alpha(t,\tau)=\tfrac12(\bar\lambda-\lambda)\min\{1,t-\tau\}.$$

Then

$$\int_0^t e^{-2\pi((\bar\lambda-\lambda)(t-\tau)-\alpha)}\,d\tau\leqslant\int_0^t e^{-\pi(\bar\lambda-\lambda)(t-\tau)}\,d\tau\leqslant\frac{1}{\pi(\bar\lambda-\lambda)}.$$

So in the end

$$\|\sigma(t)\|_{\mathcal{F}^{\lambda t}}\leqslant\frac{d\|G\|_{\mathcal{X}^{\bar\lambda}}}{\pi(\bar\lambda-\lambda)}\sup_{0\leqslant\tau\leqslant t}\|R(\tau)\|_{\mathcal{F}^{\lambda\tau-\alpha(t,\tau)}},$$

where $\|G\|_{\mathcal{X}^{\bar\lambda}}=\sup_{\eta\in\mathbb{Z}^d}|\widetilde{G}(\eta)|e^{2\pi\bar\lambda|\eta|}$.

In the preceding computation there are three important things to notice, which lie at the heart of Landau damping:

• The natural index of analytic regularity of σ in x increases linearly in time: this is an automatic consequence of the gliding regularity, already observed in §4.

• A bit $\alpha(t,\tau)$ of analytic regularity of G was transferred from G to R, however no more than a fraction of $(\bar\lambda-\lambda)(t-\tau)$. We call this the *regularity extortion*: if f *forces* $\bar f$, i.e. if it satisfies an equation of the form $\partial_t f+v\cdot\nabla_x f+F[f]\cdot\nabla_v\bar f=S$, then $\bar f$ will give away some (gliding) smoothness to $\varrho=\int_{\mathbb{R}^d} f\,dv$.

• The combination of higher regularity of G and the assumption $\widehat{R}(0)=0$ has been converted into a time-decay, so that the time-integral is bounded, uniformly as $t\to\infty$. Thus there is *decay by regularity*.

The main goal of this section is to establish quantitative variants of these effects in some general situations when G is not only a function of v and R not only a function of t and x. Note that we shall have to work with regularity indices depending on t and τ!

Regularity extortion is related to velocity-averaging regularity, well known in kinetic theory [43]; what is unusual though is that we are working in analytic regularity, and in large time, while velocity-averaging regularity is mainly a short-time effect. In fact we shall study two distinct mechanisms for the extortion: the first one will be well suited for short times ($t-\tau$ small), and will be crucial later to get rid of small deteriorations in the functional spaces due to composition; the second one will be well adapted to large times ($t-\tau\to\infty$) and will ensure convergence of the time-integrals.

The estimates in this section lead to a serious twist on the popular view on Landau damping, according to which the wave gives energy to the particles that it forces; instead, the picture here is that the wave gains regularity from the background, and regularity is converted into decay.

For the sake of pedagogy, we shall first establish the basic, simple bilinear estimate, and then discuss the two mechanisms once at a time.

6.1. Basic bilinear estimate

PROPOSITION 6.1. (Basic bilinear estimate in gliding regularity) *Let $G=G(\tau,x,v)$ and $R=R(\tau,x,v)$ be valued in $\mathbb{R}^d$,*

$$\beta(\tau,x)=\int_{\mathbb{R}^d}(G\cdot R)(\tau,x-v(t-\tau),v)\,dv,$$
$$\sigma(t,x)=\int_0^t\beta(\tau,x)\,d\tau.$$

Then

$$\|\beta(\tau,\cdot)\|_{\mathcal{F}^{\lambda t+\mu}}\leqslant d\|G\|_{\mathcal{Z}_\tau^{\lambda,\mu;1}}\|R\|_{\mathcal{Z}_\tau^{\lambda,\mu}};\tag{6.1}$$

and

$$\|\sigma(t,\cdot)\|_{\mathcal{F}^{\lambda t+\mu}}\leqslant d\int_0^t\|G\|_{\mathcal{Z}_\tau^{\lambda,\mu;1}}\|R\|_{\mathcal{Z}_\tau^{\lambda,\mu}}\,d\tau.\tag{6.2}$$

Proof. Obviously (6.2) follows from (6.1). To prove (6.1) we apply successively Propositions 4.15, 4.19 and 4.24, obtaining

$$\begin{aligned}\|\beta(\tau,\cdot)\|_{\mathcal{F}^{\lambda t+\mu}} &\leqslant \left\|\int_{\mathbb{R}^d}(G\cdot R)\circ S^0_{\tau-t}\,dv\right\|_{\mathcal{F}^{\lambda t+\mu}} \leqslant \|(G\cdot R)\circ S^0_{\tau-t}\|_{\mathcal{Z}^{\lambda,\mu;1}_t}\\ &= \|G\cdot R\|_{\mathcal{Z}^{\lambda,\mu;1}_\tau} \leqslant d\|G\|_{\mathcal{Z}^{\lambda,\mu;1}_\tau}\|R\|_{\mathcal{Z}^{\lambda,\mu}_\tau}.\end{aligned}$$

□

6.2. Short-term regularity extortion by time cheating

PROPOSITION 6.2. (Short-term regularity extortion) *Let $G{=}G(x,v)$ and $R{=}R(x,v)$ be valued in $\mathbb{R}^d$, and*

$$\beta(x) = \int_{\mathbb{R}^d}(G\cdot R)(x-v(t-\tau),v)\,dv.$$

Then for any $\lambda,\mu,t\geqslant 0$ and any $b>-1$, we have

$$\|\beta\|_{\mathcal{F}^{\lambda t+\mu}} \leqslant d\|G\|_{\mathcal{Z}^{\lambda(1+b),\mu;1}_{\tau-bt/(1+b)}}\|R\|_{\mathcal{Z}^{\lambda(1+b),\mu}_{\tau-bt/(1+b)}}. \tag{6.3}$$

Moreover, if P_k stands for the projection on the k-th Fourier mode as in (4.36)*, one has*

$$e^{2\pi(\lambda t+\mu)|k|}|\hat\beta(k)| \leqslant d\sum_{l\in\mathbb{Z}^d}\|P_lG\|_{\mathcal{Z}^{\lambda(1+b),\mu;1}_{\tau-bt/(1+b)}}\|P_{k-l}R\|_{\mathcal{Z}^{\lambda(1+b),\mu}_{\tau-bt/(1+b)}}. \tag{6.4}$$

Remark 6.3. If R only depends on t and x, then the norm of R in the right-hand side of (6.3) is $\|R\|_{\mathcal{F}^\nu}$ with

$$\nu = \lambda(1+b)\left|\tau-\frac{bt}{1+b}\right|+\mu = (\lambda\tau+\mu)-b(t-\tau),$$

as soon as $\tau\geqslant bt/(1+b)$. Thus, some regularity has been gained with respect to Proposition 6.1. Even if R is not a function of t and x alone, but rather a function of t and x *composed* with a function depending on all the variables, this gain will be preserved through the composition inequality.

Proof. By applying successively Propositions 4.15 (iv), 4.19 and 4.24 (i), we get

$$\begin{aligned}\left\|\int_{\mathbb{R}^d}(G\cdot R)(x-v(t-\tau),v)\,dv\right\|_{\mathcal{F}^{\lambda t+\mu}} &\leqslant \|(G\cdot R)(x-v(t-\tau,v))\|_{\mathcal{Z}^{\lambda(1+b),\mu;1}_{t/(1+b)}}\\ &= \|G\cdot R\|_{\mathcal{Z}^{\lambda(1+b),\mu;1}_{\tau-bt/(1+b}}\\ &\leqslant d\|G\|_{\mathcal{Z}^{\lambda(1+b),\mu;1}_{\tau-bt/(1+b)}}\|R\|_{\mathcal{Z}^{\lambda(1+b),\mu;\infty}_{\tau-bt/(1+b)}},\end{aligned}$$

which is the desired result (6.3).

Inequality (6.4) is obtained in a similar way with the help of Proposition 4.34. □

Remark 6.4. Let us sketch an alternative proof of Proposition 6.2, which is longer but has the interest to rely on commutators involving ∇_v, ∇_x and the transport semigroup, all of them classically related to hypoelliptic regularity and velocity averaging. Let $S=S^0_{\tau-t}$, so that $R\circ S(x,v)=R(x-v(t-\tau),v)$; and let

$$D=D_{\tau,t,b}:=(\tau-b(t-\tau))\nabla_x+(1+b)\nabla_v.$$

Then, by direct computation,

$$t\nabla_x(R\circ S)=(DR)\circ S-(1+b)\nabla_v(R\circ S); \tag{6.5}$$

and since ∇_x commutes with ∇_v and D, and with the composition by S as well, this can be generalized by induction into

$$(t\nabla_x)^n(R\circ S)=\sum_{\substack{m\in\mathbb{N}_0^d\\ m\leqslant n}}\binom{n}{m}[-(1+b)\nabla_v]^m((D^{n-m}R)\circ S). \tag{6.6}$$

Applying this formula with R replaced by $G\cdot R$ and integrating in v yields

$$(t\nabla_x)^n\int_{\mathbb{R}^d}(G\cdot R)\circ S^0_{\tau-t}\,dv=\int_{\mathbb{R}^d}D^n(G\cdot R)\circ S^0_{\tau-t}\,dv=\int_{\mathbb{R}^d}D^n(G\cdot R)\,dv.$$

It follows by taking Fourier transform that

$$(2i\pi tk)^n\hat\beta(k)=\int_{\mathbb{R}^d}[D^n(G\cdot R)]\hat{}\,dv=\int_{\mathbb{R}^d}((1+b)\nabla_v+2i\pi(\tau-b(t-\tau))k)^n(\widehat{G\cdot R})(k,v)\,dv,$$

whence

$$\begin{aligned}\sum_{\substack{k\in\mathbb{Z}^d\\ n\in\mathbb{N}_0^d}}&e^{2\pi\mu|k|}\frac{|2\pi\lambda tk|^n}{n!}|\hat\beta(k)|\\ &\leqslant\sum_{\substack{k\in\mathbb{Z}^d\\ n\in\mathbb{N}_0^d}}e^{2\pi\mu|k|}\frac{(\lambda(1+b))^n}{n!}\left\|\left[\nabla_v+2i\pi\left(\tau-\frac{bt}{1+b}\right)k\right]^n(\widehat{G\cdot R})(k,v)\right\|_{L^1(dv)}\\ &=\|G\cdot R\|_{\mathcal{Z}^{\lambda(1+b),\mu;1}_{\tau-bt/(1+b)}},\end{aligned}$$

and then (6.3) follows by Proposition 4.24.

Let us conclude this subsection with some comments on Proposition 6.2. When we wish to apply it, what constraints on $b(t,\tau)$ (assumed to be non-negative to fix the ideas) does this presuppose? First, b should be small, so that $\lambda(1+b)\leqslant\bar\lambda$. But most importantly, we have estimated G_τ in a norm $\mathcal{Z}_{\tau'}$ instead of $\mathcal{Z}_\tau$ (this is the time cheating), where

$|\tau'-\tau|=bt/(1+b)$. To compensate for this discrepancy, we may apply (4.19), but for this to work $bt/(1+b)$ should be small, otherwise we would lose a large index of analyticity in x, or at best we would inherit an undesirable exponentially growing constant. So all we are allowed is $b(t,\tau)=O(1/(1+t))$. This is not enough to get the time-decay which would lead to Landau damping. Indeed, if $R=R(x)$ with $\widehat{R}(0)=0$, then

$$\|R\|_{\mathcal{Z}^{\lambda(1+b),\mu}_{\tau-bt/(1+b)}} = \|R\|_{\mathcal{F}^{\lambda\tau+\mu-\lambda b(t-\tau)}} \leqslant e^{-\lambda b(t-\tau)}\|R\|_{\mathcal{F}^{\lambda\tau+\mu}};$$

so we gain a coefficient $e^{-\lambda b(t-\tau)}$, but then

$$\int_0^t e^{-\lambda b(t-\tau)}\,d\tau \geqslant \int_0^t e^{-\lambda\varepsilon(t-\tau)/t}\,d\tau = \left(\frac{1-e^{-\lambda\varepsilon}}{\lambda\varepsilon}\right)t,$$

which of course diverges in large time.

To summarize: Proposition 6.2 is helpful when $t-\tau=O(1)$, or when some extra time-decay is available. This will already be very useful; but for long-time estimates we need another, complementary mechanism.

6.3. Long-term regularity extortion

To search for the extra decay, let us refine the computation of the beginning of this section. Assume that $G_\tau=\nabla_v g_\tau$, where $(g_\tau)_{\tau\geqslant 0}$ solves a transport-like equation, so $\widetilde{G}(\tau,k,\eta)=2i\pi\eta\tilde{g}(\tau,k,\eta)$, and

$$|\widetilde{G}(\tau,k,\eta)| \lesssim 2\pi|\eta|e^{-2\pi\bar{\mu}|k|}e^{-2\pi\bar{\lambda}|\eta+k\tau|}.$$

Up to slightly increasing $\bar{\lambda}$ and $\bar{\mu}$, we may assume that

$$|\widetilde{G}(\tau,k,\eta)| \lesssim (1+\tau)e^{-2\pi\bar{\mu}|k|}e^{-2\pi\bar{\lambda}|\eta+k\tau|}. \tag{6.7}$$

Let then $\varrho(\tau,x)=\int_{\mathbb{R}^d} f(\tau,x,v)\,dv$, where also f solves a transport equation, but has a lower analytic regularity; and $R=\nabla W*\varrho$. Assuming that $|\widehat{\nabla W}(k)|=O(|k|^{-\gamma})$ for some $\gamma\geqslant 0$, we have

$$|\widehat{R}(\tau,k)| \lesssim \frac{e^{-2\pi(\lambda\tau+\mu)|k|}1_{k\neq 0}}{1+|k|^\gamma}. \tag{6.8}$$

Let again

$$\sigma(t,x) = \int_0^t \int_{\mathbb{R}^d} G(\tau,x-v(t-\tau),v)\cdot R(\tau,x-v(t-\tau))\,dv\,d\tau.$$

As $t\to\infty$, G in the integrand of σ oscillates wildly in phase space, so it is not clear that it will help at all. But let us compute

$$\begin{aligned}
\hat\sigma(t,k) &= \int_0^t \int_{\mathbb{T}^d}\int_{\mathbb{R}^d} G(\tau,x-v(t-\tau),v)\cdot R(\tau,x-v(t-\tau))e^{-2i\pi k\cdot x}\,dv\,dx\,d\tau\\
&= \int_0^t \int_{\mathbb{T}^d}\int_{\mathbb{R}^d} G(\tau,x,v)\cdot R(\tau,x)e^{-2i\pi k\cdot x}e^{-2i\pi k\cdot v(t-\tau)}\,dv\,dx\,d\tau\\
&= \int_0^t \int_{\mathbb{R}^d} \widehat{G\cdot R}(\tau,k,v)e^{-2i\pi k\cdot v(t-\tau)}\,dv\,d\tau\\
&= \int_0^t \int_{\mathbb{R}^d} \sum_{l\in\mathbb{Z}^d} \widehat{G}(\tau,l,v)\cdot\widehat{R}(\tau,k-l)e^{-2i\pi k\cdot v(t-\tau)}\,dv\,d\tau\\
&= \int_0^t \sum_{l\in\mathbb{Z}^d} \widetilde{G}(\tau,l,k(t-\tau))\cdot\widehat{R}(\tau,k-l)\,d\tau.
\end{aligned}$$

At this level, the difference with respect to the beginning of this section lies in the fact that there is a summation over $l\in\mathbb{Z}^d$, instead of just choosing $l=0$. Note that

$$\hat\sigma(t,0) = \int_0^t \int_{\mathbb{T}^d}\int_{\mathbb{R}^d} G(\tau,x,v)\cdot R(\tau,x)\,dv\,dx\,d\tau = 0,$$

because G is a v-gradient.

From (6.7) and (6.8) we deduce that

$$\begin{aligned}
&\sum_{k\in\mathbb{Z}^d} |\hat\sigma(t,k)|e^{2\pi(\lambda t+\mu)|k|}\\
&\qquad\lesssim \int_0^t (1+\tau) \sum_{\substack{k,l\in\mathbb{Z}^d\\ 0\neq k\neq l}} e^{2\pi\mu|k|}e^{2\pi\lambda t|k|}e^{-2\pi\bar\mu|l|}e^{-2\pi\bar\lambda|k(t-\tau)+l\tau|}e^{-2\pi\mu|k-l|}\frac{e^{-2\pi\lambda\tau|k-l|}}{1+|k-l|^\gamma}\,d\tau.
\end{aligned}$$

Using the inequalities

$$e^{-2\pi\mu|k-l|}e^{2\pi\mu|k|}e^{-2\pi\bar\mu|l|} \leqslant e^{-2\pi(\bar\mu-\mu)|l|}$$

and

$$e^{-2\pi\lambda\tau|k-l|}e^{2\pi\lambda t|k|}e^{-2\pi\bar\lambda|k(t-\tau)+l\tau|} \leqslant e^{-2\pi(\bar\lambda-\lambda)|k(t-\tau)+l\tau|},$$

we end up with

$$\|\sigma(t)\|_{\mathcal{F}^{\lambda t+\mu}} \lesssim \sum_{\substack{k,l\in\mathbb{Z}^d\\ 0\neq k\neq l}} \frac{e^{-2\pi(\bar\mu-\mu)|l|}}{1+|k-l|^\gamma}\int_0^t e^{-2\pi(\bar\lambda-\lambda)|k(t-\tau)+l\tau|}(1+\tau)\,d\tau.$$

If it were not for the negative exponential, the time-integral would be $O(t^2)$ as $t\to\infty$. The exponential helps only a bit: its argument vanishes e.g. for $d=1$, $k>0$, $l<0$ and $\tau=kt/(k+|l|)$. Thus we have the essentially optimal bounds

$$\int_0^t e^{-2\pi(\bar\lambda-\lambda)|k(t-\tau)+l\tau|}\,d\tau \leqslant \frac{1}{\pi(\bar\lambda-\lambda)|k-l|} \tag{6.9}$$

and

$$\int_0^t e^{-2\pi(\bar\lambda-\lambda)|k(t-\tau)+l\tau|}\,\tau\,d\tau \leqslant \frac{1}{2\pi^2(\bar\lambda-\lambda)^2|k-l|^2}+\frac{1}{\pi(\bar\lambda-\lambda)}\,\frac{|k|t}{|k-l|}. \tag{6.10}$$

From this computation we conclude that:

• The higher regularity of G has allowed us to reduce the time-integral due to a factor $e^{-\alpha|k(t-\tau)+l\tau|}$; but this factor is not small when τ/t is equal to $k/(k-l)$. As discussed in the next section, this reflects an important physical phenomenon called (plasma) echo, which can be assimilated to a resonance.

• If we had (in "gliding" norm) $\|G_\tau\|=O(1)$ this would ensure a uniform bound on the integral, as soon as $\gamma>0$, due to (6.9) and

$$\sum_{k,l\in\mathbb{Z}^d}\frac{e^{-\alpha|l|}}{(1+|k-l|)^{1+\gamma}}<\infty.$$

• But G_τ is a velocity gradient, so (unless of course G depends only on v) $\|G_\tau\|$ diverges like $O(\tau)$ as $\tau\to\infty$, which implies a divergence of our bounds in large time, as can be seen from (6.10). If $\gamma\leqslant 1$ this comes with a divergence in the k variable, since in this case

$$\sum_{k,l\in\mathbb{Z}^d}\frac{e^{-\alpha|l|}|k|}{(1+|k-l|)^{1+\gamma}}=\infty.$$

(The Coulomb case corresponds to $\gamma=1$, so in this respect it has a borderline divergence.)

The following estimate adapts this computation to the formalism of hybrid norms, and at the same time allows for a time cheating similar to the one in Proposition 6.2. Fortunately, we shall only need to treat the case when $R=R(\tau,x)$; the more general case with $R=R(\tau,x,v)$ would be much more tricky.

THEOREM 6.5. (Long-term regularity extortion) *Let* $G=G(\tau,x,v)$, $R=R(\tau,x)$ *and*

$$\sigma(t,x)=\int_0^t\int_{\mathbb{R}^d} G(\tau,x-v(t-\tau),v)\cdot R(\tau,x-v(t-\tau))\,dv\,d\tau.$$

Let λ, $\bar\lambda$, μ, $\bar\mu$, $\mu'=\mu'(t,\tau)$ *and* $M\geqslant 1$ *be such that* $(1+M)\lambda\geqslant\bar\lambda>\lambda>0$ *and* $\bar\mu\geqslant\mu'>\mu>0$, *and let* $\gamma\geqslant 0$ *and* $b=b(t,\tau)\geqslant 0$. *Then*

$$\|\sigma(t,\cdot)\|_{\dot{\mathcal{F}}^{\lambda t+\mu}}\leqslant\int_0^t K_0^G(t,\tau)\|R_\tau\|_{\mathcal{F}^\nu}\,d\tau+\int_0^t K_1^G(t,\tau)\|R_\tau\|_{\mathcal{F}^{\nu,\gamma}}\,d\tau, \tag{6.11}$$

where

$$\nu = \max\big\{\lambda\tau+\mu'-\tfrac{1}{2}\lambda b(t-\tau), 0\big\}, \tag{6.12}$$

$$K_0^G(t,\tau) = d e^{-\pi(\bar\lambda-\lambda)(t-\tau)} \left\| \int_{\mathbb{T}^d} G(\tau,x,\cdot)\,dx \right\|_{\mathcal{C}^{\bar\lambda(1+b);1}}, \tag{6.13}$$

$$K_1^G(t,\tau) = \left(\sup_{0\leqslant\tau\leqslant t} \frac{\|G_\tau\|_{\mathcal{Z}^{\bar\lambda(1+b),\bar\mu}_{\tau-bt/(1+b)}}}{1+\tau} \right) K_1(t,\tau), \tag{6.14}$$

$$\begin{aligned} &K_1(t,\tau) \\ &\quad = (1+\tau) d \sup_{k,l\in\mathbb{Z}^d_*} \frac{e^{-\pi(\bar\mu-\mu)|l|} e^{-\pi(\bar\lambda-\lambda)|k(t-\tau)+l\tau|/M} e^{-2\pi[\mu'-\mu+\lambda b(t-\tau)/2]|k-l|}}{1+|k-l|^\gamma}. \end{aligned} \tag{6.15}$$

Remark 6.6. It is essential in (6.11) to separate the contribution of $\widehat{G}(\tau,0,v)$ from the rest. Indeed, if we removed the restriction $l\neq 0$ in (6.15) the kernel K_1 would be too large to be correctly controlled in large time. What makes this separation reasonable is that, although in cases of application $G(\tau,x,v)$ is expected to grow like $O(\tau)$ in large time, the spatial average $\int_{\mathbb{T}^d} G(\tau,x,v)\,dx$ is expected to be bounded. Also, we will not need to take advantage of the parameter γ to handle this term.

Proof. Without loss of generality we may assume that G and R are scalar-valued. (This explains the constant d in front of the right-hand sides of (6.13) and (6.15).) First we assume that $\widehat{G}(\tau,0,v)=0$, and we write as before

$$\begin{aligned} \hat\sigma(t,k) &= \int_0^t \left(\sum_{l\in\mathbb{Z}^d_*} \int_{\mathbb{R}^d} \widehat{G}(\tau,l,v) \widehat{R}(\tau,k-l) e^{-2i\pi k\cdot v(t-\tau)}\,dv \right) d\tau, \\ |\hat\sigma(t,k)| &\leqslant \int_0^t \left(\sum_{l\in\mathbb{Z}^d_*} \left| \int_{\mathbb{R}^d} \widehat{G}(\tau,l,v) e^{-2i\pi k\cdot v(t-\tau)}\,dv \right| |\widehat{R}(\tau,k-l)| \right) d\tau. \end{aligned} \tag{6.16}$$

Next we let $\tau'=\tau-b(t-\tau)$ and write

$$\begin{aligned} e^{2\pi(\lambda t+\mu)|k|} \leqslant\, & e^{-2\pi(\bar\mu-\mu)|l|} e^{-2\pi\lambda(\tau-\tau')|k-l|} e^{-2\pi(\mu'-\mu)|k-l|} e^{-2\pi(\bar\lambda-\lambda)|k(t-\tau')+l\tau'|} \\ & \times e^{2\pi\bar\mu|l|} e^{2\pi(\lambda\tau+\mu')|k-l|} e^{2\pi\bar\lambda|k(t-\tau')+l\tau'|}. \end{aligned} \tag{6.17}$$

Since $0\leqslant\bar\lambda-\lambda\leqslant M\lambda$, we have

$$e^{-2\pi(\bar\lambda-\lambda)|k(t-\tau')+l\tau'|} \leqslant e^{-\pi(\bar\lambda-\lambda)|k(t-\tau)+l\tau|/M} e^{\pi\lambda(\tau-\tau')|k-l|};$$

so (6.17) implies that

$$\begin{aligned} e^{2\pi(\lambda t+\mu)|k|} \leqslant\, & e^{-2\pi(\bar\mu-\mu)|l|} e^{-\pi(\bar\lambda-\lambda)|k(t-\tau)+l\tau|/M} e^{-2\pi[\mu'-\mu+\lambda(\tau-\tau')/2]|k-l|} \\ & \times e^{2\pi\bar\mu|l|} e^{2\pi(\lambda\tau+\mu')|k-l|} \sum_{n\in\mathbb{N}^d_0} \frac{|2i\pi\bar\lambda(k(t-\tau')+l\tau')|^n}{n!}. \end{aligned} \tag{6.18}$$

For each $n\in\mathbb{N}_0^d$,

$$\begin{aligned}
&\frac{|2i\pi\bar\lambda(k(t-\tau')+l\tau')|^n}{n!}\biggl|\int_{\mathbb{R}^d}\widehat G(\tau,l,v)e^{-2i\pi k\cdot v(t-\tau)}\,dv\biggr|\\
&=\frac{\bar\lambda^n}{n!}\biggl|\int_{\mathbb{R}^d}\widehat G(\tau,l,v)[2i\pi(k(t-\tau')+l\tau')]^n e^{-2i\pi k\cdot v(t-\tau)}\,dv\biggr|\\
&=\frac{\bar\lambda^n}{n!}\biggl(\frac{t-\tau'}{t-\tau}\biggr)^n\biggl|\int_{\mathbb{R}^d}\widehat G(\tau,l,v)\biggl[2i\pi\biggl(k(t-\tau)+l\tau'\biggl(\frac{t-\tau}{t-\tau'}\biggr)\biggr)\biggr]^n e^{-2i\pi k\cdot v(t-\tau)}\,dv\biggr|\\
&=\frac{\bar\lambda^n}{n!}\biggl(\frac{t-\tau'}{t-\tau}\biggr)^n\biggl|\int_{\mathbb{R}^d}\widehat G(\tau,l,v)\biggl[-\nabla_v+2i\pi l\tau'\biggl(\frac{t-\tau}{t-\tau'}\biggr)\biggr]^n e^{-2i\pi k\cdot v(t-\tau)}\,dv\biggr|\\
&=\frac{\bar\lambda^n}{n!}\biggl(\frac{t-\tau'}{t-\tau}\biggr)^n\biggl|\int_{\mathbb{R}^d}\biggl[\nabla_v+2i\pi l\tau'\biggl(\frac{t-\tau}{t-\tau'}\biggr)\biggr]^n\widehat G(\tau,l,v)e^{-2i\pi k\cdot v(t-\tau)}\,dv\biggr|\\
&\leqslant\frac{\bar\lambda^n}{n!}\biggl(\frac{t-\tau'}{t-\tau}\biggr)^n\biggl\|\biggl(\nabla_v+2i\pi l\tau'\biggl(\frac{t-\tau}{t-\tau'}\biggr)\biggr)^n\widehat G(\tau,l,v)\biggr\|_{L^1(dv)}\\
&=\frac{\bar\lambda^n(1+b)^n}{n!}\biggl\|\biggl(\nabla_v+2i\pi l\biggl(\tau-\frac{bt}{1+b}\biggr)\biggr)^n\widehat G(\tau,l,v)\biggr\|_{L^1(dv)}.
\end{aligned}$$

Combining this with (6.16) and (6.18), and summing over k, we deduce that

$$\begin{aligned}
&\|\sigma(t,\cdot)\|_{\dot{\mathcal{F}}^{\lambda t+\mu}}\\
&\quad=\sum_{k\in\mathbb{Z}^d_*}e^{2\pi(\lambda t+\mu)|k|}|\hat\sigma(t,k)|\\
&\quad\leqslant\int_0^t\sum_{\substack{k,l\in\mathbb{Z}^d\\ n\in\mathbb{N}_0^d}}\frac{e^{-2\pi(\bar\mu-\mu)|l|}e^{-\pi(\bar\lambda-\lambda)|k(t-\tau)+l\tau|/M}e^{-2\pi[\mu'-\mu+\lambda(\tau-\tau')/2]|k-l|}}{1+|k-l|^\gamma}\\
&\qquad\times e^{2\pi\bar\mu|l|}e^{2\pi[\lambda\tau+\mu'-\lambda b(t-\tau)/2]|k-l|}\frac{\bar\lambda^n(1+b)^n}{n!}|\widehat R(\tau,k-l)|\\
&\qquad\times\biggl\|\biggl(\nabla_v+2i\pi l\biggl(\tau-\frac{bt}{1+b}\biggr)\biggr)^n\widehat G(\tau,l,v)\biggr\|_{L^1(dv)}\,d\tau,
\end{aligned}$$

and the desired estimate follows readily.

Finally we consider the contribution of

$$\widehat G(\tau,0,v)=\int_{\mathbb{T}^d}G(\tau,x,v)\,dx.$$

This is done in the same way, noting that

$$\sup_{k\in\mathbb{Z}^d_*}\frac{e^{-2\pi(\bar\lambda-\lambda)|k|(t-\tau)}}{1+|k|^\gamma}\leqslant e^{-2\pi(\bar\lambda-\lambda)(t-\tau)}.$$

□

To conclude this section we provide a "mode-by-mode" variant of Theorem 6.5; this will be useful for very singular interactions ($\gamma=1$ in Theorem 2.6).

THEOREM 6.7. *Under the same assumptions as in Theorem* 6.5, *for all* $k\in\mathbb{Z}^d$ *we have the estimate*

$$\begin{aligned} e^{2\pi(\lambda t+\mu)|k|}|\hat\sigma(t,k)| \leqslant \int_0^t K_0^G(t,\tau)(e^{2\pi\nu|k|}|\widehat R(\tau,k)|)\,d\tau \\ +\int_0^t \sum_{l\in\mathbb{Z}^d} K_{k,l}^G(t,\tau)e^{2\pi\nu|k-l|}(1+|k-l|^\gamma)|\widehat R(\tau,k-l)|\,d\tau, \end{aligned} \tag{6.19}$$

where K_0^G *is defined by* (6.13), ν *by* (6.12) *and*

$$\begin{aligned} K_{k,l}^G(t,\tau) &= \left(\sup_{0\leqslant\tau\leqslant t}\frac{\|G\|_{\mathcal{Z}^{\bar\lambda(1+b),\bar\mu;1}_{\tau-bt/(1+b)}}}{1+\tau}\right)K_{k,l}(t,\tau),\\ K_{k,l}(t,\tau) &= \frac{(1+\tau)d e^{-2\pi(\bar\mu-\mu)|l|}e^{-\pi(\bar\lambda-\lambda)|k(t-\tau)+l\tau|/M}e^{-2\pi[\mu'-\mu+\lambda b(t-\tau)/2]|k-l|}}{1+|k-l|^\gamma}. \end{aligned}$$

Proof. The proof is similar to that of Theorem 6.5, except that k is fixed and we use, for each l, the crude bound

$$\begin{aligned} e^{2\pi\bar\mu|l|}\left\|\left(\nabla_v+2i\pi l\left(\tau-\frac{bt}{1+b}\right)\right)^n\widehat G(\tau,l,v)\right\|_{L^1(dv)} \\ \leqslant \sum_{j\in\mathbb{Z}^d} e^{2\pi\bar\mu|j|}\left\|\left(\nabla_v+2i\pi j\left(\tau-\frac{bt}{1+b}\right)\right)^n\widehat G(\tau,j,v)\right\|_{L^1(dv)}. \end{aligned}$$

□

7. Control of the time-response

To motivate this section, let us start from the linearized equation (3.3), but now assume that f^0 depends on t, x and v, and that there is an extra source term S, decaying in time. Thus the equation is

$$\frac{\partial f}{\partial t}+v\cdot\nabla_x f-(\nabla W*\varrho)\cdot\nabla_v f^0=S,$$

and the equation for the density ϱ, as in the proof of Theorem 3.1, is

$$\begin{aligned} \varrho(t,x) = \int_{\mathbb{R}^d} f_i(x-vt,v)\,dv \\ +\int_0^t\int_{\mathbb{R}^d}\nabla_v f^0(\tau,x-v(t-\tau),v)\cdot(\nabla W*\varrho)(\tau,x-v(t-\tau))\,dv\,d\tau \\ +\int_0^t\int_{\mathbb{R}^d} S(\tau,x-v(t-\tau),v)\,dv\,d\tau. \end{aligned} \tag{7.1}$$

Hopefully we may apply Theorem 6.5 to deduce from (7.1) an integral inequality on $\varphi(t):=\|\varrho(t)\|_{\mathcal{F}^{\lambda t+\mu}}$, which will look like

$$\varphi(t)\leqslant A+c\int_0^t K(t,\tau)\varphi(\tau)\,d\tau, \tag{7.2}$$

where A is the contribution of the initial datum and the source term, and $K(t,\tau)$ is a kernel looking like, say, (6.15).

How do we proceed from (7.2)? Assume for a start that a smallness condition of the form (a) in Proposition 2.1 is satisfied. Then the simple and natural way, as in §3, would be to write

$$\varphi(t)\leqslant A+c\left(\int_0^t K(t,\tau)\,d\tau\right)\sup_{0\leqslant\tau\leqslant t}\varphi(\tau)$$

and deduce that

$$\varphi(t)\leqslant\frac{A}{1-c\int_0^t K(t,\tau)\,d\tau} \tag{7.3}$$

(assuming of course the denominator to be positive). However, if K is given by (6.15), it is easily seen that $\int_0^t K(t,\tau)\,d\tau\geqslant\varkappa t$ as $t\to\infty$, where $\varkappa>0$; then (7.3) is useless. In fact (7.2) *does not* prevent φ from tending to ∞ as $t\to\infty$. Nevertheless, its growth may be controlled under certain assumptions, as we shall see in this section. Before embarking on cumbersome calculations, we shall start with a qualitative discussion.

7.1. Qualitative discussion

The kernel K in (6.15) depends on the choice of $\mu'=\mu(t,\tau)$. How large $\mu'-\mu$ can be depends in turn on the amount of regularization offered by the convolution with the interaction ∇W. We shall distinguish several cases according to the regularity of the interaction.

7.1.1. Analytic interaction

If ∇W is analytic, there is $\sigma>0$ such that

$$\|\varrho*\nabla W\|_{\mathcal{F}^{\nu+\sigma}}\leqslant C\|\varrho\|_{\mathcal{F}^{\nu}}\quad\text{for all }\nu\geqslant 0;$$

then in (6.15) we can afford to choose, say, $\mu'-\mu=\sigma$ and $\gamma=0$. Thus, assuming that

$$b=\frac{B}{1+t}$$

with B so small that $(\mu'-\mu)-\lambda b(t-\tau)\geqslant\frac{1}{2}\sigma$, K is bounded by

$$\overline{K}^{(\alpha)}(t,\tau)=(1+\tau)\sup_{l,k\in\mathbb{Z}^d_*} e^{-\alpha|l|}e^{-\alpha|k-l|}e^{-\alpha|k(t-\tau)+l\tau|}, \tag{7.4}$$

where $\alpha=\frac{1}{2}\min\{\bar\lambda-\lambda,\bar\mu-\mu,\sigma\}$. To fix ideas, let us work in dimension $d=1$. The goal is to estimate solutions of

$$\varphi(t)\leqslant a+c\int_0^t \overline{K}^{(\alpha)}(t,\tau)\varphi(\tau)\,d\tau. \tag{7.5}$$

Whenever τ/t is a rational number distinct from 0 or 1, there are $k,l\in\mathbb{Z}$ such that $|k(t-\tau)+l\tau|=0$, and the size of $\overline{K}^{(\alpha)}(t,\tau)$ mainly depends on the minimum admissible values of k and $k-l$. Looking at values of τ/t of the form $1/(n+1)$ or $n/(n+1)$ suggests the approximation

$$\overline{K}^{(\alpha)}\lesssim(1+\tau)\min\{e^{-\alpha\tau/(t-\tau)}e^{-2\alpha},e^{-2\alpha(t-\tau)/\tau}e^{-\alpha}\}. \tag{7.6}$$

But this estimate is terrible: the time-integral of the right-hand side is much larger than the integral of $\overline{K}^{(\alpha)}$. In fact, the fast variation and "wiggling" behavior of $\overline{K}^{(\alpha)}$ are essential to get decent estimates.

To get a better feeling for $\overline{K}^{(\alpha)}$, let us only retain the term for $k=1$ and $l=-1$; this seems reasonable since we have an exponential decay as k or l tend to infinity (anyway, throwing away all other terms can only improve the estimates). So we look at

$$\widetilde{K}^{(\alpha)}(t,\tau)=(1+\tau)e^{-3\alpha}e^{-\alpha|t-2\tau|}.$$

Let us time-rescale by setting $k_t(\theta)=t\widetilde{K}^{(\alpha)}(t,t\theta)$ for $\theta\in[0,1]$ (the t factor appears because $d\tau=t\,d\theta$); then it is not hard to see that

$$\frac{k_t}{t}\to\frac{e^{-3\alpha}}{2\alpha}\delta_{1/2}.$$

This suggests the following baby model for (7.5):

$$\varphi(t)\leqslant a+ct\varphi\big(\tfrac{1}{2}t\big). \tag{7.7}$$

The important point in (7.7) is that, although the kernel has total mass $O(t)$, this mass is located far from the endpoint $\tau=t$; this is what prevents the fast growth of φ. Compare with the inequality $\varphi(t)\leqslant a+ct\varphi(t)$, which implies no restriction at all on φ.

To be slightly more quantitative, let us look for a power series

$$\Phi(t)=\sum_{k=0}^{\infty}a_kt^k$$

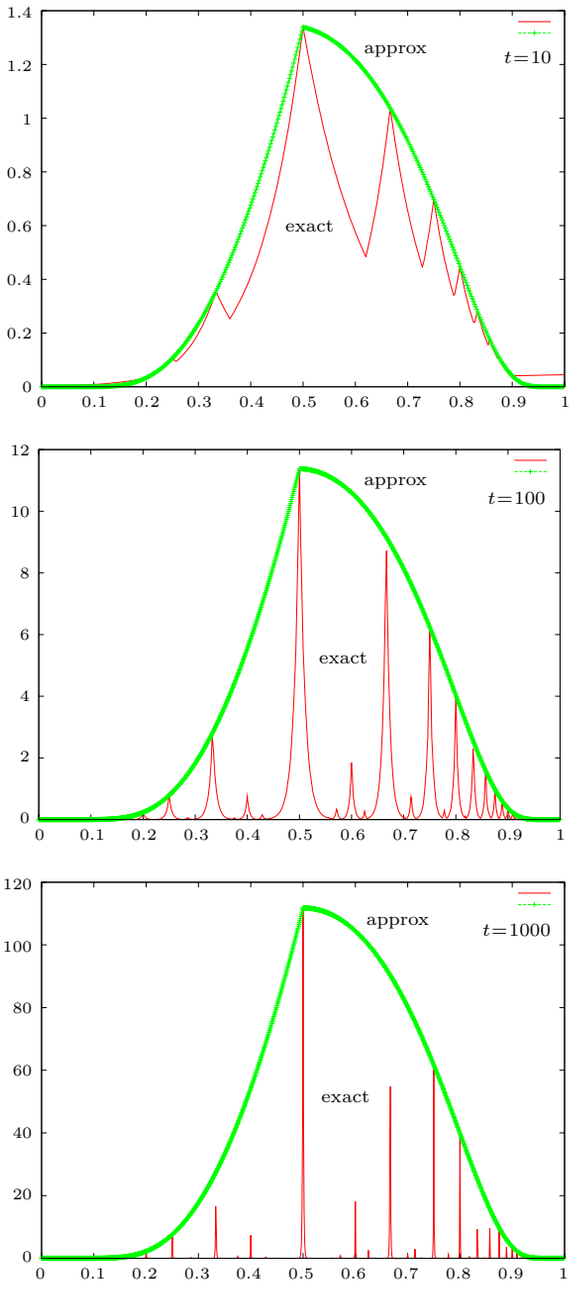

Figure 5. The kernel $\overline{K}^{(\alpha)}(t,\tau)$, together with the approximate upper bound in (7.6), for α=0.5 and t=10, t=100 and t=1000, respectively.

achieving equality in (7.7). This yields $a_0 = a$ and $a_{k+1} = c a_k 2^{-k}$, so

$$\Phi(t) = a \sum_{k=0}^{\infty} \frac{c^k t^k}{2^{k(k-1)/2}}. \tag{7.8}$$

The function Φ exhibits a truly remarkable behavior: it grows faster than any polynomial, but slower than any fractional exponential $\exp(ct^\nu)$, $\nu \in (0,1)$; essentially it behaves like $A^{(\log t)^2}$ (as can also be seen directly from (7.7)). One may conjecture that solutions of (7.5) exhibit a similar kind of growth.

Let us interpret these calculations. Typically, the kernel K controls the time variation of (say) the spatial density ϱ which is due to binary interaction of waves. When two waves of distinct frequencies interact, the effect over a long time-period is most of the time very small; this is a consequence of the oscillatory nature of the evolution, and the resulting time-averaging. But at certain particular times, the interaction becomes strong: this is known in plasma physics as the *plasma echo*, and can be thought of as a kind of resonance. Spectacular experiments by Malmberg and collaborators are based on this effect [32], [60]. Namely, if one starts a wave at frequency l at time 0, and forces it at time τ by a wave of frequency $k-l$, a strong response is obtained at time t and frequency k such that

$$k(t-\tau) + l\tau = 0 \tag{7.9}$$

(which of course is possible only if k and l are parallel to each other, with opposite directions).

In the present non-linear setting, whatever variation the density function is subject to, will result in echoes at later times. Even if each echo in itself will eventually decay, the problem is whether the accumulation of echoes will trigger an uncontrolled growth (unstability). As long as the expected growth is eaten by the time-decay coming from the linear theory, non-linear Landau damping is expected. In the present case, the growth of (7.8) is very slow in regard of the exponential time-decay due to the analytic regularity.

7.1.2. Sobolev interaction

If ∇W only has Sobolev regularity, we cannot afford in (6.15) to take $\mu'(t,\tau)$ larger than $\mu + \eta(t-\tau)/t$ (because the amount of regularity transferred in the bilinear estimates is only $O((t-\tau)/t)$, recall the discussion at the end of §6.2). On the other hand, we have $\gamma > 0$ such that

$$\|\nabla W * \varrho\|_{\mathcal{F}^{\nu,\gamma}} \leqslant C \|\varrho\|_{\mathcal{F}^\nu} \quad \text{for all } \nu \geqslant 0.$$

and then we can choose this γ in (6.15). So, assuming $b=B/(1+t)$ with B so small that $(\mu'-\mu)-\lambda b(t-\tau)\geqslant\eta(t-\tau)/2t$, K in (6.15) will be controlled by

$$K^{(\alpha),\gamma}(t,\tau)=(1+\tau)d\sup_{l,k\in\mathbb{Z}_*}\frac{e^{-\alpha|l|}e^{-\alpha(t-\tau)|k-l|/t}e^{-\alpha|k(t-\tau)+l\tau|}}{1+|k-l|^\gamma},\tag{7.10}$$

where $\alpha=\frac{1}{2}\min\{\bar\lambda-\lambda,\bar\mu-\mu,\eta\}$. The equation we are considering now is

$$\varphi(t)\leqslant a+\int_0^t K^{(\alpha),\gamma}(t,\tau)\varphi(\tau)\,d\tau.\tag{7.11}$$

For, say, $\tau\leqslant\frac{1}{2}t$, we have $K^{(\alpha)}\leqslant\overline{K}^{(\alpha/2)}$, and the discussion is similar to that in §7.1.1. But when τ aproaches t, the term $e^{-\alpha(t-\tau)|k-l|/t}$ hardly helps. Keeping only $k>0$ and $l=-1$ (because of the exponential decay in l) leads us to consider the kernel (for simplicity we set $d=1$ without loss of generality)

$$\check K^{(\alpha)}(t,\tau)=(1+\tau)\sup_{k\in\mathbb{Z}_*}\frac{e^{-\alpha|kt-(k+1)\tau|}}{1+(k+1)^\gamma}.$$

Once again we perform a time-rescaling, setting $\check k_t(\theta)=t\check K^{(\alpha)}(t,t\theta)$, and let $t\to\infty$. In this limit each exponential $e^{-\alpha|kt-(k+1)\tau|}$ becomes localized in a neighborhood of size $O(t/k)$ around $\theta=k/(k+1)$, and contributes a Dirac mass at $\theta=k/(k+1)$, with amplitude $2/\alpha(k+1)$. We get that

$$\frac{\check k_t}{t}\to\frac{2}{\alpha}\sum_{k\in\mathbb{Z}}\frac{1}{1+(k+1)^\gamma}\frac{k}{(k+1)^2}\delta_{1-1/(k+1)}\quad\text{as } t\to\infty.$$

This leads us to the following baby model for (7.11):

$$\varphi(t)\leqslant a+ct\sum_{k=1}^\infty\frac{1}{k^{1+\gamma}}\varphi\bigg(\Big(1-\frac{1}{k}\Big)t\bigg).\tag{7.12}$$

If we search for $\sum_{n=0}^\infty a_nt^n$ achieving equality, this yields

$$a_0=a,\quad a_{n+1}=c\bigg(\sum_{k=1}^\infty\frac{1}{k^{1+\gamma}}\Big(1-\frac{1}{k}\Big)^n\bigg)a_n.$$

To estimate the behavior of the $\sum_k$ above, we compare it with

$$\int_1^\infty\frac{1}{t^{1+\gamma}}\Big(1-\frac{1}{t}\Big)^n\,dt=\int_0^t u^{\gamma-1}(1-u)^n\,du=B(\gamma,n+1)=\frac{\Gamma(\gamma)\Gamma(n+1)}{\Gamma(n+\gamma+1)}=O\bigg(\frac{1}{n^\gamma}\bigg),$$

where $B(\gamma,n+1)$ is the beta function. All in all, we may expect φ in (7.11) to behave qualitatively like

$$\Phi(t)=a\sum_{n=0}^\infty\frac{c^nt^n}{(n!)^\gamma}.$$

Notice that Φ is subexponential for $\gamma>1$ (it grows essentially like the fractional exponential $\exp(t^{1/\gamma})$) and exponential for $\gamma=1$. In particular, as soon as $\gamma>1$, we expect non-linear Landau damping again.

7.1.3. Coulomb–Newton interaction ($\gamma=1$)

When $\gamma=1$, as is the case for Coulomb or Newton interaction, the previous analysis becomes borderline, since we expect (7.12) to be compatible with an exponential growth, and the linear decay is also exponential. To handle this more singular case, we shall work mode-by-mode, rather than on just one norm. Starting again from (7.1), we consider, for each $k\in\mathbb{Z}^d$,

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

and hope to get, via Theorem 6.7, an inequality which will roughly take the form

$$\varphi_k(t) \leqslant A_k + c\int_0^t \sum_{l\in\mathbb{Z}^d} K_{k,l}(t,\tau)\varphi_{k-l}(\tau)\,d\tau. \tag{7.13}$$

(Note that summing over k would yield an inequality worse than (7.11).) To fix the ideas, let us work in dimension $d=1$, and set $k\geqslant 1$ and $l=-1$. Reasoning as in §7.1.2, we obtain the baby model

$$\varphi_k(t) \leqslant A_k + \frac{ct}{(k+1)^{1+\gamma}}\varphi_{k+1}\left(\frac{kt}{k+1}\right). \tag{7.14}$$

The gain with respect to (7.12) is clear: for different values of k, the "dominant times" are distinct. From the physical point of view, we are discovering that, in some sense, *echoes occurring at distinct frequencies are asymptotically well separated*.

Let us search again for power series solutions: we set

$$\varphi_k(t) = \sum_{m=0}^{\infty} a_{k,m}t^m, \quad a_{k,0} = A_k.$$

By identification,

$$a_{k,m} = a_{k+1,m-1}c(k+1)^{-(1+\gamma)}\left(\frac{k}{k+1}\right)^{m-1},$$

and by induction

$$a_{k,m} = A_{k+m}c^m\left(\frac{k!}{(k+m)!}\right)^{1+\gamma}\frac{k^{m-1}c^m}{(k+1)(k+2)\dots(k+m)} \simeq A_{k+m}\left(\frac{k!}{(k+m)!}\right)^{\gamma+2}k^{m-1}c^m.$$

We may expect that $A_{k+m}\lesssim Ae^{-a(k+m)}$; then

$$a_{k,m} \lesssim A(ke^{-ak})k^m c^m\frac{e^{-am}}{(m!)^{\gamma+2}},$$

and in particular

$$\varphi_k(t) \lesssim Ae^{-ak/2}\sum_{m=0}^{\infty}\frac{(ckt)^m}{(m!)^{\gamma+2}} \lesssim Ae^{(1-\alpha)(ckt)^\alpha}, \quad \alpha = \frac{1}{\gamma+2}.$$

This behaves like a fractional exponential even for $\gamma=1$, and we can now believe in nonlinear Landau damping for such interactions! (The argument above works even for more singular interactions; but in the proof later the condition $\gamma\geqslant 1$ will be required for other reasons, see pp. 152 and 161.)

7.2. Exponential moments of the kernel

Now we start to estimate the kernel $K^{(\alpha),\gamma}$ from (7.10), without any approximation this time. Eventually, instead of proving that the growth is at most fractionally exponential, we shall compare it with a slow exponential $e^{\varepsilon t}$. For this, the first step consists in estimating exponential moments of the kernel $e^{-\varepsilon t}\int_0^t K(t,\tau)e^{\varepsilon\tau}\,d\tau$. (To get more precise estimates, one can study $e^{-\varepsilon t^\alpha}\int_0^t K(t,\tau)e^{\varepsilon\tau^\alpha}\,d\tau$, but such a refinement is not needed for the proof of Theorem 2.6.)

The first step consists of estimating exponential moments.

PROPOSITION 7.1. (Exponential moments of the kernel) *Let $\gamma\in[1,\infty)$ be given. For any $\alpha\in(0,1)$, let $K^{(\alpha),\gamma}$ be defined by* (7.10). *Then for any $\gamma<\infty$ there is $\bar\alpha=\bar\alpha(\gamma)>0$ such that if $\alpha\leqslant\bar\alpha$ and $\varepsilon\in(0,1)$, then for any $t>0$,*

$$e^{-\varepsilon t}\int_0^t K^{(\alpha),\gamma}(t,\tau)e^{\varepsilon\tau}\,d\tau \leqslant C\biggl(\frac{1}{\alpha\varepsilon^\gamma t^{\gamma-1}}+\frac{1}{\alpha\varepsilon^\gamma t^\gamma}\log\frac{1}{\alpha}+\frac{1}{\alpha^2\varepsilon^{1+\gamma}t^{1+\gamma}}+\biggl(\frac{1}{\alpha^3}+\frac{1}{\alpha^2\varepsilon}\log\frac{1}{\alpha}\biggr)e^{-\varepsilon t/4}+\frac{e^{-\alpha t/2}}{\alpha^3}\biggr),$$

where $C=C(\gamma)$. In particular,

- *if $\gamma>1$ and $\varepsilon\leqslant\alpha$, then*

$$e^{-\varepsilon t}\int_0^t K^{(\alpha),\gamma}(t,\tau)e^{\varepsilon\tau}\,d\tau\leqslant\frac{C(\gamma)}{\alpha^3\varepsilon^{1+\gamma}t^{\gamma-1}};$$

- *if $\gamma=1$ then*

$$e^{-\varepsilon t}\int_0^t K^{(\alpha),\gamma}(t,\tau)e^{\varepsilon\tau}\,d\tau\leqslant\frac{C}{\alpha^3}\biggl(\frac{1}{\varepsilon}+\frac{1}{\varepsilon^2 t}\biggr).$$

Remark 7.2. Much stronger estimates can be obtained if the interaction is analytic; that is, when $K^{(\alpha),\gamma}$ is replaced by $\overline{K}^{(\alpha)}$ defined in (7.4). A notable point about Proposition 7.1 is that for $\gamma=1$ we do not have any time-decay as $t\to\infty$.

Proof. To simplify notation we shall not recall the dependence of K on γ, and we shall set $d=1$ without loss of generality. We first assume that $\gamma<\infty$, and consider $\tau\leqslant\frac12 t$, which is the favorable case. We write

$$K^{(\alpha)}(t,\tau)\leqslant(1+\tau)\sup_{\substack{l\in\mathbb{Z}\\k\in\mathbb{Z}_*}}e^{-\alpha|l|}e^{-\alpha|k-l|/2}e^{-\alpha|k(t-\tau)+l\tau|}.$$

Since we got rid of the condition $l\neq0$, the right-hand side is now a non-increasing function of d. (To see this, pick up a non-zero component of k, and recall our norm conventions from Appendix A.1.) So we assume $d=1$. By symmetry we may also assume that $k>0$.

Explicit computations yield

$$\int_0^{t/2} e^{-\alpha|k(t-\tau)+l\tau|}(1+\tau)\,d\tau \leqslant \begin{cases} \dfrac{1}{\alpha(l-k)}+\dfrac{1}{\alpha^2(l-k)^2}, & \text{if } l>k, \\ e^{-\alpha kt}\left(\dfrac{t}{2}+\dfrac{t^2}{8}\right), & \text{if } l=k, \\ \dfrac{e^{-\alpha(k+l)t/2}}{\alpha|k-l|}\left(1+\dfrac{t}{2}\right), & \text{if } -k\leqslant l<k, \\ \dfrac{2}{\alpha|k-l|}+\dfrac{2kt}{\alpha|k-l|^2}+\dfrac{1}{\alpha^2|k-l|^2}, & \text{if } l<-k. \end{cases}$$

In all cases,

$$\int_0^{t/2} e^{-\alpha|k(t-\tau)+l\tau|}(1+\tau)\,d\tau \leqslant \left(\frac{3}{\alpha|k-l|}+\frac{1}{\alpha^2|k-l|^2}+\frac{2t}{\alpha|k-l|}\right)1_{k\neq l}+e^{-\alpha kt}\left(\frac{t}{2}+\frac{t^2}{8}\right)1_{l=k}.$$

So

$$\begin{aligned} e^{-\varepsilon t}\int_0^{t/2} e^{-\alpha|k(t-\tau)+l\tau|}(1+\tau)e^{\varepsilon\tau}\,d\tau &\leqslant e^{-\varepsilon t/2}\left(\frac{3}{\alpha|k-l|}+\frac{1}{\alpha^2|k-l|^2}+\frac{2t}{\alpha|k-l|}\right)1_{k\neq l}+e^{-\alpha kt}\left(\frac{t}{2}+\frac{t^2}{8}\right)1_{l=k} \\ &\leqslant e^{-\varepsilon t/4}\left(\frac{3}{\alpha|k-l|}+\frac{1}{\alpha^2|k-l|^2}+\frac{8z}{\alpha\varepsilon|k-l|}\right)1_{k\neq l}+e^{-t\alpha/2}\left(\frac{z}{\alpha}+\frac{8z^2}{\alpha^2}\right)1_{l=k}, \end{aligned}$$

where $z=\sup_x xe^{-x}=e^{-1}$. Then

$$\begin{aligned} e^{\varepsilon t}\int_0^{t/2} K^{(\alpha)}(t,\tau)e^{\varepsilon\tau}\,d\tau \leqslant e^{-\varepsilon t/4}\sum_{\substack{l,k\in\mathbb{Z}\\ 0\neq k\neq l}} e^{-\alpha|l|}e^{-\alpha|k-l|/2}\left(\frac{3}{\alpha|k-l|}+\frac{1}{\alpha^2|k-l|^2}+\frac{8z}{\alpha\varepsilon|k-l|}\right) \\ +e^{-t\alpha/2}\sum_{l\in\mathbb{Z}} e^{-\alpha|l|}\left(\frac{z}{\alpha}+\frac{8z^2}{\alpha^2}\right). \end{aligned}$$

Using the bounds (for $\alpha\sim 0^+$)

$$\sum_{l\in\mathbb{Z}} e^{-\alpha l}=O\left(\frac{1}{\alpha}\right),\quad \sum_{l\in\mathbb{Z}}\frac{e^{-\alpha l}}{l}=O\left(\log\frac{1}{\alpha}\right)\quad\text{and}\quad \sum_{l\in\mathbb{Z}}\frac{e^{-\alpha l}}{l^2}=O(1),$$

we end up, for $\alpha\leqslant\frac{1}{4}$, with a bound like

$$\begin{aligned} Ce^{-\varepsilon t/4}\left(\frac{1}{\alpha^2}\log\frac{1}{\alpha}+\frac{1}{\alpha^3}+\frac{1}{\alpha^2\varepsilon}\log\frac{1}{\alpha}\right)+Ce^{-\alpha t/2}\left(\frac{1}{\alpha^2}+\frac{1}{\alpha^3}\right) \\ \leqslant C\left[e^{-\varepsilon t/4}\left(\frac{1}{\alpha^3}+\frac{1}{\alpha^2\varepsilon}\log\frac{1}{\alpha}\right)+\frac{e^{-\alpha t/2}}{\alpha^3}\right]. \end{aligned}$$

(Note that the last term is $O(t^{-3})$, so it is anyway negligible in front of the other terms if $\gamma\leqslant 4$; in this case the restriction $\varepsilon\leqslant\alpha$ can be dispended with.)

- Next we turn to the more delicate contribution of $\tau\geqslant\frac{1}{2}t$. For this case we write

$$K^{(\alpha)}(t,\tau)\leqslant(1+\tau)\sup_{l\in\mathbb{Z}^d_*}e^{-\alpha|l|}\sup_{k\in\mathbb{Z}^d}\frac{e^{-\alpha|k(t-\tau)+l\tau|}}{1+|k-l|^\gamma},\tag{7.15}$$

and the upper bound is a non-increasing function of d, so we assume that $d=1$. Without loss of generality, we restrict the supremum to $l>0$.

The function $x\mapsto(1+|x-l|^\gamma)^{-1}e^{-\alpha|x(t-\tau)+l\tau|}$ is decreasing for $x\geqslant l$, increasing for $x\leqslant -l\tau/(t-\tau)$; and on the interval $[-l\tau/(t-\tau),l]$ its logarithmic derivative goes

$$\text{from }\left(-\alpha+\frac{\gamma/lt}{1+((t-\tau)/lt)^\gamma}\right)(t-\tau)\quad\text{to }-\alpha(t-\tau).$$

So, if $t\geqslant\gamma/\alpha$, there is a unique maximum at $x=-l\tau/(t-\tau)$, and the supremum in (7.15) is achieved for k equal to either the lower integer part, or the upper integer part of $-l\tau/(t-\tau)$. Thus a given integer k occurs in the supremum only for some times τ satisfying $k-1<-l\tau/(t-\tau)<k+1$. Since only negative values of k occur, let us change the sign so that k is non-negative. The equation

$$k-1<\frac{l\tau}{t-\tau}<k+1$$

is equivalent to

$$\frac{k-1}{k+l-1}t<\tau<\frac{k+1}{k+l+1}t.$$

Moreover, $\tau>\frac{1}{2}t$ implies that $k\geqslant l$. Thus, for $t\geqslant\gamma/\alpha$, we have

$$\begin{aligned}e^{-\varepsilon t}\int_{t/2}^t K^{(\alpha)}(t,\tau)e^{\varepsilon\tau}\,d\tau&\\ \leqslant e^{-\varepsilon t}\sum_{l=1}^\infty e^{-\alpha l}\sum_{k=l}^\infty\int_{(k-1)t/(k+l-1)}^{(k+1)t/(k+l+1)}(1+\tau)\frac{e^{-\alpha|k(t-\tau)-l\tau|}e^{\varepsilon\tau}}{1+(k+l)^\gamma}\,d\tau.&\end{aligned}\tag{7.16}$$

For $t\leqslant\gamma/\alpha$ we have the trivial bound

$$e^{-\varepsilon t}\int_{t/2}^t K^{(\alpha)}(t,\tau)e^{\varepsilon\tau}\,d\tau\leqslant\frac{\gamma}{2\alpha};$$

so in the sequel we shall just focus on the estimate of (7.16).

To evaluate the integral in the right-hand side of (7.16), we separate according to whether τ is smaller or larger than $kt/(k+l)$; we use trivial bounds for $e^{\varepsilon\tau}$ inside the integral, and in the end we get the explicit bounds

$$e^{-\varepsilon t}\int_{(k-1)t/(k+l-1)}^{kt/(k+l)}(1+\tau)e^{-\alpha|k(t-\tau)-l\tau|}e^{\varepsilon\tau}\,d\tau\leqslant e^{-\varepsilon lt/(k+l)}\left(\frac{1}{\alpha(k+l)}+\frac{kt}{\alpha(k+l)^2}\right),$$

$$e^{-\varepsilon t}\int_{kt/(k+l)}^{(k+1)t/(k+l+1)}(1+\tau)e^{-\alpha|k(t-\tau)-l\tau|}e^{\varepsilon\tau}\,d\tau\leqslant e^{-\varepsilon lt/(k+l+1)}\times\left(\frac{1}{\alpha(k+l)}+\frac{kt}{\alpha(k+l)^2}+\frac{1}{\alpha^2(k+l)^2}\right).$$

All in all, there is a numeric constant C such that (7.16) is bounded above by

$$C\sum_{l=1}^{\infty}e^{-\alpha l}\sum_{k=l}^{\infty}\left(\frac{1}{\alpha^2(k+l)^{2+\gamma}}+\frac{1}{\alpha(k+l)^{1+\gamma}}+\frac{kt}{\alpha(k+l)^{2+\gamma}}\right)e^{-\varepsilon lt/(k+l)},\tag{7.17}$$

together with an additional similar term where $e^{-\varepsilon lt/(k+l)}$ is replaced by $e^{-\varepsilon lt/(k+l+1)}$, and which will satisfy similar estimates.

We consider separately the three contributions in the right-hand side of (7.17). The first one is

$$\frac{1}{\alpha^2}\sum_{l=1}^{\infty}e^{-\alpha l}\sum_{k=l}^{\infty}\frac{e^{-\varepsilon lt/(k+l)}}{(k+l)^{2+\gamma}}.$$

To evaluate the behavior of this sum, we compare it to the 2-dimensional integral

$$I(t)=\frac{1}{\alpha^2}\int_1^{\infty}e^{-\alpha x}\int_x^{\infty}\frac{e^{-\varepsilon xt/(x+y)}}{(x+y)^{2+\gamma}}\,dy\,dx.$$

We change variables $(x,y)\mapsto(x,u)$, where $u(x,y)=\varepsilon xt/(x+y)$. This has Jacobian determinant $dx\,dy/dx\,du=\varepsilon xt/u^2$, and we find that

$$I(t)=\frac{1}{\alpha^2\varepsilon^{1+\gamma}t^{1+\gamma}}\int_1^{\infty}\frac{e^{-\alpha x}}{x^{1+\gamma}}\,dx\int_0^{\varepsilon t/2}e^{-u}u^{\gamma}\,du=O\left(\frac{1}{\alpha^2\varepsilon^{1+\gamma}t^{1+\gamma}}\right).$$

The same computation for the second integral in the right-hand side of (7.17) yields

$$\frac{1}{\alpha\varepsilon^{\gamma}t^{\gamma}}\int_1^{\infty}\frac{e^{-\alpha x}}{x^{\gamma}}\,dx\int_0^{\varepsilon t/2}e^{-u}u^{\gamma-1}\,du=O\left(\frac{1}{\alpha\varepsilon^{\gamma}\,t^{\gamma}}\log\frac{1}{\alpha}\right).$$

(The logarithmic factor arises only for $\gamma=1$.)

The third exponential in the right-hand side of (7.17) is the worst. It yields a contribution

$$\frac{t}{\alpha}\sum_{l=1}^{\infty}e^{-\alpha l}\sum_{k=l}^{\infty}\frac{e^{-\varepsilon ltk/(k+l)}}{(k+l)^{2+\gamma}}.\tag{7.18}$$

We compare this with the integral

$$\frac{t}{\alpha}\int_1^\infty e^{-\alpha x}\int_x^\infty \frac{e^{-\varepsilon x t/(x+y)}y}{(x+y)^{2+\gamma}}\,dy\,dx,$$

and the same change of variables as before equates this with

$$\frac{1}{\alpha\varepsilon^\gamma t^{\gamma-1}}\int_1^\infty \frac{e^{-\alpha x}}{x^\gamma}\,dx\int_0^{\varepsilon t/2} e^{-u}u^{\gamma-1}\,du - \frac{1}{\alpha\varepsilon^{1+\gamma}t^\gamma}\int_1^\infty \frac{e^{-\alpha x}}{x^\gamma}\,dx\int_0^{\varepsilon t/2} e^{-u}u^\gamma\,du$$
$$= O\biggl(\frac{1}{\alpha\varepsilon^\gamma t^{\gamma-1}}\log\frac{1}{\alpha}\biggr).$$

(Again the logarithmic factor arises only for $\gamma=1$.)

The proof of Proposition 7.1 follows by collecting all these bounds and keeping only the worst one. □

Remark 7.3. It is not easy to catch (say numerically) the behavior of (7.18), because it comes as a superposition of exponentially decaying modes; any truncation in k would lead to a radically different time-asymptotics.

From Proposition 7.1 we deduce L^2 exponential bounds.

COROLLARY 7.4. (L^2 exponential moments of the kernel) *With the same notation as in Proposition* 7.1,

$$e^{-2\varepsilon t}\int_0^t K^{(\alpha),\gamma}(t,\tau)^2 e^{2\varepsilon\tau}\,d\tau \leqslant \begin{cases} \dfrac{C(\gamma)}{\alpha^4\varepsilon^{1+2\gamma}t^{2(\gamma-1)}}, & \text{if } \gamma>1,\\[2ex] C\biggl(\dfrac{1}{\alpha^3\varepsilon^2}+\dfrac{1}{\alpha^2\varepsilon^3 t}\biggr), & \text{if } \gamma=1. \end{cases} \tag{7.19}$$

Proof. This follows easily from Proposition 7.1 and the obvious bound

$$K^{(\alpha),\gamma}(t,\tau)^2 \leqslant C(1+t)K^{(2\alpha),2\gamma}(t,\tau).$$ □

7.3. Dual exponential moments

PROPOSITION 7.5. *With the same notation as in Proposition* 7.1, *for any* $\gamma\geqslant 1$ *we have*

$$\sup_{\tau\geqslant 0} e^{\varepsilon\tau}\int_\tau^\infty e^{-\varepsilon t}K^{(\alpha),\gamma}(t,\tau)\,dt \leqslant C(\gamma)\biggl(\frac{1}{\alpha^2\varepsilon}+\frac{1}{\alpha\varepsilon^\gamma}\log\frac{1}{\alpha}\biggr). \tag{7.20}$$

Remark 7.6. The corresponding computation for the baby model considered in §7.1.2 is

$$e^{\varepsilon\tau}\frac{1+\tau}{\alpha}\sum_{k=1}^\infty \frac{e^{-\varepsilon(k+1)\tau/k}}{k^{1+\gamma}} \simeq \frac{1+\tau}{\alpha}\int_1^\infty \frac{e^{-\varepsilon\tau/x}}{x^{1+\gamma}}\,dx = \frac{1+\tau}{\tau^\gamma}\frac{1}{\alpha\varepsilon^\gamma}\int_0^{\varepsilon\tau} e^{-u}u^{\gamma-1}\,du.$$

So we expect the dependence upon ε in (7.20) to be sharp as $\gamma\to 1$.

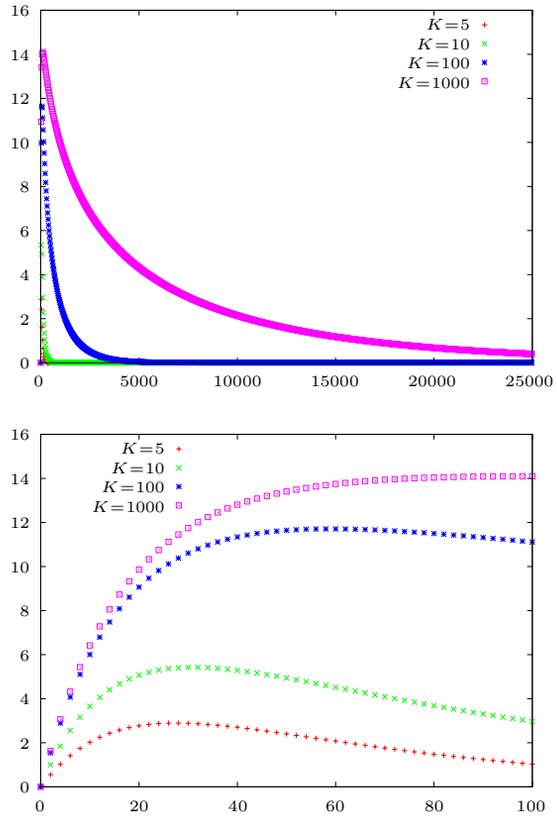

Figure 6. The function (7.18) for $\gamma=1$, truncated at $l=1$ and $k\leqslant K$, for $K=5, 10, 100, 1000$. The decay is slower and slower, but still exponential (picture above); however, the maximum value occurs on a much slower time scale and slowly increases with the truncation parameter (picture below, which is a zoom on shorter times).

Proof. We first reduce to $d=1$, and split the integral as

$$e^{\varepsilon\tau}\int_\tau^\infty e^{-\varepsilon t}K^{(\alpha),\gamma}(t,\tau)\,dt = \underbrace{e^{\varepsilon\tau}\int_{2\tau}^\infty e^{-\varepsilon t}K^{(\alpha),\gamma}(t,\tau)\,dt}_{:=I_1} + \underbrace{e^{\varepsilon\tau}\int_\tau^{2\tau} e^{-\varepsilon t}K^{(\alpha),\gamma}(t,\tau)\,dt}_{:=I_2}.$$

The first term I_1 is easy: for $2\tau\leqslant t\leqslant\infty$ we have

$$K^{(\alpha),\gamma}(t,\tau)\leqslant(1+\tau)\sum_{k=2}^\infty\sum_{l\in\mathbb{Z}_*} e^{-\alpha|l|-\alpha|k-l|/2}\leqslant\frac{C(1+\tau)}{\alpha^2},$$

and thus

$$e^{\varepsilon\tau}\int_{2\tau}^\infty e^{-\varepsilon t}K^{(\alpha),\gamma}(t,\tau)\,dt\leqslant\frac{C(1+\tau)}{\alpha^2}e^{-\varepsilon\tau}\leqslant\frac{C}{\varepsilon\alpha^2}.$$

We treat the second term I_2 as in the proof of Proposition 7.1:

$$e^{\varepsilon\tau}\int_\tau^{2\tau} K^{(\alpha),\gamma}(t,\tau)e^{-\varepsilon t}\,dt \leqslant e^{\varepsilon\tau}(1+\tau)\sum_{l=1}^{\infty} e^{-\alpha l}\sum_{k=l}^{\infty}\int_{(k+l+1)\tau/(k+1)}^{(k+l-1)\tau/(k-1)} \frac{e^{-\alpha|k(t-\tau)-l\tau|}}{1+(k+l)^\gamma} e^{-\varepsilon t}\,dt$$
$$\leqslant (1+\tau)\sum_{l=1}^{\infty} e^{-\alpha l}\sum_{k=l}^{\infty}\frac{e^{-\varepsilon l\tau/k}}{k^\gamma}\frac{2}{k\alpha}.$$

We compare this with

$$\frac{2(1+\tau)}{\alpha}\int_1^\infty e^{-\alpha x}\int_x^\infty \frac{e^{-\varepsilon x\tau/y}}{y^{1+\gamma}}\,dy\,dx = \frac{2}{\alpha\varepsilon^\gamma}\frac{1+\tau}{\tau^\gamma}\int_1^\infty\frac{e^{-\alpha x}}{x^\gamma}\int_0^{\varepsilon\tau} e^{-u}u^{\gamma-1}\,du\,dx$$
$$\leqslant \frac{C}{\alpha\varepsilon^\gamma}\log\frac{1}{\alpha},$$

where we used the change of variable $u=\varepsilon x\tau/y$. The desired conclusion follows. Note that as before the term $\log(1/\alpha)$ only occurs when $\gamma=1$, and that, for $\gamma>1$, one could improve the estimate above into a time-decay of the form $O(\tau^{-(\gamma-1)})$. □

7.4. Growth control

We will now state the main result of this section. For a sequence of functions $\Phi(k,t)$, $k\in\mathbb{Z}^d_*=\mathbb{Z}^d\setminus\{0\}$ and $t\in\mathbb{R}$, we set

$$\|\Phi(t)\|_\lambda = \sum_{k\in\mathbb{Z}^d_*}|\Phi(k,t)|e^{2\pi\lambda|k|}.$$

We shall use $K(s)\Phi(t)$ as a shorthand for $(K(k,s)\Phi(k,t))_{k\in\mathbb{Z}^d_*}$, etc.

THEOREM 7.7. (Growth control via integral inequalities) *Assume that $f^0=f^0(v)$ and $W=W(x)$ satisfy condition* **(L)** *from §2.2 with constants C_0, λ_0 and $\varkappa$; in particular $|\tilde f^0(\eta)|\leqslant C_0\,e^{-2\pi\lambda_0|\eta|}$. Let further*

$$C_W = \max\Big\{\sum_{k\in\mathbb{Z}^d_*}|\widehat W(k)|,\ \sup_{k\in\mathbb{Z}^d_*}|k|\,|\widehat W(k)|\Big\}.$$

Let $A\geqslant 0$, $\mu\geqslant 0$ and $\lambda\in(0,\lambda^*]$ with $0<\lambda^*<\lambda_0$. Let $(\Phi(k,t))_{k\in\mathbb{Z}^d_*,t\geqslant 0}$ be a continuous function of $t\geqslant 0$, valued in $\mathbb{C}^{\mathbb{Z}^d_*}$, such that

$$\Big\|\Phi(t)-\int_0^t K^0(t-\tau)\Phi(\tau)\,d\tau\Big\|_{\lambda t+\mu} \leqslant A+\int_0^t\Big(K_0(t,\tau)+K_1(t,\tau)+\frac{c_0}{(1+\tau)^m}\Big)\|\Phi(\tau)\|_{\lambda\tau+\mu}\,d\tau \quad \text{for all } t\geqslant 0, \tag{7.21}$$

where $c_0 \geqslant 0$, $m > 1$, *and* $K_0(t,\tau)$ *and* $K_1(t,\tau)$ *are non-negative kernels. Let*

$$\varphi(t) = \|\Phi(t)\|_{\lambda t+\mu}.$$

Then, we have the following:

(i) *Assume that* $\gamma > 1$ *and* $K_1 = cK^{(\alpha),\gamma}$ *for some* $c > 0$, $\alpha \in (0, \bar{\alpha}(\gamma))$, *where* $K^{(\alpha),\gamma}$ *is defined by* (7.10), *and* $\bar{\alpha}(\gamma)$ *appears in Proposition* 7.1. *Then there are positive constants* C *and* χ, *depending only on* γ, λ^*, λ_0, $\varkappa$, c_0, C_W *and* m, *uniform as* $\gamma \to 1$, *such that if*

$$\sup_{t \geqslant 0} \int_0^t K_0(t,\tau)\, d\tau \leqslant \chi \tag{7.22}$$

and

$$\sup_{t \geqslant 0} \left(\int_0^t K_0(t,\tau)^2\, d\tau \right)^{1/2} + \sup_{\tau \geqslant 0} \int_\tau^\infty K_0(t,\tau)\, dt \leqslant 1, \tag{7.23}$$

then for any $\varepsilon \in (0, \alpha)$,

$$\varphi(t) \leqslant CA \frac{1+c_0^2}{\sqrt{\varepsilon}} e^{Cc_0} \Big(1 + \frac{c}{\alpha\varepsilon}\Big) e^{CT} e^{Cc(1+T^2)} e^{\varepsilon t} \quad \textit{for all } t \geqslant 0, \tag{7.24}$$

where

$$T = C \max\left\{ \left(\frac{c^2}{\alpha^5 \varepsilon^{2+\gamma}} \right)^{1/(\gamma-1)}, \Big(\frac{c}{\alpha^2 \varepsilon^{\gamma+1/2}} \Big)^{1/(\gamma-1)}, \left(\frac{c_0^2}{\varepsilon} \right)^{1/(2m-1)} \right\}. \tag{7.25}$$

(ii) *Assume that* $K_1 = \sum_{j=1}^N c_j K^{(\alpha_j),1}$ *for some* $\alpha_j \in (0, \bar{\alpha}(1))$, *where* $\bar{\alpha}(1)$ *appears in Proposition* 7.1; *then there is a numeric constant* $\Gamma > 0$ *such that whenever*

$$1 \geqslant \varepsilon \geqslant \Gamma \sum_{j=1}^N \frac{c_j}{\alpha_j^3},$$

one has, with the same notation as in (i),

$$\varphi(t) \leqslant CA \frac{1+c_0^2}{\sqrt{\varepsilon}} e^{Cc_0} e^{CT} e^{Cc(1+T^2)} e^{\varepsilon t} \quad \textit{for all } t \geqslant 0, \tag{7.26}$$

where

$$c = \sum_{j=1}^N c_j \quad \textit{and} \quad T = C \max\left\{ \frac{1}{\varepsilon^2} \sum_{j=1}^N \frac{c_j}{\alpha_j^3}, \left(\frac{c_0^2}{\varepsilon} \right)^{1/(2m-1)} \right\}.$$

Remark 7.8. Apart from the term $c_0/(1+\tau)^m$ which will appear as a technical correction, there are three different kernels appearing in Theorem 7.7: the kernel K^0, which is associated with the linearized Landau damping; the kernel K_1, describing non-linear echoes (due to interaction between different Fourier modes); and the kernel K_0, describing the instantaneous response (due to interaction between identical Fourier modes).

We shall first prove Theorem 7.7 assuming

$$c_0 = 0 \tag{7.27}$$

and

$$\int_0^\infty \sup_{k\in\mathbb{Z}^d} |K^0(k,t)| e^{2\pi\lambda_0|k|t}\, dt \leqslant 1-\varkappa, \quad \varkappa\in(0,1), \tag{7.28}$$

which is a reinforcement of condition **(L)**. Under these assumptions the proof of Theorem 7.7 is much simpler, and its conclusion can be substantially simplified too: χ depends only on $\varkappa$; condition (7.23) on K_0 can be dropped; and the factor $e^{CT}(1+c/\alpha\varepsilon^{3/2})$ in (7.24) can be omitted. If $\widehat{W}\leqslant 0$ (as for gravitational interaction) and $\tilde{f}^0\geqslant 0$ (as for Maxwellian background), these additional assumptions do not constitute a loss of generality, since (7.28) becomes essentially equivalent to **(L)**, and for c_0 small enough the term $c_0(1+\tau)^{-m}$ can be incorporated inside K_0.

Proof of Theorem 7.7 *under* (7.27) *and* (7.28). We have

$$\varphi(t) \leqslant A + \int_0^t (|K^0|(t-\tau)+K_0(t,\tau)+K_1(t,\tau))\varphi(\tau)\, d\tau, \tag{7.29}$$

where $|K^0(t)|=\sup_{k\in\mathbb{Z}^d}|K^0(k,t)|$. We shall estimate φ by a maximum principle argument. Let $\psi(t)=Be^{\varepsilon t}$, where B will be chosen later. If ψ satisfies, for some $T\geqslant 0$,

$$\begin{cases} \varphi(t) < \psi(t), & \text{for } 0\leqslant t\leqslant T, \\ \psi(t) \geqslant A + \displaystyle\int_0^t (|K^0|(t,\tau)+K_0(t,\tau)+K_1(t,\tau))\psi(\tau)\, d\tau, & \text{for } t\geqslant T, \end{cases} \tag{7.30}$$

then $u(t):=\psi(t)-\varphi(t)$ is positive for $t\leqslant T$, and satisfies $u(t)\geqslant\int_0^t K(t,\tau)u(\tau)\,d\tau$ for $t\geqslant T$, with $K=|K^0|+K_0+K_1>0$; this prevents u from vanishing at later times, so $u\geqslant 0$ and $\varphi\leqslant\psi$. Thus it is sufficient to establish (7.30).

(i) By Proposition 7.1, and since $\int_0^t(|K^0|+K_0)\,d\tau\leqslant 1-\frac{1}{2}\varkappa$ (for $\chi\leqslant\frac{1}{2}\varkappa$),

$$\begin{aligned} A + \int_0^t (|K^0|(t,\tau)+K_0(t,\tau))\psi(\tau)\, d\tau + c\int_0^t K^{(\alpha),\gamma}(t,\tau)\psi(\tau)\, d\tau \\ \leqslant A + \left(1-\frac{\varkappa}{2}+\frac{cC(\gamma)}{\alpha^3\varepsilon^{1+\gamma}t^{\gamma-1}}\right)Be^{\varepsilon t}. \end{aligned} \tag{7.31}$$

For $t\geqslant T:=(4cC(\alpha^3\varepsilon^{1+\gamma}\varkappa))^{1/(\gamma-1)}$, this is bounded above by $A+\left(1-\frac{1}{4}\varkappa\right)Be^{\varepsilon t}$, which in turn is bounded by $Be^{\varepsilon t}$ as soon as $B\geqslant 4A/\varkappa$.

On the other hand, from the inequality

$$\varphi(t) \leqslant A + \left(1-\frac{\varkappa}{2}\right)\sup_{0\leqslant\tau\leqslant t}\varphi(\tau) + c(1+t)\int_0^t \varphi(\tau)\, d\tau$$

we deduce that

$$\varphi(t) \leqslant \frac{2A}{\varkappa}(1+t)e^{2c(t+t^2/2)/\varkappa}.$$

In particular, if

$$\frac{4A}{\varkappa}e^{c'(T+T^2)} \leqslant B$$

with $c'=c'(c,\varkappa)$ large enough, then for $0\leqslant t\leqslant T$ we have $\varphi(t)\leqslant \frac{1}{2}\psi(t)$, and (7.30) holds.

(ii) $K_1=\sum_{j=1}^N c_j K^{(\alpha_j),1}$. We use the same reasoning, replacing the right-hand side in (7.31) by

$$A+\left(1-\frac{\varkappa}{2}+C\left(\sum_{j=1}^N \frac{c_j}{\alpha_j^3\varepsilon}+\sum_{j=1}^N \frac{c_j}{\alpha_j^3\varepsilon^2 t}\right)\right)Be^{\varepsilon t}.$$

To conclude the proof, we may first impose a lower bound on ε to ensure that

$$C\sum_{j=1}^N \frac{c_j}{\alpha_j^3\varepsilon} \leqslant \frac{\varkappa}{8}, \tag{7.32}$$

and then choose t large enough to guarantee that

$$C\sum_{j=1}^N \frac{c_j}{\alpha_j^3\varepsilon^2 t} \leqslant \frac{\varkappa}{8}; \tag{7.33}$$

this yields (ii). □

Proof of Theorem 7.7 *in the general case.* We only treat (i), since the reasoning for (ii) is rather similar; and we only establish the conclusion as an a-priori estimate, skipping the continuity/approximation argument needed to turn it into a rigorous estimate. Then the proof is done in three steps.

Step 1. *Crude pointwise bounds.* From (7.21) we have

$$\begin{aligned}
\varphi(t) &= \sum_{k\in\mathbb{Z}^d_*} |\Phi(k,t)|e^{2\pi(\lambda t+\mu)|k|} \\
&\leqslant A+\sum_{k\in\mathbb{Z}^d_*}\int_0^t |K^0(k,t-\tau)|e^{2\pi(\lambda t+\mu)|k|}|\Phi(t,\tau)|\,d\tau \\
&\qquad +\int_0^t \left(K_0(t,\tau)+K_1(t,\tau)+\frac{c_0}{(1+\tau)^m}\right)\varphi(\tau)\,d\tau \\
&\leqslant A+\int_0^t \left(K_0(t,\tau)+K_1(t,\tau)+\frac{c_0}{(1+\tau)^m}+\sup_{k\in\mathbb{Z}^d_*}|K^0(k,t-\tau)|e^{2\pi\lambda(t-\tau)|k|}\right)\varphi(\tau)\,d\tau.
\end{aligned} \tag{7.34}$$

We note that for any $k\in\mathbb{Z}^d_*$ and $t\geqslant 0$,

$$\begin{aligned}|K^0(k,t-\tau)|e^{2\pi\lambda|k|(t-\tau)} &\leqslant 4\pi^2|\widehat{W}(k)|C_0e^{-2\pi(\lambda_0-\lambda)|k|t}|k|^2t\\ &\leqslant \frac{CC_0}{\lambda_0-\lambda}\sup_{k\in\mathbb{Z}^d_*}|k|\,|\widehat{W}(k)|\leqslant\frac{CC_0C_W}{\lambda_0-\lambda},\end{aligned}$$

where (here and below) C stands for a numeric constant which may change from line to line. Assuming that $\int_0^t K_0(t,\tau)\,d\tau\leqslant\frac{1}{2}$, we deduce from (7.34) that

$$\varphi(t)\leqslant A+\frac{1}{2}\sup_{0\leqslant\tau\leqslant t}\varphi(\tau)+C\int_0^t\left(\frac{C_0C_W}{\lambda_0-\lambda}+c(1+t)+\frac{c_0}{(1+\tau)^m}\right)\varphi(\tau)\,d\tau,$$

and, by Grönwall's lemma,

$$\varphi(t)\leqslant 2Ae^{C(C_0C_Wt/(\lambda_0-\lambda)+c(t+t^2)+c_0C_m)},\tag{7.35}$$

where $C_m=\int_0^\infty(1+\tau)^{-m}\,d\tau$.

Step 2. L^2 *bound.* This is the step where the smallness assumption (7.22) will be most important. For all $k\in\mathbb{Z}^d_*$ and $t\geqslant 0$ we define

$$\Psi_k(t)=e^{-\varepsilon t}\Phi(k,t)e^{2\pi(\lambda t+\mu)|k|},\tag{7.36}$$

$$\mathcal{K}^0_k(t)=e^{-\varepsilon t}K^0(k,t)e^{2\pi(\lambda t+\mu)|k|},\tag{7.37}$$

$$R_k(t)=e^{-\varepsilon t}\left(\Phi(k,t)-\int_0^t K^0(k,t-\tau)\Phi(k,\tau)\,d\tau\right)e^{2\pi(\lambda t+\mu)|k|}=(\Psi_k-\Psi_k*\mathcal{K}^0_k)(t),\tag{7.38}$$

and we extend all these functions by 0 for negative values of t. Taking Fourier transform in the time-variable yields $\widehat{R}_k=(1-\widehat{\mathcal{K}}^0_k)\widehat{\Psi}_k$; since condition **(L)** implies that $|1-\widehat{\mathcal{K}}^0_k|\geqslant\varkappa$, we deduce that $\|\widehat{\Psi}_k\|_{L^2}\leqslant\varkappa^{-1}\|\widehat{R}_k\|_{L^2}$, i.e.,

$$\|\Psi_k\|_{L^2(dt)}\leqslant\frac{\|R_k\|_{L^2(dt)}}{\varkappa}.\tag{7.39}$$

Plugging (7.39) into (7.38), we deduce that

$$\|\Psi_k-R_k\|_{L^2(dt)}\leqslant\frac{\|\mathcal{K}^0_k\|_{L^1(dt)}}{\varkappa}\|R_k\|_{L^2(dt)}\quad\text{for all } k\in\mathbb{Z}^d_*.\tag{7.40}$$

Then

$$\begin{aligned}\|\varphi(t)e^{-\varepsilon t}\|_{L^2(dt)}&=\left\|\sum_{k\in\mathbb{Z}^d_*}|\Psi_k|\right\|_{L^2(dt)}\\ &\leqslant\left\|\sum_{k\in\mathbb{Z}^d_*}|R_k|\right\|_{L^2(dt)}+\sum_{k\in\mathbb{Z}^d_*}\|R_k-\Psi_k\|_{L^2(dt)}\\ &\leqslant\left\|\sum_{k\in\mathbb{Z}^d_*}|R_k|\right\|_{L^2(dt)}\left(1+\frac{1}{\varkappa}\sum_{l\in\mathbb{Z}^d_*}\|\mathcal{K}^0_l\|_{L^1(dt)}\right).\end{aligned}\tag{7.41}$$

(Note that we bounded $\|R_l\|$ by $\big\|\sum_{k\in\mathbb{Z}^d_*}|R_k|\big\|$, which seems very crude; but the decay of $\mathcal{K}^0_k$ as a function of k will save us.) Next, we note that

$$\|\mathcal{K}^0_k\|_{L^1(dt)} \leqslant 4\pi^2|\widehat{W}(k)| \int_0^\infty C_0 e^{-2\pi(\lambda_0-\lambda)|k|t}|k|^2 t\,dt \leqslant 4\pi^2\ |\widehat{W}(k)|\frac{C_0}{(\lambda_0-\lambda)^2},$$

so

$$\sum_{k\in\mathbb{Z}^d_*}\|\mathcal{K}^0_k\|_{L^1(dt)} \leqslant 4\pi^2\bigg(\sum_{k\in\mathbb{Z}^d_*}|\widehat{W}(k)|\bigg)\frac{C_0}{(\lambda_0-\lambda)^2}.$$

Plugging this into (7.41) and using (7.21) again, we obtain

$$\begin{aligned}
&\|\varphi(t)e^{-\varepsilon t}\|_{L^2(dt)}\\
&\leqslant \bigg(1+\frac{CC_0C_W}{\varkappa(\lambda_0-\lambda)^2}\bigg)\bigg\|\sum_{k\in\mathbb{Z}^d_*}|R_k|\bigg\|_{L^2(dt)}\\
&\leqslant \bigg(1+\frac{CC_0C_W}{\varkappa(\lambda_0-\lambda)^2}\bigg)\bigg(\int_0^\infty e^{-2\varepsilon t}\bigg(A+\int_0^t\bigg(K_1+K_0+\frac{c_0}{(1+\tau)^m}\bigg)\varphi(\tau)\,d\tau\bigg)^2 dt\bigg)^{1/2}.
\end{aligned} \tag{7.42}$$

We separate this (by Minkowski's inequality) into various contributions which we estimate separately. First, of course,

$$\bigg(\int_0^\infty e^{-2\varepsilon t}A^2\,dt\bigg)^{1/2} = \frac{A}{\sqrt{2\varepsilon}}. \tag{7.43}$$

Next, for any $T\geqslant 1$, by Step 1 and

$$\int_0^t K_1(t,\tau)\,d\tau \leqslant \frac{Cc(1+t)}{\alpha},$$

we have

$$\begin{aligned}
&\bigg(\int_0^T e^{-2\varepsilon t}\bigg(\int_0^t K_1(t,\tau)\varphi(\tau)\,d\tau\bigg)^2 dt\bigg)^{1/2}\\
&\qquad\leqslant \Big(\sup_{0\leqslant t\leqslant T}\varphi(t)\Big)\bigg(\int_0^T e^{-2\varepsilon t}\bigg(\int_0^t K_1(t,\tau)\,d\tau\bigg)^2 dt\bigg)^{1/2}\\
&\qquad\leqslant CAe^{C(C_0C_WT/(\lambda_0-\lambda)+c(T+T^2))}\frac{c}{\alpha}\bigg(\int_0^\infty e^{-2\varepsilon t}(1+t)^2\,dt\bigg)^{1/2}\\
&\qquad\leqslant CA\frac{c}{\alpha\varepsilon^{3/2}}e^{C(C_0C_WT/(\lambda_0-\lambda)+c(T+T^2))}.
\end{aligned} \tag{7.44}$$

Invoking Jensen's inequality and Fubini's theorem, we also have

$$\begin{aligned}
&\biggl(\int_T^\infty e^{-2\varepsilon t}\biggl(\int_0^t K_1(t,\tau)\varphi(\tau)\,d\tau\biggr)^2\,dt\biggr)^{1/2}\\
&\quad=\biggl(\int_T^\infty\biggl(\int_0^t K_1(t,\tau)e^{-\varepsilon(t-\tau)}e^{-\varepsilon\tau}\varphi(\tau)\,d\tau\biggr)^2\,dt\biggr)^{1/2}\\
&\quad\leqslant\biggl(\int_T^\infty\biggl(\int_0^t K_1(t,\tau)e^{-\varepsilon(t-\tau)}\,d\tau\biggr)\biggl(\int_0^t K_1(t,\tau)e^{-\varepsilon(t-\tau)}e^{-2\varepsilon\tau}\varphi(\tau)^2\,d\tau\biggr)\,dt\biggr)^{1/2}\\
&\quad\leqslant\biggl(\sup_{t\geqslant T}\int_0^t e^{-\varepsilon t}K_1(t,\tau)e^{\varepsilon\tau}\,d\tau\biggr)^{1/2}\\
&\qquad\times\biggl(\int_T^\infty\int_0^t K_1(t,\tau)e^{-\varepsilon(t-\tau)}e^{-2\varepsilon\tau}\varphi(\tau)^2\,d\tau\,dt\biggr)^{1/2}\\
&\quad=\biggl(\sup_{t\geqslant T}\int_0^t e^{-\varepsilon t}K_1(t,\tau)e^{\varepsilon\tau}\,d\tau\biggr)^{1/2}\\
&\qquad\times\biggl(\int_0^\infty\int_{\max\{\tau,T\}}^\infty K_1(t,\tau)e^{-\varepsilon(t-\tau)}e^{-2\varepsilon\tau}\varphi(\tau)^2\,dt\,d\tau\biggr)^{1/2}\\
&\quad\leqslant\biggl(\sup_{t\geqslant T}\int_0^t e^{-\varepsilon t}K_1(t,\tau)e^{\varepsilon\tau}\,d\tau\biggr)^{1/2}\\
&\qquad\times\biggl(\sup_{\tau\geqslant 0}\int_\tau^\infty e^{\varepsilon\tau}K_1(t,\tau)e^{-\varepsilon t}\,dt\biggr)^{1/2}\biggl(\int_0^\infty e^{-2\varepsilon\tau}\varphi(\tau)^2\,d\tau\biggr)^{1/2}.
\end{aligned}\tag{7.45}$$

(Basically we copied the proof of Young's inequality.) Similarly,

$$\begin{aligned}
&\biggl(\int_0^\infty e^{-2\varepsilon t}\biggl(\int_0^t K_0(t,\tau)\varphi(\tau)\,d\tau\biggr)^2\,dt\biggr)^{1/2}\\
&\quad\leqslant\biggl(\sup_{t\geqslant 0}\int_0^t e^{-\varepsilon t}K_0(t,\tau)e^{\varepsilon\tau}\,d\tau\biggr)^{1/2}\\
&\qquad\times\biggl(\sup_{\tau\geqslant 0}\int_\tau^\infty e^{\varepsilon\tau}K_0(t,\tau)e^{-\varepsilon t}\,dt\biggr)^{1/2}\biggl(\int_0^\infty e^{-2\varepsilon\tau}\varphi(\tau)^2\,d\tau\biggr)^{1/2}\\
&\quad\leqslant\biggl(\sup_{t\geqslant 0}\int_0^t K_0(t,\tau)\,d\tau\biggr)^{1/2}\biggl(\sup_{\tau\geqslant 0}\int_\tau^\infty K_0(t,\tau)\,dt\biggr)^{1/2}\biggl(\int_0^\infty e^{-2\varepsilon\tau}\varphi(\tau)^2\,d\tau\biggr)^{1/2}.
\end{aligned}\tag{7.46}$$

The last term is also split, this time according to $\tau\leqslant T$ or $\tau>T$:

$$\begin{aligned}
&\biggl(\int_0^\infty e^{-2\varepsilon t}\biggl(\int_0^T\frac{c_0\varphi(\tau)}{(1+\tau)^m}\,d\tau\biggr)^2\,dt\biggr)^{1/2}\\
&\qquad\leqslant c_0\Bigl(\sup_{0\leqslant\tau\leqslant T}\varphi(\tau)\Bigr)\biggl(\int_0^\infty e^{-2\varepsilon t}\biggl(\int_0^T\frac{d\tau}{(1+\tau)^m}\biggr)^2\,dt\biggr)^{1/2}\\
&\qquad\leqslant c_0\frac{CA}{\sqrt{\varepsilon}}e^{C(C_0C_WT/(\lambda_0-\lambda)+c(T+T^2))}C_m,
\end{aligned}\tag{7.47}$$

and

$$\begin{aligned}
&\biggl(\int_0^\infty e^{-2\varepsilon t}\biggl(\int_T^t \frac{c_0\varphi(\tau)\,d\tau}{(1+\tau)^m}\biggr)^2 dt\biggr)^{1/2}\\
&\quad= c_0\biggl(\int_0^\infty\biggl(\int_T^t e^{-\varepsilon(t-\tau)}\,\frac{e^{-\varepsilon\tau}\varphi(\tau)}{(1+\tau)^m}\,d\tau\biggr)^2 dt\biggr)^{1/2}\\
&\quad\leqslant c_0\biggl(\int_0^\infty\biggl(\int_T^t \frac{e^{-2\varepsilon(t-\tau)}}{(1+\tau)^{2m}}\,d\tau\biggr)\biggl(\int_T^t e^{-2\varepsilon\tau}\varphi(\tau)^2\,d\tau\biggr)\,dt\biggr)^{1/2}\\
&\quad\leqslant c_0\biggl(\int_0^\infty e^{-2\varepsilon t}\varphi(t)^2\,dt\biggr)^{1/2}\biggl(\int_0^\infty\int_T^t \frac{e^{-2\varepsilon(t-\tau)}}{(1+\tau)^{2m}}\,d\tau\,dt\biggr)^{1/2}\\
&\quad= c_0\biggl(\int_0^\infty e^{-2\varepsilon t}\varphi(t)^2\,dt\biggr)^{1/2}\biggl(\int_T^\infty \frac{1}{(1+\tau)^{2m}}\int_\tau^\infty e^{-2\varepsilon(t-\tau)}\,dt\,d\tau\biggr)^{1/2}\\
&\quad= c_0\biggl(\int_0^\infty e^{-2\varepsilon t}\varphi(t)^2\,dt\biggr)^{1/2}\biggl(\int_T^\infty \frac{d\tau}{(1+\tau)^{2m}}\biggr)^{1/2}\biggl(\int_0^\infty e^{-2\varepsilon s}\,ds\biggr)^{1/2}\\
&\quad= \frac{C_{2m}^{1/2}c_0}{T^{m-1/2}\sqrt{\varepsilon}}\biggl(\int_0^\infty e^{-2\varepsilon t}\varphi(t)^2\,dt\biggr)^{1/2}.
\end{aligned}\tag{7.48}$$

Gathering estimates (7.43)–(7.48), we deduce from (7.42) that

$$\begin{aligned}
\|\varphi(t)e^{-\varepsilon t}\|_{L^2(dt)} \leqslant \biggl(1+\frac{CC_0C_W}{\varkappa(\lambda_0-\lambda)^2}\biggr)\frac{CA}{\sqrt{\varepsilon}}\Bigl(1+\frac{c}{\alpha\varepsilon}+c_0C_m\Bigr)\\
\times e^{C(C_0C_WT/(\lambda_0-\lambda)+c(T+T^2))}+a\|\varphi(t)e^{-\varepsilon t}\|_{L^2(dt)},
\end{aligned}\tag{7.49}$$

where

$$\begin{aligned}
a=\biggl(1+\frac{CC_0C_W}{\varkappa(\lambda_0-\lambda)^2}\biggr)\biggl[\biggl(\sup_{t\geqslant T}\int_0^t e^{-\varepsilon t}K_1(t,\tau)e^{\varepsilon\tau}d\tau\biggr)^{1/2}\biggl(\sup_{\tau\geqslant 0}\int_\tau^\infty e^{\varepsilon\tau}K_1(t,\tau)e^{-\varepsilon t}\,dt\biggr)^{1/2}\\
+\biggl(\sup_{t\geqslant 0}\int_0^t K_0(t,\tau)\,d\tau\biggr)^{1/2}\biggl(\sup_{\tau\geqslant 0}\int_\tau^\infty K_0(t,\tau)\,dt\biggr)^{1/2}+\frac{C_{2m}^{1/2}c_0}{T^{m-1/2}\sqrt{\varepsilon}}\biggr].
\end{aligned}$$

Using Propositions 7.1 (case $\gamma>1$) and 7.5, as well as assumptions (7.22) and (7.23), we see that $a\leqslant\frac{1}{2}$ for χ small enough and T satisfying (7.25). Then from (7.49) follows that

$$\|\varphi(t)e^{-\varepsilon t}\|_{L^2(dt)} \leqslant \biggl(1+\frac{CC_0C_W}{\varkappa(\lambda_0-\lambda)^2}\biggr)\frac{CA}{\sqrt{\varepsilon}}\Bigl(1+\frac{c}{\alpha\varepsilon}+c_0C_m\Bigr)e^{C(C_0C_WT/(\lambda_0-\lambda)+c(T+T^2))}.$$

Step 3. *Refined pointwise bounds.* Let us use (7.21) a third time, now for $t \geqslant T$:

$$\begin{aligned} e^{-\varepsilon t}\varphi(t) &\leqslant Ae^{-\varepsilon t} + \int_0^t \Big(\sup_{k\in\mathbb{Z}^d_*} |K^0(k,t-\tau)| e^{2\pi\lambda(t-\tau)|k|} \Big) \varphi(\tau) e^{-\varepsilon\tau}\, d\tau \\ &\quad + \int_0^t \left(K_0(t,\tau) + \frac{c_0}{(1+\tau)^m} \right) \varphi(\tau) e^{-\varepsilon\tau}\, d\tau \\ &\quad + \int_0^t e^{-\varepsilon t} K_1(t,\tau) e^{\varepsilon\tau} \varphi(\tau) e^{-\varepsilon\tau}\, d\tau \\ &\leqslant Ae^{-\varepsilon t} + \left[\left(\int_0^t \Big(\sup_{k\in\mathbb{Z}^d_*} |K^0(k,t-\tau)| e^{2\pi\lambda(t-\tau)|k|} \Big)^2 d\tau \right)^{1/2} \right. \\ &\quad + \left(\int_0^t K_0(t,\tau)^2\, d\tau \right)^{1/2} + \left(\int_0^\infty \frac{c_0^2}{(1+\tau)^{2m}}\, d\tau \right)^{1/2} \\ &\quad \left. + \left(\int_0^t e^{-2\varepsilon t} K_1(t,\tau)^2 e^{2\varepsilon\tau}\, d\tau \right)^{1/2} \right] \left(\int_0^\infty \varphi(\tau)^2 e^{-2\varepsilon\tau}\, d\tau \right)^{1/2}. \end{aligned} \tag{7.50}$$

We note that, for any $k \in \mathbb{Z}^d_*$,

$$\begin{aligned} (|K^0(k,t)| e^{2\pi\lambda|k|t})^2 &\leqslant 16\pi^4 |\widehat{W}(k)|^2\, |\tilde{f}^0(kt)|^2\, |k|^4 t^2 e^{4\pi\lambda|k|t} \\ &\leqslant CC_0^2 |\widehat{W}(k)|^2 e^{-4\pi(\lambda_0-\lambda)|k|t} |k|^4 t^2 \\ &\leqslant \frac{CC_0^2}{(\lambda_0-\lambda)^2} |\widehat{W}(k)|^2 e^{-2\pi(\lambda_0-\lambda)|k|t} |k|^2 \\ &\leqslant \frac{CC_0^2}{(\lambda_0-\lambda)^2} C_W^2 e^{-2\pi(\lambda_0-\lambda)|k|t} \\ &\leqslant \frac{CC_0^2}{(\lambda_0-\lambda)^2} C_W^2 e^{-2\pi(\lambda_0-\lambda)t}; \end{aligned}$$

so

$$\int_0^t \Big(\sup_{k\in\mathbb{Z}^d_*} |K^0(k,t-\tau)| e^{2\pi\lambda(t-\tau)|k|} \Big)^2 d\tau \leqslant \frac{CC_0^2 C_W^2}{(\lambda_0-\lambda)^3}.$$

Then the conclusion follows from (7.50), Corollary 7.4, conditions (7.25) and (7.23), and Step 2. □

Remark 7.9. Theorem 7.7 leads to enormous constants, and it is legitimate to ask about their sharpness, say with respect to the dependence in ε. We expect the constant to be roughly of the order of

$$\sup_t (e^{(ct)^{1/\gamma}} e^{-\varepsilon t}) \simeq \exp(\varepsilon^{-1/(\gamma-1)}).$$

Our bound is roughly like $\exp(\varepsilon^{-(4+2\gamma)/(\gamma-1)})$; this is worse, but displays the expected behavior as an exponential of an inverse power of ε, with a power that diverges like $O((1-\gamma)^{-1})$ as $\gamma \to 1$.

Remark 7.10. Even in the case of an analytic interaction, a similar argument suggests constants that are at best like $(\log(1/\varepsilon))^{\log(1/\varepsilon)}$, and this grows faster than any inverse power of $1/\varepsilon$.

To obtain sharper results, in §11 we shall later "break the norm" and work directly on the Fourier modes of, say, the spatial density. In this subsection we establish the estimates which will be used later; the reader who does not particularly care about the case $\gamma=1$ in Theorem 2.6 can skip them.

For any $\gamma\geqslant 1$, $\alpha>0$, $k,l\in\mathbb{Z}^d_*$ and $0\leqslant\tau\leqslant t$, we define

$$K^{(\alpha),\gamma}_{k,l}(t,\tau)=\frac{(1+\tau)de^{-\alpha|l|}e^{-\alpha(t-\tau)|k-l|/t}e^{-\alpha|k(t-\tau)+l\tau|}}{1+|k-l|^\gamma}. \tag{7.51}$$

We start by exponential moment estimates.

Proposition 7.11. *Let $\gamma\in[1,\infty)$. For any $\alpha\in(0,1)$ and $k,l\in\mathbb{Z}^d_*$, let $K^{(\alpha),\gamma}_{k,l}$ be defined by* (7.51). *Then there is $\bar\alpha=\bar\alpha(\gamma)>0$ such that if $\alpha\leqslant\bar\alpha$ and $\varepsilon\in\big(0,\frac14\alpha\big)$ then for any $t>0$,*

$$\sup_{k\in\mathbb{Z}^d_*}\sum_{l\in\mathbb{Z}^d_*}e^{-\varepsilon t}\int_0^t K^{(\alpha),\gamma}_{k,l}(t,\tau)e^{\varepsilon\tau}\,d\tau\leqslant\frac{C(d,\gamma)}{\alpha^{1+d}\varepsilon^{\gamma+1}t^\gamma}, \tag{7.52}$$

$$\sup_{k\in\mathbb{Z}^d_*}\sum_{l\in\mathbb{Z}^d_*}e^{-\varepsilon t}\bigg(\int_0^t K^{(\alpha),\gamma}_{k,l}(t,\tau)^2e^{2\varepsilon\tau}\,d\tau\bigg)^{1/2}\leqslant\frac{C(d,\gamma)}{\alpha^d\varepsilon^{\gamma+1/2}t^{\gamma-1/2}}, \tag{7.53}$$

$$\sup_{k\in\mathbb{Z}^d_*}\sum_{l\in\mathbb{Z}^d_*}\sup_{\tau\geqslant0}e^{\varepsilon\tau}\int_\tau^\infty K^{(\alpha),\gamma}_{k,l}e^{-\varepsilon t}\,dt\leqslant\frac{C(d,\gamma)}{\alpha^{2+d}\varepsilon}. \tag{7.54}$$

Proof. We first reduce to the case $d=1$. Monotonicity cannot be used now, but we note that

$$K^{(\alpha),\gamma}_{k,l}(t,\tau)\leqslant\sum_{j=1}^d e^{-\alpha|l_1|}e^{-\alpha|l_2|}\dots e^{-\alpha|l_{j-1}|}\,K^{(\alpha),\gamma}_{k_j,l_j}(t,\tau)e^{-\alpha|l_{j+1}|}\dots e^{-\alpha|l_d|},$$

where K_{k_j,l_j} stands for a 1-dimensional kernel. Thus

$$\begin{aligned}\sup_{k\in\mathbb{Z}^d_*}\sum_{l\in\mathbb{Z}^d_*}&\int_0^t e^{-\varepsilon t}K^{(\alpha),\gamma}_{k,l}(t,\tau)e^{\varepsilon\tau}\,d\tau\\&\leqslant\sup_{k\in\mathbb{Z}^d_*}\bigg(\sum_{m\in\mathbb{Z}^d}e^{-\alpha|m|}\bigg)^{d-1}\sum_{j=1}^d\sum_{l_j\in\mathbb{Z}}\int_0^t e^{-\varepsilon t}K^{(\alpha),\gamma}_{k_j,l_j}(t,\tau)e^{\varepsilon\tau}\,d\tau\\&\leqslant\frac{C(d)}{\alpha^{d-1}}\sup_{1\leqslant j\leqslant d}\sup_{k_j\in\mathbb{Z}}\sum_{l_j\in\mathbb{Z}}\int_0^t e^{-\varepsilon t}K^{(\alpha,\gamma)}_{k_j,l_j}(t,\tau)e^{\varepsilon\tau}\,d\tau.\end{aligned}$$

In other words, for (7.52) we may just consider the 1-dimensional case, provided we allow for an extra multiplicative constant $C(d)/\alpha^{d-1}$. A similar reasoning holds for (7.53) and (7.54). From now on we focus on the case $d=1$.

Without loss of generality we assume that $k>0$, and only treat the worst case $l<0$. (The other case $k,l>0$ is simpler and yields an exponential decay in time of the form $e^{-c\min\{\alpha,\varepsilon\}t}$). For simplicity we also write $K_{k,l}=K_{k,l}^{(\alpha),\gamma}$. An easy computation yields

$$e^{-\varepsilon t}\int_0^t K_{k,l}(t,\tau)e^{\varepsilon\tau}\,d\tau \leqslant \frac{Ce^{-\alpha|l|}}{1+|k-l|^\gamma}\left(\frac{1}{\alpha|k-l|}+\frac{|k|t}{\alpha|k-l|^2}+\frac{1}{\alpha^2|k-l|^2}\right)e^{-\varepsilon|l|t/|k-l|}.$$

Then, for any $k\geqslant 1$, we have (crudely writing $\alpha^2=O(\alpha)$)

$$\sum_{l=-\infty}^{-1}\int_0^t e^{-\varepsilon t}K_{k,l}(t,\tau)e^{\varepsilon\tau}\,d\tau \leqslant C\left(\sum_{l=1}^{\infty}\frac{e^{-\alpha l}e^{-\varepsilon lt/(k+l)}}{\alpha^2(k+l)^{1+\gamma}}+\sum_{l=1}^{\infty}\frac{e^{-\alpha l}e^{-\varepsilon lt/(k+l)}}{\alpha(k+l)^{2+\gamma}}kt\right). \tag{7.55}$$

For the first sum in the right-hand side of (7.55) we write

$$\sum_{l=1}^{\infty}\frac{e^{-\alpha l}e^{-\varepsilon lt/(k+l)}}{(k+l)^{1+\gamma}} \leqslant \sum_{l=1}^{\infty}\frac{e^{-\alpha l}}{l^{1+\gamma}}\left(\frac{\varepsilon lt}{k+l}\right)^{1+\gamma}e^{-\varepsilon lt/(k+l)}\frac{1}{(\varepsilon t)^{1+\gamma}} \leqslant \frac{C(\gamma)}{(\varepsilon t)^{1+\gamma}}. \tag{7.56}$$

For the second sum in the right-hand side of (7.55) we separate according to $1\leqslant l\leqslant k$ or $l\geqslant k+1$:

$$\begin{aligned}\sum_{l=1}^{k}\frac{e^{-\alpha l}e^{-\varepsilon lt/(k+l)}}{(k+l)^{2+\gamma}}kt &\leqslant \sum_{l=1}^{k}\frac{e^{-\alpha l}e^{-\varepsilon t/(k+1)}}{(k+1)^{2+\gamma}}kt\\ &\leqslant \frac{C}{\alpha}e^{-\varepsilon t/(k+1)}\left(\frac{\varepsilon t}{k+1}\right)^{1+\gamma}\frac{k}{k+1}\frac{t}{(\varepsilon t)^{1+\gamma}}\\ &\leqslant \frac{C}{\alpha\varepsilon^{1+\gamma}t^\gamma}\end{aligned} \tag{7.57}$$

and

$$\sum_{l=k+1}^{\infty}\frac{e^{-\alpha l}e^{-\varepsilon lt/(k+l)}}{(k+l)^{2+\gamma}}kt \leqslant C\sum_{l=k+1}^{\infty}\frac{e^{-\alpha l}e^{-\varepsilon t/2}}{k^{2+\gamma}}kt \leqslant \frac{C}{\alpha}\frac{e^{-\varepsilon t/4}}{\varepsilon k^{1+\gamma}} \leqslant \frac{C}{\alpha\varepsilon^{1+\gamma}t^\gamma}. \tag{7.58}$$

The combination of (7.55)–(7.58) completes the proof of (7.52).

Now we turn to (7.53). The estimates are rather similar, since

$$K_{k,l}(t,\tau)^2 \leqslant C(1+t)K_{k,l}(t,\tau)$$

with $\gamma\mapsto 2\gamma$ and $\alpha\mapsto 2\alpha$. So (7.55) should be replaced by

$$\begin{aligned}&\sum_{l\in\mathbb{Z}}e^{-\varepsilon t}\left(\int_0^t K_{k,l}(t,\tau)^2e^{2\varepsilon\tau}\,d\tau\right)^{1/2}\\ &\quad\leqslant C\left(\sum_{l=1}^{\infty}\frac{e^{-\alpha l}e^{-\varepsilon lt/(k+l)}(1+t)^{1/2}}{\alpha(k+l)^{\gamma+1/2}}+\sum_{l=1}^{\infty}\frac{e^{-\alpha l}e^{-\varepsilon lt/(k+l)}(kt)^{1/2}(1+t)^{1/2}}{\alpha^{1/2}(k+l)^{1+\gamma}}\right).\end{aligned} \tag{7.59}$$

For the first sum we use (7.56) with γ replaced by $\gamma-\frac{1}{2}$: for $t\geqslant 1$,

$$(1+t)^{1/2}\sum_{l=1}^{\infty}\frac{e^{-\alpha l}e^{-\varepsilon lt/(k+l)}}{(k+l)^{1+\gamma-1/2}}\leqslant\frac{C(\gamma)t^{1/2}}{(\varepsilon t)^{\gamma+1/2}}\leqslant\frac{C(\gamma)}{\varepsilon^{\gamma+1/2}t^{\gamma}}. \tag{7.60}$$

For the second sum in the right-hand side of (7.59) we write

$$\begin{aligned}\sum_{l=1}^{k}\frac{e^{-\alpha l}e^{-\varepsilon lt/(k+l)}k^{1/2}t}{(k+l)^{1+\gamma}}&\leqslant C\sum_{l=1}^{\infty}e^{-\alpha l}e^{-\varepsilon t/(k+1)}\left(\frac{\varepsilon t}{k+1}\right)^{\gamma+1/2}\frac{k^{1/2}}{(1+k)^{1/2}}\frac{t}{(\varepsilon t)^{\gamma+1/2}}\\&\leqslant\frac{C}{\alpha\varepsilon^{\gamma+1/2}t^{\gamma-1/2}}\end{aligned}$$

and

$$\sum_{l=k+1}^{\infty}\frac{e^{-\alpha l}e^{-\varepsilon lt/(k+l)}k^{1/2}t}{(k+l)^{1+\gamma}}\leqslant C\sum_{l=1}^{\infty}\frac{e^{-\alpha l}}{l^{\gamma+1/2}}e^{-\varepsilon t/2}t\leqslant Ce^{-\varepsilon t/4}\leqslant\frac{C}{(\varepsilon t)^{\gamma-1/2}\varepsilon}.$$

With this, (7.53) is readily obtained.

Finally we consider (7.54). As in Proposition 7.5, one easily shows that

$$\sup_{k\in\mathbb{Z}_*}\sum_{l\in\mathbb{Z}_*}\sup_{\tau\geqslant 0}e^{\varepsilon\tau}\int_{2\tau}^{\infty}K_{k,l}(t,\tau)e^{-\varepsilon t}\,dt\leqslant\frac{C}{\varepsilon\alpha^2}\sum_{l\in\mathbb{Z}_*}e^{-\alpha|l|}\leqslant\frac{C}{\varepsilon\alpha^3}.$$

Then one has

$$e^{\varepsilon\tau}\int_{\tau}^{2\tau}e^{-\alpha|k(t-\tau)+l\tau|}e^{-\varepsilon t}\,dt\leqslant\frac{C}{\alpha^2k}+\frac{C}{\alpha\varepsilon k}+\frac{C}{\alpha k}e^{-\varepsilon l\tau/k}.$$

So the problem amounts to estimate

$$\sum_{l\in\mathbb{Z}_*}\sup_{\tau\geqslant 0}(1+\tau)\frac{e^{-\alpha l}e^{-\varepsilon l\tau/k}}{\alpha k(k+l)^{\gamma}}\leqslant\sum_{l\in\mathbb{Z}_*}e^{-\alpha l}\left(\frac{1}{\alpha}+\frac{1}{\varepsilon l(k+l)^{\gamma}}e^{-\varepsilon l\tau/k}\frac{\varepsilon l\tau}{k}\right)\leqslant C\left(\frac{1}{\alpha^2}+\frac{1}{\varepsilon}\right),$$

and the proof is complete. □

We conclude this section with a mode-by-mode analogue of Theorem 7.7.

THEOREM 7.12. *Let $f^0=f^0(v)$ and $W=W(x)$ satisfy condition* **(L)** *from* §2.2 *with constants C_0, λ_0 and $\varkappa$; in particular $|\tilde{f}^0(\eta)|\leqslant C_0e^{-2\pi\lambda_0|\eta|}$. Further let*

$$C_W=\max\left\{\sum_{k\in\mathbb{Z}_*^d}|\widehat{W}(k)|,\sup_{k\in\mathbb{Z}_*^d}|k|\,|\widehat{W}(k)|\right\}.$$

Let $(A_k)_{k\in\mathbb{Z}^d_*}$, $\mu\geqslant 0$ *and* $\lambda\in(0,\lambda^*]$ *with* $0<\lambda^*<\lambda_0$. *Let* $(\Phi(k,t))_{k\in\mathbb{Z}^d,t\geqslant 0}$ *be a continuous function of* $t\geqslant 0$, *valued in* $\mathbb{C}^{\mathbb{Z}^d}$, *such that for all* $t\geqslant 0$, $\Phi(0,t)=0$ *and for any* $k\in\mathbb{Z}^d_*$,

$$\begin{aligned} e^{2\pi(\lambda t+\mu)|k|}&\biggl|\Phi(k,t)-\int_0^t K^0(k,t-\tau)\Phi(k,\tau)\,d\tau\biggr| \\ &\leqslant A_k+\int_0^t K_0(t,\tau)e^{2\pi(\lambda\tau+\mu)|k|}|\Phi(k,\tau)|\,d\tau \\ &\quad+\int_0^t\sum_{l\in\mathbb{Z}^d_*}\biggl(cK^{(\alpha),\gamma}_{k,l}(t,\tau)+\frac{c_l}{(1+\tau)^m}\biggr)e^{2\pi(\lambda\tau+\mu)|k-l|}|\Phi(k-l,\tau)|\,d\tau, \end{aligned} \tag{7.61}$$

where $c>0$, $c_l\geqslant 0$ *(for* $l\in\mathbb{Z}^d_*$*)*, $m>1$, $\gamma\geqslant 1$, $K_0(t,\tau)$ *is a non-negative kernel,* $K^{(\alpha),\gamma}_{k,l}$ *is defined by* (7.51) *and* $\alpha<\bar\alpha(\gamma)$ *is defined in Proposition* 7.11. *Then there are positive constants* C *and* χ, *depending only on* γ, λ^*, λ_0, $\varkappa$,

$$\bar c:=\max\biggl\{\sum_{l\in\mathbb{Z}^d_*}c_l,\biggl(\sum_{l\in\mathbb{Z}^d_*}c_l^2\biggr)^{1/2}\biggr\},$$

C_W *and* m *such that if*

$$\sup_{t\geqslant 0}\int_0^t K_0(t,\tau)\,d\tau\leqslant\chi \tag{7.62}$$

and

$$\sup_{t\geqslant 0}\biggl(\int_0^t K_0(t,\tau)^2\,d\tau\biggr)^{1/2}+\sup_{\tau\geqslant 0}\int_\tau^\infty K_0(t,\tau)\,dt\leqslant 1, \tag{7.63}$$

then for any $\varepsilon\in\bigl(0,\frac14\alpha\bigr)$ *and for any* $t\geqslant 0$,

$$\sup_{k\in\mathbb{Z}^d_*}|\Phi(k,t)|e^{2\pi(\lambda t+\mu)|k|}\leqslant C\bar A\frac{(1+\bar c^2)}{\sqrt{\varepsilon}}e^{C\bar c}\Bigl(1+\frac{c}{\alpha^2\varepsilon}\Bigr)e^{CT}e^{Cc(1+T^2)/\alpha}e^{\varepsilon t}, \tag{7.64}$$

where $\bar A:=\sup_k A_k$ *and*

$$T=C\max\biggl\{\biggl(\frac{c^2}{\alpha^{3+2d}\varepsilon^{\gamma+2}}\biggr)^{1/\gamma},\biggl(\frac{c}{\alpha^d\varepsilon^{\gamma+1/2}}\biggr)^{1/(\gamma-1/2)},\biggl(\frac{\bar c^2}{\varepsilon}\biggr)^{1/(2m-1)}\biggr\}. \tag{7.65}$$

Proof. The proof is quite similar to the proof of Theorem 7.7, so we shall only point out the differences. As in the proof of Theorem 7.7, we start by crude pointwise bounds obtained by Grönwall's inequality; but this time on the quantity

$$\varphi(t)=\sup_{k\in\mathbb{Z}^d_*}|\Phi(k,t)|e^{2\pi(\lambda t+\mu)|k|}.$$

Since

$$\sum_{l\in\mathbb{Z}^d_*} K_{k,l}(t,\tau) = O\biggl(\frac{1+\tau}{\alpha}\biggr),$$

we find that

$$\varphi(t) \leqslant 2\bar{A} e^{C(C_0C_Wt/(\lambda_0-\lambda)+c(t+t^2)/\alpha+\bar{c}C_m)}. \tag{7.66}$$

We next define Ψ_k, $\mathcal{K}^0_k$ and R_k as in Step 2 of the proof of Theorem 7.7, and we deduce (7.39) and (7.40). Let

$$\varphi_k(t) = |\Phi(k,t)| e^{2\pi(\lambda t+\mu)|k|}. \tag{7.67}$$

Then

$$\|\varphi_k(t)e^{-\varepsilon t}\|_{L^2(dt)} \leqslant \|R_k\|_{L^2(dt)}\biggl(1+\frac{\|\mathcal{K}^0_k\|_{L^1(dt)}}{\varkappa}\biggr) \leqslant \|R_k\|_{L^2(dt)}\biggl(1+\frac{CC_WC_0}{\varkappa}\biggr);$$

whence

$$\begin{aligned}\|\varphi_k(t)e^{-\varepsilon t}\|_{L^2(dt)} \leqslant \biggl(1+\frac{CC_0C_W}{\varkappa(\lambda_0-\lambda)^2}\biggr)\biggl[\int_0^\infty e^{-2\varepsilon t}\biggl(A_k+\int_0^t K_0(t,\tau)\varphi_k(\tau)\,d\tau \\ +\sum_{l\in\mathbb{Z}^d_*}\int_0^t\biggl(cK_{k,l}(t,\tau)+\frac{c_l}{(1+\tau)^m}\biggr)\varphi_{k-l}(\tau)\,d\tau\biggr)^2 dt\biggr]^{1/2}.\end{aligned} \tag{7.68}$$

We separate this into various contributions as in the proof of Theorem 7.7. In particular, using (7.66) and

$$\int_0^t \sum_{l\in\mathbb{Z}^d_*} K_{k,l}\,d\tau = O\biggl(\frac{1+t}{\alpha^2}\biggr),$$

we find that

$$\begin{aligned}&\biggl[\int_0^T e^{-2\varepsilon t}\biggl(\int_0^t\sum_{l\in\mathbb{Z}^d_*} K_{k,l}(t,\tau)\varphi_{k-l}(\tau)\,d\tau\biggr)^2 dt\biggr]^{1/2} \\ &\qquad\leqslant \Bigl(\sup_{k\in\mathbb{Z}^d_*}\sup_{0\leqslant t\leqslant T}\varphi_k(t)\Bigr)\biggl(\int_0^T e^{-2\varepsilon t}\biggl(\int_0^t\sum_{l\in\mathbb{Z}^d_*}K_{k,l}(t,\tau)\,d\tau\biggr)^2 dt\biggr)^{1/2} \\ &\qquad\leqslant C\bar{A}\frac{c}{\alpha^2\varepsilon^{3/2}} e^{C(C_0C_WT/(\lambda_0-\lambda)+c(T+T^2)/\alpha)}.\end{aligned} \tag{7.69}$$

Also,

$$\begin{aligned}&\biggl[\int_T^\infty e^{-2\varepsilon t}\biggl(\int_0^t\sum_{l\in\mathbb{Z}^d_*} K_{k,l}(t,\tau)\varphi_{k-l}(\tau)\,d\tau\biggr)^2 dt\biggr]^{1/2} \\ &\qquad\leqslant \biggl(\sup_{t\geqslant T}\int_0^t e^{-\varepsilon t}\sum_{l\in\mathbb{Z}^d_*}K_{k,l}(t,\tau)e^{\varepsilon\tau}\,d\tau\biggr)^{1/2} \\ &\qquad\qquad\times\biggl(\int_T^\infty\int_0^t\sum_{l\in\mathbb{Z}^d_*}K_{k,l}(t,\tau)e^{-\varepsilon(t-\tau)}e^{-2\varepsilon\tau}\varphi_{k-l}(\tau)^2\,d\tau\,dt\biggr)^{1/2},\end{aligned}$$

and the last term inside the parentheses is

$$\sum_{l\in\mathbb{Z}^d_*}\int_0^\infty\Big(\int_{\max\{\tau,T\}}^\infty K_{k,l}(t,\tau)e^{-\varepsilon(t-\tau)}\,dt\Big)e^{-2\varepsilon\tau}\varphi_{k-l}(\tau)^2\,d\tau$$
$$\leqslant\Big(\sum_{l\in\mathbb{Z}^d_*}\sup_{\tau\geqslant0}\int_\tau^\infty K_{k,l}(t,\tau)e^{-\varepsilon(t-\tau)}\,dt\Big)\Big(\sup_{l\in\mathbb{Z}^d_*}\int_0^\infty e^{-2\varepsilon\tau}\varphi_l(\tau)^2\,d\tau\Big).$$

The computation for K_0 is the same as in the proof of Theorem 7.7, and the terms in $(1+\tau)^{-m}$ are handled in essentially the same way: simple computations yield

$$\begin{aligned}\Big[\int_0^\infty e^{-2\varepsilon t}\Big(\int_0^T\frac{\sum_{l\in\mathbb{Z}^d_*}c_l\varphi_{k-l}(\tau)}{(1+\tau)^m}\,d\tau\Big)^2dt\Big]^{1/2}&\\ \leqslant\Big(\sup_{0\leqslant\tau\leqslant T}\sup_{l\in\mathbb{Z}^d_*}\varphi_l(\tau)\Big)\Big[\int_0^\infty e^{-2\varepsilon t}\Big(\int_0^T\frac{\sum_{l\in\mathbb{Z}^d_*}c_l}{(1+\tau)^m}\,d\tau\Big)^2dt\Big]^{1/2}&\\ \leqslant\bar c\frac{C_m\bar A}{\sqrt\varepsilon}e^{C(C_0C_WT/(\lambda_0-\lambda)+c(T+T^2)/\alpha)}&\end{aligned}\tag{7.70}$$

and

$$\begin{aligned}\Big[\int_0^\infty e^{-2\varepsilon t}\Big(\int_T^t\frac{\sum_{l\in\mathbb{Z}^d_*}c_l\varphi_{k-l}(\tau)\,d\tau}{(1+\tau)^m}\Big)^2dt\Big]^{1/2}&\\ \leqslant\Big[\sup_{t\geqslant0}\sup_{l\in\mathbb{Z}^d_*}\Big(\int_T^t e^{-2\varepsilon\tau}\varphi_l(\tau)^2\,d\tau\Big)\Big(\sum_{l\in\mathbb{Z}^d_*}c_l\Big)^2\int_0^\infty\int_T^t\frac{e^{-2\varepsilon(t-\tau)}}{(1+\tau)^{2m}}\,d\tau\,dt\Big]^{1/2}&\\ \leqslant\bar c\Big(\frac{C_{2m}}{\varepsilon T^{2m-1}}\Big)^{1/2}\Big(\sup_{l\in\mathbb{Z}^d_*}\int_0^\infty e^{-2\varepsilon\tau}\varphi_l(\tau)^2\,d\tau\Big)^{1/2}.&\end{aligned}\tag{7.71}$$

All in all, we end up with

$$\begin{aligned}&\sup_{k\in\mathbb{Z}^d_*}\|\varphi_k(t)e^{-\varepsilon t}\|_{L^2(dt)}\\ &\quad\leqslant\Big(1+\frac{CC_0C_W}{\varkappa(\lambda_0-\lambda)^2}\Big)\frac{CA}{\sqrt\varepsilon}\Big(1+\frac{c}{\alpha^2\varepsilon}+\bar cC_m\Big)e^{C(C_0C_WT/(\lambda_0-\lambda)+c(T+T^2)/\alpha)}\\ &\qquad+a\sup_{k\in\mathbb{Z}^d_*}\|\varphi_k(t)e^{-\varepsilon t}\|_{L^2(dt)},\end{aligned}\tag{7.72}$$

where

$$\begin{aligned}a=\Big(1+\frac{CC_0C_W}{\varkappa(\lambda_0-\lambda)^2}\Big)\Big[&c^2\Big(\sup_{t\geqslant T}\sum_{l\in\mathbb{Z}^d_*}\int_0^t e^{-\varepsilon t}K_{k,l}(t,\tau)e^{\varepsilon\tau}\,d\tau\Big)^{1/2}\\ &\times\Big(\sum_{l\in\mathbb{Z}^d_*}\sup_{\tau\geqslant0}\int_\tau^\infty e^{\varepsilon\tau}K_{k,l}(t,\tau)e^{-\varepsilon t}\,dt\Big)^{1/2}\\ &+\Big(\sup_{t\geqslant0}\int_0^t K_0(t,\tau)\,d\tau\Big)^{1/2}\Big(\sup_{\tau\geqslant0}\int_\tau^\infty K_0(t,\tau)\,dt\Big)^{1/2}+\frac{C_{2m}^{1/2}\bar c_0}{T^{m-1/2}\sqrt\varepsilon}\Big].\end{aligned}$$

Applying Proposition 7.11, we see that $a \leqslant \frac{1}{2}$, as soon as T satisfies (7.65), and then we deduce from (7.72) a bound on $\sup_{k\in\mathbb{Z}^d_*} \|\varphi_k(t)e^{-\varepsilon t}\|_{L^2(dt)}$.

Finally we conclude as in Step 3 of the proof of Theorem 7.7: from (7.61), we get

$$\begin{aligned} e^{-\varepsilon t}\varphi_k(t) \leqslant A_k e^{-\varepsilon t} + \Bigg[\Bigg(\int_0^t \Big(\sup_{k\in\mathbb{Z}^d_*} |K^0(k,t-\tau)| e^{2\pi\lambda(t-\tau)|k|} \Big)^2 d\tau \Bigg)^{1/2} \\ + \Bigg(\int_0^t K_0(t,\tau)^2 \, d\tau \Bigg)^{1/2} + \bar{c} \Bigg(\int_0^\infty \frac{d\tau}{(1+\tau)^{2m}} \Bigg)^{1/2} \\ + c \sum_{l\in\mathbb{Z}^d_*} \Bigg(\int_0^t e^{-2\varepsilon t} K_{k,l}(t,\tau)^2 e^{2\varepsilon\tau} \, d\tau \Bigg)^{1/2} \Bigg] \\ \times \Bigg(\sup_{k\in\mathbb{Z}^d_*} \int_0^\infty \varphi_k(\tau)^2 e^{-2\varepsilon\tau} \, d\tau \Bigg)^{1/2}, \end{aligned} \tag{7.73}$$

and the conclusion follows by a new application of Proposition 7.11. □

8. Approximation schemes

Having defined a functional setting (§4) and identified several mathematical/physical mechanisms (§§5–7), we are prepared to fight the Landau damping problem. For that we need an approximation scheme solving the non-linear Vlasov equation. The problem is not to prove the existence of solutions (this is much easier), but to devise the scheme in such a way that it leads to relevant estimates for our study.

The first idea which may come to mind is a classical Picard scheme for quasilinear equations:

$$\partial_t f^{n+1} + v\cdot\nabla_x f^{n+1} + F[f^n]\cdot\nabla_v f^{n+1} = 0. \tag{8.1}$$

This has two drawbacks: first, f^{n+1} evolves by the characteristics created by $F[f^n]$, and this will deteriorate the estimates in analytic regularity. Secondly, there is no hope to get a closed (or approximately closed) equation on the density associated with f^{n+1}. More promising, and more in the spirit of the linearized approach, would be a scheme like

$$\partial_t f^{n+1} + v\cdot\nabla_x f^{n+1} + F[f^{n+1}]\cdot\nabla_v f^n = 0. \tag{8.2}$$

(Physically, f^{n+1} forces f^n, and the question is whether the reaction will exhaust f^{n+1} in large time.) But when we write (8.2) we are implicitly treating a higher-order term $(\nabla_v f)$ of the equation in a perturbative way; so this has no reason to converge.

To circumvent these difficulties, we shall use a *Newton iteration*: not only will this provide more flexibility in the regularity indices, but at the same time it will yield an extremely fast rate of convergence (something like $O(\varepsilon^{2^n})$) which will be most welcome to absorb the large constants coming from Theorem 7.7 or 7.12.

8.1. The natural Newton scheme

Let us adapt the abstract Newton scheme to an abstract evolution equation in the form

$$\frac{\partial f}{\partial t} = Q(f),$$

around a stationary solution f^0 (so $Q(f^0)=0$). Write the Cauchy problem with initial datum $f_i \simeq f^0$ in the form

$$\Phi(f) := (\partial_t f - Q(f), f(0,\cdot)) - (0, f_i).$$

Starting from f^0, the Newton iteration consists in solving inductively

$$\Phi(f^{n-1}) + \Phi'(f^{n-1})\cdot(f^n - f^{n-1}) = 0 \quad \text{for } n \geqslant 1.$$

More explicitly, writing $h^n = f^n - f^{n-1}$, we should solve

$$\begin{cases} \partial_t h^1 = Q'(f^0)\cdot h^1, \\ h^1(0,\cdot) = f_i - f^0, \end{cases}$$

and, for all $n \geqslant 1$,

$$\begin{cases} \partial_t h^{n+1} = Q'(f^n)\cdot h^{n+1} - [\partial_t f^n - Q(f^n)], \\ h^{n+1}(0,\cdot) = 0. \end{cases}$$

By induction, for $n \geqslant 1$, this is the same as

$$\begin{cases} \partial_t h^{n+1} = Q'(f^n)\cdot h^{n+1} + [Q(f^{n-1}+h^n) - Q(f^{n-1}) - Q'(f^{n-1})\cdot h^n], \\ h^{n+1}(0,\cdot) = 0. \end{cases}$$

This is easily applied to the non-linear Vlasov equation, for which the non-linearity is quadratic. So we define the *natural Newton scheme for the non-linear Vlasov equation* as follows:

$$f^0 = f^0(v) \text{ is given (homogeneous stationary state)}$$

and

$$f^n = f^0 + h^1 + \ldots + h^n,$$

where

$$\begin{cases} \partial_t h^1 + v\cdot\nabla_x h^1 + F[h^1]\cdot\nabla_v f^0 = 0, \\ h^1(0,\cdot) = f_i - f^0, \end{cases} \tag{8.3}$$

and, for all $n \geqslant 1$,

$$\begin{cases} \partial_t h^{n+1} + v\cdot\nabla_x h^{n+1} + F[f^n]\cdot\nabla_v h^{n+1} + F[h^{n+1}]\cdot\nabla_v f^n = -F[h^n]\cdot\nabla_v h^n, \\ h^{n+1}(0,\cdot) = 0. \end{cases} \tag{8.4}$$

Here $F[f]$ is the force field created by the particle distribution f, namely

$$F[f](t,x) = -\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} \nabla W(x-y) f(t,y,w)\, dw\, dy. \tag{8.5}$$

Note also that all the $\varrho^n = \int_{\mathbb{R}^d} h^n\, dv$ for $n \geqslant 1$ have zero spatial average.

8.2. Battle plan

The treatment of (8.3) was performed in §4.12. Now the problem is to handle all equations appearing in (8.4). This is much more complicated, because for $n \geqslant 1$ the background density f^n depends on t and x, instead of just v. As a consequence,

(a) equation (8.4) cannot be considered as a perturbation of free transport, because of the presence of $\nabla_v h^{n+1}$ in the left-hand side;

(b) the reaction term $F[h^{n+1}]\cdot\nabla_v f^n$ no longer has the simple product structure (function of x)$\times$(function of v), so it becomes harder to get hands on the homogenization phenomenon;

(c) because of spatial inhomogeneities, echoes will appear; they are all the more dangerous that, $\nabla_v f^n$ is unbounded as $t\to\infty$, *even in gliding regularity*. (It grows like $O(t)$, which is reminiscent of the observation made by Backus [4].)

The estimates in §§5–7 have been designed precisely to overcome these problems; however we still have a few conceptual difficulties to solve before applying these tools.

Recall the discussion in §4.11: the natural strategy is to propagate the bound

$$\sup_{\tau\geqslant 0} \|f_\tau\|_{\mathcal{Z}_\tau^{\lambda,\mu;1}} < \infty \tag{8.6}$$

along the scheme; this estimate contains in particular two crucial pieces of information:

- a control of $\varrho_\tau = \int_{\mathbb{R}^d} f_\tau\, dv$ in $\mathcal{F}^{\lambda\tau+\mu}$ norm,
- a control of $\langle f_\tau\rangle = \int_{\mathbb{T}^d} f_\tau\, dx$ in $\mathcal{C}^{\lambda;1}$ norm.

So the plan would be to try to inductively get estimates of each h^n in a norm like the one in (8.6), in such a way that h^n is extremely small as $n\to\infty$, and allowing a slight deterioration of the indices λ and μ as $n\to\infty$. Let us try to see how this would work: assuming that

$$\sup_{\tau\geqslant 0} \|h_\tau^k\|_{\mathcal{Z}_\tau^{\lambda_k,\mu_k;1}} \leqslant \delta_k \quad \text{for all } 0\leqslant k\leqslant n,$$

we should try to bound h_τ^{n+1}. To "solve" (8.4), we apply the classical method of characteristics: as in §5, we define $(X_{\tau,t}^n, V_{\tau,t}^n)$ as the solution of

$$\begin{cases} \dfrac{d}{dt} X_{\tau,t}^n(x,v) = V_{\tau,t}^n(x,v), & \dfrac{d}{dt} V_{\tau,t}^n(x,v) = F[f^n](t, X_{\tau,t}^n(x,v)), \\ X_{\tau,\tau}^n(x,v) = x, & V_{\tau,\tau}^n(x,v) = v. \end{cases}$$

Then (8.4) is equivalent to

$$\frac{d}{dt} h^{n+1}(t, X_{0,t}^n, V_{0,t}^n(x,v)) = \Sigma^{n+1}(t, X_{0,\tau}^n(x,v), V_{0,\tau}^n(x,v)), \tag{8.7}$$

where

$$\Sigma^{n+1}(t,x,v) = -F[h^{n+1}]\cdot\nabla_v f^n - F[h^n]\cdot\nabla_v h^n. \tag{8.8}$$

Integrating (8.7) in time and recalling that $h^{n+1}(0,\cdot)=0$, we get

$$h^{n+1}(t,X^n_{0,t}(x,v),V^n_{0,t}(x,v))=\int_0^t \Sigma^{n+1}(\tau,X^n_{0,\tau}(x,v),V^n_{0,\tau}(x,v))\,d\tau.$$

Composing with $(X^n_{t,0},V^n_{t,0})$ and using (5.2) yields

$$h^{n+1}(t,x,v)=\int_0^t \Sigma^{n+1}(\tau,X^n_{t,\tau}(x,v),V^n_{t,\tau}(x,v))\,d\tau.$$

We rewrite this using the deflection map

$$\Omega^n_{t,\tau}(x,v)=(X^n_{t,\tau},V^n_{t,\tau})(x+v(t-\tau),v)=S^n_{t,\tau}\circ S^0_{\tau,t};$$

then we finally obtain

$$\begin{aligned} h^{n+1}(t,x,v)&=\int_0^t(\Sigma^{n+1}_\tau\circ\Omega^n_{t,\tau})(x-v(t-\tau),v)\,d\tau\\ &=-\int_0^t[(F[h^{n+1}_\tau]\circ\Omega^n_{t,\tau})\cdot((\nabla_v f^n_\tau)\circ\Omega^n_{t,\tau})](x-v(t-\tau),v)\,d\tau\\ &\quad-\int_0^t[(F[h^n_\tau]\circ\Omega^n_{t,\tau})\cdot((\nabla_v h^n_\tau)\circ\Omega^n_{t,\tau})](x-v(t-\tau),v)\,d\tau. \end{aligned}\tag{8.9}$$

Since the unknown h^{n+1} appears on both sides of (8.9), we need to get a self-consistent estimate. For this we have little choice but to integrate in v and get an integral equation on $\varrho[h^{n+1}]=\int_{\mathbb{R}^d}h^n\,dv$, namely

$$\begin{aligned} \varrho[h^{n+1}](t,x)=\int_0^t\int_{\mathbb{R}^d}[((\varrho[h^{n+1}_\tau]*\nabla W)\circ\Omega^n_{t,\tau})\cdot G^n_{\tau,t}]\circ S^0_{\tau-t}(x,v)\,dv\,d\tau\\ +(\text{terms from stage } n), \end{aligned}\tag{8.10}$$

where $G^n_{\tau,t}=\nabla_v f^n_\tau\circ\Omega^n_{t,\tau}$. By induction hypothesis, $G^n_{\tau,t}$ is smooth with regularity indices roughly equal to λ_n and μ_n; so if we accept to lose just a bit more on the regularity we may hope to apply the long-term regularity extortion and decay estimates from §6, and then time-response estimates of §7, and get the desired damping.

However, we are facing a major problem: composition of $\varrho[h^{n+1}_\tau]*\nabla W$ by $\Omega^n_{t,\tau}$ implies a loss of regularity in the right-hand side with respect to the left-hand side, which is of course unacceptable if one wants a closed estimate. The short-term regularity extortion from §6 remedies this, but the price to pay is that G^n should now be estimated at time $\tau'=\tau-bt/(1+b)$ instead of τ, and with index of gliding analytic regularity roughly equal to $\lambda_n(1+b)$ rather than λ_n. Now the catch is that the error induced by composition

by Ω^n depends on *the whole distribution* f^n, not just h^n. Thus, if the parameter b should control this error it should stay of order 1 as $n\to\infty$, instead of converging to 0.

So it seems we are sentenced to lose a fixed amount of regularity (or rather of radius of convergence) in the transition from stage n to stage $n+1$; this is reminiscent of the "Nash–Moser syndrom" [2]. The strategy introduced by Nash [70] to remedy such a problem (in his case arising in the construction of C^∞ isometric embeddings) consisted of combining a Newton scheme with regularization; his method was later developed by Moser [67] for the C^∞ KAM theorem (see [68, pp. 19–21] for some interesting historical comments). A clear and concise proof of the Nash–Moser implicit function theorem, together with its application to the C^∞ embedding problem, can be found in [83]. The Nash–Moser method is arguably the most powerful perturbation technique known to this day. However, despite significant effort, we were unable to set up any relevant regularization procedure (in gliding regularity, of course) which could be used in Nash–Moser style, because of three serious problems:

• The convergence of the Nash–Moser scheme is no longer as fast as that of the "raw" Newton iteration; instead, it is determined by the regularity of the data, and the resulting rates would be unlikely to be fast enough to win over the gigantic constants coming from §7.

• Analytic regularization in the v variable is extremely costly, especially if we wish to keep a good localization in velocity space, as the one appearing in Theorem 4.20 (iii), that is exponential integrability in v; then the uncertainty principle basically forces us to pay $O(e^{C/\varepsilon^2})$, where ε is the strength of the regularization.

• Regularization comes with an increase of amplitude (there is as usual a trade-off between size and regularity); if we regularize before composition by Ω^n, this will devastate the estimates, because the analytic regularity of $f\circ g$ depends not only on the regularity of f and g, but also on the amplitude of $g-\mathrm{Id}$.

Fortunately, it turned out that a "raw" Newton scheme could be used; but this required us to give up the natural estimate (8.6), and replace it by the *pair* of estimates

$$\begin{cases} \sup\limits_{\tau\geqslant 0} \|\varrho_\tau\|_{\mathcal{F}^{\lambda\tau+\mu}} < \infty, \\ \sup\limits_{0\leqslant\tau\leqslant t} \|f_\tau\circ\Omega_{t,\tau}\|_{\mathcal{Z}^{\bar\lambda(1+b),\bar\mu;1}_{\tau-bt/(1+b)}} < \infty. \end{cases} \tag{8.11}$$

Here $b=b(t)$ takes the form constant$/(1+t)$, and is kept fixed all along the scheme; moreover λ and μ will be slightly larger than $\bar\lambda$ and $\bar\mu$, so that none of the two estimates in (8.11) implies the other one. Note carefully that there are now two times (t and τ) explicitly involved, so this is much more complex than (8.6). Let us explain why this strategy is nonetheless workable.

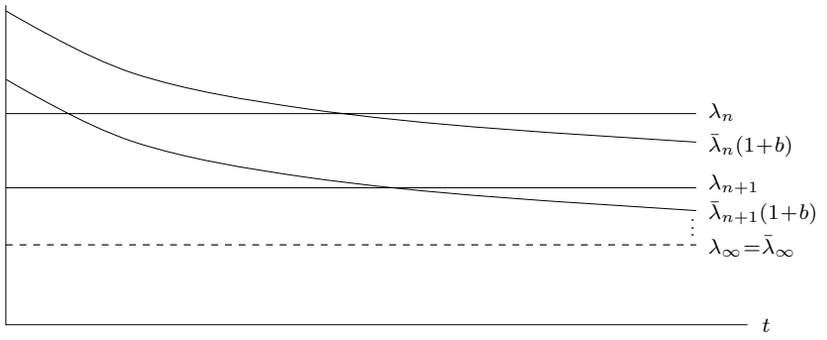

Figure 7. Indices of gliding regularity appearing throughout our Newton scheme, in the norm of $\varrho[h_\tau]$ and in the norm of $h_\tau\circ\Omega_{t,\tau}$, respectively, plotted as functions of t.

First, the density $\varrho^n=\int_{\mathbb{R}^d} f^n\,dv$ determines the characteristics at stage n, and therefore the associated deflection Ω^n. If ϱ^n_τ is bounded in $\mathcal{F}^{\lambda_n\tau+\mu}$, then by Theorem 5.2 we can estimate $\Omega^n_{t,\tau}$ in $\mathcal{Z}^{\lambda'_n,\mu'_n}_{\tau'}$, as soon as (essentially) $\lambda'_n\tau'+\mu'_n\leqslant\lambda_n\tau+\mu_n$ and $\lambda'_n<\lambda_n$, and these bounds are uniform in t.

Of course, we cannot apply this theorem in the present context, because $\bar\lambda_n(1+b)$ is not bounded above by λ_n. However, *for large times* t we may afford $\bar\lambda_n(1+b(t))<\lambda_n$, while $\bar\lambda_n(1+b)(\tau-bt/(1+b))\leqslant\lambda_n\tau$ for all times; this will be sufficient to repeat the arguments in §5, getting uniform estimates *in a regularity which depends on* t. (The constants are uniform in t; but the index of regularity goes decreases with t.) We can also do this while preserving the other good properties from Theorem 5.2, namely exponential decay in τ, and vanishing near $\tau=t$.

Figure 7 summarizes schematically the way we choose and estimate the gliding regularity indices.

Besides being uniform in t, our bounds need to be uniform in n. For this we shall have to *stratify* all our estimates, that is decompose $\varrho[f^n]=\varrho[h^1]+...+\varrho[h^n]$, and consider separately the influence of each term in the equations for characteristics. This can work only if the scheme converges very fast.

Once we have estimates on $\Omega^n_{t,\tau}$ in a time-varying regularity, we can work with the kinetic equation to derive estimates on $h^n_\tau\circ\Omega^n_{t,\tau}$; and then on all $h^k_\tau\circ\Omega^n_{t,\tau}$, also in a norm of time-varying regularity. We can also estimate their spatial average, in a norm $\mathcal{C}^{\bar\lambda(1+b);1}$; due to the exponential convergence of the deflection map as $\tau\to\infty$ these estimates will turn out to be uniform in τ.

Next, we can use all this information, in conjunction with Theorem 6.5, to get an integral inequality on the norm of $\varrho[h^{n+1}_\tau]$ in $\mathcal{F}^{\lambda\tau+\mu}$, where λ and μ are only slightly smaller than λ_n and μ_n. Then we can go through the response estimates of §7, this gives us an arbitrarily small loss in the exponential decay rate, at the price of a huge constant

which will eventually be wiped out by the fast convergence of the scheme. So we have an estimate on $\varrho[h^{n+1}]$, and we are in business to continue the iteration. (To ensure the propagation of the linear damping condition, or equivalently of the smallness of K_0 in Theorem 7.7, throughout the scheme, we shall have to stratify the estimates once more.)

9. Local-in-time iteration

Before working out the core of the proof of Theorem 2.6 in §10, we shall need a short-time estimate, which will act as an "initial regularity layer" for the Newton scheme. (This will give us room later to allow the regularity index to depend on t.) So we run the whole scheme once in this section, and another time in the next section.

Short-time estimates in the analytic class are not new for the non-linear Vlasov equation: see in particular the work of Benachour [8] on the Vlasov–Poisson equation. His arguments can probably be adapted for our purpose; also the Cauchy–Kovalevskaya method could certainly be applied. We shall provide here an alternative method, based on the analytic function spaces from §4, but not needing the apparatus from §§5–7. Unlike the more sophisticated estimates which will be performed in §10, these ones are "almost" Eulerian (the only characteristics are those of free transport). The main tool is given by the following lemma.

LEMMA 9.1. *Let f be an analytic function, $\lambda(t)=\lambda-Kt$ and $\mu(t)=\mu-Kt$; let $T>0$ be so small that $\lambda(t),\mu(t)>0$ for $0\leqslant t\leqslant T$. Then for any $\tau\in[0,T]$ and any $p\geqslant 1$,*

$$\left.\frac{d^+}{dt}\right|_{t=\tau} \|f\|_{\mathcal{Z}_\tau^{\lambda(t),\mu(t);p}} \leqslant -\frac{K}{1+\tau}\|\nabla f\|_{\mathcal{Z}_\tau^{\lambda(\tau),\mu(\tau);p}}, \tag{9.1}$$

where d^+/dt stands for the upper right derivative.

Remark 9.2. Time-differentiating Lebesgue integrability exponents is common practice in certain areas of analysis; see e.g. [33]. Time-differentiation with respect to regularity exponents is less common; however, as pointed out to us by Strain, Lemma 9.1 is strongly reminiscent of a method recently used by Chemin [17] to derive local analytic regularity bounds for the Navier–Stokes equation. We expect that similar ideas can be applied to more general situations of Cauchy–Kovalevskaya type, especially for first-order equations, and maybe this has already been done.

Proof. For notational simplicity, let us assume $d=1$. The left-hand side of (9.1) is

$$\begin{aligned}
&\sum_{\substack{k\in\mathbb{Z}^d\\ n\in\mathbb{N}_0^d}} e^{2\pi\mu(\tau)|k|}2\pi\dot\mu(\tau)|k|\frac{\lambda^n(\tau)}{n!}\|(\nabla_v+2i\pi k\tau)^n\hat f(k,v)\|_{L^p(dv)}\\
&\qquad+\sum_{\substack{k\in\mathbb{Z}^d\\ n\in N_0^d}} e^{2\pi\mu(\tau)|k|}\dot\lambda(\tau)\frac{\lambda^{n-1}(\tau)}{(n-1)!}\|(\nabla_v+2i\pi k\tau)^n\hat f(k,v)\|_{L^p(dv)}\\
&\qquad\qquad\leqslant -K\sum_{\substack{k\in\mathbb{Z}^d\\ n\in\mathbb{N}_0^d}} e^{2\pi\mu(\tau)|k|}2\pi|k|\frac{\lambda^n(\tau)}{n!}\|(\nabla_v+2i\pi k\tau)^n\hat f(k,v)\|_{L^p(dv)}\\
&\qquad\qquad\quad -K\sum_{\substack{k\in\mathbb{Z}^d\\ n\in\mathbb{N}_0^d}} e^{2\pi\mu(\tau)|k|}\frac{\lambda^n(\tau)}{n!}\|(\nabla_v+2i\pi k\tau)^{n+1}\hat f(k,v)\|_{L^p(dv)}\\
&\qquad\qquad\leqslant -K\sum_{\substack{k\in\mathbb{Z}^d\\ n\in\mathbb{N}_0^d}} e^{2\pi\mu(\tau)|k|}\frac{\lambda^n(\tau)}{n!}\|(\nabla_v+2i\pi k\tau)^n\widehat{\nabla_x f}(k,v)\|_{L^p(dv)}\\
&\qquad\qquad\quad +\frac{K\tau}{1+\tau}\sum_{\substack{k\in\mathbb{Z}^d\\ n\in\mathbb{N}_0^d}} e^{2\pi\mu(\tau)|k|}\frac{\lambda^n(\tau)}{n!}\|(\nabla_v+2i\pi k\tau)^n\widehat{\nabla_x f}(k,v)\Big\|_{L^p(dv)}\\
&\qquad\qquad\quad -\frac{K}{1+\tau}\sum_{\substack{k\in\mathbb{Z}^d\\ n\in\mathbb{N}_0^d}} e^{2\pi\mu(\tau)|k|}\frac{\lambda^n(\tau)}{n!}\|(\nabla_v+2i\pi k\tau)^n\widehat{\nabla_v f}(k,v)\|_{L^p(dv)},
\end{aligned}$$

where in the last step we used that

$$\|(\nabla_v+2i\pi k\tau)h\|\geqslant\frac{\|\nabla_v h\|-\tau\|2i\pi kh\|}{1+\tau}.$$

The conclusion follows. □

Now let us see how to propagate estimates through the Newton scheme described in §10. The first stage of the iteration (h^1 in the notation of (8.3)) was considered in §4.12, so we only need to care about higher orders. For any $k\geqslant 1$, we solve

$$\partial_t h^{k+1}+v\cdot\nabla_x h^{k+1}=\widetilde\Sigma^{k+1},$$

where

$$\widetilde\Sigma^{k+1}=-(F[h^{k+1}]\cdot\nabla_v f^k+F[f^k]\cdot\nabla_v h^{k+1}+F[h^k]\cdot\nabla_v h^k)$$

(note the difference with (8.7)–(8.8)). Recall that $f^k=f^0+h^1+\ldots+h^k$. We define

$$\lambda_k(t)=\lambda_k-2Kt\quad\text{and}\quad\mu_k(t)=\mu_k-Kt,$$

where $(\lambda_k)_{k=1}^\infty$ and $(\mu_k)_{k=1}^\infty$ are decreasing sequences of positive numbers.

We assume inductively that at stage n of the iteration, we have constructed $(\lambda_k)_{k=1}^n$, $(\mu_k)_{k=1}^n$ and $(\delta_k)_{k=1}^n$ such that

$$\sup_{0\leqslant t\leqslant T} \|h^k(t,\cdot)\|_{\mathcal{Z}_t^{\lambda_k(t),\mu_k(t);1}} \leqslant \delta_k \quad \text{for all } 1\leqslant k\leqslant n$$

for some fixed $T>0$. The issue is to construct λ_{n+1}, μ_{n+1} and δ_{n+1} so that the induction hypothesis is satisfied at stage $n+1$.

At $t=0$, $h^{n+1}=0$. Then we estimate the time-derivative of $\|h^{n+1}\|_{\mathcal{Z}_t^{\lambda_{n+1}(t),\mu_{n+1}(t);1}}$. Let us first pretend that the regularity indices λ_{n+1} and μ_{n+1} do not depend on t; then

$$h^{n+1}(t) = \int_0^t \widetilde{\Sigma}^{n+1}\circ S^0_{-(t-\tau)}\,d\tau,$$

so, by Proposition 4.19,

$$\begin{aligned}\|h^{n+1}\|_{\mathcal{Z}_t^{\lambda_{n+1},\mu_{n+1};1}} &\leqslant \int_0^t \|\widetilde{\Sigma}^{n+1}_\tau\circ S^0_{-(t-\tau)}\|_{\mathcal{Z}_t^{\lambda_{n+1},\mu_{n+1};1}}\,d\tau\\ &\leqslant \int_0^t \|\widetilde{\Sigma}^{n+1}_\tau\|_{\mathcal{Z}_\tau^{\lambda_{n+1},\mu_{n+1};1}}\,d\tau,\end{aligned}$$

and thus

$$\frac{d^+}{dt}\|h^{n+1}\|_{\mathcal{Z}_t^{\lambda_{n+1},\mu_{n+1};1}} \leqslant \|\widetilde{\Sigma}^{n+1}_t\|_{\mathcal{Z}_t^{\lambda_{n+1},\mu_{n+1};1}}.$$

Finally, according to Lemma 9.1, to this estimate we should add a negative multiple of the norm of ∇h^{n+1} to take into account the time-dependence of λ_{n+1} and μ_{n+1}.

All in all, after application of Proposition 4.24, we get

$$\begin{aligned}\frac{d^+}{dt}\|h^{n+1}(t,\cdot)\|_{\mathcal{Z}_t^{\lambda_{n+1}(t),\mu_{n+1}(t);1}} \leqslant{}& \|F[h_t^{n+1}]\|_{\mathcal{F}^{\lambda_{n+1}t+\mu_{n+1}}}\,\|\nabla_v f_t^n\|_{\mathcal{Z}_t^{\lambda_{n+1},\mu_{n+1};1}}\\ &+\|F[f_t^n]\|_{\mathcal{F}^{\lambda_{n+1}t+\mu_{n+1}}}\,\|\nabla_v h_t^{n+1}\|_{\mathcal{Z}_t^{\lambda_{n+1},\mu_{n+1};1}}\\ &+\|F[h_t^n]\|_{\mathcal{F}^{\lambda_{n+1}t+\mu_{n+1}}}\,\|\nabla_v h_t^n\|_{\mathcal{Z}_t^{\lambda_{n+1},\mu_{n+1};1}}\\ &-K\|\nabla_x h_t^{n+1}\|_{\mathcal{Z}_t^{\lambda_{n+1},\mu_{n+1};1}} - K\|\nabla_v h_t^{n+1}\|_{\mathcal{Z}_t^{\lambda_{n+1},\mu_{n+1};1}},\end{aligned}$$

where $K>0$, t is sufficiently small, and all exponents λ_{n+1} and μ_{n+1} in the right-hand side actually depend on t.

From Proposition 4.15 (iv) we easily get $\|F[h]\|_{\mathcal{F}^{\lambda t+\mu}}\leqslant C\|\nabla h\|_{\mathcal{Z}_t^{\lambda,\mu;1}}$. Moreover, by Proposition 4.10,

$$\|\nabla f^n\|_{\mathcal{Z}_t^{\lambda_{n+1},\mu_{n+1};1}} \leqslant \sum_{k=1}^n \|\nabla h^k\|_{\mathcal{Z}_t^{\lambda_{n+1},\mu_{n+1};1}} \leqslant C\sum_{k=1}^n \frac{\|h^k\|_{\mathcal{Z}_t^{\lambda_{k+1},\mu_{k+1};1}}}{\min\{\lambda_k-\lambda_{n+1},\mu_k-\mu_{n+1}\}}.$$

We end up with the bound

$$\frac{d^+}{dt}\|h^{n+1}(t,\cdot)\|_{\mathcal{Z}_t^{\lambda_{n+1}(t),\mu_{n+1}(t);1}}$$
$$\leqslant \left(C\sum_{k=1}^{n}\frac{\delta_k}{\min\{\lambda_k-\lambda_{n+1},\mu_k-\mu_{n+1}\}}-K\right)\|\nabla h^{n+1}\|_{\mathcal{Z}_t^{\lambda_{n+1}(t),\mu_{n+1}(t);1}}$$
$$+\frac{\delta_n^2}{\min\{\lambda_n-\lambda_{n+1},\mu_n-\mu_{n+1}\}}.$$

We conclude that if

$$\sum_{k=1}^{n}\frac{\delta_k}{\min\{\lambda_k-\lambda_{n+1},\mu_k-\mu_{n+1}\}}\leqslant\frac{K}{C}, \tag{9.2}$$

then we may choose

$$\delta_{n+1}=\frac{\delta_n^2}{\min\{\lambda_n-\lambda_{n+1},\mu_n-\mu_{n+1}\}}. \tag{9.3}$$

This is our first encounter with the principle of "stratification" of errors, which will be crucial in the next section: to control the error at stage $n+1$, we use not only the smallness of the error from stage n, but also an information about all previous errors; namely the fact that the convergence of the size of the error is much faster than the convergence of the regularity loss. Let us see how this works. We choose $\lambda_k-\lambda_{k+1}=\mu_k-\mu_{k+1}=\Lambda/k^2$, where $\Lambda>0$ is arbitrarily small. Then for $k\leqslant n$, $\lambda_k-\lambda_{n+1}\geqslant\Lambda/k^2$, and therefore $\delta_{n+1}\leqslant\delta_n^2 n^2/\Lambda$. The problem is to check that

$$\sum_{n=1}^{\infty}n^2\delta_n<\infty. \tag{9.4}$$

Indeed, then we can choose K large enough for (9.2) to be satisfied, and then T small enough that, say, $\lambda^*-2KT\geqslant\lambda^\sharp$ and $\mu^*-KT\geqslant\mu^\sharp$, where $\lambda^\sharp<\lambda^*$ and $\mu^\sharp<\mu^*$ have been fixed in advance.

If $\delta_1=\delta$, the general term in the series of (9.4) is

$$n^2\frac{\delta^{2^n}}{\Lambda^n}(2^2)^{2^{n-1}}(3^2)^{2^{n-2}}(4^2)^{2^{n-2}}\dots((n-1)^2)^2n^2.$$

To prove the convergence for δ small enough, we assume by induction that $\delta_n\leqslant z^{a^n}$, where a is fixed in the interval $(1,2)$ (say $a=1.5$); and we claim that this condition propagates if $z>0$ is small enough. Indeed,

$$\delta_{n+1}\leqslant\frac{z^{2a^n}}{\Lambda}n^2\leqslant z^{a^{n+1}}\frac{z^{(2-a)a^n}n^2}{\Lambda},$$

and this is bounded above by $z^{a^{n+1}}$ if z is so small that

$$z^{(2-a)a^n} \leqslant \frac{\Lambda}{n^2} \quad \text{for all } n \in \mathbb{N}.$$

This concludes the iteration argument. Note that the convergence is still extremely fast—like $O(z^{a^n})$ for any $a<2$. (Of course, when a approaches 2, the constants become huge, and the restriction on the size of the perturbation becomes more and more stringent.)

Remark 9.3. The method used in this section can certainly be applied to more general situations of Cauchy–Kovalevskaya type. Actually, as pointed out to us by Bony and Gérard, the use of a regularity index which decays linearly in time, combined with a Newton iteration, was used by Nirenberg [73] to prove an abstract Cauchy–Kovalevskaya theorem. Nirenberg uses a time-integral formulation, so there is nothing in [73] comparable to Lemma 9.1, and the details of the proof of convergence differ from ours; but the general strategy is similar. Nirenberg's proof was later simplified by Nishida [74] with a clever fixed-point argument; in the present section anyway, our final goal is to provide short-term estimates for the successive corrections arising from the Newton scheme.

10. Global in time iteration

Now let us implement the scheme described in §8, with some technical modifications. If f is a given kinetic distribution, we write

$$\varrho[f] = \int_{\mathbb{R}^d} f \, dv \quad \text{and} \quad F[f] = -\nabla W * \varrho[f].$$

We let

$$f^n = f^0 + h^1 + \ldots + h^n, \tag{10.1}$$

where the successive corrections h^k are defined by the natural Newton scheme introduced in §8. As in §5, we define $\Omega^k_{t,\tau}$ as the deflection from time t to time τ, generated by the force field $F[f^k] = -\nabla W * \varrho[f^k]$. (Note that $\Omega^0 = \mathrm{Id}$.)

10.1. The statement of the induction

We shall fix $\bar{p} \in [1, \infty]$ and make the following assumptions:

- *Regularity of the background*: there are $\lambda > 0$ and $C_0 > 0$ such that

$$\|f^0\|_{\mathcal{C}^{\lambda;p}} \leqslant C_0 \quad \text{for all } p \in [1, \bar{p}].$$

• *Linear damping condition*: The stability condition **(L)** from §2.2 holds with parameters C_0, λ (the same as above) and $\varkappa>0$.

• *Regularity of the interaction*: There are $\gamma>1$ and $C_F>0$ such that for any $\nu>0$,

$$\|\nabla W*\varrho\|_{\mathcal{F}^{\nu,\gamma}}\leqslant C_F\|\varrho\|_{\dot{\mathcal{F}}^{\nu}}. \tag{10.2}$$

• *Initial layer of regularity* (coming from §9): having chosen $\lambda^\sharp<\lambda$ and $\mu^\sharp<\mu$, we assume that for all $p\in[1,\bar{p}]$,

$$\sup_{0\leqslant t\leqslant T}\big(\|h_t^k\|_{\mathcal{Z}^{\lambda^\sharp,\mu^\sharp;p}}+\|\varrho[h_t^k]\|_{\mathcal{F}^{\mu^\sharp}}\big)\leqslant\zeta_k\quad\text{for all } k\geqslant 1, \tag{10.3}$$

where T is *some* positive time, and ζ_k converges to zero extremely fast: $\zeta_k=O(z_I^{a_I^k})$, $z_I\leqslant C\delta<1$, $1<a_I<2$ (a_I chosen in advance, arbitrarily close to 2).

• *Smallness of the solution of the linearized equation* (coming from §4.12): given $\lambda_1<\lambda^\sharp$ and $\mu_1<\mu^\sharp$, we assume that

$$\begin{cases}\sup\limits_{\tau\geqslant 0}\|\varrho[h_\tau^1]\|_{\mathcal{F}^{\lambda_1\tau+\mu_1}}\leqslant\delta_1,\\ \sup\limits_{0\leqslant\tau\leqslant t}\|h_\tau^1\|_{\mathcal{Z}^{\lambda_1(1+b),\mu_1;p}_{\tau-bt/(1+b)}}\leqslant\delta_1\quad\text{for all } p\in[1,\bar{p}],\end{cases} \tag{10.4}$$

where $\delta_1\leqslant C\delta$.

Then we prove the following induction for any $n\geqslant 1$,

$$\begin{cases}\sup\limits_{\tau\geqslant 0}\|\varrho[h_\tau^k]\|_{\mathcal{F}^{\lambda_k\tau+\mu_k}}\leqslant\delta_k,\\ \sup\limits_{0\leqslant\tau\leqslant t}\|h_\tau^k\circ\Omega_{t,\tau}^{k-1}\|_{\mathcal{Z}^{\lambda_k(1+b),\mu_k;p}_{\tau-bt/(1+b)}}\leqslant\delta_k,\end{cases}\qquad\text{for all } k\in\{1,...,n\}\text{ and } p\in[1,\bar{p}], \tag{10.5}$$

where

• $(\delta_k)_{k=1}^\infty$ is a sequence satisfying $0<C_F\zeta_k\leqslant\delta_k$, and $\delta_k=O(z^{a^k})$, $z<z_I$, $1<a<a_I$ (a arbitrarily close to a_I);

• (λ_k,μ_k) are decreasing to $(\lambda_\infty,\mu_\infty)$, where $(\lambda_\infty,\mu_\infty)$ are arbitrarily close to (λ_1,μ_1); in particular we impose

$$\lambda^\sharp-\lambda_\infty\leqslant\min\{1,\tfrac{1}{2}\lambda_\infty\}\quad\text{and}\quad\mu^\sharp-\mu_\infty\leqslant\min\{1,\tfrac{1}{2}\mu_\infty\}; \tag{10.6}$$

• T is some small positive time in (10.3); we impose

$$\lambda^\#T\leqslant\tfrac{1}{2}(\mu^\sharp-\mu_1); \tag{10.7}$$

• $b=b(t)=\dfrac{B}{1+t}$, where $B\in(0,T)$ is a (small) constant.

10.2. Preparatory remarks

As announced in (10.5), we shall propagate the following "primary" controls on the density and distribution:

$$\sup_{\tau\geqslant 0} \|\varrho[h_\tau^k]\|_{\mathcal{F}^{\lambda_k\tau+\mu_k}} \leqslant \delta_k \quad \text{for all } k\in\{1,...,n\} \tag{$\mathbf{E}^{\mathbf{n}}_{\varrho}$}$$

and

$$\sup_{0\leqslant\tau\leqslant t} \|h_\tau^k\circ\Omega_{t,\tau}^{k-1}\|_{\mathcal{Z}^{\lambda_k(1+b),\mu_k;p}_{\tau-bt/(1+b)}} \leqslant \delta_k \quad \text{for all } k\in\{1,...,n\} \text{ and } p\in[1,\bar{p}]. \tag{$\mathbf{E}^{\mathbf{n}}_{\mathbf{h}}$}$$

Estimate ($\mathbf{E}^{\mathbf{n}}_{\varrho}$) obviously implies, via (10.2), up to a multiplicative constant,

$$\sup_{\tau\geqslant 0} \|F[h_\tau^k]\|_{\mathcal{F}^{\lambda_k\tau+\mu_k,\gamma}} \leqslant \delta_k \quad \text{for all } k\in\{1,...,n\}. \tag{$\widetilde{\mathbf{E}}^{\mathbf{n}}_{\varrho}$}$$

Before we can go from there to stage $n+1$, we need an additional set of estimates on the deflection maps $(\Omega^k)_{k=1}^n$, which will be used to

(1) update the control on $\Omega^k_{t,\tau}-\mathrm{Id}$;

(2) establish the needed control along the characteristics for the background

$$(\nabla_v f_\tau^n)\circ\Omega^n_{t,\tau}$$

(same index for the distribution and the deflection);

(3) update some technical controls allowing us to exchange (asymptotically) gradient and composition by $\Omega^k_{t,\tau}$; this will be crucial to handle the contribution of the zero mode of the background after composition by characteristics.

This set of deflection estimates falls into three categories. The first group expresses the closeness of Ω^k to Id:

$$\begin{cases} \sup\limits_{0\leqslant\tau\leqslant t} \|\Omega^k X_{t,\tau}-\mathrm{Id}\,\|_{\mathcal{Z}^{\lambda_k^*(1+b),(\mu_k^*,\gamma)}_{\tau-bt/(1+b)}} \leqslant 2\mathcal{R}_2^k(\tau,t), \\ \sup\limits_{0\leqslant\tau\leqslant t} \|\Omega^k V_{t,\tau}-\mathrm{Id}\,\|_{\mathcal{Z}^{\lambda_k^*(1+b),(\mu_k^*,\gamma)}_{\tau-bt/(1+b)}} \leqslant \mathcal{R}_1^k(\tau,t), \end{cases} \quad \text{for all } k\in\{1,...,n\}, \tag{$\mathbf{E}^{\mathbf{n}}_{\mathbf{\Omega}}$}$$

with $\lambda_k>\lambda_k^*>\lambda_{k+1}$, $\mu_k>\mu_k^*>\mu_{k+1}$ and

$$\begin{cases} \mathcal{R}_1^k(\tau,t) = \left(\sum\limits_{j=1}^k \dfrac{\delta_j e^{-2\pi(\lambda_j-\lambda_j^*)\tau}}{2\pi(\lambda_j-\lambda_j^*)}\right) \min\{t-\tau,1\} \\ \mathcal{R}_2^k(\tau,t) = \left(\sum\limits_{j=1}^k \dfrac{\delta_j e^{-2\pi(\lambda_j-\lambda_j^*)\tau}}{(2\pi(\lambda_j-\lambda_j^*))^2}\right) \min\left\{\dfrac{(t-\tau)^2}{2},1\right\}. \end{cases} \tag{10.8}$$

The second group of estimates expresses the fact that $\Omega^n - \Omega^k$ is very small when k is large:

$$\left\{\begin{array}{l} \sup\limits_{0\leqslant\tau\leqslant t} \|\Omega^n X_{t,\tau} - \Omega^k X_{t,\tau}\|_{\mathcal{Z}^{\lambda_n^*(1+b),(\mu_n^*,\gamma)}_{\tau-bt/(1+b)}} \leqslant 2\mathcal{R}_2^{k,n}(\tau,t), \\ \sup\limits_{0\leqslant\tau\leqslant t} \|\Omega^n V_{t,\tau} - \Omega^k V_{t,\tau}\|_{\mathcal{Z}^{\lambda_n^*(1+b),(\mu_n^*,\gamma)}_{\tau-bt/(1+b)}} \leqslant \mathcal{R}_1^{k,n}(\tau,t) + \mathcal{R}_2^{k,n}(\tau,t), \\ \sup\limits_{0\leqslant\tau\leqslant t} \|(\Omega_{t,\tau}^k)^{-1}\circ\Omega_{t,\tau}^n - \mathrm{Id}\,\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{\tau-bt/(1+b)}} \leqslant 4(\mathcal{R}_1^{k,n}(\tau,t) + \mathcal{R}_2^{k,n}(\tau,t)), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for all } k\in\{0,...,n-1\}, \end{array}\right. \tag{$\widetilde{\mathbf{E}}^{\mathbf{n}}_{\mathbf{\Omega}}$}$$

with

$$\left\{\begin{array}{l} \mathcal{R}_1^{k,n}(\tau,t) = \Big(\sum\limits_{j=k+1}^{n} \dfrac{\delta_j e^{-2\pi(\lambda_j-\lambda_j^*)\tau}}{2\pi(\lambda_j-\lambda_j^*)}\Big) \min\{t-\tau,1\} \\ \mathcal{R}_2^{k,n}(\tau,t) = \Big(\sum\limits_{j=k+1}^{n} \dfrac{\delta_j e^{-2\pi(\lambda_j-\lambda_j^*)\tau}}{(2\pi(\lambda_j-\lambda_j^*))^2}\Big) \min\Big\{\dfrac{(t-\tau)^2}{2},1\Big\}. \end{array}\right. \tag{10.9}$$

(Choosing $k=0$ brings us back to the previous estimates ($\mathbf{E}^{\mathbf{n}}_{\mathbf{\Omega}}$).)

The last group of estimates expresses the fact that the differential of the deflection map is uniformly close to the identity (in a way which is more precise than what would follow from the first group of estimates):

$$\left\{\begin{array}{l} \sup\limits_{0\leqslant\tau\leqslant t} \|\nabla\Omega^k X_{t,\tau} - (I,0)\|_{\mathcal{Z}^{\lambda_k^*(1+b),\mu_k^*}_{\tau-bt/(1+b)}} \leqslant 2\mathcal{R}_2^{k}(\tau,t), \\ \sup\limits_{0\leqslant\tau\leqslant t} \|\nabla\Omega^k V_{t,\tau} - (0,I)\|_{\mathcal{Z}^{\lambda_k^*(1+b),\mu_k^*}_{\tau-bt/(1+b)}} \leqslant \mathcal{R}_1^{k}(\tau,t) + \mathcal{R}_2^{k}(\tau,t), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for all } k\in\{1,...,n\}, \end{array}\right. \tag{$\mathbf{E}^{\mathbf{n}}_{\nabla\mathbf{\Omega}}$}$$

where $\nabla = (\nabla_x, \nabla_v)$, and I is the identity matrix.

An important property of the functions $\mathcal{R}_1^{k,n}(\tau,t)$ and $\mathcal{R}_2^{k,n}(\tau,t)$ is their fast decay as $\tau\to\infty$ and as $k\to\infty$, uniformly in $n\geqslant k$; this is due to the fast convergence of the sequence $(\delta_k)_{k=1}^{\infty}$. Eventually, if $r\in\mathbb{N}$ is given, we shall have

$$\mathcal{R}_1^{k,n}(\tau,t) \leqslant \omega_{k,n}^{r,1}(\tau,t) \text{ and } \mathcal{R}_2^{k,n}(\tau,t) \leqslant \omega_{k,n}^{r,2}(\tau,t) \quad \text{for all } r\geqslant 1, \tag{10.10}$$

with

$$\omega_{k,n}^{r,1}(\tau,t) := C_\omega^r \Big(\sum_{j=k+1}^{n} \frac{\delta_j}{(2\pi(\lambda_j-\lambda_j^*))^{1+r}}\Big) \frac{\min\{t-\tau,1\}}{(1+\tau)^r}$$

and

$$\omega_{k,n}^{r,2}(\tau,t) := C_\omega^r \Big(\sum_{j=k+1}^{n} \frac{\delta_j}{(2\pi(\lambda_j-\lambda_j^*))^{2+r}}\Big) \frac{\min\big\{\frac{1}{2}(t-\tau)^2,1\big\}}{(1+\tau)^r}$$

for some absolute constant C_ω^r depending only on r (we also denote $\omega_{0,n}^{r,1} = \omega_n^{r,1}$ and $\omega_{0,n}^{r,2} = \omega_n^{r,2}$).

From the estimates on the characteristics and $(\mathbf{E^n_h})$ will follow the following "secondary controls" on the distribution function:

$$\left\{\begin{array}{l}
\sup_{0\leqslant\tau\leqslant t} \|(\nabla_x h^k_\tau)\circ\Omega^{k-1}_{t,\tau}\|_{\mathcal{Z}^{\lambda_k(1+b),\mu_k;p}_{\tau-bt/(1+b)}} \leqslant \delta_k,\\
\sup_{0\leqslant\tau\leqslant t} \|\nabla_x(h^k_\tau\circ\Omega^{k-1}_{t,\tau})\|_{\mathcal{Z}^{\lambda_k(1+b),\mu_k;p}_{\tau-bt/(1+b)}} \leqslant \delta_k,\\
\|((\nabla_v+\tau\nabla_x)h^k_\tau)\circ\Omega^{k-1}_{t,\tau}\|_{\mathcal{Z}^{\lambda_k(1+b),\mu_k;p}_{\tau-bt/(1+b)}} \leqslant \delta_k,\\
\|(\nabla_v+\tau\nabla_x)(h^k_\tau\circ\Omega^{k-1}_{t,\tau})\|_{\mathcal{Z}^{\lambda_k(1+b),\mu_k;p}_{\tau-bt/(1+b)}} \leqslant \delta_k,\\
\sup_{0\leqslant\tau\leqslant t} \frac{1}{(1+\tau)^2}\|(\nabla\nabla h^k_\tau)\circ\Omega^{k-1}_{t,\tau}\|_{\mathcal{Z}^{\lambda_k(1+b),\mu_k;1}_{\tau-bt/(1+b)}} \leqslant \delta_k,\\
\sup_{0\leqslant\tau\leqslant t} (1+\tau)^2\|(\nabla h^k_\tau)\circ\Omega^{k-1}_{t,\tau} - \nabla(h^k_\tau\circ\Omega^{k-1}_{t,\tau})\|_{\mathcal{Z}^{\lambda_k(1+b),\mu_k;1}_{\tau-bt/(1+b)}} \leqslant \delta_k.
\end{array}\right. \qquad (\widetilde{\mathbf{E}}^{\mathbf{n}}_{\mathbf{h}})$$

for all $k\in\{1,...,n\}$ and $p\in[1,\bar p]$.

The transition from stage n to stage $n+1$ can be summarized as follows:

$$(\widetilde{\mathbf{E}}^{\mathbf{n}}_{\varrho}) \overset{(\mathbf{A_n})}{\Longrightarrow} (\mathbf{E}^{\mathbf{n}}_{\mathbf{\Omega}})+(\widetilde{\mathbf{E}}^{\mathbf{n}}_{\mathbf{\Omega}})+(\mathbf{E}^{\mathbf{n}}_{\nabla\mathbf{\Omega}}),$$

$$(\mathbf{E}^{\mathbf{n}}_{\varrho})+(\mathbf{E}^{\mathbf{n}}_{\mathbf{\Omega}})+(\widetilde{\mathbf{E}}^{\mathbf{n}}_{\mathbf{\Omega}})+(\mathbf{E}^{\mathbf{n}}_{\nabla\mathbf{\Omega}})+(\mathbf{E}^{\mathbf{n}}_{\mathbf{h}})+(\widetilde{\mathbf{E}}^{\mathbf{n}}_{\mathbf{h}}) \overset{(\mathbf{B_n})}{\Longrightarrow} (\mathbf{E}^{\mathbf{n+1}}_{\varrho})+(\widetilde{\mathbf{E}}^{\mathbf{n+1}}_{\varrho})+(\mathbf{E}^{\mathbf{n+1}}_{\mathbf{h}})+(\widetilde{\mathbf{E}}^{\mathbf{n+1}}_{\mathbf{h}}).$$

The first implication $(\mathbf{A_n})$ is proven by an amplification of the technique used in §5; ultimately, it relies on repeated application of Picard's fixed-point theorem in analytic norms. The second implication $(\mathbf{B_n})$ is the harder part; it uses the machinery from §6 and §7, together with the idea of simultaneously propagating a *shifted* $\mathcal{Z}$ norm for the kinetic distribution and an $\mathcal{F}$ norm for the density.

In both implications, the stratification of error estimates will prevent the blow up of constants. So we shall decompose the force field F^n generated by f^n as

$$F^n = F[f^n] = E^1+...+E^n,$$

where $E^k = F[h^k] = -\nabla W * \varrho[h^k]$.

The plan of the estimates is as follows. We shall inductively construct a sequence of constant coefficients

$$\lambda^\sharp > \lambda_1 > \lambda_1^* > \lambda_2 > ... > \lambda_n > \lambda_n^* > \lambda_{n+1} > ...,$$
$$\mu^\sharp > \mu_1 > \mu_1^* > \mu_2 > ... > \mu_n > \mu_n^* > \mu_{n+1} > ...$$

(where λ_n and μ_n will be fixed in the proof of $(\mathbf{A_n})$, and λ_{n+1} and μ_{n+1} in the proof of $(\mathbf{B_n})$) converging respectively to λ_∞ and μ_∞; and a sequence $(\delta_k)_{k=1}^\infty$ decreasing very fast to zero. For simplicity we shall let

$$\mathcal{R}^n(\tau,t) = \mathcal{R}^n_1(\tau,t)+\mathcal{R}^n_2(\tau,t) \quad \text{and} \quad \mathcal{R}^{k,n}(\tau,t) = \mathcal{R}^{k,n}_1(\tau,t)+\mathcal{R}^{k,n}_2(\tau,t),$$

and assume that $2\pi(\lambda_j-\lambda_j^*)\leqslant 1$; so

$$\mathcal{R}^{k,n}(\tau,t)\leqslant C_\omega^r\left(\sum_{j=k+1}^{n}\frac{\delta_j}{(2\pi(\lambda_j-\lambda_j^*))^{2+r}}\right)\frac{\min\{t-\tau,1\}}{(1+\tau)^r}\quad\text{and}\quad\mathcal{R}^{0,n}=\mathcal{R}^n. \tag{10.11}$$

It will be sufficient to work with some fixed r, large enough (as we shall see, $r=4$ will do).

To go from stage n to stage $n+1$, we shall do as follows:

Implication $(\mathbf{A_n})$ (§10.3):

Step 1. Estimate $\Omega^n-\mathrm{Id}$ (the bound should be uniform in n).

Step 2. Estimate $\Omega^n-\Omega^k$ ($k\leqslant n-1$; the error should be small when $k\to\infty$).

Step 3. Estimate $\nabla\Omega^n-I$.

Step 4. Estimate $(\Omega^k)^{-1}\circ\Omega^n$.

Implication $(\mathbf{B_n})$ (§10.4):

Step 5. Estimate h^k and its derivatives along the composition by Ω^n.

Step 6. Estimate $\varrho[h^{n+1}]$, using §6 and §7.

Step 7. Estimate $F[h^{n+1}]$ from $\varrho[h^{n+1}]$.

Step 8. Estimate $h^{n+1}\circ\Omega^n$.

Step 9. Estimate derivatives of h^{n+1} composed with Ω^n.

Step 10. Show that for h^{n+1}, ∇ and composition by Ω^n asymptotically commute.

10.3. Estimates on the characteristics

In this subsection, we assume that $(\mathbf{E}_\varrho^{\mathbf{n}})$ is proven, and establish $(\mathbf{E}_{\mathbf{\Omega}}^{\mathbf{n}})+(\widetilde{\mathbf{E}}_{\mathbf{\Omega}}^{\mathbf{n}})+(\mathbf{E}_{\mathbf{\nabla\Omega}}^{\mathbf{n}})$. Let $\lambda_n^*<\lambda_n$ and $\mu_n^*<\mu_n$ to be fixed later on.

10.3.1. Step 1. Estimate of $\Omega^n-\mathrm{Id}$

This is the first and archetypal estimate. We shall bound $\Omega^n X_{t,\tau}-x$ in the hybrid norm $\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{\tau-bt/(1+b)}$. The Sobolev correction γ will play no role here in the proofs, and for simplicity we shall forget it in the computations, just recall it in the final results. (Use Proposition 4.32 whenever needed.)

Since we expect the characteristics for the force field F^n to be close to the free transport characteristics, it is natural to write

$$X^n_{t,\tau}(x,v)=x-v(t-\tau)+Z^n_{t,\tau}(x,v), \tag{10.12}$$

where $Z^n_{t,\tau}$ solves

$$\begin{cases}\dfrac{\partial^2}{\partial\tau^2}Z^n_{t,\tau}(x,v)=F^n(\tau,x-v(t-\tau)+Z^n_{t,\tau}(x,v))\\ Z^n_{t,t}(x,v)=0,\quad \partial_\tau Z^n_{t,\tau}|_{t=\tau}(x,v)=0.\end{cases} \tag{10.13}$$

(With respect to §5 we have dropped the parameter ε, to take advantage of the "stratified" nature of F^n; anyway this parameter was cosmetic.) So if we fix $t>0$, $(Z^n_{t,\tau})_{0\leqslant\tau\leqslant t}$ is a fixed point of the map

$$\Psi\colon (W_{t,\tau})_{0\leqslant\tau\leqslant t} \longmapsto (Z_{t,\tau})_{0\leqslant\tau\leqslant t}$$

defined by

$$\begin{cases} \dfrac{\partial^2}{\partial\tau^2} Z_{t,\tau} = F^n(\tau, x-v(t-\tau)+W_{t,\tau}) \\ Z_{t,t}=0, \quad \partial_\tau Z_{t,\tau}|_{\tau=t}=0. \end{cases} \tag{10.14}$$

The goal is to estimate $Z^n_{t,\tau}-x$ in the hybrid norm $\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{t-bt/(1+b)}$.

We first bound $(Z^n_0)_{t,\tau}=\Psi(0)$. Explicitly,

$$(Z^n_0)_{t,\tau}(x,v) = \int_\tau^t (s-\tau)\, F^n(s, x-v(t-s))\, ds.$$

By Propositions 4.15 (i) and 4.19,

$$\begin{aligned} \|(Z^n_0)_{t,\tau}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{t-bt/(1+b)}} &\leqslant \int_\tau^t (s-\tau) \|F^n(s, x-v(t-s))\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{t-bt/(1+b)}}\, ds \\ &= \int_\tau^t (s-\tau) \|F^n(s,\cdot)\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{s-bt/(1+b)}}\, ds \\ &= \int_\tau^t (s-\tau) \|F^n(s,\cdot)\|_{\mathcal{F}^{\nu(s,t)}}\, ds, \end{aligned} \tag{10.15}$$

where

$$\nu(s,t) = \lambda^*_n |s-b(t-s)| + \mu^*_n. \tag{10.16}$$

Case 1. If $s\geqslant bt/(1+b)$, then

$$\nu(s,t) \leqslant \lambda^*_n s + \mu^*_n \leqslant \lambda_k\, s + \mu_k - (\lambda_k - \lambda^*_n)s, \quad k\in\{1,\dots,n\}. \tag{10.17}$$

Case 2. If $s<bt/(1+b)$, then necessarily $s\leqslant B\leqslant T$. Taking into account (10.6), we have

$$\nu(s,t) = \lambda^*_n\, bt + \mu^*_n - \lambda^*_n(1+b)s \tag{10.18}$$

$$\leqslant \lambda^*_n B + \mu^*_n - (\lambda_k - \lambda^*_n)s. \tag{10.19}$$

(Of course, the assumption $\lambda^\sharp - \lambda_\infty \leqslant \min\{1, \frac12\lambda_\infty\}$ implies that $\lambda_k - \lambda^*_n \leqslant \lambda^*_n$.) In particular, by (10.7),

$$\nu(s,t) \leqslant \mu^\sharp - (\lambda_k - \lambda^*_n)s, \quad k\in\{1,\dots,n\}. \tag{10.20}$$

We plug these bounds into (10.15), then use $\widehat{E}^k(s,0)=0$ and the bounds $(\widetilde{\mathbf{E}}^{\mathbf{n}}_{\varrho})$ and (10.17) (for large times), and (10.3) and (10.20) (for short times). This yields

$$\begin{aligned}
&\|(Z_0^n)_{t,\tau}\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{t-bt/(1+b)}}\\
&\quad\leqslant\sum_{k=1}^n\biggl(\int_{\tau\vee bt/(1+b)}^t(s-\tau)\|E^k(s,\cdot)\|_{\mathcal{F}^{\lambda_k s+\mu_k-(\lambda_k-\lambda_n^*)s}}\,ds\\
&\qquad\qquad+\int_\tau^{\tau\vee bt/(1+b)}(s-\tau)\|E^k(s,\cdot)\|_{\mathcal{F}^{\mu^\sharp-(\lambda_k-\lambda_n^*)s}}\,ds\biggr)\\
&\quad\leqslant\sum_{k=1}^n\biggl(\int_{\tau\vee bt/(1+b)}^t(s-\tau)e^{-2\pi(\lambda_k-\lambda_n^*)s}\|E^k(s,\cdot)\|_{\mathcal{F}^{\lambda_k s+\mu_k}}\,ds\\
&\qquad\qquad+\int_\tau^{\tau\vee bt/(1+b)}(s-\tau)e^{-2\pi(\lambda_k-\lambda_n^*)s}\|E^k(s,\cdot)\|_{\mathcal{F}^{\mu^\sharp}}\,ds\biggr)\\
&\quad\leqslant\sum_{k=1}^n\delta_k\int_\tau^t(s-\tau)e^{-2\pi(\lambda_k-\lambda_n^*)s}\,ds\\
&\quad\leqslant\sum_{k=1}^n\delta_k e^{-2\pi(\lambda_k-\lambda_n^*)\tau}\min\biggl\{\frac{(t-\tau)^2}{2},\frac{1}{(2\pi(\lambda_k-\lambda_n^*))^2}\biggr\}\\
&\quad\leqslant\mathcal{R}_2^n(\tau,t).
\end{aligned}\tag{10.21}$$

Let us define the norm

$$|\!|\!|(Z_{t,\tau})_{0\leqslant\tau\leqslant t}|\!|\!|_n:=\sup_{0\leqslant\tau\leqslant t}\frac{\|Z_{t,\tau}\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{t-bt/(1+b)}}}{\mathcal{R}_2^n(\tau,t)}.$$

(Note the difference with §5: now the regularity exponents depend on time(s).) Inequality (10.21) means that $|\!|\!|\Psi(0)|\!|\!|_n\leqslant 1$. We shall check that Ψ is $\frac{1}{2}$-Lipschitz on the ball $B(0,2)$ in the norm $|\!|\!|\cdot|\!|\!|_n$. This will be subtle: the uniform bounds on the size of the force field, coming from the preceding steps, will allow us to get good decaying exponentials, which in turn will imply uniform error bounds at the present stage.

So let $W,\widetilde{W}\in B(0,2)$, and let $Z=\Psi(W)$ and $\widetilde{Z}=\Psi(\widetilde{W})$. As in §5, we write

$$Z_{t,\tau}-\widetilde{Z}_{t,\tau}=\int_0^1\int_\tau^t(s-\tau)\nabla_x F^n(s,x-v(t-s)+\theta W_{t,s}+(1-\theta)\widetilde{W}_{t,s})\cdot(W_{t,s}-\widetilde{W}_{t,s})\,ds\,d\theta,$$

and deduce that

$$|\!|\!|(Z_{t,\tau}-\widetilde{Z}_{t,\tau})_{0\leqslant\tau\leqslant t}|\!|\!|_n\leqslant A(t)|\!|\!|(W_{t,s}-\widetilde{W}_{t,s})_{0\leqslant s\leqslant t}|\!|\!|_n,$$

where

$$\begin{aligned}
A(t)=&\sup_{0\leqslant\tau\leqslant s\leqslant t}\frac{\mathcal{R}_2^n(s,t)}{\mathcal{R}_2^n(\tau,t)}\\
&\times\int_0^1\int_\tau^t(s-\tau)\|\nabla_x F^n(s,x-v(t-s)+\theta W_{t,s}+(1-\theta)\widetilde{W}_{t,s})\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{t-bt/(1+b)}}\,ds\,d\theta.
\end{aligned}$$

For $\tau \leqslant s$ we have $\mathcal{R}_2^n(s,t) \leqslant \mathcal{R}_2^n(\tau,t)$. Also, by Proposition 4.25 (applied with $V=0$, $b=-(t-s)$ and $\sigma=0$ in that statement) and Proposition 4.15,

$$A(t) \leqslant \sup_{0\leqslant\tau\leqslant t} \int_\tau^t (s-\tau) \|\nabla_x F^n(s,\cdot)\|_{\mathcal{F}^{\nu(s,t)+e(s,t)}}\, ds,$$

where ν is defined by (10.16) and the "error" $e(s,t)$ arising from composition is given by

$$e(s,t) = \sup_{0\leqslant\theta\leqslant 1} \|\theta W_{t,s} + (1-\theta)\widetilde{W}_{t,s}\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{t-bt/(1+b)}} \leqslant 2\mathcal{R}_2^n(s,t).$$

Since

$$\mathcal{R}_2^n(s,t) \leqslant \omega_n^{1,2}(s,t) := C_\omega^1 \left(\sum_{k=1}^n \frac{\delta_k}{(2\pi(\lambda_k - \lambda_k^*))^3} \right) \frac{\min\left\{\frac12 (t-s)^2, 1\right\}}{1+s},$$

we have, for all $0 \leqslant s \leqslant t$,

$$2\mathcal{R}_2^n(s,t) \leqslant \frac{\lambda_n^* b(t-s)}{2}\, 1_{s\geqslant bt/(1+b)} + \frac{\mu^\sharp - \mu_n^*}{2}\, 1_{s\leqslant bt/(1+b)}, \tag{10.22}$$

as soon as

$$2C_\omega^1 \left(\sum_{k=1}^n \frac{\delta_k}{(2\pi(\lambda_k - \lambda_n^*))^3} \right) \leqslant \min\left\{ \frac{\lambda_n^* B}{6}, \frac{\mu^\sharp - \mu_n^*}{2} \right\} \quad \text{for all } n \geqslant 1. \tag{$\mathbf{C_1}$}$$

We shall check later in §10.5 the feasibility of condition ($\mathbf{C_1}$)—as well as a number of other forthcoming ones.

The extra error term in the exponent is sufficiently small to be absorbed by what we throw away in (10.17) or in (10.18)–(10.20). So we obtain, as in the estimate of Z_0^n, for any $k \in \{1, ..., n\}$,

$$(\nu+e)(s,t) \leqslant \begin{cases} \lambda_k s + \mu_k - (\lambda_k - \lambda_n^*)s & \text{for } s \geqslant bt/(1+b), \\ \mu^\sharp - (\lambda_k - \lambda_n^*)s & \text{for } s \leqslant bt/(1+b), \end{cases}$$

and we deduce (using $(\widetilde{\mathbf{E}}_\varrho^{\mathbf{n}})$ and $\gamma \geqslant 1$) that

$$\begin{aligned} A(t) &\leqslant \sup_{0\leqslant\tau\leqslant t} \sum_{k=1}^n \left(\int_{\tau\vee bt/(1+b)}^t (s-\tau) \|\nabla_x E^k(s,\cdot)\|_{\mathcal{F}^{\lambda_k s + \mu_k - (\lambda_k - \lambda_n^*)\, s}}\, ds \right. \\ &\qquad \left. + \int_\tau^{\tau\vee bt/(1+b)} (s-\tau) \|\nabla_x E^k(s,\cdot)\|_{\mathcal{F}^{\mu^\sharp - (\lambda_k - \lambda_n^*)s}}\, ds \right) \\ &\leqslant \sup_{0\leqslant\tau\leqslant t} \sum_{k=1}^n \delta_k \int_\tau^t (s-\tau) e^{-(\lambda_k - \lambda_n^*)s}\, ds \\ &\leqslant \sup_{0\leqslant\tau\leqslant t} \mathcal{R}_2^n(\tau,t) = \mathcal{R}_2^n(0,t) \\ &\leqslant \sum_{k=1}^n \frac{\delta_k}{(2\pi(\lambda_k - \lambda_n^*))^2}. \end{aligned}$$

If the latter quantity is bounded above by $\frac{1}{2}$, then Ψ is $\frac{1}{2}$-Lipschitz and we may apply the fixed-point result from Theorem A.2. Therefore, under the condition (whose feasibility will be checked later)

$$\sum_{k=1}^{n} \frac{\delta_k}{(2\pi(\lambda_k-\lambda_n^*))^2} \leqslant \frac{1}{2} \quad \text{for all } n \geqslant 1, \tag{$\mathbf{C_2}$}$$

we deduce that

$$\|Z_{t,\tau}^n\|_{\mathcal{Z}_{t-bt/(1+b)}^{\lambda_n^*(1+b),\mu_n^*}} \leqslant 2\mathcal{R}_2^n(\tau,t).$$

After that, the estimates on the deflection map are obtained exactly as in §5: writing $\Omega_{t,\tau}^n=(\Omega^n X_{t,\tau}, \Omega^n V_{t,\tau})$, recalling the dependence on γ again, we end up with

$$\begin{cases} \|\Omega^n X_{t,\tau}-x\|_{\mathcal{Z}_{\tau-bt/(1+b)}^{\lambda_n^*(1+b),(\mu_n^*,\gamma)}} \leqslant 2\mathcal{R}_2^n(\tau,t), \\ \|\Omega^n V_{t,\tau}-v\|_{\mathcal{Z}_{\tau-bt/(1+b)}^{\lambda_n^*(1+b),(\mu_n^*,\gamma)}} \leqslant \mathcal{R}_1^n(\tau,t). \end{cases} \tag{10.23}$$

10.3.2. Step 2. Estimate of $\Omega^n-\Omega^k$

In this step our goal is to estimate $\Omega^n-\Omega^k$ for $1\leqslant k\leqslant n-1$. The point is that the error should be small as $k\to\infty$, uniformly in n, so we cannot just write

$$\|\Omega^n-\Omega^k\| \leqslant \|\Omega^n-\mathrm{Id}\,\|+\|\Omega^k-\mathrm{Id}\,\|.$$

Instead, we start again from the differential equation satisfied by Z^k and Z^n:

$$\begin{aligned} \frac{\partial^2}{\partial\tau^2}(Z_{t,\tau}^n-Z_{t,\tau}^k)(x,v) &= F^n(\tau,x-v(t-\tau)+Z_{t,\tau}^n(x,v))-F^k(\tau,x-v(t-\tau)+Z_{t,\tau}^k(x,v)) \\ &= F^n(\tau,x-v(t-\tau)+Z_{t,\tau}^n)-F^n(\tau,x-v(t-\tau)+Z_{t,\tau}^k) \\ &\quad +(F^n-F^k)(\tau,x-v(t-\tau)+Z_{t,\tau}^k). \end{aligned}$$

This, together with the boundary conditions $Z_{t,t}^n-Z_{t,t}^k=0$ and $\partial_\tau(Z_{t,\tau}^n-Z_{t,\tau}^k)|_{\tau=t}=0$, implies that

$$\begin{aligned} Z_{t,\tau}^n-Z_{t,\tau}^k &= \int_0^1\int_\tau^t (s-\tau)\nabla_x F^n(s,x-v(t-s)+\theta Z_{t,s}^k+(1-\theta)Z_{t,s}^n)\cdot(Z_{t,s}^n-Z_{t,s}^k)\,ds\,d\theta \\ &\quad +\int_\tau^t (s-\tau)(F^n-F^k)(s,x-v(t-s)+Z_{t,s}^k(x,v))\,ds. \end{aligned}$$

We fix t and define the norm

$$|||(Z_{t,\tau})_{0\leqslant\tau\leqslant t}|||_{k,n} := \sup_{0\leqslant\tau\leqslant t} \frac{\|Z_{t,\tau}\|_{\mathcal{Z}_{t-bt/(1+b)}^{\lambda_n^*(1+b),\mu_n^*}}}{\mathcal{R}_2^{k,n}(\tau,t)},$$

where $\mathcal{R}_2^{k,n}$ is defined in (10.9). Using the bounds on Z^n and Z^k in $|\!|\!| \cdot |\!|\!|_n$ (since $|\!|\!| \cdot |\!|\!|_n \leqslant |\!|\!| \cdot |\!|\!|_k$ by using the fact that $\mathcal{R}_2^k \leqslant \mathcal{R}_2^n$) and proceeding as before, we get that

$$\begin{aligned} &|\!|\!|(Z^n_{t,\tau}-Z^k_{t,\tau})_{0\leqslant\tau\leqslant t}|\!|\!|_{k,n} \\ &\quad\leqslant \tfrac{1}{2}|\!|\!|(Z^n_{t,\tau}-Z^k_{t,\tau})_{0\leqslant\tau\leqslant t}|\!|\!|_{k,n} \\ &\qquad + \left|\!\left|\!\left| \left(\int_\tau^t (s-\tau)(F^n-F^k)(s,x-v(t-s)+Z^k_{t,s})\,ds \right)_{0\leqslant\tau\leqslant t} \right|\!\right|\!\right|_{k,n}. \end{aligned} \tag{10.24}$$

Next we estimate

$$\begin{aligned} \|(F^n-F^k)(s,x-v(t-s)+Z^k_{t,s})\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{t-bt/(1+b)}} &= \|(F^n-F^k)(s,X^k_{t,s})\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{t-bt/(1+b)}} \\ &= \|(F^n-F^k)(s,\Omega^k_{t,s})\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{s-bt/(1+b)}} \\ &\leqslant \|(F^n-F^k)(s,\cdot)\|_{\mathcal{F}^{\nu(s,t)+e(s,t)}}, \end{aligned}$$

where the last inequality follows from Proposition 4.25, ν is again given by (10.16), and

$$e(s,t) = \|\Omega^k X_{t,s} - \mathrm{Id}\,\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{s-bt/(1+b)}} \leqslant 2\mathcal{R}_2^k(s,t) \leqslant 2\mathcal{R}_2^n(s,t).$$

The same reasoning as in Step 1 yields, under assumptions ($\mathbf{C_1}$) and ($\mathbf{C_2}$), for $k+1 \leqslant j \leqslant n$,

$$(\nu+e)(s,t) \leqslant \begin{cases} \lambda_j s + \mu_j - (\lambda_j - \lambda_n^*)s & \text{for } s \geqslant bt/(1+b), \\ \mu^\sharp - (\lambda_j-\lambda_n^*)s & \text{for } s \leqslant bt/(1+b), \end{cases}$$

and so

$$\|F^n_s - F^k_s\|_{\mathcal{F}^{\nu+e}} \leqslant \sum_{j=k+1}^n \delta_j e^{-2\pi(\lambda_j-\lambda_n^*)s}.$$

For any $\tau \geqslant 0$, by integrating in time we find that

$$\begin{aligned} &\left\| \int_\tau^t (s-\tau)(F^n-F^k)(s,x-v(t-s)+Z^k_{t,s})\,ds \right\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{t-bt/(1+b)}} \\ &\qquad\qquad \leqslant \int_\tau^t (s-\tau) \sum_{j=k+1}^n \delta_j e^{-2\pi(\lambda_j-\lambda_n^*)s}\,ds \leqslant \mathcal{R}_2^{k,n}(\tau,t). \end{aligned}$$

Therefore

$$\left|\!\left|\!\left| \left(\int_\tau^t (s-\tau)(F^n-F^k)(s,x-v(t-s)+Z^k_{t,s})\,ds \right)_{0\leqslant\tau\leqslant t} \right|\!\right|\!\right|_{k,n} \leqslant 1$$

and, by(10.24),

$$|\!|\!|(Z^n_{t,\tau}-Z^k_{t,\tau})_{0\leqslant\tau\leqslant t}|\!|\!|_{k,n} \leqslant 2.$$

Recalling the Sobolev correction, we conclude that

$$\|\Omega^n X_{t,\tau} - \Omega^k X_{t,\tau}\|_{\mathcal{Z}^{\lambda_n^*(1+b),(\mu_n^*,\gamma)}_{\tau-bt/(1+b)}} \leqslant 2\mathcal{R}_2^{k,n}(\tau,t). \tag{10.25}$$

For the velocity component, say U, we write

$$\begin{aligned}\frac{\partial}{\partial\tau}(U^n_{t,\tau}-U^k_{t,\tau})(x,v) &= F^n(\tau,x-v(t-\tau)+Z^n_{t,\tau}(x,v))-F^k(\tau,x-v(t-\tau)+Z^k_{t,\tau}(x,v))\\ &= F^n(\tau,x-v(t-\tau)+Z^n_{t,\tau})-F^n(\tau,x-v(t-\tau)+Z^k_{t,\tau})\\ &\quad +(F^n-F^k)(\tau,x-v(t-\tau)+Z^k_{t,\tau}),\end{aligned}$$

where Z^n and Z^k were estimated above, and the boundary conditions are $U^n_{t,t}-U^k_{t,t}=0$. Thus

$$\begin{aligned}U^n_{t,\tau}-U^k_{t,\tau} &= \int_0^1\int_\tau^t \nabla_x F^n(s,x-v(t-s)+\theta Z^k_{t,s}+(1-\theta)Z^n_{t,s})\cdot(Z^n_{t,s}-Z^k_{t,s})\,ds\,d\theta\\ &\quad +\int_\tau^t (F^n-F^k)(s,x-v(t-s)+Z^k_{t,s}(x,v))\,ds,\end{aligned}$$

and from this one easily derives the similar estimates

$$\begin{cases}\|\Omega^n X_{t,\tau}-\Omega^k X_{t,\tau}\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{\tau-bt/(1+b)}} \leqslant 2\mathcal{R}_2^{k,n}(t,\tau),\\ \|\Omega^n V_{t,\tau}-\Omega^k V_{t,\tau}\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{\tau-bt/(1+b)}} \leqslant \mathcal{R}_1^{k,n}(t,\tau)+\mathcal{R}_2^{k,n}(t,\tau).\end{cases}$$

10.3.3. Step 3. Estimate of $\nabla\Omega^n$

We now establish a control on the derivative of the deflection map. Of course, we could deduce such a control from the bound on $\Omega^n-\mathrm{Id}$ and Proposition 4.32 (vi): for instance, if $\lambda_n^{**}<\lambda_n^*$ and $\mu_n^{**}<\mu_n^*$, then

$$\|\nabla\Omega^n_{t,\tau}-I\|_{\mathcal{Z}^{\lambda_n^{**}(1+b),(\mu_n^{**},\gamma)}_{\tau-bt/(1+b)}} \leqslant \frac{C\mathcal{R}_2^n(\tau,t)}{\min\{\lambda_n^*-\lambda_n^{**},\mu_n^*-\mu_n^{**}\}}. \tag{10.26}$$

But this bound involves very large constants, and is useless in our argument. Better estimates can be obtained by using again equation (10.13). Writing

$$(\Omega^n_{t,\tau}-\mathrm{Id})(x,v) = (Z^n_{t,\tau}(x+v(t-\tau),v),\dot{Z}^n_{t,\tau}(x+v(t-\tau),v)),$$

where the dot stands for $\partial/\partial\tau$, we get by differentiation

$$\begin{aligned}\nabla_x\Omega^n_{t,\tau}-(I,0) &= \big(\nabla_x Z^n_{t,\tau}(x+v(t-\tau),v\big),\nabla_x\dot{Z}^n_{t,\tau}(x+v(t-\tau),v)),\\ \nabla_v\Omega^n_{t,\tau}-(0,I) &= ((\nabla_v+(t-\tau)\nabla_x)Z^n_{t,\tau}(x+v(t-\tau),v),\\ &\qquad (\nabla_v+(t-\tau)\nabla_x)\dot{Z}^n_{t,\tau}(x+v(t-\tau),v)).\end{aligned}$$

Let us estimate for instance $\nabla_x\Omega - (I,0)$, or equivalently $\nabla_x Z^n_{t,\tau}$. By differentiating (10.13), we obtain

$$\frac{\partial^2}{\partial\tau^2}\nabla_x Z^n_{t,\tau}(x,v) = \nabla_x F^n(\tau, x-v(t-\tau)+Z^n_{t,\tau}(x,v))\cdot(\mathrm{Id}+\nabla_x Z^n_{t,\tau}).$$

So $\nabla_x Z^n_{t,\tau}$ is a fixed point of $\Psi: W\mapsto Q$, where W and Q are functions of $\tau\in[0,t]$ satisfying

$$\begin{cases} \dfrac{\partial^2 Q}{\partial\tau^2} = \nabla_x F^n(\tau, x-v(t-\tau)+Z^n_{t,\tau})(I+W), \\ Q(t)=0, \quad \partial_\tau Q(t)=0. \end{cases}$$

We treat this in the same way as in Steps 1 and 2, and find for Q_x (the x component of Q) the same estimates as we had previously on the x component of Ω. For the velocity component, a direct estimate from the integral equation expressing the velocity in terms of F yields a control by $\mathcal{R}^n_1+\mathcal{R}^n_2$. Finally for $\nabla_v\Omega$ this is similar, noting that $(\nabla_v+(t-\tau)\nabla_x)(x-v(t-\tau))=0$, the differential equation being for instance:

$$\frac{\partial^2}{\partial\tau^2}(\nabla_v+(t-\tau)\nabla_x)Z^n_{t,\tau}(x,v) = \nabla_x F^n(\tau, x-v(t-\tau)+Z^n_{t,\tau}(x,v))\cdot((\nabla_v+(t-\tau)\nabla_x)Z^n_{t,\tau}).$$

In the end we obtain

$$\begin{cases} \displaystyle\sup_{0\leqslant\tau\leqslant t} \|\nabla\Omega^n X_{t,\tau}-(I,0)\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}} \leqslant 2\mathcal{R}^n_2(\tau,t), \\ \displaystyle\sup_{0\leqslant\tau\leqslant t} \|\nabla\Omega^n V_{t,\tau}-(0,I)\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}} \leqslant \mathcal{R}^n_1(\tau,t)+\mathcal{R}^n_2(\tau,t). \end{cases} \tag{10.27}$$

10.3.4. Step 4. Estimate of $(\Omega^k)^{-1}\circ\Omega^n$

We do this by applying Proposition 4.28 with $F=\Omega^k$ and $G=\Omega^n$. (Note that we cannot exchange the roles of Ω^k and Ω^n in this step, because we have better information on the regularity of Ω^k.) Let $\varepsilon=\varepsilon(d)$ be the small constant appearing in Proposition 4.28. If

$$3\mathcal{R}^k_2(\tau,t)+\mathcal{R}^k_1(\tau,t)\leqslant\varepsilon \quad \text{for all } k\geqslant 1, \tag{$\mathbf{C_3}$}$$

then $\|\nabla\Omega^k_{t,\tau}-I\|_{\mathcal{Z}^{\lambda^*_k(1+b),\mu^*_k}_{\tau-bt/(1+b)}}\leqslant\varepsilon$; if in addition

$$\begin{aligned} 2(1+\tau)(1+B)(3\mathcal{R}^{k,n}_2+\mathcal{R}^{k,n}_1)(\tau,t) &\leqslant \max\{\lambda^*_k-\lambda^*_n, \mu^*_k-\mu^*_n\} \\ &\qquad \text{for all } k\in\{1,\dots,n-1\} \text{ and all } t\geqslant\tau, \end{aligned} \tag{$\mathbf{C_4}$}$$

then

$$\begin{cases} \lambda^*_n(1+b)+2\|\Omega^n-\Omega^k\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}} \leqslant \lambda^*_k(1+b), \\ \mu^*_n+2\left(1+\left|\tau-\dfrac{bt}{1+b}\right|\right)\|\Omega^n-\Omega^k\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}} \leqslant \mu^*_k. \end{cases}$$

(Once again, short times should be treated separately. Further note that the need for the factor $1+\tau$ in $(\mathbf{C_4})$ ultimately comes from the fact that we are composing also in the v variable, see the coefficient σ in the last norm of (4.30).) Then Proposition 4.28 (ii) yields

$$\|(\Omega^k_{t,\tau})^{-1}\circ\Omega^n_{t,\tau}-\mathrm{Id}\,\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}}\leqslant 2\|\Omega^k_{t,\tau}-\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}}\leqslant 4(\mathcal{R}^{k,n}_1+\mathcal{R}^{k,n}_2)(\tau,t).$$

10.3.5. Partial conclusion

At this point we have established $(\mathbf{E^n_\Omega})+(\widetilde{\mathbf{E}}^{\mathbf{n}}_{\mathbf{\Omega}})+(\mathbf{E^n_{\nabla\Omega}})$.

10.4. Estimates on the density and distribution along characteristics

In this subsection we establish $(\mathbf{E^{n+1}_\varrho})+(\widetilde{\mathbf{E}}^{\mathbf{n+1}}_{\varrho})+(\mathbf{E^{n+1}_h})+(\widetilde{\mathbf{E}}^{\mathbf{n+1}}_{\mathbf{h}})$.

10.4.1. Step 5. Estimate of $h^k\circ\Omega^n$ and $(\nabla h^k)\circ\Omega^n$, $k\leqslant n$.

Let $k\in\{1,...,n\}$. Since

$$h^k_\tau\circ\Omega^n_{t,\tau}=(h^k_\tau\circ\Omega^{k-1}_{t,\tau})\circ((\Omega^{k-1}_{t,\tau})^{-1}\circ\Omega^n_{t,\tau}),$$

the control on $h^k\circ\Omega^n$ will follow from the control on $h^k\circ\Omega^{k-1}$ in $(\mathbf{E^n_h})$, together with the control on $(\Omega^{k-1})^{-1}\circ\Omega^n$ in $(\widetilde{\mathbf{E}}^{\mathbf{n}}_{\mathbf{\Omega}})$. If

$$(1+\tau)\|(\Omega^{k-1}_{t,\tau})^{-1}\circ\Omega^n_{t,\tau}-\mathrm{Id}\,\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}}\leqslant\min\{\lambda_k-\lambda^*_n,\mu_k-\mu^*_n\},\tag{10.28}$$

then we can apply Proposition 4.25 and get, for any $p\in[1,\bar p]$, and $t\geqslant\tau\geqslant 0$,

$$\|h^k_\tau\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n;p}_{\tau-bt/(1+b)}}\leqslant\|h^k_\tau\circ\Omega^{k-1}_{t,\tau}\|_{\mathcal{Z}^{\lambda_k(1+b),\mu_k;p}_{\tau-bt/(1+b)}}\leqslant\delta_k.\tag{10.29}$$

In turn, (10.28) is satisfied if

$$\begin{aligned}4(1+\tau)(\mathcal{R}^{k,n}_1(\tau,t)+\mathcal{R}^{k,n}_2(\tau,t))&\leqslant\min\{\lambda_k-\lambda^*_n,\mu_k-\mu^*_n\}\\&\qquad\text{for all } k\in\{1,...,n\}\text{ and all }\tau\in[0,t];\end{aligned}\tag{$\mathbf{C_5}$}$$

we shall check later the feasibility of this condition.

Then, by the same argument, we also have

$$\sup_{0\leqslant\tau\leqslant t}\|(\nabla_x h^k_\tau)\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n;p}_{\tau-bt/(1+b)}}+\|((\nabla_v+\tau\nabla_x)h^k_\tau)\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n;p}_{\tau-bt/(1+b)}}\leqslant\delta_k$$
$$\text{for all } k\in\{1,...,n\}\text{ and all }p\in[1,\bar p].$$

10.4.2. Step 6. Estimate on $\varrho[h^{n+1}]$

This step is the first where we shall use the Vlasov equation. Starting from (8.4), we apply the method of characteristics to get, as in §8,

$$h^{n+1}(t, X^n_{0,t}(x,v), V^n_{0,t}(x,v)) = \int_0^t \Sigma^{n+1}(\tau, X^n_{0,\tau}(x,v), V^n_{0,\tau}(x,v))\, d\tau, \tag{10.30}$$

where

$$\Sigma^{n+1} = -(F[h^{n+1}]\cdot\nabla_v f^n + F[h^n]\cdot\nabla_v h^n).$$

We compose this with $(X^n_{t,0}, V^n_{t,0})$ and apply (5.2) to get

$$h^{n+1}(t,x,v) = \int_0^t \Sigma^{n+1}(\tau, X^n_{t,\tau}(x,v), V^n_{t,\tau}(x,v))\, d\tau,$$

and so, by integration in the v variable,

$$\begin{aligned}\varrho[h^{n+1}](t,x) &= \int_0^t \int_{\mathbb{R}^d} \Sigma^{n+1}(\tau, X^n_{t,\tau}(x,v), V^n_{t,\tau}(x,v))\, dv\, d\tau \\ &= -\int_0^t \int_{\mathbb{R}^d} (R^{n+1}_{\tau,t}\cdot G^n_{\tau,t})(x - v(t-\tau), v)\, dv\, d\tau \\ &\quad -\int_0^t \int_{\mathbb{R}^d} (R^n_{\tau,t}\cdot H^n_{\tau,t})(x - v(t-\tau), v)\, dv\, d\tau,\end{aligned} \tag{10.31}$$

where (with a slight inconsistency in the notation)

$$\begin{cases} R^{n+1}_{\tau,t} = F[h^{n+1}]\circ\Omega^n_{t,\tau}, & R^n_{\tau,t} = F[h^n]\circ\Omega^n_{t,\tau}, \\ G^n_{\tau,t} = (\nabla_v f^n)\circ\Omega^n_{t,\tau}, & H^n_{\tau,t} = (\nabla_v h^n)\circ\Omega^n_{t,\tau}. \end{cases} \tag{10.32}$$

Since the free transport semigroup and $\Omega^n_{t,\tau}$ are measure-preserving, for all $0 \leqslant \tau \leqslant t$ we have

$$\begin{aligned}\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} (R^{n+1}_{\tau,t}\cdot G^n_{\tau,t})(x - v(t-\tau), v)\, dv\, dx &= \int_{\mathbb{T}^d}\int_{\mathbb{R}^d} R^{n+1}_{\tau,t}\cdot G^n_{\tau,t}\, dv\, dx \\ &= \int_{\mathbb{T}^d}\int_{\mathbb{R}^d} F[h^{n+1}]\cdot\nabla_v f^n\, dv\, dx \\ &= \int_{\mathbb{T}^d}\int_{\mathbb{R}^d} \nabla_v\cdot(F[h^{n+1}] f^n)\, dv\, dx \\ &= 0,\end{aligned}$$

and similarly

$$\int_{\mathbb{T}^d}\int_{\mathbb{R}^d} (R^n_{\tau,t}\cdot H^n_{\tau,t})(x - v(t-\tau), v)\, dv\, dx = 0 \quad \text{for all } 0 \leqslant \tau \leqslant t.$$

This will allow us to apply the inequalities from §6.

Substep a. Let us first deal with the source term

$$\sigma^{n,n}(t,x) := \int_0^t \int_{\mathbb{R}^d} (R^n_{\tau,t}\cdot H^n_{\tau,t})(x-v(t-\tau),v)\,dv\,d\tau. \tag{10.33}$$

By Proposition 6.2,

$$\|\sigma^{n,n}(t,\cdot)\|_{\mathcal{F}^{\lambda^*_n t+\mu^*_n}} \leqslant \int_0^t \|R^n_{\tau,t}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}} \|H^n_{\tau,t}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n;1}_{\tau-bt/(1+b)}}\,d\tau. \tag{10.34}$$

On the one hand, we have from Step 5 that

$$\|H^n_{\tau,t}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n;1}_{\tau-bt/(1+b)}} \leqslant 2(1+\tau)\delta_n.$$

On the other hand, under condition $(\mathbf{C_1})$, we may apply Proposition 4.25 (with $\sigma=0$) to get

$$\|R^n_{\tau,t}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}} \leqslant \|F[h^n_\tau]\|_{\mathcal{F}^{\nu_n}},$$

where

$$\begin{aligned}\nu_n(t,\tau) &= \mu^*_n+\lambda^*_n(1+b)\left|\tau-\frac{bt}{1+b}\right| + \|\Omega^n X_{t,\tau}-\mathrm{Id}\,\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}}\\ &\leqslant \mu^*_n+\lambda^*_n(1+b)\left|\tau-\frac{bt}{1+b}\right| + 2\mathcal{R}^n_2(\tau,t).\end{aligned}$$

Proceeding as in Step 1 (treating small times separately), we deduce that

$$\begin{aligned}\|R^n_{\tau,t}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}} &\leqslant \|F[h^n_\tau]\|_{\mathcal{F}^{\nu_n}} \leqslant e^{-2\pi(\lambda_n-\lambda^*_n)\tau}\|F[h^n_\tau]\|_{\mathcal{F}^{\bar\nu_n}}\\ &\leqslant C_F e^{-2\pi(\lambda_n-\lambda^*_n)\tau}\|\varrho[h^n_\tau]\|_{\mathcal{F}^{\bar\nu_n}} \leqslant C_F e^{-2\pi(\lambda_n-\lambda^*_n)\tau}\delta_n,\end{aligned}$$

with

$$\bar\nu_n(\tau,t) := \begin{cases} \mu^\sharp, & \text{when } 0\leqslant\tau\leqslant\dfrac{bt}{1+b},\\ \lambda_n\tau+\mu_n, & \text{when } \tau\geqslant\dfrac{bt}{1+b}.\end{cases} \tag{10.35}$$

(We have used the gradient structure of the force to convert (gliding) regularity into decay.) Thus

$$\begin{aligned}\int_0^t \|R^n_{\tau,t}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}} \|H^n_{\tau,t}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n;1}_{\tau-bt/(1+b)}}\,d\tau &\leqslant 2C_F\delta_n^2\int_0^t e^{-2\pi(\lambda_n-\lambda^*_n)\tau}(1+\tau)\,d\tau\\ &\leqslant \frac{2C_F\delta_n^2}{(\pi(\lambda_n-\lambda^*_n))^2}.\end{aligned} \tag{10.36}$$

(Note that this is the power 2 which is responsible for the very fast convergence of the Newton scheme.)

Substep b. Now let us handle the term

$$\sigma^{n,n+1}(t,x) := \int_0^t \int_{\mathbb{R}^d} (R^{n+1}_{\tau,t} \cdot G^n_{\tau,t})(x - v(t-\tau), v)\, dv\, d\tau. \tag{10.37}$$

This is the focal point of all our analysis, because it is in this term that the self-consistent nature of the Vlasov equation appears. In particular, we will make crucial use of the time-cheating trick to overcome the loss of regularity implied by composition; and also the other bilinear estimates (regularity extortion) from §6, as well as the time-response study from §7. Particular care should be given to the zero spatial mode of G^n, which is associated with instantaneous response (no echo). In the linearized equation we did not see this problem because the contribution of the zero mode was vanishing!

We start by introducing

$$\overline{G}^n_{\tau,t} = \nabla_v f^0 + \sum_{k=1}^{n} \nabla_v (h^k_\tau \circ \Omega^{k-1}_{t,\tau}), \tag{10.38}$$

and we decompose $\sigma^{n,n+1}$ as

$$\sigma^{n,n+1} = \overline{\sigma}^{n,n+1} + \mathcal{E} + \overline{\mathcal{E}}, \tag{10.39}$$

where

$$\overline{\sigma}^{n,n+1}(t,x) = \int_0^t \int_{\mathbb{R}^d} F[h^{n+1}_\tau] \cdot \overline{G}^n_{\tau,t}(x - v(t-\tau), v)\, dv\, d\tau \tag{10.40}$$

and the error terms $\mathcal{E}$ and $\overline{\mathcal{E}}$ are defined by

$$\mathcal{E}(t,x) = \int_0^t \int_{\mathbb{R}^d} ((F[h^{n+1}_\tau] \circ \Omega^n_{t,\tau} - F[h^{n+1}_\tau]) \cdot G^n)(\tau, x - v(t-\tau), v)\, dv\, d\tau, \tag{10.41}$$

$$\overline{\mathcal{E}}(t,x) = \int_0^t \int_{\mathbb{R}^d} (F[h^{n+1}_\tau] \cdot (G^n - \overline{G}^n))(\tau, x - v(t-\tau), v)\, dv\, d\tau. \tag{10.42}$$

We shall first estimate $\mathcal{E}$ and $\overline{\mathcal{E}}$.

Control of $\mathcal{E}$. This is based on the time-cheating trick from §6, and the regularity of the force. By Proposition 6.2,

$$\|\mathcal{E}(t,\cdot)\|_{\mathcal{F}^{\lambda^*_n t + \mu^*_n}} \leqslant \int_0^t \|F[h^{n+1}_\tau] \circ \Omega^n_{t,\tau} - F[h^{n+1}_\tau]\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau - bt/(1+b)}} \|G^n\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n;1}_{\tau - bt/(1+b)}}\, d\tau. \tag{10.43}$$

From (10.1) and Step 5,

$$\begin{aligned}\|G^n\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*;1}_{\tau-bt/(1+b)}} &\leqslant \|\nabla_v f^0\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*;1}_{\tau-bt/(1+b)}}+\sum_{k=1}^n \|\nabla_v h^k_\tau\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*;1}_{\tau-bt/(1+b)}}\\ &\leqslant C_0'+\biggl(\sum_{k=1}^n\delta_k\biggr)(1+\tau),\end{aligned} \tag{10.44}$$

where C_0' comes from the contribution of f^0.

Next, by Propositions 4.24 and 4.25 (with $V=0$, $\tau=\sigma$ and $b=0$),

$$\begin{aligned}&\|F[h^{n+1}_\tau]\circ\Omega^n_{t,\tau}-F[h^{n+1}_\tau]\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{\tau-bt/(1+b)}}\\ &\quad\leqslant \|\Omega^n_{t,\tau}-\mathrm{Id}\,\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{\tau-bt/(1+b)}}\int_0^1 \|\nabla F[h^{n+1}_\tau]\circ(\mathrm{Id}+\theta(\Omega^n_{t,\tau}-\mathrm{Id}))\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{\tau-bt/(1+b)}}\,d\theta\\ &\quad\leqslant \|\nabla F[h^{n+1}_\tau]\|_{\mathcal{F}^{\nu_n}}\|\Omega^n_{t,\tau}-\mathrm{Id}\,\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{\tau-bt/(1+b)}},\end{aligned} \tag{10.45}$$

where

$$\nu_n=\mu_n^*+\lambda_n^*(1+b)\biggl|\tau-\frac{bt}{1+b}\biggr|+\|\Omega^n X_{t,\tau}-x\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{\tau-bt/(1+b)}}.$$

Small times are taken care of, as usual, by the initial regularity layer, so we only focus on the case $\tau\geqslant bt/(1+b)$; then

$$\begin{aligned}\nu_n&\leqslant\lambda_n^*\tau+\mu_n^*-\lambda_n^*b(t-\tau)+2\mathcal{R}^n(\tau,t)\\ &\leqslant\lambda_n^*\tau+\mu_n^*-\lambda_n^*\frac{B(t-\tau)}{1+t}+4C^1_\omega\biggl(\sum_{k=1}^n\frac{\delta_k}{(2\pi(\lambda_k-\lambda_k^*))^3}\biggr)\frac{\min\{t-\tau,1\}}{1+\tau}.\end{aligned}$$

To make sure that $\nu_n\leqslant\lambda_n^*\tau+\mu_n^*$, we assume that

$$4C^1_\omega\sum_{k=1}^n\frac{\delta_k}{(2\pi(\lambda_k-\lambda_k^*))^3}\leqslant\frac{\lambda_\infty^*B}{3}, \tag{$\mathbf{C_6}$}$$

and we note that

$$\frac{\min\{t-\tau,1\}}{1+\tau}\leqslant 3\frac{t-\tau}{1+t}.$$

(This is easily seen by separating four cases: (a) $t\leqslant 2$, (b) $t\geqslant 2$ and $t-\tau\leqslant 1$, (c) $t\geqslant 2$, $t-\tau\geqslant 1$ and $\tau\leqslant\frac12 t$, (d) $t\geqslant 2$, $t-\tau\geqslant 1$ and $\tau\geqslant\frac12 t$.)

Then, since $\gamma\geqslant 1$, we have

$$\begin{aligned}\|\nabla F[h^{n+1}_\tau]\|_{\mathcal{F}^{\nu_n}}&\leqslant\|\nabla F[h^{n+1}_\tau]\|_{\mathcal{F}^{\lambda_n^*\tau+\mu_n^*}}\\ &\leqslant\|F[h^{n+1}_\tau]\|_{\mathcal{F}^{\lambda_n^*\tau+\mu_n^*,\gamma}}\leqslant C_F\|\varrho[h^{n+1}_\tau]\|_{\mathcal{F}^{\lambda_n^*\tau+\mu_n^*}}.\end{aligned} \tag{10.46}$$

(Note that applying Proposition 4.10 instead of the regularity coming from the interaction would consume more regularity than we can afford to.)

Plugging this back into (10.45), we get

$$\begin{aligned}\|F[h_\tau^{n+1}]\circ\Omega_{t,\tau}^n - F[h_\tau^{n+1}]\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{\tau-bt/(1+b)}}&\\ \leqslant 2\mathcal{R}^n(\tau,t)C_F\|\varrho[h_\tau^{n+1}]\|_{\mathcal{F}^{\lambda_n^*\tau+\mu_n^*}}&\\ \leqslant 2C_\omega^3 C_F\left(\sum_{k=1}^n \frac{\delta_k}{(2\pi(\lambda_k-\lambda_k^*))^5}\right)\frac{1}{(1+\tau)^3}\|\varrho[h_\tau^{n+1}]\|_{\mathcal{F}^{\lambda_n^*\tau+\mu_n^*}}.&\end{aligned}$$

Recalling (10.41) and (10.44), applying Proposition 4.24, we conclude that

$$\begin{aligned}&\|\mathcal{E}(t,\cdot)\|_{\mathcal{F}^{\lambda_n^* t+\mu_n^*}}\\ &\leqslant 2C_\omega^3 C_F\left(C_0'+\sum_{k=1}^n\delta_k\right)\left(\sum_{k=1}^n\frac{\delta_k}{(2\pi(\lambda_k-\lambda_k^*))^5}\right)\int_0^t\|\varrho[h_\tau^{n+1}]\|_{\mathcal{F}^{\lambda_n^*\tau+\mu_n^*}}\frac{d\tau}{(1+\tau)^2}.\end{aligned}\tag{10.47}$$

(We could be a bit more precise; anyway we cannot go further since we do not yet have an estimate on $\varrho[h^{n+1}]$. Recall that the latter quantity has zero mean, so the $\mathcal{F}$ norm above could be replaced by a $\dot{\mathcal{F}}$ norm.)

Control of $\bar{\mathcal{E}}$. This will use the control on the derivatives of h^k. We start again from Proposition 6.2,

$$\|\bar{\mathcal{E}}(t,\cdot)\|_{\mathcal{F}^{\lambda_n^* t+\mu_n^*}}\leqslant\int_0^t\|G^n-\overline{G}^n\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*;1}_{\tau-bt/(1+b)}}\|F[h_\tau^{n+1}]\|_{\mathcal{F}^{\beta_n}}\,d\tau,\tag{10.48}$$

where

$$\beta_n=\lambda_n^*(1+b)\left|\tau-\frac{bt}{1+b}\right|+\mu_n^*.$$

We focus again on the case $\tau\geqslant bt/(1+b)$, so that (with crude estimates)

$$\|F[h_\tau^{n+1}]\|_{\mathcal{F}^{\beta_n}}\leqslant\|F[h_\tau^{n+1}]\|_{\mathcal{F}^{\lambda_n^*\tau+\mu_n^*}}\leqslant C_F\|\varrho[h_\tau^{n+1}]\|_{\mathcal{F}^{\lambda_n^*\tau+\mu_n^*}},$$

and the problem is to control $G^n-\overline{G}^n$,

$$\begin{aligned}\|G^n-\overline{G}^n\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*;1}_{\tau-bt/(1+b)}}\leqslant&\|(\nabla_v f^0)\circ\Omega_{t,\tau}^n-\nabla_v f^0\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*;1}_{\tau-bt/(1+b)}}\\ &+\sum_{k=1}^n\|(\nabla_v h_\tau^k)\circ\Omega_{t,\tau}^n-(\nabla_v h_\tau^k)\circ\Omega_{t,\tau}^{k-1}\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*;1}_{\tau-bt/(1+b)}}\\ &+\sum_{k=1}^n\|(\nabla_v h_\tau^k)\circ\Omega_{t,\tau}^{k-1}-\nabla_v(h_\tau^k\circ\Omega_{t,\tau}^{k-1})\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*;1}_{\tau-bt/(1+b)}}.\end{aligned}\tag{10.49}$$

By induction hypothesis $(\widetilde{\mathbf{E}}^{\mathbf{n}}_{\mathbf{h}})$, and since the $\mathcal{Z}^{\lambda,\mu}_\tau$ norms are increasing as a function of λ and μ,

$$\sum_{k=1}^{n} \|(\nabla_v h^k_\tau)\circ\Omega^{k-1}_{t,\tau} - \nabla_v(h^k_\tau\circ\Omega^{k-1}_{t,\tau})\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n;1}_{\tau-bt/(1+b)}} \leqslant \left(\sum_{k=1}^{n}\delta_k\right)\frac{1}{(1+\tau)^2}.$$

It remains to treat the first and second terms in the right-hand side of (10.49). This is done by inversion/composition as in Step 5; let us consider for instance the contribution of h^k, $k\geqslant 1$,

$$\begin{aligned}
&\|\nabla_v h^k_\tau\circ\Omega^n_{t,\tau} - \nabla_v h^k_\tau\circ\Omega^{k-1}_{t,\tau}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n;1}_{\tau-bt/(1+b)}}\\
&\quad\leqslant \int_0^1 \|\nabla\nabla_v h^k_\tau\circ((1-\theta)\Omega^n_{t,\tau}+\theta\Omega^{k-1}_{t,\tau})\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n;1}_{\tau-bt/(1+b)}}\|\Omega^n_{t,\tau}-\Omega^{k-1}_{t,\tau}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}}\,d\theta\\
&\quad\leqslant 2\|\nabla\nabla_v h^k_\tau\circ\Omega^{k-1}_{t,\tau}\|_{\mathcal{Z}^{\lambda^*_k(1+b),\mu^*_k;1}_{\tau-bt/(1+b)}}\|\Omega^n_{t,\tau}-\Omega^{k-1}_{t,\tau}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}}\\
&\quad\leqslant 4\delta_k(1+\tau)^2\mathcal{R}^{k-1,n}(\tau,t)\\
&\quad\leqslant 4C^4_\omega\delta_k\left(\sum_{j=k}^{n}\frac{\delta_j}{(2\pi(\lambda_j-\lambda^*_j))^6}\right)\frac{1}{(1+\tau)^2},
\end{aligned}$$

where in the second-last step we used $(\widetilde{\mathbf{E}}^{\mathbf{n}}_{\mathbf{\Omega}})$, $(\widetilde{\mathbf{E}}^{\mathbf{n}}_{\varrho})$, Propositions 4.24 and 4.28, condition $(\mathbf{C_5})$ and the same reasoning as in Step 5.

Summing up all contributions and inserting into (10.48) yields

$$\begin{aligned}
\|\bar{\mathcal{E}}(t,\cdot)\|_{\mathcal{F}^{\lambda^*_n t+\mu^*_n}} \leqslant 4C_F&\left[C^4_\omega\left(C'_0+\sum_{k=1}^{n}\delta_k\right)\left(\sum_{j=1}^{n}\frac{\delta_j}{(2\pi(\lambda_j-\lambda^*_n))^6}\right)+\sum_{k=1}^{n}\delta_k\right]\\
&\times\int_0^t \|\varrho[h^{n+1}_\tau]\|_{\mathcal{F}^{\lambda^*_n\tau+\mu^*_n}}\frac{d\tau}{(1+\tau)^2}.
\end{aligned}\tag{10.50}$$

Main contribution. Now we consider $\bar\sigma^{n,n+1}$, which we decompose as

$$\bar\sigma^{n,n+1}_t = \bar\sigma^{n,n+1}_{t,0}+\sum_{k=1}^{n}\bar\sigma^{n,n+1}_{t,k},$$

where

$$\begin{aligned}
\bar\sigma^{n,n+1}_{t,0}(x) &= \int_0^t\int_{\mathbb{R}^d} F[h^{n+1}](\tau,x-v(t-\tau),v)\cdot\nabla_v f^0(v)\,dv\,d\tau,\\
\bar\sigma^{n,n+1}_{t,k}(x) &= \int_0^t\int_{\mathbb{R}^d} (F[h^{n+1}_\tau]\cdot\nabla_v(h^k_\tau\circ\Omega^{k-1}_{t,\tau}))(\tau,x-v(t-\tau),v)\,dv\,d\tau.
\end{aligned}$$

Note that their zero modes vanish. For any $k \geqslant 1$, we apply Theorem 6.5 (with $M=1$) to get

$$\|\bar{\sigma}_{t,k}^{n,n+1}\|_{\mathcal{F}^{\lambda_n^* t+\mu_n^*}} \leqslant \int_0^t K_1^{n,h^k}(t,\tau)\|F[h_\tau^{n+1}]\|_{\mathcal{F}^{\nu'_n,\gamma}}\,d\tau + \int_0^t K_0^{n,h^k}(t,\tau)\|F[h_\tau^{n+1}]\|_{\mathcal{F}^{\nu'_n,\gamma}}\,d\tau,$$

where

$$\nu'_n = \lambda_n^*(1+b)\left|\tau - \frac{bt}{1+b}\right| + \mu'_n$$

and

$$K_1^{n,h^k}(t,\tau) = K_1^{n,k}(t,\tau)\sup_{0\leqslant s\leqslant t}\frac{\|\nabla_v(h_s^k\circ\Omega_{t,s}^{k-1}) - \langle\nabla_v(h_s^k\circ\Omega_{t,s}^{k-1})\rangle\|_{\mathcal{Z}^{\lambda_k(1+b),\mu_k}_{s-bt/(1+b)}}}{1+s},$$

$$K_1^{n,k}(t,\tau) = (1+\tau)d\sup_{l,m\in\mathbb{Z}^d_*} e^{-\pi(\mu_k-\mu_n^*)|m|}\frac{e^{-2\pi(\mu'_n-\mu_n^*)|l-m|}}{1+|l-m|^\gamma}e^{-\pi(\lambda_k-\lambda_n^*)|l(t-\tau)+m\tau|},$$

$$K_0^{n,h^k}(t,\tau) = K_0^{n,k}(t,\tau)\sup_{0\leqslant s\leqslant t}\|\nabla_v\langle h_s^k\circ\Omega_{t,s}^{k-1}\rangle\|_{\mathcal{C}^{\lambda_k(1+b);1}},$$

$$K_0^{n,k}(t,\tau) = de^{-\pi(\lambda_k-\lambda_n^*)(t-\tau)}.$$

We assume that

$$\mu'_n = \mu_n^* + \eta\left(\frac{t-\tau}{1+t}\right), \quad \eta > 0 \text{ small}, \tag{10.51}$$

and check that $\nu'_n \leqslant \lambda_n^*\tau + \mu_n^*$. Leaving apart the small-time case, we assume $\tau \geqslant bt/(1+b)$, so that

$$\nu'_n = (\lambda_n^*\tau+\mu_n^*) - \frac{B\lambda_n^*(t-\tau)}{1+t} + \eta\left(\frac{t-\tau}{1+t}\right),$$

which is indeed bounded above by $\lambda_n^*\tau+\mu_n^*$ as soon as

$$\eta \leqslant B\lambda_\infty^*. \tag{10.52}$$

Then, with the notation (7.10),

$$K_1^{n,k}(t,\tau) \leqslant K_1^{(\alpha_{n,k}),\gamma}(t,\tau), \tag{10.53}$$

with

$$\alpha_{n,k} = \pi\min\{\mu_k-\mu_n^*, \lambda_k-\lambda_n^*, 2\eta\}. \tag{10.54}$$

From the controls on h^k (assumption $(\widetilde{\mathbf{E}}^{\mathbf{n}}_{\mathbf{h}})$) we have that

$$\|\nabla_v(h_\tau^k\circ\Omega_{t,\tau}^{k-1}) - \langle\nabla_v(h_\tau^k\circ\Omega_{t,\tau}^{k-1})\rangle\|_{\mathcal{Z}^{\lambda_k(1+b),\mu_k;1}_{\tau-bt/(1+b)}} \leqslant \|\nabla_v(h_\tau^k\circ\Omega_{t,\tau}^{k-1})\|_{\mathcal{Z}^{\lambda_k(1+b),\mu_k;1}_{\tau-bt/(1+b)}} \leqslant \delta_k(1+\tau)$$

and

$$\begin{aligned}\|\langle\nabla_v(h_\tau^k\circ\Omega_{t,\tau}^{k-1})\rangle\|_{\mathcal{C}^{\lambda_k(1+b);1}}&=\|\langle(\nabla_v+\tau\nabla_x)(h_\tau^k\circ\Omega_{t,\tau}^{k-1})\rangle\|_{\mathcal{C}^{\lambda_k(1+b);1}}\\&\leqslant\|(\nabla_v+\tau\nabla_x)(h_\tau^k\circ\Omega_{t,\tau}^{k-1})\|_{\mathcal{Z}^{\lambda_k(1+b);1}_{\tau-bt/(1+b)}}\leqslant\delta_k.\end{aligned}$$

After controlling $F[h^{n+1}]$ by $\varrho[h^{n+1}]$, we end up with

$$\begin{aligned}\|\overline{\sigma}_{t,k}^{n,n+1}\|_{\mathcal{F}^{\lambda_n^*t+\mu_n^*}}&\leqslant C_F\int_0^t\biggl(\sum_{k=1}^n\delta_kK_1^{(\alpha_{n,k}),\gamma}(t,\tau)\biggr)\|\varrho[h_\tau^{n+1}]\|_{\mathcal{F}^{\lambda_n^*\tau+\mu_n^*}}\,d\tau\\&\quad+C_F\int_0^t\biggl(\sum_{k=1}^n\delta_k\;e^{-\pi(\lambda_k-\lambda_n^*)(t-\tau)}\biggr)\|\varrho[h_\tau^{n+1}]\|_{\mathcal{F}^{\lambda_n^*\tau+\mu_n^*}}\,d\tau,\end{aligned}\tag{10.55}$$

with $\alpha_{n,k}$ defined by (10.54).

Substep c. Gathering all previous controls, we obtain the following integral inequality for $\varphi=\varrho[h^{n+1}]$:

$$\begin{aligned}&\biggl\|\varphi(t,x)-\int_0^t\int_{\mathbb{R}^d}(\nabla W*\varphi)(\tau,x-v(t-\tau))\cdot\nabla_vf^0(v)\,dv\,d\tau\biggr\|_{\mathcal{F}^{\lambda_n^*t+\mu_n^*}}\\&\qquad\leqslant A_n+\int_0^t\biggl(K_1^n(t,\tau)+K_0^n(t,\tau)+\frac{c_0^n}{(1+\tau)^2}\biggr)\|\varphi(\tau,\cdot)\|_{\mathcal{F}^{\lambda_n^*\tau+\mu_n^*}}\,d\tau,\end{aligned}\tag{10.56}$$

where, by (10.36), (10.47) and (10.50),

$$A_n=\sup_{t\geqslant0}\|\sigma^{n,n}(t,\cdot)\|_{\mathcal{F}^{\lambda_n^*t+\mu_n^*}}\leqslant\frac{2C_F\delta_n^2}{(\pi(\lambda_n-\lambda_n^*))^2},\tag{10.57}$$

$$\begin{aligned}K_1^n(t,\tau)&=\biggl(C_F\sum_{k=1}^n\delta_k\biggr)K_1^{(\alpha_n),\gamma}(t,\tau),\\\alpha_n&=\alpha_{n,n}=\pi\min\{(\mu_n-\mu_n^*),(\lambda_n-\lambda_n^*),2\eta\},\\K_0^n(t,\tau)&=C_F\sum_{k=1}^n\delta_ke^{-\pi(\lambda_k-\lambda_n^*)(t-\tau)},\\c_0^n&=3C_FC_\omega^4\biggl(C_0'+\sum_{k=1}^n\delta_k\biggr)\biggl(\sum_{k=1}^n\frac{\delta_k}{(2\pi(\lambda_k-\lambda_k^*))^6}\biggr)+\sum_{k=1}^n\delta_k.\end{aligned}$$

(We are cheating a bit when writing (10.56), because in fact one should take into account small times separately; but this does not cause any difficulty.)

We easily estimate K_0^n:

$$\begin{aligned}\int_0^tK_0^n(t,\tau)\,d\tau&\leqslant C_F\sum_{k=1}^n\frac{\delta_k}{\pi(\lambda_k-\lambda_n^*)},\\\int_\tau^\infty K_0^n(t,\tau)\,dt&\leqslant C_F\sum_{k=1}^n\frac{\delta_k}{\pi(\lambda_k-\lambda_n^*)},\\\biggl(\int_0^tK_0^n(t,\tau)^2\,d\tau\biggr)^{1/2}&\leqslant C_F\sum_{k=1}^n\frac{\delta_k}{\sqrt{2\pi(\lambda_k-\lambda_n^*)}}.\end{aligned}$$

Let us assume that α_n is smaller than $\bar{\alpha}(\gamma)$ appearing in Theorem 7.7, and that

$$3C_F C_\omega^4 \left(C_0' + \sum_{k=1}^n \delta_k + 1\right)\left(\sum_{k=1}^n \frac{\delta_k}{(2\pi(\lambda_k - \lambda_k^*))^6}\right) \leqslant \frac{1}{4}, \tag{$\mathbf{C_7}$}$$

$$C_F \sum_{k=1}^n \frac{\delta_k}{\sqrt{2\pi(\lambda_k - \lambda_k^*)}} \leqslant \frac{1}{2}, \tag{$\mathbf{C_8}$}$$

$$C_F \sum_{k=1}^n \frac{\delta_k}{\pi(\lambda_k - \lambda_k^*)} \leqslant \max\left\{\frac{1}{4}, \chi\right\}, \tag{$\mathbf{C_9}$}$$

(note that in these conditions we have strenghtened the inequalities by replacing $\lambda_k - \lambda_n^*$ by $\lambda_k - \lambda_k^*$ where $\chi > 0$ is also defined by Theorem 7.7). Applying Theorem 7.7 with $\lambda_0 = \lambda$ and $\lambda^* = \lambda_1$, we deduce that for any $\varepsilon \in (0, \alpha_n)$ and $t \geqslant 0$,

$$\|\varrho_t^{n+1}\|_{\mathcal{F}^{\lambda_n^* t + \mu_n^*}} \leqslant C A_n \frac{(1+c_0^n)^2}{\sqrt{\varepsilon}} e^{C c_0^n} \left(1 + \frac{c_n}{\alpha_n \varepsilon}\right) e^{C T_{\varepsilon,n}} e^{C c_n (1 + T_{\varepsilon,n}^2)} e^{\varepsilon t}, \tag{10.58}$$

where

$$c_n = 2C_F \sum_{k=1}^n \delta_k$$

and

$$T_{\varepsilon,n} = C_\gamma \max\left\{\left(\frac{c_n^2}{\alpha_n^5 \varepsilon^{2+\gamma}}\right)^{1/(\gamma-1)}, \left(\frac{c_n}{\alpha_n^2 \varepsilon^{\gamma+1/2}}\right)^{1/(\gamma-1)}, \frac{(c_0^n)^{2/3}}{\varepsilon^{1/3}}\right\}.$$

Pick up $\lambda_n^\dagger < \lambda_n^*$ such that $2\pi(\lambda_n^* - \lambda_n^\dagger) \leqslant \alpha_n$, and choose $\varepsilon = 2\pi(\lambda_n^* - \lambda_n^\dagger)$; recalling that $\hat{\varrho}^{n+1}(t,0) = 0$, and that our conditions imply an upper bound on c_n and c_0^n, we deduce the *uniform* control

$$\begin{aligned} \|\varrho_t^{n+1}\|_{\mathcal{F}^{\lambda_n^\dagger t + \mu_n^*}} &\leqslant e^{-2\pi(\lambda_n^* - \lambda_n^\dagger)t} \|\varrho_t^{n+1}\|_{\mathcal{F}^{\lambda_n^* t + \mu_n^*}} \\ &\leqslant C A_n \left(1 + \frac{1}{\alpha_n (\lambda_n^* - \lambda_n^\dagger)^{3/2}}\right) e^{C T_n^2}, \end{aligned} \tag{10.59}$$

where

$$T_n = C\left(\frac{1}{\alpha_n^5 (\lambda_n^* - \lambda_n^\dagger)^{2+\gamma}}\right)^{1/(\gamma-1)}. \tag{10.60}$$

10.4.3. Step 7. Estimate on $F[h^{n+1}]$

As an immediate consequence of (10.2) and (10.59), we have

$$\sup_{t \geqslant 0} \|F[\varrho_t^{n+1}]\|_{\mathcal{F}^{\lambda_n^\dagger t + \mu_n^*, \gamma}} \leqslant C A_n \left(1 + \frac{1}{\alpha_n (\lambda_n^* - \lambda_n^\dagger)^{3/2}}\right) e^{C T_n^2}. \tag{10.61}$$

10.4.4. Step 8. Estimate of $h^{n+1}\circ\Omega^n$

In this step we shall again use the Vlasov equation. We rewrite (10.30) as

$$h^{n+1}(\tau, X^n_{0,\tau}(x,v), V^n_{0,\tau}(x,v)) = \int_0^\tau \Sigma^{n+1}(s, X^n_{0,s}(x,v), V^n_{0,s}(x,v))\, ds;$$

but now we compose with $(X^n_{t,0}, V^n_{t,0})$, *where* $t \geqslant \tau$ *is arbitrary.* This gives

$$h^{n+1}(\tau, X^n_{t,\tau}(x,v), V^n_{t,\tau}(x,v)) = \int_0^\tau \Sigma^{n+1}(s, X^n_{t,s}(x,v), V^n_{t,s}(x,v))\, ds.$$

Then for any $p \in [1, \bar{p}]$ and $\lambda^\flat_n < \lambda^\dagger_n$, using Propositions 4.19 and 4.24, and the notation (10.32), we get

$$\begin{aligned}
\|h^{n+1}_\tau \circ \Omega^n_{t,\tau}\|_{\mathcal{Z}^{(1+b)\lambda^\flat_n, \mu^*_n; p}_{\tau - bt/(1+b)}} &= \|h^{n+1}_\tau \circ (X^n_{t,\tau}, V^n_{t,\tau})\|_{\mathcal{Z}^{(1+b)\lambda^\flat_n, \mu^*_n; p}_{t - bt/(1+b)}} \\
&\leqslant \int_0^\tau \|\Sigma^{n+1}(s, X^n_{t,s}, V^n_{t,s})\|_{\mathcal{Z}^{(1+b)\lambda^\flat_n, \mu^*_n; p}_{t - bt/(1+b)}}\, ds \\
&= \int_0^\tau \|\Sigma^{n+1}(s, \Omega^n_{t,s})\|_{\mathcal{Z}^{(1+b)\lambda^\flat_n, \mu^*_n; p}_{s - bt/(1+b)}}\, ds \\
&\leqslant \int_0^\tau \|R^{n+1}_{s,t}\|_{\mathcal{Z}^{(1+b)\lambda^\flat_n, \mu^*_n}_{s - bt/(1+b)}} \|G^n_{s,t}\|_{\mathcal{Z}^{(1+b)\lambda^\flat_n, \mu^*_n; p}_{s - bt/(1+b)}}\, ds \\
&\quad + \int_0^\tau \|R^n_{s,t}\|_{\mathcal{Z}^{(1+b)\lambda^\flat_n, \mu^*_n}_{s - bt/(1+b)}} \|H^n_{s,t}\|_{\mathcal{Z}^{(1+b)\lambda^\flat_n, \mu^*_n; p}_{s - bt/(1+b)}}\, ds.
\end{aligned}$$

Then (proceeding as in Step 6 to check that the exponents lie in the appropriate range)

$$\|R^{n+1}_{s,t}\|_{\mathcal{Z}^{(1+b)\lambda^\flat_n, \mu^*_n}_{s - bt/(1+b)}} \leqslant C_F e^{-2\pi(\lambda^\dagger_n - \lambda^\flat_n)s} \|\varrho^{n+1}_s\|_{\mathcal{F}^{\bar{\nu}_n(s)}}$$

and

$$\|R^n_{s,t}\|_{\mathcal{Z}^{(1+b)\lambda^\flat, \mu^*_n}_{s - bt/(1+b)}} \leqslant C_F e^{-2\pi(\lambda^\dagger_n - \lambda^\flat_n)s} \|\varrho^n_s\|_{\mathcal{F}^{\bar{\nu}_n(s)}} \leqslant C_F e^{-2\pi(\lambda^\dagger_n - \lambda^\flat_n)s} \delta_n,$$

with

$$\bar{\nu}_n(s,t) := \begin{cases} \mu^\sharp, & \text{when } s \leqslant \dfrac{bt}{1+b}, \\ \lambda^\dagger_n s + \mu^*_n, & \text{when } s \geqslant \dfrac{bt}{1+b}. \end{cases}$$

On the other hand, from the induction assumption $(\mathbf{E^n_h})$–$(\widetilde{\mathbf{E}}^{\mathbf{n}}_{\mathbf{h}})$ (and again control of composition via Proposition 4.25),

$$\|H^n_{s,t}\|_{\mathcal{Z}^{(1+b)\lambda^\flat_n, \mu^*_n; p}_{s - bt/(1+b)}} \leqslant 2(1+s)\delta_n$$

and

$$\|G^n_{s,t}\|_{\mathcal{Z}^{(1+b)\lambda^\flat_n,\mu^*_n;p}_{s-bt/(1+b)}} \leqslant 2(1+s)\sum_{k=1}^n \delta_k.$$

We deduce that

$$y(t,\tau) := \|h^{n+1}_\tau \circ \Omega^n_{t,\tau}\|_{\mathcal{Z}^{(1+b)\lambda^\flat_n,\mu^*_n;p}_{\tau-bt/(1+b)}}$$

satisfies

$$\begin{aligned} y(t,\tau) &\leqslant 2C_F\left(\sum_{k=1}^n \delta_k\right)\int_0^\tau e^{-2\pi(\lambda^\dagger_n-\lambda^\flat_n)s}\|\varrho^{n+1}_s\|_{\mathcal{F}^{\bar\nu_n(s)}}(1+s)\,ds \\ &\quad + 2C_F\delta_n^2\int_0^\tau e^{-2\pi(\lambda^\dagger_n-\lambda^\flat_n)s}(1+s)\,ds; \end{aligned}$$

so, for all $0\leqslant\tau\leqslant t$,

$$\|h^{n+1}_\tau\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{(1+b)\lambda^\flat_n,\mu^*_n;p}_{\tau-bt/(1+b)}} \leqslant \frac{4C_F\max\{\sum_{k=1}^n\delta_k,1\}}{(2\pi(\lambda^\dagger_n-\lambda^\flat_n))^2}\Big(\delta_n^2+\sup_{s\geqslant 0}\|\varrho^{n+1}_s\|_{\mathcal{F}^{\bar\nu_n(s)}}\Big). \tag{10.62}$$

10.4.5. Step 9. Crude estimates on the derivatives of h^{n+1}

Again we choose $p\in[1,\bar p]$. From the previous step and Proposition 4.27 we deduce, for any $\lambda^\ddagger_n$ such that $\lambda^\ddagger_n<\lambda^\flat_n<\lambda^\ddagger_n$, and any $\mu^\ddagger_n<\mu^*_n$, that

$$\begin{aligned} &\|\nabla_x(h^{n+1}_\tau\circ\Omega^n_{t,\tau})\|_{\mathcal{Z}^{\lambda^\ddagger_n(1+b),\mu^\ddagger_n;p}_{\tau-bt/(1+b)}} + \|(\nabla_v+\tau\nabla_x)(h^{n+1}_\tau\circ\Omega^n_{t,\tau})\|_{\mathcal{Z}^{\lambda^\ddagger_n(1+b),\mu^\ddagger_n;p}_{\tau-bt/(1+b)}} \\ &\qquad\leqslant \frac{C(d)}{\min\{\lambda^\flat_n-\lambda^\ddagger_n,\mu^*_n-\mu^\ddagger_n\}}\|h^{n+1}_\tau\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^\flat_n(1+b),\mu^*_n;p}_{\tau-bt/(1+b)}} \end{aligned} \tag{10.63}$$

and

$$\|\nabla(h^{n+1}_\tau\circ\Omega^n_{t,\tau})\|_{\mathcal{Z}^{\lambda^\ddagger_n(1+b),\mu^\ddagger_n;p}_{\tau-bt/(1+b)}} \leqslant \frac{C(d)(1+\tau)}{\min\{\lambda^\flat_n-\lambda^\ddagger_n,\mu^*_n-\mu^\ddagger_n\}}\|h^{n+1}_\tau\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^\flat_n(1+b),\mu^*_n;p}_{\tau-bt/(1+b)}}. \tag{10.64}$$

Similarly,

$$\|\nabla\nabla(h^{n+1}_\tau\circ\Omega^n_{t,\tau})\|_{\mathcal{Z}^{\lambda^\ddagger_n(1+b),\mu^\ddagger_n;p}_{\tau-bt/(1+b)}} \leqslant \frac{C(d)(1+\tau)^2}{\min\{\lambda^\flat_n-\lambda^\ddagger_n,\mu^*_n-\mu^\ddagger_n\}^2}\|h^{n+1}_\tau\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^\flat_n(1+b),\mu^*_n;p}_{\tau-bt/(1+b)}}. \tag{10.65}$$

10.4.6. Step 10. Chain-rule and refined estimates on derivatives of h^{n+1}

From Step 3 we have

$$\|\nabla\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}} + \|(\nabla\Omega^n_{t,\tau})^{-1}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}} \leqslant C(d) \tag{10.66}$$

and (via Proposition 4.27)

$$\begin{aligned}\|\nabla\nabla\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^\ddagger_n(1+b),\mu^\ddagger_n}_{\tau-bt/(1+b)}} &\leqslant \frac{C(d)(1+\tau)}{\min\{\lambda^*_n-\lambda^\ddagger_n,\mu^*_n-\mu^\ddagger_n\}}\|\nabla\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^*_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}}\\ &\leqslant \frac{C(d)(1+\tau)}{\min\{\lambda^*_n-\lambda^\ddagger_n,\mu^*_n-\mu^\ddagger_n\}}.\end{aligned} \tag{10.67}$$

Combining these bounds with Step 9, Proposition 4.24 and the identities

$$\begin{cases}(\nabla h)\circ\Omega=(\nabla\Omega)^{-1}\nabla(h\circ\Omega),\\ (\nabla^2 h)\circ\Omega=(\nabla\Omega)^{-2}\nabla^2(h\circ\Omega)-(\nabla\Omega)^{-1}\nabla^2\Omega(\nabla\Omega)^{-1}(\nabla h\circ\Omega),\end{cases} \tag{10.68}$$

we get

$$\begin{aligned}\|(\nabla h^{n+1}_\tau)\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^\ddagger_n(1+b),\mu^\ddagger_n;1}_{\tau-bt/(1+b)}} &\leqslant C(d)\|\nabla(h^{n+1}_\tau\circ\Omega^n_{t,\tau})\|_{\mathcal{Z}^{\lambda^\ddagger_n(1+b),\mu^\ddagger_n;1}_{\tau-bt/(1+b)}}\\ &\leqslant \frac{C(d)(1+\tau)}{\min\{\lambda^\flat_n-\lambda^\ddagger_n,\mu^*_n-\mu^\ddagger_n\}}\|h^{n+1}_\tau\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^\flat_n(1+b),\mu^*_n;1}_{\tau-bt/(1+b)}}\end{aligned} \tag{10.69}$$

and

$$\begin{aligned}&\|(\nabla^2 h^{n+1}_\tau)\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^\ddagger_n(1+b),\mu^\ddagger_n;1}_{\tau-bt/(1+b)}}\\ &\quad\leqslant C(d)\Big[\|\nabla^2(h^{n+1}_\tau\circ\Omega^n_{t,\tau})\|_{\mathcal{Z}^{\lambda^\ddagger_n(1+b),\mu^\ddagger_n;1}_{\tau-bt/(1+b)}}\\ &\qquad\qquad+\|\nabla^2\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^\ddagger_n(1+b),\mu^\ddagger_n;1}_{\tau-bt/(1+b)}}\|(\nabla h^{n+1}_\tau)\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^\ddagger_n(1+b),\mu^\ddagger_n;1}_{\tau-bt/(1+b)}}\Big]\\ &\quad\leqslant \frac{C(d)(1+\tau)^2}{\min\{\lambda^\flat_n-\lambda^\ddagger_n,\mu^*_n-\mu^\ddagger_n\}^2}\|h^{n+1}_\tau\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^\flat_n(1+b),\mu^*_n;p}_{\tau-bt/(1+b)}}.\end{aligned} \tag{10.70}$$

This gives us the bounds

$$\|(\nabla h^{n+1})\circ\Omega^n\|=O(1+\tau)\quad\text{and}\quad\|(\nabla^2 h^{n+1})\circ\Omega^n\|=O((1+\tau)^2),$$

which are optimal if one does not distinguish between the x and v variables. We shall now refine these estimates. We first write

$$\nabla(h^{n+1}_\tau\circ\Omega^n_{t,\tau})-(\nabla h^{n+1}_\tau)\circ\Omega^n_{t,\tau}=\nabla(\Omega^n_{t,\tau}-\mathrm{Id})\cdot[(\nabla h^{n+1}_\tau)\circ\Omega^n_{t,\tau}],$$

and we deduce (via Propositions 4.24 and 4.27)

$$\begin{aligned}&\|\nabla(h^{n+1}_\tau\circ\Omega^n_{t,\tau})-(\nabla h^{n+1}_\tau)\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^\ddagger_n(1+b),\mu^\ddagger_n;p}_{\tau-bt/(1+b)}}\\ &\quad\leqslant\|\nabla(\Omega^n_{t,\tau}-\mathrm{Id})\|_{\mathcal{Z}^{\lambda^\ddagger_n(1+b),\mu^\ddagger_n}_{\tau-bt/(1+b)}}\|(\nabla h^{n+1}_\tau)\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^\ddagger_n(1+b),\mu^\ddagger_n;p}_{\tau-bt/(1+b)}}\\ &\quad\leqslant C(d)\left(\frac{1+\tau}{\min\{\lambda^\flat_n-\lambda^\ddagger_n,\mu^*_n-\mu^\ddagger_n\}}\right)^2\|\Omega^n_{t,\tau}-\mathrm{Id}\,\|_{\mathcal{Z}^{\lambda^\flat_n(1+b),\mu^*_n}_{\tau-bt/(1+b)}}\|h^{n+1}_\tau\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^\flat_n(1+b),\mu^*_n;p}_{\tau-bt/(1+b)}}\\ &\quad\leqslant\frac{C(d)C^4_\omega}{\min\{\lambda^\flat_n-\lambda^\ddagger_n,\mu^*_n-\mu^\ddagger_n\}^2}\left(\sum_{k=1}^n\frac{\delta_k}{(2\pi(\lambda_k-\lambda^*_k))^6}\right)(1+\tau)^{-2}\|h^{n+1}_\tau\circ\Omega^n_{t,\tau}\|_{\mathcal{Z}^{\lambda^\flat_n(1+b),\mu^*_n;p}_{\tau-bt/(1+b)}}.\end{aligned} \tag{10.71}$$

(Note that $\Omega^n - \mathrm{Id}$ brings the time-decay, while h^{n+1} brings the smallness.)

This shows that $(\nabla h^{n+1}) \circ \Omega^n \simeq \nabla (h^{n+1} \circ \Omega^n)$ as $\tau \to \infty$. In view of Step 9, this also implies the refined gradient estimates

$$\begin{aligned} \|(\nabla_x h_\tau^{n+1}) \circ \Omega_{t,\tau}^n\|_{\mathcal{Z}^{\lambda_n^\ddagger, \mu_n^\ddagger; p}_{\tau - bt/(1+b)}} + \|((\nabla_v + \tau \nabla_x) h_\tau^{n+1}) \circ \Omega_{t,\tau}^n\|_{\mathcal{Z}^{\lambda_n^\ddagger, \mu_n^\ddagger; p}_{\tau - bt/(1+b)}} \\ \leqslant \overline{C} \|h_\tau^{n+1} \circ \Omega_{t,\tau}^n\|_{\mathcal{Z}^{\lambda_n^\flat (1+b), \mu_n^*; p}_{\tau - bt/(1+b)}}, \end{aligned} \tag{10.72}$$

with

$$\overline{C} = C(d) \left(\frac{C_\omega^4}{\min\{\lambda_n^\flat - \lambda_n^\ddagger, \mu_n^* - \mu_n^\ddagger\}^2} \sum_{k=1}^n \frac{\delta_k}{(2\pi(\lambda_k - \lambda_k^*))^6} + \frac{1}{\min\{\lambda_n^\flat - \lambda_n^\ddagger, \mu_n^* - \mu_n^\ddagger\}} \right).$$

10.4.7. Conclusion

Given $\lambda_{n+1} < \lambda_n^*$ and $\mu_{n+1} < \mu_n^*$, we define

$$\lambda_{n+1} = \lambda_n^\ddagger \quad \text{and} \quad \mu_{n+1} = \mu_n^\ddagger,$$

and we impose that

$$\lambda_n^* - \lambda_n^\dagger = \lambda_n^\dagger - \lambda_n^\flat = \lambda_n^\flat - \lambda_n^\ddagger = \tfrac{1}{3}(\lambda_n^* - \lambda_{n+1}) \quad \text{and} \quad \mu_n^* - \mu_n^\ddagger = \mu_n^* - \mu_{n+1}.$$

Then from (10.59), (10.61)–(10.63) and (10.70)–(10.72) we see that $(\mathbf{E}_\varrho^{\mathbf{n+1}})$, $(\widetilde{\mathbf{E}}_\varrho^{\mathbf{n+1}})$, $(\mathbf{E}_{\mathbf{h}}^{\mathbf{n+1}})$ and $(\widetilde{\mathbf{E}}_{\mathbf{h}}^{\mathbf{n+1}})$ have all been established in the present subsection, with

$$\delta_{n+1} = \frac{C(d) C_F (1 + C_F)(1 + C_\omega^4) e^{C\, T_n^2}}{\min\{\lambda_n^* - \lambda_{n+1}, \mu_n^* - \mu_{n+1}\}^9} \max\left\{1, \sum_{k=1}^n \delta_k, \right\} \left(1 + \sum_{k=1}^n \frac{\delta_k}{(2\pi(\lambda_k - \lambda_k^*))^6}\right) \delta_n^2. \tag{10.73}$$

10.5. Convergence of the scheme

For any $n \geqslant 1$, we set

$$\lambda_n - \lambda_n^* = \lambda_n^* - \lambda_{n+1} = \mu_n - \mu_n^* = \mu_n^* - \mu_{n+1} = \frac{\Lambda}{n^2} \tag{10.74}$$

for some $\Lambda > 0$. By choosing Λ small enough, we can make sure that the conditions

$$2\pi(\lambda_k - \lambda_k^*) < 1 \quad \text{and} \quad 2\pi(\mu_k - \mu_k^*) < 1$$

are satisfied for all k, as well as the other smallness assumptions made throughout this section. Moreover, we have $\lambda_k - \lambda_k^* \geqslant \Lambda/k^2$, so conditions ($\mathbf{C_1}$)–($\mathbf{C_9}$) will be satisfied if

$$\sum_{k=1}^{n} k^{12}\delta_k \leqslant \Lambda^6\omega \quad \text{and} \quad \sum_{j=k+1}^{n} j^6\delta_j \leqslant \Lambda^3\omega\left(\frac{1}{k^2}-\frac{1}{n^2}\right)$$

for some small explicit constant $\omega>0$, depending on the other constants appearing in the problem. Both conditions are satisfied if

$$\sum_{k=1}^{\infty} k^{12}\delta_k \leqslant \Lambda^6\omega. \tag{10.75}$$

Then from (10.60) we have that $T_n \leqslant C_\gamma(n^2/\Lambda)^{(7+\gamma)/(\gamma-1)}$, so the induction relation on δ_n gives

$$\delta_1 \leqslant C\delta \quad \text{and} \quad \delta_{n+1} = C\left(\frac{n^2}{\Lambda}\right)^9 e^{C(n^2/\Lambda)^{(14+2\gamma)/(\gamma-1)}}\delta_n^2. \tag{10.76}$$

To establish this relation we also assumed that δ_n is bounded below by $C_F\zeta_n$, the error coming from the short-time iteration; but this follows easily by construction, since the constraints imposed on δ_n are much worse than those on ζ_n.

Having fixed Λ, we will check that for δ small enough, (10.76) implies both the fast convergence of $(\delta_k)_{k=1}^{\infty}$, and the condition (10.75), which will justify a posteriori the derivation of (10.76). (An easy induction is enough to turn this into a rigorous reasoning.)

For this we fix $a\in(1,a_I)$, $0<z<z_I<1$, and we check by induction that

$$\delta_n \leqslant \Delta z^{a^n} \quad \text{for all } n \geqslant 1. \tag{10.77}$$

If Δ is given, (10.77) holds for $n=1$ as soon as $\delta \leqslant (\Delta/C)z^a$. Then, to go from stage n to stage $n+1$, we should check that

$$\frac{Cn^{18}}{\Lambda^9} e^{Cn^{(28+4\gamma)/(\gamma-1)}/\Lambda^{(14+2\gamma)/(\gamma-1)}} \Delta^2 z^{2a^n} \leqslant \Delta z^{a^{n+1}};$$

this is true if

$$\frac{1}{\Delta} \geqslant \frac{C}{\Lambda^9} \sup_{n\in\mathbb{N}} n^{19} e^{Cn^{(28+4\gamma)/(\gamma-1)}/\Lambda^{(14+2\gamma)/(\gamma-1)}} z^{(2-a)a^n}.$$

Since $a<2$, the supremum on the right-hand side is finite, and we just have to choose Δ small enough. Then, reducing Δ further if necessary, we can ensure (10.75). This concludes the proof.

Remark 10.1. This argument almost fully exploits the bi-exponential convergence of the Netwon scheme: a convergence like, say, $O(e^{-n^{1000}})$ would not be enough to treat values of γ which are close to 1. In §11.2 we shall present a more cumbersome approach which is less greedy in the convergence rate, but still needs convergence like $O(e^{-n^\alpha})$ for α large enough.

11. Coulomb–Newton interaction

In this section we modify the scheme of §10 to treat the case $\gamma=1$. We provide two different strategies. The first one is quite simple and will only come close to treat this case, since it will hold on (nearly) *exponentially large times* in the inverse of the perturbation size. The second one, somewhat more involved, will hold up to infinite times.

11.1. Estimates on exponentially large times

In this subsection we adapt the estimates of §10 to the case $\gamma=1$, under the additional restriction that $0\leqslant t\leqslant A^{1/\delta(\log\delta)^2}$ for some constant $A>1$.

In the iterative scheme, the only place where we used $\gamma>1$ (and not just $\gamma\geqslant 1$) is in Step 6, when it comes to the echo response via Theorem 7.7. Now, in the case $\gamma=1$, the formula for K_1^n should be

$$K_1^n(t,\tau)=\sum_{k=1}^{n}\delta_k K_1^{(\alpha_{n,k}),1}(t,\tau),$$

with $\alpha_{n,k}=\pi\min\{\mu_k-\mu_n^*,\lambda_k-\lambda_n^*,2\eta\}$. By Theorem 7.7 (ii) this induces, in addition to other well-behaved factors, an uncontrolled exponential growth $O(e^{\varepsilon_n t})$, with

$$\varepsilon_n=\Gamma\sum_{k=1}^{n}\frac{\delta_k}{\alpha_{n,k}^3};$$

in particular ε_n will remain bounded and $O(\delta)$ throughout the scheme.

Let us replace (10.74) by

$$\lambda_n-\lambda_n^*=\lambda_n^*-\lambda_{n+1}=\mu_n-\mu_n^*=\mu_n^*-\mu_{n+1}=\frac{\Lambda}{n(\log(e+n))^2},$$

where $\Lambda>0$ is very small. (This is allowed since the series

$$\sum_{n=1}^{\infty}\frac{1}{n(\log(e+n))^2}$$

converges—the power 2 could of course be replaced by any $r>1$.) Then during the first stages of the iteration we can absorb the $O(e^{\varepsilon_n t})$ factor by the loss of regularity if, say,

$$\varepsilon_n\leqslant\frac{\Lambda}{2n(\log(e+n))^2}.$$

Recalling that $\varepsilon_n=O(\delta)$, this is satisfied as soon as

$$n\leqslant N:=\frac{K}{\delta(\log(1/\delta))^2},\tag{11.1}$$

where $K>0$ is a positive constant depending on the other parameters of the problem but of course not on δ. So during these first stages we get the same long-time estimates as in §10.

For $n>N$ we cannot rely on the loss of regularity any longer; at this stage the error is about

$$\delta_N \leqslant C\delta^{a^N},$$

where $1<a<2$. To get the bounds for larger values of n, we impose a restriction on the time-interval, say $0\leqslant t\leqslant T_{\max}$. Allowing for a degradation of the rate δ^{a^n} into $\delta^{\underline{a}^n}$ with $\underline{a}<a$, we see that the new factor $e^{\varepsilon_n T_{\max}}$ can be eaten up by the scheme if

$$e^{\varepsilon_n T_{\max}}\delta^{(a-\underline{a})\underline{a}^n} \leqslant 1 \quad \text{for all } n \geqslant N.$$

This is satisfied if

$$T_{\max} = O\left(\underline{a}^N \frac{1}{\delta}\log\frac{1}{\delta}\right).$$

Recalling (11.1), we see that the latter condition holds true if

$$T_{\max} = O\left(A^{1/\delta(\log\delta)^2}\frac{1}{\delta}\log\frac{1}{\delta}\right)$$

for some well-chosen constant $A>1$. Then we can complete the iteration, and end up with a bound like

$$\|f_t - f_i\|_{\mathcal{Z}_t^{\lambda',\mu'}} \leqslant C\delta \quad \text{for all } t\in[0,T_{\max}],$$

where C is another constant independent of δ. The conclusion follows easily.

11.2. Mode-by-mode estimates

Now we shall change the estimates of §10 a bit more in depth to treat arbitrarily large times for $\gamma=1$. The main idea is to work *mode-by-mode* in the estimate of the spatial density, instead of looking directly for norm estimates.

Steps 1–5 remain the same, and the changes mainly occur in Step 6.

Substep 6 (a) is unchanged, but we only retain from that substep

$$e^{2\pi(\lambda_n^* t+\mu_n^*)|l|}\left|\widehat{\sigma_t^{n,n}}(l)\right| \leqslant \frac{2C_F\delta_n^2}{(\pi(\lambda_n-\lambda_n^*))^2} \quad \text{for all } l\in\mathbb{Z}^d. \tag{11.2}$$

Substep (b) is more deeply changed. Let $\hat{\mu}_n<\mu_n^*$.

• First, for each $l\in\mathbb{Z}^d$, we have, by Propositions 4.34 and 4.35, and the last part of Proposition 6.2,

$$
\begin{aligned}
&e^{2\pi(\lambda_n^* t+\hat\mu_n)|l|}|\widehat{\mathcal{E}}(t,l)|\\
&\quad\leqslant\int_0^t\sum_{m\in\mathbb{Z}^d}\|P_m(F[h_\tau^{n+1}]\circ\Omega_{t,\tau}^n-F[h_\tau^{n+1}])\|_{\mathcal{Z}^{\lambda_n^*(1+b),\hat\mu_n}_{\tau-bt/(1+b)}}\\
&\qquad\times\|P_{l-m}G_{\tau,t}^n\|_{\mathcal{Z}^{\lambda_n^*(1+b),\hat\mu_n;1}_{\tau-bt/(1+b)}}\,d\tau\\
&\quad\leqslant\int_0^t\sum_{m,m'\in\mathbb{Z}^d}\left\|P_{m-m'}\int_0^1\nabla F[h_\tau^{n+1}]\circ(\mathrm{Id}+\theta(\Omega_{t,\tau}^n-\mathrm{Id}))\,d\theta\right\|_{\mathcal{Z}^{\lambda_n^*(1+b),\hat\mu_n}_{\tau-bt/(1+b)}}\\
&\qquad\times\|P_{m'}(\Omega_{t,\tau}^n-\mathrm{Id})\|_{\mathcal{Z}^{\lambda_n^*(1+b),\hat\mu_n}_{\tau-bt/(1+b)}}\|P_{l-m}G_{\tau,t}^n\|_{\mathcal{Z}^{\lambda_n^*(1+b),\hat\mu_n;1}_{\tau-bt/(1+b)}}\,d\tau\\
&\quad\leqslant\int_0^1\int_0^t\|G_{\tau,t}^n\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*;1}_{\tau-bt/(1+b)}}\left\|\Omega_{t,\tau}^n-\mathrm{Id}\right\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{\tau-bt/(1+b)}}\sum_{m,m'\in\mathbb{Z}^d}e^{-2\pi(\mu_n^*-\hat\mu_n)|l-m|}\\
&\qquad\times e^{-2\pi(\mu_n^*-\hat\mu_n)|m'|}\|P_{m-m'}(\nabla F[h_\tau^{n+1}]\circ(\mathrm{Id}+\theta(\Omega_{t,\tau}^n-\mathrm{Id})))\|_{\mathcal{Z}^{\lambda_n^*(1+b),\hat\mu_n}_{\tau-bt/(1+b)}}\,d\tau\,d\theta\\
&\quad\leqslant\int_0^t|G^n\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*;1}_{\tau-bt/(1+b)}}\|\Omega_{t,\tau}^n-\mathrm{Id}\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{\tau-bt/(1+b)}}\\
&\qquad\times\sum_{m,m',q\in\mathbb{Z}^d}e^{-2\pi(\mu_n^*-\hat\mu_n)|l-m|}e^{-2\pi(\mu_n^*-\hat\mu_n)|m'|}\\
&\qquad\times e^{-2\pi(\mu_n^*-\hat\mu_n)|m-m'-q|}\|P_q(\nabla F[h_\tau^{n+1}])\|_{\mathcal{F}^{\hat\nu_n}}\,d\tau,
\end{aligned}
$$

where

$$\hat\nu_n=\hat\mu_n+\lambda_n^*(1+b)\left|\tau-\frac{bt}{1+b}\right|+\|\Omega^n X_{t,\tau}-x\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*}_{\tau-bt/(1+b)}}.$$

For $\alpha\leqslant1$ we have

$$\sum_{m,m'\in\mathbb{Z}^d}e^{-2\pi\alpha|l-m|}e^{-2\pi\alpha|m'|}e^{-2\pi\alpha|m-m'-q|}\leqslant\frac{C(d)}{\alpha^d}e^{-\pi\alpha|l-q|},$$

and we can argue as in Substep 6 (b) of §10 to get

$$
\begin{aligned}
&e^{2\pi(\lambda_n^* t+\mu_n^*)|l|}|\widehat{\mathcal{E}}(t,l)|\\
&\quad\leqslant\frac{C}{(\mu_n^*-\hat\mu_n)^d}\left(C_0'+\sum_{k=1}^n\delta_k\right)\left(\sum_{k=1}^n\frac{\delta_k}{(2\pi(\lambda_k-\lambda_k^*))^5}\right)\\
&\qquad\times\sum_{q\in\mathbb{Z}^d}e^{-\pi(\mu_n^*-\hat\mu_n)|l-q|}\int_0^t e^{2\pi(\lambda_n^*\tau+\hat\mu_n)|q|}\big|\widehat{\varrho[h_\tau^{n+1}]}(q)\big|\frac{d\tau}{(1+\tau)^2}.
\end{aligned}
\tag{11.3}
$$

- Next, we use again Proposition 4.34 and simple estimates to bound $\overline{\mathcal{E}}$,

$$e^{2\pi(\lambda_n^*t+\hat\mu_n)|l|}\big|\widehat{\overline{\mathcal{E}}}(t,l)\big| \leqslant \int_0^t \|G^n-\overline{G}^n\|_{\mathcal{Z}^{\lambda_n^*(1+b),\mu_n^*;1}_{\tau-bt/(1+b)}} \times \sum_{m\in\mathbb{Z}^d} e^{-2\pi(\mu_n^*-\hat\mu_n)|m|}\|P_{l-m}(F[h_\tau^{n+1}])\|_{\mathcal{F}^{\hat\beta_n}}\,d\tau,$$

where

$$\hat\beta_n = \lambda_n^*(1+b)\left|\tau-\frac{bt}{1+b}\right|+\hat\mu_n.$$

Reasoning as in Substep 6 (b) of §10, we arrive at

$$\begin{aligned} &e^{2\pi(\lambda_n^*t+\hat\mu_n)|l|}\big|\widehat{\overline{\mathcal{E}}}(t,l)\big| \\ &\quad\leqslant C\bigg(C_0'+\sum_{k=1}^n\delta_k\bigg)\bigg(\sum_{j=1}^n\frac{\delta_j}{(2\pi(\lambda_j-\lambda_n^*))^6}+\sum_{k=1}^n\delta_k\bigg) \\ &\qquad\times\sum_{m\in\mathbb{Z}^d} e^{-2\pi(\mu_n^*-\hat\mu_n)|m|}\int_0^t e^{2\pi(\lambda_n^*\tau+\hat\mu_n)|l-m|}\big|\widehat{\varrho[h_\tau^{n+1}]}(l-m)\big|\frac{d\tau}{(1+\tau)^2}. \end{aligned} \tag{11.4}$$

- Then we consider the "main contribution" $\overline{\sigma}^{n,n+1}$, which we decompose as in §10:

$$\overline{\sigma}_t^{n,n+1} = \overline{\sigma}_{t,0}^{n,n+1}+\sum_{k=1}^n\overline{\sigma}_{t,k}^{n,n+1},$$

and we write, for $k\geqslant 1$,

$$\begin{aligned} e^{2\pi(\lambda_n^*t+\hat\mu_n)}\big|\widehat{\overline{\sigma}_{t,k}^{n,n+1}}(l)\big| &\leqslant \sum_{m\in\mathbb{Z}^d}\int_0^t K_{l,m}^{n,h^k}(t,\tau)\|P_{l-m}(F[h_\tau^{n+1}])\|_{\mathcal{F}^{\nu_n',\gamma}}\,d\tau \\ &\quad+\int_0^t K_0^{n,h^k}(t,\tau)\|P_l(F[h_\tau^{n+1}])\|_{\mathcal{F}^{\nu_n',\gamma}}\,d\tau, \end{aligned}$$

where

$$\begin{aligned} \nu_n' &= \lambda_n^*(1+b)\left|\tau-\frac{bt}{1+b}\right|+\mu_n', \\ K_{l,m}^{n,h^k}(t,\tau) &= K_{l,m}^{n,k}(t,\tau)\sup_{0\leqslant s\leqslant t}\frac{\|\nabla_v(h_s^k\circ\Omega_{t,s}^{k-1})-\langle\nabla_v(h_s^k\circ\Omega_{t,s}^{k-1})\rangle\|_{\mathcal{Z}^{\lambda_k(1+b),\mu_k}_{s-bt/(1+b)}}}{1+s}, \\ K_{l,m}^{n,k}(t,\tau) &= (1+\tau)e^{-\pi(\mu_k-\hat\mu_n)|m|}\frac{e^{-2\pi(\mu_n'-\hat\mu_n)|l-m|}}{1+|l-m|^\gamma}e^{-\pi(\lambda_k-\lambda_n^*)|l(t-\tau)+m\tau|}, \end{aligned}$$

and the formula for K_0^{n,h^k} is unchanged with respect to §10.

Assuming that $\mu_n'=\hat\mu_n+\eta(t-\tau)/(1+t)$ and reasoning as in Substep 6 (b) of §10, we end up with the following estimate on the "main term":

$$\begin{aligned}
&e^{2\pi(\lambda_n^*t+\hat\mu_n)|l|}\big|\widehat{\bar\sigma_{t,k}^{n,n+1}}(l)\big|\\
&\leqslant C\sum_{m\in\mathbb{Z}^d}\int_0^t\bigg(\sum_{k=1}^n\delta_k K_{l,m}^{(\alpha_{n,k}),\gamma}(t,\tau)\bigg)e^{2\pi(\lambda_n^*\tau+\hat\mu_n)|l-m|}\big|\widehat{\varrho[h_\tau^{n+1}]}(l-m)\big|\,d\tau\\
&\quad+C\sum_{m\in\mathbb{Z}^d}\int_0^t\bigg(\sum_{k=1}^n\delta_k e^{-\pi(\lambda_k-\lambda_n^*)(t-\tau)}\bigg)e^{2\pi(\lambda_n^*\tau+\hat\mu_n)|l|}\big|\widehat{\varrho[h_\tau^{n+1}]}(l-m)\big|\,d\tau.
\end{aligned}\tag{11.5}$$

Then Substep 6 (c) becomes, with $\Phi(l,\tau)=\widehat{\varrho[h_\tau^{n+1}]}(l)$,

$$\begin{aligned}
&e^{2\pi(\lambda_n^*t+\hat\mu_n)|l|}\bigg|\Phi(l,t)-\int_0^t K^0(l,t-\tau)\Phi(l,\tau)\,d\tau\bigg|\\
&\quad\leqslant\frac{C\delta_n^2}{(\lambda_n-\lambda_n^*)^2}+\sum_{m\in\mathbb{Z}^d}\int_0^t\bigg(K_{l,m}^n(t,\tau)+\frac{c_m^n}{(1+\tau)^2}\bigg)e^{2\pi(\lambda_n^*\tau+\hat\mu_n)|l-m|}|\Phi(l-m,\tau)|\,d\tau\\
&\qquad+\int_0^t K_0^n(t,\tau)e^{2\pi(\lambda_n^*\tau+\hat\mu_n)|l|}|\Phi(l,\tau)|\,d\tau,
\end{aligned}$$

with

$$\begin{aligned}
K_{l,m}^n(t,\tau)&=C\sum_{k=1}^n\delta_k K_{l,m}^{(\widehat\alpha_n),\gamma}(t,\tau),\\
\widehat\alpha_n&=\pi\min\{\mu_n-\hat\mu_n,\lambda_n-\lambda_n^*,2\eta\},\\
K_0^n(t,\tau)&=C\sum_{k=1}^n\delta_k e^{-\pi(\lambda_k-\lambda_n^*)(t-\tau)},\\
c_m^n&=\frac{C}{(\mu_n^*-\hat\mu_n)^d}\bigg(C_0'+\sum_{k=1}^n\delta_k\bigg)\bigg(1+\sum_{k=1}^n\frac{\delta_k}{(2\pi(\lambda_k-\lambda_k^*))^6}\bigg)e^{-\pi(\mu_n^*-\hat\mu_n)|m|}.
\end{aligned}$$

Note that

$$\sum_{m\in\mathbb{Z}^d}c_m^n+\bigg(\sum_{m\in\mathbb{Z}^d}(c_m^n)^2\bigg)^{1/2}\leqslant\frac{C}{(\mu_n^*-\hat\mu_n)^{2d}}\bigg(C_0'+\sum_{k=1}^n\delta_k\bigg)\bigg(1+\sum_{k=1}^n\frac{\delta_k}{(2\pi(\lambda_k-\lambda_k^*))^6}\bigg).$$

Then, being $\Phi(0,t)=0$, we can apply Theorem 7.12 and deduce (taking already into account, for the sake of readability of the formula, that $\sum_{k=1}^n\delta_k$ and $\sum_{k=1}^n\delta_k/(\lambda_k-\lambda_k^*)$ are uniformly bounded)

$$e^{2\pi(\lambda_n^*t+\hat\mu_n)|l|}\big|\widehat{\varrho_t^{n+1}}(l)\big|\leqslant C\frac{\delta_n^2}{(\lambda_n-\lambda_n^*)^2\widehat\alpha_n\varepsilon^{3/2}(\mu_n^*-\hat\mu_n)^{2d}}\exp\bigg(\frac{C(1+\widehat T_{\varepsilon,n}^2)}{(\mu_n^*-\hat\mu_n)^{2d}}\bigg)e^{\varepsilon t},$$

where

$$\widehat{T}_{\varepsilon,n} = C\max\left\{\left(\frac{1}{\alpha_n^{3+2d}\varepsilon^{\gamma+2}}\right)^{1/\gamma}, \left(\frac{1}{\alpha_n^{d}\varepsilon^{\gamma+1/2}}\right)^{1/(\gamma-1/2)}, \left(\frac{1}{\varepsilon}\Big(\sum_{m\in\mathbb{Z}^d} c_m^n\Big)^2\right)^{1/3}\right\}.$$

If $\lambda_n^\dagger < \lambda_n^*$ and $\mu_n^\dagger < \hat{\mu}_n$ are chosen as before and $\varepsilon = 2\pi(\lambda_n^* - \lambda^\dagger)$, this implies a uniform bound on

$$\|\varrho_t^{n+1}\|_{\mathcal{F}^{\lambda_n^\dagger t+\mu_n^\dagger}} \leqslant \frac{C}{(\hat{\mu}_n-\mu_n^\dagger)^d} \sup_{l\in\mathbb{Z}^d} e^{2\pi(\lambda_n^* t+\hat{\mu}_n)|l|}\big|\widehat{\varrho_t^{n+1}}(l)\big|$$

obtained with the formula above with

$$\widehat{T}_{\varepsilon,n} = \widehat{T}_n = C\max\left\{\frac{1}{\lambda_n^*-\lambda_n^\dagger}, \frac{1}{\lambda_n-\lambda_n^*}, \frac{1}{\mu_n-\mu_n^*}, \frac{1}{\mu_n^*-\hat{\mu}_n}\right\}^a,$$

where

$$a = \max\left\{\frac{5+\gamma+2d}{\gamma}, \frac{d+\gamma+\frac{1}{2}}{\gamma-\frac{1}{2}}, \frac{4d+1}{3}\right\}.$$

Then Steps 7–10 of the iteration can be repeated with the only modification that μ_n^* is replaced by $\mu_n^\dagger$.

The convergence (§10.5) works just the same, except that now we need more intermediate regularity indices μ_n:

$$\mu_{n+1} = \mu_n^\ddagger < \mu_n^\dagger < \hat{\mu}_n < \mu_n^*;$$

the obvious choice being to let $\mu_n^\dagger - \mu_n^\ddagger = \hat{\mu}_n - \mu_n^\dagger = \mu_n^* - \hat{\mu}_n$.

Choosing $\lambda_n - \lambda_{n+1}$ and $\mu_n - \mu_{n+1}$ of the order of Λ/n^2, we arrive in the end at the induction

$$\delta_{n+1} \leqslant C\left(\frac{n^2}{\Lambda}\right)^{9+6d} e^{C(n^2/\Lambda)^{\xi(d,\gamma)}}\delta_n^2,$$

where

$$\xi(d,\gamma) := 2d+2\max\left\{\frac{5+\gamma+2d}{\gamma}, \frac{d+\gamma+\frac{1}{2}}{\gamma-\frac{1}{2}}, \frac{4d+1}{3}\right\}. \tag{11.6}$$

Then the convergence of the scheme (and a-posteriori justification of all the assumptions) is done exactly as in §10.

12. Convergence in large time

In this section we prove Theorem 2.6 as a simple consequence of the uniform bounds established in §10 and §11.

So let f^0, L and W satisfy the assumptions of Theorem 2.6. To simplify the notation we assume that $L=1$.

The second part of Assumption (2.14) precisely means that $f^0 \in \mathcal{C}^{\lambda;1}$. We shall actually assume a slightly more precise condition, namely that for some $\bar{p} \in [1,\infty]$,

$$\sum_{n \in \mathbb{N}_0^d} \frac{\lambda^n}{n!} \|\nabla_v^n f^0\|_{L^p(\mathbb{R}^d)} \leqslant C_0 < \infty \quad \text{for all } p \in [1,\bar{p}]. \tag{12.1}$$

(It is sufficient to take $\bar{p}=1$ to get Theorem 2.6; but if this bound is available for some $\bar{p}>1$ then it will be propagated by the iteration scheme, and will result in more precise bounds.) Then we pick up $\underline{\lambda} \in (0,\lambda)$, $\underline{\mu} \in (0,\mu)$, $\beta>0$ and $\beta' \in (0,\beta)$. By symmetry, we only consider non-negative times.

If f_i is an initial datum satisfying the smallness condition (2.15), then by Theorem 4.20, we have a smallness estimate on $\|f_i - f^0\|_{\mathcal{Z}^{\lambda',\mu';p}}$ for all $p \in [1,\bar{p}]$, $\lambda'<\lambda$ and $\mu'<\mu$. Then, as in §4.12 we can estimate the solution h^1 to the linearized equation

$$\begin{cases} \partial_t h^1 + v \cdot \nabla_x h^1 + F[h^1] \cdot \nabla_v f^0 = 0, \\ h^1(0,\cdot) = f_i - f^0, \end{cases} \tag{12.2}$$

and we recover uniform bounds in $\mathcal{Z}^{\hat{\lambda},\hat{\mu};p}$ spaces, for any $\hat{\lambda} \in (\underline{\lambda},\lambda)$ and $\hat{\mu} \in (\underline{\mu},\mu)$. More precisely,

$$\sup_{t \geqslant 0} \|\varrho[h_t^1]\|_{\mathcal{F}^{\hat{\lambda}t+\hat{\mu}}} + \sup_{t \geqslant 0} \|h^1(t,\cdot)\|_{\mathcal{Z}_t^{\hat{\lambda},\hat{\mu};p}} \leqslant C\delta, \tag{12.3}$$

with $C = C(d,\lambda',\hat{\lambda},\mu',\hat{\mu},W,f^0)$ (this is of course assuming ε in Theorem 2.6 to be small enough).

We now set $\lambda_1 = \lambda'$, and we run the iterative scheme of §§9–11 for all $n \geqslant 2$. If ε is small enough, up to slightly lowering λ_1, we may choose all parameters in such a way that

$$\lambda_k, \lambda_k^* \to \lambda_\infty > \underline{\lambda} \text{ and } \mu_k, \mu_k^* \to \mu_\infty > \underline{\mu} \quad \text{as } k \to \infty;$$

then we pick up $B>0$ such that

$$\mu_\infty - \lambda_\infty (1+B) B \geqslant \mu'_\infty > \underline{\mu},$$

and we let $b(t) = B/(1+t)$.

As a result of the scheme, we have, for all $k \geqslant 2$,

$$\sup_{0 \leqslant \tau \leqslant t} \|h_\tau^k \circ \Omega_{t,\tau}^{k-1}\|_{\mathcal{Z}^{\lambda_\infty(1+b),\mu_\infty;1}_{\tau - bt/(1+b)}} \leqslant \delta_k, \tag{12.4}$$

where $\sum_{k=2}^\infty \delta_k \leqslant C\delta$ and Ω^k is the deflection associated with the force field generated by $h^1 + \ldots + h^k$. Choosing $t=\tau$ in (12.4) yields

$$\sup_{t \geqslant 0} \|h_t^k\|_{\mathcal{Z}^{\lambda_\infty(1+B),\mu_\infty;1}_{t - Bt/(1+B+t)}} \leqslant \delta_k.$$

By Proposition 4.17, this implies that

$$\sup_{t\geqslant 0}\|h_t^k\|_{\mathcal{Z}_t^{\lambda_\infty(1+B),\mu_\infty-\lambda_\infty(1+B)B;1}}\leqslant\delta_k.$$

In particular, we have a uniform estimate on h_t^k in $\mathcal{Z}_t^{\lambda_\infty,\mu'_\infty;1}$. Summing up over k yields for $f=f^0+\sum_{k=1}^\infty h^k$ the estimate

$$\sup_{t\geqslant 0}\|f(t,\cdot)-f^0\|_{\mathcal{Z}_t^{\lambda_\infty,\mu'_\infty;1}}\leqslant C\delta. \tag{12.5}$$

Passing to the limit in the Newton scheme, one shows that f solves the non-linear Vlasov equation with initial datum f_i. (Once again we do not check the details; to be rigorous one would need to establish moment estimates, locally in time, before passing to the limit.) This implies in particular that f stays non-negative at all times.

Applying Theorem 4.20 again, we deduce from (12.5) that

$$\sup_{t\geqslant 0}\|f(t,\cdot)-f^0\|_{\mathcal{Y}_t^{\underline{\lambda},\underline{\mu}}}\leqslant C\delta;$$

or equivalently, with the notation used in Theorem 2.6,

$$\sup_{t\geqslant 0}\|f(t,x-vt,v)-f^0(v)\|_{\underline{\lambda},\underline{\mu}}\leqslant C\delta. \tag{12.6}$$

Moreover, $\varrho=\int_{\mathbb{R}^d} f\,dv$ satisfies similarly

$$\sup_{t\geqslant 0}\|\varrho(t,\cdot)\|_{\mathcal{F}^{\lambda_\infty t+\mu_\infty}}\leqslant C\delta.$$

It follows that $|\hat\varrho(t,k)|\leqslant C\delta e^{-2\pi\lambda_\infty|k|t}e^{-2\pi\mu_\infty|k|}$ for any $k\neq 0$. On the one hand, by Sobolev embedding, we deduce that for any $r\in\mathbb{N}$,

$$\|\varrho(t,\cdot)-\langle\varrho\rangle\|_{C^r(\mathbb{T}^d)}\leqslant C_r\delta e^{-2\pi\lambda' t};$$

on the other hand, multiplying $\hat\varrho$ by the Fourier transform of ∇W, we see that the force $F=F[f]$ satisfies

$$|\widehat{F}(t,k)|\leqslant C\delta e^{-2\pi\lambda'|k|t}e^{-2\pi\mu'|k|}\quad\text{for all } t\geqslant 0 \text{ and all } k\in\mathbb{Z}^d, \tag{12.7}$$

for some $\lambda'>\underline{\lambda}$ and $\mu'>\underline{\mu}$.

Now, from (12.6) we have, for any $(k,\eta)\in\mathbb{Z}^d\times\mathbb{R}^d$ and any $t\geqslant 0$,

$$|\tilde f(t,k,\eta+kt)-\tilde f^0(\eta)|\leqslant C\delta e^{-2\pi\mu'|k|}e^{-2\pi\lambda'|\eta|}; \tag{12.8}$$

so

$$|\tilde{f}(t,k,\eta)| \leqslant |\tilde{f}^0(\eta+kt)| + C\delta e^{-2\pi\mu'|k|} e^{-2\pi\lambda'|\eta+kt|}. \tag{12.9}$$

In particular, for any $k \neq 0$, and any $\eta \in \mathbb{R}^d$,

$$\tilde{f}(t,k,\eta) = O(e^{-2\pi\lambda' t}). \tag{12.10}$$

Thus f is asymptotically close (in the weak topology) to its spatial average

$$g = \langle f \rangle = \int_{\mathbb{T}^d} f \, dx.$$

Taking $k=0$ in (12.8) shows that, for any $\eta \in \mathbb{R}^d$,

$$|\tilde{g}(t,\eta) - \tilde{f}^0(\eta)| \leqslant C\delta e^{-2\pi\lambda'|\eta|}. \tag{12.11}$$

Also, from the non-linear Vlasov equation, for any $\eta \in \mathbb{R}^d$ we have

$$\begin{aligned} \tilde{g}(t,\eta) &= \tilde{f}_i(0,\eta) - \int_0^t \int_{\mathbb{T}^d} \int_{\mathbb{R}^d} F(\tau,x) \cdot \nabla_v f(\tau,x,v) e^{-2i\pi\eta\cdot v} \, dv \, dx \, d\tau \\ &= \tilde{f}_i(0,\eta) - 2i\pi \sum_{l\in\mathbb{Z}^d} \int_0^t \widehat{F}(\tau,l) \cdot \eta \tilde{f}(\tau,-l,\eta) \, d\tau. \end{aligned}$$

Using the bounds (12.7) and (12.10), it is easily shown that the above time-integral converges exponentially fast as $t \to \infty$, with rate $O(e^{-\lambda'' t})$ for any $\lambda'' < \lambda'$, to its limit

$$\tilde{g}_\infty(\eta) = \tilde{f}_i(0,\eta) - 2i\pi \sum_{l\in\mathbb{Z}^d} \int_0^\infty \widehat{F}(\tau,l) \cdot \eta \tilde{f}(\tau,-l,\eta) \, d\tau. \tag{12.12}$$

By passing to the limit in (12.11) we see that

$$|\tilde{g}_\infty(\eta) - \tilde{f}^0(\eta)| \leqslant C\delta e^{-2\pi\lambda'|\eta|},$$

and this concludes the proof of Theorem 2.6.

13. Non-analytic perturbations

Although the vast majority of studies of Landau damping assume that the perturbation is analytic, it is natural to ask whether this condition can be relaxed. As we noticed in Remark 3.5, this is the case for the linear problem. As for non-linear Landau damping, once the analogy with KAM theory has been identified, it is anybody's guess that the

answer might come from a Moser-type argument. However, this is not so simple, because the "loss of convergence" in our argument is much more severe than the "loss of regularity" which Moser's scheme allows one to overcome.

For instance, the second-order correction h^2 satisfies

$$\partial_t h^2 + v\cdot\nabla_x h^2 + F[f^1]\cdot\nabla_v h^2 + F[h^2]\cdot\nabla_v f^1 = -F[h^1]\cdot\nabla_v h^1.$$

The action of $F[f^1]$ is to curve trajectories, which does not help in our estimates. Discarding this effect and solving by Duhamel's formula and Fourier transform, we obtain, with $S=-F[h^1]\cdot\nabla_v h^1$ and $\varrho^2=\int_{\mathbb{R}^d} h^2\,dv$,

$$\begin{aligned}\hat\varrho^2(t,k) \simeq \int_0^t K^0(t-\tau,k)\hat\varrho^2(\tau,k)\,d\tau \\ +2i\pi\int_0^t \sum_{l\in\mathbb{Z}^d} (k-l)\widehat{W}(k-l)\hat\varrho^2(\tau,k-l)\widetilde{\nabla_v h^1}(\tau,l,k(t-\tau))\,d\tau \\ +\int_0^t \widetilde{S}(\tau,k,k(t-\tau))\,d\tau.\end{aligned} \tag{13.1}$$

(The term with K^0 includes the contribution of $\nabla_v f^0$.)

Our regularity/decay estimates on h^1 will never be better than those on the solution of the free transport equation, i.e., $h_i(x-vt,v)$, where $h_i=f_i-f^0$. Let us forget about the effect of K^0 in (13.1), replace the contribution of S by a decaying term $A(kt)$. Let us choose $d=1$ and assume that $\hat h_i(l,\cdot)=0$ if $l\neq\pm1$. For $k>0$, let us use the long-time approximation

$$\tilde h_i(-1,k(t-\tau)-\tau)\,1_{[0,t]}(\tau)\,d\tau \simeq \frac{c}{k+1}\delta_{kt/(k+1)}, \quad c=\int_{\mathbb{R}}\tilde h_i(-1,s)\,ds=\hat h_i(-1,0).$$

Note that $c\neq0$ in general. Plugging all these simplifications into estimate (13.1) and choosing $\widehat{W}(k)=1/|k|^{1+\gamma}$ suggests the a-priori simpler equation

$$\varphi(t,k)=A(kt)+\frac{ckt}{(k+1)^{\gamma+2}}\varphi\biggl(\frac{kt}{k+1},k+1\biggr). \tag{13.2}$$

Replacing $\varphi(t,k)$ by $\varphi(t,k)/A(kt)$ reduces to $A=1$, and then we can solve this equation by power series as in §7.1.3, obtaining

$$\varphi(t,k)\simeq A(kt)e^{(ckt)^{1/(\gamma+2)}}. \tag{13.3}$$

With a polynomial deterioration of the rate, we could use a regularization argument, but the fractional exponential is much worse.

However, our bounds are still good enough to establish decay for Gevrey perturbations. Let us agree that a function $f=f(x,v)$ lies in the Gevrey class $\mathcal{G}^\nu$, $\nu\geqslant 1$, if

$$|\tilde{f}(k,\eta)| = O(\exp(-c|(k,\eta)|^{1/\nu})) \quad \text{for some } c>0;$$

in particular $\mathcal{G}^1$ means analytic. (An alternative convention would be to require the nth derivative to be $O((n!)^\nu)$.) As we shall explain, we can still get non-linear Landau damping if the initial datum f_i lies in $\mathcal{G}^\nu$ for ν close enough to 1. Although this is still quite demanding, it already shows that non-linear Landau damping is not tied to analyticity or quasi-analyticity, and holds for a large class of compactly supported perturbations.

THEOREM 13.1. *Let $\lambda>0$. Let $f^0=f^0(v)\geqslant 0$ be an analytic homogeneous profile such that*

$$\sum_{n\in\mathbb{N}_0^d} \frac{\lambda^n}{n!} \|\nabla_v^n f^0\|_{L^1(\mathbb{R}^d)} < \infty,$$

and let $W=W(x)$ satisfy $|\widehat{W}(k)|=O(1/|k|^2)$. Assume that condition **(L)** *from §2.2 holds. Let $\nu\in(1,1+\theta)$ with $\theta=1/\xi(d,\gamma)$, where ξ was defined in* (11.6). *Let $\beta>0$ and $\alpha<1/\nu$. Then there is $\varepsilon>0$ such that if*

$$\delta := \sup_{\substack{k\in\mathbb{Z}^d\\ \eta\in\mathbb{R}^d}} |(\tilde{f}_i-\tilde{f}^0)(k,\eta)| e^{\lambda|\eta|^{1/\nu}} e^{\lambda|k|^{1/\nu}} + \int_{\mathbb{T}^d}\int_{\mathbb{R}^d} |(f_i-f^0)(x,v)| e^{\beta|v|}\, dv\, dx \leqslant \varepsilon,$$

then as $t\to+\infty$ the solution $f=f(t,x,v)$ of the non-linear Vlasov equation on $\mathbb{T}^d\times\mathbb{R}^d$ with interaction potential W and initial datum f_i satisfies

$$|\tilde{f}(t,k,\eta)-\tilde{f}_\infty(\eta)| = O(\delta e^{-ct^\alpha}) \quad \textit{for all } (k,\eta)\in\mathbb{Z}^d\times\mathbb{R}^d$$

and

$$\|F(t,\cdot)\|_{C^r(\mathbb{T}^d)} = O(\delta e^{-ct^\alpha}) \quad \textit{for all } r\in\mathbb{N}$$

for some $c>0$ and some homogeneous Gevrey profile f_∞, where F stands for the self-consistent force.

Remark 13.2. In view of (13.3), one may hope that the result remains true for $\theta=2$. Proving this would require much more precise estimates, including among other things a qualitative improvement of the constants in Theorem 4.20 (recall Remark 4.23).

Remark 13.3. One could also relax the analyticity of f^0, but there is little incentive to do so.

Sketch of proof. We first decompose $h_i = f_i - f^0$, using truncation by a smooth partition of unity in Fourier space,

$$h_i = \sum_{n=0}^{\infty} \mathcal{F}^{-1}(\tilde{h}_i \chi_n) \equiv \sum_{n=0}^{\infty} h_i^n,$$

where $\mathcal{F}$ is the Fourier transform. Each bump function χ_n should be localized around the domain (in Fourier space)

$$D_n = \{(k,\eta) \in \mathbb{Z}^d \times \mathbb{R}^d : n^K \leqslant |(k,\eta)| \leqslant (n+1)^K\},$$

for some exponent $K > 1$; but at the same time $\mathcal{F}^{-1}(\chi_n)$ should be exponentially decreasing in v. To achieve this, we let

$$\chi_n = 1_{D_n} * \gamma \quad \text{and} \quad \gamma(\eta) = e^{-\pi|\eta|^2}.$$

Then $\mathcal{F}^{-1}(\chi_n) = \mathcal{F}^{-1}(1_{D_n})\,\gamma$ has Gaussian decay, independently of n; so there is a uniform bound on

$$\int_{\mathbb{T}^d} \int_{\mathbb{R}^d} |h_i^n(x,v)| e^{\beta|v|}\, dv\, dx$$

for some $\beta > 0$.

On the other hand, if $(k,\eta) \in D_n$ and $(k',\eta') \notin D_{n-1} \cup D_n \cup D_{n+1}$, then

$$|k-k'| + |\eta - \eta'| \geqslant c n^{K-1}$$

for some $c > 0$. From this one obtains, after simple computations,

$$|\chi_n(k,\eta)| \leqslant 1_{(n-1)^K \leqslant |(k,\eta)| \leqslant (n+2)^K} + C e^{-c n^{2(K-1)}} e^{-c(|k|^2 + |\eta|^2)}.$$

So (with constants C and c changing from line to line)

$$\begin{aligned} |\tilde{h}_i^n(k,\eta)| &\leqslant C e^{-\lambda|k|^{1/\nu}} e^{-\lambda|\eta|^{1/\nu}} 1_{(n-1)^K \leqslant |(k,\eta)| \leqslant (n+2)^K} + C e^{-c n^{2(K-1)}} e^{-c(|k|+|\eta|)} \\ &\leqslant C \max\{e^{-\lambda(n-1)^{K/\nu}/2}, e^{-c n^{2(K-1)}}\} e^{-\bar{\lambda}_n(|k|+|\eta|)}, \end{aligned}$$

where

$$\bar{\lambda}_n \sim \tfrac{1}{2} \lambda (n+2)^{-(1-1/\nu)K}.$$

If $K \geqslant 2$ then $2(K-1) > K/\nu$; so

$$\|h_i^n\|_{\mathcal{Y}^{\bar{\lambda}_n, \bar{\lambda}_n}} \leqslant C e^{-\lambda n^{K/\nu}/2}.$$

Then we may apply Theorem 4.20 to get a bound on h_i^n in the space $\mathcal{Z}^{\hat\lambda_n,\hat\lambda_n;1}$ with $\hat\lambda_n=\frac12\bar\lambda_n$, at the price of a constant $\exp(C(n+2)^{(1-1/\nu)K})$. Assuming $K\nu>(1-1/\nu)K$, i.e., $\nu<2$, we end up with

$$\|h_i^n\|_{\mathcal{Z}^{\hat\lambda_n,\hat\lambda_n;1}}=O\big(e^{-cn^{K/\nu}}\big),\quad \hat\lambda_n=\tfrac12\bar\lambda_n. \tag{13.4}$$

Then we run the iteration scheme of §8 with the following modifications: (1) instead of $h^n(0,\cdot)=0$, choose $h^n(0,\cdot)=h_i^n$, and (2) choose regularity indices $\lambda_n\sim\hat\lambda_n$ which tend to zero as n tends to infinity. This generates an additional error term of size $O\big(e^{-cn^{K/\nu}}\big)$, and imposes that $\lambda_n-\lambda_{n+1}$ be of order $n^{-[(1-1/\nu)K+1]}$. When we apply the bilinear estimates from §6, we can take $\bar\lambda-\lambda$ to be of the same order; so $\alpha=\alpha_n$ and $\varepsilon=\varepsilon_n$ can be chosen proportional to $n^{-[(1-1/\nu)K+1]}$. Then the large constants coming from the time-response will be, as in §11, of order $n^qe^{cn^r}$, with $q\in\mathbb{N}$ and $r=[(1-1/\nu)K+1]\xi$, and the scheme will still converge like $O(e^{-cn^s})$ for any $s<K/\nu$, provided that $K/\nu>r$, i.e.,

$$\nu-1+\frac{\nu}{K}<\frac{1}{\xi}.$$

The rest of the argument is similar to what we did in §§10–12. In the end the decay rate of any non-zero mode of the spatial density ϱ is controlled by

$$\sum_{n=0}^{\infty}e^{-cn^s}e^{-\lambda_nt}\leqslant\left(\sum_{n=0}^{\infty}e^{-cn^s}\right)\sup_{n\geqslant0}e^{-cn^s}e^{-cn^{-(1-1/\nu)}t}\leqslant Ce^{-ct^{s/K}},$$

and the result follows since s/K is arbitrarily close to $1/\nu$. □

Remark 13.4. An alternative approach to Gevrey regularity consists in rewriting the whole proof with the help of Gevrey norms such as

$$\|f\|_{\mathcal{C}^\lambda_\nu}=\sum_{n=0}^{\infty}\frac{\lambda^n\|f^{(n)}\|_\infty}{n!^\nu}\quad\text{and}\quad\|f\|_{\mathcal{F}^\lambda_\nu}=\sum_{k\in\mathbb{Z}}e^{2\pi\lambda|k|^{1/\nu}}|\hat f(k)|,$$

which satisfy the algebra property for any $\nu\geqslant1$. Then one can hybridize these norms, rewrite the time-response in this setting, estimate fractional exponential moments of the kernel, etc. The only part which does not seem to adapt to this strategy is §9 where the analyticity is crucially used for the local result.

Remark 13.5. In a more general C^r context, we do not know whether decay holds for the non-linear Vlasov–Poisson equation. Speculations about this issue can be found in [55] where it is shown that (unlike in the linearized case) one needs more than one derivative on the perturbation. As a first step in this direction, we mention that our methods imply a bound like $O(\delta/(1+t)^{r-\bar r})$ for times $t=O(1/\delta)$, where $\bar r$ is a constant

and $r>\bar{r}$, as soon as the initial perturbation has norm δ in a functional space $\mathcal{W}^r$ involving r derivatives in a certain sense. The reason why this is non-trivial is that the natural time scale for non-linear effects in the Vlasov–Poisson equation is not $O(1/\delta)$, but $O(1/\sqrt{\delta}\,)$, as predicted by O'Neil [75] and very well checked in numerical simulations [61].([17]) Let us sketch the argument in a few lines. Assume that (for some positive constants c and C)

$$h_i = \sum_{n=0}^{\infty} h_i^n, \quad \|h_i^n\|_{\mathcal{Z}^{\lambda_n,\lambda_n;1}} \leqslant \frac{C^n}{2^{rn}} \quad \text{and} \quad \lambda_n = \frac{cn}{2^n}. \tag{13.5}$$

Then we may run the Newton scheme again choosing $\alpha_n \sim cn/2^n$, $c_n = O(\delta 2^{-(r-r_1)n})$ and $\varepsilon_n = c'\delta$. Over a time-interval of length $O(1/\delta)$, Theorem 7.7 (ii) only yields a multiplicative constant $O(e^{c\delta t}/\alpha_n^9) = O(2^{10n})$. In the end, after Sobolev injection again, we recover a time-decay on the force F like

$$\delta \sum_{n=0}^{\infty} 2^{nr_2} 2^{-nr} e^{-\lambda_n t} \leqslant C\delta \sup_{n\geqslant 0} (2^{-n(r-r_3)} e^{-\lambda_n t}) \leqslant \frac{C\delta}{(1+t)^{r-r_4}},$$

as desired. Equation (13.5) means that h_i is of size $O(\delta)$ in a functional space $\mathcal{W}^r$ whose definition is close to the Littlewood–Paley characterization of a Sobolev space with r derivatives. In fact, if the conjecture formulated in Remark 4.23 holds true, then it can be shown that $\mathcal{W}^r$ contains all functions in the Sobolev space $W^{r+r_0,2}$ satisfying an adequate moment condition, for some constant r_0.

14. Expansions and counterexamples

A most important consequence of the proof of Theorem 2.6 is that the asymptotic behavior of the solution of the non-linear Vlasov equation can in principle be determined at arbitrary precision as the size of the perturbation tends to 0. Indeed, if we define $g_\infty^k(v)$ as the large-time limit of h^k (say in positive time), then $\|g^k\| = O(\delta_k)$, so $f^0 + g_\infty^1 + \ldots + g_\infty^n$ converges very fast to f_∞. In other words, to investigate the properties of the time-asymptotics of the system, we may freely exchange the limits $t\to\infty$ and $\delta\to 0$, perform expansions, etc. This at once puts on rigorous grounds many asymptotic expansions used by various authors—who so far implicitly postulated the possibility of this exchange.

With this in mind, let us estimate the first corrections to the linearized theory, in the regime of a very small perturbation and small interaction strength (which can be achieved by a proper scaling of physical quantities). We shall work in dimension $d=1$ and in a periodic box of length $L=1$.

([17]) Passing from $O(1/\sqrt{\delta}\,)$ to $O(1/\delta)$ is arguably an infinite-dimensional counterpart of Laplace's averaging principle, which yields stability for certain Hamiltonian systems over time-intervals $O(1/\delta^2)$ rather than $O(1/\delta)$.

14.1. Simple excitation

For a start, let us consider the case where the perturbation affects only the first spatial frequency. We let

- $f^0(v)=e^{-\pi v^2}$ be the homogeneous (Maxwellian) distribution;
- $\varepsilon\varrho_i(x)=\varepsilon\cos(2\pi x)$ be the initial space density perturbation;
- $\varepsilon\varrho_i(x)\theta(v)$ be the initial perturbation of the distribution function; we denote by φ the Fourier transform of θ;
- αW be the interaction potential, with $W(-x)=W(x)$. We do not specify its form, but it should satisfy the assumptions in Theorem 2.6.

We work in the asymptotic regime $\varepsilon\to0$ and $\alpha\to0$. The parameter ε measures the size of the perturbation, while α measures the strength of the interaction; after dimensional change, if W is an inverse power, α can be thought of as an inverse power of the ratio (Debye length)/(perturbation wavelength). We will not write norms explicitly, but all our computations can be made in the norms introduced in §4, with small losses in the regularity indices—as we have done throughout the paper.

The first-order correction $h^1=O(\varepsilon)$ to f^0 is provided by the solution of the linearized equation (3.3), here taking the form

$$\partial_t h^1+v\cdot\nabla_x h^1+F[h^1]\cdot\nabla_v f^0=0,$$

with initial datum $h^1(0,\cdot)=h_i:=f_i-f^0$. As in §3 we get a closed equation for the associated density $\varrho[h^1]$,

$$\hat\varrho[h^1](t,k)=\tilde h_i(k,kt)-4\pi^2\alpha\widehat{W}(k)\int_0^t\hat\varrho[h^1](\tau,k)e^{-\pi(k(t-\tau))^2}(t-\tau)k^2\,d\tau.$$

It follows that $\hat\varrho[h^1](t,k)=0$ for $k\neq\pm1$, so the behavior of $\hat\varrho[h^1]$ is entirely determined by $u_1(t)=\hat\varrho[h^1](t,1)$ and $u_{-1}(t)=\hat\varrho[h^1](t,-1)$, which satisfy

$$u_1(t)=\frac{\varepsilon}{2}\varphi(t)-4\pi^2\alpha\widehat{W}(1)\int_0^t u_1(\tau)e^{-\pi(t-\tau)^2}(t-\tau)\,d\tau=\frac{\varepsilon}{2}[\varphi(t)+O(\alpha)].\tag{14.1}$$

(This equation can be solved explicitly [11, equation (6)], but we only need the expansion.) Similarly,

$$u_{-1}(t)=\frac{\varepsilon}{2}\varphi(-t)-4\pi^2\alpha\widehat{W}(1)\int_0^t u_{-1}(\tau)e^{-\pi(t-\tau)^2}(t-\tau)\,d\tau=\frac{\varepsilon}{2}[\varphi(-t)+O(\alpha)].\tag{14.2}$$

The corresponding force, in Fourier transform, is given by

$$\widehat{F}^1(t,1)=-2i\pi\alpha\widehat{W}(1)u_1(t)\quad\text{and}\quad\widehat{F}^1(t,-1)=2i\pi\alpha\widehat{W}(1)u_{-1}(t).$$

From this we also deduce the Fourier transform of h^1 itself:

$$\tilde h^1(t,k,\eta)=\tilde h_i(k,\eta+kt)-4\pi^2\alpha\widehat{W}(k)\int_0^t \hat\varrho[h^1](\tau,k)e^{-\pi(\eta+k(t-\tau))^2}(\eta+k(t-\tau))\cdot k\,d\tau. \tag{14.3}$$

This is 0 if $k\neq\pm1$, while

$$\begin{aligned}\tilde h^1(t,1,\eta)&=\frac{\varepsilon}{2}\varphi(\eta+t)-4\pi^2\alpha\widehat{W}(1)\int_0^t u_1(\tau)e^{-\pi(\eta+(t-\tau))^2}(\eta+(t-\tau))\,d\tau\\&=\frac{\varepsilon}{2}[\varphi(\eta+t)+O(\alpha)]\end{aligned} \tag{14.4}$$

and

$$\begin{aligned}\tilde h^1(t,-1,\eta)&=\frac{\varepsilon}{2}\varphi(\eta-t)+4\pi^2\alpha\widehat{W}(1)\int_0^t u_{-1}(\tau)e^{-\pi(\eta-(t-\tau))^2}(\eta-(t-\tau))\,d\tau\\&=\frac{\varepsilon}{2}[\varphi(\eta-t)+O(\alpha)].\end{aligned} \tag{14.5}$$

To get the next order correction, we solve, as in §10,

$$\partial_t h^2+v\cdot\nabla_x h^2+F[h^1]\cdot\nabla_v h^2+F[h^2]\cdot(\nabla_v f^0+\nabla_v h^1)=-F[h^1]\cdot\nabla_v h^1,$$

with zero initial datum. Since $h^2=O(\varepsilon^2)$, we may neglect the terms $F[h^1]\cdot\nabla_v h^2$ and $F[h^2]\cdot\nabla_v h^1$ which are both $O(\alpha\varepsilon^3)$. So it is sufficient to solve

$$\partial_t h_2'+v\cdot\nabla_x h_2'+F[h_2']\cdot\nabla_v f^0=-F[h^1]\cdot\nabla_v h^1 \tag{14.6}$$

with vanishing initial datum. As $t\to\infty$, we know that the solution $h_2'(t,x,v)$ is asymptotically close to its spatial average $\langle h_2'\rangle=\int_{\mathbb{T}^d}h_2'\,dx$. Taking the integral over $\mathbb{T}^d$ in (14.6) yields

$$\partial_t\langle h_2'\rangle=-\langle F[h^1]\cdot\nabla_v h^1\rangle.$$

Since h^1 converges to $\langle h_i\rangle$, the deviation of f to $\langle f_i\rangle$ is given, at order ε^2, by

$$g(v)=-\int_0^\infty\langle F[h_1]\cdot\nabla_v h^1\rangle(t,v)\,dt=-\int_0^\infty\sum_{k\in\mathbb{Z}}\widehat{F}[h^1](t,-k)\cdot\nabla_v\hat h^1(t,k,v)\,dt.$$

Applying the Fourier transform and using (14.1), (14.2), (14.4) and (14.5), we deduce

$$\begin{aligned}\tilde g(\eta)&=-\int_0^\infty\sum_{k\in\mathbb{Z}}\widehat{F}[h^1](t,-k)\cdot\widetilde{\nabla_v h^1}(t,k,\eta)\,dt\\&=-\int_0^\infty\widehat{F}[h^1](t,-1)(2i\pi\eta)\tilde h^1(t,1,\eta)\,dt-\int_0^\infty\widehat{F}[h^1](t,1)(2i\pi\eta)\tilde h^1(t,-1,\eta)\,dt\\&=\pi^2\varepsilon^2\alpha\widehat{W}(1)\eta\bigg(\int_0^\infty\varphi(-t)\varphi(\eta+t)\,dt-\int_0^\infty\varphi(t)\varphi(\eta-t)\,dt+O(\alpha)\bigg)\\&=-\pi^2\varepsilon^2\alpha\widehat{W}(1)\eta\bigg(\int_{-\infty}^\infty\varphi(t)\varphi(\eta-t)\,\mathrm{sign}(t)\,dt+O(\alpha)\bigg).\end{aligned}$$

Summarizing,

$$\begin{cases} \lim_{t\to\infty} \tilde f(t,k,\eta) = 0, \quad \text{if } k\neq 0, \\ \lim_{t\to\infty} \tilde f(t,0,\eta) = \tilde f_i(t,0,\eta) - \varepsilon^2\alpha(\pi^2\widehat{W}(1))\eta\Big(\displaystyle\int_{-\infty}^{\infty} \varphi(t)\varphi(\eta-t)\,\mathrm{sign}(t)\,dt + O(\alpha)\Big). \end{cases} \tag{14.7}$$

Since φ is an arbitrary analytic profile, this simple calculation already shows that the asymptotic profile is not necessarily the spatial mean of the initial datum.

Assuming $\varepsilon \ll \alpha$, higher-order expansions in α can be obtained by bootstrap on the equations (14.1), (14.2), (14.4) and (14.5); for instance,

$$\begin{aligned} &\lim_{t\to\infty} \tilde f(t,0,\eta) \\ &\quad = \tilde f_i(0,\eta) - \varepsilon^2\alpha(\pi^2\widehat{W}(1))\eta \int_{-\infty}^{\infty} \varphi(t)\varphi(\eta-t)\,\mathrm{sign}(t)\,dt \\ &\qquad -\varepsilon^2\alpha^2(2\pi^2\widehat{W}(1))^2\eta \int_0^{\infty}\int_0^t \big[(\varphi(\eta+t)\varphi(-\tau) - \varphi(\eta-t)\varphi(\tau))e^{-\pi(t-\tau)^2}(t-\tau) \\ &\qquad\quad +\varphi(\tau)\varphi(-t)e^{-\pi(\eta+(t-\tau))^2}(\eta+t-\tau) \\ &\qquad\quad +\varphi(-\tau)\varphi(t)e^{-\pi(\eta-(t-\tau))^2}(\eta-t-\tau)\big]\,d\tau\,dt + O(\varepsilon^2\alpha^3). \end{aligned}$$

What about the limit in negative time? Reversing time is equivalent to changing $f(t,x,v)$ into $f(t,x,-v)$ and letting time go forward. So we define $S(v) := -v$ and

$$T(\varphi)(\eta) := \varepsilon^2\alpha\pi^2\widehat{W}(1)\eta \int_{-\infty}^{\infty} \varphi(t)\varphi(\eta-t)\,\mathrm{sign}(t)\,dt;$$

then $T(\varphi\circ S) = T(\varphi)\circ S$, which means that the solutions constructed above are always *homoclinic* at order $O(\varepsilon^2\alpha)$. The same is true for the more precise expansions at order $O(\varepsilon^2\alpha^2)$, and in fact it can be checked that the whole distribution f^2 is homoclinic; in other words, f is homoclinic up to possible corrections of order $O(\varepsilon^4)$. To exhibit heteroclinic deviations, we shall consider more general perturbations.

14.2. General perturbation

Let us now consider a "general" initial datum $f_i(x,v)$ close to $f^0(v)$, and expand the solution f. We write $\varepsilon\varphi_k(\eta) = (f_i - f^0)^{\tilde{}}\,(k,\eta)$ and $\varrho^m = \varrho[h^m]$. The interaction potential is assumed to be of the form αW with $\alpha \ll 1$ and $W(x) = W(-x)$. The first equations of the Newton scheme are

$$\hat\varrho^1(t,k) = \varepsilon\varphi_k(kt) - 4\pi^2\alpha\widehat{W}(k)\int_0^t \hat\varrho^1(\tau,k)\tilde f^0(k(t-\tau))|k|^2(t-\tau)\,d\tau, \tag{14.8}$$

$$\tilde{h}^1(t,k,\eta) = \varepsilon\varphi_k(\eta+kt) - 4\pi^2\alpha\widehat{W}(k)\int_0^t \hat{\varrho}^1(\tau,k)\tilde{f}^0(\eta+k(t-\tau))k\cdot(\eta+k(t-\tau))\,d\tau, \quad (14.9)$$

$$\begin{aligned}\tilde{h}^2(t,k,\eta) = {} & -4\pi^2\alpha\widehat{W}(k)\int_0^t \hat{\varrho}^2(\tau,k)\tilde{f}^0(\eta+k(t-\tau))k\cdot(\eta+k(t-\tau))\,d\tau \\ & -4\pi^2\alpha\int_0^t \sum_{l\in\mathbb{Z}^d} \widehat{W}(l)\hat{\varrho}^1(\tau,l)\tilde{h}^1(\tau,k-l,\eta+k(t-\tau))l\cdot(\eta+k(t-\tau))\,d\tau \\ & -4\pi^2\alpha\int_0^t \sum_{l\in\mathbb{Z}^d} \widehat{W}(l)\hat{\varrho}^2(\tau,l)\tilde{h}^1(\tau,k-l,\eta+k(t-\tau))l\cdot(\eta+k(t-\tau))\,d\tau \\ & -4\pi^2\alpha\int_0^t \sum_{l\in\mathbb{Z}^d} \widehat{W}(l)\hat{\varrho}^1(\tau,l)\tilde{h}^2(\tau,k-l,\eta+k(t-\tau))l\cdot(\eta+k(t-\tau))\,d\tau,\end{aligned} \quad (14.10)$$

$$\begin{aligned}\hat{\varrho}^2(t,k) = {} & -4\pi^2\alpha\widehat{W}(k)\int_0^t \hat{\varrho}^2(\tau,k)\tilde{f}^0(k(t-\tau))|k|^2(t-\tau)\,d\tau \\ & -4\pi^2\alpha\int_0^t \sum_{l\in\mathbb{Z}^d} \widehat{W}(l)\hat{\varrho}^1(\tau,l)\tilde{h}^1(\tau,k-l,k(t-\tau))l\cdot k(t-\tau)\,d\tau \\ & -4\pi^2\alpha\int_0^t \sum_{l\in\mathbb{Z}^d} \widehat{W}(l)\hat{\varrho}^2(\tau,l)\tilde{h}^1(\tau,k-l,k(t-\tau))l\cdot k(t-\tau)\,d\tau \\ & -4\pi^2\alpha\int_0^t \sum_{l\in\mathbb{Z}^d} \widehat{W}(l)\hat{\varrho}^1(\tau,l)\tilde{h}^2(\tau,k-l,k(t-\tau))l\cdot k(t-\tau)\,d\tau.\end{aligned} \quad (14.11)$$

Here k runs over $\mathbb{Z}^d$.

From (14.8) and (14.9) we see that ϱ^1 and h^1 depend *linearly* on ε, and that

$$\hat{\varrho}^1(t,k) = \varepsilon[\varphi_k(kt)+O(\alpha)] \quad \text{and} \quad \tilde{h}^1(t,k,\eta) = \varepsilon[\varphi_k(\eta+kt)+O(\alpha)]. \quad (14.12)$$

Then from (14.10) and (14.11), ϱ^2 and h^2 are $O(\varepsilon^2\,\alpha)$; so by plugging (14.12) into these equations we obtain

$$\hat{\varrho}^2(t,k) = -4\pi^2\varepsilon^2\alpha\int_0^t \sum_{l\in\mathbb{Z}^d} \widehat{W}(l)\varphi_l(l\tau)\varphi_{k-l}(kt-l\tau)l\cdot k(t-\tau)\,d\tau + O(\varepsilon^2\alpha^2) + O(\varepsilon^3\alpha) \quad (14.13)$$

and

$$\begin{aligned}\tilde{h}^2(t,k,\eta) = {} & -4\pi^2\varepsilon^2\alpha\int_0^t \sum_{l\in\mathbb{Z}^d} \widehat{W}(l)\varphi_l(l\tau)\varphi_{k-l}(\eta+kt-l\tau)l\cdot(\eta+k(t-\tau))\,d\tau \\ & +O(\varepsilon^2\alpha^2)+O(\varepsilon^3\alpha).\end{aligned} \quad (14.14)$$

We plug these bounds again into the right-hand side of (14.10) to find that

$$\tilde{h}^2(t,0,\eta) = \mathrm{I\!I}_\varepsilon(t,\eta) + \mathrm{I\!I\!I}_\varepsilon(t,\eta) + O(\varepsilon^3\,\alpha^3), \tag{14.15}$$

where

$$\mathrm{I\!I}_\varepsilon(t,\eta) = -4\pi^2\alpha \int_0^t \sum_{l\in\mathbb{Z}^d} (l\cdot\eta)\widehat{W}(l)\hat{\varrho}^1(\tau,l)\tilde{h}^1(\tau,-l,\eta)\,d\tau$$

is quadratic in ε, and $\mathrm{I\!I\!I}_\varepsilon(t,\eta)$ is a third-order correction:

$$\begin{aligned}\mathrm{I\!I\!I}_\varepsilon(t,\eta) = 16\pi^4\varepsilon^3\alpha^2 \sum_{m,l\in\mathbb{Z}^d} \widehat{W}(l)\widehat{W}(m) \\ \times \int_0^t \int_0^\tau \varphi_m(ms)\big[\varphi_{l-m}(l\tau-ms)\varphi_{-l}(\eta-l\tau)(l\cdot m)(\tau-s) \\ +\varphi_l(l\tau)\varphi_{-l-m}(\eta-l\tau-ms)m\cdot(\eta-l(\tau-s))\big](l\cdot\eta)\,ds\,d\tau.\end{aligned} \tag{14.16}$$

If $\tilde{f}^0$ is even, changing φ_k into $\varphi_k(-\cdot)$ and η into $-\eta$ amounts to change k into $-k$ at the level of (14.8) and (14.9); but then $\mathrm{I\!I}_\varepsilon$ is invariant under this operation. We conclude that f is always homoclinic at second order in ε, and we consider the influence of the third-order term (14.16). Let

$$C[\varphi](\eta) := \lim_{t\to\infty} \mathrm{I\!I\!I}_\varepsilon(t,\eta).$$

After some relabeling, we find that

$$\begin{aligned}C[\varphi](\eta) = 16\pi^4\varepsilon^3\alpha^2 \sum_{k,l\in\mathbb{Z}^d} \widehat{W}(k)\,\widehat{W}(l) \\ \times \int_0^\infty \int_0^t \varphi_l(l\tau)\big[\varphi_{k-l}(kt-l\tau)\varphi_{-k}(\eta-kt)(k\cdot l)(t-\tau) \\ +\varphi_k(kt)\varphi_{-k-l}(\eta-kt-l\tau)l\cdot(\eta-k(t-\tau))\big](k\cdot\eta)\,d\tau\,dt.\end{aligned} \tag{14.17}$$

Now assume that $\varphi_{-k}=\sigma\varphi_k$ with $\sigma=\pm1$ ($\sigma=1$ means that the perturbation is even in x, $\sigma=-1$ that it is odd.) Using the symmetry $(k,l)\leftrightarrow(-k,-l)$ one can check that

$$C[\varphi\circ S]\circ S = \sigma C[\varphi],$$

where $S(z)=-z$. In particular, if the perturbation is odd in x, then the third-order correction imposes a heteroclinic behavior for the solution, as soon as $C[\varphi]\neq0$.

To construct an example where $C[\varphi]\neq0$, we set $d=1$, $f^0=$Gaussian,

$$f_i - f^0 = \sin(2\pi x)\theta_1(v) + \sin(4\pi x)\theta_2(v),$$

$\varphi_1 = -\varphi_{-1} = \frac{1}{2}\tilde{\theta}_1$ and $\varphi_2 = -\varphi_{-2} = \frac{1}{2}\tilde{\theta}_2$. The six pairs (k,l) contributing to (14.17) are $(-1,1)$, $(1,-1)$, $(1,2)$, $(2,1)$, $(-1,-2)$ and $(-2,-1)$, By playing with the respective sizes of $\widehat{W}(1)$ and $\widehat{W}(2)$ (which amounts in fact to changing the size of the box), it is sufficient to consider the terms with coefficient $\widehat{W}(1)^2$, i.e., the pairs $(-1,1)$ and $(1,-1)$. Then the corresponding bit of $C[\varphi](\eta)$ is

$$-16\pi^4\varepsilon^3\alpha^2\widehat{W}(1)^2\eta\int_0^\infty\int_0^t \big[\varphi_1(\tau)\varphi_1(\eta+t)\varphi_2(-t+\tau)(t-\tau) + \varphi_1(\tau)\varphi_1(t)\varphi_2(\eta+t-\tau)(\eta+t-\tau) + \varphi_1(-\tau)\varphi_1(\eta-t)\varphi_2(t+\tau)(t-\tau) + \varphi_1(-\tau)\varphi_1(t)\varphi_2(\eta-t+\tau)(t-\tau-\eta)\big]\,d\tau\,dt.$$

If we let φ_1 and φ_2 vary in such a way that they become positive and almost concentrated on $\mathbb{R}_+$, the only remaining term is the one in $\varphi_1(\tau)\varphi_2(\eta+t-\tau)\varphi_1(t)$, and its contribution is negative for $\eta>0$. So, at least for certain values of $W(1)$ and $W(2)$ there is a choice of analytic functions φ_1 and φ_2, such that $C[\varphi]\neq 0$. This demonstrates the existence of heteroclinic trajectories.

To summarize: At first order in ε, the convergence is to the spatial average; at second order there is a homoclinic correction; at third order, if at least three modes with zero sum are excited, there is a possibility for heteroclinic behavior.

Remark 14.1. As pointed out to us by Bouchet, the existence of heteroclinic trajectories implies that the asymptotic behavior cannot be predicted on the basis of the invariants of the equation and the interaction; indeed, the latter do not distinguish between the forward and backward solutions.

15. Beyond Landau damping

We conclude this paper with some general comments about the physical implications of Landau damping.

Remark 14.1 show in particular that there is no "universal" large-time behavior of the solution of the non-linear Vlasov equation in terms of just, say, conservation laws and the initial datum; the *dynamics* also have to enter explicitly. One can also interpret this as a lack of ergodicity: the non-linearity is not sufficient to make the system explore the space of all "possible" distributions and to choose the most favorable one, whatever this means. Failure of ergodicity for a system of finitely many particles was already known to occur, in relation to the KAM theorem; this is mentioned e.g. in [62, p. 257] for the vortex system. There it is hoped that such behavior disappears as the dimension tends

to infinity; but now we see that it also exists even in the infinite-dimensional setting of the Vlasov equation.

At first, this seems to be bad news for the statistical theory of the Vlasov equation, pioneered by Lynden-Bell [58] and explored by various authors [16], [65], [80], [89], [94], since even the sophisticated variants of this theory try to predict the likely final states in terms of just the characteristics of the initial data. In this sense, our results provide support for an objection raised by Isichenko [42, p. 2372] against the statistical theory.

However, looking more closely at our proofs and results, proponents of the statistical theory will have a lot to rejoice about.

To start with, our results are the first to rigorously establish that the non-linear Vlasov equation does enjoy some asymptotic "stabilization" property in large time, without the help of any extra diffusion or ensemble averaging.

Next, the whole analysis is perturbative: each stable spatially homogeneous distribution will have its small "basin of damping", and it may be that some distributions are "much more stable" than others, say in the sense of having a larger basin.

Even more importantly, in §7 we have crucially used the smoothness to overcome the potentially destabilizing non-linear effects. So any theory based on non-smooth functions might not be constrained by Landau damping. This certainly applies to a statistical theory, for which smooth functions should be a zero-probability set.

Finally, to overcome the non-linearity, we had to cope with huge constants (even qualitatively larger than those appearing in classical KAM theory). If one believes in the explanatory virtues of proofs, these large constants might be the indication that Landau damping is a thin effect, which might be neglected when it comes to predict the "final" state in a "turbulent" situation.

Further work needs to be done to understand whether these considerations apply equally to the electrostatic and gravitational cases, or whether the electrostatic case is favored in these respects; and what happens in "low" regularity.

Although the underlying mathematical and physical mechanisms differ, non-linear Landau damping (as defined by Theorem 2.6) may arguably be to the theory of the Vlasov equation what the KAM theorem is to the theory of Hamiltonian systems. Like the KAM theorem, it might be conceptually important in theory and practice, and still be severely limited.([18])

Beyond the range of application of KAM theory lies the softer, more robust *weak KAM theory* developed by Fathi [26] in relation to Aubry–Mather theory. By a nice co-

([18]) It is a well-known scientific paradox that the KAM theorem was at the same time tremendously influential in the science of the twentieth century, and so restrictive that its assumptions are essentially never satisfied in practice.

incidence, a Vlasov version of the weak KAM theory has just been developed by Gangbo and Tudorascu [29], although with no relation to Landau damping. Making the connection is just one of the many developments which may be explored in the future.

Appendix

In this appendix we gather some elementary tools, our conventions, and some reminders about calculus. We write $\mathbb{N}_0 = \{0, 1, 2, \dots\}$.

A.1. Calculus in dimension d

For $n \in \mathbb{N}_0^d$ we define

$$n! = n_1! \dots n_d!$$

and for $n, m \in \mathbb{N}_0^d$ we set

$$\binom{n}{m} = \binom{n_1}{m_1} \dots \binom{n_d}{m_d}.$$

For $z \in \mathbb{C}^d$ and $n \in \mathbb{Z}^d$ we let

$$\|z\| = |z_1| + \dots + |z_d|, \quad z^n = z_1^{n_1} \dots z_n^{n_d} \in \mathbb{C} \quad \text{and} \quad |z|^n = |z^n|.$$

In particular, for $z \in \mathbb{C}^d$ we have

$$e^{\|z\|} = e^{|z_1| + \dots + |z_d|} = \sum_{n \in \mathbb{N}_0^d} \frac{\|z\|^n}{n!}.$$

We may write $e^{|z|}$ instead of $e^{\|z\|}$.

A.2. Multi-dimensional differential calculus

The Leibniz formula for functions $f, g : \mathbb{R} \to \mathbb{R}$ is

$$(fg)^{(n)} = \sum_{m=0}^{n} \binom{n}{m} f^{(m)} g^{(n-m)},$$

where of course $f^{(n)} = d^n f / dx^n$. The expression of derivatives of composed functions is given by the Faà di Bruno formula:

$$(f \circ G)^{(n)} = \sum_{\sum_j j m_j = n} \frac{n!}{m_1! \dots m_n!} (f^{(m_1 + \dots + m_n)} \circ G) \prod_{j=1}^{n} \left(\frac{G^{(j)}}{j!} \right)^{m_j}.$$

These formulae remain valid in several dimensions, provided that one defines, for a multi-index $n=(n_1,...,n_d)$,

$$f^{(n)}=\frac{\partial^{n_1}}{\partial x_1^{n_1}}\cdots\frac{\partial^{n_d}}{\partial x_d^{n_d}}f.$$

They also remain true if $(\partial_1,...,\partial_d)$ is replaced by a d-tuple of *commuting* derivation operators.

As a consequence, we shall establish the following Leibniz-type formula for operators which are combinations of gradients and multiplications.

LEMMA A.1. *Let f and g be functions of $v\in\mathbb{R}^d$ and $a,b\in\mathbb{C}^d$. Then for any $n\in\mathbb{N}^d$,*

$$(\nabla_v+(a+b))^n(fg)=\sum_{0\leqslant m\leqslant n}\binom{n}{m}(\nabla_v+a)^m f(\nabla_v+b)^{n-m}g.$$

Proof. The right-hand side is equal to

$$\sum_{\substack{0\leqslant q\leqslant m\leqslant n\\0\leqslant r\leqslant n-m}}\binom{n}{m}\binom{m}{q}\binom{n-m}{r}\nabla_v^q f\nabla_v^r g a^{m-q}b^{n-m-r}.$$

After changing indices $p=q+r$ and $s=m-q$, this becomes

$$\begin{aligned}\sum_{\substack{0\leqslant r\leqslant p\leqslant n\\0\leqslant s\leqslant n-p}}\binom{n}{p}\binom{p}{r}\binom{n-p}{s}\nabla_v^r g\nabla_v^{p-r}f a^s b^{n-p-s}&=\sum_p\binom{n}{p}\nabla_v^p(fg)(a+b)^{n-p}\\&=(\nabla_v+(a+b))^n(fg).\end{aligned}$$

□

A.3. Fourier transform

For a function $f:\mathbb{R}^d\to\mathbb{R}$, we define

$$\tilde f(\eta)=\int_{\mathbb{R}^d}e^{-2i\pi\eta\cdot v}f(v)\,dv. \tag{A.1}$$

Then we have the usual formulae

$$f(v)=\int_{\mathbb{R}^d}\tilde f(\eta)e^{2i\pi\eta\cdot v}\,d\eta\quad\text{and}\quad\widetilde{\nabla f}(\eta)=2i\pi\eta\tilde f(\eta).$$

Let $\mathbb{T}_L^d=\mathbb{R}^d/L\mathbb{Z}^d$. For a function $f:\mathbb{T}_L^d\to\mathbb{R}$, we define

$$\hat f^{(L)}(k)=\int_{\mathbb{T}_L^d}e^{-2i\pi k\cdot x/L}f(x)\,dx. \tag{A.2}$$

Then we have

$$f(x)=\frac{1}{L^d}\sum_{k\in\mathbb{Z}^d}\hat{f}^{(L)}(k)e^{2i\pi k\cdot x/L}\quad\text{and}\quad\widehat{\nabla f}^{(L)}(k)=2i\pi\frac{k}{L}\hat{f}^{(L)}(k).$$

For a function $f\colon\mathbb{T}^d_L\times\mathbb{R}^d\to\mathbb{R}$, we define

$$\tilde{f}^{(L)}(k,\eta)=\int_{\mathbb{T}^d_L}\int_{\mathbb{R}^d}e^{-2i\pi k\cdot x/L}e^{-2i\pi\eta\cdot v}f(x,v)\,dv\,dx,\tag{A.3}$$

so that the reconstruction formula reads

$$f(x,v)=\frac{1}{L^d}\sum_{k\in\mathbb{Z}^d}\int_{\mathbb{R}^d}\tilde{f}^{(L)}(k,\eta)e^{2i\pi k\cdot x/L}e^{2i\pi\eta\cdot v}\,dv.$$

When $L=1$ we do not specify it: so we just write

$$\hat{f}=\hat{f}^{(1)}\quad\text{and}\quad\tilde{f}=\tilde{f}^{(1)}.$$

(There is no risk of confusion since, in that case, (A.3) and (A.1) coincide.)

A.4. Fixed-point theorem

The following theorem is one of the many variants of the Picard fixed-point theorem. We write $B(0,R)$ for the closed ball of center 0 and radius R.

THEOREM A.2. (Fixed-point theorem) *Let E be a Banach space, $F\colon E\to E$ and $R=2\|F(0)\|$. If $F\colon B(0,R)\to E$ is $\frac{1}{2}$-Lipschitz, then it has a unique fixed point in $B(0,R)$.*

Proof. Uniqueness is obvious. To prove existence, run the classical Picard iterative scheme initialized at 0: $x_0=0$, $x_1=F(0)$, $x_2=F(F(0))$, etc. It is clear that $(x_n)_{n=1}^{\infty}$ is a Cauchy sequence and $\|x_n\|\leqslant\|F(0)\|(1+...+1/2^n)\leqslant 2\|F(0)\|$, so x_n converges in $B(0,R)$ to a fixed point of F. □

A.5. Plemelj formula

The Plemelj formula states that, in $\mathcal{D}'(\mathbb{R})$,

$$\frac{1}{x-i0}=\text{p.v.}\left(\frac{1}{x}\right)+i\pi\,\delta_0,\tag{A.4}$$

or, equivalently,

$$\frac{1}{x+i0}=\text{p.v.}\left(\frac{1}{x}\right)-i\pi\,\delta_0.\tag{A.5}$$

Let us give a complete proof of this formula which plays a crucial role in plasma physics.

Proof of (A.4). First we recall that

$$\text{p.v.}\left(\frac{1}{x}\right) = \lim_{\varepsilon \to 0} \frac{1_{|x| \geqslant \varepsilon}}{x}.$$

A more explicit expression can be given for this limit. Let χ be an even function on $\mathbb{R}$ with $\chi(0)=1$, $\chi \in \text{Lip} \cap L^1(\mathbb{R})$; in particular

$$\int_{\mathbb{R}} \frac{1_{|x| \geqslant \varepsilon}}{x} \chi(x)\, dx = 0$$

and $(1-\chi(x))/x$ is bounded. Then for any $\varphi \in \text{Lip}(\mathbb{R}) \cap L^1(\mathbb{R})$,

$$\begin{aligned} \int_{|x| \geqslant \varepsilon} \frac{\varphi(x)}{x}\, dx &= \int_{|x| \geqslant \varepsilon} \frac{\varphi(x)}{x} \chi(x)\, dx + \int_{|x| \geqslant \varepsilon} \varphi(x) \frac{1-\chi(x)}{x}\, dx \\ &= \int_{|x| \geqslant \varepsilon} \frac{\varphi(x)-\varphi(0)}{x} \chi(x)\, dx + \int_{|x| \geqslant \varepsilon} \varphi(x) \frac{1-\chi(x)}{x}\, dx \qquad \text{(A.6)} \\ &\to \int_{\mathbb{R}} \frac{\varphi(x)-\varphi(0)}{x} \chi(x)\, dx + \int_{\mathbb{R}} \varphi(x) \frac{1-\chi(x)}{x}\, dx \quad \text{as } \varepsilon \to 0. \end{aligned}$$

Then (A.4) can be rewritten in the following more explicit way: for any $\varphi \in \text{Lip}(\mathbb{R}) \cap L^1(\mathbb{R})$ and any χ satisfying the above assumptions,

$$\lim_{\lambda \to 0^+} \int_{\mathbb{R}} \frac{\varphi(x)}{x - i\lambda}\, dx = \int_{\mathbb{R}} \frac{\varphi(x)-\varphi(0)}{x} \chi(x)\, dx + \int_{\mathbb{R}} \varphi(x) \frac{1-\chi(x)}{x}\, dx + i\pi\varphi(0). \qquad \text{(A.7)}$$

Now, to prove (A.7), let us pick such a function χ and write, for $\lambda > 0$,

$$\int_{\mathbb{R}} \frac{\varphi(x)}{x-i\lambda}\, dx = \int_{\mathbb{R}} \frac{\varphi(x)-\varphi(0)}{x-i\lambda} \chi(x)\, dx + \int_{\mathbb{R}} \varphi(x) \frac{1-\chi(x)}{x-i\lambda}\, dx + \varphi(0) \int_{\mathbb{R}} \frac{\chi(x)}{x-i\lambda}\, dx.$$

As $\lambda \to 0$, the first two integrals on the right-hand side converge to the right-hand side of (A.6), and there is an extra term proportional to $\varphi(0)$; so it only remains to check that

$$\int_{\mathbb{R}} \frac{\chi(x)}{x-i\lambda}\, dx \to i\pi \quad \text{as } \lambda \to 0^+. \qquad \text{(A.8)}$$

If (A.8) holds for some particular χ satisfying the requested conditions, then (A.4) follows and it implies that (A.8) holds for any such χ. So let us pick one particular χ, say $\chi(x) = e^{-x^2}$, which can be extended throughout the complex plane into a holomorphic function. Then, since the complex integral is invariant under contour deformation,

$$\int_{\mathbb{R}} \frac{e^{-x^2}}{x-i\lambda}\, dx = \int_{C_\varepsilon} \frac{e^{-z^2}}{z-i\lambda}\, dz,$$

where C_ε is the complex contour made of the straight line $(-\infty, -\varepsilon)$, followed by the half-circle $D_\varepsilon = \{-\varepsilon e^{i\theta} : 0 \leqslant \theta \leqslant \pi\}$, followed by the straight line (ε, ∞). The contributions of both straight lines cancel each other by symmetry, and only the integral on the half-circle D_ε remains. As $\lambda \to 0$ this integral approaches $\int_{D_\varepsilon} e^{-z^2}\, dz/z$, which as $\varepsilon \to 0$ becomes close to $\int_{D_\varepsilon} dz/z = i\pi$, as was claimed. □

References

[1] Akhiezer, A., Akhiezer, I., Polovin, R., Sitenko, A. & Stepanov, K., *Plasma Electrodynamics.* Vol. I: *Linear Theory.* Vol. II: *Non-Linear Theory and Fluctuations.* Pergamon Press, Oxford–New York, 1975.

[2] Alinhac, S. & Gérard, P., *Pseudo-Differential Operators and the Nash–Moser Theorem.* Graduate Studies in Mathematics, 82. Amer. Math. Soc., Providence, RI, 2007.

[3] Bach, V., Fröhlich, J. & Sigal, I. M., Spectral analysis for systems of atoms and molecules coupled to the quantized radiation field. *Comm. Math. Phys.*, 207 (1999), 249–290.

[4] Backus, G., Linearized plasma oscillations in arbitrary electron velocity distributions. *J. Math. Phys.*, 1 (1960), 178–191; erratum, 559.

[5] Balescu, R., *Statistical Mechanics of Charged Particles.* Monographs in Statistical Physics and Thermodynamics, 4. Wiley, London–New York–Sydney, 1963.

[6] Batt, J. & Rein, G., Global classical solutions of the periodic Vlasov–Poisson system in three dimensions. *C. R. Acad. Sci. Paris Sér. I Math.*, 313 (1991), 411–416.

[7] Belmont, G., Mottez, F., Chust, T. & Hess, S., Existence of non-Landau solutions for Langmuir waves. *Phys. of Plasmas*, 15 (2008), 052310, 1–14.

[8] Benachour, S., Analyticité des solutions des équations de Vlasov–Poisson. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 16 (1989), 83–104.

[9] Binney, J. & Tremaine, S., *Galactic Dynamics.* First edition. Princeton Series in Astrophysics. Princeton University Press, Princeton, 1987.

[10] — *Galactic Dynamics.* Second edition. Princeton Series in Astrophysics. Princeton University Press, Princeton, 2008.

[11] Bouchet, F., Stochastic process of equilibrium fluctuations of a system with long-range interactions. *Phys. Rev. E*, 70 (2004), 036113, 1–4.

[12] Bourgain, J., Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations. *Geom. Funct. Anal.*, 3 (1993), 107–156.

[13] Caglioti, E. & Maffei, C., Time asymptotics for solutions of Vlasov–Poisson equation in a circle. *J. Stat. Phys.*, 92 (1998), 301–323.

[14] Case, K. M., Plasma oscillations. *Ann. Physics*, 7 (1959), 349–364.

[15] Chavanis, P. H., Quasilinear theory of the 2D Euler equation. *Phys. Rev. Lett.*, 84:24 (2000), 5512–5515.

[16] Chavanis, P. H., Sommeria, J., & Robert, R., Statistical mechanics of two-dimensional vortices and collisionless stellar systems. *Astrophys. J.*, 471 (1996), 385–399.

[17] 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*, Sémin. Congr., 9, pp. 99–123. Soc. Math. France, Paris, 2004.

[18] 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*, 98 (1954), 527–530]. *Regul. Chaotic Dyn.*, 13 (2008), 130–139.

[19] Chust, T., Belmont, G., Mottez, F. & Hess, S., Landau and non-Landau linear damping: Physics of the dissipation. *Phys. Plasmas*, 16 (2009), 092104, 13 pp.

[20] Degond, P., Spectral theory of the linearized Vlasov–Poisson equation. *Trans. Amer. Math. Soc.*, 294 (1986), 435–453.

[21] Dereziński, J. & Gérard, C., *Scattering Theory of Classical and Quantum N-Particle Systems.* Texts and Monographs in Physics. Springer, Berlin–Heidelberg, 1997.

[22] Desvillettes, L. & Villani, C., On the trend to global equilibrium in spatially inhomogeneous entropy-dissipating systems: the linear Fokker–Planck equation. *Comm. Pure Appl. Math.*, 54 (2001), 1–42.

[23] Elskens, Y., Irreversible behaviours in Vlasov equation and many-body Hamiltonian dynamics: Landau damping, chaos and granularity, in *Topics in Kinetic Theory*, Fields Inst. Commun., 46, pp. 89–108. Amer. Math. Soc., Providence, RI, 2005.

[24] Elskens, Y. & Escande, D. F., *Microscopic Dynamics of Plasmas and Chaos*. Institute of Physics, Bristol, 2003.

[25] Escande, D. F., Wave-particle interaction in plasmas: a qualitative approach, in *Long-Range Interacting Systems*, pp. 13–14, 469–506. Oxford University Press, Oxford, 2010.

[26] Fathi, A., *Weak KAM Theory in Lagrangian Dynamics*. Cambridge University Press, Cambridge, 2010.

[27] Filbet, F., Numerical simulations. Available online at `http://math.univ-lyon1.fr/~filbet/publication.html`.

[28] Fridman, A. M. & Polyachenko, V. L., *Physics of Gravitating Systems*. Vol. I: *Equilibrium and Stability*. Vol. II: *Nonlinear Collective Processes. Astrophysical Applications*. Springer, New York, 1984.

[29] Gangbo, W. & Tudorascu, A., Lagrangian dynamics on an infinite-dimensional torus; a weak KAM theorem. *Adv. Math.*, 224 (2010), 260–292.

[30] Glassey, R. & Schaeffer, J., Time decay for solutions to the linearized Vlasov equation. *Transport Theory Statist. Phys.*, 23 (1994), 411–453.

[31] — On time decay rates in Landau damping. *Comm. Partial Differential Equations*, 20 (1995), 647–676.

[32] Gould, R., O'Neil, T. & Malmberg, J., Plasma wave echo. *Phys. Rev. Letters*, 19:5 (1967), 219–222.

[33] Gross, L., Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97 (1975), 1061–1083.

[34] Guo, Y. & Rein, G., A non-variational approach to non-linear stability in stellar dynamics applied to the King model. *Comm. Math. Phys.*, 271 (2007), 489–509.

[35] Guo, Y. & Strauss, W. A., Nonlinear instability of double-humped equilibria. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 12 (1995), 339–352.

[36] ter Haar, D., *Men of Physics*: *L. D. Landau*. Selected Reading of Physics, 2. Pergamon Press, Oxford–New York, 1969.

[37] Hauray, M. & Jabin, P. E., N-particles approximation of the Vlasov equations with singular potential. *Arch. Ration. Mech. Anal.*, 183 (2007), 489–524.

[38] Hayes, J. N., On non-Landau damped solutions to the linearized Vlasov equation. *Nuovo Cimento*, 10:30 (1963), 1048–1063.

[39] Heath, R., Gamba, I., Morrison, P. & Michler, C., A discontinuous Galerkin method for the Vlasov–Poisson system. To appear in *J. Comput. Phys.*

[40] Horst, E., On the asymptotic growth of the solutions of the Vlasov–Poisson system. *Math. Methods Appl. Sci.*, 16 (1993), 75–86.

[41] Hwang, H. J. & Velázquez, J. J. L., On the existence of exponentially decreasing solutions of the nonlinear Landau damping problem. *Indiana Univ. Math. J.*, 58:6 (2009), 2623–2660.

[42] Isichenko, M., Nonlinear Landau damping in collisionless plasma and inviscid fluid. *Phys. Rev. Lett.*, 78:12 (1997), 2369–2372.

[43] Jabin, P. E., Averaging lemmas and dispersion estimates for kinetic equations. *Riv. Mat. Univ. Parma*, 1 (2009), 71–138.

[44] Kaganovich, I. D., Effects of collisions and particle trapping on collisionless heating. *Phys. Rev. Lett.*, 82:2 (1999), 327–330.

[45] VAN KAMPEN, N. G., On the theory of stationary waves in plasmas. *Physica*, 21 (1955), 949–963.

[46] KANDRUP, H., Violent relaxation, phase mixing, and gravitational Landau damping. *Astrophys. J.*, 500 (1998), 120–128.

[47] KIESSLING, M. K.-H., The "Jeans swindle": a true story—mathematically speaking. *Adv. in Appl. Math.*, 31 (2003), 132–149.

[48] — Personal communication, 2009.

[49] KRALL, N. & TRIVELPIECE, A., *Principles of Plasma Physics*. San Francisco Press, San Francisco, 1986.

[50] KUKSIN, S. B., *Nearly Integrable Infinite-Dimensional Hamiltonian Systems*. Lecture Notes in Mathematics, 1556. Springer, Berlin–Heidelberg, 1993.

[51] — *Analysis of Hamiltonian PDEs*. Oxford Lecture Series in Mathematics and its Applications, 19. Oxford University Press, Oxford, 2000.

[52] LANDAU, L., On the vibrations of the electronic plasma. *Akad. Nauk SSSR. Zhurnal Eksper. Teoret. Fiz.*, 16 (1946), 574–586 (Russian); English translation in *Acad. Sci. USSR. J. Phys.*, 10 (1946), 25–34. Reproduced in [36].([19])

[53] LEMOU, M., MÉHATS, F. & RAPHAËL, P., Orbital stability of spherical galactic models. To appear in *Invent. Math.*

[54] LIFSHITZ, E. M. & PITAEVSKIĬ, L. P., *Course of Theoretical Physics* ["Landau–Lifshits"]. Vol. 10. Nauka, Moscow, 1979. English translation in Pergamon International Library of Science, Technology, Engineering and Social Studies. Pergamon Press, Oxford–New York, 1981.

[55] LIN, Z. & ZENG, C., BGK waves and non-linear Landau damping. To appear in *Comm. Math. Phys.*

[56] LIONS, P. L. & PERTHAME, B., Propagation of moments and regularity for the 3-dimensional Vlasov–Poisson system. *Invent. Math.*, 105 (1991), 415–430.

[57] LYNDEN-BELL, D., The stability and vibrations of a gas of stars. *Mon. Not. R. Astr. Soc.*, 124 (1962), 279–296.

[58] — Statistical mechanics of violent relaxation in stellar systems. *Mon. Not. R. Astr. Soc.*, 136 (1967), 101–121.

[59] MALMBERG, J. & WHARTON, C., Collisionless damping of electrostatic plasma waves. *Phys. Rev. Lett.*, 13:6 (1964), 184–186.

[60] MALMBERG, J., WHARTON, C., GOULD, R. & O'NEIL, T., Plasma wave echo experiment. *Phys. Rev. Letters*, 20:3 (1968), 95–97.

[61] MANFREDI, G., Long-time behavior of non-linear Landau damping. *Phys. Rev. Lett.*, 79:15 (1997), 2815–2818.

[62] MARCHIORO, C. & PULVIRENTI, M., *Mathematical Theory of Incompressible Nonviscous Fluids*. Applied Mathematical Sciences, 96. Springer, New York, 1994.

[63] MASLOV, V. P. & FEDORYUK, M. V., The linear theory of Landau damping. *Mat. Sb.*, 127(169) (1985), 445–475, 559 (Russian); English translation in *Math. USSR–Sb.*, 55 (1986), 437–465.

[64] MEDVEDEV, M. V., DIAMOND, P. H., ROSENBLUTH, M. N. & SHEVCHENKO, V. I., Asymptotic theory of non-linear Landau damping and particle trapping in waves of finite amplitude. *Phys. Rev. Lett.*, 81:26 (1998), 5824–5827.

[65] MILLER, J., Statistical mechanics of Euler equations in two dimensions. *Phys. Rev. Lett.*, 65:17 (1990), 2137–2140.

([19]) There is a misprint in formula (17) of this reference (p. 104): replace $e^{-(ka)^2/2}$ by $e^{-1/2(ka)^2}$.

[66] Morrison, P. J., Hamiltonian description of Vlasov dynamics: action-angle variables for the continuous spectrum, in *Proceedings of the Fifth International Workshop on Mathematical Aspects of Fluid and Plasma Dynamics* (Maui, HI, 1998), Transport Theory Statist. Phys., 29, pp. 397–414. Taylor & Francis, Philadelphia, PA, 2000.

[67] Moser, J., A rapidly convergent iteration method and non-linear differential equations. II. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 20 (1966), 499–535.

[68] — Recollections, in *The Arnoldfest* (Toronto, ON, 1997), Fields Inst. Commun., 24, pp. 19–21. Amer. Math. Soc., Providence, RI, 1999.

[69] Mouhot, C. & Villani, C., Landau damping. *J. Math. Phys.*, 51 (2010), 015204, 7.

[70] Nash, J., The imbedding problem for Riemannian manifolds. *Ann. of Math.*, 63 (1956), 20–63.

[71] Nekhoroshev, N. N., An exponential estimate of the time of stability of nearly integrable Hamiltonian systems. *Uspekhi Mat. Nauk*, 32 (1977), 5–66, 287 (Russian); English translation in *Russian Math. Surveys*, 32 (1977), 1–65.

[72] — An exponential estimate of the time of stability of nearly integrable Hamiltonian systems. II. *Trudy Sem. Petrovsk.*, (1979), 5–50 (Russian); English translation in *Topics in Modern Mathematics*, pp. 1–58, Contemporary Soviet Mathematics, Consultants Bureau, New York, 1985.

[73] Nirenberg, L., An abstract form of the nonlinear Cauchy–Kowalewski theorem. *J. Differential Geom.*, 6 (1972), 561–576.

[74] Nishida, T., A note on a theorem of Nirenberg. *J. Differential Geom.*, 12 (1977), 629–633 (1978).

[75] O'Neil, T. M., Collisionless damping of nonlinear plasma oscillations. *Phys. Fluids*, 8 (1965), 2255–2262.

[76] O'Neil, T. M. & Coroniti, F. V., The collisionless nature of high-temperature plasmas. *Rev. Modern Phys.*, 71:2 (1999), S404–S410.

[77] Penrose, O., Electrostatic instability of a non-Maxwellian plasma. *Phys. Fluids*, 3 (1960), 258–265.

[78] Pfaffelmoser, K., Global classical solutions of the Vlasov–Poisson system in three dimensions for general initial data. *J. Differential Equations*, 95 (1992), 281–303.

[79] Rein, G., Personal communication, 2008.

[80] Robert, R., Statistical mechanics and hydrodynamical turbulence, in *Proceedings of the International Congress of Mathematicians* (Zürich, 1994), Vol. 2, pp. 1523–1531. Birkhäuser, Basel, 1995.

[81] Ryutov, D. D., Landau damping: half a century with the great discovery. *Plasma Phys. Control. Fusion*, 41 (1999), A1–A12.

[82] Sáenz, A. W., Long-time behavior of the electric potential and stability in the linearized Vlasov theory. *J. Math. Phys.*, 6 (1965), 859–875.

[83] Saint Raymond, X., A simple Nash–Moser implicit function theorem. *Enseign. Math.*, 35 (1989), 217–226.

[84] Schaeffer, J., Global existence of smooth solutions to the Vlasov–Poisson system in three dimensions. *Comm. Partial Differential Equations*, 16 (1991), 1313–1335.

[85] Soffer, A. & Weinstein, M. I., Multichannel nonlinear scattering for nonintegrable equations. *Comm. Math. Phys.*, 133 (1990), 119–146.

[86] Spentzouris, L., Ostiguy, J. & Colestock, P., Direct measurement of diffusion rates in high energy synchrotrons using longitudinal beam echoes. *Phys. Rev. Lett.*, 76:4 (1996), 620–623.

[87] Stahl, B., Kiessling, M. K.-H. & Schindler, K., Phase transitions in gravitating systems and the formation of condensed objects. *Planet. Space Sci.*, 43 (1995), 271–282.
[88] Stix, T. H., *The Theory of Plasma Waves*. McGraw-Hill, New York, 1962.
[89] Tremaine, S., Hénon, M. & Lynden-Bell, D., H-functions and mixing in violent relaxation. *Mon. Not. R. Astr. Soc.*, 219 (1986), 285–297.
[90] Turkington, B., Statistical equilibrium measures and coherent states in two-dimensional turbulence. *Comm. Pure Appl. Math.*, 52 (1999), 781–809.
[91] Vekstein, G. E., Landau resonance mechanism for plasma and wind-generated water waves. *Amer. J. Phys.*, 66:10 (1998), 886–892.
[92] Villani, C., A review of mathematical topics in collisional kinetic theory, in *Handbook of Mathematical Fluid Dynamics*, Vol. I, pp. 71–305. North-Holland, Amsterdam, 2002.
[93] — Hypocoercivity. *Mem. Amer. Math. Soc.*, 202 (2009), iv+141.
[94] Wiechen, H., Ziegler, H. J. & Schindler, K., Relaxation of collisionless self-gravitating matter – the lowest energy state. *Mon. Not. R. Astr. Soc.*, 232 (1988), 623–646.
[95] Zhou, T., Guo, Y. & Shu, C.-W., Numerical study on Landau damping. *Phys. D*, 157 (2001), 322–333.

Clément Mouhot
University of Cambridge
DPMMS, Centre for Mathematical Sciences
Wilberforce Road
Cambridge CB3 0WA
U.K.
C.Mouhot@dpmms.cam.ac.uk

Cédric Villani
Institut Henri Poincaré & Université de Lyon
Institut Camille Jordan, Université Claude Bernard
43 Boulevard du 11 novembre 1918
FR-69622 Villeurbanne Cedex
France
villani@math.univ-lyon1.fr

Received December 10, 2009
Received in revised form July 10, 2011