

Landau Damping: Paraproducts and Gevrey Regularity

Jacob Bedrossian¹ · Nader Masmoudi² ·
Clément Mouhot³

Received: 31 March 2015 / Accepted: 8 March 2016 / Published online: 21 April 2016
© Springer International Publishing AG 2016

Abstract We give a new, simpler, but also and most importantly more general and robust, proof of nonlinear Landau damping on \mathbb{T}^d in Gevrey $-\frac{1}{s}$ regularity ($s > 1/3$) which matches the regularity requirement predicted by the formal analysis of Mouhot and Villani [67]. Our proof combines in a novel way ideas from the original proof of Landau damping Mouhot and Villani [67] and the proof of inviscid damping in 2D Euler Bedrossian and Masmoudi [10]. As in Bedrossian and Masmoudi [10], we use paraproduct decompositions and controlled regularity loss along time to replace the Newton iteration scheme of Mouhot and Villani [67]. We perform time-response estimates adapted from Mouhot and Villani [67] to control the plasma echoes and couple them to energy estimates on the distribution function in the style of the work Bedrossian and Masmoudi [10]. We believe the work is an important step forward in developing a systematic theory of phase mixing in infinite dimensional Hamiltonian systems.

Keywords Landau damping · Vlasov equations · Gevrey class · Plasma physics · Nonlinear stability

✉ Jacob Bedrossian
jacob@cscamm.umd.edu
Nader Masmoudi
masmoudi@cims.nyu.edu
Clément Mouhot
c.mouhot@dpms.cam.ac.uk

¹ University of Maryland, College Park, USA

² Courant Institute of Mathematical Sciences, New York, USA

³ Centre for Mathematical Sciences, University of Cambridge, Cambridge, UK

1 Introduction

Landau damping is a collisionless relaxation mechanism discovered by Landau [46] (after preliminary works of Vlasov [86]) predicting the decay of spatial oscillations in a plasma when perturbed around certain stable spatially homogeneous distributions. This effect is now considered fundamental to modern plasma physics (see e.g. [17,67,77,80]). It was later “exported” to galactic dynamics by Lynden-Bell [55,56] where it is thought to play a key role in the stability of galaxies. Moreover, Landau damping is one example of a more general effect usually referred to as “phase mixing”, which arises in many physical settings; see [9,10,67] and the references therein. See also [10,20,23,67] for a discussion about the differences and similarities with dispersive phenomena.

The physical model we will be focusing on is the nonlinear Vlasov equations, which is posed in the periodic box $x \in \mathbb{T}_L^d := [-L, L]^d$ with size $L > 0$:

$$\begin{cases} \partial_t f + v \cdot \nabla_x f + F(t, x) \cdot \nabla_v f = 0, \\ F(t, x) = -\nabla_x W *_x \left(\rho_f(t, x) - L^{-d} \int_y \rho_f(t, y) dy \right), \\ \rho_f(t, x) = \int_{\mathbb{R}^d} f(t, x, v) dv, \\ f(t = 0, x, v) = f_{in}(x, v), \end{cases} \tag{1.1}$$

with $f(t, x, v) : \mathbb{R} \times \mathbb{T}_L^d \times \mathbb{R}^d \rightarrow [0, \infty)$, the distribution function in phase space, and W the non-local interaction potential. We are interested in solutions of the form $f(t, x, v) = f^0(v) + h(t, x, v)$, where $f^0(v)$ is a spatially homogeneous background distribution and h is a mean-zero perturbation. If we denote simply the (perturbation) density $\rho(t, x)$, then the Vlasov equations can be written as

$$\begin{cases} \partial_t h + v \cdot \nabla_x h + F(t, x) \cdot \nabla_v h + F(t, x) \cdot \nabla_v f^0 = 0, \\ F(t, x) = -\nabla_x W *_x \rho(t, x), \\ \rho(t, x) = \int_{\mathbb{R}^d} h(t, x, v) dv, \\ h(t = 0, x, v) = h_{in}(x, v). \end{cases} \tag{1.2}$$

The potential $W(x)$ describes the mean-field interaction between particles; the cases of most physical interest are (1) Coulomb repulsive interactions $F = e \nabla_x \Delta_x^{-1} \rho$ between electrons in plasmas (where $e > 0$ is the electron charge-to-mass ratio) and Newtonian attractive interactions $F = -m \mathcal{G} \nabla_x \Delta_x^{-1} \rho$ between stars in galaxies (where $m > 0$ is the mass of the identical stars and \mathcal{G} is the gravitational constant). In Fourier variables (see later for the notation) these two cases correspond respectively to $\widehat{W}(k) = (2\pi)^{-2} e L^2 |k|^{-2}$ and $\widehat{W}(k) = -(2\pi)^{-2} m \mathcal{G} L^2 |k|^{-2}$. The former arises in plasma physics where (1.1) describes the distribution of electrons in a plasma interacting with a spatially homogeneous background of ions ensuring global electrical neutrality, after neglecting magnetic effects and ion acceleration. The latter arises in galactic dynamics

where (1.1) describes a distribution of stars interacting via Newtonian gravitation, neglecting smaller planetary objects as well as relativistic effects, and assuming Jean’s swindle (see [67] and the references therein).

By re-scaling t, x and W , we may normalize the size of the box to $L = 2\pi$ without loss of generality and write $\mathbb{T}^d = \mathbb{T}_{2\pi}^d$ (see Remark 2). For simplicity of notation and mathematical generality, we consider a general class of potentials with Coulomb/Newton representing the most singular examples. Specifically, we only require that there exists $C_W < \infty$ and $\gamma \geq 1$ such that

$$\forall k \in \mathbb{Z}^d \setminus \{0\}, \quad |\widehat{W}(k)| \leq \frac{C_W}{|k|^{1+\gamma}}. \tag{1.3}$$

This paper is concerned with the mathematically rigorous treatment of Landau damping for the full nonlinear mean-field dynamics, as initiated in [20,42,67].

1.1 Historical Context

The foundation of classical mechanics are the Newton’s laws and one of the most fundamental issues is the understanding of irreversibility at a macroscopical level based on these (reversible) laws. Landau damping is, as we shall see, part of the quest for the understanding of irreversibility at a statistical level, despite the fact that it is a *reversible* process. Consider the N particle system governed by binary interactions,

$$\dot{p}_j = -\frac{\partial \mathcal{H}}{\partial q_j}, \quad \dot{q}_j = \frac{\partial \mathcal{H}}{\partial p_j}, \quad \mathcal{H}(q, p) := \sum_{j=1}^N \frac{|p_j|^2}{2} + \sum_{j_1 \neq j_2} \phi(q_{j_1} - q_{j_2}) \tag{1.4}$$

with some radially symmetric interaction potential ϕ (note we are taking an electrostatic approximation by neglecting magnetic effects). Consider now indistinguishable particles with identical mass, (say normalized to 1), and reduce the canonical Hamiltonian variables to simply $q_j = x_j$ and $p_j = v_j$ (position and velocity of the j -th particle, both belonging to \mathbb{R}^3). Denote $X = (x_1, \dots, x_N)$ and $V = (v_1, \dots, v_N)$. *Liouville’s theorem* states that the distribution function $F^N(t, X, V)$ is constant along any trajectory in phase space. This translates into *Liouville equation*

$$0 = \frac{\partial F^N}{\partial t} + \{F^N, \mathcal{H}\} = \frac{\partial F^N}{\partial t} + V \cdot \nabla_X F^N - \sum_{j_1=1}^N \left(\sum_{j_2 \neq j_1} \nabla \phi(x_{j_2} - x_{j_1}) \cdot \nabla_{v_{j_1}} F^N \right). \tag{1.5}$$

Observe that this evolution equation preserves the *symmetry* of the distribution function (i.e. the fact that it is invariant under permutation of the particles). Liouville’s theorem also implies in particular that along this microscopic evolution the Boltzmann entropy is preserved:

$$\frac{d}{dt} \int_{\mathbb{R}^{6N}} F^N \log F^N \, dX \, dV = 0. \tag{1.6}$$

This highlights the fact that the microscopic evolution is *reversible* and that no loss of information occurs along time.

In the many-particle limit $N \rightarrow +\infty$, one tries to close an equation on the *reduced one-particle distribution function*

$$f(t, x_1, v_1) := \int_{\mathbb{R}^{6(N-1)}} F^N(t, X, V) dx_2 \cdots dx_N dv_2 \cdots dv_N. \quad (1.7)$$

Maxwell proposed in [61] the so-called *Boltzmann equation* for the statistical evolution of f . In the case of short-range interactions (e.g. ϕ has compact support), Boltzmann [14] then suggested a formal derivation in the limit $N \rightarrow +\infty$, with the so-called *Boltzmann-Grad* scaling: $Nr^2 = O(1)$, and discovered the celebrated *H-theorem*, which shows that this collisional equation is *irreversible* and thus quite different from (1.2).

In 1936, Landau [47] proposed a modification of the Boltzmann equation in order to model collisional gases of charged particles, such as electrons in a plasma. However, it was noted by Vlasov [85] that collisions are relatively weak in plasmas (see also [17]) and hence, for many phenomena, it makes sense to consider only *mean-field* electromagnetic fields. Assuming a *mean-field scale* $N\phi = O(1)$ where ϕ is the interaction potential between particles, he derived the *Vlasov-Poisson equations* (1.2).

Observe that in both cases (Coulomb or Newton forces) Vlasov-Poisson is *time-reversible*: it is invariant under the change of variable $(t, x, v) \rightarrow (-t, x, -v)$. This of course reflects the time-reversibility of the Newton equations at a microscopic level, which unlike collisional kinetic models, has been preserved in the mean-field limit. Hence, it follows that the entropy is *constant*.

A few years later, Landau [47] discovered a (then) mysterious “damping” effect: the exponential decay of longitudinal electrostatic waves, predicted by linearizing (1.2) around a homogeneous Maxwellian steady state. This provides a certain “asymptotic stability” for the spatially homogeneous steady state in the sense of the asymptotic stability of the spatial density. It was later argued by Lynden-Bell [55] that a similar phenomenon was occurring in galactic dynamics, where the gas of electrons interacting by electro-static forces is replaced by a “gas of stars” interacting by gravitation forces. See §2.1 below for review for how to derive this on the linear level. In the case of plasmas, it was confirmed experimentally in [57].

Landau damping at first appears at odds with the reversibility of (1.2). It was van Kampen in [81] who pointed out that Landau damping is actually a *phase-mixing* effect, that is, the streaming of particles creates rapid oscillations (in v) of the distribution function which are averaged away by the v integral in the non-local law for F . Information is *not* lost, but is simply transferred to the small scales in v , and hence while the behavior appears to be irreversible, in fact, it is completely reversible, as demonstrated by the weakly nonlinear *plasma echoes* [58]. Mixing as a reversible relaxation mechanism had already been discovered in the context of 2D Euler by Orr in 1907 [71] (where the linear problem is much easier), which is now known as *inviscid damping* (see also [10]). This effect was also demonstrated to be reversible by ‘Euler echo’ experiments [88, 89].

The original works in physics neglected nonlinear effects, which lead to some speculation, and Landau himself was seemingly very prudent about the exact validity of his analysis for the nonlinear equation and remained silent on this point. The mathematically rigorous theory of the linear damping was pioneered by Backus [4] and Penrose [73] as discussed above, and further clarified by many mathematicians, see e.g. [23,59]. Penrose seemed optimistic that Landau damping should occur in the nonlinear equations, and highlighted the fact that near equilibrium, the linear evolution would cause the nonlinear electric field to decay and that the nonlinear equations would be asymptotically linear. Such a situation is very reminiscent of *scattering* in dispersive equations, as discussed in more detail in e.g. [10,20,23,67]. However, Backus [4] made the important point that while the electric field may decay, the gradients $\nabla_v f$ will grow, and hence perhaps the linearization will cease to make good predictions.

The first nonlinear results were obtained by [20,42] which showed that Landau damping was at least possible in (1.2) for analytic data (see also [52] for a negative result). However, Landau damping was not shown to hold for all small data until [67]. The results therein hold in analytic spaces or in close to analytic Gevrey spaces [29] and the authors made heuristic conjectures about the minimal regularity required. The proof involved an intricate use of Eulerian and Lagrangian coordinates combined with a global-in-time Newton iteration reminiscent of the proof of the KAM theorem (see therein for a discussion about broader analogies with KAM theory). A key step in the analysis of [67] was controlling the potentially destabilizing influence of the plasma echoes, a weakly nonlinear effect discovered by Malmberg et. al. in [58]. The basic mechanism of the plasma echo is as follows. The force field is damped due to mixing in phase space transferring information from large spatial scales to small scales in the velocity distribution, however, as mixing is time-reversible, un-mixing will then create growth in the force field by transferring information back from the velocity distribution to the large spatial scales. Therefore, any nonlinear effect which transfers information to modes which are un-mixing will lead to a large force field in the future when that information un-mixes (hence ‘echo’). These can potentially chain into a cascade of echoes, as demonstrated experimentally in plasmas [58] and 2D Euler [88,89], hence we cannot expect to rule them, or the potential cascade, out. See [67] and below for more discussion of the role this resonance plays in the nonlinear theory. In fluid mechanics, the growth caused by un-mixing was noticed earlier by Orr [71] and is known there as the *Orr mechanism*; see [10] for more discussion.

1.2 Objective of New Work

In this work, we provide a new proof of Landau damping for (1.2) that nearly obtains the “critical” Gevrey regularity predicted in [67]. Obviously, both works have common themes (as these are coming from the physics), namely,

- (1) the abstract linear stability condition is the same,
- (2) the physical mechanism of phase mixing transferring regularity to decay is the same,
- (3) the isolation and control of the plasma echoes is still the main challenge.

However, on a mathematical level, the two proofs are quite different; see §2 for a full discussion.

Our proof is shorter, and arguably simpler—assuming some knowledge in paradifferential calculus—than that of [67]. More importantly, this new approach is more robust for developing a general theory. The paper [67] was a breakthrough that potentially opens the way for mathematical studies on phase mixing for nonlinear PDEs, but the main drawback was that the conceptual and technical construction of the proof was so long and mixed so many different abstract viewpoints that it has proved hard to use in a wider context. Our new approach, by avoiding the mixture of Lagrangian and Eulerian viewpoints, avoiding the Newton scheme, and using paradifferential energy estimates which take full advantage of time varying multipliers, makes the proof significantly less “rigid”. There is a pressing need for more flexible and powerful analytical tools as the ‘basic’ situation studied in [67] is only a tiny part of what remains to be understood and we believe that our work is a key step to addressing many interesting and qualitatively different questions ahead. Indeed, the tools developed in this paper have already been adapted to prove a nonlinear Landau damping result in relativistic plasmas [87]. Other examples where our methods may prove useful are damping in un-confined geometries by separating scales in frequencies, geometries and cyclotron effects imposed by external magnetic fields or more subtle self-created magnetic effects, damping in weakly collisional plasmas and finally the stability of self-created gravitational geometries [49,66] or nonlinear BGK waves [13].

Our proof combines the viewpoint in the original work [67] with the recent work on the 2D Euler equations in [10]. This latter work proves the asymptotic stability (in a suitable sense) of sufficiently smooth shear flows near Couette flow in $\mathbb{T} \times \mathbb{R}$ via inviscid damping (the hydrodynamic analogue of Landau damping). The proof in [10] uses a number of ideas specific to the structure of 2D Euler, however, some aspects of the viewpoint taken therein will be useful here as well (when suitably combined with ideas from [67]).

One of the main ingredients of our proof, employed also in [10], is the use of the paradifferential calculus to split nonlinear terms into one that carries the transport structure and another that is analogous to the nonlocal interaction term referred to as ‘reaction’ in [67]. It has been long believed that Nash-Moser or similar Newton iterations (see the classical work [64,65]) can generally be replaced by a more standard fixed point argument if one uses better all of the structure in the equation. This has been the case in most examples in the literature (e.g. Nash’s isometric embedding theorem [68] by Günther [34–36]), with maybe the only exception being KAM theory. Other examples can be found in [40], where Hörmander specifically points out paradifferential calculus as a useful tool in this context.

It is well known that one of the main physical barriers to Landau damping is nonlinear particle trapping, whereby particles are trapped in the potential wells of (say) electrostatic waves. Exact traveling wave solutions of this type exist in plasma physics and are known as BGK waves [13]. They were used by Lin and Zeng in [52] to show that one needs at least H^σ with $\sigma > 3/2$ regularity on the distribution function to expect Landau damping in (1.2) in the neighborhood of Landau-stable stationary solutions (see Definition 1.1 below). The plasma echoes provide a natural nonlinear bootstrap mechanism by which the electrostatic waves can persist long enough to

trap particles. After modding out by particle free streaming, the bootstrap appears as a cascade of information to high frequencies and the regularity requirement of Gevrey- $(2 + \gamma)^{-1}$ comes from formal ‘worst-case’ calculations carried out in [67]. Mathematically, the same requirement arises here in §6. Lower regularity is an open question: it seems plausible that Theorem 1 is false for all Gevrey- $\frac{1}{s}$ with $s < (2+\gamma)^{-1}$, however there might be additional cancellations that could allow it to hold in lower regularities. Finally, in weakly collisional plasmas, the requirement could perhaps be relaxed in some suitable sense (e.g. permitting data which is Gevrey plus a smaller rough contribution that will be instantly regularized, as in the recent work on the Navier-Stokes equations [11]).

It is well known that many areas in physics present striking analogies with each other and many important developments have come from a good understanding of these analogies. This work is an example where the analogy between 2D incompressible Euler and Vlasov-Poisson proved fruitful. The connection between inviscid damping and Landau damping has been acknowledged by a number of authors, for example [6, 16, 19, 32, 51, 79]. Both have similar weakly nonlinear echoes [82, 83, 88, 89], moreover, the work of Lin and Zeng [51, 52] and [9] show that particle trapping and vortex roll-up may in fact be essentially the same ‘universal’ over-turning instability. On a more general level, both systems are conservative transport equations governed by a single scalar quantity (the vorticity in the 2D Euler equation and the distribution function in Vlasov-Poisson equations). Both equations have a large set of stationary solutions: the shear flows for the Euler equation and spatially homogeneous distributions for Vlasov-Poisson equations being the simplest. Each can be viewed as Hamiltonian systems and variational methods have been used for both to provide non-linear stability results in low-regularity spaces (i.e. functional spaces invariant under the free-streaming operator): we refer for instance, among a huge literature, to [3] in the context of the Euler equation, and to the recent remarkable results [37, 49] in the context of the gravitational Vlasov-Poisson equation (see also the review paper [66] and the references therein). One can also derive the incompressible Euler system from the Vlasov-Poisson system in the quasi-neutral regime for cold electrons (vanishing initial temperature) [18, 60].

1.3 The Main Result

In what follows, we denote the Gevrey- ν^{-1} norms, $\nu \in (0, 1]$, with Sobolev corrections $\sigma \in \mathbb{R}$ (see §3.1 for Fourier analysis conventions and §1.3.1 for other notations),

$$\|f\|_{\mathcal{G}^{\lambda, \sigma; \nu}}^2 = \sum_{k \in \mathbb{Z}^d} \int_{\eta} |\hat{f}_k(\eta)|^2 \langle k, \eta \rangle^{\sigma} e^{2\lambda \langle k, \eta \rangle \nu} d\eta, \tag{1.8a}$$

$$\|\rho(t)\|_{\mathcal{F}^{\lambda, \sigma; \nu}}^2 = \sum_{k \in \mathbb{Z}^d} |\hat{\rho}_k(t)|^2 \langle k, kt \rangle^{\sigma} e^{2\lambda \langle k, kt \rangle \nu}. \tag{1.8b}$$

When $\sigma = 0$ or $\nu = s$, (defined below) these parameters are usually omitted.

Recall the sufficient linear stability condition (L) introduced in [67] for analytic background distributions, which we slightly adapt to the norms we are using.

Definition 1.1 Given a homogeneous analytic distribution $f^0(v)$ we say it satisfies *stability condition (L)* if there exist constants $C_0, \bar{\lambda}, \kappa > 0$ and an integer $M > d/2$ such that

$$\sum_{\alpha \in \mathbb{N}^d : |\alpha| \leq M} \|v^\alpha f^0\|_{\mathcal{G}^{\bar{\lambda},1}}^2 \leq C_0, \tag{1.9}$$

and for all $\xi \in \mathbb{C}$ with $\text{Re } \xi < \bar{\lambda}$,

$$\inf_{k \in \mathbb{Z}^d} |\mathcal{L}(\xi, k) - 1| \geq \kappa, \tag{1.10}$$

where \mathcal{L} is defined by the following (where $\bar{\xi}$ denotes the complex conjugate of ξ),

$$\mathcal{L}(\xi, k) = - \int_0^\infty e^{\bar{\xi}|k|t} \widehat{f^0}(kt) \widehat{W}(k) |k|^2 t \, dt. \tag{1.11}$$

Remark 1 Note (1.9) implies the integral in (1.11) is absolutely convergent by the $H^{d/2+} \hookrightarrow C^0$ embedding theorem.

In [67], it is shown that (L) is practically equivalent to several well known stability criteria in plasma and galactic dynamics (see §4.3 for a completion of the proof that the Penrose condition [73] implies (L)).

We prove the following nonlinear Landau damping result, which for Coulomb/Newton interaction nearly obtains the Gevrey-3 regularity predicted heuristically in [67]. In §2 we give the outline of the proof and discuss the relationship with the original proof in [67] and the proof of inviscid damping in 2D Euler [10].

Theorem 1 *Let f^0 be given which satisfies stability condition (L) with constants $M, \bar{\lambda}, C_0$ and κ . Let $\frac{1}{(2+\gamma)} < s \leq 1$ and $\lambda_0 > \lambda' > 0$ be arbitrary (if $s = 1$ we require $\bar{\lambda} > \lambda_0$). Then there exists an $\epsilon_0 = \epsilon_0(d, M, \bar{\lambda}, C_0, \kappa, \lambda_0, \lambda', s)$ such that if h_{in} is mean zero (that is, $\int \int h_{\text{in}}(x, v) dx dv = 0$) and*

$$\sum_{\alpha \in \mathbb{N}^d : |\alpha| \leq M} \|v^\alpha h_{\text{in}}\|_{\mathcal{G}^{\lambda_0, s}}^2 < \epsilon^2 \leq \epsilon_0^2,$$

then there exists a mean zero $h_{+\infty} \in \mathcal{G}^{\lambda', s}$ satisfying,

$$\|h(t, x + vt, v) - h_{+\infty}(x, v)\|_{\mathcal{G}^{\lambda', s}} \lesssim e^{-\frac{1}{2}(\lambda_0 - \lambda')t^s}, \tag{1.12a}$$

$$\|\rho(t)\|_{\mathcal{F}^{\lambda', s}} \lesssim e^{-\frac{1}{2}(\lambda_0 - \lambda')t^s}, \tag{1.12b}$$

for all $t \geq 0$.

Remark 2 Through the rescaling on W , our estimate of ϵ_0 in Theorem 1 is a decreasing function of the side-length of the original torus, L . That is, the restriction for nonlinear

stability becomes more stringent as the confinement is removed. Moreover, through the rescaling on time, Theorem 1 predicts damping on a characteristic time-scale of $O(L)$. See [30, 31, 67, 84] for more discussion about what can happen without confinement.

Remark 3 It is immediate to deduce estimates on the complete distribution $f(t, x, v) = f^0(v) + h(t, x, v)$ of the original equation (1.1). In particular, Theorem 1 shows that all solutions to (1.1) with non-trivial spatial dependence close to f^0 satisfy $\|f(t)\|_{H^N} \approx \langle t \rangle^N$ (the same as free transport).

Remark 4 From (1.12a) we have the homogenization $h(t, x, v) \rightharpoonup \langle h_{+\infty}(\cdot, v) \rangle_x$ and the exponential decay of the electrical or gravitational field $F[h](t, x)$. Note however that the estimates on h and ρ are more precise as they show decay rates which increase with k (the spatial frequency).

Remark 5 The theorem also holds backwards in time for some $h_{-\infty} \in \mathcal{G}^{\lambda'; s}$.

Remark 6 The asymptotic distribution function $f_{+\infty}(x, v) := f^0(v) + h_{+\infty}(x, v)$ depends on the entire nonlinear evolution, however, at least one can show $f_{+\infty}$ is within $O(\varepsilon^2)$ of the distribution predicted by the linear theory in $\mathcal{G}^{\lambda''; s}$ for any $\lambda'' < \lambda'$ (similar to [67]). See §7 for a sketch.

Remark 7 Though f^0 is analytic, the statement shows the asymptotic stability of homogeneous distributions within a small neighborhood of f^0 in Gevrey- s^{-1} since we are really making a perturbation of $f^0(v) + \langle h_{\text{in}}(\cdot, v) \rangle_x$. However, the size of that neighborhood depends on the parameters in Definition 1.1, so it still must be close to an analytic distribution satisfying **(L)**.

Remark 8 Requiring h_{in} to be average zero does not lose any generality. Indeed, if h_{in} were not mean zero we may apply Theorem 1 to $\tilde{f}^0(v) = f^0(v) + \langle h_{\text{in}}(\cdot, v) \rangle_{<1}$ and $\tilde{h}_{\text{in}} = h_{\text{in}} - \langle h_{\text{in}}(\cdot, v) \rangle_{<1}$, where $g_{<1}$ denotes projection onto frequencies less than one. Since **(L)** is open, for ε_0 sufficiently small, \tilde{f}^0 still satisfies **(L)** with slightly adjusted parameters C_0 and κ .

Remark 9 If W is in the Schwartz space, then Theorem 1 holds for all $0 < s < 1$ (although ε_0 goes to zero as $s \searrow 0$). The heuristics of [67] suggest that even in the case of analytic W , we cannot hope to work in the Sobolev scale without major new ideas, if such a result holds at all.

1.3.1 Notation and Conventions

We denote $\mathbb{N} = \{0, 1, 2, \dots\}$ (including zero) and $\mathbb{Z}_* = \mathbb{Z} \setminus \{0\}$. For $\xi \in \mathbb{C}$ we use $\bar{\xi}$ to denote the complex conjugate. We will denote the ℓ^1 vector norm $|k, \eta| = |k| + |\eta|$, which by convention is the norm taken in our work. We denote

$$\langle v \rangle = \left(1 + |v|^2\right)^{1/2}.$$

We use the multi-index notation: given $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$ and $v = (v_1, \dots, v_d) \in \mathbb{R}^d$ then

$$v^\alpha = v_1^{\alpha_1} v_2^{\alpha_2} \dots v_d^{\alpha_d}, \quad D_\eta^\alpha = (i \partial_{\eta_1})^{\alpha_1} \dots (i \partial_{\eta_d})^{\alpha_d}.$$

We denote Lebesgue norms for $p, q \in [1, \infty]$ and a, b either in $\mathbb{R}^d, \mathbb{Z}^d$ or \mathbb{T}^d as

$$\|f\|_{L_a^p L_b^q} = \left(\int_a \left(\int_b |f(a, b)|^q db \right)^{p/q} da \right)^{1/p}$$

and Sobolev norms (usually applied to Fourier transforms) as

$$\|\hat{f}\|_{H_\eta^M}^2 = \sum_{\alpha \in \mathbb{N}^d; |\alpha| \leq M} \|D_\eta^\alpha \hat{f}\|_{L_\eta^2}^2.$$

We will often use the short-hand $\|\cdot\|_2$ for $\|\cdot\|_{L_{z,v}^2}$ or $\|\cdot\|_{L_v^2}$ depending on the context.

See §3 for the Fourier analysis conventions we are taking. A convention we generally use is to denote the discrete x (or z) frequencies as subscripts. We use Greek letters such as η and ξ to denote frequencies in \mathbb{R}^d and lowercase Latin characters such as k and ℓ to denote frequencies in \mathbb{Z}^d . Another convention we use is to denote N, N' as dyadic integers $N \in \mathbb{D}$ where

$$\mathbb{D} = \left\{ \frac{1}{2}, 1, 2, \dots, 2^j, \dots \right\}.$$

When a sum is written with indices N or N' it will always be over a subset of \mathbb{D} . This will be useful when defining Littlewood-Paley projections and paraproduct decompositions, see §3. Given a function $A = A(k, \eta) \in L^\infty$, we define the Fourier multiplier operator $\mathcal{A}g = A(\nabla_{z,v})g$ by

$$\left(A(\widehat{\nabla_{z,v}})g \right)_k(\eta) := A((ik, i\eta))\hat{g}_k(\eta),$$

and we also define the corresponding operator $\mathcal{A}_l(\rho) = A(\nabla_{z,zt})\rho =: A_l(\nabla_z)\rho$ when acting on functions of x only along the frequencies space (k, kt) given by the velocity averages along the moving frame:

$$A(\widehat{\nabla_{z,zt}})\rho_k(t) = A((ik, ikt))\hat{\rho}_k(t), \quad \mathcal{A}_l(\nabla_z) = A(\nabla_{z,zt}). \tag{1.13}$$

We use the notation $f \lesssim g$ when there exists a constant $C > 0$ independent of the parameters of interest such that $f \leq Cg$ (we analogously define $f \gtrsim g$). Similarly, we use the notation $f \approx g$ when there exists $C > 0$ such that $C^{-1}g \leq f \leq Cg$. We sometimes use the notation $f \lesssim_\alpha g$ if we want to emphasize that the implicit constant depends on some parameter α .

2 Outline of the Proof

2.1 Linear Behavior

In [85], Vlasov sought to understand (1.2) for small h by first linearizing the equations around the steady state $f^0(v) = e^{-\frac{|v|^2}{2T}} / (2\pi T)^{3/2}$:

$$\begin{cases} \partial_t h + v \cdot \nabla_x h + F(t, x) \cdot \nabla_v f^0 = 0, \\ F(t, x) = -\nabla_x W *_{x} \rho(t, x), \\ h(t = 0, x, v) = h_{\text{in}}(x, v). \end{cases} \quad (2.1)$$

However, Vlasov searched for plane wave solutions via a normal mode method, and in general the spectrum of (2.1) turns out to be purely continuous. Landau understood that the normal mode method was bound to fail for (2.1) (indeed he described the approach of Vlasov as being “without any foundation”) and instead argued for studying and solving the *Cauchy problem* for given analytic initial data, i.e. the “vibrations for a given initial distribution”. He then solved (2.1) via a Fourier-Laplace transform method. Interestingly, a very similar exchange had already occurred in fluid mechanics long before Vlasov and Landau: when studying the stability of shear flows, Lord Rayleigh had also attempted to take a normal mode approach [76], but it was subsequently understood by Lord Kelvin and Orr [43, 71] that this would not work, for reasons that are indeed the same. However, we do note that by properly dealing with “singular eigenfunctions” for the continuous spectrum, one can actually solve (2.1) with a spectral method, see e.g. [6–8, 62, 63, 81] for more discussion.

Landau’s argument is subtle and strongly reminiscent of the by-now classical Gearhart-Prüss-Greinier Theorem (see for instance [24, Theorem V.1.11, page 302]; see also below in §4). However, as pointed by Penrose [73], Backus [4] and others, it is not completely rigorous. For our discussion of linear stability, let us follow a variant of the approach of Penrose, whose method is the most straightforward and flexible. By taking the Fourier transform of (2.1) in x and then integrating in v and time, one derives the set of decoupled *Volterra equations*:

$$\hat{\rho}(t, k) = \hat{h}_{\text{in}}(k, kt) - \int_0^t \widehat{W}(k) k \cdot k t \widehat{f^0}(kt) \hat{\rho}(s, k) ds, \quad (2.2)$$

where here $\hat{\rho}(t, k)$ denotes the Fourier transform in x , $\hat{h}_{\text{in}}(k, \eta)$ denotes the Fourier transform in both x and v . Upon extending all the functions by zero for negative times and (formally) taking the Fourier transform in time, we have

$$\tilde{\rho}(\omega, k) = A(\omega, k) + \mathcal{L} \left(\frac{i\omega}{|k|}, k \right) \tilde{\rho}(\omega, k), \quad (2.3)$$

where \mathcal{L} is given above in (1.11) and

$$A(\omega, k) = \int_0^\infty e^{-i\omega t} \hat{h}_{\text{in}}(k, kt) dt.$$

Therefore, if (1.10) holds then we can write

$$\tilde{\rho}(\omega, k) = \frac{A(\omega, k)}{1 - \mathcal{L}\left(\frac{i\omega}{|k|}, k\right)},$$

then using that the Fourier transform is unitary on L^2 we have, at least,

$$\int_0^\infty |\hat{\rho}(t, k)|^2 dt \lesssim \kappa^{-1} \int_0^\infty \left| \hat{h}_{in}(k, kt) \right|^2 dt.$$

The RHS is finite as soon as $|v|^{\frac{d-1}{2}} + h_{in} \in L^2$ by taking the Sobolev space restriction on the Fourier side; see Lemma 3.4 below. To obtain higher decay rates one can use several methods. To obtain polynomial rates of decay, the simplest approach is to take derivatives in ω of (2.3). For example, one derivative is enough to prove that $\hat{\rho}(t, k)$ is integrable in time, allowing to prove that solutions to (2.1) are asymptotically ballistic, i.e. $h(t, x, v) \sim h_\infty(x - tv, v)$ for some h_∞ as $t \rightarrow \infty$. Hence, by considering $h_\infty(\widehat{x - tv}, v)(k, \eta) = \widehat{h_\infty}(k, \eta + kt)$, we see that the information in the distribution function is being moved to higher frequencies/smaller scales in v . Moreover, we can see that $h(t) \rightarrow \langle h_\infty \rangle_x$, and so the evolution is weakly (but not strongly) returning to a spatially homogeneous equilibrium. For analytic regularity, the approach taken in [67] was to multiply (2.2) by $e^{|k|t}$ and deduce an estimate on $\Phi(t, k) = e^{|k|t} \hat{\rho}(t, k)$. This inspires what is done below in §4 in Gevrey class.

Several comments are left to be made. First, the above argument is only formal, as in order to take the Fourier transform in time to deduce (2.3), one already needed to have some a priori decay estimate on $\hat{\rho}(t, k)$. This can be resolved in several ways, see §4 below for a full discussion. Second, Penrose derived a reasonably practical way of checking (1.10) for general background distributions, known as the *Penrose condition*, which is also known to be essentially sharp [52]. See §4 below for a proof of this.

2.2 Summary, Weakly Nonlinear Heuristics, and Comparison with Original Proof [68]

Next, let us consider how to prove Theorem 1 and extend the linear results to the nonlinear level. Landau damping predicts that the solution evolves by kinetic free transport as $t \rightarrow \infty$:

$$h(t, x, v) \sim h_\infty(x - vt, v).$$

We ‘mod out’ by the characteristics of free transport and work in the coordinates $z = x - vt$ by making the definition $g(t, z, v) := h(t, z + tv, v)$. Then (1.12a) becomes equivalent to $g(t) \rightarrow h_\infty$ strongly in Gevrey $-\frac{1}{5}$. This coordinate shift was used in [20, 42] and is related to the notion of ‘gliding regularity’ used in [67] (although we will not have to use the time-shifting tricks employed therein). Moreover, it is also

closely related to the notion of ‘profile’ used in nonlinear dispersive equations (see e.g. [28]). A related coordinate change is also used in [10].

A straightforward computation gives the evolution equation:

$$\begin{cases} \partial_t g + F(t, z + vt) \cdot (\nabla_v - t \nabla_z)g + F(t, z + vt) \cdot \nabla_v f^0 = 0, \\ g(t = 0, z, v) = h_{in}(z, v). \end{cases} \tag{2.4}$$

Note that the density satisfies $\hat{\rho}_k(t) = \hat{g}_k(t, kt)$; by the $H^{d/2+} \hookrightarrow C^0$ embedding theorem and the requirement $M > d/2$, this formula at least makes sense pointwise in time provided $\sum_{\alpha \leq M} \|v^\alpha g\|_2$ is finite. Moreover, from this formula we see that a uniform bound on the regularity of f translates directly into decay of the density: this is precisely the phase mixing mechanism. Hence, to prove (1.12b), we are aiming for a uniform-in-time bound on the regularity in a velocity polynomial-weighted space (so that we may restrict the Fourier transform using $H^{d/2+} \hookrightarrow C^0$).

All of our analysis is on the Fourier side; taking Fourier transforms and using $[F(\widehat{F(t, z + vt)})](t, k, \eta) = -(2\pi)^d i k \widehat{W}(k) \hat{\rho}_k(t) \delta_{\eta=kt}$, (2.4) becomes

$$\begin{aligned} \partial_t \hat{g}_k(t, \eta) + \hat{\rho}_k(t) \widehat{W}(k) k \cdot (\eta - tk) \hat{f}^0(\eta - kt) \\ + \sum_{\ell \in \mathbb{Z}_*^d} \hat{\rho}_\ell(t) \widehat{W}(\ell) \ell \cdot [\eta - tk] \hat{g}_{k-\ell}(t, \eta - t\ell) = 0. \end{aligned} \tag{2.5}$$

Note in the summation that $\eta - tk = (\eta - t\ell) - t(k - \ell)$. By the formula $\hat{\rho}_k(t) = \hat{g}_k(t, kt)$, (2.5) is a closed equation for f , however, we will derive a separate integral equation for $\rho(t)$ and look at (g, ρ) as a coupled system (similar to [67]). Integrating (2.5) in time and evaluating at $\eta = kt$ gives

$$\begin{aligned} \hat{\rho}_k(t) = \hat{h}_{in}(k, kt) - \int_0^t \hat{\rho}_k(\tau) \widehat{W}(k) k \cdot k(t - \tau) f^0(k(t - \tau)) \, d\tau \\ - \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \hat{\rho}_\ell(\tau) \widehat{W}(\ell) \ell \cdot k(t - \tau) \hat{g}_{k-\ell}(\tau, kt - \tau\ell) \, d\tau. \end{aligned} \tag{2.6}$$

As in [67], our goal now is to use the system (2.5)–(2.6) to derive a uniform control on the regularity of g in the moving frame (referred to as ‘gliding regularity’ in [67]). The linear term in (2.6) is handled with the help of the abstract stability condition (L); the difference here with [67] is that we must adapt this to the Gevrey norms we are using, which is done using a slightly technical decomposition technique similar to one which appeared in [67] to treat Gevrey data (carried out below in §4). The main point of departure from [67] is our treatment of the nonlinear terms in (2.5)–(2.6). There are schematically two mechanisms of potential loss of regularity in (2.5)–(2.6) and one potential loss of localization in velocity space:

1. Equation (2.5) describes a transport structure in $\mathbb{T}^d \times \mathbb{R}^d$, and hence we can expect this to induce the loss of regularity usually associated with transport by controlled coefficients. A different loss of regularity occurs in (2.6) due to the

$k(t - \tau)\hat{g}_{k-\ell}(\tau, kt - \ell\tau)$: here there is a derivative of g appearing but no transport structure to take advantage of. However, we still refer to this effect as ‘transport’ and remark that it seems related to the beam instability [17].

2. We can see from (2.6) that $\hat{\rho}_\ell(\tau)$ has a strong effect on $\hat{\rho}_k(t)$ when $kt \sim \ell\tau$, which is referred to as *reaction*. These nonlinear resonances are exactly the plasma echoes of [58], and arise due to interaction with the oscillations in the velocity variable. These effects can potentially amplify high frequencies (e.g. *costing regularity*) at localized times of strong interaction. A related effect will appear in (2.5) when ρ forces g via interaction with lower frequencies, and we will refer to this as reaction as well. It was shown in [67] via formal heuristics that an infinite cascade of echoes could lose Gevrey- $(2 + \gamma)^{-1}$ regularity, and hence the restriction $s > (2 + \gamma)^{-1}$.
3. The density ρ is a restriction of the Fourier transform of g and the nonlinear terms in (2.5) and (2.6) each involve Fourier restrictions. This issue was treated in [67] by adapting carefully the norms used in order to keep under control some L^1 -based norms of regularity.

The proof of [67] employs a global-in-time Newton scheme which loses a decreasing amount of analytic regularity at each step. The linearization of the Newton scheme provides a natural way to isolate transport effects from reaction effects. The transport is treated by Lagrangian methods to estimate analytic regularity along the characteristic trajectories. The reaction effects are treated via time-integrated estimates on (2.6) that account carefully for the localized, time-delayed effects of the plasma echoes. The proof treats (2.5) and (2.6) as a coupled system in the sense that the main estimates are two coupled but separate controls, one on g and one on ρ . To extend the results to Gevrey regularity, a frequency decomposition of the initial data is employed so that at each step in the scheme everything is analytic.

Here the two different mechanisms of *transport* and *reaction* are recovered by a rather different approach, employed recently in the proof of inviscid damping [10]. We use a paraproduct decomposition in order to split the bilinear terms as

$$G_1 G_2 = (G_1)_{\text{lower}}(G_2)_{\text{higher}} + (G_1)_{\text{higher}}(G_2)_{\text{lower}} + (G_1 G_2)_{\text{similar frequencies}}.$$

In a general sense, one of the first two terms on the RHS will capture the transport effects and the other will capture the reaction effects. The last term is a remainder which roughly corresponds to the quadratic error term in the Newton iteration. Indeed, the paraproduct decomposition can be thought of formally linearizing the evolution of higher frequencies around lower frequencies.

As in [10], the transport terms in (2.5) are treated via an adaptation of the Gevrey energy methods of [25, 50]. The essential content is a *commutator estimate* to take advantage of the cancellations inherent in the natural transport structure. For this step to work we need to use L^2 based norms, and so to deal with the Fourier restrictions, we use norms with polynomial weights in velocity.

The reaction effects in (2.6) are treated here by making use of a refinement of the integral estimates on ρ of [67] but with some important conceptual changes inspired from [10]: the loss of Gevrey regularity can occur along time rather than iteratively in a Newton scheme. Together with the paraproduct decompositions, this allows us to significantly shorten the proof.

Since we do not use a Newton scheme, we have an additional new constraint: we are not allowed to lose any derivatives in our coupled estimates on ρ and g , a problem due to the derivative of g appearing in (2.5)–(2.6). However, since regularity can be traded for decay, we solve this problem by propagating controls on both “high” and “low” norms of regularity – the “high” ones being mildly growing in time, a general scheme which is common in the study of wave and dispersive equations (see e.g. [28,44,53] and the references therein).

2.3 Gevrey Functional Setting

As discussed in the previous section, our goal now is to use the system (2.5)–(2.6) to derive a uniform control on the regularity of g . Unlike the norms used in [10] and [67] we will only need the standard norm $\mathcal{G}^{\lambda,\sigma;s}$ (and the variant $\mathcal{F}^{\lambda,\sigma;s}$) defined in (1.8) with time-dependent $\lambda(t)$. For future notational convenience, we define the Fourier multiplier $A(t)$ such that $\|A(t)g\|_2 = \|g\|_{\mathcal{G}^{\lambda(t),\sigma;s}}$:

$$A_k(t, \eta) = e^{\lambda(t)\langle k, \eta \rangle^s} \langle k, \eta \rangle^\sigma, \tag{2.7}$$

where $\sigma > d + 8$ is fixed and $\lambda(t)$ is an index (or ‘radius’) of regularity which is decreasing in time. In the sequel, we also denote for $\sigma' \in \mathbb{R}$:

$$A_k^{(\sigma')}(t, \eta) = e^{\lambda(t)\langle k, \eta \rangle^s} \langle k, \eta \rangle^{\sigma+\sigma'}.$$

We will choose λ and s so as to absorb the potential loss of regularity due to the plasma echoes. In particular, the choice $s > (2 + \gamma)^{-1}$ will be necessary to ensure that the nonlinear plasma echoes do not destabilize the phase mixing mechanism. This restriction is used only in equations (6.6) and (6.9) of §6. Additionally, in order to absorb the loss of regularity from the echoes, we will need to choose $\lambda(t)$ to decay slowly in time; this restriction is also only used in (6.6) and (6.9) of §6. It will suffice to make the following choices:

$$s > (2 + \gamma)^{-1}, \quad \alpha_0 = \frac{\lambda_0}{2} + \frac{\lambda'}{2}, \quad a = \frac{(2 + \gamma)s - 1}{(1 + \gamma)},$$

(note $0 < a < s$ if $s < 1$) and

$$\lambda(t) = \frac{1}{8} (\lambda_0 - \lambda') (1 - t)_+ + \alpha_0 + \frac{1}{4} (\lambda_0 - \lambda') \min \left(1, \frac{1}{t^a} \right). \tag{2.8}$$

Then $\alpha_0 \leq \lambda(t) \leq 7\lambda_0/8 + \lambda'/8$ and the derivative never vanishes:

$$\dot{\lambda}(t) \lesssim -\frac{a(\lambda_0 - \lambda')}{\langle t \rangle^{1+a}}. \tag{2.9}$$

Norms such as $\mathcal{G}^{\lambda(t),\sigma;s}$ are common when dealing with analytic or Gevrey regularity, for example, see the works [10,21,22,25,27,45,50,67]. The Sobolev correction σ is included mostly for technical convenience. This correction allows to avoid needing to pay Gevrey regularity where Sobolev regularity would suffice; for example, as $\lambda(t) \gtrsim 1$ and $s < 1$, $\mathcal{G}^{\lambda(t),0;s}$ is an algebra, however it is simpler to use that $H^\sigma(\mathbb{T}^d \times \mathbb{R}^d)$ is an algebra for $\sigma > d$ instead. In order to study the analytic case, $s = 1$, we would need to add an additional Gevrey- $\frac{1}{\zeta}$ correction with $(2 + \gamma)^{-1} < \zeta < 1$ as an intermediate regularity (and define a in terms of ζ) so that we may take advantage of beneficial properties of Gevrey norms (see, for example, Lemma 3.3). For the duration of the proof we assume $s < 1$ and do not address the additional technical minor issue in the limit case $s = 1$.

The reason for using $\langle \cdot \rangle$ in the definition of A_k and $A_k^{(\sigma')}$ in (2.7), as opposed to $|\cdot|$, is so that for all $\alpha \in \mathbb{N}^d$ and $\sigma' \in \mathbb{R}$,

$$\left| D_\eta^\alpha A_k^{(\sigma')}(t, \eta) \right| \lesssim_{|\alpha|,\lambda_0,\sigma'} \frac{1}{\langle k, \eta \rangle^{|\alpha|(1-s)}} A_k^{(\sigma')}(t, \eta), \tag{2.10}$$

which is useful when estimating velocity moments.

2.4 Uniform in Time Regularity Estimates

In this section we set up the continuity argument we use to derive a uniform bound on (ρ, g) . In order to ensure the formal computations are rigorous, we first regularize the initial data to be analytic. The following standard lemma provides local existence of a unique analytic solution which remains analytic as long as a suitable Sobolev norm remains finite. The local existence of analytic solutions can be proved with an abstract Cauchy-Kovalevskaya theorem, see for example [69,70]. The propagation of analyticity by Sobolev regularity can be proved by a variant of the arguments in [50] along with the inequality $\|B\rho\|_2 \lesssim \sum_{\alpha \leq M} \|v^\alpha Bg\|_2$ for all Fourier multipliers B (with our notation (1.13)) and all integers $M > d/2$. We omit the proof of Lemma 2.1 for brevity.

Lemma 2.1 (Local existence and propagation of analyticity) *Let $M > d/2$ be an integer and $\tilde{\lambda} > 0$. Suppose h_{in} is analytic such that*

$$\sum_{\alpha \in \mathbb{N}^d; |\alpha| \leq M} \|v^\alpha h_{\text{in}}\|_{\mathcal{G}^{\tilde{\lambda};1}} < \infty.$$

Then there exists some $T_0 > 0$ such that there exists a unique analytic solution $g(t)$ to (2.4) on $[0, T]$ for all $T < T_0$ and for some index $\tilde{\lambda}(t)$ with $\inf_{t \in [0, T]} \tilde{\lambda}(t) > 0$ we have,

$$\sup_{t \in [0, T]} \sum_{\alpha \in \mathbb{N}^d; |\alpha| \leq M} \|v^\alpha g(t)\|_{\mathcal{G}^{\tilde{\lambda}(t);1}} < \infty.$$

Moreover, if for some $T \leq T_0$ and $\sigma > d/2$, $\limsup_{t \nearrow T} \sum_{\alpha \in \mathbb{N}^d; |\alpha| \leq M} \|v^\alpha g(t)\|_{H_{x,v}^\sigma} < \infty$, then $T < T_0$.

Remark 10 If $d \leq 3$ and the solution has finite kinetic energy, then global analytic solutions to (1.2) exist even for large data. See the classical results [12, 41, 54, 74, 78] for the global existence of strong solutions (from which analyticity can be propagated by a variant of e.g. [50]). We remark that in $d \geq 4$, finite time blow-up is possible at least for gravitational interactions [48], however, this case is still covered by Theorem 1.

Remark 11 To treat the case $s = 1$ in Theorem 1, we would need to be slightly more careful in applying Lemma 2.1. In this case, we may regularize the data to a larger radius of analyticity $\tilde{\lambda} > \lambda_0$, perform our a priori estimates until $\tilde{\lambda}(t) = \lambda(t)$, at which point we may stop, re-regularize and restart iteratively.

Lemma 2.1 implies that as long as we retain control on the moments and regularity of the regularized solutions, they exist and remain analytic. We perform the a priori estimates on these solutions, for which the computations are rigorous, and then we may pass to the limit to show that the original solutions exist globally and satisfy the same estimates as the regularized solutions. For the remainder of the paper, we omit these details and discuss only the a priori estimates.

For constants $K_i, 1 \leq i \leq 3$ fixed in the proof depending only on $C_0, \bar{\lambda}, \kappa, s, d, \lambda_0$ and λ' , let $I \subset \mathbb{R}_+$ be the largest interval of times such that $0 \in I$ and the following controls hold for all $t \in I$:

$$\sum_{\alpha \in \mathbb{N}^d: |\alpha| \leq M} \|\langle \nabla_{z,v} \rangle A(v^\alpha g)(t)\|_2^2 = \sum_{\alpha \in \mathbb{N}^d: |\alpha| \leq M} \|A^{(1)}(v^\alpha g)(t)\|_2^2 \leq 4K_1(t)^7 \epsilon^2 \tag{2.11a}$$

$$\sum_{\alpha \in \mathbb{N}^d: |\alpha| \leq M} \|\langle \nabla_{z,v} \rangle^{-\beta} A(v^\alpha g)(t)\|_2^2 = \sum_{\alpha \in \mathbb{N}^d: |\alpha| \leq M} \|A^{(-\beta)}(v^\alpha g)(t)\|_2^2 \leq 4K_2 \epsilon^2 \tag{2.11b}$$

$$\int_0^t \|A\rho(\tau)\|_2^2 d\tau \leq 4K_3 \epsilon^2, \tag{2.11c}$$

where we may fix $\beta > 2$ arbitrary and ϵ satisfies a certain smallness assumption as in Theorem 1. Recall the definition of A in (2.7). It is clear from the assumptions that if $K_i \geq 1$ then $0 \in I$. The primary step in the proof of Theorem 1 is to show that $I = [0, \infty)$.

For the regularized solutions it will be clear from the ensuing arguments that the quantities on the LHS of (2.11) are continuous in time, from which it follows that I is relatively closed in \mathbb{R}_+ . Hence define $T^* \leq \infty$ such that $I = [0, T^*]$. In order to prove that $T^* = \infty$ it suffices to establish that I is also relatively open, which is implied by the following bootstrap.

Proposition 2.2 (Bootstrap) *There exists ϵ_0, K_i depending only on $d, M, \bar{\lambda}, C_0, \kappa, \lambda_0, \lambda'$ and s such that if (2.11) holds on some time interval $[0, T^*)$ and $\epsilon < \epsilon_0$, then for $t \leq T^*$,*

$$\sum_{\alpha \in \mathbb{N}^d: |\alpha| \leq M} \|\langle \nabla_{z,v} \rangle A(v^\alpha g)(t)\|_2^2 = \sum_{\alpha \in \mathbb{N}^d: |\alpha| \leq M} \|A^{(1)}(v^\alpha g)(t)\|_2^2 < 2K_1 \langle t \rangle^7 \epsilon^2 \tag{2.12a}$$

$$\sum_{\alpha \in \mathbb{N}^d: |\alpha| \leq M} \|\langle \nabla_{z,v} \rangle^{-\beta} A(v^\alpha g)(t)\|_2^2 = \sum_{\alpha \in \mathbb{N}^d: |\alpha| \leq M} \|A^{(-\beta)}(v^\alpha g)(t)\|_2^2 < 2K_2 \epsilon^2 \tag{2.12b}$$

$$\int_0^t \|A\rho(\tau)\|_2^2 \, d\tau < 2K_3 \epsilon^2, \tag{2.12c}$$

from which it follows that $T^* = \infty$.

Once Proposition 2.2 is deduced, Theorem 1 follows quickly. This is carried out in §7.

Remark 12 In order to close the bootstrap in Proposition 2.2 we need to understand how, or if, the constants K_i depend on each other so that we can be sure that they can be chosen self-consistently. In fact, K_3 is determined by the linearized evolution (from C_{LD} in Lemma 4.1) then K_1 is fixed in §5.3 depending on K_3 (and $s, d, \lambda_0, \lambda'$) and K_2 is analogously fixed in §5.4 depending on K_3 (and $s, d, \lambda_0, \lambda'$ but not directly K_1). Finally, ϵ_0 is chosen small with respect to everything.

Remark 13 The unbalance of a whole derivative between (2.12c) and (2.12a) uses the regularization of the interaction potential and is the only aspect of the proof which requires $\gamma \geq 1$.

The purpose of the weights in velocity is to control derivatives of the Fourier transform so that the trace Lemma 3.4 and the $H^{d/2+} \hookrightarrow C^0$ embedding theorem can be applied to restrict the Fourier transform along rays and pointwise. Both are necessary to deduce the controls on the density in §5.1 and §5.2. Moreover, from (2.11b) and (2.10) we have,

$$\begin{aligned} \|A^{(-\beta)} \hat{g}\|_{L_k^2 H_\eta^M} &\lesssim \sum_{\alpha \in \mathbb{N}^d: |\alpha| \leq M} \|D_\eta^\alpha A^{(-\beta)} \hat{g}\|_{L_k^2 L_\eta^2} \\ &\leq \sum_{\alpha \in \mathbb{N}^d: |\alpha| \leq M} \sum_{j \leq \alpha} \frac{\alpha!}{j!(\alpha-j)!} \|(D_\eta^{\alpha-j} A^{(-\beta)})(D_\eta^j \hat{g})\|_{L_k^2 L_\eta^2} \\ &\lesssim_M \sqrt{K_2} \epsilon. \end{aligned} \tag{2.13}$$

Similarly, (2.11a) implies

$$\|A^{(1)} \hat{g}\|_{L_k^2 H_\eta^M} \lesssim_M \sqrt{K_1} \epsilon \langle t \rangle^{7/2}. \tag{2.14}$$

Let us briefly summarize how Proposition 2.2 is proved. The main step is the time-integral estimate on the L^2 norm of the density in (2.12c), deduced in §5.1. This is done

by analyzing (2.6). The linear term in (2.6) is treated using a Fourier-Laplace transform and (\mathbf{L}) as in [67] with a technical time decomposition in order to get an estimate in Gevrey regularity using $A(t)$. The nonlinear term (2.6) is decomposed using a paraproduct into *reaction*, *transport* and *remainder* terms. As discussed above in §2.2, the reaction term in (2.6) is connected to the plasma echoes. Once the paraproduct decomposition and (2.11b) have allowed us to isolate this effect, it is treated with an adaptation of §7 in [67], carried out in §6 (our treatment is in the same spirit but not quite the same). The transport term describes the interaction of ρ with ‘higher’ frequencies of g ; this effect is controlled using (2.11a). Once the time-integral estimate has been established, we also derive a relatively straightforward pointwise-in-time control in §5.2.

The estimate (2.12a) is deduced in §5.3 via an energy estimate in the spirit of [10]. A paraproduct is again used to decompose the nonlinearity into *reaction*, *transport* and *remainder* terms. As in [10], the transport term is treated using an adaptation of the methods of [25, 45, 50]. However, perhaps more like [67], the reaction term is treated using (2.11c). The time growth is due to the fact that there is no regularity available to transfer to decay; that the estimate is closable at all requires the regularization from $\gamma \geq 1$. The low norm estimate (2.12b) is proved in §5.4 in a fashion similar to that of (2.12a) (the uniform bound is possible due to the regularity gap of $\beta > 2$ derivatives between (2.11c) and (2.12b), which can be into time decay on ρ).

3 Toolbox

In this section we review some of the technical tools used in the proof of Theorem 1: the Littlewood-Paley dyadic decomposition, paraproducts and useful inequalities for working in Gevrey regularity.

3.1 Fourier Analysis Conventions

For a function $g = g(z, v)$ we write its Fourier transform $\hat{g}_k(\eta)$ where $(k, \eta) \in \mathbb{Z}^d \times \mathbb{R}^d$ with

$$\hat{g}_k(\eta) := \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d \times \mathbb{R}^d} e^{-izk - iv\eta} g(z, v) \, dz \, dv,$$

$$g(z, v) := \frac{1}{(2\pi)^d} \sum_{k \in \mathbb{Z}^d} \int_{\mathbb{R}^d} e^{izk + iv\eta} \hat{g}_k(\eta) \, d\eta.$$

We use an analogous convention for Fourier transforms to functions of x or v alone.

With these conventions we have the following relations:

$$\left\{ \begin{aligned} \int_{\mathbb{T}^d \times \mathbb{R}^d} g(z, v) \overline{g}(z, v) \, dz \, dv &= \sum_{k \in \mathbb{Z}^d} \int_{\mathbb{R}^d} \hat{g}_k(\eta) \overline{\hat{g}_k(\eta)} \, d\eta, \\ \widehat{g^1 g^2} &= \frac{1}{(2\pi)^d} \hat{g}^1 * \hat{g}^2, \\ (\widehat{\nabla g})_k(\eta) &= (ik, i\eta) \hat{g}_k(\eta), \\ (\widehat{v^\alpha g})_k(\eta) &= D_\eta^\alpha \hat{g}_k(\eta). \end{aligned} \right.$$

The following versions of Young’s inequality occur frequently in the proof.

Lemma 3.1 (a) *Let $g_k^1(\eta), g_k^1(\eta) \in L^2(\mathbb{Z}^d \times \mathbb{R}^d)$ and $\langle k \rangle^\sigma r_k(t) \in L^2(\mathbb{Z}^d)$ for $\sigma > d/2$. Then, for any $t \in \mathbb{R}$,*

$$\left| \sum_{k,\ell} \int_{\eta} g_k^1(\eta) r_{\ell}(t) g_{k-\ell}^1(\eta - t\ell) \, d\eta \right| \lesssim_{d,\sigma} \|g^1\|_{L^2_{k,\eta}} \|g^1\|_{L^2_{k,\eta}} \|\langle k \rangle^\sigma r(t)\|_{L^2_k}. \tag{3.1}$$

(b) *Let $g_k^1(\eta), \langle k \rangle^\sigma g_k^2(\eta) \in L^2(\mathbb{Z}^d \times \mathbb{R}^d)$ and $r_k \in L^2(\mathbb{Z}^d)$ for $\sigma > d/2$. Then, for any $t \in \mathbb{R}$,*

$$\left| \sum_{k,\ell} \int_{\eta} g_k^1(\eta) h_{\ell}(t) g_{k-\ell}^2(\eta - t\ell) \, d\eta \right| \lesssim_{d,\sigma} \|g^1\|_{L^2_{k,\eta}} \|\langle k \rangle^\sigma g^2\|_{L^2_{k,\eta}} \|r(t)\|_{L^2_k}. \tag{3.2}$$

Proof To prove (a):

$$\begin{aligned} & \left| \sum_{k,\ell} \int_{\eta} g_k^1(\eta) r_{\ell} g_{k-\ell}^2(\eta - t\ell) \, d\eta \right| \\ & \leq \sum_k \left(\int_{\eta} |g_k^1(\eta)|^2 \, d\eta \right)^{1/2} \sum_{\ell} r_{\ell} \left(\int_{\eta} |g_{k-\ell}^2(\eta - t\ell)|^2 \, d\eta \right)^{1/2} \\ & = \sum_k \left(\int_{\eta} |g_k^1(\eta)|^2 \, d\eta \right)^{1/2} \sum_{\ell} r_{\ell} \left(\int_{\eta} |g_{k-\ell}^2(\eta)|^2 \, d\eta \right)^{1/2} \\ & \leq \left(\sum_k \int_{\eta} |g_k^1(\eta)|^2 \, d\eta \right)^{1/2} \left[\sum_k \left(\sum_{\ell} r_{\ell} \left(\int_{\eta} |g_{k-\ell}^2(\eta)|^2 \, d\eta \right)^{1/2} \right)^2 \right]^{1/2} \\ & \leq \|g^1\|_{L^2_{k,\eta}} \|g^2\|_{L^2_{k,\eta}} \sum_{\ell} |r_{\ell}| \\ & \lesssim_{d,\sigma} \|g^1\|_{L^2_{k,\eta}} \|g^2\|_{L^2_{k,\eta}} \|\langle k \rangle^\sigma r_k\|_{L^2_k}, \end{aligned}$$

where the penultimate line followed from the $L^2 \times L^1 \mapsto L^2$ Young’s inequality. The proof of (b) is analogous, simply putting g^2 rather than r in L^1_k . \square

3.2 Littlewood-Paley Decomposition

This work makes use of the Littlewood-Paley dyadic decomposition. Here we fix conventions and review the basic properties of this classical theory (see e.g. [5] for more details).

Define the joint variable $\Xi := (k, \eta) \in \mathbb{Z}^d \times \mathbb{R}^d$. Let $\psi \in C_0^\infty(\mathbb{Z}^d \times \mathbb{R}^d; \mathbb{R})$ be a radially symmetric non-negative function such that $\psi(\Xi) = 1$ for $|\Xi| \leq 1/2$ and $\psi(\Xi) = 0$ for $|\Xi| \geq 3/4$. Then we define $\phi(\Xi) := \psi(\Xi/2) - \psi(\Xi)$, a non-negative, radially symmetric function supported in the annulus $\{1/2 \leq |\Xi| \leq 3/2\}$. Then we define the rescaled functions $\phi_N(\Xi) = \phi(N^{-1}\Xi)$, which satisfy

$$\text{supp } \phi_N(\Xi) = \{N/2 \leq |\Xi| \leq 3N/2\}$$

and we have classically the partition of unity,

$$1 = \psi(\Xi) + \sum_{N \in 2^{\mathbb{N}}} \phi_N(\Xi),$$

(observe that there are always at most two non-zero terms in this sum), where we mean that the sum runs over the dyadic numbers $N = 1, 2, 4, 8, \dots, 2^j, \dots$.

For $g \in L^2(\mathbb{T}^d \times \mathbb{R}^d)$ we define

$$\begin{cases} g_N & := \phi_N(\nabla_{z,v})g, \\ g_{\frac{1}{2}} & := \psi(\nabla_{z,v})g, \\ g_{<N} & := g_{\frac{1}{2}} + \sum_{N' \in 2^{\mathbb{N}} : N' < N} g_{N'}, \end{cases}$$

and we have the natural decomposition,

$$g = \sum_{N \in \mathbb{D}} g_N = g_{\frac{1}{2}} + \sum_{N \in 2^{\mathbb{N}}} g_N, \quad \mathbb{D} := \left\{ \frac{1}{2}, 1, 2, \dots, 2^j, \dots \right\}.$$

Normally one would use g_0 or g_{-1} rather than the slightly inconsistent $g_{1/2}$, however here g_0 is reserved for the much more commonly used projection onto the zero mode only in z or x .

There holds the almost orthogonality and the approximate projection property

$$\begin{cases} \sum_{N \in \mathbb{D}} \|g_N\|_2^2 \leq \|g\|_2^2 \leq 2 \sum_{N \in \mathbb{D}} \|g_N\|_2^2 \\ \|(g_N)_N\|_2 \leq \|g_N\|_2. \end{cases} \tag{3.3}$$

Similarly to (3.3) but more generally, if $g = \sum_{\mathbb{D}} g_N$ with $\text{supp } g_N \subset \{C^{-1}N \leq |\Xi| \leq CN\}$ for $N \geq 1$ and $\text{supp } g_{\frac{1}{2}} \subset \{|\Xi| \leq C\}$ then we have

$$\|g\|_2^2 \approx_C \sum_{N \in \mathbb{D}} \|g_N\|_2^2. \tag{3.4}$$

Moreover, the dyadic decomposition behaves nicely with respect to differentiation:

$$\| \langle \nabla_{z,v} \rangle g_N \|_2 \approx N \|g_N\|_2. \tag{3.5}$$

We also define the notation

$$g_{\sim N} = \sum_{N' \in \mathbb{D} : C^{-1}N \leq N' \leq CN} g_{N'},$$

for some constant C which is independent of N . Generally the exact value of C which is being used is not important; what is important is that it is finite and independent of N .

In some steps of the proof, we will apply the Littlewood-Paley decomposition to the spatial density $\rho(t, x)$. In this case it will be more convenient to use the following definition that uses kt in place of the frequency in v , a natural convention when one recalls that $\hat{\rho}_k(t) = \hat{g}_k(t, kt)$:

$$\begin{cases} \widehat{\rho(t)}_N &= \phi_N(|k, kt|) \hat{\rho}_k(t), \\ \widehat{\rho(t)}_{\frac{1}{2}} &= \psi(|k, kt|) \hat{\rho}_k(t), \\ \widehat{\rho(t)}_{<N} &= \rho(t)_{\frac{1}{2}} + \sum_{N' \in 2^{\mathbb{N}} : N' < N} \rho(t)_{N'}. \end{cases}$$

Remark 14 We have opted to use the compact notation above, rather than the commonly used alternatives $\Delta_j g = g_{2^j}$ and $S_j g = g_{<2^j}$, in order to reduce the number of characters in long formulas.

3.3 The Paraproduct Decomposition

Another key Fourier analysis tool employed in this work is the *paraproduct decomposition*, introduced by Bony [15] (see also [5]). Given suitable functions g^1, g^2 we may define the paraproduct decomposition (in either (z, v) or just v),

$$\begin{aligned} g^1 g^2 &= \sum_{N \geq 8} g^1_{<N/8} g^2_N + \sum_{N \geq 8} g^1_N g^2_{<N/8} + \sum_{N \in \mathbb{D}} \sum_{N/8 \leq N' \leq 8N} g^1_N g^2_{N'} \\ &:= T_{g^1} g^2 + T_{g^2} g^1 + \mathcal{R}(g^1, g^2) \end{aligned} \tag{3.6}$$

where all the sums are understood to run over \mathbb{D} .

The advantage of this decomposition in the energy estimates is that the first term $T_{g^1} g^2$ contains the highest derivatives on g^2 but allows to take advantage of the frequency cutoff on the first function g^1 , whereas the second term $T_{g^2} g^1$ contains the highest derivatives on g^1 but allows to take advantage of the frequency cutoff on the second function g^2 . Finally the last “remainder” term contains the contribution from comparable frequencies which allows to split regularity evenly between the factors.

3.4 Elementary Inequalities for Gevrey Regularity

In this section we discuss a set of elementary, but crucial, inequalities for working in Gevrey regularity spaces. First we point out that Gevrey and Sobolev regularities can be related with the following two inequalities.

(i) For $x \geq 0, \alpha > \beta \geq 0, C, \delta > 0$,

$$e^{Cx^\beta} \leq e^{C\left(\frac{\alpha-\beta}{\alpha}\right)\left(\frac{C}{\delta}\right)^{\frac{\beta}{\alpha-\beta}}} e^{\delta x^\alpha}. \quad (3.7)$$

(ii) For $x \geq 0, \alpha > 0, \sigma, \delta > 0$,

$$e^{-\delta x^\alpha} \lesssim_{\sigma, \alpha} \frac{1}{\delta^{\frac{\sigma}{\alpha}} \langle x \rangle^\sigma}. \quad (3.8)$$

Next, we state several useful inequalities regarding the weight $\langle x \rangle = (1 + |x|^2)^{1/2}$. In particular, the improvements to the triangle inequality for $s < 1$ given in (3.10), (3.11) and (3.12) are important for getting useful bilinear (and trilinear) estimates. The proof is straightforward and is omitted.

Lemma 3.2 *Let $0 < s < 1$ and $x, y \geq 0$.*

(i) *We have the triangle inequalities (which hold also for $s = 1$),*

$$\langle x + y \rangle^s \leq \langle x \rangle^s + \langle y \rangle^s \quad (3.9a)$$

$$|\langle x \rangle^s - \langle y \rangle^s| \leq \langle x - y \rangle^s \quad (3.9b)$$

$$C_s (\langle x \rangle^s + \langle y \rangle^s) \leq \langle x + y \rangle^s, \quad (3.9c)$$

for some $C_s > 0$ depending only on s .

(ii) *In general,*

$$|\langle x \rangle^s - \langle y \rangle^s| \lesssim_s \frac{1}{\langle x \rangle^{1-s} + \langle y \rangle^{1-s}} \langle x - y \rangle. \quad (3.10)$$

(iii) *If $|x - y| \leq x/K$ for some $K > 1$, then we have the improved triangle inequality*

$$|\langle x \rangle^s - \langle y \rangle^s| \leq \frac{s}{(K-1)^{1-s}} \langle x - y \rangle^s. \quad (3.11)$$

(iv) *We have the improved triangle inequality for $x \geq y$,*

$$\langle x + y \rangle^s \leq \left(\frac{\langle x \rangle}{\langle x + y \rangle} \right)^{1-s} (\langle x \rangle^s + \langle y \rangle^s). \quad (3.12)$$

The following product lemma will be used several times in the sequel. Notice that $\tilde{c} < 1$ for $s < 1$, which shows that we gain something by working in Gevrey spaces with $s < 1$; indeed this important gain is used many times in the nonlinear estimates. We sketch the proof as it provides a representative example of arguments used several times in §5.

Lemma 3.3 (Product Lemma) *For all $0 < s < 1$, and $\sigma \geq 0$ there exists a $\tilde{c} = \tilde{c}(s, \sigma) \in (0, 1)$ such that the following holds for all $\lambda > 0$, $g^1, g^2 \in \mathcal{G}^{\lambda, \sigma; s}(\mathbb{T}^d \times \mathbb{R}^d)$ and $r(t) \in \mathcal{F}^{\lambda, \sigma; s}(\mathbb{T}^d)$:*

$$\begin{aligned} & \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}^d} \int_{\mathbb{R}^d} \langle k, \eta \rangle^{2\sigma} e^{2\lambda \langle k, \eta \rangle^s} \left| \hat{g}^1_k(\eta) \hat{r}_\ell(t) \hat{g}^2_{k-\ell}(\eta - \ell t) \right| d\eta \\ & \lesssim_{\lambda, \sigma, s, d} \|g^1\|_{\mathcal{G}^{\lambda, \sigma; s}} \|g^2\|_{\mathcal{G}^{\tilde{c}\lambda, 0; s}} \|r(t)\|_{\mathcal{F}^{\lambda, \sigma; s}} + \|g^1\|_{\mathcal{G}^{\lambda, \sigma; s}} \|g^2\|_{\mathcal{G}^{\lambda, \sigma; s}} \|r(t)\|_{\mathcal{F}^{\tilde{c}\lambda, 0; s}}. \end{aligned} \tag{3.13}$$

Moreover, we have the algebra property, for $g \in \mathcal{G}^{\lambda, \sigma; s}(\mathbb{T}^d \times \mathbb{R}^d)$ and $r(t) \in \mathcal{F}^{\lambda, \sigma; s}(\mathbb{T}^d)$,

$$\sum_{k \in \mathbb{Z}^d} \int_{\mathbb{R}^d} \left| \sum_{\ell \in \mathbb{Z}^d} e^{\lambda \langle k, \eta \rangle^s} \langle k, \eta \rangle^\sigma \hat{r}_\ell(t) \hat{g}_{k-\ell}(\eta - \ell t) \right|^2 d\eta \lesssim \|r(t)\|_{\mathcal{F}^{\lambda, \sigma; s}}^2 \|g\|_{\mathcal{G}^{\lambda, \sigma; s}}^2. \tag{3.14}$$

Proof We only prove (3.13), which is slightly harder. Denote $B_k(\eta) = \langle k, \eta \rangle^\sigma e^{\lambda \langle k, \eta \rangle^s}$. The proof proceeds by decomposing with a paraproduct:

$$\begin{aligned} & \sum_{k, \ell \in \mathbb{Z}^d} \int_{\eta} \left| B \hat{g}^1_k(\eta) B_k(\eta) \hat{r}_\ell(t) \hat{g}^2_{k-\ell}(\eta - \ell t) \right| d\eta \\ & \leq \sum_{N \geq 8} \sum_{k, \ell \in \mathbb{Z}^d_*} \int_{\eta} \left| B \hat{g}^1_k(\eta) B_k(\eta) \hat{r}_\ell(t) \hat{g}^2_{k-\ell}(\eta - \ell t)_{<N/8} \right| d\eta \\ & \quad + \sum_{N \geq 8} \sum_{k, \ell \in \mathbb{Z}^d_*} \int_{\eta} \left| B \hat{g}^1_k(\eta) B_k(\eta) \hat{r}_\ell(t)_{<N/8} \hat{g}^2_{k-\ell}(\eta - \ell t)_N \right| d\eta \\ & \quad + \sum_{N \in \mathbb{D}} \sum_{N/8 \leq N' \leq 8N} \sum_{k, \ell \in \mathbb{Z}^d_*} \int_{\eta} \left| B \hat{g}^1_k(\eta) B_k(\eta) \hat{r}_\ell(t)_{N'} \hat{g}^2_{k-\ell}(\eta - \ell t)_N \right| d\eta \\ & = R + T + \mathcal{R}. \end{aligned}$$

Note that the R and T term are almost, but not quite, symmetric. Consider first the R term. On the support of the integrand we have the frequency localizations

$$\frac{N}{2} \leq |\ell, \ell t| \leq \frac{3N}{2}, \tag{3.15a}$$

$$|k - \ell, \eta - \ell t| \leq \frac{3N}{32}, \tag{3.15b}$$

$$\frac{13}{16} \leq \frac{|k, \eta|}{|\ell, \ell t|} \leq \frac{19}{16}. \tag{3.15c}$$

Therefore, by (3.11), there exists some $c = c(s) \in (0, 1)$ such that

$$\begin{aligned}
 B_k(\eta) &\leq \langle k, \eta \rangle^\sigma e^{\lambda \langle \ell, \ell t \rangle^s} e^{c\lambda \langle k-\ell, \eta-\ell t \rangle^s} \lesssim_\sigma \langle \ell, \ell t \rangle^\sigma e^{\lambda \langle \ell, \ell t \rangle^s} e^{c\lambda \langle k-\ell, \eta-\ell t \rangle^s} \\
 &\lesssim_\lambda \langle \ell, \ell t \rangle^\sigma e^{\lambda \langle \ell, \ell t \rangle^s} \langle k-\ell, \eta-\ell t \rangle^{-\frac{d}{2}-1} e^{\frac{1}{2}(c+1)\lambda \langle k-\ell, \eta-\ell t \rangle^s},
 \end{aligned}$$

where in the last inequality we applied (3.8). Adding a frequency localization with (3.15), denoting $\tilde{c} = \frac{1}{2}(c + 1)$, and using (3.2) we have

$$\begin{aligned}
 R &\lesssim \sum_{N \geq 8} \sum_{k, \ell \in \mathbb{Z}_*^d} \int_\eta \left| B g^1_k(\eta) \sim_N B_\ell(\ell t) \hat{r}_\ell(t)_N e^{\tilde{c}\lambda \langle k-\ell, \eta-\ell t \rangle^s} \right. \\
 &\quad \left. \langle k-\ell, \eta-\ell t \rangle^{-\frac{d}{2}-1} \hat{g}^2_{k-l}(\eta-lt)_{<N/8} \right| d\eta \\
 &\lesssim \sum_{N \geq 8} \left\| g^1_{\sim N} \right\|_{\mathcal{G}^{\lambda, \sigma; s}} \|r(t)_N\|_{\mathcal{F}^{\lambda, \sigma; s}} \left\| g^2 \right\|_{\mathcal{F}^{\tilde{c}\lambda, 0; s}}.
 \end{aligned}$$

By Cauchy-Schwarz and (3.4) we have

$$R \lesssim \left\| g^1 \right\|_{\mathcal{G}^{\lambda, \sigma; s}} \|r(t)\|_{\mathcal{F}^{\lambda, \sigma; s}} \left\| g^2 \right\|_{\mathcal{F}^{\tilde{c}\lambda, 0; s}},$$

which appears on the RHS of (3.13). Treating the T term is essentially the same except reversing the role of $(\ell, \ell t)$ and $(k - \ell, \eta - \ell t)$ and applying (3.1) as opposed to (3.2).

To treat the \mathcal{R} term we use a simple variant. We claim that there exists some $c' = c'(s) \in (0, 1)$ such that on the support of the integrand we have

$$B_k(\eta) \lesssim_\sigma e^{c'\lambda \langle k-\ell, \eta-\ell t \rangle^s} e^{c'\lambda \langle \ell, \ell t \rangle^s}. \tag{3.16}$$

To see this, consider separately the cases (say) $N \geq 128$ and $N < 128$. On the latter, $B_k(\eta)$ is simply bounded by a constant. In the case $N \geq 128$ we have the frequency localizations

$$\frac{N}{2} \leq |k - \ell, \eta - \ell t| \leq \frac{3N}{2}, \tag{3.17a}$$

$$\frac{N'}{2} \leq |\ell, \ell t| \leq \frac{3N'}{2}, \tag{3.17b}$$

$$\frac{1}{24} \leq \frac{|k - \ell, \eta - \ell t|}{|\ell, \ell t|} \leq 24, \tag{3.17c}$$

and hence we may apply (3.12) since in this case $64 \leq |k - \ell, \eta - \ell t| \approx |\ell, \ell t|$. Further, we can use (3.8) to absorb the Sobolev corrections and, indeed, we have (3.16) on the support of the integrand in \mathcal{R} . Hence,

$$\mathcal{R} \lesssim \sum_{N \in \mathbb{D}} \sum_{N' \approx N} \sum_{k, \ell \in \mathbb{Z}_*^d} \int_\eta \left| B g^1_k(\eta) e^{c'\lambda \langle \ell, \ell t \rangle^s} \hat{r}_\ell(t)_{N'} e^{c'\lambda \langle k-\ell, \eta-\ell t \rangle^s} \hat{g}^2_{k-l}(\eta-lt)_N \right| d\eta.$$

Applying (3.2) followed by (3.5) and (3.8) (since $c' < 1$) implies,

$$\begin{aligned} \mathcal{R} &\lesssim \sum_{N \in \mathbb{D}} \sum_{N' \approx N} \left\| g^1 \right\|_{\mathcal{G}^{\lambda, \sigma; s}} \|r_{N'}\|_{\mathcal{F}^{c'\lambda, \frac{d}{2}+1; s}} \left\| g_N^2 \right\|_{\mathcal{G}^{c'\lambda, 0; s}} \\ &\lesssim \sum_{N \in \mathbb{D}} \frac{1}{N} \left\| g^1 \right\|_{\mathcal{G}^{\lambda, \sigma; s}} \|r_{\sim N}\|_{\mathcal{F}^{c'\lambda, \frac{d}{2}+2; s}} \left\| g_N^2 \right\|_{\mathcal{G}^{c'\lambda, 0; s}} \\ &\lesssim_{\lambda, \sigma} \left\| g^1 \right\|_{\mathcal{G}^{\lambda, \sigma; s}} \|r\|_{\mathcal{F}^{\lambda, \sigma; s}} \left\| g^2 \right\|_{\mathcal{G}^{c'\lambda, 0; s}}. \end{aligned}$$

Hence (after possibly adjusting \tilde{c}) this term appears on the RHS of (3.13). □

We also need the standard Sobolev space trace lemma, which we will apply on the Fourier side.

Lemma 3.4 (*L^2 Trace*) *Let u be smooth on \mathbb{R}^d and $C \subset \mathbb{R}^d$ be an arbitrary straight line. Then for all $\sigma \in \mathbb{R}_+$ with $\sigma > (d - 1)/2$,*

$$\|u\|_{L^2(C)} \lesssim \|u\|_{H^\sigma(\mathbb{R}^d)}.$$

Proof It follows by induction on co-dimension with the standard $H^{1/2}$ restriction theorem [1]. □

4 Linear Damping in Gevrey Regularity

The first step to proving (2.12c) is understanding (forced) linear Landau damping in the L^2 Gevrey norms we are using. This is provided by the following lemma, which also shows that (L) implies linear damping in all Gevrey regularities ($s > 1/3$ is not relevant to the proof). It is crucial that the same norm appears on both sides of (4.2) so that we may use it in the nonlinear estimate on $\rho(t)$. The main idea of the proof of Lemma 4.1 appears in [67] to treat damping in Gevrey regularity and is based on decomposing the solution into analytic/exponentially decaying sub-components. We note that Lin and Zeng in [52] have linear damping results at much lower regularities (similarly, we believe the following proof can be adapted also to Sobolev spaces). We will first give a formal proof of Lemma 4.1 as an a priori estimate in §4.1 and then explain the rigorous justification in §4.2. In §4.3, we discuss the proof that the Penrose condition [73] implies (L).

Lemma 4.1 (Linear integral-in-time control) *Let f^0 satisfy the condition (L) with constants $M > d/2$ and $C_0, \bar{\lambda}, \kappa > 0$. Let $A_k(t, \eta)$ be the multiplier defined in (2.7) for $s \in (0, 1)$, and $\lambda = \lambda(t) \in (0, \lambda_0)$ as defined in (2.8). Let $F(t)$ and $T^* > 0$ be given such that, if we denote $I = [0, T^*)$,*

$$\int_0^{T^*} \|F(t)\|_{\mathcal{F}^{\lambda(t), \sigma; s}}^2 dt = \sum_{k \in \mathbb{Z}_*^d} \|A_k(t, kt)F_k(t)\|_{L_t^2(I)}^2 < \infty.$$

Then there exists a constant $C_{LD} = C_{LD}(C_0, \bar{\lambda}, \kappa, s, d, \lambda_0, \lambda')$ such that for all $k \in \mathbb{Z}_*^d$, the solution $\phi_k(t)$ to the system

$$\phi_k(t) = F_k(t) + \int_0^t K^0(t - \tau, k)\phi_k(\tau) \, d\tau \tag{4.1}$$

in $t \in \mathbb{R}_+$ with $K^0(t, k) := -\tilde{f}^0(k) \widehat{W}(k) |k|^2 t$ satisfies the mode-by-mode estimate

$$\forall k \in \mathbb{Z}_*^d, \int_0^{T_*} A_k(t, kt)^2 |\phi_k(t)|^2 \, dt \leq C_{LD}^2 \int_0^{T_*} A_k(t, kt) |F_k(t)|^2 \, dt \tag{4.2}$$

which is equivalent to $\int_0^{T_*} \|\phi\|_{\mathcal{F}^{\lambda(t), \sigma; s}}^2 \, dt \leq C_{LD}^2 \int_0^{T_*} \|F\|_{\mathcal{F}^{\lambda(t), \sigma; s}}^2 \, dt$.

Remark 15 The proof proceeds slightly differently in the case $s = 1$ where the additional requirement $\bar{\lambda} > \lambda(0)$ occurs naturally (and the constant badly depends in the parameter $\bar{\lambda} - \lambda(0)$).

4.1 Proof of the A Priori Estimate

We only consider the $s < 1$ case; the analytic case is only a slight variation. As the hypothesis on F is known a priori only to hold on $[0, T_*)$, we simply extend $F_k(t)$ to be zero for all $t \geq T_*$.

Step 1. Rough Grönwall bound. First we deduce a rough bound using Grönwall’s inequality with no attempt to be optimal. This bound shows in particular that the integral equation (4.1) is globally well-posed (for each frequency $k \in \mathbb{Z}_*^d$) in the norm associated with the multiplier A . By (3.9), the definition of K^0 , (1.3) and that $\lambda(t)$ is non-increasing in time,

$$\begin{aligned} A_k(t, kt) |\phi_k(t)| &\leq A_k(t, kt) |F_k(t)| + \int_0^t A_k(t, kt) |K^0(t - \tau, k)| |\phi_k(\tau)| \, d\tau \\ &\lesssim A_k(t, kt) |F_k(t)| + \int_0^t \langle k(t - \tau) \rangle^\sigma e^{\lambda(t)\langle k(t-\tau) \rangle^s} \\ &\quad \times \left| \widehat{\nabla_v f^0}(k(t - \tau)) \right| A_k(\tau, k\tau) |\phi_k(\tau)| \, d\tau. \end{aligned}$$

Then by (3.8), the $H^{d/2+} \hookrightarrow C^0$ embedding theorem and (1.9),

$$\begin{aligned} A_k(t, kt) |\phi_k(t)| &\lesssim A_k(t, kt) |F_k(t)| + \left(\sup_{\eta \in \mathbb{R}^d} \langle \eta \rangle^\sigma e^{\lambda(0)\langle \eta \rangle^s} \left| \widehat{f^0}(\eta) \right| \right) \\ &\quad \times \int_0^t A_k(\tau, k\tau) |\phi_k(\tau)| \, d\tau \\ &\lesssim A_k(t, kt) |F_k(t)| + \left\| \langle \eta \rangle^\sigma e^{\lambda(0)\langle \eta \rangle^s} \widehat{f^0}(\eta) \right\|_{H^M} \\ &\quad \times \int_0^t A_k(\tau, k\tau) |\phi_k(\tau)| \, d\tau. \end{aligned}$$

By (1.9) with (3.8), it follows by Grönwall’s inequality that for some $C > 0$ we have the following (using also Cauchy-Schwarz and (3.8) in the last inequality),

$$A_k(t, kt) |\phi_k(t)| \lesssim e^{Ct} \int_0^t |A_k(\tau, k\tau) F_k(\tau)| \, d\tau \lesssim e^{2Ct} \|AF_k\|_{L_t^2(I)}. \tag{4.3}$$

Step 2. Frequency localization. We would like to use the Fourier-Laplace transform as in [67], but $F_k(t)$ does not decay exponentially in time. Instead, we deduce the estimate by decomposing the problem into a countable number of exponentially decaying contributions. Let $R \geq e$ to be a constant fixed later depending only on $\bar{\lambda}, \lambda(0)^{-1}$ and s , let $F_k^n(t) = F_k(t) \mathbf{1}_{Rn \leq |kt|^s \leq R(n+1)}$, and define ϕ_k^n as solutions to

$$\phi_k^n(t) = F_k^n(t) + \int_0^t K^0(t - \tau, k) \phi_k^n(\tau) \, d\tau. \tag{4.4}$$

Then $\phi_k(t) = \sum_{n=0}^\infty \phi_k^n(t)$ by linearity of the equation, and by the definition of $F_k^n(t)$, $\phi_k^n(t)$ is supported for $|kt|^s \geq Rn$. Moreover, obviously (4.3) holds for each ϕ_k^n . Now we will use the Fourier-Laplace transform in time to get an L_t^2 estimate on ϕ_k^n as in [67], but with a contour which gets progressively closer to the imaginary axis as n increases. Define the indices μ_n as

$$\begin{aligned} \mu_0 &= \mu_1, \\ \mu_n &= \frac{1}{(Rn)^{1/s}} \left[\lambda \left(\frac{(Rn)^{1/s}}{|k|} \right) \langle (Rn)^{1/s} \rangle^s + \sigma \log \langle (Rn)^{1/s} \rangle \right], \quad n \geq 1. \end{aligned}$$

We will have the requirement that $\mu_n < \bar{\lambda}$ so that our integration contour always lies in the half plane defined in (L). As long as $s < 1$ this only requires us to choose R large relative to $\bar{\lambda}(\lambda(0))^{-1}$, so to fix ideas suppose that R is large enough so that $\sup \mu_n < \bar{\lambda}/2$. Define the amplitude corrections

$$\begin{aligned} N_{k,0} &= \langle k \rangle^\sigma \exp \left[\lambda \left(\frac{(R)^{1/s}}{|k|} \right) \langle k \rangle^\sigma \right], \\ N_{k,n} &= \frac{\langle k, (Rn)^{1/s} \rangle^\sigma}{\langle (Rn)^{1/s} \rangle^\sigma} \exp \left[\lambda \left(\frac{(Rn)^{1/s}}{|k|} \right) \left[\langle k, (Rn)^{1/s} \rangle^s - \langle (Rn)^{1/s} \rangle^s \right] \right] \quad n \geq 1. \end{aligned}$$

The role of this correction is highlighted by the fact that when $|kt| = (Rn)^{1/s}$ and $n \geq 1$, then

$$\begin{cases} e^{\mu_n |kt|} = \langle kt \rangle^\sigma e^{\lambda(t) \langle kt \rangle^s} \\ N_{n,k} = \frac{\langle k, kt \rangle^\sigma}{\langle kt \rangle^\sigma} e^{\lambda(t) (\langle k, kt \rangle^s - \langle kt \rangle^s)} \\ e^{\mu_n |kt|} N_{n,k} = \langle k, kt \rangle^\sigma e^{\lambda(t) \langle k, kt \rangle^s} = A_k(t, kt). \end{cases}$$

Hence the index μ_n is related to the radius of analyticity in the velocity variable, whereas the correction $N_{k,n}$ measures the ratio of what is lost by not taking into

account the regularity in the space variable. A further error is introduced by the fact that these regularity weights only exactly fit the multiplier A at the left endpoint of the interval of the decomposition.

Step 3. The one-block estimate via the Fourier-Laplace transform. Now multiply (4.4) by $e^{\mu_n|k|t} N_{k,n}$, and denoting

$$\begin{cases} R_k^n(t) &= e^{\mu_n|k|t} N_{k,n} F_k^n(t) \\ \Phi_k^n(t) &= e^{\mu_n|k|t} N_{k,n} \phi_k^n(t), \end{cases}$$

we have

$$\Phi_k^n(t) = R_k^n(t) + \int_0^t e^{\mu_n|k|(t-\tau)} K^0(t - \tau, k) \Phi_k^n(\tau) \, d\tau. \tag{4.5}$$

Taking the Fourier transform in time $\hat{G}(\omega) = (1/\sqrt{2\pi}) \int_{\mathbb{R}} e^{-it\omega} G(t) \, dt$ (extending R_k^n and Φ_k^n and K^0 as zero for negative times)¹ we obtain

$$\hat{\Phi}_k^n(\omega) = \hat{R}_k^n(\omega) + \mathcal{L}\left(k, \mu_n + i \frac{\omega}{|k|}\right) \hat{\Phi}_k^n(\omega),$$

(where $\mathcal{L}(\xi, k) := \int_0^{+\infty} e^{\xi|k|t} K^0(t, k) \, dt$), which is formally solved as

$$\Phi_k^n(\omega) = \frac{\hat{R}_k^n(\omega)}{1 - \mathcal{L}\left(k, \mu_n + i \frac{\omega}{|k|}\right)}.$$

Applying the stability condition **(L)** and Plancherel’s theorem implies

$$\|\Phi_k^n(t)\|_{L^2_t(\mathbb{R})} \lesssim \frac{1}{\kappa} \|R_k^n(t)\|_{L^2_t(\mathbb{R})}. \tag{4.6}$$

Step 4. Coming back to the original multiplier A . Consider $Rn \leq |kt|^s \leq R(n + 1)$ for $n \geq 1$. From the definition (2.8) of $\lambda(t)$ we see that if R is chosen sufficiently large then for $t \geq R^{1/s}$ we have

$$\begin{aligned} \frac{d}{dt} (\lambda(t) \langle k, kt \rangle^s) &= \left(\dot{\lambda}(t) + s\lambda(t) \frac{|k| |k, kt|}{\langle k, kt \rangle^2} \right) \langle k, kt \rangle^s \\ &= \left(-\frac{a(\lambda_0 - \lambda')}{4\langle t \rangle^{1+a}} + s |k| \lambda(t) \frac{|k, kt|}{\langle k, kt \rangle^2} \right) \langle k, kt \rangle^s > 0, \end{aligned}$$

¹ The tacit assumption that the Fourier transform of $\Phi_k^n(t)$ exists is where the a priori estimate is not rigorous. We justify that the transform exists and that this computation can be made rigorous in §4.2 below.

and hence $\lambda(t)\langle k, kt \rangle^s$ is increasing since $|k| \geq 1$. Applying this gives,

$$\begin{aligned}
 N_{k,n}e^{\mu_n|kt|} &= \exp \left[\lambda \left(\frac{(Rn)^{1/s}}{|k|} \right) \left[\langle k, (Rn)^{1/s} \rangle^s - \langle (Rn)^{1/s} \rangle^s \right] \frac{\langle k, (Rn)^{1/s} \rangle^\sigma}{\langle (Rn)^{1/s} \rangle^\sigma} \right. \\
 &\quad \times \left. \exp \left[\frac{|kt|}{(Rn)^{1/s}} \left(\lambda \left(\frac{(Rn)^{1/s}}{|k|} \right) \langle (Rn)^{1/s} \rangle^s + \sigma \log \langle (Rn)^{1/s} \rangle \right) \right] \right] \\
 &\leq \exp \left[\lambda(t) \langle k, kt \rangle^s \right] \langle k, kt \rangle^\sigma \langle (Rn)^{1/s} \rangle^\sigma \left(\frac{(n+1)^{1/s}}{n^{1/s}} - 1 \right) \\
 &\quad \times \exp \left[\lambda \left(\frac{(Rn)^{1/s}}{|k|} \right) \left(\frac{(n+1)^{1/s}}{(n)^{1/s}} - 1 \right) \langle (Rn)^{1/s} \rangle^s \right] \\
 &\lesssim_{R,\lambda_0,s} A_k(t, kt).
 \end{aligned} \tag{4.7}$$

A similar result holds for $n = 0$ when $|kt|^s \leq R$ using that $\lambda(t)$ is non-increasing,

$$\begin{aligned}
 N_{k,0}e^{\mu_0|kt|} &= \exp \left[\lambda \left(\frac{R^{1/s}}{|k|} \right) \langle k \rangle^s \right] \langle k \rangle^\sigma \\
 &\quad \times \exp \left[\frac{|kt|}{R^{1/s}} \left(\lambda \left(\frac{R^{1/s}}{|k|} \right) \langle R^{1/s} \rangle^s + \sigma \log \langle R^{1/s} \rangle \right) \right] \\
 &\lesssim_R A_k(t, kt).
 \end{aligned}$$

Hence it follows from (4.6) that

$$\left\| \Phi_k^n(t) \right\|_{L_t^2(\mathbb{R})} \lesssim \frac{1}{k} \left\| AF_k(t) \mathbf{1}_{Rn \leq |kt|^s \leq R(n+1)} \right\|_{L_t^2(\mathbb{R})}. \tag{4.8}$$

Step 5. Summation of the different frequency blocks. Now we want to estimate $A\phi_k$ by summing over n which will require some almost orthogonality. This is possible as each ϕ_k^n is exponentially localized near $|kt|^s \approx Rn$. Computing, noting that by construction ϕ_k^n is only supported on $|kt|^s \geq Rn$,

$$\begin{aligned}
 \|A\phi_k\|_{L_t^2}^2 &= \int_0^{+\infty} \left| \sum_{n=0}^{+\infty} A_k(t, kt) \phi_k^n(t) \right|^2 dt \\
 &\lesssim \int_0^\infty \left| A_k(t, kt) \phi_k^0(t) \right|^2 dt + \int_0^\infty \left| \sum_{n=1}^\infty A_k(t, kt) e^{-\mu_n|kt|} N_{k,n}^{-1} \mathbf{1}_{|kt|^s \geq Rn} \Phi_k^n(t) \right|^2 dt \\
 &= \int_0^\infty \left| A_k(t, kt) \phi_k^0(t) \right|^2 dt + \int_0^\infty \sum_{n,n' \geq 1} A_k(t, kt)^2 e^{-\mu_n|kt|} \\
 &\quad \times N_{k,n}^{-1} e^{-\mu_{n'}|kt|} N_{k,n'}^{-1} \mathbf{1}_{|kt|^s \geq Rn} \mathbf{1}_{|kt|^s \geq Rn'} \Phi_k^n(t) \Phi_{k'}^{n'}(t) dt.
 \end{aligned} \tag{4.9}$$

First we will approach the infinite sum, which is the more challenging term. For this we use Schur’s test. Indeed, if we denote the interaction kernel

$$K_{n,n'}(t, k) := A_k(t, kt)^2 e^{-\mu_n|kt|} N_{k,n}^{-1} e^{-\mu_{n'}|kt|} N_{k,n'}^{-1} \mathbf{1}_{|kt|^s \geq Rn} \mathbf{1}_{|kt|^s \geq Rn'},$$

then Schur’s test (or Cauchy-Schwarz three times) implies

$$\begin{aligned}
 \int_0^{+\infty} \sum_{n,n' \geq 1} K_{n,n'}(t, k) \Phi_k^{n'}(t) \Phi_k^n(t) dt &\leq \left(\sup_{t \in [0, \infty)} \sup_{n \geq 1} \sum_{n'=1}^{\infty} K_{n,n'}(t, k) \right)^{1/2} \\
 &\times \left(\sup_{t \in [0, \infty)} \sup_{n' \geq 1} \sum_{n=1}^{\infty} K_{n,n'}(t, k) \right)^{1/2} \\
 &\times \sum_{n=1}^{\infty} \|\Phi_k^n(t)\|_{L_t^2(\mathbb{R})}^2. \tag{4.10}
 \end{aligned}$$

It remains to see that the row and column sums of the interaction kernel are uniformly bounded in time. Since the kernel is symmetric in n and n' it suffices to consider only one of the sums. The computations above to deduce (4.7) can be adapted to show at least that

$$A_k(t, kt)e^{-\mu_{n'}|kt|} N_{k,n'}^{-1} \mathbf{1}_{|kt|^s \geq Rn'} \lesssim_R 1,$$

and hence, since $\lambda(t)$ is decreasing,

$$\begin{aligned}
 \sum_{n=1}^{+\infty} K_{n,n'}(t, k) &\lesssim_R \sum_{n=1}^{+\infty} A_k(t, kt)e^{-\mu_n|kt|} N_{k,n}^{-1} \mathbf{1}_{|kt|^s \geq Rn} \\
 &= \sum_{n=1}^{+\infty} \mathbf{1}_{|kt|^s \geq Rn} e^{\lambda(t)\langle k, kt \rangle^s - \lambda\left(\frac{(Rn)^{1/s}}{|k|}\right)\langle (k, (Rn)^{1/s})^s - \langle (Rn)^{1/s} \rangle^s \rangle} \frac{\langle k, kt \rangle^\sigma \langle (Rn)^{1/s} \rangle^\sigma}{\langle k, (Rn)^{1/s} \rangle^\sigma} e^{-\mu_n|kt|} \\
 &\lesssim \sum_{n=1}^{+\infty} \mathbf{1}_{|kt|^s \geq Rn} e^{\lambda(t)\langle k, kt \rangle^s - \lambda(t)\langle (k, (Rn)^{1/s})^s - \langle (Rn)^{1/s} \rangle^s \rangle} \frac{\langle k, kt \rangle^\sigma \langle (Rn)^{1/s} \rangle^\sigma}{\langle k, (Rn)^{1/s} \rangle^\sigma} \\
 &\times e^{-\frac{|kt|}{(Rn)^{1/s}} [\lambda(t)\langle (Rn)^{1/s} \rangle^s + \sigma \log \langle (Rn)^{1/s} \rangle]}.
 \end{aligned}$$

Since $e \leq (Rn)^{1/s} \leq |kt|$, we have $-\langle k, (Rn)^{1/s} \rangle^s + \langle (Rn)^{1/s} \rangle^s \leq \langle kt \rangle^s - \langle k, kt \rangle^s$ since the LHS is increasing as a function of n and therefore,

$$\begin{aligned}
 &\sum_{n=1}^{+\infty} K_{n,n'}(t, k) \\
 &\lesssim \sum_{n=1}^{+\infty} \mathbf{1}_{|kt|^s \geq Rn} e^{\lambda(t)\langle kt \rangle^s \left[1 - \frac{|kt|\langle (Rn)^{1/s} \rangle^s}{(Rn)^{1/s}\langle kt \rangle^s} \right]} e^{\sigma \log \left[\frac{\langle k, kt \rangle \langle (Rn)^{1/s} \rangle}{\langle k, (Rn)^{1/s} \rangle} \right] - \sigma \frac{|kt|}{(Rn)^{1/s}} \log \langle (Rn)^{1/s} \rangle}.
 \end{aligned}$$

Since $\langle x \rangle \langle k, x \rangle^{-1}$ is increasing in x for $|k| \geq 1$ and $x \geq 0$, we have

$$\sum_{n=1}^{+\infty} K_{n,n'}(t, k) \lesssim \sum_{n=1}^{+\infty} \mathbf{1}_{|kt|^s \geq Rn} e^{\lambda(t)\langle kt \rangle^s \left[1 - \frac{|kt|\langle (Rn)^{1/s} \rangle^s}{(Rn)^{1/s}\langle kt \rangle^s} \right]} e^{\sigma \log \langle kt \rangle \left(1 - \frac{|kt| \log \langle (Rn)^{1/s} \rangle}{(Rn)^{1/s} \log \langle kt \rangle} \right)}.$$

Finally using that $x / \log(x)$ is increasing for $x \geq e$ we get

$$\sum_{n=1}^{+\infty} K_{n,n'}(t, k) \lesssim \sum_{n=1}^{+\infty} \mathbf{1}_{|kt|^s \geq Rn} e^{\lambda(t)\langle kt \rangle^s} \left[1 - \frac{|kt|((Rn)^{1/s})^s}{(Rn)^{1/s}\langle kt \rangle^s} \right].$$

Using that $\langle x^{1/s} \rangle^s \geq x$, the sum can be bounded by

$$\begin{aligned} \sum_{n=1}^{\infty} K_{n,n'}(t, k) &\lesssim_R e^{\lambda(t)\langle kt \rangle^s} \int_R^{|kt|^s} e^{-\lambda(t)|kt| \frac{\langle x^{1/s} \rangle^s}{x^{1/s}}} dx \\ &\lesssim e^{\lambda(t)\langle kt \rangle^s} \int_R^{|kt|^s} e^{-\lambda(t)|kt|x^{1-1/s}} dx \\ &\lesssim e^{\lambda(t)\langle kt \rangle^s} \int_{|kt|^{s-1}}^{R^{1-1/s}} e^{-\lambda(t)|kt|\tau} \tau^{\frac{1}{s-1}} d\tau \lesssim 1. \end{aligned}$$

This shows that the row sums of $K_{n,n'}(t, k)$ are uniformly bounded; by symmetry the column sums are also bounded and by (4.10), this completes the treatment of the summation in (4.9).

Now we turn our attention to the $n = 0$ term in (4.9). By (4.3),

$$\begin{aligned} \int_0^{+\infty} |A_k(t, kt)\phi_k^0(t)|^2 dt &= \int_0^{R^{1/s}} |A_k(t, kt)\phi_k^0(t)|^2 dt + \int_{R^{1/s}}^{+\infty} |A_k(t, kt)\phi_k^0(t)|^2 dt \\ &\lesssim_R \|AF_k\|_{L^2_t(\mathbb{R})}^2 + \int_{R^{1/s}}^{+\infty} |A_k(t, kt)\phi_k^0(t)|^2 dt. \end{aligned} \tag{4.11}$$

However, for $|t| \geq R^{1/s}$, we have by (3.9), that $\lambda(t)$ is non-increasing and (3.7),

$$A_k(t, kt) \leq e^{\lambda(t)\langle k \rangle^s + \lambda(t)\langle kt \rangle^s} \langle k \rangle^\sigma \langle kt \rangle^\sigma \lesssim_R e^{\lambda\left(\frac{R^{1/s}}{|k|}\right)\langle k \rangle^s} \langle k \rangle^\sigma e^{\mu_0|kt|} \lesssim N_{k,0} e^{\mu_0|kt|},$$

which implies with (4.11) and (4.8) that

$$\int_0^{+\infty} |A_k(t, kt)\phi_k^0(t)|^2 dt \lesssim \left(1 + \frac{1}{\kappa^2}\right) \|AF_k\|_{L^2_t(\mathbb{R})}^2. \tag{4.12}$$

Combining (4.9), (4.12), (4.10) with (4.8) we have

$$\begin{aligned} \|A\phi_k\|_{L^2_t(I)}^2 &\lesssim_R \left(1 + \frac{1}{\kappa^2}\right) \|AF_k\|_{L^2_t(\mathbb{R})}^2 + \sum_{n=1}^{\infty} \|\Phi_k^n(t)\|_{L^2_t(I)}^2 \\ &\lesssim \left(1 + \frac{1}{\kappa^2}\right) \sum_{n=0}^{\infty} \|AF_k(t)\mathbf{1}_{Rn \leq |kt|^s \leq R(n+1)}\|_{L^2_t(\mathbb{R})}^2, \end{aligned}$$

which completes the proof of the lemma.

4.2 Rigorous Justification of the A Priori Estimate

The reader may have noticed that in the previous subsection it seems that we only used the bound from below $|1 - \mathcal{L}(k, \xi)| \geq \kappa$ with $\xi = \mu_n + i\omega/|k|$, i.e. in the strip $\Re e \xi \in (0, \bar{\lambda}/2)$. The subtlety is that the Fourier-Laplace transform of $\Phi_k^n(t)$ is only granted to exist when some L^2 integrability as $t \rightarrow \infty$ is known.

To be more specific, consider (4.5). From the Grönwall bound (4.3) established in step 1, it is clear that the Fourier-Laplace transform would exist if one chooses $\mu_n < -2C$, however it is not clear that we can perform the computation as μ_n approaches the imaginary axis. In order to avoid a circular argument –establishing some time decay by assuming the existence of a Fourier-Laplace transform which already requires some time decay–, we can appeal to several arguments:

1. We can use as a black box the *Paley-Wiener theory* (see [72, Chap. 18] or [33, Chap. 2]): for every $f \in L^1_{loc}(\mathbb{R}_+)$ there exists a unique solution $u \in L^1_{loc}(\mathbb{R}_+)$ to the integral equation $u = f + k * u$ with $k \in L^1_t$ and with $k \leq e^{Ct}$ for some constant $C > 0$ (where the convolution over $t \in \mathbb{R}_+$ is defined as before by extending functions to zero on \mathbb{R}_-), given by $u = f - f * r$ where $r \in L^1_{loc}(\mathbb{R}_+)$ is the so-called resolvent kernel of k . The latter is the unique solution to $r = k + r * k$ and the key result of the theory is that $r \in L^1(\mathbb{R}_+)$ iff the Fourier-Laplace transform of k satisfies $\mathcal{L}[k](\xi) \neq 1$ for any $\Re e \xi \leq 0$. As soon as $r, f \in L^1(\mathbb{R}_+)$ we have $u \in L^1(\mathbb{R}_+)$. Then step 3 of Lemma 4.1 can be justified by applying this theory to $u(t) := \Phi_k^n(t)$, $f(t) := R_k^n(t)$ and $k(t) := e^{\mu_n|k|t} K^0(t, k)$.
2. A second method is to use an approximation argument, in the spirit of energy methods in PDEs, which was discussed in [84, Section3]. Define $\Phi_{k,\delta}^n(t) := \Phi_k^n(t)e^{-\delta t^2/2}$. Now we have the existence of Fourier-Laplace transform of $\Phi_{k,\delta}^n(t)$ for any μ_n thanks to the Gaussian decay in time and (4.3). Using the Fourier-Laplace transform on (4.5) and (L), we may deduce that for $\mu_n < -2C$ (where $2C$ comes from the estimate (4.3)) that the following formula holds:

$$\hat{\Phi}_{k,\delta}^n(\omega) = \left(\frac{\hat{R}_k^n(\omega)}{1 - \mathcal{L}\left(k, \mu_n + i\frac{\omega}{|k|}\right)} \right) * \gamma_\delta, \quad \gamma_\delta(\omega) := \frac{e^{-\frac{|\omega|^2}{2\delta}}}{\sqrt{2\pi\delta}}.$$

Since this is an analytic function in μ_n and ω as long as we do not approach a singularity, by analytic continuation we may deduce that this formula remains valid for all $\mu_n < \bar{\lambda}$ by (L). Therefore by Plancherel’s theorem

$$\|\Phi_{k,\delta}^n(t)\|_{L^2_t(\mathbb{R})} \lesssim \frac{1}{\kappa} \|R_k^n(t)\|_{L^2_t(\mathbb{R})} \|\gamma_\delta\|_{L^1(\mathbb{R})} \lesssim \frac{1}{\kappa} \|R_k^n(t)\|_{L^2_t(\mathbb{R})}$$

which is an estimate independent of $\delta > 0$. We then let δ go to zero and deduce by Fatou’s lemma the desired bound (4.6):

$$\|\Phi_k^n(t)\|_{L^2_t(\mathbb{R})} \lesssim \frac{1}{\kappa} \|R_k^n(t)\|_{L^2_t(\mathbb{R})}$$

(which also justifies the existence of the Fourier-Laplace transform).

Remark 16 Let us mention to finish with that the present discussion is related to the Gerhart-Herbst-Prüss-Greiner theorem [2, 26, 39, 75] (see also [24]) for semigroups in Hilbert spaces. The latter asserts that the semigroup decay is given by the spectral bound, under a sole pointwise control on the resolvent. While the constants seem to be non-constructive in the first versions of this theorem, Engel and Nagel gave a comprehensive and elementary proof with constructive constant in [24, Theorem 1.10; chapter V]. Let us also mention on the same subject subsequent more recent works like [38]. The main idea in the proof of [24, Theorem 1.10, chapter V], which is also used in [38], is to use a Plancherel identity on the resolvent in Hilbert spaces in order to obtain explicit rates of decay on the semigroup in terms of bounds on the resolvent. However in the Volterra integral equation we study here, there is no semigroup structure on the unknown $\phi_k(t)$, and we cannot appeal directly to these classical results.

4.3 The Penrose Criterion

A generalized form of the Penrose criterion [73] was given in [67] as follows:

$$(\mathbf{P}) \quad \forall k \in \mathbb{Z}_*^d \text{ and } r \in \mathbb{R} \text{ s.t. } (f_k^0)'(r) = 0, \quad \widehat{W}(k) \left(\text{p.v.} \int_{\mathbb{R}} \frac{(f_k^0)'(r)}{r - w} dr \right) < 1, \tag{4.13}$$

where f_k^0 denotes the marginals of the background f^0 along arbitrary wave vector $k \in \mathbb{Z}_*^d$:

$$f_k^0(r) := \int_{kr/|k|+k^\perp} f^0(w) dw, \quad r \in \mathbb{R}.$$

The proof that condition **(P)** implies the condition **(L)** was not quite complete in [67] as it was proved only that **(P)** implies the lower bound $|1 - \mathcal{L}(k, \xi)| \geq \kappa$ in a strip and not in a half-plane. The complete proof due to Penrose relies on the argument principle. The starting point is to observe that $\mathcal{L}(k, \xi) = \int_0^{+\infty} e^{\xi|k|t} K^0(t, k) dt$ with $\xi = \lambda + i\zeta$ and $K^0(t, k) := -\widehat{f}^0(kt) \widehat{W}(k) |k|^2 t$ is well-defined for $\lambda < \bar{\lambda}$ by the analyticity of f^0 and is small for large ζ by integration by parts. We therefore restrict ourselves to a compact interval $|\zeta| \leq C$ in ζ , and we compute by the Plemelj formula (see [67] for more details) that

$$\mathcal{L}(k, i\zeta) = \widehat{W}(k) \left[\left(\text{p.v.} \int_{\mathbb{R}} \frac{(f_k^0)'(r)}{r - \zeta} dr \right) - i\pi (f_k^0)'(\zeta) \right]. \tag{4.14}$$

Therefore, the condition **(P)** implies that $|1 - \mathcal{L}(k, \xi)| \geq 2\kappa$ for some $\kappa > 0$ at $\lambda = 0$ and for $|\zeta| \leq C$. Combined with smallness for large ζ and continuity, we deduce the lower bound $|1 - \mathcal{L}(k, \xi)| \geq \kappa$ in a strip $\Re e \xi \in [0, \lambda']$ for some $\lambda' > 0$. Since the function $\xi \mapsto \mathcal{L}(k, \xi)$ is holomorphic on $\Re e \xi < \bar{\lambda}$ and the value 1 is not taken on $i\mathbb{R}$, by the argument principle, the value 1 can only be taken on $\Re e \xi < 0$

if $\Xi : \zeta \mapsto \mathcal{L}(k, i\zeta)$ has a positive winding number around this value. However, this would imply that the curve Ξ crosses the real axis above 1, which is prohibited again by (4.14) and (P), which concludes the proof.

5 Energy Estimates on the (ρ, g) System

In this section we perform the necessary energy estimates to prove Proposition 2.2, that is, we deduce the multi-tier controls stated in (2.12) from (2.11) for suitable K_i and sufficiently small ϵ .

5.1 $L^2_t(I)$ Estimates on the Density (Equation (2.12c))

The most fundamental estimate we need to make is the $L^2_t(I)$ control (2.12c), which requires the linear damping Lemma 4.1 and the crucial plasma echo analysis carried out in §6 (which we apply as a black box in this section). The controls (2.11b) and (2.11a) were chosen specifically for this.

To highlight its primary importance, we state the inequality as a separate proposition.

Proposition 5.1 (Nonlinear control of ρ) *For suitable K_i and ϵ_0 sufficiently small, the estimate (2.12c) holds under the bootstrap hypotheses (2.11).*

Proof Proposition 5.1 requires two main controls, which follow from (2.11b) and (2.11a) respectively.

- (a) Define the time-response kernel $\bar{K}_{k,\ell}(t, \tau)$ for some $c = c(s) \in (0, 1)$ (determined by the proof):

$$\begin{aligned} \bar{K}_{k,\ell}(t, \tau) &= \frac{1}{|\ell|^\gamma} e^{(\lambda(t)-\lambda(\tau))\langle k, kt \rangle^s} e^{c\lambda(\tau)\langle k-\ell, kt-\ell\tau \rangle^s} \\ &\quad \times |k(t-\tau)\hat{g}_{k-\ell}(\tau, kt-\ell\tau)| \mathbf{1}_{\ell \neq 0}. \end{aligned} \tag{5.1}$$

Proposition 5.1 will depend on the estimate

$$\left(\sup_{t \geq 0} \sup_{k \in \mathbb{Z}_*^d} \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) \, d\tau \right) \left(\sup_{\tau \geq 0} \sup_{\ell \in \mathbb{Z}_*^d} \sum_{k \in \mathbb{Z}_*^d} \int_\tau^\infty \bar{K}_{k,\ell}(t, \tau) \, dt \right) \lesssim K_2 \epsilon^2. \tag{5.2}$$

- (b) Proposition 5.1 will also depend on the estimate

$$\sup_{\tau \geq 0} e^{(c-1)\omega_0(\tau)^s} \sum_{k \in \mathbb{Z}_*^d} \sup_{\omega \in \mathbb{Z}_*^d} \sup_{x \in \mathbb{R}^d} \int_{-\infty}^\infty \left| (A\widehat{\nabla}g)_k \left(\tau, \frac{\omega}{|\omega|}\zeta - x \right) \right|^2 d\zeta \lesssim K_1 \epsilon^2. \tag{5.3}$$

The condition (5.2) controls *reaction*: the interaction of the density with the lower frequencies of g ; condition (5.3) controls *transport*: the interaction with higher fre-

quencies of g . The latter, condition (5.3), follows from (2.11a): by Lemma 3.4 followed by (2.14),

$$\sum_{k \in \mathbb{Z}^d} \sup_{\omega \in \mathbb{Z}_*^d} \sup_{x \in \mathbb{R}^d} \int_{-\infty}^{\infty} \left| (A\widehat{\nabla}g)_k \left(\tau, \frac{\omega}{|\omega|} \zeta - x \right) \right|^2 d\zeta \lesssim_M \|A\widehat{\nabla}g(\tau)\|_{L_x^2 H_\eta^M}^2 \lesssim K_1 \epsilon^2 \langle \tau \rangle^7,$$

from which (5.3) follows by (3.8) and $c < 1$. Since condition (5.2) is much harder to verify and contains the physical mechanism of the plasma echoes, we prove Proposition 5.1 assuming (5.2). In §6 below, we prove that (5.2) follows from (2.11b).

Expanding the integral equation (2.6) using the paraproduct decomposition:

$$\begin{aligned} \hat{\rho}_k(t) &= \hat{h}_{\text{in}}(k, kt) - \int_0^t \hat{\rho}_k(\tau) |k|^2 \widehat{W}(k)(t - \tau) \hat{f}^0(k(t - \tau)) d\tau \\ &\quad - \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \sum_{N \geq 8} \hat{\rho}_\ell(\tau) \widehat{W}(\ell) \ell \cdot k(t - \tau) \hat{g}_{k-\ell}(\tau, kt - \ell\tau)_N d\tau \\ &\quad - \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \sum_{N \geq 8} \hat{\rho}_\ell(\tau)_N \widehat{W}(\ell) \ell \cdot k(t - \tau) \hat{g}_{k-\ell}(\tau, kt - \ell\tau)_{<N/8} d\tau \\ &\quad - \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \sum_{N \in \mathbb{D}} \sum_{N/8 \leq N' \leq 8N} \hat{\rho}_\ell(\tau)_{N'} \widehat{W}(\ell) \ell \cdot k(t - \tau) \hat{g}_{k-\ell}(\tau, kt - \ell\tau)_N d\tau, \\ &= \hat{h}_{\text{in}}(k, kt) - \int_0^t \hat{\rho}_k(\tau) |k|^2 \widehat{W}(k)(t - \tau) \hat{f}^0(k(t - \tau)) d\tau \\ &\quad - T_k(t) - R_k(t) - \mathcal{R}_k(t). \end{aligned} \tag{5.4}$$

Recall our convention that the Littlewood-Paley projection of $\hat{\rho}_\ell(\tau)$ treats $\ell\tau$ in place of the v frequency. We begin by applying Lemma 4.1 to (5.4), which implies for each $k \in \mathbb{Z}_*^d$,

$$\begin{aligned} \|A\hat{\rho}_k\|_{L_t^2(I)}^2 &\lesssim C_{LD} \left\| A_k(\cdot, k \cdot) \hat{h}_{\text{in}}(k, k \cdot) \right\|_{L_t^2(I)}^2 + C_{LD} \|AT_k\|_{L_t^2(I)}^2 \\ &\quad + C_{LD} \|AR_k\|_{L_t^2(I)}^2 + C_{LD} \|A\mathcal{R}_k\|_{L_t^2(I)}^2. \end{aligned} \tag{5.5}$$

First, Lemma 3.4 and a version of the argument applied in (2.13) (using also that $\lambda(t)$ is decreasing (2.8)) imply

$$\begin{aligned} \sum_{k \in \mathbb{Z}_*^d} \left\| A_k(\cdot, k \cdot) \hat{h}_{\text{in}}(k, k \cdot) \right\|_{L_t^2(I)}^2 &= \sum_{k \in \mathbb{Z}_*^d} \int_0^{T^*} \left| A_k(t, kt) \hat{h}_{\text{in}}(k, kt) \right|^2 dt \\ &\lesssim \sum_{k \in \mathbb{Z}_*^d} \left\| A_k(0, \cdot) \hat{h}_{\text{in}}(k, \cdot) \right\|_{H^M(\mathbb{R}_\eta^d)}^2 \lesssim \epsilon^2. \end{aligned} \tag{5.6}$$

□

Now we turn to the nonlinear contributions in (5.4).

5.1.1 Reaction

Our goal is to show

$$\|AR\|_{L_k^2 L_t^2(I)}^2 \lesssim K_2 \epsilon^2 \|A\hat{\rho}\|_{L_k^2 L_t^2(I)}^2, \tag{5.7}$$

since for ϵ chosen sufficiently small, this contribution can then be absorbed on the LHS of (5.5).

First, by applying (1.3),

$$\begin{aligned} \|AR\|_{L_k^2 L_t^2(I)}^2 &\lesssim \sum_{k \in \mathbb{Z}_*^d} \int_0^{T^*} \left[A_k(t, kt) \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \sum_{N \geq 8} \left| \hat{g}_{k-\ell}(\tau, kt - \ell\tau)_{<N/8} \right. \right. \\ &\quad \left. \left. \frac{|k(t - \tau)|}{|\ell|^\gamma} \hat{\rho}_\ell(\tau)_N \right| d\tau \right]^2 dt. \end{aligned}$$

By definition, the Littlewood-Paley projections imply the frequency localizations (as in (3.15)):

$$\frac{N}{2} \leq |\ell| + |\ell\tau| \leq \frac{3N}{2}, \tag{5.8a}$$

$$|k - \ell| + |kt - \ell\tau| \leq \frac{3N}{32}, \tag{5.8b}$$

$$\frac{13}{16} \leq \frac{|k, kt|}{|\ell, \tau\ell|} \leq \frac{19}{16}. \tag{5.8c}$$

From (5.8), on the support of the integrand, (3.11) implies that for some $c = c(s) \in (0, 1)$:

$$A_k(t, kt) = e^{(\lambda(t) - \lambda(\tau))\langle k, kt \rangle^s} A_k(\tau, kt) \lesssim e^{(\lambda(t) - \lambda(\tau))\langle k, kt \rangle^s} e^{c\lambda(\tau)\langle k - \ell, kt - \ell\tau \rangle^s} A_\ell(\tau, \ell\tau). \tag{5.9}$$

Therefore, by definition of \bar{K} we have (dropping the Littlewood-Paley projection on g),

$$\|AR\|_{L_k^2 L_t^2(I)}^2 \lesssim \sum_{k \in \mathbb{Z}_*^d} \int_0^{T^*} \left[\int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) A_\ell(\tau, \ell\tau) \sum_{N \geq 8} |\hat{\rho}_\ell(\tau)_N| d\tau \right]^2 dt.$$

Since the Littlewood-Paley projections define a partition of unity,

$$\|AR\|_{L_k^2 L_t^2(I)}^2 \lesssim \sum_{k \in \mathbb{Z}_*^d} \int_0^{T^*} \left[\int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) A_\ell(\tau, \ell\tau) |\hat{\rho}_\ell(\tau)| d\tau \right]^2 dt.$$

From here we may proceed analogous to §7 in [67], for which we apply Schur’s test in $L_k^2 L_t^2$. Indeed,

$$\begin{aligned} \|AR\|_{L_k^2 L_t^2(I)}^2 &\lesssim \sum_{k \in \mathbb{Z}_*^d} \int_0^{T^*} \left(\int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) \, d\tau \right) \\ &\quad \times \left(\int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) |A_\ell(\tau, \ell\tau) \hat{\rho}_\ell(\tau)|^2 \, d\tau \right) dt \\ &\leq \left(\sup_{t \geq 0} \sup_{k \in \mathbb{Z}_*^d} \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) \, d\tau \right) \\ &\quad \times \sum_{k \in \mathbb{Z}_*^d} \int_0^{T^*} \left(\int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) |A_\ell(\tau, \ell\tau) \hat{\rho}_\ell(\tau)|^2 \, d\tau \right) dt. \end{aligned}$$

By Fubini’s theorem,

$$\begin{aligned} \|AR\|_{L_k^2 L_t^2(I)}^2 &\lesssim \left(\sup_{t \geq 0} \sup_{k \in \mathbb{Z}_*^d} \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) \, d\tau \right) \\ &\quad \times \sum_{\ell \in \mathbb{Z}_*^d} \int_0^{T^*} \left(\int_\tau^{T^*} \sum_{k \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) \, dt \right) |A_\ell(\tau, \ell\tau) \hat{\rho}_\ell(\tau)|^2 \, d\tau \\ &\lesssim \left(\sup_{t \geq 0} \sup_{k \in \mathbb{Z}_*^d} \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) \, d\tau \right) \\ &\quad \times \left(\sup_{\tau \geq 0} \sup_{\ell \in \mathbb{Z}_*^d} \sum_{k \in \mathbb{Z}_*^d} \int_\tau^{T^*} \bar{K}_{k,\ell}(t, \tau) \, dt \right) \|A \hat{\rho}\|_{L_k^2 L_t^2(I)}^2. \end{aligned}$$

Hence, condition (5.2) implies (5.7).

5.1.2 Transport

As above, first apply (1.3),

$$\begin{aligned} \|AT\|_{L_k^2 L_t^2(I)}^2 &\lesssim \sum_{k \in \mathbb{Z}_*^d} \int_0^{T^*} \left[A_k(t, kt) \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \sum_{N \geq 8} \left| \hat{g}_{k-\ell}(\tau, kt - \ell\tau)_N \right. \right. \\ &\quad \left. \left. \frac{|k(t-\tau)|}{|\ell|^\gamma} \hat{\rho}_\ell(\tau)_{<N/8} \right| \, d\tau \right]^2 dt. \end{aligned}$$

By the Littlewood-Paley projections, on the support of the integrand there holds,

$$\frac{N}{2} \leq |k - \ell| + |kt - \ell\tau| \leq \frac{3N}{2} \tag{5.10a}$$

$$|\ell| + |\ell\tau| \leq \frac{3N}{32} \tag{5.10b}$$

$$\frac{13}{16} \leq \frac{|k, kt|}{|k - \ell, kt - \ell\tau|} \leq \frac{19}{16}. \tag{5.10c}$$

By (5.10), on the support of the integrand, (3.11) implies that for some $c = c(s) \in (0, 1)$:

$$\begin{aligned} A_k(t, kt) &= e^{(\lambda(t)-\lambda(\tau))\langle k, kt \rangle^s} A_k(\tau, kt) \\ &\lesssim e^{(\lambda(t)-\lambda(\tau))\langle k, kt \rangle^s} e^{c\lambda(\tau)\langle \ell, \ell\tau \rangle^s} A_{k-\ell}(\tau, kt - \ell\tau). \end{aligned}$$

Using that

$$|k(t - \tau)| \leq |kt - \ell\tau| + \tau |k - \ell| \leq \langle \tau \rangle |k - \ell, kt - \ell\tau|, \tag{5.11}$$

we have (ignoring the Littlewood-Paley projection on ρ and the $|\ell|^{-\gamma}$ which are not helpful),

$$\begin{aligned} \|AT\|_{L_k^2 L_t^2(I)} &\lesssim \sum_{k \in \mathbb{Z}_*^d} \int_0^{T^*} \left[\sum_{\ell \in \mathbb{Z}_*^d} \int_0^t \sum_{N \geq 8} |(A\widehat{\nabla}g)_{k-\ell}(\tau, kt - \ell\tau)_N| \right. \\ &\quad \left. e^{c\lambda(\tau)\langle \ell, \ell\tau \rangle^s} \langle \tau \rangle |\hat{\rho}_\ell(\tau)| \, d\tau \right]^2 dt. \end{aligned}$$

Since the Littlewood-Paley projections define a partition of unity (and using also the Cauchy-Schwarz inequality),

$$\begin{aligned} \|AT\|_{L_k^2 L_t^2(I)}^2 &\lesssim \sum_{k \in \mathbb{Z}_*^d} \int_0^{T^*} \left[\sum_{\ell \in \mathbb{Z}_*^d} \int_0^t e^{c\lambda(\tau)\langle \ell, \ell\tau \rangle^s} \langle \tau \rangle |\hat{\rho}_\ell(\tau)| \, d\tau \right] \\ &\quad \times \left[\sum_{\ell \in \mathbb{Z}_*^d} \int_0^t |(A\widehat{\nabla}g)_{k-\ell}(\tau, kt - \ell\tau)|^2 e^{c\lambda(\tau)\langle \ell, \ell\tau \rangle^s} \langle \tau \rangle |\hat{\rho}_\ell(\tau)| \, d\tau \right] dt. \end{aligned}$$

By Cauchy-Schwarz and $\sigma > \frac{d}{2} + 2$,

$$\begin{aligned} \sum_{\ell \in \mathbb{Z}_*^d} \int_0^t e^{c\lambda(\tau)\langle \ell, \ell\tau \rangle^s} \langle \tau \rangle |\hat{\rho}_\ell(\tau)| \, d\tau &\leq \left(\int_0^t \sum_{\ell \in \mathbb{Z}_*^d} e^{2(c-1)\lambda(\tau)\langle \ell, \ell\tau \rangle^s} \langle \tau \rangle^2 \langle \ell, \ell\tau \rangle^{-2\sigma} \, d\tau \right)^{1/2} \|A\hat{\rho}\|_{L_k^2 L_t^2(I)} \\ &\lesssim \|A\hat{\rho}\|_{L_k^2 L_t^2(I)}. \end{aligned} \tag{5.12}$$

Then (5.12) and Fubini’s theorem imply

$$\begin{aligned}
 & \|AT\|_{L_k^2 L_t^2(I)} \\
 & \lesssim \|A\hat{\rho}\|_{L_k^2 L_t^2(I)} \sum_{k \in \mathbb{Z}_*^d} \int_0^{T^*} \sum_{\ell \in \mathbb{Z}_*^d} \int_0^t |(A\widehat{\nabla}g)_{k-\ell}(\tau, kt - \ell\tau)|^2 e^{c\lambda(\tau)(\ell, \ell\tau)^s} \langle \tau \rangle |\hat{\rho}_\ell(\tau)| \, d\tau \, dt \\
 & \lesssim \|A\hat{\rho}\|_{L_k^2 L_t^2(I)} \sum_{\ell \in \mathbb{Z}_*^d} \int_0^{T^*} \left(\sum_{k \in \mathbb{Z}_*^d} \int_\tau^{T^*} |(A\widehat{\nabla}g)_{k-\ell}(\tau, kt - \ell\tau)|^2 \, dt \right) e^{c\lambda(\tau)(\ell, \ell\tau)^s} \langle \tau \rangle |\hat{\rho}_\ell(\tau)| \, d\tau \\
 & \leq \|A\hat{\rho}\|_{L_k^2 L_t^2(I)} \sum_{\ell \in \mathbb{Z}_*^d} \int_0^{T^*} \left(\sum_{k \in \mathbb{Z}_*^d} \int_{-\infty}^\infty |(A\widehat{\nabla}g)_{k-\ell}(\tau, kt - \ell\tau)|^2 \, dt \right) e^{c\lambda(\tau)(\ell, \ell\tau)^s} \langle \tau \rangle |\hat{\rho}_\ell(\tau)| \, d\tau \\
 & \leq \|A\hat{\rho}\|_{L_k^2 L_t^2(I)} \left(\sup_{\tau \geq 0} e^{(c-1)\alpha_0(\tau)^s} \sum_{k \in \mathbb{Z}^d} \sup_{\omega \in \mathbb{Z}_*^d} \sup_{x \in \mathbb{R}^d} \int_{-\infty}^\infty \left| (A\widehat{\nabla}g)_k \left(\tau, \frac{\omega}{|\omega|} \zeta - x \right) \right|^2 \, d\zeta \right) \\
 & \quad \times \sum_{\ell \neq 0} \int_0^{T^*} e^{\lambda(\tau)(\ell, \ell\tau)^s} \langle \tau \rangle |\hat{\rho}_\ell(\tau)| \, d\tau.
 \end{aligned}$$

Proceeding as in (5.12) then gives

$$\begin{aligned}
 & \|AT\|_{L_k^2 L_t^2(I)} \lesssim \|A\hat{\rho}\|_{L_k^2 L_t^2(I)}^2 \\
 & \quad \times \left(\sup_{\tau \geq 0} e^{(c-1)\alpha_0(\tau)^s} \sum_{k \in \mathbb{Z}^d} \sup_{\omega \in \mathbb{Z}_*^d} \sup_{x \in \mathbb{R}^d} \int_{-\infty}^\infty \left| (A\widehat{\nabla}g)_k \left(\tau, \frac{\omega}{|\omega|} \zeta - x \right) \right|^2 \, d\zeta \right).
 \end{aligned}$$

Using condition (5.3), we derive

$$\|AT\|_{L_k^2 L_t^2(I)} \lesssim_{\alpha_0} K_1 \epsilon^2 \|A\hat{\rho}\|_{L_k^2 L_t^2(I)}^2, \tag{5.13}$$

which suffices to treat transport.

5.1.3 Remainders

We treat the remainder with a variant of the method used to treat transport. First, by (1.3):

$$\begin{aligned}
 & \|A\mathcal{R}\|_{L_k^2 L_t^2(I)}^2 \lesssim \sum_{k \in \mathbb{Z}_*^d} \int_0^{T^*} \left[A_k(t, kt) \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \sum_{N \in \mathbb{D}} \sum_{N/8 \leq N' \leq 8N} |\hat{\rho}_\ell(\tau)_N| |k(t - \tau)| \right. \\
 & \quad \left. |\hat{g}_{k-\ell}(\tau, kt - \ell\tau)_N| \, d\tau \right]^2 \, dt.
 \end{aligned}$$

Next we claim that on the integrand there holds for some $c' = c'(s) \in (0, 1)$,

$$A_k(t, kt) \lesssim_{\lambda_0, \alpha_0} e^{c'\lambda(\tau)\langle k-\ell, kt-\ell\tau \rangle^s} e^{c'\lambda(\tau)\langle \ell, \ell\tau \rangle^s}. \tag{5.14}$$

Indeed, this follows simply by following the argument used to deduce (3.16), with kt replace η . Therefore by (5.14), Cauchy-Schwarz,

$$\begin{aligned} \|A\mathcal{R}\|_{L_k^2 L_t^2(I)}^2 &\lesssim \sum_{k \in \mathbb{Z}_*^d} \int_0^{T^*} \left[\int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \left(\sum_{N \in \mathbb{D}} e^{2\lambda(\tau)\langle \ell, \ell\tau \rangle^s} \langle \tau \rangle^2 |\hat{\rho}_\ell(\tau)_{\sim N}|^2 \right) d\tau \right] \\ &\times \left[\int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \left(\sum_{N \in \mathbb{D}} e^{2(c'-1)\lambda(\tau)\langle \ell, \ell\tau \rangle^s} e^{2c'\lambda(\tau)\langle k-\ell, kt-\ell\tau \rangle^s} \right. \right. \\ &\left. \left. |\widehat{\nabla} g_{k-\ell}(\tau, kt - \ell\tau)_N|^2 \right) d\tau \right] dt. \end{aligned}$$

By the almost orthogonality of the Littlewood-Paley decomposition (3.4) and $\sigma > 1$,

$$\begin{aligned} \|A\mathcal{R}\|_{L_k^2 L_t^2(I)}^2 &\lesssim \|A\hat{\rho}\|_{L_k^2 L_t^2(I)}^2 \sum_{k \in \mathbb{Z}_*^d} \int_0^{T^*} \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \left(\sum_{N \in \mathbb{D}} e^{2(c'-1)\lambda(\tau)\langle \ell, \ell\tau \rangle^s} \right. \\ &\left. \times e^{2c'\lambda(\tau)\langle k-\ell, kt-\ell\tau \rangle^s} |\hat{g}_{k-\ell}(\tau, kt - \ell\tau)_N|^2 \right) d\tau dt. \end{aligned}$$

By Fubini's theorem and Lemma 3.4,

$$\begin{aligned} \|A\mathcal{R}\|_{L_k^2 L_t^2(I)}^2 &\lesssim \|A\hat{\rho}\|_{L_k^2 L_t^2(I)}^2 \int_0^{T^*} \sum_{\ell \in \mathbb{Z}_*^d} e^{2(c'-1)\lambda(\tau)\langle \ell, \ell\tau \rangle^s} \\ &\times \left(\sum_{N \in \mathbb{D}} \sum_{k \in \mathbb{Z}_*^d} \int_\tau^{T^*} |A\widehat{\nabla} g_{k-\ell}(\tau, kt - \ell\tau)_N|^2 dt \right) d\tau \\ &\lesssim \|A\hat{\rho}\|_{L_k^2 L_t^2(I)}^2 \int_0^{T^*} \sum_{\ell \in \mathbb{Z}_*^d} e^{2(c'-1)\lambda(\tau)\langle \ell, \ell\tau \rangle^s} \\ &\times \left(\sum_{N \in \mathbb{D}} \sum_{k \in \mathbb{Z}^d} \sup_{\omega \in \mathbb{Z}_*^d} \sup_{x \in \mathbb{R}^d} \int_{-\infty}^{\infty} \left| A\widehat{\nabla} g_k \left(\tau, \frac{\omega}{|\omega|} \zeta - x \right)_N \right|^2 d\zeta \right) d\tau \end{aligned}$$

$$\begin{aligned} &\lesssim \|A\hat{\rho}\|_{L_k^2 L_t^2(I)}^2 \int_0^{T^*} \sum_{\ell \in \mathbb{Z}_*^d} e^{2(c'-1)\lambda(\tau)\langle \ell, \ell \tau \rangle^s} \\ &\quad \times \left(\sum_{N \in \mathbb{D}} \sum_{k \in \mathbb{Z}^d} \|A^{(1)} \hat{g}_k(\tau)_N\|_{H^M(\mathbb{R}_\eta^d)}^2 \right) d\tau. \end{aligned}$$

The Littlewood-Paley projections do not commute with derivatives in frequency space, however since the projections have bounded derivatives we still have (see §3),

$$\begin{aligned} \|A\mathcal{R}\|_{L_k^2 L_t^2(I)}^2 &\lesssim_M \|A\hat{\rho}\|_{L_k^2 L_t^2(I)}^2 \int_0^{T^*} \sum_{\ell \in \mathbb{Z}_*^d} e^{2(c'-1)\lambda(\tau)\langle \ell, \ell \tau \rangle^s} \\ &\quad \times \left(\sum_{N \in \mathbb{D}} \sum_{|\alpha| \leq M} \|(D_\eta^\alpha A^{(1)} \hat{g}_k)(\tau)_{\sim N}\|_{L_k^2 L_\eta^2}^2 \right) d\tau. \end{aligned}$$

Then by the almost orthogonality (3.4) with (2.14), (3.8) and $c' < 1$, we have

$$\begin{aligned} \|A\mathcal{R}\|_{L_k^2 L_t^2(I)}^2 &\lesssim K_1 \epsilon^2 \|A\hat{\rho}\|_{L_k^2 L_t^2(I)}^2 \int_0^{T^*} \sum_{\ell \in \mathbb{Z}_*^d} e^{2(c'-1)\lambda(\tau)\langle \ell, \ell \tau \rangle^s} \langle \tau \rangle^7 d\tau \\ &\lesssim K_1 \epsilon^2 \|A\hat{\rho}\|_{L_k^2 L_t^2(I)}^2, \end{aligned} \tag{5.15}$$

which suffices to treat remainder contributions.

5.1.4 Conclusion of L^2 Bound

Putting (5.6), (5.7), (5.13) and (5.15) together with (5.5) we have for some $\tilde{K} = \tilde{K}(s, d, M, \lambda_0, \alpha_0)$,

$$\|A\hat{\rho}\|_{L_k^2 L_t^2(I)}^2 \leq \tilde{K} C_{LD} \epsilon^2 + \tilde{K} C_{LD} (K_1 + K_2) \epsilon^2 \|A\hat{\rho}(t)\|_{L_k^2 L_t^2(I)}^2.$$

Therefore for $\epsilon^2 < \frac{1}{2}(\tilde{K} C_{LD} (K_1 + K_2))^{-1}$ we have

$$\|A\hat{\rho}(t)\|_{L_k^2 L_t^2(I)}^2 < 2\tilde{K} C_{LD} \epsilon^2.$$

Hence, Proposition 5.1 follows provided we fix $K_3 = \tilde{K} C_{LD}$.

5.2 Pointwise-in-time Estimate on the Density

The constant K_3 basically only depends on the linearized Vlasov equation with homogeneous background f^0 . The same is not true of the pointwise-in-time estimate we deduce next.

Lemma 5.2 (Pointwise estimate) *For ϵ_0 sufficiently small, under the bootstrap hypotheses (2.11), there exists some $K_4 = K_4(C_0, \bar{\lambda}, \kappa, M, s, d, \lambda_0, \lambda', K_1, K_2, K_3)$ such that for $t \in [0, T^*]$,*

$$\|A\rho(t)\|_2^2 \leq K_4 \langle t \rangle \epsilon^2. \tag{5.16}$$

Proof As in [67], we use the L_t^2 bound together with (2.6). Our starting point is again the paraproduct decomposition (5.4):

$$\begin{aligned} \|A\hat{\rho}(t)\|_{L_k^2} &\lesssim \sum_{k \in \mathbb{Z}_*^d} \left| A_k(t, kt) \hat{h}_{\text{in}}(k, kt) \right|^2 \\ &\quad + \sum_{k \in \mathbb{Z}_*^d} \left(A_k(t, kt) \int_0^t \hat{\rho}_k(\tau) |k|^2 \widehat{W}(k)(t - \tau) \hat{f}^0(k(t - \tau)) \, d\tau \right)^2 \\ &\quad + \sum_{k \in \mathbb{Z}_*^d} |A_k(t, kt) T_k(t)|^2 + \sum_{k \in \mathbb{Z}_*^d} |A_k(t, kt) R_k(t)|^2 \\ &\quad + \sum_{k \in \mathbb{Z}_*^d} |A_k(t, kt) \mathcal{R}_k(t)|^2. \end{aligned} \tag{5.17}$$

To treat the initial data we use the $H^{d/2+} \hookrightarrow C^0$ embedding and that $\lambda(t)$ is decreasing (2.8):

$$\sum_{k \in \mathbb{Z}_*^d} \left| A_k(t, kt) \hat{h}_{\text{in}}(k, kt) \right|^2 \leq \sum_{k \in \mathbb{Z}_*^d} \sup_{\eta \in \mathbb{R}^d} |A_k(t, \eta) h_{\text{in}}(k, \eta)|^2 \lesssim \|A(0) \hat{h}_{\text{in}}\|_{L_k^2 H_\eta^M}^2 \lesssim \epsilon^2, \tag{5.18}$$

where we used an argument analogous to (2.13) to deduce the last inequality. □

5.2.1 Linear Contribution

Next consider the term in (5.17) coming from the homogeneous background. By (1.3), (3.9), (3.8), and the $H^{d/2+} \hookrightarrow C^0$ embedding with (3.7), (1.9) and (2.12c),

$$\begin{aligned} &\sum_{k \in \mathbb{Z}_*^d} \left(A_k(t, kt) \int_0^t \hat{\rho}_k(\tau) |k|^2 \widehat{W}(k)(t - \tau) \hat{f}^0(k(t - \tau)) \, d\tau \right)^2 \\ &\lesssim \sum_{k \in \mathbb{Z}_*^d} \left(\int_0^t A_k(\tau, k\tau) |\hat{\rho}_k(\tau)| \langle k(t - \tau) \rangle^{\sigma+1} e^{\lambda(t)(k(t-\tau))^s} \left| \hat{f}^0(k(t - \tau)) \right| \, d\tau \right)^2 \\ &\lesssim \left(\sup_{\eta} e^{\lambda_0(\eta)^s} \left| \hat{f}^0(\eta) \right| \right)^2 \sum_{k \in \mathbb{Z}_*^d} \left(\int_0^t A_k(\tau, k\tau) |\hat{\rho}_k(\tau)| e^{\frac{1}{2}(\lambda(0) - \lambda_0)(t-\tau)^s} \, d\tau \right)^2 \end{aligned}$$

$$\begin{aligned}
 &\lesssim C_0^2 \left(\sum_{k \in \mathbb{Z}_*^d} \int_0^t |A_k(\tau, k\tau) \hat{\rho}_k(\tau)|^2 d\tau \right) \left(\int_0^t e^{(\lambda(0) - \lambda_0)(t-\tau)^s} d\tau \right) \\
 &\lesssim \sum_{k \in \mathbb{Z}_*^d} \|A \hat{\rho}_k\|_{L_t^2(I)}^2 \\
 &\lesssim K_3 \epsilon^2.
 \end{aligned} \tag{5.19}$$

5.2.2 Reaction

Next we treat the reaction term in (5.17), which by (1.3),

$$\begin{aligned}
 \sum_{k \in \mathbb{Z}_*^d} |A_k(kt) R_k|^2 &\lesssim \sum_{k \in \mathbb{Z}_*^d} \left[A_k(t, kt) \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \sum_{N \geq 8} |\hat{g}_{k-\ell}(\tau, kt - \ell\tau)_{<N/8}| \right. \\
 &\quad \left. \frac{|k(t-\tau)|}{|\ell|^\gamma} |\hat{\rho}_\ell(\tau)_N| d\tau \right]^2.
 \end{aligned}$$

As in the L_t^2 estimate we have by (5.9) and the definition (5.1) of \bar{K} :

$$\sum_{k \in \mathbb{Z}_*^d} |A_k(kt) R_k(t)|^2 \lesssim \sum_{k \in \mathbb{Z}_*^d} \left[\int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) |A_\ell(\tau, \ell\tau) \hat{\rho}_\ell(\tau)| d\tau \right]^2.$$

By the Cauchy-Schwarz inequality and Fubini's theorem,

$$\begin{aligned}
 \sum_{k \in \mathbb{Z}_*^d} |A_k(kt) R_k|^2 &\lesssim \sum_{k \in \mathbb{Z}_*^d} \left(\int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) d\tau \right) \\
 &\quad \times \left(\sum_{\ell \in \mathbb{Z}_*^d} \int_0^t \bar{K}_{k,\ell}(t, \tau) |A_\ell(\tau, \ell\tau) \hat{\rho}_\ell(\tau)|^2 d\tau \right) \\
 &\lesssim \left(\sup_{t \geq 0} \sup_{k \in \mathbb{Z}_*^d} \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) d\tau \right) \sum_{\ell \in \mathbb{Z}_*^d} \int_0^t \left(\sum_{k \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) \right) |A_\ell(\tau, \ell\tau) \hat{\rho}_\ell(\tau)|^2 d\tau \\
 &\lesssim \left(\sup_{t \geq 0} \sup_{k \in \mathbb{Z}_*^d} \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) d\tau \right) \left(\sup_{0 \leq \tau \leq t} \sup_{\ell \in \mathbb{Z}_*^d} \sum_{k \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) \right) \|A \hat{\rho}\|_{L_k^2 L_t^2(I)}.
 \end{aligned}$$

The first factor appears in (5.2) and is controlled by Lemma 6.1. The second factor is controlled by Lemma 6.3 and results in the power of $\langle t \rangle$ loss. Therefore by (2.12c),

$$\sum_{k \in \mathbb{Z}_*^d} |A_k(t, kt)R_k(t)|^2 \lesssim K_2 K_3(t) \epsilon^4, \tag{5.20}$$

which suffices to treat the reaction term.

5.2.3 Transport

By (1.3) and $|k(t - \tau)| \leq \langle \tau \rangle |k - \ell, kt - \ell\tau|$, the transport term is bounded by

$$\begin{aligned} & \sum_{k \in \mathbb{Z}_*^d} |A_k(t, kt)T_k(t)|^2 \\ & \lesssim \sum_{k \in \mathbb{Z}_*^d} \left[A_k(t, kt) \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \sum_{N \geq 8} \left| \widehat{\nabla_{z,v} g_{k-\ell}}(\tau, kt - \ell\tau)_N \right| \langle \tau \rangle |\hat{\rho}_\ell(\tau)_{<N/8}| \, d\tau \right]^2. \end{aligned}$$

We begin as in Proposition 5.1. By the frequency localizations (5.10) (which hold on the support of the integrand), (3.11) implies that for some $c = c(s) \in (0, 1)$ we have (using also that the Littlewood-Paley projections define a partition of unity),

$$\begin{aligned} & \sum_{k \in \mathbb{Z}_*^d} |A_k(t, kt)T_k(t)|^2 \\ & \lesssim \sum_{k \in \mathbb{Z}_*^d} \left[\sum_{\ell \in \mathbb{Z}_*^d} \int_0^t |(A\widehat{\nabla}g)_{k-\ell}(\tau, kt - \ell\tau)| e^{c\lambda(\tau)(\ell, \ell\tau)^s} \langle \tau \rangle |\hat{\rho}_\ell(\tau)| \, d\tau \right]^2. \end{aligned}$$

From Cauchy-Schwarz, (5.12) and (2.12c),

$$\begin{aligned} \sum_{k \in \mathbb{Z}_*^d} |A_k(t, kt)T_k(t)|^2 & \lesssim \sqrt{K_3} \epsilon \sum_{k \in \mathbb{Z}_*^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_0^t |(A\widehat{\nabla}g)_{k-\ell}(\tau, kt - \ell\tau)|^2 \\ & \quad \times e^{c\lambda(\tau)(\ell, \ell\tau)^s} \langle \tau \rangle |\hat{\rho}_\ell(\tau)| \, d\tau. \end{aligned}$$

By Fubini’s theorem, (3.8), the $H^{d/2+} \hookrightarrow C^0$ embedding theorem and (2.14) with (3.8),

$$\begin{aligned} & \sum_{k \in \mathbb{Z}_*^d} |A_k(t, kt)T_k(t)|^2 \\ & \lesssim \sqrt{K_3} \epsilon \sum_{\ell \in \mathbb{Z}_*^d} \int_0^t \left(\sum_{k \in \mathbb{Z}_*^d} |(A\widehat{\nabla}g)_{k-\ell}(\tau, kt - \ell\tau)|^2 e^{\frac{1}{2}(c-1)\alpha_0(\tau)^s} \right) e^{\lambda(\tau)(\ell, \ell\tau)^s} |\hat{\rho}_\ell(\tau)| \, d\tau \\ & \lesssim \sqrt{K_3} \epsilon \left(\sup_{\tau \leq t} e^{\frac{1}{2}(c-1)\alpha_0(\tau)^s} \sum_{k \in \mathbb{Z}_*^d} \sup_{\eta \in \mathbb{R}^d} |(A\widehat{\nabla}g)_k(\tau, \eta)|^2 \right) \left(\sum_{\ell \in \mathbb{Z}_*^d} \int_0^t e^{\lambda(\tau)(\ell, \ell\tau)^s} |\hat{\rho}_\ell(\tau)| \, d\tau \right) \end{aligned}$$

$$\begin{aligned} &\lesssim \sqrt{K_3} \epsilon \left(\sup_{\tau \leq t} e^{\frac{1}{2}(c-1)\alpha_0(\tau)^s} \|A^{(1)} \hat{g}\|_{L_k^2 H_t^M}^2 \right) \left(\sum_{\ell \in \mathbb{Z}_*^d} \int_0^t e^{\lambda(\tau)\langle \ell, \ell \tau \rangle^s} |\hat{\rho}_\ell(\tau)| \, d\tau \right) \\ &\lesssim K_1 \sqrt{K_3} \epsilon^3 \left(\sum_{\ell \in \mathbb{Z}_*^d} \int_0^t e^{\lambda(\tau)\langle \ell, \ell \tau \rangle^s} |\hat{\rho}_\ell(\tau)| \, d\tau \right). \end{aligned}$$

Proceeding as in (5.12) and applying (2.12c), we get

$$\sum_{k \in \mathbb{Z}_*^d} |A_k(t, kt) T_k(t)|^2 \lesssim K_3 K_1 \epsilon^4. \tag{5.21}$$

5.2.4 Remainders

The remainder follows from a slight variant of the argument used to treat transport. By (1.3),

$$\|A\mathcal{R}(t)\|_{L_k^2}^2 \lesssim \sum_{k \in \mathbb{Z}_*^d} \left[A_k(t, kt) \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \sum_{N \in \mathbb{D}} \sum_{N/8 \leq N' \leq 8N} |\hat{\rho}_\ell(\tau)_{N'}| |k(t - \tau)| |\hat{g}_{k-\ell}(\tau, kt - \ell\tau)_N| \, d\tau \right]^2.$$

As in Proposition 5.1, (5.14) holds on the support of the integrand and hence, by (5.11) and using the Cauchy-Schwarz inequality,

$$\begin{aligned} \|A\mathcal{R}(t)\|_{L_k^2}^2 &\lesssim \sum_{k \in \mathbb{Z}_*^d} \left[\int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \left(\sum_{N \in \mathbb{D}} e^{2\lambda(\tau)\langle \ell, \ell \tau \rangle^s} \langle \tau \rangle^2 |\hat{\rho}_\ell(\tau)_{\sim N}|^2 \right) \, d\tau \right] \\ &\times \left[\int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \left(\sum_{N \in \mathbb{D}} e^{2(c'-1)\lambda(\tau)\langle \ell, \ell \tau \rangle^s} e^{2c'\lambda(\tau)\langle k-\ell, kt-\ell\tau \rangle^s} |\widehat{\nabla} g_{k-\ell}(\tau, kt - \ell\tau)_N|^2 \right) \, d\tau \right]. \end{aligned}$$

By the almost orthogonality of the Littlewood-Paley decomposition (3.4) and $\sigma > 1$,

$$\begin{aligned} \|A\mathcal{R}(t)\|_{L_k^2}^2 &\lesssim \|A\hat{\rho}\|_{L_k^2 L_t^2(I)}^2 \sum_{k \in \mathbb{Z}_*^d} \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \left(\sum_{N \in \mathbb{D}} e^{2(c'-1)\lambda(\tau)\langle \ell, \ell \tau \rangle^s} e^{2c'\lambda(\tau)\langle k-\ell, kt-\ell\tau \rangle^s} \right. \\ &\quad \left. |\widehat{\nabla} g_{k-\ell}(\tau, kt - \ell\tau)_N|^2 \right) \, d\tau. \end{aligned}$$

By Fubini’s theorem, the $H^{d/2+} \hookrightarrow C^0$ embedding theorem and (3.8),

$$\begin{aligned} \|A\mathcal{R}(t)\|_{L_k^2}^2 &\lesssim \|A\hat{\rho}\|_{L_k^2 L_t^2(I)}^2 \sum_{\ell \in \mathbb{Z}_*^d} \int_0^t e^{2(c'-1)\lambda(\tau)\langle \ell, \ell \tau \rangle^s} \left(\sum_{k \in \mathbb{Z}_*^d} \sum_{N \in \mathbb{D}} e^{2c'\lambda(\tau)\langle k-\ell, kt-\ell \tau \rangle^s} \right. \\ &\quad \left. |\widehat{\nabla} g_{k-\ell}(\tau, kt - \ell \tau)_N|^2 \right) d\tau \\ &\lesssim \|A\hat{\rho}\|_{L_k^2 L_t^2(I)}^2 \sum_{\ell \in \mathbb{Z}_*^d} \int_0^t e^{2(c'-1)\lambda(\tau)\langle \ell, \ell \tau \rangle^s} \\ &\quad \times \left(\sum_{k \in \mathbb{Z}^d} \sum_{N \in \mathbb{D}} \left\| (A^{(-\beta)} \hat{g}_k)(\tau)_N \right\|_{H_\eta^M}^2 \right) d\tau. \end{aligned}$$

The Littlewood-Paley projections do not commute with derivatives in frequency space, however since the projections have bounded derivatives we still have (see §3),

$$\begin{aligned} \|A\mathcal{R}(t)\|_{L_k^2}^2 &\lesssim_M \|A\hat{\rho}\|_{L_k^2 L_t^2(I)}^2 \sum_{\ell \in \mathbb{Z}_*^d} \int_0^t e^{2(c'-1)\lambda(\tau)\langle \ell, \ell \tau \rangle^s} \\ &\quad \times \left(\sum_{N \in \mathbb{D}} \sum_{|\alpha| \leq M} \left\| (D_\eta^\alpha A^{(-\beta)} \hat{g}_k)(\tau)_{\sim N} \right\|_{L_k^2 L_\eta^2}^2 \right) d\tau. \end{aligned}$$

Hence by (3.4), (2.13), (2.12c) and (3.8),

$$\|A\mathcal{R}(t)\|_{L_k^2}^2 \lesssim K_3 K_2 \epsilon^4. \tag{5.22}$$

Summing (5.18), (5.19), (5.20), (5.21) and (5.22) implies the result with $K_4 \approx 1 + K_3 + K_2 K_3 + K_3 K_1$ (in fact we are rather sub-optimal).

5.3 Proof of High Norm Estimate (2.12a)

In this section we derive the high norm estimate on the full distribution, (2.12a). For some multi-index $\alpha \in \mathbb{N}^d$ with $|\alpha| \leq M$, compute the time-derivative

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|A^{(1)} D_\eta^\alpha \hat{g}\|_2^2 &= \sum_{k \in \mathbb{Z}^d} \int_\eta \dot{\lambda}(t)\langle k, \eta \rangle^s \left| A^{(1)} D_\eta^\alpha \hat{g}_k(\eta) \right|^2 d\eta \\ &\quad + \sum_{k \in \mathbb{Z}^d} \int_\eta A^{(1)} D_\eta^\alpha \overline{\hat{g}_k(\eta)} A^{(1)} D_\eta^\alpha \partial_t \hat{g}_k(\eta) d\eta \\ &= CK + E. \end{aligned} \tag{5.23}$$

Like similar terms appearing in [10,45,50], the CK term (for ‘Cauchy-Kovalevskaya’) is used to absorb the highest order terms coming from E .

Turning to E , we separate into the linear and nonlinear contributions

$$\begin{aligned}
 E &= - \sum_{k \in \mathbb{Z}_*^d} \int_{\eta} A^{(1)} D_{\eta}^{\alpha} \overline{\hat{g}_k(\eta)} A_k^{(1)}(t, \eta) D_{\eta}^{\alpha} \left[\hat{\rho}_k(t) \widehat{W}(k) k \cdot (\eta - tk) \widehat{f^0}(\eta - kt) \right] d\eta \\
 &\quad - \sum_{k \in \mathbb{Z}^d} \int_{\eta} A^{(1)} D_{\eta}^{\alpha} \overline{\hat{g}_k(\eta)} A_k^{(1)}(t, \eta) D_{\eta}^{\alpha} \left[\sum_{\ell \in \mathbb{Z}_*^d} \hat{\rho}_{\ell}(t) \widehat{W}(\ell) \ell \cdot (\eta - tk) \hat{g}_{k-\ell}(t, \eta - t\ell) \right] d\eta \\
 &= -E_L - E_{NL}.
 \end{aligned} \tag{5.24}$$

5.3.1 Linear Contribution

The linear contribution E_L is easier from a regularity standpoint than E_{NL} since we may lose regularity when estimating f^0 . However, E_L has one less power of ϵ which requires some care to handle and is the reason we cannot just take K_1 in (2.12a) to be $O(1)$. The treatment of E_L begins with the product lemma (3.13):

$$\begin{aligned}
 |E_L| &\lesssim \left\| A^{(1)} D_{\eta}^{\alpha} \hat{g} \right\|_2 \left\| A^{(1)} D_{\eta}^{\alpha} (\eta \hat{f}^0(\eta)) \right\|_{L_{\eta}^2} \|\nabla_x W *_{x} \rho\|_{\mathcal{F}^{\tilde{c}\lambda(t);s}} \\
 &\quad + \left\| A^{(1)} D_{\eta}^{\alpha} \hat{g} \right\|_2 \left\| v^{\alpha} (\nabla_v f^0) \right\|_{\mathcal{G}^{\tilde{c}\lambda(t);s}} \left\| A^{(1)} \nabla_x W *_{x} \rho(t) \right\|_2.
 \end{aligned}$$

By (1.9), (3.8) and (3.7),

$$\begin{aligned}
 \left\| A^{(1)} D_{\eta}^{\alpha} (\eta \hat{f}^0(\eta)) \right\|_2 &\leq \left\| A^{(1)} \left(\eta D_{\eta}^{\alpha} \hat{f}^0(\eta) \right) \right\|_2 + \sum_{|j|=1; j \leq \alpha} \left\| A^{(1)} \left(\eta D_{\eta}^{\alpha-j} \hat{f}^0(\eta) \right) \right\|_2 \\
 &\lesssim C_0.
 \end{aligned} \tag{5.25}$$

Next we use $\gamma \geq 1$ to deduce from (1.3),

$$A_k^{(1)}(t, kt) |\widehat{W}(k)k| = \langle k, kt \rangle |\widehat{W}(k)k| A_k(t, kt) \lesssim \langle t \rangle A_k(t, kt),$$

which implies (also using $\tilde{c} < 1$ and (3.8)),

$$\begin{aligned}
 |E_L| &\lesssim e^{(\tilde{c}-1)\alpha_0(t)^s} \left\| A^{(1)} D_{\eta}^{\alpha} \hat{g} \right\|_2 \|A\rho(t)\|_2 + \langle t \rangle \left\| A^{(1)} D_{\eta}^{\alpha} \hat{g} \right\|_2 \|A\rho(t)\|_2 \\
 &\lesssim \langle t \rangle \left\| A^{(1)} D_{\eta}^{\alpha} \hat{g} \right\|_2 \|A\rho(t)\|_2.
 \end{aligned} \tag{5.26}$$

5.3.2 Commutator Trick for the Nonlinear Term

Turn now to the nonlinear term in (5.24), E_{NL} . Here we cannot lose much regularity on any of the factors involved, however we have additional powers of ϵ which will eliminate the large constants. First, we expand the D_η^α derivative

$$\begin{aligned}
 E_{NL} &= \sum_{k \in \mathbb{Z}^d} \int_\eta A^{(1)} D_\eta^\alpha \overline{\hat{g}_k(\eta)} \left(A_k^{(1)}(t, \eta) \left[\sum_{\ell \in \mathbb{Z}_*^d} \hat{\rho}_\ell(t) \widehat{W}(\ell) \ell \cdot (\eta - t\ell) D_\eta^\alpha \hat{g}_{k-\ell}(t, \eta - t\ell) \right] \right) d\eta \\
 &+ \sum_{k \in \mathbb{Z}^d} \int_\eta A^{(1)} D_\eta^\alpha \overline{\hat{g}_k(\eta)} \left(A_k^{(1)}(t, \eta) \sum_{|j|=1; j \leq \alpha} \left[\sum_{\ell \in \mathbb{Z}_*^d} \hat{\rho}_\ell(t) \widehat{W}(\ell) \ell_j D_\eta^{\alpha-j} \hat{g}_{k-\ell}(t, \eta - t\ell) \right] \right) d\eta \\
 &= E_{NL}^1 + E_{NL}^2.
 \end{aligned}$$

Consider first E_{NL}^1 , as this contains an extra derivative which results in a loss of regularity that must be balanced by the CK term in (5.23). To gain from the cancellations inherent in transport we follow the commutator trick used in (for example) [10, 25, 45, 50] by applying the identity,

$$\frac{1}{2} \int_{\mathbb{T}^d \times \mathbb{R}^d} F(t, z + tv) \cdot (\nabla_v - t \nabla_z) \left[A^{(1)}(v^\alpha g) \right]^2 dz dv = 0, \tag{5.27}$$

to write,

$$\begin{aligned}
 E_{NL}^1 &= \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_\eta A^{(1)} D_\eta^\alpha \overline{\hat{g}_k(\eta)} \hat{\rho}_\ell(t) \widehat{W}(\ell) \ell \cdot (\eta - kt) \\
 &\quad \times \left[A_k^{(1)}(t, \eta) - A_{k-\ell}^{(1)}(t, \eta - t\ell) \right] \left(D_\eta^\alpha \hat{g}_{k-\ell}(t, \eta - t\ell) \right) d\eta.
 \end{aligned}$$

We divide further via paraproduct:

$$E_{NL}^1 = \sum_{N \geq 8} T_N^1 + \sum_{N \geq 8} R_N^1 + \mathcal{R}^1, \tag{5.28}$$

where the *transport* term is given by

$$\begin{aligned}
 T_N^1 &= \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_\eta A^{(1)} D_\eta^\alpha \overline{\hat{g}_k(\eta)} \hat{\rho}_\ell(t) \widehat{W}(\ell) \ell \cdot (\eta - kt) \\
 &\quad \times \left[A_k^{(1)}(t, \eta) - A_{k-\ell}^{(1)}(t, \eta - t\ell) \right] \left(D_\eta^\alpha \hat{g}_{k-\ell}(t, \eta - t\ell) \right)_N d\eta, \tag{5.29}
 \end{aligned}$$

and the reaction term by

$$\begin{aligned}
 R_N^1 &= \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} A^{(1)} D_{\eta}^{\alpha} \overline{\hat{g}_k(\eta)} \hat{\rho}_{\ell}(t)_N \widehat{W}(\ell) \ell \cdot (\eta - kt) \\
 &\quad \times \left[A_k^{(1)}(t, \eta) - A_{k-\ell}^{(1)}(t, \eta - t\ell) \right] \left(D_{\eta}^{\alpha} \hat{g}_{k-\ell}(t, \eta - t\ell) \right)_{<N/8} d\eta. \quad (5.30)
 \end{aligned}$$

The remainder, \mathcal{R}^1 , is whatever is left over.

5.3.3 Transport

On the support of the integrand in (5.29) we have

$$\frac{N}{2} \leq |k - \ell, \eta - t\ell| \leq \frac{3N}{2}, \quad (5.31a)$$

$$|\ell, \ell t| \leq \frac{3N}{32}, \quad (5.31b)$$

$$\frac{13}{16} \leq \frac{|k, \eta|}{|k - \ell, \eta - t\ell|} \leq \frac{19}{16}. \quad (5.31c)$$

By (5.31a) we can gain from the multiplier:

$$\begin{aligned}
 \left| \frac{A_k^{(1)}(\eta)}{A_{k-\ell}^{(1)}(\eta - t\ell)} - 1 \right| &= \left| \frac{e^{\lambda \langle k, \eta \rangle^s} \langle k, \eta \rangle^{\sigma+1}}{e^{\lambda \langle k-\ell, \eta-t\ell \rangle^s} \langle k-\ell, \eta-t\ell \rangle^{\sigma+1}} - 1 \right| \\
 &\leq \left| e^{\lambda \langle k, \eta \rangle^s - \lambda \langle k-\ell, \eta-t\ell \rangle^s} - 1 \right| \\
 &\quad + e^{\lambda \langle k, \eta \rangle^s - \lambda \langle k-\ell, \eta-t\ell \rangle^s} \left| \frac{\langle k, \eta \rangle^{\sigma+1}}{\langle k-\ell, \eta-t\ell \rangle^{\sigma+1}} - 1 \right|. \quad (5.32)
 \end{aligned}$$

By $|e^x - 1| \leq xe^x$, (3.10) and (3.11) (using (5.31a)), there is some $c = c(s) \in (0, 1)$ such that:

$$\begin{aligned}
 \left| e^{\lambda \langle k, \eta \rangle^s - \lambda \langle k-\ell, \eta-t\ell \rangle^s} - 1 \right| &\leq \lambda \left| \langle k, \eta \rangle^s - \langle k-\ell, \eta-t\ell \rangle^s \right| e^{\lambda \langle k, \eta \rangle^s - \lambda \langle k-\ell, \eta-t\ell \rangle^s} \\
 &\lesssim \frac{\langle \ell, \ell t \rangle}{\langle k, \eta \rangle^{1-s} + \langle k-\ell, \eta-t\ell \rangle^{1-s}} e^{c\lambda \langle \ell, \ell t \rangle^s}. \quad (5.33)
 \end{aligned}$$

The other term in (5.32) can be treated with the mean-value theorem and (3.11), resulting in a bound not worse than (5.33). Therefore, applying (1.3), (5.32), (5.33) and adding a frequency localization by (5.31a) to T_N^1 implies

$$\begin{aligned}
 |T_N^1| &\lesssim \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} \left| A^{(1)} D_{\eta}^{\alpha} \hat{g}_k(\eta) \right| |\hat{\rho}_{\ell}(t)_{<N/8}| \frac{|\eta - \ell t - t(k - \ell)| \langle \ell, \ell t \rangle}{\langle k, \eta \rangle^{1-s} + \langle k - \ell, \eta - \ell t \rangle^{1-s}} e^{c\lambda(\ell, \ell t)s} \\
 &\quad \times A_{k-\ell}^{(1)}(t, \eta - t\ell) \left| \left(D_{\eta}^{\alpha} \hat{g}_{k-\ell}(t, \eta - t\ell) \right)_N \right| d\eta \\
 &\lesssim \langle t \rangle^2 \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} \left| \left(A^{(1)} D_{\eta}^{\alpha} \hat{g}_k(\eta) \right)_{\sim N} \right| |\hat{\rho}_{\ell}(t)| \langle \ell \rangle |k - \ell, \eta - t\ell|^{s/2} |k, \eta|^{s/2} e^{c\lambda(\ell, \ell t)s} \\
 &\quad \times A_{k-\ell}^{(1)}(t, \eta - t\ell) \left| \left(D_{\eta}^{\alpha} \hat{g}_{k-\ell}(t, \eta - t\ell) \right)_N \right| d\eta.
 \end{aligned}$$

Applying (3.1) implies

$$|T_N^1| \lesssim \langle t \rangle^2 \left\| |\nabla_{z,v}|^{s/2} A^{(1)}(v^{\alpha} g)_{\sim N} \right\|_2 \left\| |\nabla_{z,v}|^{s/2} A^{(1)}(v^{\alpha} g)_N \right\|_2 \|\rho(t)\|_{\mathcal{F}^{c\lambda(t), \frac{d}{2}+2;s}}.$$

Using the regularity gap provided by $c < 1$ and (3.8) (also $\sigma > \frac{d}{2} + 2$),

$$\begin{aligned}
 |T_N^1| &\lesssim \langle t \rangle^2 e^{(c-1)\lambda(t)(t)^s} \|A\rho(t)\|_2 \left\| \langle \nabla_{z,v} \rangle^{s/2} A^{(1)}(v^{\alpha} g)_{\sim N} \right\|_2 \left\| \langle \nabla_{z,v} \rangle^{s/2} A^{(1)}(v^{\alpha} g)_N \right\|_2 \\
 &\lesssim_{\alpha_0} e^{\frac{1}{2}(c-1)\alpha_0(t)^s} \|A\rho(t)\|_2 \left\| \langle \nabla_{z,v} \rangle^{s/2} A^{(1)}(v^{\alpha} g)_{\sim N} \right\|_2^2. \tag{5.34}
 \end{aligned}$$

We will find that this term is eventually absorbed by the CK term in (5.23).

5.3.4 Reaction

Next we consider the reaction contribution, where the commutator introduced by the identity (5.27) to deal with transport will not be useful. Hence, write $R_N^1 = R_N^{1;1} + R_N^{1;2}$ where

$$\begin{aligned}
 R_N^{1;1} &= \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} A^{(1)} D_{\eta}^{\alpha} \overline{\hat{g}_k(\eta)} A_k^{(1)}(t, \eta) \hat{\rho}_{\ell}(t)_N \widehat{W}(\ell) \ell \cdot (\eta - kt) \\
 &\quad \times \left(D_{\eta}^{\alpha} \hat{g}_{k-\ell}(t, \eta - t\ell) \right)_{<N/8} d\eta.
 \end{aligned}$$

We focus on $R_N^{1;1}$ first; $R_N^{1;2}$ is easier as the norm is landing on the ‘low frequency’ factor. On the support of the integrand, we have the frequency localizations (3.15), from which it follows by (3.11) that there exists some $c = c(s) \in (0, 1)$ such that

$$\begin{aligned}
 |R_N^{1;1}| &\lesssim \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} \left| A^{(1)} D_{\eta}^{\alpha} \hat{g}_k(\eta) \right| A_{\ell}^{(1)}(t, \ell t) |\widehat{W}(\ell) \ell \hat{\rho}_{\ell}(t)_N| \\
 &\quad \times e^{c\lambda(k-\ell, \eta-t\ell)s} \left| [\eta - tk] \left(D_{\eta}^{\alpha} \hat{g}_{k-\ell}(t, \eta - t\ell) \right)_{<N/8} \right| d\eta. \tag{5.35}
 \end{aligned}$$

Now again we have the crucial use of the assumption $\gamma \geq 1$ as in E_L^1 :

$$A_\ell^{(1)}(t, \ell t) |\widehat{W}(\ell)\ell| \lesssim \frac{\langle \ell, \ell t \rangle}{|\ell|} A_\ell(t, \ell t) \lesssim \langle t \rangle A_\ell(t, \ell t).$$

Therefore (adding a frequency localization by (3.15)), by $|\eta - kt| \leq \langle t \rangle |k - \ell, \eta - t\ell|$ and (3.8)

$$\begin{aligned} |R_N^{1;1}| &\lesssim \langle t \rangle \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} \left| A^{(1)} \left(D_\eta^\alpha \hat{g}_k(\eta) \right)_{\sim N} \right| |A_\ell(t, t\ell) \hat{\rho}_\ell(t)_N| e^{c\lambda \langle k-\ell, \eta-t\ell \rangle^s} \\ &\quad \left| (\eta - tk) \left(D_\eta^\alpha \hat{g}_{k-\ell}(t, \eta - t\ell) \right) \right| d\eta \\ &\lesssim \langle t \rangle^2 \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} \left| A^{(1)} \left(D_\eta^\alpha \hat{g}_k(\eta) \right)_{\sim N} \right| |A_\ell(t, t\ell) \hat{\rho}_\ell(t)_N| e^{\lambda \langle k-\ell, \eta-t\ell \rangle^s} \\ &\quad \left| \left(D_\eta^\alpha \hat{g}_{k-\ell}(t, \eta - t\ell) \right) \right| d\eta. \end{aligned}$$

Applying (3.2) and $\sigma - \beta > \frac{d}{2} + 1$,

$$\begin{aligned} |R_N^{1;1}| &\lesssim \langle t \rangle^2 \left\| A^{(1)}(v^\alpha g)_{\sim N} \right\|_2 \|A\rho_N\|_2 \left\| A^{(-\beta)} v^\alpha g \right\|_2 \\ &\lesssim \frac{\left\| A^{(-\beta)} v^\alpha g \right\|_2}{\langle t \rangle^2} \left\| A^{(1)}(v^\alpha g)_{\sim N} \right\|_2^2 + \left\| A^{(-\beta)} v^\alpha g \right\|_2 \langle t \rangle^6 \|A\rho_N\|_2^2, \end{aligned} \tag{5.36}$$

which will suffice to treat this term.

Next turn to $R_N^{1;2}$, which is easier. By (1.3),

$$\begin{aligned} |R_N^{1;2}| &\lesssim \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} \left| A^{(1)} D_\eta^\alpha \hat{g}_k(\eta) \right| |\hat{\rho}_\ell(t)_N| |\eta - kt| \\ &\quad \left| A_{k-\ell}^{(1)}(t, \eta - t\ell) \left(D_\eta^\alpha \hat{g}_{k-\ell}(t, \eta - t\ell) \right)_{<N/8} \right| d\eta. \end{aligned}$$

Since the frequency localizations (3.15) hold also on the support of the integrand of $R_N^{1;2}$ (in particular $|\eta - kt| \leq \langle t \rangle |k - \ell, \eta - t\ell| \lesssim \langle t \rangle |\ell, \ell t|$),

$$\begin{aligned} |R_N^{1;2}| &\lesssim \langle t \rangle \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} \left| A^{(1)} D_\eta^\alpha \hat{g}_k(\eta)_{\sim N} \right| \langle \ell, \ell t \rangle |\hat{\rho}_\ell(t)_N| A_{k-\ell}^{(1)}(t, \eta - t\ell) \\ &\quad \left| \left(D_\eta^\alpha \hat{g}_{k-\ell}(t, \eta - t\ell) \right)_{<N/8} \right| d\eta. \end{aligned}$$

Therefore, by (3.1), (3.5), (3.8), (3.15) and $\sigma > \frac{d}{2} + 3$,

$$\begin{aligned} \left| \mathcal{R}_N^{1;2} \right| &\lesssim \langle t \rangle \left\| A^{(1)}(v^\alpha g)_{\sim N} \right\|_2 \|\rho(t)_N\|_{\mathcal{F}^{0, \frac{d}{2}+2}} \left\| A^{(1)}(v^\alpha g) \right\|_2 \\ &\lesssim \frac{\langle t \rangle}{N} \left\| A^{(1)}(v^\alpha g)_{\sim N} \right\|_2 \|\rho(t)_N\|_{\mathcal{F}^{0, \frac{d}{2}+3}} \left\| A^{(1)}(v^\alpha g) \right\|_2 \\ &\lesssim \frac{e^{-\alpha_0 \langle t \rangle^s}}{N} \left\| A^{(1)}(v^\alpha g) \right\|_2^2 \|A\rho(t)\|_2, \end{aligned} \tag{5.37}$$

which suffices to treat this term.

5.3.5 Remainders

In order to complete the treatment of E_{NL}^1 it remains to estimate the remainder \mathcal{R} . Like \mathcal{R}_N^1 , the commutator introduced by (5.27) is not helpful, so divide into two pieces:

$$\begin{aligned} \mathcal{R}^1 &= \sum_{N \in \mathbb{D}} \sum_{N/8 \leq N' \leq 8N} \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} A^{(1)} D_{\eta}^{\alpha} \overline{\hat{g}_k(\eta)} \hat{\rho}_{\ell}(t)_{N'} \widehat{W}(\ell) \ell \cdot (\eta - kt) \\ &\quad \times \left[A_k^{(1)}(t, \eta) - A_{k-\ell}^{(1)}(t, \eta - t\ell) \right] \left(D_{\eta}^{\alpha} \hat{g}_{k-\ell}(t, \eta - t\ell) \right)_N d\eta \\ &= \mathcal{R}^{1;1} + \mathcal{R}^{1;2}. \end{aligned}$$

Analogous to (5.14), we claim that on the integrand there holds for some $c' = c'(s) \in (0, 1)$,

$$A_k^{(1)}(t, \eta) \lesssim_{\lambda_0, \alpha_0} e^{c' \lambda(t) \langle k-\ell, \eta-t\ell \rangle^s} e^{c' \lambda(t) \langle \ell, t\ell \rangle^s}, \tag{5.38}$$

which again follows by the argument used to deduce (3.16). Therefore, (5.38) implies (using also (1.3) and $|\eta - kt| \leq \langle t \rangle \langle k - \ell, \eta - t\ell \rangle$),

$$\begin{aligned} \left| \mathcal{R}^{1;1} \right| &\lesssim \langle t \rangle \sum_{N \in \mathbb{D}} \sum_{N' \approx N} \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} \left| A^{(1)} D_{\eta}^{\alpha} \hat{g}_k(\eta) \right| e^{c' \lambda(t) \langle \ell, t\ell \rangle^s} \left| \hat{\rho}_{\ell}(t)_{N'} \right| \\ &\quad \times \langle k - \ell, \eta - t\ell \rangle e^{c' \lambda(t) \langle k-\ell, \eta-t\ell \rangle^s} \left| \left(D_{\eta}^{\alpha} \hat{g}_{k-\ell}(t, \eta - t\ell) \right)_N \right| d\eta. \end{aligned}$$

Applying (3.2), $\sigma > \frac{d}{2} + 2$, (3.5) and (3.3), we have

$$\begin{aligned} \left| \mathcal{R}^{1;1} \right| &\lesssim e^{\frac{1}{2}(c'-1)\alpha_0 \langle t \rangle^s} \sum_{N \in \mathbb{D}} \sum_{N' \approx N} \left\| A^{(1)}(v^\alpha g) \right\|_2 \frac{1}{N'} \|A\rho(t)_{N'}\|_2 \left\| (v^\alpha g)_N \right\|_{\mathcal{G}^{c' \lambda(t); \frac{d}{2}+2; s}} \\ &\lesssim \|A\rho(t)\|_2 e^{\frac{1}{2}(c'-1)\alpha_0 \langle t \rangle^s} \sum_{N \in \mathbb{D}} \left\| A^{(1)}(v^\alpha g) \right\|_2 \frac{1}{N} \left\| A^{(1)}(v^\alpha g)_N \right\|_2 \\ &\lesssim \|A\rho(t)\|_2 \langle t \rangle^{-1} e^{\frac{1}{4}(c'-1)\alpha_0 \langle t \rangle^s} \left\| A^{(1)}(v^\alpha g) \right\|_2^2, \end{aligned} \tag{5.39}$$

which will suffice to treat this term.

Treating $\mathcal{R}^{1;2}$ is very similar to $\mathcal{R}^{1;1}$. Indeed, on the support of the integrand $\langle k - \ell, \eta - t\ell \rangle \lesssim \langle \ell, t\ell \rangle$ by the same logic used to deduce (5.38) and hence (1.3) and (3.8) imply

$$\begin{aligned} |\mathcal{R}^{1;2}| &\lesssim \langle t \rangle \sum_{N \in \mathbb{D}} \sum_{N' \approx N} \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} \left| A^{(1)} D_{\eta}^{\alpha} \hat{g}_k(\eta) \right| |\hat{\rho}_{\ell}(t)_{N'}| \\ &\quad \times \langle k - \ell, \eta - t\ell \rangle A_{k-\ell}^{(1)}(t, \eta - t\ell) \left| \left(D_{\eta}^{\alpha} \hat{g}_{k-\ell}(t, \eta - t\ell) \right)_N \right| d\eta \\ &\lesssim \langle t \rangle \sum_{N \in \mathbb{D}} \sum_{N' \approx N} \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} \left| A^{(1)} D_{\eta}^{\alpha} \hat{g}_k(\eta) \right| e^{\frac{1}{2}\lambda(t)(\ell, \ell t)^s} |\hat{\rho}_{\ell}(t)_{N'}| \\ &\quad \times A_{k-\ell}^{(1)}(t, \eta - t\ell) \left| \left(D_{\eta}^{\alpha} \hat{g}_{k-\ell}(t, \eta - t\ell) \right)_N \right| d\eta, \end{aligned}$$

which implies by (3.1), (3.8), (3.3) and $\sigma > d/2 + 2$,

$$\begin{aligned} |\mathcal{R}^{1;2}| &\lesssim \langle t \rangle e^{-\frac{1}{2}\alpha_0(t)^s} \sum_{N \in \mathbb{D}} \sum_{N' \approx N} \left\| A^{(1)}(v^{\alpha} g) \right\|_2 \frac{1}{N'} \|A\hat{\rho}(t)_{N'}\|_2 \left\| A^{(1)}(v^{\alpha} g)_N \right\|_2 \\ &\lesssim \langle t \rangle e^{-\frac{1}{2}\alpha_0(t)^s} \|A\rho(t)\|_2 \sum_{N \in \mathbb{D}} \left\| A^{(1)}(v^{\alpha} g) \right\|_2 \frac{1}{N} \left\| A^{(1)}(v^{\alpha} g)_N \right\|_2 \\ &\lesssim \langle t \rangle^{-1} \|A\rho(t)\|_2 e^{-\frac{1}{4}\alpha_0(t)^s} \left\| A^{(1)}(v^{\alpha} g) \right\|_2^2, \end{aligned} \tag{5.40}$$

which suffices to treat $\mathcal{R}^{1;2}$.

5.3.6 Treatment of Lower Moments

Next we turn to the treatment of E_{NL}^2 . First apply (1.3) (using $\gamma \geq 1$),

$$\begin{aligned} |E_{NL}^2| &\lesssim \sum_{|j|=1; j \leq \alpha} \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} \left| A^{(1)} D_{\eta}^{\alpha} \hat{g}_k(\eta) \right| \\ &\quad A_k^{(1)}(t, \eta) \left| \langle \ell \rangle^{-1} \hat{\rho}_{\ell}(t) D_{\eta}^{\alpha-j} \hat{g}_{k-\ell}(t, \eta - t\ell) \right| d\eta. \end{aligned}$$

Then we apply (3.13)

$$\begin{aligned} |E_{NL}^2| &\lesssim \sum_{|j|=1; j \leq \alpha} \left\| A^{(1)} v^{\alpha} g \right\|_2 \left\| A^{(1)} v^{\alpha-j} g \right\|_2 \|\rho(t)\|_{\mathcal{F}^{\bar{c}\lambda(t), 0; s}} \\ &\quad + \sum_{|j|=1; j \leq \alpha} \left\| A^{(1)} v^{\alpha} g \right\|_2 \left\| v^{\alpha-j} g \right\|_{\mathcal{G}^{\bar{c}\lambda, 0; s}} \left\| \langle k \rangle^{-1} A_k^{(1)} \hat{\rho}_k(t) \right\|_{L_k^2}. \end{aligned}$$

From here, we take advantage of the regularity gap and $\langle k \rangle^{-1} A_k^{(1)}(k, kt) \lesssim \langle t \rangle A_k(t, kt)$ to deduce

$$\begin{aligned} |E_{NL}^2| &\lesssim \sum_{|j|=1; j \leq \alpha} e^{(\tilde{c}-1)\alpha_0 \langle t \rangle^s} \|A^{(1)} v^\alpha g\|_2 \|A^{(1)} v^{\alpha-j} g\|_2 \|A\rho(t)\|_2 \\ &+ \sum_{|j|=1; j \leq \alpha} \langle t \rangle \|A^{(1)} v^\alpha g\|_2 \|A^{(-\beta)} v^{\alpha-j} g\|_2 \|A\rho(t)\|_2, \end{aligned} \tag{5.41}$$

which suffices to treat this contribution.

5.3.7 Conclusion of High Norm Estimate

Denote $\delta = -\frac{1}{4} \min(c - 1, \tilde{c} - 1, c' - 1) \alpha_0$. Collecting the contributions of (5.23), (5.26) (5.34), (5.36), (5.37), (5.39), (5.40) and (5.41) then summing in N with (3.4) (note we used $1 \lesssim \langle k, \eta \rangle^{s/2}$ to group (5.37), (5.39) and (5.40) with (5.34)), we have the following for some $\tilde{K} = \tilde{K}(s, M, \sigma, \lambda_0, \lambda', C_0, d)$,

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|A^{(1)} v^\alpha g\|_2^2 &\leq \left(\tilde{K} \langle t \rangle^{-1} e^{-\delta \langle t \rangle^s} \|A\rho(t)\|_2 + \dot{\lambda}(t) \right) \|\langle \nabla_{z,v} \rangle^{s/2} A^{(1)}(v^\alpha g)\|_2^2 \\ &+ \tilde{K} \langle t \rangle \|A^{(1)} v^\alpha g\|_2 \|A\rho(t)\|_2 \\ &+ \tilde{K} \frac{\|A^{(-\beta)} v^\alpha g\|_2}{\langle t \rangle^2} \|A^{(1)} v^\alpha g\|_2^2 + \tilde{K} \langle t \rangle^6 \|A^{(-\beta)} v^\alpha g\|_2 \|A\rho\|_2^2 \\ &+ \tilde{K} \sum_{|j|=1; j \leq \alpha} e^{-\delta \langle t \rangle^s} \|A^{(1)} v^\alpha g\|_2 \|A^{(1)} v^{\alpha-j} g\|_2 \|A\rho(t)\|_2 \\ &+ \tilde{K} \sum_{|j|=1; j \leq \alpha} \langle t \rangle \|A^{(1)} v^\alpha g\|_2 \|A^{(-\beta)} v^{\alpha-j} g\|_2 \|A\rho(t)\|_2. \end{aligned}$$

Introducing a small parameter b to be fixed depending only on \tilde{K} and $\lambda(t)$,

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|A^{(1)} v^\alpha g\|_2^2 &\leq \left(\tilde{K} e^{-\delta \langle t \rangle^s} \langle t \rangle^{-1} \|A\rho(t)\|_2 + \frac{b\tilde{K}}{\langle t \rangle^2} + \dot{\lambda}(t) \right) \|\langle \nabla \rangle^{s/2} A^{(1)}(v^\alpha g)\|_2^2 \\ &+ \frac{\tilde{K}}{b} \langle t \rangle^4 \|A\rho(t)\|_2^2 \\ &+ \tilde{K} \frac{\|A^{(-\beta)} v^\alpha g\|_2}{\langle t \rangle^2} \|A^{(1)} v^\alpha g\|_2^2 + \tilde{K} \langle t \rangle^6 \|A^{(-\beta)} v^\alpha g\|_2 \|A\rho\|_2^2 \\ &+ \tilde{K} \sum_{|j|=1; j \leq \alpha} e^{-\delta \langle t \rangle^s} \|A^{(1)} v^\alpha g\|_2 \|A^{(1)} v^{\alpha-j} g\|_2 \|A\rho(t)\|_2 \\ &+ \tilde{K} \sum_{|j|=1; j \leq \alpha} \langle t \rangle \|A^{(1)} v^\alpha g\|_2 \|A^{(-\beta)} v^{\alpha-j} g\|_2 \|A\rho(t)\|_2. \end{aligned} \tag{5.42}$$

By (3.8) and (2.9) we may fix b and ϵ small such that

$$\tilde{K} e^{-\delta \langle t \rangle^s} \sqrt{K_4} \epsilon + \frac{b \tilde{K}}{\langle t \rangle^2} \leq \frac{1}{2} |\dot{\lambda}(t)|.$$

Note this requires fixing ϵ small relative to K_4 but not b . Then by (5.16) we deduce that the first term in (5.42) is negative. Therefore, summing in α , integrating and applying the bootstrap hypotheses (2.11) and (5.16) implies (adjusting \tilde{K} to \tilde{K}')

$$\begin{aligned} \sum_{|\alpha| \leq M} \left\| A^{(1)} v^\alpha g \right\|_2^2 &\leq \epsilon^2 + \tilde{K}' \langle t \rangle^5 K_3 \epsilon^2 + \tilde{K}' K_1 \sqrt{K_2} \langle t \rangle^6 \epsilon^3 + \tilde{K}' \sqrt{K_2} \langle t \rangle^7 K_3 \epsilon^3 \\ &\quad + \tilde{K}' K_1 \sqrt{K_4} \epsilon^3 + \tilde{K}' \sqrt{K_1 K_2 K_4} \langle t \rangle^6 \epsilon^3. \end{aligned}$$

Hence we may take $K_1 = \tilde{K}' K_3 + 1$ and we have (2.12a) by choosing

$$\epsilon < K_1 \left(4 \tilde{K}' \right)^{-1} \left(K_1 \sqrt{K_2} + \sqrt{K_2} K_3 + K_1 \sqrt{K_4} + \sqrt{K_1 K_2 K_4} \right)^{-1}.$$

5.4 Proof of Low Norm Estimate (Equation (2.12b))

This proof proceeds analogously to (2.12a) replacing $A^{(1)}$ with $A^{(-\beta)}$. First compute the derivative as in (5.23),

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \left\| A^{(-\beta)} D_\eta^\alpha \hat{g} \right\|_2^2 &= \sum_{k \in \mathbb{Z}^d} \int_\eta \dot{\lambda}(t) \langle k, \eta \rangle^s \left| A^{(-\beta)} D_\eta^\alpha \hat{g}_k(\eta) \right|^2 d\eta \\ &\quad + \sum_{k \in \mathbb{Z}^d} \int_\eta A^{(-\beta)} D_\eta^\alpha \overline{\hat{g}_k(\eta)} A^{(-\beta)} D_\eta^\alpha \partial_t \hat{g}_k(\eta) d\eta \\ &= C K_{\mathcal{L}} + E_{\mathcal{L}}, \end{aligned} \tag{5.43}$$

where

$$\begin{aligned} E_{\mathcal{L}} &= - \sum_{k \in \mathbb{Z}^d} \int_\eta A^{(-\beta)} D_\eta^\alpha \overline{\hat{g}_k(\eta)} A_k^{(-\beta)}(t, \eta) \hat{\rho}_k(t) \widehat{W}(k) \\ &\quad \times D_\eta^\alpha \left[k \cdot (\eta - tk) \hat{f}^0(\eta - kt) \right] d\eta \\ &\quad - \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_\eta A^{(-\beta)} D_\eta^\alpha \overline{\hat{g}_k(\eta)} A_k^{(-\beta)}(t, \eta) \hat{\rho}_\ell(t) \widehat{W}(\ell) \\ &\quad \times D_\eta^\alpha \left[\ell \cdot (\eta - tk) \hat{g}_{k-\ell}(t, \eta - t\ell) \right] d\eta \\ &= -E_{\mathcal{L};L} - E_{\mathcal{L};NL}. \end{aligned} \tag{5.44}$$

As in the treatment of E_L in §5.3.1, we may use the product lemma (3.13) and (1.3) to deduce

$$|E_{\mathcal{L};L}| \lesssim \left\| A^{(-\beta)} D_\eta^\alpha \hat{g} \right\|_2 \left\| A^{(-\beta)} D_\eta^\alpha (\eta \hat{f}^0(\eta)) \right\|_{L_\eta^2} \|\rho(t)\|_{\mathcal{F}^{\tilde{c}\lambda(t),0;s}} + \left\| A^{(-\beta)} D_\eta^\alpha \hat{g} \right\|_2 \left\| A^{(-\beta)} D_\eta^\alpha (\eta \hat{f}^0(\eta)) \right\|_{L_\eta^2} \left\| A^{(-\beta)} \rho(t) \right\|_2.$$

By the analogue of (5.25), $\tilde{c} < 1$ and the regularity gap between $A^{(-\beta)}$ and A and (3.8)

$$|E_{\mathcal{L};L}| \lesssim e^{(\tilde{c}-1)\alpha_0(t)^s} \left\| A^{(-\beta)} D_\eta^\alpha \hat{g} \right\|_2 \|A\rho(t)\|_2 + \langle t \rangle^{-\beta} \left\| A^{(-\beta)} D_\eta^\alpha \hat{g} \right\|_2 \|A\rho(t)\|_2 \lesssim \langle t \rangle^{-\beta} \left\| A^{(-\beta)} v^\alpha g \right\|_2 \|A\rho(t)\|_2, \tag{5.45}$$

which suffices to treat this term.

We now turn to the treatment of $E_{\mathcal{L};NL}$, which as in §5.3.2 is expanded by

$$\begin{aligned} E_{\mathcal{L};NL} &= \sum_{k \in \mathbb{Z}^d} \int_\eta A^{(-\beta)} D_\eta^\alpha \overline{\hat{g}_k(\eta)} \\ &\times \left(A_k^{(-\beta)}(t, \eta) \left[\sum_{\ell \in \mathbb{Z}_*^d} \hat{\rho}_\ell(t) \widehat{W}(\ell) \ell \cdot (\eta - tk) D_\eta^\alpha \hat{g}_{k-\ell}(t, \eta - t\ell) \right] \right) d\eta \\ &+ \sum_{k \in \mathbb{Z}^d} \int_\eta A^{(-\beta)} D_\eta^\alpha \overline{\hat{g}_k(\eta)} \\ &\times \left(A_k^{(-\beta)}(t, \eta) \sum_{|j|=1; j \leq \alpha} \left[\sum_{\ell \in \mathbb{Z}_*^d} \hat{\rho}_\ell(t) \widehat{W}(\ell) \ell_j D_\eta^{\alpha-j} \hat{g}_{k-\ell}(t, \eta - t\ell) \right] \right) d\eta \\ &= E_{\mathcal{L};NL}^1 + E_{\mathcal{L};NL}^2. \end{aligned}$$

First consider $E_{\mathcal{L};NL}^1$, to which we apply (5.27) (with $A^{(-\beta)}$ instead of $A^{(1)}$) and then decompose via paraproduct as in (5.28):

$$E_{\mathcal{L};NL}^1 = \sum_{N \geq 8} T_{\mathcal{L};N}^1 + \sum_{N \geq 8} R_{\mathcal{L};N}^1 + \mathcal{R}_{\mathcal{L}}^1, \tag{5.46}$$

where the *transport* term is given by

$$T_{\mathcal{L};N}^1 = \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_\eta A^{(-\beta)} D_\eta^\alpha \overline{\hat{g}_k(\eta)} \hat{\rho}_\ell(t) \widehat{W}(\ell) \ell \cdot (\eta - kt) \times \left[A_k^{(-\beta)}(t, \eta) - A_{k-\ell}^{(-\beta)}(t, \eta - t\ell) \right] \left(D_\eta^\alpha \hat{g}_{k-\ell}(t, \eta - t\ell) \right)_N d\eta, \tag{5.47}$$

and the reaction term by

$$R_{\mathcal{L};N}^1 = \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} A^{(-\beta)} D_{\eta}^{\alpha} \overline{\hat{g}_k(\eta)} \hat{\rho}_{\ell}(t)_N \widehat{W}(\ell) \ell \cdot (\eta - kt) \times \left[A_k^{(-\beta)}(t, \eta) - A_{k-\ell}^{(-\beta)}(t, \eta - t\ell) \right] \left(D_{\eta}^{\alpha} \hat{g}_{k-\ell}(t, \eta - t\ell) \right)_{<N/8} d\eta, \tag{5.48}$$

and as before the remainder is whatever is left over. The treatment of the transport term $T_{\mathcal{L};N}^1$ and the remainder $\mathcal{R}_{\mathcal{L}}^1$ is unchanged from the corresponding treatments of T_N^1 and \mathcal{R}^1 in §5.3.3 and §5.3.5 respectively. Hence, we omit it and simply conclude as in (5.34), (5.39) and (5.40):

$$\left| T_{\mathcal{L};N}^1 \right| \lesssim \langle t \rangle^{-1} e^{\frac{1}{2}(c-1)\alpha_0 \langle t \rangle^s} \|A\rho(t)\|_2 \left\| \langle \nabla_{z,v} \rangle^{s/2} A^{(-\beta)}(v^{\alpha} g)_{\sim N} \right\|_2^2, \tag{5.49}$$

$$\left| \mathcal{R}_{\mathcal{L}}^1 \right| \lesssim \|A\rho(t)\|_2 \langle t \rangle^{-1} e^{\frac{1}{4}(c'-1)\alpha_0 \langle t \rangle^s} \left\| A^{(-\beta)}(v^{\alpha} g) \right\|_2^2. \tag{5.50}$$

The reaction term is slightly altered to gain from the regularity gap and get a uniform bound (as in the linear contribution $E_{\mathcal{L};L}$). As in the treatment of reaction in E_{NL}^1 in §5.3.4, we separate into $R_{\mathcal{L};N}^1 = R_{\mathcal{L};N}^{1;1} + R_{\mathcal{L};N}^{1;2}$ where the leading order reaction term is given by

$$R_{\mathcal{L};N}^{1;1} = - \sum_{k \in \mathbb{Z}^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} A^{(-\beta)} D_{\eta}^{\alpha} \overline{\hat{g}_k(\eta)} A_k^{(-\beta)}(t, \eta) \hat{\rho}_{\ell}(t)_N \widehat{W}(\ell) \ell \cdot [\eta - t\ell] \times \left(D_{\eta}^{\alpha} \hat{g}_{k-\ell}(t, \eta - t\ell) \right)_{<N/8} d\eta.$$

By the frequency localizations (3.15), (3.11) implies for some $c = c(s) \in (0, 1)$ (using also (1.3) and $|\eta - kt| \leq \langle t \rangle \langle k - \ell, \eta - t\ell \rangle$),

$$\left| R_{\mathcal{L};N}^{1;1} \right| \lesssim \langle t \rangle \sum_{k \in \mathbb{Z}_*^d} \sum_{\ell \in \mathbb{Z}_*^d} \int_{\eta} \left| A^{(-\beta)} \left(D_{\eta}^{\alpha} \hat{g}_k(\eta) \right)_{\sim N} \right| \left| A_{\ell}^{(-\beta)}(t, t\ell) \right| \left| \hat{\rho}_{\ell}(t)_N \right| \times e^{c\lambda \langle t \rangle \langle k-\ell, \eta-t\ell \rangle^s} \langle k - \ell, \eta - t\ell \rangle \left| \left(D_{\eta}^{\alpha} \hat{g}_{k-\ell}(t, \eta - t\ell) \right)_{<N/8} \right| d\eta.$$

Proceeding as in the proof of (5.36), applying (3.2) (along with $\sigma > d/2 + 2$) and using the regularity gap between $A^{(-\beta)}$ and A implies

$$\begin{aligned} \left| R_{\mathcal{L};N}^{1;1} \right| &\lesssim \langle t \rangle \left\| A^{(-\beta)}(v^{\alpha} g)_{\sim N} \right\|_2 \left\| A^{(-\beta)}\rho_N \right\|_2 \left\| A^{(-\beta)}v^{\alpha} g \right\|_2 \\ &\lesssim \langle t \rangle^{1-\beta} \left\| A^{(-\beta)}(v^{\alpha} g)_{\sim N} \right\|_2 \|A\rho_N\|_2 \left\| A^{(-\beta)}v^{\alpha} g \right\|_2 \\ &\lesssim \langle t \rangle^{1-\beta} \left\| A^{(-\beta)}v^{\alpha} g \right\|_2 \left\| A^{(-\beta)}(v^{\alpha} g)_{\sim N} \right\|_2^2 + \langle t \rangle^{1-\beta} \left\| A^{(-\beta)}v^{\alpha} g \right\|_2 \|A\rho_N\|_2^2, \end{aligned} \tag{5.51}$$

which will be sufficient for the proof of (2.12b). The term $R_{\mathcal{L};N}^{1;2}$ can be treated exactly as $R_N^{1;2}$ and hence we omit and simply conclude

$$\left| R_{\mathcal{L};N}^{1;2} \right| \lesssim \frac{e^{-\alpha_0 \langle t \rangle^s}}{N} \left\| A^{(-\beta)}(v^\alpha g) \right\|_2^2 \|A\rho(t)\|_2. \tag{5.52}$$

The term $E_{\mathcal{L};NL}^2$ is treated as in §5.3.6. By (1.3), (3.13) and the regularity gap between $A^{(-\beta)}$ and A (also (3.8) in the last line),

$$\begin{aligned} \left| E_{\mathcal{L};NL}^2 \right| &\lesssim \sum_{|j|=1; j \leq \alpha} \left\| A^{(-\beta)} v^\alpha g \right\|_2 \left\| A^{(-\beta)} v^{\alpha-j} g \right\|_2 \| \rho(t) \|_{\mathcal{F}^{\tilde{c}\lambda, 0; s}} \\ &\quad + \sum_{|j|=1; j \leq \alpha} \left\| A^{(-\beta)} v^\alpha g \right\|_2 \left\| v^{\alpha-j} g \right\|_{\mathcal{G}^{\tilde{c}\lambda, \sigma; s}} \left\| A^{(-\beta)} \rho(t) \right\|_2 \\ &\lesssim e^{(\tilde{c}-1)\alpha_0 \langle t \rangle^s} \sum_{|j|=1; j \leq \alpha} \left\| A^{(-\beta)} v^\alpha g \right\|_2 \left\| A^{(-\beta)} v^{\alpha-j} g \right\|_2 \|A\rho(t)\|_2 \\ &\quad + \langle t \rangle^{-\beta} \sum_{|j|=1; j \leq \alpha} \left\| A^{(-\beta)} v^\alpha g \right\|_2 \left\| A^{(-\beta)} v^{\alpha-j} g \right\|_2 \|A\rho(t)\|_2 \\ &\lesssim \langle t \rangle^{-\beta} \sum_{|j|=1; j \leq \alpha} \left\| A^{(-\beta)} v^\alpha g \right\|_2 \left\| A^{(-\beta)} v^{\alpha-j} g \right\|_2 \|A\rho(t)\|_2. \end{aligned} \tag{5.53}$$

Denote $\delta = -\frac{1}{4} \min(c - 1, \tilde{c} - 1, c' - 1) \alpha_0$. Collecting (5.45) (5.49), (5.50), (5.51), (5.52) and (5.53) and summing in N , splitting the linear terms with a small parameter b and combining (5.50) and (5.52) with (5.49) as in (5.42) (using also (3.8)), we have the following for some $\tilde{K} = \tilde{K}(s, \sigma, \alpha_0, C_0, d)$ (not the same as the \tilde{K} in (5.42) but this is irrelevant),

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \left\| A^{(-\beta)} v^\alpha g \right\|_2^2 &\leq \left(\tilde{K} \langle t \rangle^{-1} e^{-\delta \langle t \rangle^s} \|A\rho(t)\|_2 + \tilde{K} \langle t \rangle^{-\beta} b + \dot{\lambda}(t) \right) \\ &\quad \times \left\| \langle \nabla_{z,v} \rangle^{s/2} A^{(-\beta)}(v^\alpha g) \right\|_2^2 + \frac{\tilde{K}}{b} \langle t \rangle^{-\beta} \|A\rho(t)\|_2^2 \\ &\quad + \tilde{K} \langle t \rangle^{1-\beta} \left\| A^{(-\beta)}(v^\alpha g) \right\|_2 \left(\left\| A^{(-\beta)}(v^\alpha g) \right\|_2^2 + \|A\rho(t)\|_2^2 \right) \\ &\quad + \tilde{K} \langle t \rangle^{-\beta} \sum_{|j|=1; j \leq \alpha} \left\| A^{(-\beta)} v^\alpha g \right\|_2 \left\| A^{(-\beta)} v^{\alpha-j} g \right\|_2 \|A\rho(t)\|_2. \end{aligned} \tag{5.54}$$

By (3.8) and (2.9) we may fix b and ϵ small such that

$$\tilde{K} \sqrt{K_4} \epsilon e^{-\delta \langle t \rangle^s} + \tilde{K} \langle t \rangle^{-\beta} b \leq \frac{1}{2} |\dot{\lambda}(t)|,$$

which by (5.16), implies that the first term in (5.54) is non-positive. Therefore, summing in α , integrating with $\beta > 2$ and applying the bootstrap hypotheses (2.11) and (5.16) implies (adjusting \tilde{K} to \tilde{K}'),

$$\sum_{|\alpha| \leq M} \left\| A^{(-\beta)} v^\alpha g \right\|_2^2 \leq \epsilon^2 + \tilde{K}' K_3 \epsilon^2 + \tilde{K}' K_2^{3/2} \epsilon^3 + \tilde{K}' \sqrt{K_2} K_3 \epsilon^3 + \tilde{K}' K_2 \sqrt{K_4} \epsilon^3.$$

Hence, we take $K_2 = 1 + \tilde{K}' K_3$ and $\epsilon < K_2 \left(3\tilde{K}'\right)^{-1} \left(K_2^{3/2} + \sqrt{K_2} K_3 + K_2 \sqrt{K_4}\right)^{-1}$ to deduce (2.12b).

6 Analysis of the Plasma Echoes

The most important step to pushing linear Landau damping to the nonlinear level is analyzing and controlling the dominant weakly nonlinear effect: *the plasma echo*. Mathematically, this comes down to verifying condition (5.2) on the time-response kernels, crucial to the proof of (2.12c) in §5.1. Our choices of $\lambda(t)$ for $t \gg 1$ (in particular the choice of a) and $s > 1/(2 + \gamma)$ are both determined in this section. The analysis in this section is similar to the moment estimates carried out on the time-response kernels in §7 of [67] except with the regularity loss encoded by our choice of $\lambda(t)$ taking the place of amplitude growth. The distinction is arguably minor, but this increased precision allows for a slightly cleaner treatment and highlights more clearly the origin of the regularity requirement.

Lemma 6.1 (Time response estimate I) *Under the bootstrap hypotheses (2.11), there holds*

$$\sup_{t \in [0, T^*]} \sup_{k \in \mathbb{Z}_*^d} \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \bar{K}_{k, \ell}(t, \tau) \, d\tau \lesssim_{a, s, d, \lambda_0, \lambda'} \sqrt{K_2} \epsilon.$$

Proof Consider first the effect of g_0 , the homogeneous part of g , which corresponds to $\bar{K}_{k, k}(t, \tau)$:

$$\begin{aligned} \mathcal{I}_{inst}(t) &:= \int_0^t e^{(\lambda(t) - \lambda(\tau)) \langle k, k \rangle^s} e^{c\lambda(\tau) \langle k(t-\tau) \rangle^s} \frac{|k(t-\tau)|}{|k|^\gamma} |\widehat{g}_0(\tau, k(t-\tau))| \, d\tau \\ &\leq \int_0^t e^{c\lambda(\tau) \langle k(t-\tau) \rangle^s} \frac{|k(t-\tau)|}{|k|^\gamma} |\widehat{g}_0(\tau, k(t-\tau))| \, d\tau. \end{aligned}$$

Here *inst* stands for ‘instantaneous’ as this effect has no time delay (unlike $k \neq \ell$ below); this terminology was borrowed from [67]. Also note that this is only controlling the effect of ‘low’ frequencies in g_0 . From the $H^{d/2+} \hookrightarrow C^0$ embedding, $\sigma > \beta + 1$ and (2.13),

$$\begin{aligned} \mathcal{I}_{inst}(t) &\leq \int_0^t e^{(c-1)\lambda(\tau)\langle k(t-\tau)\rangle^s} \left(\sup_{\eta \in \mathbb{R}^d} e^{\lambda(\tau)\langle \eta \rangle^s} |\eta| |\widehat{g}_0(\tau, \eta)| \right) d\tau \\ &\lesssim_M \int_0^t e^{(c-1)\lambda(\tau)\langle k(t-\tau)\rangle^s} \left\| A^{(-\beta)} g_0(\tau) \right\|_{H^M_\eta} d\tau \\ &\lesssim_{\alpha_0} \sqrt{K_2} \epsilon. \end{aligned}$$

Next turn to the contributions from the case $k \neq \ell$, which is the origin of the plasma echoes. Using $|k(t - \tau)| \leq \langle \tau \rangle |k - \ell, kt - \ell\tau|$ and the definition of \bar{K} in (5.1),

$$\mathbf{1}_{k \neq \ell} \bar{K}_{k, \ell}(t, \tau) \lesssim e^{(\lambda(t) - \lambda(\tau))\langle k, kt \rangle^s} e^{c\lambda(\tau)\langle k - \ell, kt - \ell\tau \rangle^s} \frac{\langle \tau \rangle}{|\ell|^\gamma} |\widehat{\nabla} g_{k - \ell}(\tau, kt - \ell\tau)| \mathbf{1}_{k \neq \ell \neq 0}.$$

In what follows denote

$$-v(t, \tau) = \lambda(t) - \lambda(\tau).$$

Then using that $\lambda(t) \geq \alpha_0$ and $c < 1$, if we write $\delta = (1 - c)\alpha_0$ we are left to estimate,

$$\begin{aligned} \mathcal{I}(t) &:= \int_0^t \sum_{\ell \in \mathbb{Z}^d_*} e^{-v(t, \tau)\langle k, kt \rangle^s} e^{c\lambda(\tau)\langle k - \ell, kt - \ell\tau \rangle^s} \frac{\langle \tau \rangle}{|\ell|^\gamma} |\widehat{\nabla} g_{k - \ell}(\tau, kt - \ell\tau)| \mathbf{1}_{k \neq \ell} d\tau \\ &\lesssim \int_0^t \sum_{\ell \in \mathbb{Z}^d_*} e^{-v(t, \tau)\langle k, kt \rangle^s} \frac{\langle \tau \rangle}{|\ell|^\gamma} e^{-\delta\langle k - \ell, kt - \ell\tau \rangle^s} \\ &\quad \left| e^{\lambda(\tau)\langle k - \ell, kt - \ell\tau \rangle^s} \widehat{\nabla} g_{k - \ell}(\tau, kt - \ell\tau) \right| \mathbf{1}_{k \neq \ell} d\tau. \end{aligned} \tag{6.1}$$

By $\sigma \geq \beta + 1$, the $H^{d/2+} \hookrightarrow C^0$ embedding and (2.13),

$$\begin{aligned} \left| e^{\lambda(\tau)\langle k - \ell, kt - \ell\tau \rangle^s} \widehat{\nabla} g_{k - \ell}(\tau, kt - \ell\tau) \right| &\leq \sup_{\eta \in \mathbb{R}^d} e^{\lambda(\tau)\langle k - \ell, \eta \rangle^s} \langle k - \ell, \eta \rangle |\widehat{g}_{k - \ell}(\tau, \eta)| \\ &\leq \left(\sum_{k \in \mathbb{Z}^d} \sup_{\eta \in \mathbb{R}^d} e^{2\lambda(\tau)\langle k, \eta \rangle^s} \langle k, \eta \rangle^2 |\widehat{g}_k(\tau, \eta)|^2 \right)^{1/2} \\ &\lesssim \left\| A^{(-\beta)} g(\tau) \right\|_{L^2_k H^M_\eta} \\ &\lesssim_M \sqrt{K_2} \epsilon. \end{aligned}$$

Applying this to (6.1) implies

$$\mathcal{I}(t) \lesssim \sqrt{K_2} \epsilon \int_0^t \sum_{\ell \in \mathbb{Z}^d_*} e^{-v(t, \tau)\langle k, kt \rangle^s} \frac{\langle \tau \rangle}{|\ell|^\gamma} e^{-\delta\langle k - \ell, kt - \ell\tau \rangle^s} \mathbf{1}_{\ell \neq k} d\tau. \tag{6.2}$$

Following an argument similar to that in [67] we may reduce to the $d = 1$ case. By (3.9c),

$$\begin{aligned} \mathcal{I}(t) &\lesssim \sqrt{K_2} \epsilon \int_0^t \sum_{\ell \in \mathbb{Z}_*^d} \sum_{j: \ell_j \neq k_j} e^{-\nu(t, \tau) \langle k_j, k_j t \rangle^s} \frac{\langle \tau \rangle}{|\ell|^\gamma} e^{-C_s \delta \langle k_j - \ell_j, k_j t - \ell_j \tau \rangle^s} \\ &\quad \prod_{\substack{i=1 \\ i \neq j}}^d e^{-C_s^{d-1} \delta \langle k_i - \ell_i \rangle^s} \mathbf{1}_{\ell \neq k} \, d\tau \\ &\lesssim \frac{\sqrt{K_2} \epsilon}{\delta^{\frac{d-1}{s}}} \sum_{1 \leq j \leq d} \int_0^t \sum_{\ell_j \in \mathbb{Z}} e^{-\nu(t, \tau) \langle k_j, k_j t \rangle^s} \frac{\langle \tau \rangle}{\langle \ell_j \rangle^\gamma} e^{-C_s \delta \langle k_j - \ell_j, k_j t - \ell_j \tau \rangle^s} \mathbf{1}_{\ell_j \neq k_j} \, d\tau. \end{aligned}$$

Notice that we may not assert that both k_j and ℓ_j are non-zero. However, if either k_j or ℓ_j is zero we have by (3.9c), (3.8) and $\tau \leq t$,

$$\begin{aligned} \langle \tau \rangle e^{-C_s \delta \langle k_j - \ell_j, k_j t - \ell_j \tau \rangle^s} &\leq \langle \tau \rangle e^{-C_s^2 \delta \langle k_j - \ell_j \rangle^s - C_s^2 \delta \langle k_j t - \ell_j \tau \rangle^s} \\ &\lesssim \delta^{-1/s} e^{-C_s^2 \delta \langle k_j - \ell_j \rangle^s - \frac{1}{2} C_s^2 \delta \langle \tau \rangle^s}. \end{aligned}$$

Hence, we see that such cases cannot contribute anything to the sum in k of $\mathcal{I}(t)$ which is not bounded uniformly in time. Therefore, up to adjusting the definition of δ by a constant, we may focus on the cases such that both $k, \ell \in \mathbb{Z}_*$ and $k \neq \ell$. Let us now focus on one such choice:

$$\mathcal{I}_{k, \ell}(t) := \int_0^t e^{-\nu(t, \tau) \langle k, k t \rangle^s} \frac{\langle \tau \rangle}{|\ell|^\gamma} e^{-\delta \langle k - \ell, k t - \ell \tau \rangle^s} \, d\tau.$$

This term isolates a single possible echo at $\tau = tk/\ell$: notice how the integrand is sharply localized near this time which accounts for the effect $\rho_\ell(\tau\ell)$ has on the behavior of $\rho_k(kt)$. Summing them deals with the cumulative effect of all the echoes. See [67] for more discussion. By symmetry we need only consider the case $k \geq \ell$.

Let us first eliminate the irrelevant early times; indeed by (3.9c),

$$\begin{aligned} \int_0^{\min(1, t)} e^{-\nu(t, \tau) \langle k, k t \rangle^s} \frac{\langle \tau \rangle}{|\ell|^\gamma} e^{-\delta \langle k - \ell, k t - \ell \tau \rangle^s} \, d\tau &\lesssim \int_0^{\min(1, t)} \frac{1}{|\ell|^\gamma} e^{-\delta \langle k - \ell, k t - \ell \tau \rangle^s} \, d\tau \\ &\lesssim \frac{e^{-C_s \delta \langle k - \ell \rangle^s}}{\delta^{1/s} |\ell|^{1+\gamma}}. \end{aligned} \tag{6.3}$$

Now let us turn to the more interesting contributions of $t \geq \tau \geq 1$. Given t, k and ℓ , define the *resonant interval* as

$$I_R = \left\{ \tau \in [1, t] : |k t - \ell \tau| < \frac{t}{2} \right\}$$

and divide $\mathcal{I}_{k,\ell}$ into three contributions (one from (6.3)):

$$\begin{aligned} \mathcal{I}_{k,\ell}(t) &\lesssim \frac{1}{\delta^{1/s} |\ell|^{1+\gamma}} e^{-C_s \langle k-\ell \rangle^s} + \left(\int_{[1,t] \cap I_R} + \int_{[1,t] \setminus I_R} \right) \\ &\quad \times \frac{\langle \tau \rangle}{|\ell|^\gamma} e^{-\delta \langle k-\ell, kt-\ell\tau \rangle^s} e^{-\nu(t,\tau) \langle k,kt \rangle^s} d\tau \\ &= \frac{1}{\delta^{1/s} |\ell|^{1+\gamma}} e^{-C_s \delta \langle k-\ell \rangle^s} + \mathcal{I}_R + \mathcal{I}_{NR}. \end{aligned}$$

Here ‘NR’ stands for ‘non-resonant’. Note that if $\ell \leq k-1$ then in fact $[0, t] \cap I_R = \emptyset$.

Consider first the easier case of \mathcal{I}_{NR} . Since $|kt - \ell t| \geq t/2$ on the support of the integrand, by (3.9c) and (3.8) we have

$$\begin{aligned} \mathcal{I}_{NR} &\leq \frac{\langle t \rangle}{|\ell|^\gamma} \int_{[1,t] \setminus I_R} e^{-C_s \delta \langle k-\ell \rangle^s - C_s \delta \langle kt-\ell\tau \rangle^s} e^{-\nu(\tau,t) \langle k,kt \rangle^s} d\tau \\ &\leq \frac{\langle t \rangle}{|\ell|^\gamma} e^{-C_s \delta \langle k-\ell \rangle^s - \frac{1}{2} C_s \delta \langle \frac{t}{2} \rangle^s} \int_0^t e^{-\frac{1}{2} C_s \delta \langle kt-\ell\tau \rangle^s} d\tau \\ &\lesssim \frac{\langle t \rangle}{\delta^{1/s} |\ell|^{1+\gamma}} e^{-C_s \delta \langle k-\ell \rangle^s - \frac{1}{2} C_s \delta \langle \frac{t}{2} \rangle^s} \\ &\lesssim \frac{1}{\delta^{2/s} |\ell|^{1+\gamma}} e^{-C_s \delta \langle k-\ell \rangle^s}, \end{aligned} \tag{6.4}$$

which suffices to treat all of the non-resonant contributions.

Now focus on the resonant contribution \mathcal{I}_R , which as pointed out above, is only present if $\ell \geq k+1$ due to the echo at $\tau = tk/\ell \in (0, t)$. Since we are interested in $t \geq \tau \geq 1$, by the definition of $\lambda(t)$ in (2.8), there exists some constant δ' (possibly adjusted by the reduction to one dimension) which is proportional to $\lambda_0 - \lambda'$ such that on $[1, t] \cap I_R$,

$$\nu(t, \tau) = \delta' \left(\tau^{-a} - t^{-a} \right) = \delta' \left(\frac{t^a - \tau^a}{\tau^a t^a} \right).$$

For t and τ well separated, this provides a gap of regularity that helps us to control \mathcal{I}_R . Hence, we see that the most dangerous echoes occur for $\ell \approx k$ as these echoes are stacking up near t and the regularity gap provided by ν becomes very small. From the formal analysis of [67] we expect to find the requirement $s > 1/(2 + \gamma)$ due precisely to this effect. Indeed we will see that is the case, in fact, here is the only place in the proof of Theorem 1 where this requirement is used (also at the analogous step in the proof of Lemma 6.2 below). By the mean-value theorem and the restriction that $\tau \in I_R$ (also $\tau \leq \frac{3kt}{2\ell}$ and $\ell - k \geq 1$), we have

$$\begin{aligned} \nu(t, \tau) &\geq a\delta' \frac{t - \tau}{\tau^a t} = \frac{a\delta'}{\tau^a t} \left[t - \frac{kt}{\ell} \right] - \frac{a\delta'}{\tau^a t} \left[\tau - \frac{kt}{\ell} \right] \\ &\geq \frac{a\delta'}{\tau^a} \left[1 - \frac{k}{\ell} \right] - \frac{a\delta'}{2\tau^a \ell} \end{aligned}$$

$$\begin{aligned} &\geq \frac{a\delta'}{2\tau^a\ell} \\ &\geq \frac{a\delta'}{2^{1-a}3^a(k\tau)^a\ell^{1-a}}. \end{aligned} \tag{6.5}$$

Let $\tilde{\delta}' = \frac{a\delta'}{2^{1-a}3^a}$. The lower bound (6.5) precisely measures the usefulness of ν . Indeed, by (6.5), (3.9c), (3.8) and $(2 + \gamma)(s - a) = 1 - a$ we have

$$\begin{aligned} \mathcal{I}_R &\lesssim \int_{I_R} \frac{kt}{\ell^{1+\gamma}} e^{-\delta(k-\ell, kt-\ell\tau)^s} e^{-\frac{\tilde{\delta}'}{\ell^{1-a}}|kt|^{s-a}} d\tau \\ &\lesssim \frac{kt}{\delta^{1/s}\ell^{2+\gamma}} e^{-\frac{\tilde{\delta}'}{\ell^{1-a}}|kt|^{s-a}} e^{-C_s\delta(k-\ell)^s} \\ &\lesssim \frac{kt}{\delta^{1/s}\ell^{2+\gamma}} \left(\frac{\ell^{\frac{1-a}{s-a}}}{(\tilde{\delta}')^{\frac{1}{s-a}}kt} \right) e^{-C_s\delta(k-\ell)^s} \\ &\lesssim_{s,a} e^{-C_s\delta(k-\ell)^s} \frac{1}{\delta^{1/s}(a\delta')^{\frac{1}{s-a}}}. \end{aligned} \tag{6.6}$$

The use of $(2 + \gamma)(s - a) \geq 1 - a$ above is exactly the mathematical origin of the requirement $s > (2 + \gamma)^{-1}$. Notice also that (6.6) can be summed in either k or l , but not in both.

Assembling (6.3), (6.4) and (6.6) implies the lemma after summing in ℓ and taking the supremum in t and k . □

The next estimate is in some sense the ‘dual’ of Lemma 6.1 and is proved in the same way.

Lemma 6.2 (Time response estimate II) *Under the bootstrap hypotheses (2.11) there holds*

$$\sup_{\tau \in [0, T^*]} \sup_{\ell \in \mathbb{Z}_*^d} \sum_{k \in \mathbb{Z}_*^d} \int_{\tau}^{T^*} \bar{K}_{k,\ell}(t, \tau) dt \lesssim_{a,s,d,\lambda_0,\lambda'} \sqrt{K_2} \epsilon.$$

Proof First consider $\bar{K}_{k,k}(t, \tau)$, which corresponds to the homogeneous part of g :

$$\mathcal{I}_{inst}(\tau) := \int_{\tau}^{T^*} e^{(\lambda(t)-\lambda(\tau))(k,kt)^s} e^{c\lambda(\tau)(k(t-\tau))^s} \frac{|k(t-\tau)|}{|k|^\gamma} |\widehat{g}_0(\tau, k(t-\tau))| dt.$$

By the same argument as used in Lemma 6.1, it is straightforward to show

$$\mathcal{I}_{inst}(\tau) \lesssim \sqrt{K_2} \epsilon.$$

Next consider the case $k \neq \ell$. By following the analysis of Lemma 6.1 the problem reduces to analyzing the analogue of (6.2):

$$\mathcal{I}(\tau) = \sqrt{K_2} \epsilon \int_{\tau}^{T^*} \sum_{k \in \mathbb{Z}_*^d} e^{-\nu(t, \tau) \langle k, kt \rangle^s} e^{-\delta \langle k - \ell, kt - \ell \tau \rangle^s} \frac{\langle \tau \rangle}{|\ell|^\gamma} \mathbf{1}_{\ell \neq k} dt$$

where $\nu(t, \tau) = \lambda(\tau) - \lambda(t)$ and δ are defined as in Lemma 6.1. As before we may reduce to the one dimensional case at the cost of adjusting the constant and the definition of δ . Hence consider the one dimensional integrals with $k, \ell \in \mathbb{Z}_*$, $k \neq \ell$ and $k \geq 1$ (by symmetry):

$$\mathcal{I}_{k, \ell}(\tau) = \int_{\tau}^{T^*} e^{-\nu(t, \tau) \langle k, kt \rangle^s} e^{-\delta \langle k - \ell, kt - \ell \tau \rangle^s} \frac{\langle \tau \rangle}{|\ell|^\gamma} dt. \tag{6.7}$$

As in the proof of Lemma 6.1, we may eliminate early times; we omit the details and conclude

$$\int_{\tau}^{\max(\tau, \min(1, T^*))} e^{-\nu(t, \tau) \langle k, kt \rangle^s} e^{-\delta \langle k - \ell, kt - \ell \tau \rangle^s} \frac{\langle \tau \rangle}{|\ell|^\gamma} dt \lesssim \frac{1}{\delta^{1/s} |\ell|^\gamma |k|} e^{-C_s \delta \langle k - \ell \rangle^s}.$$

For the remainder of the proof, we will henceforth just assume $T^* > \tau \geq 1$. Following the proof of Lemma 6.1, define the resonant interval as

$$I_R = \left\{ t \in [\tau, T^*] : |kt - \ell \tau| < \frac{\tau}{2} \right\}$$

and divide the integral into two main contributions:

$$\begin{aligned} \mathcal{I}_{k, \ell}(\tau) &= \left(\int_{[\tau, T^*] \cap I_R} + \int_{[\tau, T^*] \setminus I_R} \right) \frac{\langle \tau \rangle}{|\ell|^\gamma} e^{-\delta \langle k - \ell, kt - \ell \tau \rangle^s} e^{-\nu(t, \tau) \langle k, kt \rangle^s} dt \\ &= \mathcal{I}_R + \mathcal{I}_{NR}. \end{aligned}$$

The non-resonant contribution follows essentially the same proof as (6.4) in Lemma 6.1; we omit the details and conclude

$$\mathcal{I}_{NR} \lesssim \frac{1}{\delta^{2/s} |\ell|^\gamma k} e^{-C_s \delta \langle k - \ell \rangle^s}. \tag{6.8}$$

Turn now to the resonant integral, in which case $\ell \geq k + 1$, and there is an echo at $\ell \tau / k = t \in (\tau, \infty)$. Since we are interested in $t \geq \tau \geq 1$, by the definition of $\lambda(t)$ in (2.8), there exists some constant δ' (possibly adjusted by the reduction to one dimension) which is proportional to $\lambda_0 - \lambda'$ such that by the mean-value theorem and the restriction that $t \in I_R$ (also since $\frac{kt}{2\ell} \leq \tau$),

$$\begin{aligned}
 v(t, \tau) &\geq a\delta' \frac{t - \tau}{\tau^\alpha t} \geq \frac{a\delta'}{\tau^\alpha t} \left[\frac{\ell\tau}{k} - \tau \right] - \frac{a\delta'}{\tau^\alpha t} \left[t - \frac{\ell\tau}{k} \right] \\
 &\geq \frac{a\delta' \tau^{1-a}}{2tk} \\
 &\geq \frac{a\delta'}{2^{2-a} \ell^{1-a} (kt)^a}.
 \end{aligned}$$

If we now let $\tilde{\delta}' = \frac{a\delta'}{2^{2-a}}$ and apply (3.9c), (3.8) and $(2 + \gamma)(s - a) = 1 - a$ then we have

$$\begin{aligned}
 \mathcal{I}_R &\lesssim \int_{I_R} \frac{kt}{\ell^{1+\gamma}} e^{-\delta(k-\ell, kt-\ell\tau)^s} e^{-\frac{\tilde{\delta}'}{\ell^{1-a}} |kt|^{s-a}} dt \\
 &\lesssim \int_{I_R} \frac{\ell^{\frac{1-a}{s-a}}}{\ell^{1+\gamma} (\tilde{\delta}')^{\frac{1}{s-a}}} e^{-C_s \delta(k-\ell)^s - C_s \delta(kt-\ell\tau)^s} dt \\
 &\lesssim \frac{\ell^{\frac{1-a}{s-a}}}{\ell^{2+\gamma} \delta^{1/s} (\tilde{\delta}')^{\frac{1}{s-a}}} \left(\frac{\ell e^{-C_s \delta(k-\ell)^s}}{k} \right) \\
 &\lesssim \frac{1}{\delta^{2/s} (\tilde{\delta}')^{\frac{1}{s-a}}} e^{-\frac{1}{2} C_s \delta(k-\ell)^s}, \tag{6.9}
 \end{aligned}$$

which is summable in k uniformly in l (the extra power of k in the denominator of the penultimate line came from the time integration).

Assembling the contributions of (6.8) and (6.9), summing in k and taking the supremum in ℓ and $\tau \leq \infty$ completes the proof of Lemma 6.2. \square

The following simple lemma is used in §5.2 above to deduce the pointwise-in-time control on ρ .

Lemma 6.3 *Under the bootstrap hypotheses (2.11) we have*

$$\sup_{0 \leq \tau \leq t} \sup_{\ell \in \mathbb{Z}_*^d} \sum_{k \in \mathbb{Z}_*^d} \bar{K}_{k,\ell}(t, \tau) \lesssim \sqrt{K_2} \epsilon(t).$$

Proof As in the proof of Lemmas 6.1 and 6.2, we may control g by (2.11b) and reduce to dimension one, leaving us to analyze the analogue of (6.7) except without the time integral:

$$\mathcal{I}_{k,\ell}(t, \tau) = e^{-v(t,\tau)(k,kt)^s} e^{-\delta(k-\ell, kt-\ell\tau)^s} \frac{\langle \tau \rangle}{|\ell|^\gamma}.$$

By using (3.9c) we have,

$$\mathcal{I}_{k,\ell}(t, \tau) \lesssim e^{-C_s \delta(k-\ell)^s} \langle \tau \rangle,$$

which after summing in k and taking the supremum in ℓ and $\tau \leq t$ gives the lemma. \square

7 Final Steps of Proof

By Proposition 2.2, (2.12) holds uniformly in time. By (1.3) and the algebra property (3.14),

$$\begin{aligned} & \int_0^T \left\| F(t, z + vt) \cdot (\nabla_v - t\nabla_z)(f^0 + g)(t) \right\|_{\mathcal{G}^{\alpha_0}} dt \\ & \lesssim \int_0^T \|\rho(t)\|_{\mathcal{F}^{\alpha_0}} \left\| (\nabla_v - t\nabla_z)(f^0 + g)(t) \right\|_{\mathcal{G}^{\alpha_0}} dt. \end{aligned}$$

Therefore, (2.12), (3.8), $\lambda(t) \geq \alpha_0$, (1.9) and $\sigma > \beta + 1$ imply

$$\begin{aligned} & \int_0^T \left\| F(t, z + vt) \cdot (\nabla_v - t\nabla_z)(f^0 + g)(t) \right\|_{\mathcal{G}^{\alpha_0}} dt \\ & \lesssim \int_0^T \langle t \rangle^{-\sigma+1} \|A\rho(t)\|_2 \left\| A^{(-\beta)}(f^0 + g)(t) \right\|_2 dt \\ & \lesssim \left(\int_0^T \|A\rho(t)\|_2^2 dt \right)^{1/2} \left(\int_0^T \langle t \rangle^{-2\sigma+2} \left\| A^{(-\beta)}(f^0 + g)(t) \right\|_2^2 dt \right)^{1/2} \\ & \lesssim \epsilon. \end{aligned}$$

Therefore, we may define g_∞ satisfying $\|g_\infty\|_{\mathcal{G}^{\alpha_0}} \lesssim \epsilon$ by the absolutely convergent integral

$$g_\infty = h_{in} - \int_0^\infty F(\tau, z + v\tau) \cdot (\nabla_v - \tau\nabla_z)g(\tau) d\tau.$$

Moreover, again by (3.14), (2.12) and (3.8),

$$\begin{aligned} \|g(t) - g_\infty\|_{\mathcal{G}^{\lambda'}} & \lesssim \int_t^\infty e^{(\lambda' - \alpha_0)\langle \tau \rangle^s} \langle \tau \rangle^{-\sigma+1} \|A\rho(\tau)\|_2 \left\| A^{(-\beta)}(f^0 + g)(\tau) \right\|_2 d\tau \\ & \lesssim \epsilon e^{\frac{1}{2}(\lambda' - \lambda_0)\langle t \rangle^s}, \end{aligned}$$

which implies (1.12a). Then Lemma 5.2 implies (1.12b) (since $\sigma > 1/2$), completing the proof of Theorem 1.

We briefly sketch the refinement mentioned in Remark 6. Specifically we verify that the final state predicted by the linear theory is accurate to within $O(\epsilon^2)$. Indeed, let g^L be the solution to

$$\begin{cases} \partial_t g^L + F^L(t, z + vt) \cdot \nabla_v f^0 = 0, \\ F^L(\widehat{t, z + vt})(t, k, \eta) = -ik \widehat{W}(k) \widehat{g^L}_k(t, kt) \delta_{\eta=kt}, \\ g^L(t = 0, z, v) = h_{in}(z, v). \end{cases} \tag{7.1}$$

By the analysis of §4 we have that $h^L(t, x, v) = g^L(t, x - tv, v)$ satisfies the conclusions of Theorem 1 for $h_\infty^L = h_\infty^L(z, v)$ given by

$$h_\infty^L(z, v) = h_{\text{in}}(z, v) - \int_0^\infty F^L(t, z + vt) \cdot \nabla_v f^0(v) dt.$$

Consider next the PDE

$$\partial_t \left(g - g^L \right) + \left(F - F^L \right) (t, z + vt) \cdot \nabla_v f^0 = -F(t, z + vt) \cdot (\nabla_v - t \nabla_z) g.$$

By treating the right-hand side as a decaying external force, the analysis of §4 with λ' replaced by $\lambda'' < \lambda'$, then implies

$$\left\| g(t) - g^L(t) \right\|_{\lambda''} \lesssim_{\lambda' - \lambda''} \epsilon^2,$$

which shows that the nonlinearity changes the linear behavior at the expected $O(\epsilon^2)$ order. Justifying higher order expansions should also be possible, but justifying the convergence of a Newton iteration is significantly more challenging as the constants would need to be quantified carefully.

Acknowledgments J. Bedrossian research partially supported by NSF Postdoctoral Fellowship in Mathematical Sciences, DMS-1103765. N. Masmoudi research partially supported by NSF grant DMS-1211806. C. Mouhot research partially funded by ERC grant MATKIT. The authors would like to thank the following people for references and suggestions: Antoine Cerfon, Yan Guo, George Hagstrom, Vladimir Šverák, Cédric Villani and Walter Strauss.

References

1. Adams, R.A., Fournier, J.J.F.: Sobolev Spaces. Pure and Applied Mathematics, vol. 140, 2nd edn. Elsevier/Academic Press, Amsterdam (2003)
2. Arendt, W., Grabosch, A., Greiner, G., Groh, U., Lotz, H.P., Moustakas, U., Nagel, R., Neubrander, F., Schlotterbeck, U.: One-Parameter Semigroups of Positive Operators. Lecture Notes in Mathematics, vol. 1184. Springer, Berlin (1986)
3. Arnold, V.I., Khesin, B.A.: Topological Methods in Hydrodynamics. Applied Mathematical Sciences, vol. 125. Springer, New York (1998)
4. Backus, G.: Linearized plasma oscillations in arbitrary electron distributions. *J. Math. Phys.* **1**, 178–191 (1960). 559
5. Bahouri, H., Chemin, J.-Y., Danchin, R.: Fourier Analysis and Nonlinear Partial Differential Equations. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 343. Springer, Heidelberg (2011)
6. Balmforth, N.J., Morrison, P.J.: Normal modes and continuous spectra. *Ann. N. Y. Acad. Sci.* **773**(1), 80–94 (1995)
7. Balmforth, N.J., Morrison, P.J.: Singular eigenfunctions for shearing fluids I. Institute for Fusion Studies Report. University of Texas-Austin, vol. 692, pp. 1–80 (1995)
8. Balmforth, N.J., Morrison, P.J.: Hamiltonian description of shear flow. In: Large-Scale Atmosphere–Ocean Dynamics, vol. II, pp. 117–142. Cambridge University Press, Cambridge (2002)
9. Balmforth, N.J., Morrison, P.J., Thiffeault, J.-L.: Pattern formation in Hamiltonian systems with continuous spectra; a normal-form single-wave model. Preprint (2013)
10. Bedrossian, J., Masmoudi, N.: Inviscid damping and the asymptotic stability of planar shear flows in the 2D Euler equations. [arXiv:1306.5028](https://arxiv.org/abs/1306.5028)

11. Bedrossian, J., Masmoudi, N., Vicol, V.: Enhanced dissipation and inviscid damping in the inviscid limit of the Navier–Stokes equations near the 2D Couette flow. [arXiv:1408.4754](https://arxiv.org/abs/1408.4754) 2014, 2014
12. Benachour, S.: Analyticité des solutions des équations de Vlasov–Poisson. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **16**(1), 83–104 (1989)
13. Bernstein, I.B., Greene, J.M., Kruskal, M.D.: Exact nonlinear plasma oscillations. *Phys. Rev.* **108**, 546 (1957)
14. Boltzmann, L.: Weitere studien über das wärme gleichgewicht unter gasmolekular. *Sitz. Akad. Wiss.* **66**, 275–370 (1872). Translation: Further studies on the thermal equilibrium of gas molecules. In: Brush, S.G. (ed.) *Kinetic Theory 2*, pp. 88–174. Pergamon, Oxford (1966)
15. Bony, J.-M.: Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires. *Ann. Sci. École Norm. Sup.* **14**, 209–246 (1981)
16. Bouchet, F., Morita, H.: Large time behavior and asymptotic stability of the 2D Euler and linearized Euler equations. *Physica D* **239**, 948–966 (2010)
17. Boyd, T.J.M., Sanderson, J.J.: *The Physics of Plasmas*. Cambridge University Press, Cambridge (2003)
18. Brenier, Y.: Convergence of the Vlasov–Poisson system to the incompressible Euler equations. *Commun. Partial Differ. Equ.* **25**(3–4), 737–754 (2000)
19. Briggs, R.J., Daugherty, J.D., Levy, R.H.: Role of Landau damping in crossed-field electron beams and inviscid shear flow. *Phys. Fluids* **13**(2), 421–432 (1970)
20. Caglioti, E., Maffei, C.: Time asymptotics for solutions of Vlasov–Poisson equation in a circle. *J. Stat. Phys.* **92**(1/2), 301–323 (1998)
21. 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.*, vol. 9, pp. 99–123. Société Mathématique de France, Paris (2004)
22. Chemin, J.-Y., Gallagher, I., Paicu, M.: Global regularity for some classes of large solutions to the Navier–Stokes equations. *Ann. Math. (2)* **173**(2), 983–1012 (2011)
23. Degond, P.: Spectral theory of the linearized Vlasov–Poisson equation. *Trans. Am. Math. Soc.* **294**(2), 435–453 (1986)
24. Engel, K.-J., Nagel, R.: *One-Parameter Semigroups for Linear Evolution Equations*. Graduate Texts in Mathematics, vol. 194. Springer, New York (2000). With contributions by Brendle, S., Campiti, M., Hahn, T., Metafun, G., Nickel, G., Pallara, D., Perazzoli, C., Rhandi, A., Romanelli, S., Schnaubelt, R
25. Foias, C., Temam, R.: Gevrey class regularity for solutions of the Navier–Stokes equations. *J. Funct. Anal.* **87**, 359–369 (1989)
26. Gearhart, L.: Spectral theory for contraction semigroups on Hilbert space. *Trans. Am. Math. Soc.* **236**, 385–394 (1978)
27. Gérard-Varet, D., Masmoudi, N.: Well-posedness for the Prandtl system without analyticity or monotonicity. [arXiv:1305.0221](https://arxiv.org/abs/1305.0221), (2013)
28. Germain, P., Masmoudi, N., Shatah, J.: Global solutions for the gravity water waves equation in dimension 3. *Ann. Math. (2)* **175**(2), 691–754 (2012)
29. Gevrey, M.: Sur la nature analytique des solutions des équations aux dérivées partielles. Premier mémoire. *Ann. Sci. École Norm. Sup. (3)* **35**, 129–190 (1918)
30. Glassey, R., Schaeffer, J.: Time decay for solutions to the linearized Vlasov equation. *Transp. Theory Stat. Phys.* **23**(4), 411–453 (1994)
31. Glassey, R., Schaeffer, J.: On time decay rates in Landau damping. *Commun. Partial Differ. Equ.* **20**(3–4), 647–676 (1995)
32. Glatt-Holtz, N., Sverak, V., Vicol, V.: On inviscid limits for the stochastic Navier–Stokes equations and related models. *Eprint*. [arXiv:1302.0542](https://arxiv.org/abs/1302.0542) (2013)
33. Gripenberg, G., Londen, S.-O., Staffans, O.: *Volterra Integral and Functional Equations*. Encyclopedia of Mathematics and Its Applications, vol. 34. Cambridge University Press, Cambridge (1990)
34. Günther, M.: On the perturbation problem associated to isometric embeddings of Riemannian manifolds. *Ann. Glob. Anal. Geom.* **7**(1), 69–77 (1989)
35. Günther, M.: Zum Einbettungssatz von J. Nash. *Math. Nachr.* **144**, 165–187 (1989)
36. Günther, M.: Isometric embeddings of Riemannian manifolds. In: *Proceedings of the International Congress of Mathematicians*, vol. I, II (Kyoto, 1990), Tokyo, pp. 1137–1143. Mathematical Society, Japan (1991)
37. Guo, Y., Lin, Z.: Unstable and stable galaxy models. *Commun. Math. Phys.* **279**(3), 789–813 (2008)

38. Helffer, B., Sjöstrand, J.: From resolvent bounds to semigroup bounds. Actes du colloque d'Évian 2009. Preprint. [arXiv:1001.4171](https://arxiv.org/abs/1001.4171) (2010)
39. Herbst, I.: The spectrum of Hilbert space semigroups. *J. Oper. Theory* **10**(1), 87–94 (1983)
40. Hörmander, L.: The Nash–Moser theorem and paradifferential operators. In: *Analysis, et cetera*, pp. 429–449. Academic Press, Boston (1990)
41. Horst, E.: On the asymptotic growth of the solutions of the Vlasov–Poisson system. *Math. Methods Appl. Sci.* **16**(2), 75–86 (1993)
42. Hwang, H.J., Velaázquez, J.-L.L.: On the existence of exponentially decreasing solutions of the nonlinear Landau damping problem. *Indiana Univ. Math. J.* **58**(6), 2623–2660 (2009)
43. Kelvin, W.T.: Stability of fluid motion-rectilinear motion of viscous fluid between two parallel plates. *Philos. Mag.* **24**, 188 (1887)
44. Klainerman, S.: Long-time behavior of solutions to nonlinear evolution equations. *Arch. Ration. Mech. Anal.* **78**(1), 73–98 (1982)
45. Kukavica, I., Vicol, V.: On the radius of analyticity of solutions to the three-dimensional Euler equations. *Proc. Am. Math. Soc.* **137**(2), 669–677 (2009)
46. Landau, L.: On the vibration of the electronic plasma. *J. Phys. USSR* **10**, 25 (1946)
47. Landau, L.D.: On the vibration of the electronic plasma. *J. Phys. USSR* **10**, 25 (1946). English translation in *JETP* **16**, 574. Reproduced in *Collected Papers of L.D. Landau*, edited and with an introduction by ter Haar, D., pp. 445–460. Pergamon Press, Oxford (1965); and in ter Haar, D. (ed.) *Men of Physics: L.D. Landau*, vol. 2. Pergamon Press, Oxford (1965)
48. Lemou, M., Méhats, F., Raphaël, P.: The orbital stability of the ground states and the singularity formation for the gravitational Vlasov Poisson system. *Arch. Ration. Mech. Anal.* **189**(3), 425–468 (2008)
49. Lemou, M., Méhats, F., Raphaël, P.: Orbital stability of spherical galactic models. *Invent. Math.* **187**(1), 145–194 (2012)
50. Levermore, D., Oliver, M.: Analyticity of solutions for a generalized Euler equation. *J. Differ. Equ.* **133**, 321–339 (1997)
51. Lin, Z., Zeng, C.: Inviscid dynamic structures near Couette flow. *Arch. Ration. Mech. Anal.* **200**, 1075–1097 (2011)
52. Lin, Z., Zeng, C.: Small BGK waves and nonlinear Landau damping. *Commun. Math. Phys.* **306**(2), 291–331 (2011)
53. Lindblad, H., Rodnianski, I.: Global existence for the Einstein vacuum equations in wave coordinates. *Commun. Math. Phys.* **256**(1), 43–110 (2005)
54. Lions, P.-L., Perthame, B.: Propagation of moments and regularity for the 3-dimensional Vlasov–Poisson system. *Invent. Math.* **105**(2), 415–430 (1991)
55. Lynden-Bell, D.: The stability and vibrations of a gas of stars. *Mon. Not. R. Astron. Soc.* **124**(4), 279–296 (1962)
56. Lynden-Bell, D.: Statistical mechanics of violent relaxation in stellar systems. *Mon. Not. R. Astron. Soc.* **136**, 101–121 (1967)
57. Malmberg, J., Wharton, C.: Collisionless damping of electrostatic plasma waves. *Phys. Rev. Lett.* **13**(6), 184–186 (1964)
58. Malmberg, J., Wharton, C., Gould, C., O’Neil, T.: Plasma wave echo. *Phys. Rev. Lett.* **20**(3), 95–97 (1968)
59. Maslov, V.P., Fedoryuk, M.V.: The linear theory of Landau damping. *Mat. Sb. (N.S.)* 127(169):445–475, 559 (1985)
60. Masmoudi, N.: From Vlasov–Poisson system to the incompressible Euler system. *Commun. Partial Differ. Equ.* **26**(9–10), 1913–1928 (2001)
61. Maxwell, J.C.: On the dynamical theory of gases. *Philos. Trans. R. Soc. Lond.* **157**, 49–88 (1867)
62. Morrison, P.J.: Hamiltonian description of the ideal fluid. *Rev. Mod. Phys.* **70**(2), 467–521 (1998)
63. Morrison, P.J.: Hamiltonian description of Vlasov dynamics: action-angle variables for the continuous spectrum. *Trans. Theory Stat. Phys.* **29**(3–5), 397–414 (2000)
64. Moser, J.: A rapidly convergent iteration method and non-linear partial differential equations. I. *Ann. Scuola Norm. Sup. Pisa* **20**(3), 265–315 (1966)
65. Moser, J.: A rapidly convergent iteration method and non-linear partial differential equations. II. *Ann. Scuola Norm. Sup. Pisa* **20**(3), 499–535 (1966)

66. Mouhot, C., Stabilité orbitale pour le système de Vlasov–Poisson gravitationnel [d’après Lemou–Méhats–Raphaël, Guo, Lin, Rein et al.]. *Astérisque* 352 (2013). Séminaire Bourbaki, année 2011–2012. Exp. No. 1044
67. Mouhot, C., Villani, C.: On Landau damping. *Acta Math.* **207**, 29–201 (2011)
68. Nash, J.: The imbedding problem for Riemannian manifolds. *Ann. Math.* **2**(63), 20–63 (1956)
69. Nirenberg, L.: An abstract form of the nonlinear Cauchy–Kowalewski theorem. *J. Differ. Geom.* **6**, 561–576 (1972)
70. Nishida, T.: A note on a theorem of Nirenberg. *J. Differ. Geom.* **12**, 629–633 (1977)
71. Orr, W.: The stability or instability of steady motions of a perfect liquid and of a viscous liquid. Part I: A perfect liquid. *Proc. R. Ir. Acad. Sect. A Math. Phys. Sci.* **27**, 9–68 (1907)
72. Paley, R.E.A.C., Wiener, N.: *Fourier Transforms in the Complex Domain*. American Mathematical Society Colloquium Publications, vol. 19. American Mathematical Society, Providence (1987). Reprint of the 1934 original
73. Penrose, O.: Electrostatic instability of a uniform non-Maxwellian plasma. *Phys. Fluids* **3**, 258–265 (1960)
74. Pfaffelmoser, K.: Global classical solutions of the Vlasov–Poisson system in three dimensions for general initial data. *J. Differ. Equ.* **95**(2), 281–303 (1992)
75. Prüss, J.: On the spectrum of C_0 -semigroups. *Trans. Am. Math. Soc.* **284**(2), 847–857 (1984)
76. Rayleigh, J.W.S.: On the stability, or instability, of certain fluid motions. *Proc. Lond. Math. Soc.* **S1–11**(1), 57 (1880)
77. Ryutov, D.D.: Landau damping: half a century with the great discovery. *Plasma Phys. Control. Fusion* **41**(3A), A1 (1999)
78. Schaeffer, J.: Global existence of smooth solutions to the Vlasov–Poisson system in three dimensions. *Commun. Partial Differ. Equ.* **16**(8–9), 1313–1335 (1991)
79. Schecter, D.A., Dubin, D., Cass, A.C., Driscoll, C.F., Lansky, I.M., et al.: Inviscid damping of asymmetries on a two-dimensional vortex. *Phys. Fluids* **12**(10), 2397–2412 (2000)
80. Stix, T.: *Waves in Plasmas*. Springer, New York (1992)
81. van Kampen, N.G.: On the theory of stationary waves in plasmas. *Physica* **21**, 949–963 (1955)
82. Vanneste, J.: Nonlinear dynamics of anisotropic disturbances in plane Couette flow. *SIAM J. Appl. Math.* **62**(3), 924–944 (2002). (electronic)
83. Vanneste, J., Morrison, P.J., Warn, T.: Strong echo effect and nonlinear transient growth in shear flows. *Phys. Fluids* **10**, 1398 (1998)
84. Villani, C.: Landau damping. In: *Numerical models for fusion*. Panoramic Synthèses, vol. 39–40, pp. 237–326. Société Mathématique de France, Paris (2013)
85. Vlasov, A.: On vibration properties of electron gas. *J. Exp. Theor. Phys.* **8**(3), 291 (1938)
86. Vlasov, A.A.: The vibrational properties of an electron gas. *Zh. Eksp. Teor. Fiz.* **291**(8) (1938) (in Russian). Translation in English in *Sov. Phys. Uspekhi* **93**(3, 4) (1968)
87. Young, B.: Landau damping in relativistic plasmas. Eprint. [arXiv:1408.2666](https://arxiv.org/abs/1408.2666) (2014)
88. Yu, J.H., Driscoll, C.F.: Diocotron wave echoes in a pure electron plasma. *IEEE Trans. Plasma Sci.* **30**(1), 24–25 (2002)
89. Yu, J.H., O’Neil, T.M., Driscoll, C.F.: Phase mixing and echoes in a pure electron plasma. *Phys. Plasmas* **12**, 055701 (2005)