J Stat Phys (2009) 136: 643–665
DOI 10.1007/s10955-009-9800-7

On the Fluctuations about the Vlasov Limit for N-particle Systems with Mean-Field Interactions

Carlo Lancellotti

Received: 30 April 2009 / Accepted: 30 July 2009 / Published online: 13 August 2009

Abstract Whereas the Vlasov (a.k.a. "mean-field") limit for N-particle systems with sufficiently smooth potentials has been the subject of many studies, the literature on the dynamics of the fluctuations around the limit is sparse and somewhat incomplete. The present work fulfills two goals: 1) to provide a complete, simple proof of a general theorem describing the evolution of a given initial fluctuation field for the particle density in phase space, and 2) to characterize the most general class of initial symmetric probability measures that lead (in the infinite-particle limit) to the same Gaussian random field that arises when the initial phase space coordinates of the particles are assumed to be i.i.d. random variables (so that the standard central limit theorem applies). The strategy of the proof of the fluctuation evolution result is to show first that the deviations from mean-field converge for each individual system, in a purely deterministic context. Then, one obtains the corresponding probabilistic result by a modification of the continuous mapping theorem. The characterization of the initial probability measures is in terms of a higher-order chaoticity condition (a.k.a. "Boltzmann property").

Keywords Vlasov equation · Mean-field limit · Central limit theorems · Fluctuation dynamics · Sznitman-Tanaka theorem

1 Introduction

Among the various models of classical kinetic theory, one of the earliest to be justified in terms of a rigorous derivation from microscopic N-body dynamics was the Vlasov equation for a classical plasma with sufficiently regular[1] interparticle potential, see [2, 3, 14, 17, 18].

[1] Typically, "sufficiently regular" means bounded and uniformly Lipshitz continuous. For recent efforts to extend the theory to other situations, see [5, 13] and references therein.

This work was supported by the National Science Foundation DMS Grant No. 0604946.

C. Lancellotti (✉)
Department of Mathematics, City University of New York-CSI, Staten Island, NY 10314, USA
e-mail: lancellotti@mail.csi.cuny.edu

Since the Vlasov limit can be easily given a probabilistic interpretation as a law of large numbers [2], a natural question to ask is whether one can establish a central limit theorem (CLT) for the fluctuations. The first result of this kind was established in 1977 by Braun and Hepp [2]. Their CLT, however, does not concern the fluctuations of the phase-space density of particles, which is the principal physical observable for the model under consideration. Rather, they studied the fluctuations of the trajectory of a test particle,[2] by comparing the phase flow determined by the force field of N interacting particles with the Vlasov flow obtained in the limit $N \to \infty$. Of course, in the special case when the test particle coincides with one of the N interacting particles, this amounts to describing the fluctuations of each N-particle orbit from its mean field approximation. While it may well be possible to use the Braun-Hepp CLT for the characteristics in order to prove a CLT for the phase space density, no such proof is found in Ref. [2]. An argument of this kind is implied in the book by Spohn [20, §7.5], which describes a full-fledged CLT for the Vlasov density fluctuations but gives no proof other than referring back to the work of Braun and Hepp [2].

Despite its somewhat unfinished state, this part of the theory has not attracted much interest ever since. A development came a few years later when Sznitman [21] proved a CLT for Vlasov-McKean stochastic processes that includes a Vlasov CLT as a special case. We recall that in a Vlasov-McKean process for N particles [15, 16, 21, 22] the dynamics of each particle is driven not only by a mean-field potential but also by one or more Brownian motion terms, whose diffusion coefficients also depend on the positions of the other particles in a mean-field fashion. The N-particle Newtonian equations associated with the Vlasov limit are a very simple example of Vlasov-McKean process, in which the diffusion coefficients are exactly zero and the stochasticity may come in exclusively from the initial condition, whereas the time evolution is purely deterministic. Sznitman's proof does go through in this special case; however, being designed to control a truly stochastic dynamics, it involves a large amount of probabilistic machinery and estimates that turn out to be unnecessary in the absence of Brownian terms. Moreover, Sznitman considered only the special case of initial probability measures that are products of N single-particle measures.

Rebus sic stantibus a revisitation seems to be in order. The first goal of the present study will be to work out in detail a simple, economical proof of a general CLT for fluctuations about the Vlasov limit. Such proof will not rely on the Braun-Hepp CLT for test particle trajectories and will apply to a broader class of initial conditions, with factorized initial measures as a special case. The guiding principle will be that, as long as the N-particle evolution equations are deterministic, the simplest and most easily extendable results are obtained by keeping the dynamical and probabilistic aspects as separate as possible. Specifically, it will be shown that if the deviations of the particle density from the limiting value converge at time $t = 0$ as $N \to \infty$ (in an appropriate weak topology), then they will converge at subsequent times *for any individual system*—very much like the particle density of an individual system converges to its infinite-particle limit at time t if it does so initially. Then, if instead of "deviations" (deterministic objects) one considers "fluctuations" (random variables), all is left to do is to study the propagation in time of whatever CLT is known to hold at $t = 0$. Here, this is done by a simple adaptation of the continuous mapping theorem of probability theory [12]. The observation that Vlasov fluctuations can be expected to be simply transported by the (linearized) Vlasov dynamics was already made by Spohn [20]. However, to the best of my knowledge it was never properly implemented in the literature, perhaps because of a widespread opinion that all major questions about Vlasov fluctuations had been answered by the work of Braun and Hepp [2].

[2] This fact was pointed out to me by Michael Kiessling.

The fact that the infinite-particle limit for an individual system is treated separately from the probabilistic analysis for an ensemble means that the two can also be upgraded separately. On one hand, if one can control the Vlasov limit (including the deviations) for more general interparticle potentials, then the probabilistic treatment described here will apply without changes. On the other hand, if one proves some different, possibly more sophisticated CLT at $t = 0$, the fluctuation field will still propagate in time according to the same dynamical laws described below. An interesting generalization of the CLT along these lines occurs if one studies probability measures that are not factorized but satisfy only a higher-order version ("strong μ-chaoticity") of the familiar "Boltzmann property" [11]. The second goal of the present work is to prove that strong μ-chaoticity characterizes completely the sequences of symmetric probability measures that lead as $N \to \infty$ to the same Gaussian random field as in the standard CLT for i.i.d. random variables. Then, if the initial condition is strongly μ-chaotic, our general propagation result yields at $t > 0$ the same CLT that holds for factorized initial measures [21]. It turns out, however, that strong μ-chaoticity does not propagate in time.

The rest of the article is structured as follows. The model under consideration and much of the mathematical notation is introduced in Sect. 2, in the context of a concise review of the Vlasov limit. For the sake of generality, and for notational convenience, we consider not the Vlasov case *per se* but a general system of N ordinary differential equations with mean-field coupling [3]; the Vlasov formulas will appear as examples. The original part of the work starts in Sect. 3.1, with the study of the $N \to \infty$ limit for the deviations from mean field for individual systems. The probabilistic analysis (fluctuations) is the subject of Sect. 3.2, leading to a general CLT which includes Sznitman's result [21] (without Brownian motions) as a special case. Section 4 presents the characterization of strongly μ-chaotic sequences and extends the CLT for fluctuations around the Vlasov limit to this class of initial probability measures.

2 Review of the Vlasov Limit

In this section we review very succinctly some basic results about the Vlasov limit. For the sake of simplicity we assume the force between pairs of particles to be smooth (bounded and globally Lipschitz continuous). For a broader review of the field see [13]. All these results are by now classic and divide naturally in two parts. In the first subsection the Vlasov limit is carried out for *individual* N-particle systems, in a purely deterministic setting. No probabilistic concepts are introduced and all that needs to be assumed is a sequence of initial N-particle configurations that approximate a given density in phase space as $N \to \infty$. Statistical ensembles of systems are considered in the second subsection, where the Vlasov limit takes the form of a law of large numbers for the ensemble.

2.1 Convergence of Densities

Consider a system of N objects ("particles"), described by coordinate vectors $\mathbf{z}_i \in \mathbb{R}^d$, $i = 1, \ldots, N$, whose time evolution is determined by the $d \times N$ ODEs

$$\dot{\mathbf{z}}_i = \mathbf{G}(\mathbf{z}_i) + \frac{1}{N}\sum_{j=1}^{N} \mathbf{F}(\mathbf{z}_i - \mathbf{z}_j), \tag{1}$$

with given initial conditions $\mathbf{z}_i(0) \equiv \mathbf{z}_{i,0}$. Throughout this paper $\mathbf{F}$ and $\mathbf{G}$ are taken to be in Lip($\mathbb{R}^d$), the space of globally Lipschitz continuous functions from $\mathbb{R}^d$ to $\mathbb{R}^d$, thus ensuring that the Cauchy problem for (1) is well-posed on any time-interval $[0, T]$. The most important example, of course, is that of a mechanical system of N particles with $d = 6$, $\mathbf{z}_i \equiv [\mathbf{q}_i, \mathbf{p}_i] \in \mathbb{R}^3 \times \mathbb{R}^3$, $\mathbf{G}(\mathbf{z}) \equiv [\mathbf{p}, \mathbf{K}(\mathbf{q})]$ and $\mathbf{F}(\mathbf{z}) \equiv [\mathbf{0}, -\nabla\phi(\mathbf{q})]$, where $\mathbf{K}$ is a given external force and ϕ is a central potential[3] describing the interaction between pairs of particles.

In the context of the Vlasov limit, it is convenient to consider not the state vector $(\mathbf{z}_1, \dots, \mathbf{z}_N)$ but rather the corresponding "empirical measure" on $\mathbb{R}^d$

$$\mu_t^N = \frac{1}{N} \sum_{j=1}^N \delta_{\mathbf{z}_j(t)}. \tag{2}$$

For each fixed t, μ_t^N belongs to the set $\mathcal{M}_1^+$ of all probability measures on the Borel σ-algebra $\mathcal{B}(\mathbb{R}^d)$, endowed with the topology of weak convergence, to be denoted by the symbol $\stackrel{w}{\to}$. As is well known, weak convergence on $\mathcal{M}_1^+$ can be metrized with the bounded Lipschitz (BL) metric (see [4, p. 394]). For $\mu, \nu \in \mathcal{M}_1^+$ the BL distance is given by

$$d(\mu, \nu) \stackrel{\text{def}}{=} \sup_{\|f\|_{BL} \le 1} \left| \int f \mathrm{d}(\mu - \nu) \right|, \tag{3}$$

where

$$\|f\|_{BL} \stackrel{\text{def}}{=} \sup_{\mathbf{z} \in \mathbb{R}^d} |f(\mathbf{z})| + \sup_{\substack{\mathbf{y},\mathbf{z} \in \mathbb{R}^d \\ \mathbf{y} \ne \mathbf{z}}} \frac{|f(\mathbf{y}) - f(\mathbf{z})|}{|\mathbf{y} - \mathbf{z}|}. \tag{4}$$

In what follows $\mathcal{C}_b(\mathbb{R}^d)$ (respectively, $\mathcal{C}_b^k(\mathbb{R}^d)$) will denote the Banach space of continuous and bounded (respectively, k-times continuously differentiable with bounded derivatives) functions from $\mathbb{R}^d$ to $\mathbb{R}$ with the usual norm. In order to study the Vlasov limit for a sequence of trajectories $\{\mu_t^n\}$, with $n = 1, 2, \dots$, and $t \in [0, T]$, it is natural to consider the space $\mathcal{C}_w([0, T], \mathcal{M}_1^+)$ of all the time-dependent probability measures that are weakly continuous, meaning that for all test functions $g \in \mathcal{C}_b(\mathbb{R}^d)$ the quantity $\langle \mu_t, g \rangle \equiv \int_{\mathbb{R}^d} g(\mathbf{z}) \mu_t(\mathrm{d}\mathbf{z})$ is a continuous function of t. Then, one wants to prove that if initially $\mu_0^N \stackrel{w}{\to} \mu_0$, then at subsequent times μ_t^N is also weakly convergent to a limiting evolution $\mu_t \in \mathcal{C}_w([0, T], \mathcal{M}_1^+)$. In order to characterize the limit one observes [18] that μ_t^N is a solution to the fixed-point equation in $\mathcal{C}_w([0, T], \mathcal{M}_1^+)$

$$\sigma_t = \sigma_0 \circ T_t^0[\sigma] \tag{5}$$

where σ denotes the curve $t \Rightarrow \sigma_t$, weakly continuous in t, and $T_s^t[\sigma]$ is the flow in $\mathbb{R}^d$ from time s to time t determined by

$$\dot{\mathbf{z}} = \mathbf{G}(\mathbf{z}) + \int_{\mathbb{R}^d} \mathbf{F}(\mathbf{z} - \mathbf{z}') \sigma_t(\mathrm{d}\mathbf{z}'). \tag{6}$$

That μ_t^N solves (5) follows easily from the fact that the evolution equations (1) are just a special case (with $\sigma_t = \mu_t^N$) of (6). The idea is to show that the limiting evolution for the density is also a solution to (5) but with initial condition μ_0. This is a corollary to the following

[3] Together with the continuity of $\mathbf{F}$, this ensures that $\mathbf{F}(\mathbf{0}) = \mathbf{0}$.

Theorem 1 *Let* $\mathbf{F} \in \mathcal{C}_b(\mathbb{R}^d) \bigcap \mathrm{Lip}(\mathbb{R}^d)$, $\mathbf{G} \in \mathrm{Lip}(\mathbb{R}^d)$. *Then,*

1. *For every* $\mu_0 \in \mathcal{M}_1^+$ (5) *has a unique solution* μ_t *in* $\mathcal{C}_w([0,T], \mathcal{M}_1^+)$.
2. *If* $\mu_0, \nu_0 \in \mathcal{M}_1^+$ *and if* $\mu_t, \nu_t \in \mathcal{C}_w([0,T], \mathcal{M}_1^+)$ *are the corresponding solutions to* (5), *then*

$$d(\mu_t, \nu_t) \leq e^{Ct} d(\mu_0, \nu_0), \tag{7}$$

where C *is a constant that depends only on the choice of* $\mathbf{F}$ *and* $\mathbf{G}$.

Proof The proof is found in [17, 18] (with trivial modifications to include the term $\mathbf{G}$). See also [2, 3, 14] and the reviews in [20, Chap. 5] and [8]. □

From (7) follows immediately the desired result:

Corollary 1 *Consider a sequence of initial measures* $\mu_0^{(N)} \stackrel{w}{\to} \mu_0$ *in* $\mathcal{M}_1^+$, *and let* $\mu_t^{(N)}$ *and* μ_t *be the corresponding solutions to* (5) *in* $\mathcal{C}_w([0,T], \mathcal{M}_1^+)$. *Then* $\mu_t^{(N)} \stackrel{w}{\to} \mu_t$ *as* $N \to \infty$ *for all* $t \in [0,T]$.

For future reference, observe that once the existence and uniqueness of the limiting measure μ_t has been established, also the corresponding initial value problem for the characteristics, (6) with $\sigma_t = \mu_t$, is well-posed. Moreover, the characteristics themselves converge. Precisely, let $\mathbf{z}(t; \mathbf{z}_0, \sigma_0)$ indicate the characteristic curve (a test particle's trajectory, if you like) starting at $\mathbf{z}_0 \in \mathbb{R}^d$ for a chosen initial density $\sigma_0 \in \mathcal{M}_1^+$; then

$$\lim_{N\to\infty} \sup_{\mathbf{z}_0 \in \mathbb{R}^d} \|\mathbf{z}(t; \mathbf{z}_0, \mu_0^N) - \mathbf{z}(t; \mathbf{z}_0, \mu_0)\|_{\mathbb{R}^d} = 0 \tag{8}$$

because of $\mathbf{F}, \mathbf{G} \in \mathrm{Lip}(\mathbb{R}^d)$ and Gronwall's inequality (see Appendix A.1).

Even though the "continuity" property expressed by Corollary 1 holds for any limiting measure $\mu_0 \in \mathcal{M}_1^+$, in applications the most interesting situation occurs when μ_0 is continuous with respect to Lebesgue, $\mu_0(\mathrm{d}\mathbf{z}) = f_0(\mathbf{z})\mathrm{d}\mathbf{z}$ for some $f_0 \in \mathrm{L}_+^1(\mathbb{R}^d)$. This corresponds to the physical notion of a continuous density of particles in phase space (what physicists call a "distribution function," i.e. a "mesoscopic" description of a fluid). Then, μ_t is also absolutely continuous [18] and the Vlasov limit takes the intuitive meaning that as $N \to \infty$ the "discrete" density μ_t^N approaches the "continuous" density $f_t(\mathbf{z}) \in \mathrm{L}_+^1(\mathbb{R}^d)$. Moreover, $f_t(\mathbf{z})$ is easily shown [2, 17, 18] to be a solution (either classical or weak, depending on the smoothness of $f_0(\mathbf{z})$) to the partial differential equation

$$\frac{\partial f_t}{\partial t} + \frac{\partial}{\partial \mathbf{z}} \cdot \left[\left(\mathbf{G}(\mathbf{z}) + \int_{\mathbb{R}^d} \mathbf{F}(\mathbf{z} - \mathbf{z}') f_t(\mathbf{z}') \mathrm{d}\mathbf{z}' \right) f_t(\mathbf{z}) \right] = 0. \tag{9}$$

For a system of N mechanical particles with pair potential ϕ (and $\mathbf{K} \equiv 0$) this is, of course, the Vlasov equation

$$\frac{\partial f_t}{\partial t} + \mathbf{p} \cdot \nabla_{\mathbf{q}} f_t - \left[\int_{\mathbb{R}^d} \nabla\phi(\mathbf{q} - \mathbf{q}') f_t(\mathbf{q}', \mathbf{p}') \mathrm{d}\mathbf{q}' \mathrm{d}\mathbf{p}' \right] \cdot \nabla_{\mathbf{p}} f_t = 0. \tag{10}$$

2.2 Propagation of Chaos and the Law of Large Numbers

It must be emphasized that the results cited so far are purely deterministic, and that the use of probability measures was just a matter of normalization of the initial densities without any probabilistic meaning. In particular, no statistical assumptions were made on the initial conditions, which were only required to approximate the desired infinite-particle limit. A truly probabilistic theory is obtained by assigning a statistical superposition of N-particle initial conditions and studying the time evolution of an ensemble probability density on the phase space $\mathbb{R}^{dN}$. The empirical densities μ_t^N in (2) are then regarded as $\mathcal{M}_1^+$-valued random variables on $\mathbb{R}^d$ endowed with the Borel σ-algebra and a permutation invariant probability measure $P_t^N(\mathrm{d}\mathbf{z}_1, \dots, \mathrm{d}\mathbf{z}_N)$. If the initial measure P_0^N is given, the probability measure at time t will be

$$P_t^N = P_0^N \circ \mathbf{T}_t^0, \tag{11}$$

where $\mathbf{T}_s^t$ denotes the flow on $\mathbb{R}^{dN}$ from time s to time t associated with the evolution equations for $(\mathbf{z}_1, \dots, \mathbf{z}_N)$, (1). Thus, the dynamics remains deterministic and the evolution of the probability density on $\mathbb{R}^{dN}$ is simply induced by the flow $\mathbf{T}_s^t$ acting on each "pure state." In order to study the $N \to \infty$ limit, of course, one needs to consider the complete sequence of symmetric probabilities $\{P_t^n\}$, $n = 1, 2, \dots$, which determines a (symmetric) probability measure P_t on $\Gamma \equiv \prod_{n>0} \mathbb{R}^{dN}$ (e.g. see [4, p. 255]).

The analysis of the convergence of the sequences μ_t^N is very simple in the light of Corollary 1, which is time-reversible and thus ensures a one-to-one correspondence between elements of the ensemble that converge at time zero and at time t. Thus, if one assumes that $\{\mu_0^n\}$ converges in probability, or almost surely, or even at every point in the probability space $(\Gamma, \mathcal{B}_0, P_0)$, the same will be true for $\{\mu_t^n\}$. What needs to be clarified is the relationship between the statistics for convergence of the empirical measures and the probability measures $P_t^N(\mathrm{d}\mathbf{z}_1, \dots \mathrm{d}\mathbf{z}_N)$. The idea behind the following definition goes back to Boltzmann but was given a precise mathematical expression by Kac [11], who called it the "Boltzmann property":

Definition 1 Given $\mu \in \mathcal{M}_1^+$, a sequence $\{P^n\}$ of symmetric probabilities is said to be μ-*chaotic* if, for all $k = 1, 2, \dots$ and $g_1, \dots, g_k \in \mathcal{C}_b(\mathbb{R}^d)$,

$$\lim_{N\to\infty} \int_{\mathbb{R}^{dk}} \prod_{i=i}^{k} g_i(\mathbf{z}_i) P^{k,N}(\mathrm{d}\mathbf{z}_1 \dots \mathrm{d}\mathbf{z}_k) = \prod_{i=i}^{k} \int_{\mathbb{R}^d} g_i(\mathbf{z})\mu(\mathrm{d}\mathbf{z}), \tag{12}$$

where $P^{k,N}$ denotes the k-th marginal of P^N.

Then, the crucial connection is given by the following

Theorem 2 (Sznitman-Tanaka) *A sequence* $\{P^n\}$ *is* μ*-chaotic* if and only if *the associated sequence of empirical measures converges to* μ, in probability.

Proof See [22], where it is shown that this holds even if the convergence is in law rather than in probability. See also [8, 9]. □

In other words, μ-chaoticity is equivalent to the validity *in probability* (and also in law) of the law of large numbers for the empirical measures. Since the latter propagates in time, so does the former, and if *either* property holds at $t = 0$, then *both* of them will hold at later times. In summary, we have:

Theorem 3 *Let μ_t be the solution to the weak Vlasov equation* (35), *with initial condition μ_0. Let the sequence $\{P_0^n\}$ be μ_0-chaotic* (*or, equivalently, let the sequence of empirical measures at $t = 0$ converge weakly to μ_0 in probability*). *Then the following statements hold*:

1. (*Law of large numbers*) *For all $g \in \mathcal{C}_b(\mathbb{R}^d)$ and $t > 0$, in probability*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=i}^{N} g(\mathbf{z}_i(t)) = \int_{\mathbb{R}^d} g(\mathbf{z})\mu_t(d\mathbf{z}). \tag{13}$$

2. (*Propagation of chaos*) *For all $k = 1, 2, \dots$ and $g_1, \dots, g_k \in \mathcal{C}_b(\mathbb{R}^d)$*

$$\lim_{N\to\infty} \int_{\mathbb{R}^{dk}} \prod_{i=i}^{k} g_i(\mathbf{z}_i) P_t^{k,N}(d\mathbf{z}_1 \dots d\mathbf{z}_k) = \prod_{i=i}^{k} \int_{\mathbb{R}^d} g_i(\mathbf{z})\mu_t(d\mathbf{z}). \tag{14}$$

Among μ_0-chaotic sequences, an important role is played by the product measures

$$\widetilde{P}_0^N(d\mathbf{z}_1 \dots d\mathbf{z}_N) = \bigotimes_{i=1}^{N} \mu_0(d\mathbf{z}_i). \tag{15}$$

In this case the $\mathbf{z}_i$ are initially i.i.d., and a classical probability result ([4], Theorem 11.4.1) ensures that the empirical distributions converge almost surely (not just in probability) in $(\Gamma, \mathcal{B}_0, \widetilde{P}_0)$. However, the opposite implication is not true. As a consequence, the law of large numbers for $t > 0$, (13), also holds almost surely, but in general the factorization in (15) does not propagate in time and the sequence $\{\widetilde{P}_t^n\}$ will be only μ_t-chaotic for $t > 0$.

3 Deviation/Fluctuation Theory

Having established the limit $\mu_t^N \to \mu_t$, one can seek information "at next order" by studying the convergence of quantities of the general form

$$\zeta_t^N = \frac{\mu_t^N - \mu_t}{\alpha_N}, \tag{16}$$

where $\{\alpha_n\}$ is a suitable numerical sequence such that convergence holds at $t = 0$. Like for the Vlasov limit itself, it turns out that the problem can be first studied for sequences of *individual* N-particle system. This will be done in the first subsection, where the ζ_t^N in (16) will be regarded as deterministic objects and called "deviations." The term "fluctuations" will be reserved for the probabilistic theory, to be developed in the second subsection, in which the ζ_t^N will be random variables. Even if the limit will be studied in the general case, the physically interesting situation is still when the μ_t^N are empirical measures for N particles and μ_t their limiting density, whereas α_N typically will be the familiar factor $1/\sqrt{N}$ that appears in the classical CLT for i.i.d. random variables.

3.1 Convergence of Deviations

The first decision in order to study the convergence of a sequence of deviations of the form in (16) is what topology to use. Since $\zeta_t^N \in \mathcal{M}_b$, the vector space of totally bounded signed measures, and $\mathcal{M}_b = \mathcal{C}_b(\mathbb{R}^d)^*$, it is natural to consider once again the topology of weak-$*$ convergence with respect to $\mathcal{C}_b(\mathbb{R}^d)$. However, when one takes a closer look at the time evolution of ζ_t^N it quickly becomes apparent that a simpler and more natural theory is obtained by using test functions that are not only continuous and bounded but also differentiable. Thus, we will study the convergence of ζ_t^N in the larger space $\mathcal{M}_b^1 \equiv \mathcal{C}_b^1(\mathbb{R}^d)^*$ with the weak-$*$ topology. In fact, it will be enough to work in the linear subspace $\mathcal{M}_{b,0}^1$ of functionals that satisfy the additional condition $\langle \zeta, 1\rangle = 0$, as the ζ_t^N do. Accordingly, we will identify the elements of $\mathcal{C}_b^1(\mathbb{R}^d)$ that differ by a constant; one way to do so is by restriction to the subsets $\mathcal{C}_{b,0}^1(\mathbb{R}^d)$ of functions such that $\langle \mu_0, g\rangle = 0$. Often we will stretch the notation by using the integral notation $\int \zeta(\mathrm{d}\mathbf{u}_0)g(\mathbf{u}_0)$ to indicate the action of a functional $\zeta \in \mathcal{M}_b^1$ on $g \in \mathcal{C}_b^1$. When $\zeta \in \mathcal{M}_b$, of course, $\langle \zeta, g\rangle$ is indeed an integral with respect to ζ (the difference of the integrals with respect to the Hahn-Jordan components). It should stressed once more that at this point the ζ_t^N are *not* random variables and the analysis is not probabilistic. All we are studying in this section is the *deterministic* evolution of the deviations of a sequence of *individual* measures μ_t^N from the limiting measure μ_t. With this in mind, here is the main result to be established in this subsection:

Theorem 4 *Let* $\mathbf{F} \in \mathcal{C}_b^2(\mathbb{R}^d)$, $\mathbf{G} \in \mathcal{C}^2(\mathbb{R}^d)$ *with bounded first and second derivatives* ($\mathbf{G}$ *itself need not be bounded*). *If at* $t = 0$ $\zeta_0^N \xrightarrow{w} \zeta_0 \in \mathcal{M}_{b,0}^1$, *then* $\zeta_t^N \xrightarrow{w} \zeta_t \in \mathcal{M}_{b,0}^1$ *for* $t \in [0, T]$, *weakly continuous in* t. *For all* $g \in \mathcal{C}_{b,0}^1(\mathbb{R}^d)$

$$\langle \zeta_t, g\rangle = \langle \zeta_0, \mathcal{T}_t(g)\rangle, \tag{17}$$

where $\mathcal{T}_t : \mathcal{C}_{b,0}^1(\mathbb{R}^d) \to \mathcal{C}_b^1(\mathbb{R}^d)$ *is the continuous linear operator given*[4] *by*

$$\mathcal{T}_t(g) = g(\mathbf{z}(t; \mathbf{w}_0, \mu_0)) + \int_{\mathbb{R}^d} \mathbf{k}(t; \mathbf{z}_0, \mathbf{w}_0, \mu_0) \cdot \nabla g(\mathbf{z}(t; \mathbf{z}_0, \mu_0))\mu_0(\mathrm{d}\mathbf{z}_0). \tag{18}$$

The integration kernel $\mathbf{k}(t; \mathbf{z}_0, \mathbf{w}_0, \mu_0)$ *is the unique solution to the* (*generalized*) *Braun-Hepp integral equations*,

$$\begin{aligned}\mathbf{k}(t; \mathbf{z}_0, \mathbf{w}_0, \mu_0) = &\int_0^t \mathrm{d}s \mathbf{F}(\mathbf{z}(s; \mathbf{z}_0, \mu_0) - \mathbf{z}(s; \mathbf{w}_0, \mu_0)) \\ &+ \int_0^t \mathrm{d}s \mathbf{k}(s; \mathbf{z}_0, \mathbf{w}_0, \mu_0) \cdot \nabla \mathbf{G}(\mathbf{z}(s; \mathbf{z}_0, \mu_0)) \\ &+ \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \mu_0(\mathrm{d}\mathbf{u}_0)(\mathbf{k}(s; \mathbf{z}_0, \mathbf{w}_0, \mu_0) - \mathbf{k}(s; \mathbf{u}_0, \mathbf{w}_0, \mu_0)) \\ &\times \nabla \mathbf{F}(\mathbf{z}(s; \mathbf{z}_0, \mu_0) - \mathbf{z}(s; \mathbf{u}_0, \mu_0));\end{aligned} \tag{19}$$

$\mathbf{k}(t; \mathbf{z}_0, \mathbf{w}_0, \mu_0)$ *is bounded and in* $\mathcal{C}^2([0, T])$ *with respect to* t, *and in* $\mathcal{C}_b^1(\mathbb{R}^d)$ *with respect to* $\mathbf{z}_0, \mathbf{w}_0$.

Equations (19) (for the Vlasov case) appear already in the work of Braun and Hepp [2], but in a rather different context, namely the calculation of the Gâteaux derivative of a characteristic along perturbations of μ_0 *for* N *fixed* (including $N = \infty$). Specifically, they showed

[4] Of course, just by subtracting $\langle \mu_0, \mathcal{T}_t(g)\rangle$ one can always modify $\mathcal{T}_t$ so that $\mathcal{T}_t(g) \in \mathcal{C}_{b,0}^1$.

that $\mathbf{k}(t; \mathbf{z}_0, \mathbf{w}_0, \mu_0)$ is related to $\mathbf{z}(t; \mathbf{z}_0, \mu_0)$ by the formula[5]

$$\int_{\mathbb{R}^d} \nu_0(\mathrm{d}\mathbf{w}_0)\mathbf{k}(t; \mathbf{z}_0, \mathbf{w}_0, \mu_0) = \lim_{s\to 0} \frac{1}{s}[\mathbf{z}(t; \mathbf{z}_0, \mu_0 + s\nu_0) - \mathbf{z}(t; \mathbf{z}_0, \mu_0)], \tag{20}$$

where ν_0 is a *positive* measure. Here, instead, we will prove that

$$\int_{\mathbb{R}^d} \zeta_0(\mathrm{d}\mathbf{w}_0)\mathbf{k}(t; \mathbf{z}_0, \mathbf{w}_0, \mu_0) = \lim_{N\to\infty} \frac{1}{\alpha_N}[\mathbf{z}(t; \mathbf{z}_0, \mu_0^N) - \mathbf{z}(t; \mathbf{z}_0, \mu_0)], \tag{21}$$

where $\zeta_0 \in \mathcal{M}^1_{b,0}$, which includes *signed* measures as well as more general functionals on $\mathcal{C}^1_{b,0}(\mathbb{R}^d)$. This is no longer a Gâteaux derivative, but rather an *ad hoc* functional derivation of $\mathbf{z}$ about μ_0 along the sequence of signed measures $\alpha_N\zeta_0^N = \mu_0^N - \mu_0$. Such reinterpretation of $\mathbf{k}(t; \mathbf{z}_0, \mathbf{w}_0, \mu_0)$ has important consequences. Since Braun and Hepp's functional differentiation of the trajectories does not involve N, they need to carry out the limit $N \to \infty$ later, during the proof of their central limit theorem. This requires some additional, fairly complicated asymptotic estimates. The same situation occurs in [21]. Here, on the contrary, the infinite-particle limit is already included in the deterministic analysis of how the characteristic curves depend on variations in the initial condition. This will make the proof of the central limit theorem much simpler.

To start, consider the Vlasov characteristic equations (6) in t-integral form, with the σ_t integration transformed backward to a σ_0 integration

$$\begin{aligned}\mathbf{z}(t; \mathbf{z}_0, \sigma_0) - \mathbf{z}_0 = &\int_0^t \mathrm{d}s\ \mathbf{G}(\mathbf{z}(s; \mathbf{z}_0, \sigma_0)) \\ &+ \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \sigma_0(\mathrm{d}\mathbf{u}_0)\mathbf{F}(\mathbf{z}(s; \mathbf{z}_0, \sigma_0) - \mathbf{z}(s; \mathbf{u}_0, \sigma_0)).\end{aligned} \tag{22}$$

If the two equations with σ_0 equal, respectively, to μ_0 and $\mu_0^N = \mu_0 + \alpha_N\zeta_0^N$ are subtracted, and if the operator

$$D_{\zeta_0^N} f(\mu_0) \equiv \frac{1}{\alpha_N}[f(\mu_0 + \alpha_N\zeta_0^N) - f(\mu_0)] \tag{23}$$

is introduced, one gets

$$\begin{aligned}D_{\zeta_0^N}\mathbf{z}(t; \mathbf{z}_0, \mu_0) = &\int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \zeta_0^N(\mathrm{d}\mathbf{u}_0)\mathbf{F}(\mathbf{z}(s; \mathbf{z}_0, \mu_0) - \mathbf{z}(s; \mathbf{u}_0, \mu_0)) \\ &+ \int_0^t \mathrm{d}s\, D_{\zeta_0^N}\mathbf{G}(\mathbf{z}(s; \mathbf{z}_0, \mu_0)) \\ &+ \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \mu_0^N(\mathrm{d}\mathbf{u}_0) D_{\zeta_0^N}\mathbf{F}(\mathbf{z}(s; \mathbf{z}_0, \mu_0) - \mathbf{z}(s; \mathbf{u}_0, \mu_0)).\end{aligned} \tag{24}$$

[5] See Braun and Hepp's equation (2.13), which contains a minor notational mistake: on the left-hand side t is used to indicate both time and the parameter in the limit. They are clearly meant to be two different symbols, since the right-hand side still depends on t.

In order to study the limit $N \to \infty$, let us define a new function $D_{\zeta_0}\mathbf{z}(t;\mathbf{z}_0,\mu_0)$ in terms of the linear integral equations

$$\begin{aligned} D_{\zeta_0}\mathbf{z}(t;\mathbf{z}_0,\mu_0) &= \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \zeta_0(\mathrm{d}\mathbf{u}_0)\mathbf{F}(\mathbf{z}(s;\mathbf{z}_0,\mu_0) - \mathbf{z}(s;\mathbf{u}_0,\mu_0)) \\ &\quad + \int_0^t \mathrm{d}s\, D_{\zeta_0}\mathbf{z}(s;\mathbf{z}_0,\mu_0)\cdot \nabla\mathbf{G}(\mathbf{z}(s;\mathbf{z}_0,\mu_0)) \\ &\quad + \int_0^t \mathrm{d}s \int_{\mathbb{R}^d} \mu_0(\mathrm{d}\mathbf{u}_0)[D_{\zeta_0}\mathbf{z}(s;\mathbf{z}_0,\mu_0) - D_{\zeta_0}\mathbf{z}(s;\mathbf{u}_0,\mu_0)] \\ &\quad \times \nabla\mathbf{F}(\mathbf{z}(s;\mathbf{z}_0,\mu_0) - \mathbf{z}(s;\mathbf{u}_0,\mu_0)). \end{aligned} \tag{25}$$

Under the given smoothness assumption on $\mathbf{F}$ and $\mathbf{G}$, the unique solution to these Volterra-Fredholm type equations for $D_{\zeta_0}\mathbf{z}(t;\mathbf{z}_0,\mu_0)$ is easily obtained in series form. An important feature is that the initial deviation ζ_0 appears only in the inhomogeneous term. Hence, the solution has the form

$$D_{\zeta_0}\mathbf{z}(t;\mathbf{z}_0,\mu_0) = \int_{\mathbb{R}^d} \zeta_0(\mathrm{d}\mathbf{w}_0)\mathbf{k}(t;\mathbf{z}_0,\mathbf{w}_0,\mu_0), \tag{26}$$

where $\mathbf{k}$ does not depend on ζ_0 and satisfies (19), whose (unique) solution is also easily obtained as a series and shown to have the desired regularity.

Since $D_{\zeta_0}\mathbf{z}(t;\mathbf{z}_0,\mu_0)$ is well-defined, the next step is to prove convergence:

Lemma 1 *Under the assumptions in Theorem 4*

$$\lim_{N\to\infty} D_{\zeta_0^N}\mathbf{z}(t;\mathbf{z}_0,\mu_0) = D_{\zeta_0}\mathbf{z}(t;\mathbf{z}_0,\mu_0) \tag{27}$$

pointwise in $\mathbf{z}_0 \in \mathbb{R}^d$, *uniformly in* $t \in [0,T]$. *Convergence holds uniformly in* $\mathbf{z}_0$ *as well if* ζ_0 *is a totally bounded signed measure.*

Proof Pointwise convergence with respect to $\mathbf{z}_0$ is easily obtained by subtracting (24) and (25), exploiting the regularity of $\mathbf{F}$, $\mathbf{G}$ and applying Gronwall's inequality. Proving uniform convergence requires some extra measure-theoretical work. A complete proof is presented in Appendix A.3. Note that pointwise convergence is all is needed in the proof of Theorem 4 here below. □

We are finally ready to prove Theorem 4:

Proof Consider the fixed-point equations for μ_t^N and μ_t ((5) with $\sigma_t = \mu_t^N, \mu_t$)

$$\mu_t^N = \mu_0^N \circ T_t^0[\mu^N], \qquad \mu_t = \mu_0 \circ T_t^0[\mu]. \tag{28}$$

Subtracting and rearranging gives

$$\frac{\mu_t^N - \mu_t}{\alpha_N} = \frac{\mu_0^N - \mu_0}{\alpha_N} \circ T_t^0[\mu^N] + \mu_0 \circ \frac{T_t^0[\mu^N] - T_t^0[\mu]}{\alpha_N} \tag{29}$$

that is

$$\zeta_t^N = \zeta_0^N \circ T_t^0[\mu^N] + \mu_0 \circ \frac{T_t^0[\mu^N] - T_t^0[\mu]}{\alpha_N}. \tag{30}$$

Now, multiply both sides by $g \in \mathcal{C}^1_{b,0}(\mathbb{R}^d)$ and integrate in $\mathbf{z}$. Changing the integration variables on the right-hand side to $\mathbf{z}_0$ gives

$$\langle \zeta_t^N, g\rangle = \langle \zeta_0^N, g \circ T_0^t[\mu^N]\rangle + \int_{\mathbb{R}^d} D_{\zeta_0^N} g(\mathbf{z}(t;\mathbf{z}_0,\mu_0)\mu_0(\mathrm{d}\mathbf{z}_0), \tag{31}$$

where $\langle \zeta_0^N, g \circ T_0^t[\mu^N]\rangle$ converges by hypothesis. The integrand also converges pointwise to $D_{\zeta_0}\mathbf{z}(t;\mathbf{z}_0,\mu_0,\zeta_0)\cdot\nabla g(\mathbf{z}(t;\mathbf{z}_0,\mu_0))$ by Lemma 1 and a standard chain rule argument. Moreover, since $g \in \mathcal{C}^1_{b,0}(\mathbb{R}^d)$

$$|D_{\zeta_0^N} g(\mathbf{z}(t;\mathbf{z}_0,\mu_0)| \le \|g\|_1 \|D^N_{\zeta_0}\mathbf{z}(t;\mathbf{z}_0,\mu_0,\zeta_0)\|_{\mathbb{R}^d}, \tag{32}$$

where $\sup_{\mathbf{z}_0} \|D^N_{\zeta_0}\mathbf{z}(t;\mathbf{z}_0,\mu_0,\zeta_0)\|_{\mathbb{R}^d}$ is finite by one more application of Gronwall's inequality (Appendix A.2). Thus, dominated convergence ensures that

$$\lim_{N\to\infty}\langle \zeta_t^N, g\rangle = \langle \zeta_0, g\circ T_0^t[\mu]\rangle + \int_{\mathbb{R}^d} D_{\zeta_0}\mathbf{z}(t;\mathbf{z}_0,\mu_0,\zeta_0)\cdot\nabla g(\mathbf{z}(t;\mathbf{z}_0,\mu_0))\mu_0(\mathrm{d}\mathbf{z}_0). \tag{33}$$

Substituting (26) this can be written in the form

$$\lim_{N\to\infty}\langle \zeta_t^N, g\rangle = \langle \zeta_0, \mathcal{T}_t(g)\rangle, \tag{34}$$

where $\mathcal{T}_t$ is the linear operator in (18). □

As was observed by Spohn [20] (in a probabilistic context), the convergence of $\langle \zeta_t^N, g\rangle$ allows one to write down an evolution equation for ζ_t. Note first that (9) can be generalized to a partial differential equation for measures (in weak form)

$$\frac{\partial}{\partial t}\langle \sigma_t, g\rangle - \left\langle \sigma_t, \left(\mathbf{G}(\mathbf{z}) + \int_{\mathbb{R}^d}\mathbf{F}(\mathbf{z}-\mathbf{z}')\sigma_t(\mathrm{d}\mathbf{z}')\right)\cdot\frac{\partial g}{\partial \mathbf{z}}\right\rangle = 0, \tag{35}$$

where now $g \in \mathcal{C}^1_b(\mathbb{R}^d)$. Since $\langle \sigma_t, g\rangle$ may still not be differentiable in time, its derivative must be understood in the usual weak sense [18]. With that in mind, subtracting the equations with $\sigma_t = \mu_t, \mu_t^N$ gives

$$\begin{aligned}\frac{\partial}{\partial t}\langle \zeta_t^N, g\rangle &- \left\langle \zeta_t^N, \mathbf{G}\cdot\frac{\partial g}{\partial \mathbf{z}}\right\rangle \\ &+ \left\langle \zeta_t^N\times\mu_t + \mu_t\times\zeta_t^N - \alpha_N\zeta_t^N\times\zeta_t^N, \mathbf{F}(\mathbf{z}-\mathbf{z}')\cdot\frac{\partial g}{\partial \mathbf{z}}\right\rangle = 0.\end{aligned} \tag{36}$$

Taking the limit $N\to\infty$ we have:

Corollary 2 *For all $g \in \mathcal{C}^1_{b,0}(\mathbb{R}^d)$, ζ_t given by (33) is the (unique) solution to the initial value problem for the linearized Vlasov equation about μ_t*

$$\frac{\partial}{\partial t}\langle \zeta_t, g\rangle - \left\langle \zeta_t, \mathbf{G}\cdot\frac{\partial g}{\partial \mathbf{z}}\right\rangle + \left\langle \zeta_t\times\mu_t + \mu_t\times\zeta_t, \mathbf{F}(\mathbf{z}-\mathbf{z}')\cdot\frac{\partial g}{\partial \mathbf{z}}\right\rangle = 0 \tag{37}$$

with initial condition $\zeta_0 \in \mathcal{M}^1_{b,0}$.

3.2 Convergence of Fluctuations

We now consider the fluctuations from the mean field, i.e. the deviations ζ_t^N in (16) regarded as random variables taking values in the topological vector space $\mathcal{M}_{b,0}^1(\mathbb{R}^d)$ with the weak-$*$ topology. In the previous subsection we proved that if any one realization of $\{\zeta_0^n\}$ converges to ζ_0, then the corresponding sequence $\{\zeta_t^n\}$ also converges to ζ_t in (17). This makes it fairly easy to prove convergence results for the random variables, if good convergence properties at $t=0$ are given. For example, assume that the initial sequence of laws $\{\mathcal{L}(\zeta_0^n)\}$ converges[6] in the usual weak sense. Because of Theorem 4 we know that for each n the map $F_n : \zeta_0^n \mapsto \zeta_t^n$ is well-defined. We also know that the sequence of the F_n's is "continuous," in the sense that if $\zeta_0^n \stackrel{w}{\to} \zeta_0$ for $\zeta_0 \in \mathcal{M}_{b,0}^1(\mathbb{R}^6)$ then $F_n(\zeta_0^n) \stackrel{w}{\to} F(\zeta_0)$, where F indicates the map $\zeta_0 \mapsto \zeta_t$ with ζ_t given by (17). In this framework, (weak) convergence of the laws $\{\mathcal{L}(\zeta_t^n)\}$ follows by the continuous mapping theorem[7] in its more general form that allows for multiple maps F_n.

The applicability of this type of result is limited because usually one does not want just to postulate the convergence properties of ζ_0^N, but rather deduce them from some assumptions on P_0 via a central limit theorem. In fact, given that the convergence $\mu_0^n \stackrel{w}{\to} \mu_0$ was understood as a law of large numbers associated with a μ_0-chaotic initial sequence of symmetric probabilities $\{P_0^n\}$, it is natural to try and prove a central limit theorem for the fluctuations $\{\zeta_0^n\}$. Unfortunately, it is usually very difficult to do so in terms of convergence (in law or otherwise) of random functionals, e.g. random signed measures with the weak-$*$ topology (although work has been done on sequences of signed measures with other, carefully chosen topologies, see [6]). The more common approach is to study the convergence in law of the *finite-dimensional distributions* of quantities like the ζ_0^N's, regarded as a random fields (or "empirical processes") on a suitable set of test functions. In our case this means integrating ζ_t^N against finite sets of parameter functions $g_1, g_2, \ldots, g_k$ in $\mathcal{C}_{b,0}^1(\mathbb{R}^d)$, and considering the joint probabilities for $\langle \zeta_t^N, g_1\rangle, \ldots, \langle \zeta_t^N, g_k\rangle$. Then, what one needs to prove is the following

Theorem 5 *Let* **F**, **G** *satisfy the same hypotheses as in Theorem* 4. *If the fluctuation field* $\zeta_t^N \in \mathcal{M}_{b,0}^1$ *converges in law at* $t=0$, *in the sense of the finite-dimensional distributions, then it will converge in the same fashion for* $t \in [0, T]$. *For every* $g \in \mathcal{C}_{b,0}^1$ *the limiting field* $\zeta_t \in \mathcal{M}_{b,0}^1$ *satisfies* (17), *and is the* (*unique*) *solution to the linearized Vlasov equation*, (37), *with initial condition* ζ_0.

This is similar to a theorem stated without proof by Spohn [20, Theorem 7.4], the main difference being that he asks for stronger regularity on the part of both the test functions and the interaction force. His requirements appear to be motivated by those in the CLT for test particle trajectories by Braun and Hepp [2]. Sznitman [21], in the Vlasov-McKean context, finds that the natural index set for the process is $\mathcal{C}_{b,0}^1$, as I do. Note that the linearized Vlasov equation is now understood as a stochastic PDE, although the evolution of each realization of the fluctuation field ζ_t is still completely deterministic.

[6] Of course, weak-$*$ convergence of laws $\{\mathcal{L}(\zeta_t^n)\}$ *on* $\mathcal{M}_{b,0}^1(\mathbb{R}^d)$ should not to be confused with weak-$*$ convergence *in* $\mathcal{M}_{b,0}^1(\mathbb{R}^d)$.

[7] See [12, Theorem 16.16] with the *caveat* that $\mathcal{M}_b^1$ is not a metric space but just a Hausdorff topological vector space. However, inspection of the proof suggests that the continuous mapping theorem generalizes easily if only the probability measures are τ-*smooth*. See [19] for the definition of τ-smoothness and for a closely related result.

Proof The following is an adaptation of the proof of the continuous mapping theorem [12, Theorem 4.27], to account for the fact that our assumptions do not guarantee[8] that the sequence of random measures ζ_0^n converges in law. Rather, we only have convergence in law of the random vector $(\langle\zeta_0^n, g_1\rangle, \dots, \langle\zeta_0^n, g_k\rangle) \in \mathbb{R}^k$, for any given choice of $g_j \in \mathcal{C}^1_{b,0}(\mathbb{R}^d)$, $j = 1, \dots, k$. We will use the symbol π_k to denote the projection $\pi_k\zeta = (\langle\zeta, g_1\rangle, \dots, \langle\zeta, g_k\rangle)$. Again based on the analysis of Sect. 3.1, we know that for each n the map $F_n : \zeta_0^n \mapsto \zeta_t^n$ is well-defined and that if $\zeta_0^n \xrightarrow{w} \zeta_0$, with $\zeta_0 \in \mathcal{M}^1_{b,0}(\mathbb{R}^6)$, then $F_n(\zeta_0^n) \xrightarrow{w} F(\zeta_0)$, with $F(\zeta_0) \equiv \zeta_t$ defined by (17) and satisfying (37). Let G be a fixed open set in $\mathbb{R}^k$ and let $s \in F^{-1} \circ \pi_k^{-1} G$; since $\pi_k \circ F_n(\zeta_0^n) \to \pi_k \circ F(\zeta_0)$ if $\zeta_0^n \xrightarrow{w} \zeta_0$, for any given neighborhood of s we can choose m large enough that $\pi_k \circ F_l$ maps the whole neighborhood into G for all $l \geq m$. Hence,

$$F^{-1} \circ \pi_k^{-1} G \subset \bigcup_m \left[\bigcap_{l\geq m} F_l^{-1} \circ \pi_k^{-1} G\right]^\circ. \tag{38}$$

By applying π_k to both sides of this inclusion it is then easy to see that

$$\pi_k \circ F^{-1} \circ \pi_k^{-1} G \subset \bigcup_m \left[\bigcap_{l\geq m} \pi_k \circ F_l^{-1} \circ \pi_k^{-1} G\right]^\circ. \tag{39}$$

Let P_0^n and P_0 be, respectively, the probability distributions of ζ_0^n and ζ_0. Then,

$$\begin{aligned} P_0(F^{-1} \circ \pi_k^{-1} G) &\leq P_0 \circ \pi_k^{-1} \bigcup_m \left[\bigcap_{l\geq m} \pi_k \circ F_l^{-1} \circ \pi_k^{-1} G\right]^\circ \\ &= \sup_m P_0 \circ \pi_k^{-1} \left[\bigcap_{l\geq m} \pi_k \circ F_l^{-1} \circ \pi_k^{-1} G\right]^\circ. \end{aligned} \tag{40}$$

By hypothesis $\mathcal{L}(\pi_k\zeta_0^n) \to \mathcal{L}(\pi_k\zeta_0)$ as $n \to \infty$, i.e. $P_0^n \circ \pi_k^{-1} \xrightarrow{w} P_0 \circ \pi_k^{-1}$. Using one of the Portmanteau characterizations of convergence in law, the right-hand side of the last equation is dominated by

$$\sup_m \liminf_{n\to\infty} P_0^n \circ \pi_k^{-1} \left[\bigcap_{l\geq m} \pi_k \circ F_l^{-1} \circ \pi_k^{-1} G\right]^\circ; \tag{41}$$

$\liminf_{n\to\infty}$ does not change if we only consider $n \geq m$, and then $\sup_m \bigcap_{l\geq m}$ must be less or equal to the value that occurs when $l = m = n$, which is

$$\liminf_{n\to\infty} P_0^n \circ \pi_k^{-1} [\pi_k \circ F_n^{-1} \circ \pi_k^{-1} G]^\circ \leq \liminf_{n\to\infty} P_0^n (F_n^{-1} \circ \pi_k^{-1} G). \tag{42}$$

Finally, since the left-most term in (40) is less or equal to the right-most term in (42), it follows by another application of the Portmanteau theorem that

$$P_0^n \circ F_n^{-1} \circ \pi_k^{-1} \xrightarrow{\mathcal{L}} P_0 \circ F^{-1} \circ \pi_k^{-1} \tag{43}$$

which is equivalent to say that $\mathcal{L}(\pi_k\zeta_t^n) \to \mathcal{L}(\pi_k\zeta_t)$. □

[8] We remark that for (positive) random measures, convergence in law of the finite dimensional distributions does imply convergence in law of the measures, see [12], Theorem 16.16. However, as far as I know, no similar theorem is available for elements of $\mathcal{M}^1_{b,0}$ or even just for signed measures. For work in this direction see [10].

The simplest example to which this general result applies is that of factorized initial probability measures like those considered by Sznitman [21]. In this case, it follows immediately from the standard central limit theorem for $\mathbb{R}^k$-valued random vectors that as $N \to \infty$ the random field ζ_0^N (with $\alpha_N = 1/\sqrt{N}$) converges in law, in the sense of the finite-dimensional distributions with respect to the parameter space $\mathcal{C}^1_{b,0}(\mathbb{R}^d)$, to a Gaussian random field ζ_0 with mean zero and covariance

$$\mathrm{E}(\langle \zeta_0, g_l\rangle\langle \zeta_0, g_m\rangle) = \langle \mu_0, g_l g_m\rangle - \langle \mu_0, g_l\rangle\langle \mu_0, g_m\rangle. \tag{44}$$

Then, Sznitman's CLT (for the "pure Vlasov" case, without diffusion terms) is recovered from Theorem 5 without any further effort.

Theorem 6 *Let* **F**, **G** *satisfy the same hypotheses as in Theorem* 4, *let the initial sequence of symmetric probability measures be* $\{\widetilde{P}_0^n\}$ *of the form in* (15), *and let* $\alpha_N = 1/\sqrt{N}$. *Then, as* $N \to \infty$ *the random field* ζ_t^N *converges in law, in the sense of the finite-dimensional distributions, to a Gaussian random field* ζ_t *for* $t \in [0, T]$; *the field* ζ_t *has mean zero and covariance*

$$\mathrm{E}(\langle \zeta_t, g_l\rangle\langle \zeta_t, g_m\rangle) = \langle \mu_0, \mathcal{T}_t(g_l)\mathcal{T}_t(g_m)\rangle - \langle \mu_0, \mathcal{T}_t(g_l)\rangle\langle \mu_0, \mathcal{T}_t(g_m)\rangle, \tag{45}$$

where $\mathcal{T}_t$ *is the operator defined in* (18).

Proof The limiting field ζ_t has zero mean, like ζ_0, and the covariance is obtained from (17) and (44). □

4 Strongly μ-Chaotic Sequences

Clearly the hypothesis that the initial probability measures factorize for all N is a very strong assumption. It is not easily justified on physical grounds, especially because we cannot expect it to propagate in time. Thus, we cannot argue that an initially factorized probability measure is the result of some previous evolution of the system, since in general interactions between particles create statistical correlations. We encountered a somewhat analogous situation when reviewing the law of large numbers, which is known to be equivalent to a weaker form of statistical independence, called μ-chaoticity, which does propagate in time and so determines a more natural class of initial probability measures for the mean-field dynamics. One wonders if anything similar holds true for the central limit theorem, but the answer is only partially in the affirmative. In this last section we show that there is indeed a stronger μ-chaoticity-type condition on the N-particle probability measures that is both necessary and sufficient for the validity of a CLT in which the limiting field is Gaussian with mean zero and covariance of the form in (44). In other words, we are going to prove a "higher-order" version of the Sznitman-Tanaka Theorem, Theorem 2. If one assumes "strong" μ-chaoticity at $t = 0$, our general propagation result, Theorem 5, ensures the validity of a similar CLT for $t > 0$. However, the sequence of ensemble probability measures at $t > 0$ turns out to be μ-chaotic but not strongly μ-chaotic, because the covariance, which is given again by (45), is not of the required form.

From now on let $\alpha_N = 1/\sqrt{N}$. For a given $\mu \in \mathcal{M}_1^+$, let $\{P^n\}$ be a μ-chaotic sequence of symmetric probability measures for the random vectors $(\mathbf{z}_1, \ldots, \mathbf{z}_N) \in \mathbb{R}^{dN}$, and let $\{\widetilde{P}^n\}$ be the sequence of product measures $\widetilde{P}^N = \bigotimes_{i=1}^N \mu(\mathrm{d}\tilde{\mathbf{z}}_i)$, with $(\tilde{\mathbf{z}}_1, \ldots, \tilde{\mathbf{z}}_N)$ the corresponding random vectors. Let μ^N, ζ^N and $\tilde{\mu}^N$, $\tilde{\zeta}^N$ be the empirical measures and the fluctuations associated, respectively, with P^N and $\widetilde{P}^N$. Let $\mathcal{P}^N \equiv P^N \times \widetilde{P}^N$ be the joint probability (law) of $(\mathbf{z}_1, \ldots, \mathbf{z}_N, \tilde{\mathbf{z}}_1, \ldots, \tilde{\mathbf{z}}_N)$, and $\mathcal{P} \equiv P \times \widetilde{P}$ the probability measure for infinite sequences determined by the hierarchy $\{\mathcal{P}^n\}$. Expectations with respect to $\mathcal{P}^N$ and $\mathcal{P}$ will be denoted by $\mathcal{E}_N$ and $\mathcal{E}$. For $g, h \in \mathcal{C}_b(\mathbb{R}^d)$ let

$$C_\mu[g,h] \stackrel{\text{def}}{=} \langle \mu, gh\rangle - \langle \mu, g\rangle\langle \mu, h\rangle. \tag{46}$$

Of course, this is just the bilinear form that appears in the expression for the covariance of the limiting Gaussian law in the central limit theorem for empirical densities associated to i.i.d. random variables, see (44).

Definition 2 A μ-chaotic sequence $\{P^n\}$ of symmetric probabilities is said to be *strongly μ-chaotic* if, for $k = 1, 2, \ldots$ and $g_1, \ldots, g_k \in \mathcal{C}_b(\mathbb{R}^d)$, it satisfies

$$\lim_{N\to\infty} N^{\frac{k}{2}} \int_{\mathbb{R}^{dk}} \prod_{i=1}^{k} (g_i(\mathbf{z}_i) - \langle \mu, g_i\rangle) P^{k,N}(\mathrm{d}\mathbf{z}_1 \ldots \mathrm{d}\mathbf{z}_k) = 0. \tag{47}$$

To put this in context, note that μ-chaoticity by itself is equivalent to

$$\lim_{N\to\infty} \int_{\mathbb{R}^{dk}} \prod_{i=1}^{k} (g_i(\mathbf{z}_i) - \langle \mu, g_i\rangle) P^{k,N}(\mathrm{d}\mathbf{z}_1 \ldots \mathrm{d}\mathbf{z}_k) = 0. \tag{48}$$

Equation (47) simply requires a higher rate of convergence for the same limit. Obviously, sequences of factorized probability densities are strongly μ-chaotic.

Lemma 2 *A μ-chaotic sequence of symmetric probability measures $\{P^n\}$ is strongly μ-chaotic if and only if, for $k = 1, 2, \ldots$ and $g_1, \ldots, g_k \in \mathcal{C}_b(\mathbb{R}^d)$,*

$$\lim_{N\to\infty} \mathcal{E}_N\left[\prod_{i=1}^{k} (\langle \zeta^N, g_i\rangle - \langle \tilde{\zeta}^N, g_i\rangle)\right] = \Theta_k, \tag{49}$$

where

$$\Theta_k = \begin{cases} 0 & \text{for } k \text{ odd} \\ \sum_{\mathcal{A}\in\Lambda_k} \prod_{\{i,j\}\in\mathbb{P}_\mathcal{A}} 2C_\mu[g_i, g_j] & \text{for } k \text{ even}; \end{cases} \tag{50}$$

here Λ_k indicates the set of all allocations $\mathcal{A}$ of the set $\{1, 2, \ldots, k\}$ into $k/2$ unordered pairs $\{i, j\}$, and $\mathbb{P}_\mathcal{A} = \mathrm{Range}(\mathcal{A})$.

Proof If $\phi_j(\mathbf{z}_i, \tilde{\mathbf{z}}_i) \equiv g_j(\mathbf{z}_i) - g_j(\tilde{\mathbf{z}}_i)$, $j = 1, \ldots, k$, the left-hand side in (49) can be written in the form

$$\int_{\mathbb{R}^{dN}} \int_{\mathbb{R}^{dN}} \frac{1}{N^{k/2}} \prod_{j=1}^{k} \sum_{i=1}^{N} \phi_j(\mathbf{z}_i, \tilde{\mathbf{z}}_i) \mathcal{P}^N(\mathrm{d}\mathbf{z}_1 \ldots \mathrm{d}\mathbf{z}_N \mathrm{d}\tilde{\mathbf{z}}_1 \ldots \mathrm{d}\tilde{\mathbf{z}}_N). \tag{51}$$

Expanding the product turns the integrand into a sum of products of the form

$$\sum_{i_1=1}^{N} \cdots \sum_{i_k=1}^{N} \phi_1(\mathbf{z}_{i_1}, \tilde{\mathbf{z}}_{i_1}) \ldots \phi_k(\mathbf{z}_{i_k}, \tilde{\mathbf{z}}_{i_k}). \tag{52}$$

By permutation symmetry, all the index vectors $(i_1 \dots i_k)$ that are related by permutations give identical contributions to the expectation value. Let $\mathbb{I}^k$ indicate all such equivalence classes of index vectors with respect to the symmetric group. A representative element for each class can be selected, for instance, by stipulating that $i_1 = 1$ and then moving from left to right and attributing the next integer value to any index whose value never appeared before. Elementary combinatorics shows that the number of elements in each class $\mathcal{I}$ is given by

$$\frac{N!}{(N - \mathrm{Rk}(\mathcal{I}))!} \prod_{\ell=1}^{\mathrm{Rk}(\mathcal{I})} n_\ell !,$$

where $\mathrm{Rk}(\mathcal{I})$ is the number of non-identical entries in each index vector belonging to the class and n_ℓ is the number of factors that depend on $(\mathbf{z}_\ell, \tilde{\mathbf{z}}_\ell)$. The notation reflects the fact that if the elements of $\mathcal{I}$ are viewed as a maps from $(1, \dots, k)$ to $(1, \dots, N)$ then $\mathrm{Rk}(\mathcal{I})$ is the dimension of their range. In order to study the asymptotics as $N \to \infty$, it is convenient to distinguish the subsets $\mathbb{I}^k_m \equiv \{\mathcal{I} \in \mathbb{I}^k : \mathrm{Rk}(\mathcal{I}) = m\}$. Then, (51) becomes

$$\sum_{m=1}^{k} \sum_{\mathcal{I}\in\mathbb{I}^k_m} \frac{1}{N^{k/2}} \frac{N!}{(N-m)!} \prod_{\ell=1}^{m} n_\ell ! \mathcal{E}_{m,N}[\phi_1(\mathbf{z}_{i_1}, \tilde{\mathbf{z}}_{i_1}) \dots \phi_k(\mathbf{z}_{i_k}, \tilde{\mathbf{z}}_{i_k})], \tag{53}$$

$\mathcal{E}_{m,N}$ being the expectation with respect to $\mathcal{P}^{m,N} \equiv P^{m,N} \times \widetilde{P}^{m,N}$.

Now, we are ready to prove our statement. Assume first that (49) holds, so that also (53) goes to Θ_k; μ-chaoticity will be proven by induction. For $k = 1$, $\Theta_1 = 0$, and (53) gives

$$\lim_{N\to\infty} \sqrt{N} \mathcal{E}_{1,N}[\phi_1(\mathbf{z}_1, \tilde{\mathbf{z}}_1)] = 0. \tag{54}$$

so that the μ-chaoticity condition is satisfied. Now, assume that (47) has been verified up to $k - 1$, and take the limit as $N \to \infty$ in (53); note that the factors that multiply $\mathcal{E}_{m,N}$ in (53) are of order $N^{m-k/2}$.

- The dominant term occurs for $m = k$, when $(i_1 i_2 \dots i_k) = (12 \dots k)$; this leading term is asymptotically equal to

$$N^{\frac{k}{2}} \mathrm{E}_{k,N}\left[\prod_{j=1}^{k} (g_i(\mathbf{z}_i) - \langle \mu, g_i \rangle)\right] \tag{55}$$

  which is the quantity inside the limit in (47).
- Clearly, all terms with $m < k/2$ vanish as $N \to \infty$.
- Let us consider the terms with $k > m > k/2$. We want to use the induction hypothesis of lower-order strong μ-chaoticity. Among the functions $\phi_1, \dots, \phi_k$ at least $m - (k - m) = 2m - k$ do not share their independent variable with anybody else. If we assign to each of these a factor $\sqrt{N}$, we "use" the existing factor $N^{m-k/2}$ completely (we may even "create" some power of $1/\sqrt{N}$ if more that $2m - k$ ϕ_j's do not share independent variables). The remaining ϕ_j's can be organized in products of those that share the same independent variables, say $(\mathbf{z}_i, \tilde{\mathbf{z}}_i)$. Take the expectation of each product with respect to the $\mu(\mathrm{d}\tilde{\mathbf{z}}_i)$ factor from $\widetilde{P}_N$ to produce a new test function $h_i(\mathbf{z}_i)$. Add and subtract the μ average of h_i with respect to $\mathbf{z}_i$: $h_i(\mathbf{z}_i) = (h_i(\mathbf{z}_i) - \langle \mu, h_i \rangle) + \langle \mu, h_i \rangle$. Both types of factors $h_i(\mathbf{z}_i) - \langle \mu, h \rangle$ and $\langle \mu, h \rangle$ allow one to apply strong μ-chaoticity after the product of the various "blocks" of ϕ_j's is expanded, and it is easily seen that everything vanishes—the dominant term being the one where the product of all the constants $\langle \mu, h_i \rangle$ appears. If k is

odd, we have established strong μ-chaoticity, since (51) vanishes by (49) and it has been shown to have the same limit as (55).

– If k is even, we still have to consider the term with $m = k/2$ in (53). If any one among the ϕ_j's does not share its independent variable with anybody else, again strong μ-chaoticity at order lower than k implies that the corresponding term vanishes. Hence, one only needs consider the equivalence classes in $\mathbb{I}^k_{k/2}$ where each independent variable appears in exactly *two* ϕ_j's, i.e. $n_\ell = 2$ for $\ell = 1, 2, \ldots, k/2$; it is easily seen that there are $(k-1)!! \equiv 1 \times 3 \times 5 \times \cdots \times (k-1)$ such classes. Again regarding the $\widetilde{P}_N$ expectation of products $\phi_i \phi_j$ of pairs of functions with the same variable as new test functions, adding and subtracting the μ-averages and using standard μ-chaoticity shows that as $N \to \infty$ all is left is

$$2^{\frac{k}{2}} \sum_{\mathcal{A} \in \Lambda_k} \prod_{\{i,j\} \in \mathbb{P}_{\mathcal{A}}} \langle \mu, (g_i - \langle \mu, g_i \rangle)(g_j - \langle \mu, g_j \rangle) \rangle = \Theta_k \tag{56}$$

Combining the last equation with (55) and (49) yields strong μ-chaoticity also for k even.

Conversely, starting from strong μ-chaoticity the same calculation shows that the quantity in (51) goes to Θ_k as $N \to \infty$, which proves (49). □

It should be mentioned that if one applies a similar argument to (51) and (53) *without* the factors $1/N^{k/2}$, one recovers the Sznitman-Tanaka theorem (by proving that $\mu^N - \tilde{\mu}^N \to 0$ in law if and only if the sequence $\{P^n\}$ is μ-chaotic.) By now, observant readers will have recognized the quantities Θ_k in (50) as the k-th order central moments of a Gaussian random vector. By well-known theorems of probability theory, this observation leads to:

Lemma 3 *A μ-chaotic sequence of symmetric probability measures $\{P^n\}$ is strongly μ-chaotic if and only if, as $N \to \infty$, $\zeta^N - \tilde{\zeta}^N$ converges in law, in the sense of the finite dimensional distributions, to a Gaussian process with mean zero and covariance $2C_\mu[g_i, g_j]$.*

It is now easy to see that strong μ-chaoticity is both necessary and sufficient for the validity of the "classical" central limit theorem with a Gaussian limit

Theorem 7 *Let $\{P^n\}$ be a sequence of μ-chaotic symmetric probability measures. The fluctuations ζ^N converge (in law, in the sense of the finite dimensional distributions) to a Gaussian field with mean zero and covariance $C_\mu[g_i, g_j]$, $g_i, g_j \in \mathcal{C}_b(\mathbb{R}^d)$, if and only if $\{P^n\}$ is strongly μ-chaotic.*

Proof For $g_1, \ldots, g_k \in \mathcal{C}_b(\mathbb{R}^d)$ consider

$$\mathbf{Z}_N = (\langle \zeta^N, g_1 \rangle, \ldots, \langle \zeta^N, g_k \rangle),$$

$$\widetilde{\mathbf{Z}}_N = (\langle \tilde{\zeta}^N, g_1 \rangle, \ldots, \langle \tilde{\zeta}^N, g_k \rangle),$$

$$\mathbf{D}_N = \mathbf{Z}_N - \widetilde{\mathbf{Z}}_N$$

which we regard as $\mathbb{R}^k$-valued random vectors on the probability space $\mathbb{R}^{dN} \times \mathbb{R}^{dN}$ with the Borel sets and the product probability measure $\mathcal{P}^N$. We know from the classical central limit theorem for empirical measures for i.i.d. random variables that $\widetilde{\mathbf{Z}}_N$ converges in law to a Gaussian random vector $\widetilde{\mathbf{Z}}_\infty$ with mean zero and covariance matrix $C_\mu[g_i, g_j]$. We also know from Theorem 3 that strong μ-chaoticity is equivalent to convergence in law of $\mathbf{D}_N$ to a Gaussian random vector with mean zero and covariance matrix $2C_\mu[g_i, g_j]$. Hence, strong μ-chaoticity implies that also $\mathbf{Z}_N$ converges in law to $\mathbf{Z}_\infty$, and vice versa convergence of $\mathbf{Z}_N$ ensures convergence of $\mathbf{D}_N$ to $\mathbf{D}_\infty$, with $\mathbf{Z}_\infty - \widetilde{\mathbf{Z}}_\infty = \mathbf{D}_\infty$. In order to identify the limits $\mathbf{Z}_\infty$, $\mathbf{D}_\infty$, observe that $\mathcal{L}(-\widetilde{\mathbf{Z}}_\infty) = \mathcal{L}(\widetilde{\mathbf{Z}}_\infty)$ and that $\mathbf{Z}_\infty$ and $\widetilde{\mathbf{Z}}_\infty$ are independent. Hence, standard theorems give the relation among characteristic functions

$$\mathcal{E}[e^{i\mathbf{Z}_\infty\cdot\mathbf{W}}]\mathcal{E}[e^{i\widetilde{\mathbf{Z}}_\infty\cdot\mathbf{W}}] = \mathcal{E}[e^{i\mathbf{D}_\infty\cdot\mathbf{W}}] \tag{57}$$

from which it follows easily that if two of the limiting vectors are Gaussian so is the third one, with the desired covariance matrices. □

Both this theorem and the original Sznitman-Tanaka one, Theorem 2, hold true if the index functions are taken from $\mathcal{C}^1_{b,0}(\mathbb{R}^d)$ instead of $\mathcal{C}_b(\mathbb{R}^d)$. In particular, the definitions of μ-chaoticity and strong μ-chaoticity can be modified by restricting the set of index functions to $\mathcal{C}^1_{0,b}(\mathbb{R}^d)$. With that in mind, Theorems 3 and 5 apply to strongly μ-chaotic initial conditions:

Theorem 8 *Let μ_t be the solution to (5), with initial condition μ_0. Let the sequence $\{P^n_0\}$ be strongly μ_0-chaotic. Then, for $t \in [0, T]$*

1. (*Central Limit Theorem*) *The fluctuations ζ^N_t converge (in law, in the sense of the finite dimensional distributions on $\mathcal{C}^1_{b,0}$) to the Gaussian field with mean zero and covariance $C_{\mu_0}[\mathcal{T}_t(g_i), \mathcal{T}_t(g_j)]$, $g_i, g_j \in \mathcal{C}^1_{b,0}(\mathbb{R}^d)$.*
2. (*Propagation of chaos*) *The sequence $\{P^n_t\}$ is μ-chaotic.*

Since in general $C_{\mu_0}[\mathcal{T}_t(g_i), \mathcal{T}_t(g_j)] \neq C_{\mu_t}[g_i, g_j]$, *strong* μ-chaoticity *does not* propagate under the Vlasov evolution. One way to look at this situation is that the covariance matrix at $t > 0$ of the Vlasov fluctuations reflects higher-order statistical correlations between particles that “survive” the infinite particle limit. One wonders if strong μ-chaoticity does propagate for more complicated “collisional” models—such as the Landau or Balescu-Guernsey-Lenard equations of classical plasma physics—which may be able to wipe out statistical correlations more efficiently than mean-field dynamics. This is, of course, a very difficult question to be left for future studies.

Acknowledgements I would like to thank Michael Kiessling for many helpful discussions.

Appendix A: Applications of Gronwall’s Inequality

Gronwall’s inequality plays an important role in the mathematical study of the mean field limit, and fluctuation theory is no exception. Here we only need the inequality in its simplest form, which says that if a continuously differentiable function $f : [0, T] \to \mathbb{R}$ satisfies

$$\frac{\mathrm{d}f}{\mathrm{d}t} \le A + Bf, \tag{58}$$

where A and B are constants, then

$$f(t) \le f(0)e^{Bt} + A\frac{e^{Bt} - 1}{B}. \tag{59}$$

In all cases of interest here $f(0)=0$. To simplify the notation, everywhere in this appendix the symbol $\|\cdot\|$ indicates the Euclidean norm in $\mathbb{R}^d$. We shall use repeatedly that fact that for any $f\in\mathcal{C}^1([0,T];\mathbb{R}^d)$

$$\frac{\mathrm{d}}{\mathrm{d}t}\|f(t)\|\le\left\|\frac{\mathrm{d}f}{\mathrm{d}t}\right\| \tag{60}$$

which follows easily from the triangle inequality (it is, in fact, true for functions valued in any Banach space, not just $\mathbb{R}^d$). Also, when we talk about uniform convergence as $N\to\infty$, it will be understood that we are concerned about uniformity with respect to the initial point $\mathbf{z}_0\in\mathbb{R}^d$. Uniformity with respect to t will always be trivially true as long as t lies in the finite interval $[0,T]$.

A.1 Uniform Convergence of Characteristic Curves

We use the shorthand

$$\Delta_N f(\mu_0)\equiv f(\mu_0^N)-f(\mu_0). \tag{61}$$

Subtracting the two Vlasov characteristic equations, (6), with $\sigma_t=\mu_t,\mu_t^N$ and transforming the integrations back to μ_0,μ_0^N yields

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Delta_N\mathbf{z}(t;\mathbf{z}_0,\mu_0)=&\int_{\mathbb{R}^d}(\mu_0^N-\mu_0)(\mathrm{d}\mathbf{u}_0)\mathbf{F}(\mathbf{z}(t;\mathbf{z}_0,\mu_0)-\mathbf{z}(t;\mathbf{u}_0,\mu_0))\\&+\Delta_N\mathbf{G}(\mathbf{z}(t;\mathbf{z}_0,\mu_0))\\&+\int_{\mathbb{R}^d}\mu_N(\mathrm{d}\mathbf{u}_0)\Delta_N\mathbf{F}(\mathbf{z}(t;\mathbf{z}_0,\mu_0)-\mathbf{z}(t;\mathbf{u}_0,\mu_0)).\end{aligned} \tag{62}$$

Let $L_{\mathbf{F}}$, $L_{\mathbf{G}}$ be the Lipschitz constants of $\mathbf{F}$ and $\mathbf{G}$, respectively, and $B_{\mathbf{F}}$ the upper bound on $\|\mathbf{F}\|$. By a simple trick (see [18, p. 68]) the first term on the right-hand side can be bounded in terms of the BL distance between μ_0^N and μ_0. Using the Lipshitz continuity of $\mathbf{F}$ and $\mathbf{G}$ and the triangle inequality

$$\frac{\mathrm{d}}{\mathrm{d}t}\sup_{\mathbf{z}_0\in\mathbb{R}^d}\|\Delta_N\mathbf{z}(t;\mathbf{z}_0,\mu_0)\|\le 2B_{\mathbf{F}}L_{\mathbf{F}}d(\mu_0^N,\mu_0)+L\sup_{\mathbf{z}_0\in\mathbb{R}^d}\|\Delta_N\mathbf{z}(t;\mathbf{z}_0,\mu_0)\|, \tag{63}$$

where $L\equiv 2L_{\mathbf{F}}+L_{\mathbf{G}}$. Then, Gronwall's inequality gives

$$\sup_{\mathbf{z}_0\in\mathbb{R}^d}\|\Delta_N\mathbf{z}(t;\mathbf{z}_0,\mu_0)\|\le 2B_{\mathbf{F}}L_{\mathbf{F}}\frac{e^{Lt}-1}{L}d(\mu_0^N,\mu_0) \tag{64}$$

which implies (8).

A.2 Uniform upper Bound on $D_{\zeta_0^N}\mathbf{z}(t;\mathbf{z}_0,\mu_0)$

If one divides (62) by α_N and exploits like before the Lipschitz continuity of $\mathbf{F}$ and $\mathbf{G}$ one gets

$$\frac{\mathrm{d}}{\mathrm{d}t}\|D_{\zeta_0^N}\mathbf{z}(t;\mathbf{z}_0,\mu_0)\| \leq \left\|\int_{\mathbb{R}^d}\zeta_0^N(\mathrm{d}\mathbf{u}_0)\mathbf{F}(\mathbf{z}(t;\mathbf{z}_0,\mu_0)-\mathbf{z}(t;\mathbf{u}_0,\mu_0))\right\| + L\|D_{\zeta_0^N}\mathbf{z}(t;\mathbf{z}_0,\mu_0)\|. \tag{65}$$

Since the sequence $\{\zeta_0^n\}$ converges in the weak-$*$ topology of $\mathcal{M}_b^1$, it is norm-bounded. Thus the first term on the right has a uniform upper bound M, obtained by multiplying the bound on the norm of the ζ_N times the $\mathcal{C}_b^1$ norms of the components of $\mathbf{F}$ and taking the Euclidean norm:

$$\frac{\mathrm{d}}{\mathrm{d}t}\sup_{\mathbf{z}_0\in\mathbb{R}^d}\|D_{\zeta_0^N}\mathbf{z}(t;\mathbf{z}_0,\mu_0)\| \leq M + L\sup_{\mathbf{z}_0\in\mathbb{R}^d}\|D_{\zeta_0^N}\mathbf{z}(t;\mathbf{z}_0,\mu_0)\|. \tag{66}$$

Hence, Gronwall's inequality gives

$$\sup_{\mathbf{z}_0\in\mathbb{R}^d}\|D_{\zeta_0^N}\mathbf{z}(t;\mathbf{z}_0,\mu_0)\| \leq MB\frac{e^{Lt}-1}{L}. \tag{67}$$

A.3 Proof of Lemma 1

In this case the quantity that needs to be majorized is

$$\eta_N(t;\mathbf{z}_0,\mu_0,\zeta_0) \equiv \|D_{\zeta_0^N}\mathbf{z}(t;\mathbf{z}_0,\mu_0) - D_{\zeta_0}\mathbf{z}(t;\mathbf{z}_0,\mu_0)\|, \tag{68}$$

which is uniformly bounded for $\mathbf{z}_0\in\mathbb{R}^d$ and $t\in[0,T]$ (since both $D_{\zeta_0^N}$ and D_{ζ_0} are). Using the multivariate Taylor formula (with remainder) gives

$$\begin{aligned}&\|D_{\zeta_0^N}\mathbf{G}(\mathbf{z}(t;\mathbf{z}_0,\mu_0)) - D_{\zeta_0}\mathbf{z}(t;\mathbf{z}_0,\mu_0)\cdot\nabla\mathbf{G}(\mathbf{z}(t;\mathbf{z}_0,\mu_0))\|\\ &\quad\leq L_{\mathbf{G}}\eta_N(t;\mathbf{z}_0,\mu_0,\zeta_0) + C_{\mathbf{G}}\|D_{\zeta_0^N}\mathbf{z}(t;\mathbf{z}_0,\mu_0)\|\|\Delta_N\mathbf{z}(t;\mathbf{z}_0,\mu_0)\|,\end{aligned} \tag{69}$$

where now $L_{\mathbf{G}} \equiv \sup_{\mathbb{R}^d}\|\nabla\mathbf{G}\|$ and $C_{\mathbf{G}}$ is a constant whose value reflects the upper bounds on the second derivatives of $\mathbf{G}$. In the same manner

$$\begin{aligned}&\|D_{\zeta_0^N}\mathbf{F}(\mathbf{z}(t;\mathbf{z}_0,\mu_0)-\mathbf{z}(t;\mathbf{u}_0,\mu_0))\\ &\qquad-[D_{\zeta_0}\mathbf{z}(t;\mathbf{z}_0,\mu_0)-D_{\zeta_0}\mathbf{z}(t;\mathbf{u}_0,\mu_0)]\cdot\nabla\mathbf{F}(\mathbf{z}(t;\mathbf{z}_0,\mu_0)-\mathbf{z}(t;\mathbf{u}_0,\mu_0))\|\\ &\quad\leq L_{\mathbf{F}}[\eta_N(t;\mathbf{z}_0,\mu_0,\zeta_0)+\eta_N(t;\mathbf{u}_0,\mu_0,\zeta_0)]\\ &\qquad+C_{\mathbf{F}}\|D_{\zeta_0^N}\mathbf{z}(t;\mathbf{z}_0,\mu_0)-D_{\zeta_0^N}\mathbf{z}(t;\mathbf{u}_0,\mu_0)\|\|\Delta_N\mathbf{z}(t;\mathbf{z}_0,\mu_0)-\Delta_N\mathbf{z}(t;\mathbf{u}_0,\mu_0)\|.\end{aligned} \tag{70}$$

Using these bounds in the equation obtained by differentiating and subtracting (24) and (25) leads to

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\eta_N(t;\mathbf{z}_0,\mu_0,\zeta_0) \leq{}& A_N(t;\mathbf{z}_0,\mu_0,\zeta_0)+(L_{\mathbf{f}}+L_{\mathbf{G}})\eta_N(t;\mathbf{z}_0,\mu_0,\zeta_0)\\ &+L_{\mathbf{f}}\int_{\mathbb{R}^d}\mu_0(\mathrm{d}\mathbf{u}_0)\eta_N(t;\mathbf{u}_0,\mu_0,\zeta_0),\end{aligned} \tag{71}$$

where

$$
\begin{aligned}
A_N(t;\mathbf{z}_0,\mu_0,\zeta_0) = & \left\| \int_{\mathbb{R}^d} (\zeta_0^N - \zeta_0)(\mathrm{d}\mathbf{u}_0)\mathbf{F}(\mathbf{z}(t;\mathbf{z}_0,\mu_0) - \mathbf{z}(t;\mathbf{u}_0,\mu_0)) \right\| \\
& + C_{\mathbf{G}} \| D_{\zeta_0^N}\mathbf{z}(t;\mathbf{z}_0,\mu_0)\| \|\Delta_N \mathbf{z}(t;\mathbf{z}_0,\mu_0)\| \\
& + C_{\mathbf{F}} \int_{\mathbb{R}^d} \mu_0(\mathrm{d}\mathbf{u}_0) \| D_{\zeta_0^N}\mathbf{z}(t;\mathbf{z}_0,\mu_0) - D_{\zeta_0^N}\mathbf{z}(t;\mathbf{u}_0,\mu_0)\| \\
& \times \|\Delta_N \mathbf{z}(t;\mathbf{z}_0,\mu_0) - \Delta_N \mathbf{z}(t;\mathbf{u}_0,\mu_0)\|. \qquad (72)
\end{aligned}
$$

Clearly, $A_N(t;\mathbf{z}_0,\mu_0,\zeta_0)$ goes to zero as $N\to\infty$ for each $\mathbf{z}_0$. It follows that $\eta_N(t;\mathbf{z}_0,\mu_0,\zeta_0)\to 0$ pointwise via a double application of Gronwall's inequality, first to (71) integrated with respect to $\mu_0(\mathrm{d}\mathbf{z}_0)$,

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d}\mu_0(\mathrm{d}\mathbf{u}_0)\eta_N(t;\mathbf{u}_0,\mu_0,\zeta_0) \le & \int_{\mathbb{R}^d}\mu_0(\mathrm{d}\mathbf{u}_0)A_N(t;\mathbf{u}_0,\mu_0,\zeta_0) \\
& + L\int_{\mathbb{R}^d}\mu_0(\mathrm{d}\mathbf{u}_0)\eta_N(t;\mathbf{u}_0,\mu_0,\zeta_0). \qquad (73)
\end{aligned}
$$

Dominated convergence ensures that

$$
\lim_{N\to\infty}\int_{\mathbb{R}^d}\mu_0(\mathrm{d}\mathbf{u}_0)A_N(t;\mathbf{u}_0,\mu_0,\zeta_0) = 0 \qquad (74)
$$

and then (73) implies that also $\int_{\mathbb{R}^d}\mu_0(\mathrm{d}\mathbf{u}_0)\eta_N(t;\mathbf{u}_0,\mu_0,\zeta_0)\to 0$. Finally, Gronwall's inequality must be applied again to (71) itself in order to bound $\eta_N(t;\mathbf{z}_0,\mu_0,\zeta_0)$ in terms of $A_N(t;\mathbf{z}_0,\mu_0,\zeta_0) + L_{\mathbf{F}}\int_{\mathbb{R}^d}\mu_0(\mathrm{d}\mathbf{u}_0)\eta_N(t;\mathbf{u}_0,\mu_0,\zeta_0)$.

While pointwise convergence is all is needed in the proof of Theorem 4, it is interesting that whenever $\zeta_0\in\mathcal{M}_b$ one can show that convergence is in fact uniform with respect to $\mathbf{z}_0$. Equation (71) implies

$$
\frac{\mathrm{d}}{\mathrm{d}t}\sup_{\mathbf{z}_0\in\mathbb{R}^d}\eta_N(t;\mathbf{z}_0,\mu_0,\zeta_0) \le \sup_{\mathbf{z}_0\in\mathbb{R}^d}A_N(t;\mathbf{z}_0,\mu_0,\zeta_0) + L\sup_{\mathbf{z}_0\in\mathbb{R}^d}\eta_N(t;\mathbf{z}_0,\mu_0,\zeta_0) \qquad (75)
$$

and uniform convergence will follow by Gronwall's inequality if only one can establish that $\sup_{\mathbf{z}_0\in\mathbb{R}^d}A_N(t;\mathbf{z}_0,\mu_0,\zeta_0)\to 0$. From (64) and (67) it is easy to see that the terms on the second and third lines in (72) are dominated by

$$
2M(C_{\mathbf{G}}+4C_{\mathbf{F}})B_{\mathbf{F}}^2 L_{\mathbf{F}}\left[\frac{e^{Lt}-1}{L}\right]^2 d(\mu_0^N,\mu_0) \qquad (76)
$$

which does not depend on $\mathbf{z}_0$. The main difficulty lies with the integral with respect to $\zeta_0^N-\zeta_0$ in (72). One would like to have an uniform estimate in terms of some "distance" between ζ_0^N and ζ_0, like it was done for the inhomogeneous term in (62), but this idea cannot be immediately implemented for signed measures since the weak-$*$ topology in $\mathcal{M}_b$ is not metrizable. Note, however, that the definition of the BL distance for positive measures, (3), extends without any trouble also to signed measures. Clearly, the condition $\lim_{N\to\infty}d(\zeta_N,\zeta)=0$ implies that $\int f\mathrm{d}\zeta_N\to\int f\mathrm{d}\zeta$ for all f that are bounded and uniformly Lipschitz, but for signed measures this is not enough to establish weak-$*$ convergence (see [1, § 8.3.1]). However, the opposite implication is still valid:

Lemma 4 *If* $\zeta_N \stackrel{w}{\to} \zeta$ *in* $\mathcal{M}_b$, *then* $\lim_{N\to\infty} d(\zeta_N, \zeta) = 0$.

Proof Since $\mathbb{R}^d$ with the Euclidean topology is a complete metric space and $\mathcal{Q} = \{\zeta_N\} \cup \zeta$ is weakly-$*$ compact, general theorems of measure theory (see [7, Theorems 4 and 5] and also [1, Sec. 8.6]) ensure that $\mathcal{Q}$ is norm-bounded and *uniformly tight*, meaning that for each $\varepsilon > 0$ there is a compact subset $K \subset \mathbb{R}^d$ such that

$$\sup_{\xi\in\mathcal{Q}}\{|\xi|(\mathbb{R}^d/K)\} < \varepsilon. \tag{77}$$

Then, the same type of argument as in [4, Theorem 11.3.3] applies. Consider the set B of functions such that $\|f\|_{BL} \leq 1$. Their restrictions to K form a compact subset of $\mathcal{C}_b(K)$ by the Arzelà-Ascoli Theorem. Hence, one can extract a finite family $f_1, \dots, f_k$ of functions in B such that for any $f \in B$ and some $j \leq k$

$$\sup_{\mathbf{y}\in K} |f(\mathbf{y}) - f_j(\mathbf{y})| < \varepsilon. \tag{78}$$

We write:

$$\left|\int_{\mathbb{R}^d} f \mathrm{d}(\zeta_n - \zeta)\right| \leq \int_{\mathbb{R}^d} |f - f_j| \mathrm{d}(|\zeta_N| + |\zeta|) + \left|\int_{\mathbb{R}^d} f_j \mathrm{d}(\zeta_N - \zeta)\right|. \tag{79}$$

Once the first term on the right-hand side is broken into two integrals on K and $\mathbb{R}^d/K$, it is easily majorized (independently of f) by a constant times ε thanks to (77) and (78) (recall that $\mathcal{Q}$ is bounded in norm). The second term is smaller than ε if N is large enough, regardless of f. The conclusion follows. □

At last, since

$$\left\|\int_{\mathbb{R}^d} (\zeta_0^N - \zeta_0)(\mathbf{u}_0)\mathbf{F}(\mathbf{z}(t;\mathbf{z}_0,\mu_0) - \mathbf{z}(t;\mathbf{u}_0,\mu_0))\right\| \leq 2B_{\mathbf{F}}L_{\mathbf{F}}d(\zeta_0^N, \zeta_0) \tag{80}$$

and $\zeta_0^N \stackrel{w}{\to} \zeta_0$, the Lemma ensures that the left-hand side goes to zero uniformly with respect to $\mathbf{z}_0$, and we can conclude that

$$\lim_{N\to\infty} \sup_{\mathbf{z}_0\in\mathbb{R}^d} \eta_N(t;\mathbf{z}_0,\mu_0,\zeta_0) = 0. \tag{81}$$

References

1. Bogachev, V.I.: Measure Theory. Springer, New York (2006)
2. Braun, W., Hepp, K.: The Vlasov dynamics and its fluctuations in the $1/N$ limit of interacting classical particles. Commun. Math. Phys. **56**, 101–113 (1977)
3. Dobrushin, R.L.: Vlasov equations. Funct. Anal. Appl. **13**, 115–123 (1979)
4. Dudley, R.M.: Real Analysis and Probability. Cambridge University Press, Cambridge (2002)
5. Elskens, Y., Kiessling, M.K.-H., Ricci, V.: The Vlasov limit for a system of particles which interact with a wave field. Commun. Math. Phys. **285**, 673–712 (2009)
6. Fernandez, B., Méléard, S.: A Hilbertian approach for fluctuations on the McKean-Vlasov model. Stoch. Proc. Appl. **71**, 33–53 (1997)
7. Fremlin, D.H., Garling, D.J.H., Haydon, R.G.: Bounded measures on topological spaces. Proc. Lond. Math. Soc. **3**, 115–136 (1972)

8. Golse, F.: The mean-field limit for the dynamics of large particle systems. In: Journées Équations aux dérivées partielles GDR 2434 CNRS (2003)
9. Gottlieb, A.: Markov transitions and the propagation of chaos. Lawrence Berkeley National Laboratory Report LBNL-42839 (1998)
10. Jacob, P., Oliveira, P.E.: A representation of infinitely divisible signed random measures. Port. Math. **52**, 211–219 (1995)
11. Kac, M.: Foundations of kinetic theory. In: Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, pp. 171–197. University of California Press, Berkeley (1956)
12. Kallenberg, O.: Foundations of Modern Probability. Springer, New York (2001)
13. Kiessling, M.K.-H.: Microscopic derivations of Vlasov equations. Commun. Nonlinear Sci. Numer. Simul. **13**, 106–113 (2008)
14. Maslov, V.P.: Self-consistent field equations. In: Contemporary Problems in Mathematics. vol. 11, pp. 153–234. VINITI, Moscow (1978)
15. McKean, H.P. Jr.: A class of Markov processes associated with nonlinear parabolic equations. Proc. Natl. Acad. Sci. **56**, 1907–1911 (1966)
16. McKean, H.P. Jr.: Propagation of chaos for a class of nonlinear parabolic equations. In: Lecture Series in Differential Equations, vol. 7, pp. 41–57. Catholic University of America, Washington (1967)
17. Neunzert, H.: Mathematical investigations on particle-in-cell methods. Fluid Dyn. Trans. **9**, 229–254 (1978)
18. Neunzert, H.: An introduction to the nonlinear Boltzmann-Vlasov equation. In: Cercignani, C. (ed.) Kinetic Theories and the Boltzmann Equation. Lecture Notes in Mathematics, vol. 1048, pp. 60–110. Springer, New York (1984)
19. Ressel, P.: A topological version of Slutsky’s theorem. Proc. Am. Math. Soc. **85**, 272–274 (1982)
20. Spohn, H.: Large Scale Dynamics of Interacting Particles. Springer, New York (1991)
21. Sznitman, A.S.: A fluctuation result for nonlinear diffusions. In: Albeverio, S. (ed.) Infinite Dimensional Analysis and Stochastic Processes. Research Notes in Mathematics, vol. 124, pp. 145–160. Pitman, Boston (1985)
22. Sznitman, A.S.: Topics in propagation of chaos. In: Ecole d’été de Probabilités de Saint-Flour XIX 1989. Lecture Notes in Mathematics, vol. 1464, pp. 165–251. Springer, New York (1991)