Kinetic Theory and Renormalization Group Methods for Time Dependent Stochastic Systems

Matteo Marcozzi

Academic dissertation presented for the degree of Doctor of Philosophy

To be presented for public examination with the permission of the Faculty of Science of the University of Helsinki in Auditorium A111 in the Kumpula Campus (Gustaf Hällströmin katu 2b, Helsinki) on 30th September 2016 at 12 o'clock.

Department of Mathematics and Statistics
Faculty of Science
University of Helsinki
2016
Acknowledgements

First and foremost, I would like to thank my supervisors Antti Kupiainen and Jani Lukkarinen. The first for giving me the opportunity to come to Helsinki for my doctoral programme and for his extremely valuable insights into mathematical physics and beyond. The second for guiding me through my PhD with infinite patience and helpfulness and for largely enriching my scientific tool box.

Secondly, I would like to thank my pre-examiners Professor Cédric Bernardin and Doctor Giada Basile and my opponent Professor Thierry Bodineau, for carefully reading my thesis and permitting me to defend it in a tight schedule. I would like to express my gratitude also to Doctor Alessia Nota for our fruitful collaboration during the year she spent at the University of Helsinki as a postdoctoral researcher.

I am also indebted to the Centre of Excellence in Analysis and Dynamics Research and to the Doctoral School of Mathematics and Statistics of the University of Helsinki for providing me with financial support during my studies.

Last but not least, I wish to express my deep gratitude to my family and friends in Italy and in Finland, who have been encouraging and supporting me during this long scientific marathon.

Saint-Pierre-Quiberon, 31.08.2016

Matteo Marcozzi
Contents

Contents

1 Included Articles
  1.1 Contribution of the Author ................................. 5

2 Prologue and interrelations of [I]-[V] .......................... 6

3 Wick polynomials and cumulants as dynamical variables in statistical mechanics .... 8
  3.1 Cumulants vs moments ......................................... 8
  3.2 Wick polynomials and cumulants .............................. 9
  3.3 Possible applications in kinetic theory ........................ 11

4 Velocity flip model: thermalization and kinetic theory ....................... 13
  4.1 The model and the physical picture ........................... 13
  4.2 Kinetic theory predictions for the thermalization of the velocity flip model .... 14

5 Kinetic equations for the Lorentz model with an external magnetic field .......... 17
  5.1 Lorentz model and scaling limits ............................ 17
  5.2 Effect of an external magnetic field on the two-dimensional Lorentz gas ........ 19

6 Renormalization group for SPDEs and nonlinear fluctuating hydrodynamic ....... 23
  6.1 An overview about KPZ equation ............................. 23
  6.2 Fluctuating hydrodynamics for anharmonic particle chains ......................... 25
  6.3 Renormalization of the coupled KPZ ................................ 27

References .................................................................. 33

7 Articles [I]-[V] .......................................................... 36
1 Included Articles

The dissertation consists of this introduction and the following five articles. The aim of the introductory part is to give a clear overview of the content of the papers with minimal emphasis on technicalities. The articles are referenced in the text by Roman numerals [I]-[V]


[IV] M. Marcozzi, A. Nota, Derivation of the linear Landau equation and of linear Boltzmann equation from the Lorentz model with magnetic field, Journal of Statistical Physics, 162(6), 2016².


1.1 Contribution of the Author

I have been carrying out most of the computations and the technical details in [I], [III] and [V], while for paper [II] and [IV] my collaborators and I equally shared the amount of work.

¹The final submitted manuscript of articles [I] is reprinted by the kind permission of AIP.
²Article [IV] is reprinted by the kind permission of Springer.
2 Prologue and interrelations of [I]-[V]

The purpose of the present thesis is to study some examples of time dependent stochastic systems of physical relevance. By this terminology we indicate two classes of models which are characterized by the different role of randomness:

1. deterministic evolution with random initial data;

2. truly stochastic evolution, namely driven by some sort of random force, with either deterministic or random initial data.

Before presenting some instances of these two scenarios, let us try to explain why such stochastic systems are so extensively investigated in the mathematical physics and theoretical physics literature. Suppose we want to study the dynamics of a physical phenomenon by looking at some kind of deterministic (discretized) differential equation describing the time evolution of a certain observable. To be more concrete the reader could think about the energy density in the context of the heat conduction in solid crystals or the growth process of a surface which separates two different physical phases of some material.

It turns out that in such complex systems a deterministic equation is not enough to capture all the salient features and, consequently, to obtain sensible predictions to be contrasted with the experiments. In some of these cases an affordable way to circumvent this impasse is to introduce some randomness in the initial conditions or in the time evolution equation.

In fact, in many circumstances it is impossible to have a complete control of the initial state of the system, therefore it is natural to consider reasonable random initial data in order to handle this uncertainty. Moreover, sometimes the time evolution is driven by some complicated mechanism for which we do not want to build a detailed description, but it suffices to model it in the evolution equation by a suitable stochastic force designed to retain just few important characteristics of our setting.

In other words, the time dependent stochastic systems are effective models for physical phenomena where the stochasticity takes into account some features whose analytic control is unattainable and/or unnecessary.

As an example of the setting (1) in this thesis we will deal with the discrete nonlinear Schrödinger equation (DNLS) with random initial data. This model is widely used in a variety of fields like nonlinear optics, nonlinear acoustics and quantum condensates, but here we will mainly focus on its applications concerning the study of transport coefficients in lattice systems. Since the seminal work by Green and Kubo [26,39] in the mid 50’s, when they discovered that transport coefficients for simple fluids can be obtained through a time integral over the respective total current correlation function, the mathematical physics community has been trying to rigorously validate these predictions and extend them also to solids.

In particular, the main technical difficulty is to obtain at least a reliable asymptotic form of the time behaviour of the Green-Kubo correlation. To do this, one of the possible approaches is kinetic theory, a branch of the modern mathematical physics stemmed from the challenge of deriving the classical laws of thermodynamics from microscopic systems. The paradigmatic case is the gas of free particles in a box, where the particles travel with constant velocity, i.e. ballistically, until they collide with the box walls.

Generalizing the simple case of the free gas, nowadays kinetic theory deals with models whose dynamics is transport dominated in the sense that typically the solutions to the kinetic equations, whose prototype is the Boltzmann equation, correspond to ballistic motion intercepted by collisions whose frequency is order one on the kinetic space-time scale.
In [I] and [II] DNLS equation is considered as a microscopic model from which a kinetic equation arises in a suitable limit, namely the Boltzmann-Peierls equation. This was introduced by Peierls [49] to understand the energy transport in solid insulators: in particular, his idea was to describe the crystalline lattice vibrations as a gas of interacting phonons whose distribution function satisfies a Boltzmann type equation. The main goal of [I] and [II] is to build some technical tools, namely Wick polynomials and their connection with cumulants, to pave the way towards the rigorous derivation of kinetic equations from the DNLS model.

The paper [III] can be contextualized in the same framework of kinetic predictions for transport coefficients. In particular, we consider the velocity flip model which belongs to the family (2) of our previous classification, since it consists of a particle chain with harmonic interaction and a stochastic term which flips the velocity of the particles. This model was introduced in [21] and it is one of the simplest one-dimensional model with finite conductivity. In [III] we perform a detailed study of the position-momentum correlation matrix via two different methods (one of them being kinetic theory) and we get an explicit formula for the thermal conductivity.

Moreover, in [IV] we consider the Lorentz model [41] perturbed by an external magnetic field which can be categorized in the class (1): it is a gas of non interacting particles colliding with obstacles located at random positions in the plane, therefore, once the configuration of obstacles is fixed, the evolution is deterministic. Here we show that under a suitable scaling limit the system is described by a kinetic equation where the magnetic field affects only the transport term, but not the collisions.

Finally, in [IV] we studied a generalization of the famous Kardar-Parisi-Zhang (KPZ) equation which falls into the category (2) being a nonlinear stochastic partial differential equation driven by a space-time white noise. The KPZ equation [35] was originally proposed as a model for surface growth, but lately it has turned out to describe a broad class of phenomena. For instance, Spohn [57] has recently introduced a generalized vector valued KPZ equation in the framework of nonlinear fluctuating hydrodynamics for anharmonic particle chains, a research field which is again strictly connected to the investigation of transport coefficients.

The problem with the KPZ equation is that it is ill-posed, as originally formulated in [35]. However, in 2013 Hairer [29], combining the rough path theory with ideas from quantum field theory, succeeded to give a rigorous mathematical meaning to the solution of the KPZ via an approximation scheme involving the renormalization of the nonlinear term by a formally infinite constant.

In [V] we tackle a vector valued generalization of the KPZ and we prove local in time well-posedness. We employ the method recently formulated by Kupiainen [40] to treat singular stochastic PDEs: it is inspired by the so-called Wilsonian Renormalization Group approach [62] which was extensively used in the 70’s in the study of critical exponents in statistical mechanics and later on also in many other fields, from quantum field theory [23] and KAM theory [10].

The following sections are dedicated to providing the necessary background to contextualize the articles [I]-[V] and to reviewing the main results.
3 Wick polynomials and cumulants as dynamical variables in statistical mechanics

3.1 Cumulants vs moments

Suppose that we want to study a certain physical system which is modelled by a random field $\psi_x$ where $x \in X$ is some discrete or continuum index encoding the variables on which the field might depend. In order to get some information on the features of such a system, a natural starting point is to look at the moments (or simple correlation in the jargon of statistical physics):

$$\mathbb{E} \left[ \prod_{i=1}^{n} \psi_{x_i} \right]$$

(3.1)

where $\mathbb{E}[\cdot]$ denotes the expectation with respect to the probability measure associated to $\psi_x$. To be more concrete, we take $X = \mathbb{Z}^d$, so that $\psi_x$ is a lattice random field in $d$ dimensions and we set $i = 2$ in (3.1). We note that if $\psi_{x_1}, \psi_{x_2}$ happen to be independent, then $\mathbb{E}[\psi_{x_1}\psi_{x_2}] = \mathbb{E}[\psi_{x_1}][\psi_{x_2}]$. To emphasize this property one could consider the quantity

$$\kappa(x_1, x_2) := \mathbb{E}[\psi_{x_1}\psi_{x_2}] - \mathbb{E}[\psi_{x_1}][\psi_{x_2}],$$

(3.2)

which is defined such that it vanishes when the two random fields are independent. Despite the fact that mathematically this might seem just a trivial rearrangement of terms, from the physical point of view $\kappa(x_1, x_2)$ might capture quite easily some features which are not evident by looking at moments. For example, a large family of interesting physical systems are “asymptotically independent”, i.e. $\psi_{x_1}$ and $\psi_{x_2}$ are almost independent when $|x_1 - x_2|$ is large, thus this property can be readily related to the smallness of $\kappa(x_1, x_2)$, while $\mathbb{E}[\psi_{x_1}\psi_{x_2}]$ is not very informative in this respect because in general it is not small even when $|x_1 - x_2| \rightarrow \infty$.

$\kappa(x_1, x_2)$ is called second order cumulant or two-points connected correlation of the fields $\psi$ and it turns out to be an ubiquitous tool in statistical physics along with its generalization to higher orders.

In order to study more systematically the properties of cumulants and their relation with moments, let us fix some notation: consider a collection $y_j, j \in J$ where $J$ is some fixed nonempty index set, of real or complex random variables on some probability space $(\Omega, \mathcal{A}, \mu)$. Then for any sequence of indices, $I = (i_1, i_2, \ldots, i_n) \in J^n$, we denote monomials of the above random variables by

$$y^I = y_{i_1}y_{i_2}\cdots y_{i_n} = \prod_{k=1}^{n} y_{i_k}, \quad y^\emptyset := 1 \quad \text{if} \quad I = \emptyset.$$  

(3.3)

We consider sequences of indices instead of sets of indices, so that we do not need to relabel elements which are repeated in a sequence, but we will stick to a set-like notation for subsequences and partitions of sequences.

For any $I$ we denote the corresponding moment by $\mathbb{E}[y^I]$ and the related cumulant by

$$\kappa[y_I] = \kappa_{\mu}[y_I] = \kappa[y_{i_1}, y_{i_2}, \cdots, y_{i_n}].$$

(3.4)

If $\mathbb{E}[|y_E|^2] < \infty$ for all $E \subset I$, they can be defined recursively as

$$\kappa[y_I] = \mathbb{E}[y^I] - \sum_{E : x \in E \subset I} \mathbb{E}[y^{I \setminus E}]\kappa[y_E]$$

(3.5)
where $x$ is some element of $I$. Furthermore, cumulants are multilinear and permutation invariant and if the random variables $y_j$, $j = 1, 2, \ldots, n$, have joint exponential moments, i.e. if there exists $\beta > 0$ such that $\mathbb{E}[e^{\beta \sum_j |y_j|}] < \infty$, then moments and cumulants can be computed as

$$\mathbb{E}[y^I] = \partial^I G_m(0), \quad \kappa[y^I] = \partial^I g_c(0),$$

(3.6)

where $G_m(\lambda)$ and $g_c(\lambda)$ are generating functions defined by

$$G_m(\lambda) := \mathbb{E}[e^{\lambda \cdot x}], \quad g_c(\lambda) := \ln G_m(\lambda)$$

(3.7)

and “$\partial^I$” denotes $\partial_{\lambda_1} \partial_{\lambda_2} \cdots \partial_{\lambda_n}$. Note that (3.5) can be inverted yielding the famous “moments-to-cumulants” formula

$$\mathbb{E}[y^I] = \sum_{\pi \in \mathcal{P}(I)} \prod_{A \in \pi} \kappa[y_A],$$

(3.8)

where $\mathcal{P}(I)$ denotes the collection of partitions of $I$. When the random variables $y$ are time dependent and they describe a physical field, the above discussion about the decay property of cumulants indicates that it is interesting to look at their time evolution, namely trying to write an equation for $\partial_t \kappa[y_E]$ with $E \subset I$. One would like to achieve this goal without resorting to the moments, since the moments-to-cumulants formula, even though very useful in some situations, it is a nonlinear relation with a nontrivial combinatorial structure: one way out is to consider Wick polynomials (WP), which are polynomials of random variables with nice combinatorial properties.

In the next section we will summarize the content of [I] and [II]: the main features of WP are reviewed and their strict relation with cumulants is investigated. In particular, WP turn out to be a valuable tool to study the decay property of cumulants since they allow to prove a nontrivial summability result for cumulants of random lattice fields.

### 3.2 Wick polynomials and cumulants

Wick polynomials, also called Wick products, appeared in the 50’s in quantum field theory as a tool to systematically replace monomial of field operators with state dependent coefficients in order to cancel singular terms in perturbation theory [61]. In the seminal work by Wick and in the ensuing physics literature WP are defined with respect to a Gaussian measure, but the construction can be extended to any measure for which moments exist.

Within the same setting and notation of the former section, we denote the WP corresponding to the moment $\mathbb{E}[y^I]$ by

$$:y^I: = :y^I;\mu = :y_{i_1}y_{i_2}\cdots y_{i_n}: .$$

(3.9)

When $\mathbb{E}[|y^E|^2] < \infty$ for all $E \subset I$, then $:y^I:$ is defined recursively as

$$:y^I: = y^I - \sum_{E \subset I} \mathbb{E}[y^I|E] :y^E: .$$

(3.10)

If the random variables $y_j$, $j = 1, 2, \ldots, n$ have joint exponential moments, then we also have

$$:y^I: = \partial^I \lambda G_w(0; y)$$

(3.11)

where

$$G_w(\lambda; y) := \frac{e^{\lambda \cdot y}}{\mathbb{E}[e^{\lambda \cdot x}]} = e^{\lambda \cdot y - g_\lambda(y)}$$
is the WP generating function. The main combinatorial property of WP is that expectations of their products can be expanded in terms of cumulants, yielding a formula with the same structure as the standard moments-to-cumulants expansion. In fact, in [I] the following result is proven:

**Proposition 3.1.** Assume that the measure \( \mu \) has all moments of order \( N \), i.e., suppose that \( \mathbb{E}[|y'|] < \infty \) for all \( I \in \mathcal{I} \) with \( |I| \leq N \). Suppose \( L \geq 1 \) is given and consider a collection of \( L + 1 \) index sequences \( J', J_\ell \in \mathcal{I} \), \( \ell = 1, \ldots, L \), such that \( |J'| + \sum \ell |J_\ell| \leq N \). Then for \( I := \sum_{\ell=1}^L J_\ell + J' \) with (the set of its labels in \( I \)) we have

\[
\mathbb{E}\left[ \prod_{\ell=1}^L :y^{J_\ell}: :y^{J'}: \right] = \sum_{\pi \in \mathcal{P}(I)} \prod_{A \in \pi} (\kappa[y_A] \mathbb{I}(A \not\subset J_\ell \forall \ell)) \, .
\] (3.12)

In words, the constraint determined by the characteristic functions on the right hand side of (3.12) amounts to removing from the standard cumulant expansion all terms which have any clusters internal to one of the sets \( J_\ell \).

Proposition 3.1 implies a simple representation of cumulants in terms of WP, in fact we have

\[
\kappa[y(t)_{I+(j)}] = \mathbb{E}[y(t)^{I_1} : y_j(t):] = \mathbb{E}[y(t)^{I_2}: y_j(t)]
\] (3.13)

from which one gets

\[
\partial_t \kappa[y(t)_I] = \sum_{i \in I} \mathbb{E}[\partial_i y_j(t) : y(t)^{I\setminus{i}}].
\] (3.14)

To have a more explicit form for (3.14), we consider a deterministic evolution of the fields driven by a polynomial potential with random initial data: for \( j \) belonging to the index set \( J \) we assume

\[
\partial_t y_j(t) = \sum_{I \in \mathcal{F}_j} M^t_I(t) : y(t)^{I;}. \tag{3.15}
\]

where the functions \( M^t_I(t) \) are “interaction amplitudes” from the \( I \)-th Wick polynomial of \( y(t) \) to \( y_j(t) \) and the set \( \mathcal{F}_j \) collects those \( I \) which have a nonzero amplitude for each \( j \in J \). Even though the evolution is deterministic, \( y_j(t) \) are still random variables because they inherit the randomness of the initial data \( y_i(0) \). In this setting (3.14) becomes

\[
\partial_t \kappa[y(t)_I] = \sum_{i \in I} \sum_{I' \in \mathcal{F}_j} M^t_I(t) \mathbb{E}[y(t)^{I_1} : y(t)^{I'\setminus{i}}]. \tag{3.16}
\]

Since the expectation of the right hand side of (3.16) involves a product of WP, we can apply 3.1 and expand it in cumulants, so that we end up with an evolution equation for cumulants without resorting to moments, as we wanted.

Moreover, in [I], starting from formula (3.16), several hierarchies for cumulants and WP are derived and a direct application of (3.16) is provided for the heuristic derivation of the Boltzmann-Peierls equation from the DNLS.

As we have already mentioned, it turns out that the decay properties of cumulants can be studied via WP. Sticking to the case of a random lattice field \( \psi_x \) with \( x \in \mathbb{Z}^d \), one way to investigate how fast the connected correlations decay is to consider the so-called \( \ell_p \)-clustering norms defined as follows: for \( 1 \leq p < \infty \) and \( n \in \mathbb{N}_+ \) we set

\[
||\psi||_p^{(n)} := \sup_{x_0 \in \mathbb{Z}^d} \left[ \sum_{x \in \mathbb{Z}^{n-1}} |\kappa[\psi(x_0), \psi(x_1), \ldots, \psi(x_{n-1})]|^p \right]^{1/p}, \tag{3.17}
\]
and define analogously $\|\psi\|_p^{(1)} := \sup_{x_0 \in \mathbb{Z}^d} |\mathbb{E}[\psi(x_0)]|$. The norm (3.17) is designed in such a way that it is still well defined for translation invariant measures, for which the ordinary $\ell_p$-norm of cumulants would be proportional to an infinite volume factor.

We will say that the field $\psi$ is $\ell_p$-clustering up to order $m$ if $\|\psi\|_p^{(n)} < \infty$ for all $n \leq m$ and simply $\ell_p$-clustering if the same holds for any $m$. The main result of [II] can be summarized in the following theorem:

**Theorem 3.2.** Suppose $\psi$ and $\phi$ are random lattice fields which are closed under complex conjugation and defined on the same probability space. Assume that $\phi$ is $\ell_1$-clustering and $\psi$ is $\ell_\infty$-clustering, both up to order $2N$ for some $N \in \mathbb{N}_+$. Then their joint cumulants satisfy the following $\ell_2$-estimate for any $n,m \in \mathbb{N}_+$ for which $n,m \leq N$,

$$
\sup_{x' \in \mathbb{Z}^m} \left[ \sum_{x \in \mathbb{Z}^n} |x[\psi(x_1'), \ldots, \psi(x_m'), \phi(x_1), \ldots, \phi(x_n)]|^2 \right]^{1/2} \leq C_{m,n}^{n+m} (n+m)!,
$$

(3.18)

where $C_{m,n}$ is a constant depending on $\|\psi\|_\infty$ and $\|\phi\|_1^{(k)}$ for $1 \leq k \leq 2N$. In particular, all of the above sums are then finite.

The idea of the proof in a nutshell is to use the representation of the cumulants in terms of Wick polynomials and then to employ a Cauchy-Schwarz inequality with respect to probability measure in order to factorize the contributions coming from the two different fields $\psi$ and $\phi$. In the next section we will discuss how this summability result could be applied in the context of kinetic theory focusing in particular on the DNLS.

### 3.3 Possible applications in kinetic theory

Both articles [I] and [II] are part of a project aiming at a rigorous derivation of a Boltzmann transport equation in the kinetic scaling limit of the weakly nonlinear discrete Schrödinger equation with random initial data. This system consists of a complex lattice field $\psi_t(x)$, with $x \in \mathbb{Z}^d$, governed for $t \geq 0$ by the following equation

$$
i \partial_t \psi_t(x) = \sum_{y \in \mathbb{Z}^d} \alpha(x - y)\psi_t(y) + \lambda |\psi_t(x)|^2 \psi_t(x),
$$

(3.19)

where the function $\alpha$ determines the “hopping amplitudes” and $\lambda > 0$ is a coupling constant. Lukkarinen and Spohn [44] have proven that, given a spatially homogeneous time stationary initial data, under the kinetic scaling limit, i.e. as $\lambda \to 0$ for $t = \lambda^{-2}\tau$ with $\tau$ sufficiently small, and under suitable conditions on the dispersion relation $\omega(k) = \hat{\alpha}(k)$, the space-time covariance $W_t(k)$ defined as

$$
\mathbb{E}[\hat{\psi}_0(k') \hat{\psi}_t(k)] = \delta(k' - k)W_t(k)
$$

is governed by a Boltzmann-Peierls equation whose solution has the form

$$
W_t(k) = W_0(k)e^{-t\Gamma(k)}
$$

(3.20)

where

$$
\Gamma(k) = 2 \int_0^\infty dt \int_{\mathbb{T}^d} dk_1dk_2dk_3 \delta(k + k_1 - k_2 - k_3)e^{i\omega(k) + i\omega(k_1) - i\omega(k_2) - i\omega(k_3)} \times [W_0(k_1)W_0(k_3) + W_0(k_1)W_0(k_2) - W_0(k_2)W_0(k_3)].
$$
The perturbative method used in that article relies on a careful decomposition of the nonlinear term which they called “pair truncation operation” (see Lemma 3.2 in [44]) designed to control the oscillatory integrals related to the perturbative expansion. In fact, this decomposition is nothing but a Wick polynomial expansion and in [I] we show at a heuristic level how the combinatorial properties of WP can simplify the derivation of (3.20).

The completely rigorous result in [44] is a contribution towards the mathematical validation of the main physical conjecture for the DNLS, namely that under time stationary random initial data and under the kinetic scaling limit the time correlation function of the particle density $\rho_t(x) = |\psi_t(x)|^2$ of this system should satisfy a linearized phonon Boltzmann equation, see [43, 56] for more details.

Beside the above mentioned tricky treatment of the oscillatory integrals, the major obstacle in attacking this claim has been the fact that no a priori bounds for the space-time correlation functions were available. In [II] we provide a class of uniform in time estimates that could serve to fill this gap. Consider for example the DNLS random field $\psi_t(x, \sigma)$ introduced above with a stationary $\ell_1$-clustering initial measure. Note that this assumption is physically legitimate because the Gibbs equilibrium measure of the DNLS equation has been proven to be $\ell_1$-clustering in [1]. The perturbative expansion of the particle density correlation involves time correlation functions of the form

$$\kappa[\psi_0(0, \sigma_0), \ldots, \psi_0(x_{m-1}, \sigma_{m-1}), \psi_t(x_m, \sigma_m), \ldots, \psi_t(x_{n+m}, \sigma_{n+m})]. \quad (3.21)$$

Theorem 3.2 immediately implies that the $\ell_2$-clustering norm of functions like (3.21) are bounded uniformly in time in term of the $\ell_1$-clustering norm of the fields $\psi_0$. This can be easily seen by indentifying $\psi_0$ with $\psi$ and $\psi_t$ with $\phi$ in theorem 3.2.

Note that this result is actually independent of the dynamics of $\psi_t$, but relies only on the features of the initial measure and on the Wick polynomial representation of cumulants, so it could be applied to a large class of different models as well.
4 Velocity flip model: thermalization and kinetic theory

4.1 The model and the physical picture

The velocity flip model consists of one-dimensional particle chain with harmonic interaction where each particle is endowed with its own Poissonian clock which flips the particle velocity whenever it rings.

More specifically, we consider a periodic chain of $L$ identical particles indexed by $x \in \Lambda_L$ with an interaction potential $\Phi : \mathbb{Z} \rightarrow \mathbb{R}$ which is assumed to be symmetric ($\Phi(-x) = \Phi(x)$) and with finite range $r_\Phi$. Furthermore, $\Phi$ is such that the corresponding dispersion relation $\omega := \sqrt{\Phi}$ is a smooth function on the circle $T := \mathbb{R}/\mathbb{Z}$ where $\hat{\Phi}$ is the Fourier transform of the potential.

The harmonic Hamiltonian is

$$H_L(X) := \sum_{x \in \Lambda_L} \frac{1}{2}(X_x^2)^2 + \sum_{x', x \in \Lambda_L} \frac{1}{2}X_{x'}^1X_x^1\Phi(|x' - x|_1)$$

where positions $X_x^1 := q(x)$ and momenta $X_x^2 := p(x)$ are vectors in $\mathbb{R}^L$.

By adding to the Hamiltonian evolution a velocity-flip noise, the system can be identified with a Markov process $X(t)$ and the process generates a Feller semigroup on the space of observables vanishing at infinity. For $t > 0$ and any $f$ in the domain of the generator $\mathcal{L} = \mathcal{A} + \mathcal{S}$ of the Feller process the expectation values of $F(X(t))$ satisfy an evolution equation $\partial_t(F(X(t))) = \langle (\mathcal{L}F)(X(t)) \rangle$

where

$$\mathcal{A} := \sum_{x \in \Lambda_L} \left( X_x^2 \partial_{X_x^1} - (\Phi_L X_x^1) x \partial_{X_x^2} \right)$$

$$(\mathcal{S}F)(X) := \frac{\gamma}{2} \sum_{x_0 \in \Lambda_L} (F(S_{x_0}X) - F(X)),$$

$$(S_{x_0}X)_x^i := \begin{cases} 
-X_x^i, & \text{if } i = 2 \text{ and } x = x_0, \\
X_x^i, & \text{otherwise}
\end{cases},$$

with $\gamma > 0$ being the Poissonian clock rate, i.e. the strength of the noise contribution.

This model first appeared in [21] in the context of the study of energy transport properties in lattice systems. Let us recall that when energy is the only relevant conserved quantity, for a large class of three or higher dimensional systems energy transport one expects energy to exhibit a diffusive behaviour, i.e. the Fourier’s law of heat conduction should hold. On the other hand, for many one-dimensional systems energy transport deviates from the Fourier’s law, see [9, 57] for more details.

In contrast, the velocity flip model is one of the simplest known one-dimensional models which has a finite heat conductivity and satisfies the time dependent Fourier’s law as proven recently with different approaches by Simon [53] and Lukkarinen [42].

Moreover, the model has been studied also when the extremities of the chain are coupled with thermostats at different temperatures as in [4, 18]. In this case it turns out that the second order correlations in the steady state coincide with those of a similar system called self-consistent heat bath model [8] for which the stationary Fourier’s law $J = \kappa \Delta T$ (being $J$ the current and $\Delta T$ the thermostats temperature difference) is satisfied with an explicit formula for the thermal conductivity $\kappa$. Since the current $J$ is a function of the second order correlations, the same expression for the conductivity is expected to hold for the velocity flip model.

In [III] we try to clarify the physical meaning of the results in [42] by computing the full spatial covariance matrix of positions and momenta at the diffusive time scale $t = O(L^2)$ including the first order corrections to the local thermal equilibrium via suitably defined Wigner-like transforms. This task is achieved with different methods: the first one is through the strong pointwise control of the
kinetic temperature \( T_t(x) = \mathbb{E}p_t(x)^2 \) found in [42], while the second one is based on a kinetic theory approach without taking any scaling limit. The two final results coincide (as they should) and the two different perspectives allow to assign a convincing physical interpretation to the contributions arising in the first order corrections and to formulate a rather precise conjecture about the non equilibrium steady state correlation of the system when its boundaries are coupled with thermostats at different temperatures.

4.2 Kinetic theory predictions for the thermalization of the velocity flip model

As mentioned above, our goal is to get an explicit formula for the full second order correlation of positions and momenta up to \( O(L^{-2}) \). We consider the system with pinning, namely \( \omega(k) > 0 \) for any \( k \in \mathbb{T} \) and in the first part we set \( \gamma > 2\omega(k) \), as assumed in [42].

We denote the full second moments matrix by

\[
C_t^{ij}(x,y) := \mathbb{E}[X^i_x(t)X^j_y(t)].
\]

The decay properties of the semigroup associated to Hamiltonian derived in [42] imply that \( \mathbb{E}X_t^i(t) \) decay to zero exponentially fast on time scale \( O(\gamma^{-1}) \), so that

\[
C_t^{ij}(x,y) - \text{Cov}(X^i_x(t),X^j_y(t)) = O(e^{-\delta t^2})
\]
on the diffusive time scales \( t = O(L^2) \). Therefore, given our approximation scheme, we will study the matrix \( C_t \). In particular, we look at a variant of the classical Wigner transform, namely

\[
U_t(x,k) := \sum_{y \in \Lambda_L} e^{-i2\pi k \cdot y} C_t(x,x + y),
\]

which can be thought as a partial Fourier transform with respect to the spacial displacement of the fields \( y \). Note also that \( U_t(x,k) \) depends in a nontrivial way on the kinetic temperature \( T_t(x) = \mathbb{E}[p_t(x)^2] \) which was investigated in detail in [42] via an accurate control of the associate renewal equation. The main result in [42] can be summarized by saying that for a large class of initial conditions (even far from the local thermal equilibrium) for times larger than \( O(L^{2/3}) \), \( T_t(x) \) is well approximated uniformly in \( x \) by the solution of a discrete diffusion equation: in formulae

\[
|T_t(x) - (e^{-tD} \tau)(x)| \leq C L(1+ t)^{-3/2}, \tag{4.1}
\]

where \( C \) is a constant and \( D \) is a discrete diffusion operator and \( \tau \) an initial condition for the discrete diffusion equation.

Using this tool, at the diffusive scale \( t = O(L^2) \) one obtains

\[
U_t(x,k) = T_t(x) \begin{pmatrix} \omega(k)^{-2} & 0 \\ 0 & 1 \end{pmatrix} - \frac{i\nu(k)\nabla_x T_t(x)}{\omega(k)} \begin{pmatrix} \omega(k)^{-2} & -\gamma^{-1} \\ \gamma^{-1} & 0 \end{pmatrix} + O(L^{-2}), \tag{4.2}
\]

where \( \nabla_x T_t(x) := T_t(x+1) - T_t(x) \) and \( \nu(k) = \partial_k \omega(k)/2\pi \). The first term gives the local thermal equilibrium correlations since the \((q,p)\)-correlation matrix of the equilibrium Gibbs state at temperature \( T \) is

\[
T \begin{pmatrix} \Phi^{-1} & 0 \\ 0 & 1 \end{pmatrix},
\]

while the second term is an \( O(L^{-1}) \) correction to the local equilibrium correlations. Its off-diagonal components can be interpreted as “current terms” while the kinetic theory approach will shed light on the meaning of the diagonal ones.
In the second part of [III] we define another modified Wigner-like transform \( W_t^{\sigma_1, \sigma_2}(\xi, k) \): in contrast with the former one \( U_t(x, k) \), \( W_t^{\sigma_1, \sigma_2}(\xi, k) \) involves a spacial averaging kernel and it is constructed from the phonon eigenmodes instead of from the positions and momenta. Note that in this second part we are considering general spacial dimension \( d \), while in the first one the computations hold only for \( d = 1 \) as required by the assumptions in [42]. In formulae we have

\[
W_t^{\sigma_1, \sigma_2}(\xi, k) := e^{i\omega(k)(\sigma_1 + \sigma_2)} \sum_{x \in \Lambda_L} \varphi(\xi - x) \sum_{y \in \Lambda_L} e^{-2\pi i y \cdot k} \mathbb{E}[\psi_t(x, \sigma_1)\psi_t(x + y, \sigma_2)]
\]

where \( \psi_t(x, 1) = \psi_t(x) \) and \( \psi_t(x, -1) = \psi_t^*(x) \) are the normal modes of the harmonic evolution obtained by setting \( \gamma \) to zero and defined in Fourier space by

\[
\widehat{\psi}_t(k, \sigma) = \frac{1}{\sqrt{2}} (\omega(k)\widehat{q}(k) + i\sigma\widehat{p}(k)).
\]

Here \( \varphi \) is the spacial averaging kernel which is defined in such a way that it varies on scales \( O(R^{-1}) \) for some \( R < L \), namely \( \partial^\alpha \varphi(\xi) = O(R^{-|\alpha|}) \) for any multi-index \( \alpha \), so that a convolution with \( \varphi \) amounts to a spacial averaging over a volume of order \( O(R^d) \).

Adopting a kinetic theory approach, we want to derive a Boltzmann-like equation for \( W_t^{\sigma_1, \sigma_2}(\xi, k) \) which we expect to be composed of a collision term and a transport contribution, because the measure \( \mathbb{E}[-] \) in general is not translation invariant. The function of the oscillating prefactor \( e^{i\omega(k)(\sigma_1 + \sigma_2)} \) is to cancel out the fast oscillations resulting from the free evolution when \( \sigma_1 = \sigma_2 \) and to give rise for all \( \sigma_1, \sigma_2 \) to the usual transport term of the form \( \partial_t \omega(k) \partial_k W_t^{\sigma_1, \sigma_2}(\xi, k) \). In fact, up to order \( O(R^{-2}) \), \( W_t^{\sigma_1, \sigma_2}(\xi, k) \) satisfies the following Boltzmann-like equation

\[
\partial_t W_t^{\sigma_1, \sigma_2}(\xi, k) = \sigma_1 \frac{\nabla \omega(k)}{2\pi} \cdot \nabla_\xi W_t^{\sigma_1, \sigma_2}(\xi, k) + \mathscr{C}[W_t(\xi, \cdot)]^{\sigma_1, \sigma_2}(k) + O(R^{-2})
\]

where \( \mathscr{C}[W_t(\xi, \cdot)]^{\sigma_1, \sigma_2}(k) \) is a linear collision operator whose strength is proportional to \( \gamma \). We point out that the closest quantity to the standard Wigner function among the four components of \( W_t^{\sigma_1, \sigma_2} \) is \( W_t := W_t^{-+, +} \), which indeed can be interpreted as an energy quasi-density in \( (\xi, k) \), as the ordinary Wigner transform in quantum mechanics which is a quasi-probability distribution on the phase space. In fact, we get

\[
\int_{LT} d\xi \int_{\Lambda_L} dk W_t(\xi, k) = H(q_0, p_0)
\]

where \( H(q_0, p_0) \) is the initial Hamiltonian. Consequently, we can also identify the energy density on spacial regions \( O(R^d) \) as \( E_t(\xi) = \int_{\Lambda_L} dk W_t(\xi, k) \).

Moreover, one can tune the scale parameter \( R \) so that the collision term \( O(\gamma) \) is dominant with respect to the transport term \( O(R^{-1}) \), i.e. we set \( R \gg \gamma^{-1} \nu(k) \). In this regime, where the two scales are well separated, the only \( O(1) \) entries are the ones with \( \sigma_1 = -\sigma_2 \), while the other ones are subdominant. In this setting we get for \( W_t(\xi, k) \) a very simple linear Boltzmann equation from which one can infer that, if the final phase of equilibration is slower than ballistic, i.e. \( \partial_t W_t = o(R^{-1}) \), then the energy must relax diffusively:

\[
\partial_t E_t(\xi) \simeq \kappa \nabla_\xi^2 E_t(\xi),
\]

where \( \kappa = \kappa_L := \gamma^{-1} \int_{\Lambda_L} dk \nu(k)^2 \). It is worth noting that this explicit formula matches the one derived [8] for the conductivity of the stationary Fourier’s law for the self-consistent heat baths model, as we expected from the discussion in the former section.
Finally, let us compute the correlation matrix and compare the result with (4.2). To do this we need to perform a change of variable in $\mathcal{W}$ from the normal modes to positions and momenta, namely we consider the observable

$$U(t)(\xi, k) = \sum_{x, y \in \Lambda_L} \varphi(\xi - x) e^{-2\pi i y k} \psi_t(x, x + y) \equiv \sum_{x \in \Lambda_L} \varphi(\xi - x) U_t(x, k)$$

which is nothing but a spacially averaged version of the former $U_t$. Again relying on the separation of scales between trasport and collisions, up to $O(R^{-2})$ one gets

$$\mathcal{W}_t(\xi, k) = E_t(\xi) \begin{pmatrix} \omega(k)^{-2} & 0 \\ 0 & 1 \end{pmatrix} - \frac{i \nu(k)}{\omega(k)} \cdot \nabla_{\xi} E_t(\xi) \begin{pmatrix} \omega(k)^{-2} & -y^{-1} \\ y^{-1} & 0 \end{pmatrix}$$

(4.4)

that matches the former formula after replacing the lattice kinetic temperature profile $T_t(x)$, $x \in \Lambda_L$, with its averaged version which in this model coincides with energy density at thermal equilibrium $E_t(\xi)$, $\xi \in \mathbb{R}$.

Since $\mathcal{W}$ is basically $\mathcal{W}$ after the change of variables $\psi \to (q, p)$, we can now interpret the first order correction to $(q, q)$-correlations in (4.2) and (4.4) as arising from this basis change. We also note that the whole first order correction would vanish for a homogeneous energy distribution because it is proportional to $\nabla_{\xi} E_t(\xi)$.

Moreover, it is natural to expect that formula (4.2) describes also the dominant contribution to the correlation in the bulk for the non equilibrium steady states with the extremities of the chain attached to two thermostats at different temperatures, since typically the effect of the thermostats remains concentrated to the boundary.

As a concluding remark, we observe that the explicit estimates used to derive (4.2) are unlikely to work for more complicated models, while the kinetic theory analysis based on the scale separation between transport and collisions seems to be applicable to other systems as well: for instance, it would be interesting to test it with the three-dimensional crystals with weak anharmonic interaction.
5 Kinetic equations for the Lorentz model with an external magnetic field

5.1 Lorentz model and scaling limits

The Lorentz model [41] is a gas of noninteracting particles in a fixed configuration of scatterers which was introduced in 1905 to corroborate the qualitative theory for electrical conduction in metals developed by Paul Drude few years before. The basic idea is that electrons are described by light particles whose trajectory is affected only by the collisions with the atoms, represented by the scatterers, while the electron-electron interaction is considered to be negligible.

The evolution of the system is governed by the Newton’s law in form given by the following set of ordinary differential equations

\[ \begin{align*}
\dot{x} &= v \\
\dot{v} &= -\sum_i \nabla \phi(|x - c_i|),
\end{align*} \tag{5.1} \]

where \((x, v)\) are the phase space coordinates of a light particle, \(\phi\) is a potential describing the interaction between scatterers and light particles and \((c_1, \ldots, c_N) =: c_N\) is a configuration of scatterers’ centers which we assume to be randomly distributed according to the Poisson distribution: the probability of finding \(N\) obstacles in a bounded set \(A \subset \mathbb{R}^d\) with intensity \(\mu > 0\) is

\[ P(dc_N) = e^{-\mu|A|} \frac{\mu^N}{N!} dc_1 \cdots dc_N. \tag{5.2} \]

Despite its simplicity, Lorentz model has been deeply investigated in the context of kinetic theory since it is one of the few reversible microscopic models which has been connected to macroscopic irreversibility through mathematically rigorous scaling limits. Moreover, as we will see later, from the perspective of probability theory, these scaling limits can be seen as a procedure to approximate the non-Markovian microscopic processes (like the Lorentz gas) with macroscopic Markovian processes. See [55] for an extensive review on the topic.

First, let us consider the Boltzmann-Grad limit or low density limit. Let be \(\epsilon\) the ratio between the macroscopic and the microscopic scales, then we take the total density of scatterers proportional to \(\epsilon\), which corresponds to a mean free time and mean free path of the test particles of order \(\epsilon^{-1}\). This suggests the scaling \(x \to \epsilon x, t \to \epsilon t\), that can be translated into a scaling of the range of the interaction and the intensity of the Poisson distribution of the scatterers as

\[ \phi_{\epsilon}(x) = \phi \left( \frac{x}{\epsilon} \right), \quad \mu_{\epsilon} = \mu \epsilon^{-d+1}. \tag{5.3} \]

Therefore, (5.1) becomes

\[ \begin{align*}
\dot{x} &= v \\
\dot{v} &= -\epsilon^{-1} \sum_i \nabla \phi(\epsilon^{-1}|x - c_i|).
\end{align*} \tag{5.4} \]

Under the scaling (5.3) the stochastic process of the Lorentz particle, converges to a process where the velocity is a Markov jump process and the position is an additive functional of the velocity. Furthermore, consider the one-particle density

\[ f_\epsilon(x, v, t) := \mathbb{E}_\epsilon[f_0(T^{-t}(x, v))] \tag{5.5} \]
where $T^t(x,v)$ is the Hamiltonian flow associated to (5.9), $f_0$ is a given initial distribution and $\mathbb{E}_\varepsilon$ is the expectation with respect to the Poisson distribution of scatterers with intensity $\mu_\varepsilon$. Under suitable hypothesis on $f_0$ one can show that the limit $\lim_{\varepsilon \to 0} f_\varepsilon(x,v,t)$ is governed by the following linear Boltzmann equation:

$$
\partial_t f(x,v,t) + v \cdot \nabla_x f(x,v,t) = Lf(x,v,t)
$$

(5.6)

where

$$
Lf(x,v,t) = \pi \mu |v| \int_{S|v|} dv' k(v',v)(f(x,v',t) - f(x,v,t)).
$$

(5.7)

Here $k(v',v)dv'$ is the probability that the velocity of the Lorentz particle jumps instantaneously from $v$ to $v'$ in a collision and it is proportional to the scattering cross section of the interaction potential. In an unpublished paper [22] (re-printed in [24]) Gallavotti gave the first rigorous contribution to the derivation of the linear Boltzmann equation from the Lorentz gas in the above setting with a hard core interaction potential and the constructive technique employed there is essentially the same as in [IV]. Gallavotti’s result was improved and extended later by using different methods in [7,16,17,54].

Here we want to emphasize that the randomness of the distribution of the scatterers is crucial in the derivation of the linear Boltzmann equation. In fact, Marklof and Strömbergsson [45] showed that for a periodic nonrandom configuration of scatterers under the Boltzmann-Grad limit is not possible to obtain a linear equation like (5.6) in the limit.

Another interesting scaling is the weak coupling limit: the interaction potential is rescaled as $\varepsilon^{1/2}\phi_\varepsilon(x)$, so that the typical velocity variation per collision is order $\varepsilon^{1/2}$ and time and space are scaled as $t \to \varepsilon t$ and $x \to \varepsilon x$ to have $\varepsilon^{-1}$ collisions per unit time interval. Heuristically, the weak strength of the interaction potential combined with huge number of collisions is responsible for a central limit mechanism which is supposed to lead a diffusion type evolution in the limit.

As before, one can rephrase the scaling by rescaling the intensity of the Poisson distribution and the potential range:

$$
\phi_\varepsilon(x) = \varepsilon^{1/2} \phi(\frac{x}{\varepsilon})
$$

$$
\mu_\varepsilon = \mu \varepsilon^{-d}
$$

(5.8)

then (5.1) becomes

$$
\begin{align*}
\dot{x} &= v \\
\dot{v} &= -\varepsilon^{-1/2} \sum_i \nabla \phi(\varepsilon^{-1}|x - c_i|),
\end{align*}
$$

(5.9)

In this setting the limit distribution (5.5) evolves in time according to the linear Landau equation

$$
(\partial_t + v \cdot \nabla_x) f(x,v,t) = \eta \Delta_{|v|} f(x,v,t)
$$

(5.10)

where $\Delta_{|v|}$ is the Laplace-Beltrami operator on the $d$-dimensional sphere of radius $|v|$ and $\eta$ is the diffusion coefficient which depends on the interaction potential. In this case in the limit the velocity process is a Brownian motion on the energy sphere and the position is an additive functional of the velocity process.

The first rigorous result in the direction of the derivation of the linear Landau equation was a paper by Kesten and Papanicolaou [36] where the authors studied the so called “stochastic acceleration problem”, i.e. they considered the motion of particle evolving according to a random force and they showed that under the weak coupling limit in $\mathbb{R}^d$, $d \geq 3$ the system converges in law to a diffusion
process. The restriction on the space dimension is due to the fact that for \( d \geq 3 \) the trajectories of the limit process do not intersect themselves, giving rise to a Markovian process, while for \( d = 2 \) it was not clear whether this was the case. Later Dürr, Goldstein and Lebowitz [19] proved that in \( \mathbb{R}^2 \) the velocity process of the light particle converges in distribution to a Brownian motion on a surface of constant speed for the specific case of the Lorentz gas with Poisson distributed obstacles and a smooth interaction potential. Moreover, in 2006 Komorowski and Ryzhik [38] extended the result in [36] to two dimensions.

Finally, the linear Landau equation arises for the usual setting of the Lorentz gas also in the intermediate scaling between the low density and the weak-coupling regime, i.e.

\[
\phi \rightarrow e^{\alpha \phi(x/\epsilon)}, \quad \mu_{\epsilon} = e^{-2\alpha(d-1)\mu}
\]

for \( \alpha \in (0, 1/2) \). In fact, Kirkpatrick [37] showed the convergence in law of the system to a diffusion process using the same tools of [19], while Desvillettes and Ricci [17] employed the constructive Gallavotti’s method to show convergence in expectation to the Landau equation for the interval \( \alpha \in (0, 1/8) \).

5.2 Effect of an external magnetic field on the two-dimensional Lorentz gas

The paper [IV] considers the microscopic Lorentz gas with Poissonian distribution of scatterers in two spacial dimensions with a constant and uniform external magnetic field orthogonal to the plane where the gas lives.

The goal is to investigate how the magnetic field comes into play when deriving effective equations for the expectation of the one-particle density under suitable scaling limits, namely the intermediate scaling (5.11) and the Boltzmann-Grad scaling (5.3). In particular, one is interested in understanding under which hypotheses (i.e. basically which kind of interaction potential, initial datum and scaling of the magnetic field) the limit equation is still of Markovian type, like the linear Boltzmann equation and the linear Landau equation.

This question was motivated by some earlier results by Bobylev et al. [6] where it was shown that in the Boltzmann-Grad scaling the Lorentz gas with Poissonian distribution of scatterers, hard core interaction potential and magnetic field is described by a non-Markovian equation, that the authors called “generalized Boltzmann equation”.

In fact, the presence of the magnetic field yields in the microscopic system some non-Markovian configurations which fail to vanish even for \( \epsilon \rightarrow 0 \), so that these pathologies reverberate in the limit equations in the form of memory terms. An example of these bad configurations is the one where a light particle follows forever a circular closed trajectory (the so-called Larmor orbit) under the effect of the Lorentz force without hitting any scatterer. Let be \( R \) the radius of the closed orbit, then the probability \( P_R \) of this event for a finite range potential can be estimated by using (5.2) as follows:

\[
P_R \approx e^{-\mu_{\epsilon} \text{Area}(\mathcal{A}_\epsilon)} \approx e^{-2\pi \mu R^2},
\]

where \( v \) is the velocity of the particle and \( \mathcal{A}_\epsilon(R) \) is the annulus of radius \( R \) and width \( \epsilon \). Hence, as anticipated, \( P_R \) stays \( O(1) \) in the limit \( \epsilon \rightarrow 0 \).

However, one could try to tune the scaling in order to make the non-Markovian configurations vanish in the limit. For example, using the intermediate scaling (5.11), the probability of the event considered above becomes

\[
P_{R_L} \approx e^{-\mu_{\epsilon} 2\pi R^2} \approx e^{-2\pi R^2 \mu e^{-2\alpha}}.
\]
which vanishes as $\epsilon \to 0$. Furthermore, one could consider a long range inverse power law interaction potential, truncated at distance $\epsilon^{-1}$ with $\gamma \in (0, 1)$ in the low density regime $\alpha = 0$: in this case the probability of the former event is

$$P_R \approx e^{-2\pi R_\gamma \mu \epsilon^{-1}}$$

which again is zero in the limit. This simple heuristic argument provides some evidence to believe that the non-Markovian kinetic behaviour of the Lorentz gas with magnetic field discovered in [6] is unstable and that Markovianity can be recovered in specific cases. In [IV] we make rigorous these considerations by studying the two regimes described above (intermediate scaling limit and truncated potential).

To state the results, consider the Lorentz gas with a Poissonian distribution of scatterers on the plane with a uniform, constant, magnetic field orthogonal to the plane. The equations of motion are

$$\begin{aligned}
\dot{x} &= u \\
\dot{u} &= B v^\perp - \epsilon^{\alpha-1} \sum_i \nabla \phi (\frac{x - c_i}{\epsilon}) ,
\end{aligned} \tag{5.12}$$

where $B$ is the magnitude of the magnetic field, $v^\perp = (v_2, -v_1)$ and the potential $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ is smooth and of range 1. Set $|v| = 1$ (note that $|v|$ is a conserved quantity) and take the initial distribution $f_0$ to be compactly supported and smooth enough. Let $f_\epsilon(x, v)$ be defined as in (5.5) where the intensity of Poisson distribution is $\mu e^{-1-2\alpha}$ with $\alpha \in (0, 1)$. Then for $t \in [0, T], T > 0$,

$$\lim_{\epsilon \to 0} f_\epsilon(\cdot; t) = g(\cdot; t)$$

where the convergence is in $L^2(\mathbb{R}^2 \times S_1)$ and $g$ is the unique solution to the Landau equation with magnetic field

$$\begin{aligned}
(\partial_t + u \cdot \nabla_x + B v^\perp \cdot \nabla_v) g(x, v, t) &= \eta \Delta_{S_1} g(x, v, t) \\
g(x, v, 0) &= f_0(x, v),
\end{aligned} \tag{5.13}$$

$\Delta_{S_1}$ is the Laplace-Beltrami operator on the circle $S_1$ and $\eta$ is the diffusion coefficient

$$\eta = \frac{\mu}{2} \int_{-1}^1 \left( \int_{\rho}^1 \frac{\rho'}{u} \phi'(\frac{\rho}{u}) \frac{du}{\sqrt{1-u^2}} \right)^2 \, d\rho. \tag{5.14}$$

Note that in the limit the magnetic field affects just the transport term, but does not enter the “collision term” represented by the Laplace-Beltrami operator. Following the Gallavotti’s method in the formulation given in [3, 17], the strategy of the proof consists in (1) showing the asymptotic equivalence between $f_\epsilon$ and $h_\epsilon$, the latter being the solution of a linear Boltzmann equation, and (2) checking that the $h_\epsilon$ converge to $g$, solution of the linear Landau equation. In particular, $h_\epsilon$ solves

$$\begin{aligned}
(\partial_t + u \cdot \nabla_x + B v^\perp \cdot \nabla_v) h_\epsilon(x, v, t) &= L_\epsilon h_\epsilon(x, v, t) \\
h_\epsilon(x, v, 0) &= f_0(x, v)
\end{aligned} \tag{5.15}$$

where

$$L_\epsilon h_\epsilon(v) = \mu_\epsilon \int_{-\epsilon}^\epsilon \, dp [h_\epsilon(v') - h_\epsilon(v)]. \tag{5.16}$$

Here $v' = v - 2(\omega \cdot v)\omega$ is the outgoing velocity after a scattering with incoming velocity $v$ and impact parameter $\rho \in [-\epsilon, \epsilon]$ generated by the potential $\epsilon^\alpha \phi(\frac{x}{\epsilon})$. Moreover, $\omega = \omega(\rho)$ is the versor bisecting the angle between the incoming and outgoing velocity and $\theta_\epsilon$ is the scattering angle.
The technical core of the proof is the first step which is tackled, according to the Gallavotti’s idea, by a change of variables in the description of the light particle trajectories: from the sequence of obstacles \( e_1, \ldots, e_n \) hit by the particle we pass to the sequence of impact times and impact factors \( (t_1, \rho_1), \ldots, (t_n, \rho_n) \) corresponding to each scattering event in the trajectory. It turns out that this change of variables is well defined only outside a set of pathological events, which are exactly the ones that prevent the microscopic rescaled system from being Markovian. Thus, the asymptotic equivalence between \( f_e \) and \( h_e \) is achieved by showing that the probability of those nasty configurations vanishes as \( \epsilon \to 0 \). From these explicit estimates we see also why the magnetic field does not spoil the Markovianity in the limit: in fact, the probability that a particle does not hit any obstacle for times \( t \geq O(\epsilon^\nu) \), with \( \nu < 2\alpha \), vanishes as \( \epsilon \to 0 \), then, since the Larmor period is \( O(1) \), in the limit no closed orbits are left.

Moreover, we note that the “grazing collision” property is preserved also with the presence of the magnetic field, i.e. the scattering angles stay small: \( \theta_\epsilon \leq C\epsilon^\alpha \). This observation makes the proof of the second part of our strategy, namely the convergence towards the Landau equation, essentially identical to the one carried out in [3, 17].

For the second regime, we consider the system (5.12) for \( \alpha = 0 \) where we replace \( \phi(x/\epsilon) \) by an inverse power law potential truncated at large distances \( \psi_\epsilon(|x|/\epsilon) \) like the one introduced in [16]:

\[
\psi_\epsilon(x) = \begin{cases} 
\frac{1}{|x|^{s\gamma-1}} & |x| < \epsilon^{\gamma-1} \\
| |x| \geq \epsilon^{\gamma-1} 
\end{cases}
\]

with \( \gamma \in (6/7, 1) \) and \( s > 2 \). Under suitable assumption on the initial condition \( f_0 \), one has that the expectation of the one-particle distribution \( f_e \) is such that for all \( t \in [0, T] \), \( T > 0 \),

\[
\lim_{\epsilon\to 0} f_e(\cdot; t) = f(\cdot; t)
\]

where the convergence is in \( \mathcal{S}'(\mathbb{R}^2 \times S_1) \) and \( f \) is the unique solution to the linear Boltzmann equation with magnetic field

\[
\begin{cases}
(\partial_t + v \cdot \nabla_x + B v_\perp \cdot \nabla_v) f(t,x,v) = L f(t,x,v) \\
f(x,v,0) = f_0(x,v),
\end{cases}
\]

(5.17)

with

\[
L f(v) = \mu \int_{-\pi}^\pi \Gamma(\theta) [f(\mathcal{R}(\theta)v) - f(v)] \, d\theta.
\]

(5.18)

Here \( \Gamma(\theta) = d \rho / d \theta \) is the differential cross section associated to the long range potential \( \psi(|x|) = |x|^{-s} \) and the operator \( \mathcal{R}(\theta) \) rotates the velocity \( v \) by the angle \( \theta \). The proof proceeds as for the former result, i.e. \( f_e \) is shown to have the same asymptotic behaviour of \( h_{e,Y} \) which is the solution of a linear Boltzmann equation and the convergence of \( h_{e,Y} \) towards \( f \) completes the argument. More in detail, \( h_{e,Y} \) satisfies

\[
\begin{cases}
(\partial_t + v \cdot \nabla_x + B v_\perp \cdot \nabla_v) h_{e,Y}(t,x,v) = \tilde{L} h_{e,Y}(t,x,v) \\
h_{e,Y}(x,v,0) = f_0(x,v),
\end{cases}
\]

(5.19)

with

\[
\tilde{L} f(v) = \mu \int_{-\pi}^\pi \Gamma_{e,Y}^{(B)}(\theta) [f(\mathcal{R}(\theta)v) - f(v)] \, d\theta.
\]

and \( \Gamma_{e,Y}^{(B)} \) is the differential cross section associated to the unrescaled potential \( \psi_\epsilon \) with magnetic field.
In this case the “grazing collisions” property does not hold, so that in the second step of the proof, following [16], one needs to study the convergence of the the cross section $\Gamma^{(B)}_{e_1}$ to $\Gamma$ in (5.18). In this analysis we rely on the technique employed in the Appendix of [16] where the assumption $s > 2$ enters. Anyway, this is believed to be just a technical condition that one could relax with some appropriate trick.
6 Renormalization group for SPDEs and nonlinear fluctuating hydrodynamic

6.1 An overview about KPZ equation

In this section we want to give a non technical introduction about the Kardar-Parisi-Zhang (KPZ) equation. For a much more extended survey on the topic see the reviews [13], [50], [51], [59].

At a formal level, this equation reads

$$\partial_t h = \lambda (\partial_x h)^2 + \partial_x^2 h + \xi$$

where $\lambda > 0$ is a “coupling strength”, $h(x,t)$ is a continuous stochastic process with $x \in \mathbb{R}$, and $\xi$ is the space-time white noise which is a distribution-valued Gaussian field with correlation given by

$$E \xi(t,x)\xi(s,y) = \delta(x-y)\delta(t-s).$$

This equation appeared first in a paper by Kardar, Parisi and Zhang [35] and it was introduced to model the large scale fluctuations of a growing interface. The physical derivation proceeds as follows: the evolution of $h$ is assumed to be governed by a slope dependent lateral growth, relaxation and random forcing. Then one can write

$$\partial_t h = \lambda F(\partial_x h) + \partial_x^2 h + \xi$$

where $\partial_x^2 h$ is the simplest possible form of relaxation, the random forcing $\xi$ is assumed to be roughly independent at different positions and different times (which leads to the choice of space-time white noise) and the nonlinear term is considered to be a symmetric function of the slope, i.e. it is assumed to depend on $(\partial_x h)^2$. If we suppose that $(\partial_x h)^2$ is small, i.e. the interface $h$ is nearly flat, then one can formally expand $F(\partial_x h)$ as

$$F(y) = F(0) + F'(0)y + \frac{1}{2} F''(0)y^2 + \cdots.$$ 

The first term in the expansion can be reabsorbed with a time shift and the second one vanishes because $F$ is even, then, by keeping the first nontrivial term $\frac{1}{2} F''(0)y^2$, one gets indeed an equation of the form (6.1).

Kardar, Parisi and Zhang employed non rigorous perturbative dynamical renormalization group arguments to predict that, under the so-called “1-2-3” rescaling and after a suitable recentering, $h$ must converge to a scale invariant random field $H$ (the KPZ fixed point) obtained as

$$H(x,t) = \lim_{\epsilon \to 0} \epsilon^4 h(\epsilon^{-2} x, \epsilon^{-3} t) - c(\epsilon) t.$$ 

This is the so called “strong KPZ universality conjecture” and it means that on the scale where $h \sim t^{1/3}$ and $x \sim t^{2/3}$ we expect to see some nontrivial fluctuations. Its rigorous proof is still an open problem, see [14] for some recent progress.

Moreover, KPZ equation is expected to be an effective model for the fluctuations of a large class of systems characterized by the dynamical scaling exponents $3/2$. For example, it is known rigorously to arise as the fluctuation process for the weakly asymmetric simple exclusion process [5] and the partition function for directed polymer models [2]. This type of predictions are commonly referred to as the “weak KPZ universality conjecture”.

23
Since the appearance of the paper by Kardar, Parisi and Zhang in 1986 the main difficulty in addressing these problems in a rigorous way has been the fact that the KPZ equation itself is ill-posed as stated in (6.1). In fact, for any fixed time $t > 0$ the space regularity of the $h(t,x)$ cannot be better than Brownian, i.e. $\mathcal{C}^{\gamma}$ with $\gamma < 1/2$ in the scale of Hölder spaces, so the quadratic term $(\partial_x h(t,x))^2$ is not well defined. Coming back to the physical derivation of the KPZ, the ill-posedness of the nonlinear term is reminiscent of the fact that we performed a Taylor expansion without actually knowing whether $\partial_x h$ was small. In fact, this term is huge at small scales because of the white noise, then $(\partial_x h)$ can be considered small only when interpreted as a slope variation at large scale. In this way one can justify the presence of the counterterm $c(\epsilon)$ in the formulation of the (6.5): its function is to cancel out the small scale divergence (“ultraviolet divergence” in the jargon of quantum field theory) produced by $(\partial_x h)^2$. Therefore, we expect $c(\epsilon) \to \infty$ as $\epsilon \to \infty$.

The first rigorous result concerning the KPZ equation is due to Bertini and Giacomin [5] who introduced the notion of “Cole-Hopf solution” of the KPZ: formally one can see that the transformation $z := e^h$ maps the KPZ equation to the linear multiplicative stochastic heat equation

$$dz = \partial_x^2 z dt + \lambda z dw(t),$$

where $w(t,x) = \int_0^t ds \xi(s,x)$ is a cylindrical Brownian motion on $L^2(\mathbb{R})$ and the term $z dw(t)$ should be interpreted in the Itô sense. (6.6) admits a unique positive solution in a suitable space of adapted processes, so one can define the Cole-Hopf solution of the KPZ as $h(x,t) := \lambda^{-1} \log z(x,t)$.

In [5] the authors proved that the the Gärtner’s microscopic Cole-Hopf transform [25] of the height function of the weakly asymmetric simple exclusion process (WASEP) converges to the solution of the stochastic heat equation (6.6). Even though the convergence result shows that any possible candidate for the solution of the KPZ equation should be such that $e^h$ solves (6.6), the microscopic Cole–Hopf transform works just for few specific models, so this method in not robust enough to attack the universality conjecture mentioned above.

However, there has been a tremendous amount of work in understanding the probability distribution of the Cole-Hopf solution of the KPZ equation that culminated in 2010 when two independent groups (see [2, 52]) found for suitable initial conditions the exact formula for the its one-point statistics. Quite surprisingly, this involves probability distributions already known in the field of random matrices, the so-called Tracy-Widow distribution. Moreover, these results showed that, as expected, the statistics of the Cole-Hopf solution of the KPZ scales like $t^{1/3}$ for large $t$. The exact formulæ allowed also to explore the short time limit of the statistics which turns out to scale like $t^{1/4}$, as for systems in the Edwards Wilkinson universality class [20]. Models in this class are governed by the additive stochastic heat equation and hence have a Gaussian one-point distribution. This observation sheds a new light on the KPZ equation: in fact, it can be interpreted as a mechanism for crossing over between two universality classes – the KPZ class in long-time and the Edwards Wilkinson class in short-time.

It is worth emphasizing that the result in [2, 52] deals with the Cole-Hopf solution of the KPZ, so no approximation theory for the actual KPZ equation is employed in those articles. The breakthrough in this direction was due to Hairer [29] who proved by rough paths techniques that solutions to equation (6.1) (with $\lambda = 1$) are limits for $\epsilon \to 0$ of the approximate solutions $h_\epsilon$ of the equation

$$\partial_t h_\epsilon(t,x) = \partial_x^2 h_\epsilon(t,x) + (\partial_x h_\epsilon(t,x))^2 - C_\epsilon + \xi_\epsilon(t,x), \quad x \in T, \quad t \geq 0$$

where $T$ is the one-dimensional torus, $\xi_\epsilon$ is a Gaussian process obtained as a mollification of the white noise and the constant $C_\epsilon$ is a counterterm that diverges as $\epsilon \to 0$ and cancels the ultraviolet divergence of the nonlinear term. Moreover, this limit coincides with the Cole-Hopf solution. Recently, Hairer
used his theory of regularity structures [30] to prove in collaboration with Quastel [33] a universality result for the KPZ equation: given any even polynomial $P$ and a mollified version of the white noise $\xi$, they showed that under the right rescaling the solution to the equation

$$\partial_t h = \partial_x^2 h + \sqrt{\epsilon} P(\partial_x h) + \xi$$

converges to the KPZ equation with a constant in front of the nonlinearity which in general depends on the polynomial $P$.

After the publication of Hairer’s paper [29], other methods have been developed to treat singular nonlinear stochastic PDEs: Gubinelli and Perkowski [28] recovered Hairer’s result on the KPZ by using paracontrolled distributions [27] and the same method is employed by Catellier and Chouk [12] to prove existence and uniqueness of a local in time solution of the dynamical stochastic $\Phi^4$ theory in 3 dimensions on the torus. Moreover, Mourrat and Weber managed to established a global in time well-posedness theory for the dynamical stochastic $\Phi^4$ theory in two dimensions on the whole real line [47] and in three dimensions on the torus [48].

Recently Kupiainen obtained a result similar to the one by Catellier and Chouk with Renormalization Group (RG) techniques [40]. In [V] we use the same RG machinery to construct local in time solution to a generalized KPZ equation involving a vector valued “height field”: component-wise it reads

$$\partial_t h_\alpha = \partial_x^2 h_\alpha + (\partial_x h, M^\alpha \partial_x h) + \xi_\alpha$$

where $\alpha = 1, 2, 3$, $h(t,x) : \mathbb{R}_+ \times \mathbb{T} \to \mathbb{R}^3$, $M$ is a symmetric matrix, $(\cdot, \cdot)$ is the ordinary scalar product in $\mathbb{R}^3$ and $\xi$ is a vector valued white noise such that

$$\mathbb{E}[\xi_\alpha(t,x)\xi_\beta(s,y)] = \delta_{\alpha\beta}\delta(t-s)\delta(x-y).$$

In the ensuing section we are going to outline the physical context in which equation (6.9) arises, namely the theory of nonlinear fluctuating hydrodynamics for anharmonic chains.

### 6.2 Fluctuating hydrodynamics for anharmonic particle chains

Over the last few years Spohn and collaborators [15, 46, 57, 58] have been studying the fluctuations of the hydrodynamical fields for one-dimensional fluids. As a microscopic model they adopted the particle anharmonic chains whose fluctuating hydrodynamics is in fact indistinguishable from the one of the one-dimensional fluids.

These works were motivated by the fact that, while the hydrodynamics of analogous systems in higher dimensions is well described by linear equations leading to good predictions for the asymptotic time behaviour of the Green-Kubo correlation of order $t^{-d/2}$, in one dimension it turns out that a linear description fails to match the related numerical simulations. We will see how resorting to a nonlinear approach will lead to an equation of the form (6.9).

Concretely, we consider a chain consisting of particles with unitary mass and we denote their positions and momenta by $q_j$ and $p_j$ respectively with $j = 1, \ldots, N$. The Hamiltonian is

$$H_N = \sum_{j=1}^N \frac{1}{2}p_j^2 + \sum_{j=1}^{N-1} V(q_{j+1} - q_j)$$

where we take $V$ to be a stable nearest neighbour potential, so that we could interpret $H_N$ as describing an anharmonic chain, namely particles in one dimension coupled through anharmonic springs.
Assuming the boundary conditions to be $q_1 = 0$ and $q_{N+1} = \ell N$, the conserved quantities are the total length of the chain, total momentum and total energy:

\[
\sum_{j=1}^{N} r_j = \ell N, \quad \sum_{j=1}^{N} p_j = u N, \quad \sum_{j=1}^{N} e_j = e N
\]  

(6.12)

where $r_j = q_j - q_{j-1}$ is the stretch, $e_j = \frac{1}{2} p_j^2 + V(r_j)$ is the local energy, while $\ell$, $u$ and $e$ are the average stretch, momentum and energy per particle. We also assume that there are no further local conservation laws.

Collecting these local fields in the 3-dimensional vector $g(j, t) = (r_j(t), p_j(t), e_j(t))$, the conservation laws reads

\[
\frac{dg(j, t)}{dt} + J(j+1, t) - J(j, t) = 0
\]  

(6.13)

where $J(j) = (p_j, -V'(r_j-1), -p_jV'(r_{j-1}))$ is the vector of the local current functions.

The main object of interest is the large scale behaviour of the equilibrium time correlations of the conserved fields which encodes the propagation of local perturbations of the equilibrium state:

\[
S_{\alpha\alpha'}(j, t) = \langle g_\alpha(j, t)g_{\alpha'}(0, 0) \rangle - \langle g_\alpha(0, 0) \rangle \langle g_{\alpha'}(0, 0) \rangle = \kappa [g_\alpha(j, t), g_{\alpha'}(0, 0)],
\]  

(6.14)

where $\alpha, \alpha' = 1, 2, 3$ and $\langle \cdot \rangle$ denote the expectation with respect to the canonical equilibrium measure associated to $H_N$.

From numerical simulations, after a suitable rotation of the matrix $S$, for large time scales we expect to see three peaks (see for example Fig. 1 in [58]): two symmetric peaks (sound peaks) moving with a certain sound speed $c$ and a central peak (heat peak) sitting at the spacial origin. More interestingly, the heat peak broadens in time as $t^{2/5}$, while the sound peaks as $t^{2/3}$, which is the conjectured scaling for the KPZ two point correlation.

An explicit computation of $S(j, t)$ is unattainable, so one could try to describe the large scale structure of $S(j, t)$ by considering the fields in $q$ as slowly varying function of the continuum spacial variable $x$, see [58] for more details about this continuum limit. Since we are interested here only in small deviations from equilibrium, we can write the hydrodynamical fields as

\[
\ell + \eta_1(x, t), \quad u + \eta_2(x, t), \quad e + \eta_3(x, t)
\]

and then write $S(x, t)$, the continuum limit of $S(j, t)$, as a function of the fluctuation fields $\eta(x, t) = (\eta_1(x, t), \eta_2(x, t), \eta_3(x, t))$. Then a good ansatz for the time evolution of the fluctuation fields is the following nonlinear equation

\[
\partial_t \eta(x, t) + \partial_x \left( A \eta(x, t) \right) + \frac{1}{2} \left( E^\alpha \eta(x, t) \right) - \partial_x (D \eta(x, t)) + (B \xi(x, t)) = 0,
\]  

(6.15)

where the term proportional to the matrix $A$ comes from the crude linearization of (6.13) around the equilibrium average $(\ell, u, e)$, $\xi_\alpha$ is the space time white noise defined in (6.10) and it models the random fluctuation of the current fields, while the contribution proportional to the matrix $D$ is a dissipative term. Here the nonlinear coupling $E^\alpha_{\beta\gamma}$ is a three-indices tensor and it represents the Hessian of the currents with derivatives evaluated at $(\ell, u, e)$. Moreover, the matrices $A, B$ and $D$ are assumed to be independent from $x$ and $t$ in first approximation and their explicit form as a function of the equilibrium parameters can be found in [58].
We now diagonalize the system with respect to the matrix $A$: let be $R$ the rotation matrix such that $RAR^{-1} = \text{diag}(c, 0, -c)$ where $c$ is the speed of the sound peaks, then for the rotated field $\zeta = R\eta$ we get
\begin{align*}
\partial_t \zeta_\alpha + \partial_x \left[c_\alpha \zeta_\alpha + (\zeta, K^\alpha) - \partial_x (D^\# \zeta) + (B^\# \zeta)_\alpha\right] = 0
\end{align*}
(6.16)
where $B^\#, D^\#$ and $K$ are the rotated version of $B, D$ and $E$ respectively. Given the characteristics of the chain, $K$ has some symmetry properties such that $K^\alpha_\beta = K^\alpha_\gamma, K^0_0 = 0$ and $K^1_1 = -K^{-1}_{-1}$ which is generically different from 0. This signals that the heat peak will behave differently from the sound peaks.

Note that (6.16) is of the form of the coupled KPZ equation which is studied in [V]: in fact, set $\partial_x h_\alpha = \zeta_\alpha$, then
\begin{align*}
\partial_t h_\alpha + c_\alpha \partial_x h_\alpha + (\partial_x h, K^\alpha \partial_x h) - (D^\# \partial_x^2 h) + (B^\# h)_\alpha = 0,
\end{align*}
(6.17)
which boils down to the ordinary (uncoupled) KPZ equation if $K^\alpha_\beta = \delta_\alpha_\beta \delta_\beta_\gamma$.

Since the definition of the fluctuation correlations in (6.14) involves the equilibrium measure, one should try to find the stationary measure associated to (6.17). It turns out that the Gaussian stationary measure of the linear Langevin equation corresponding to (6.17) is stationary also under (6.17) when the cyclicity condition holds:
\begin{align*}
K^\alpha_\beta = K^\beta_\alpha,
\end{align*}
(6.18)
i.e. when $K^\alpha_\beta$ is totally symmetric under the permutation of its indices. The stationary measure of (6.17) is not know in the general case, but in [58] Spohn argues that the contributions coming from the off-diagonal matrix entries of the nonlinear coupling are irrelevant for the large scale behaviour, so that the stationary measure of the linear equation is still a good approximation.

However, this discussion shows that, from the point of view of the nonlinear fluctuation hydrodynamics, it would be very interesting to construct rigorously the stationary measure of the coupled KPZ equation (6.17) when $K^\alpha_\beta \neq K^\beta_\alpha$. The article [V] represents a contribution for attacking this problem: in fact, even though the solution to (6.17) is constructed just for short times, the cyclicity condition plays an important role also in [V] since, whenever it holds, the counterterms needed to make sense out of the coupled KPZ is the same as for the simple KPZ.

This indicates that the symmetry breaking affects in a nontrivial way even the small time scales, while its effect at large times requires further investigations.

### 6.3 Renormalization of the coupled KPZ

Consider the coupled KPZ equation in (6.17): we can absorb the ballistic transport term $c_\alpha \partial_x h_\alpha$ by passing to the moving reference frame, i.e. $h_\alpha(t,x) \rightarrow h_\alpha(t,x - c_\alpha t)$, so that
\begin{align*}
\partial_t h_\alpha + (\partial_x h, K^\alpha \partial_x h) - (D^\# \partial_x^2 h) + (B^\# h)_\alpha = 0.
\end{align*}
(6.19)

In [V] we construct the solution for short times of a simplified version of (6.19), namely
\begin{align*}
\partial_t h_\alpha = \partial_x^2 h_\alpha + (\partial_x h, M^\alpha \partial_x h) + \Xi_\alpha
\end{align*}
(6.20)
where $h(t,x) : \mathbb{R}_+ \times \mathbb{T} \rightarrow \mathbb{R}^3, M^\alpha_\beta = M^\alpha_\beta$ and $\Xi$ is the vector valued white noise with covariance given by (6.10).
Note that the spacial domain of solution of (6.19) is the entire real line $\mathbb{R}$, while (6.20) is defined on the torus. The problem with the equation on the full real line is that the white noise is not uniformly Hölder continuous on $\mathbb{R}$, thus one needs to weight the Hölder spaces of functions/distributions to control the behaviour at infinity. In this direction recently Hairer and Labbé in [31, 32] studied the parabolic Anderson model on $\mathbb{R}^2$ and $\mathbb{R}^3$ and the simple KPZ equation on $\mathbb{R}$ via the Cole-Hopf transform. Another work dealing with SPDEs on unbounded spacial domains is [47], where the authors show the global well-posedness of the dynamic $\Phi^4$ model on the full plane.

As we explained in section 6.1, equation (6.20) is ill-posed as it stands and in [V] we provided an approximation theory to make sense of it. We study its integral (mild) form

$$
 h(t) = \int_0^t e^{(t-s)\Delta}(V(h) + \Xi(s))ds + e^{t\Delta}h_0
$$

where $V(h) = (\partial_x h, M\partial_x h)$, and $e^{t\Delta}$ denotes the semigroup associated to the heat equation such that in Fourier space $e^{t\Delta}f(p) = e^{-tp^2}f(p)$. We take the initial condition $h_0$ to be the stationary solution to the linear problem with $V = 0$, which we can construct from $\Xi$ as

$$
 h_0 = \int_{-\infty}^0 e^{(t-s)\Delta}\Xi(s)ds.
$$

In the following we will use the more compact notation

$$
 h(t) = G * [(V(h) + \Xi)1(t \geq 0)] + e^{t\Delta}h_0
$$

where $G(t, x) = e^{t\Delta}(x, 0)$ and $(G*f)(t) = \int_{-\infty}^t e^{(t-s)\Delta}f(s)$. Thus the initial condition can be absorbed in the noisy contribution by rewriting (6.21) as

$$
 h = G * (V(h)1(t \geq 0) + \Xi).
$$

In order to set up an approximation theory for (6.9), we choose not to smooth out the noise in space as in [29], but we regularize the convolution with a smooth cutoff on the small time scale:

$$
 h_\epsilon = G_\epsilon * (V^{(\epsilon)}(h_\epsilon) + \Xi)
$$

where $G_\epsilon(t, x) = e^{t\Delta}(x, 0)(1 - \chi(\epsilon^{-2}t))$ with $\chi \geq 0$ being a smooth bump, $\chi(t) = 1$ for $t \in [0, 1]$ and $\chi(t) = 0$ for $t \in [2, \infty)$ and $V^{(\epsilon)}(u) = [(\partial u, M\partial u) + C_\epsilon]1(t \geq 0)$. As mentioned earlier, the constant $C_\epsilon$ is needed to cancel the diverging small scale variation of $h$ and we will adjust it such that (6.24) has a unique solution $h_\epsilon$ which converges to a nontrivial limit as $\epsilon \to 0$.

We can now state the main result of [V].

**Theorem 6.1.** There exists $C_\epsilon$ such that the following holds. For almost all realizations of the white noise $\Xi$ there exists $t(\Xi) > 0$ such that the equation (6.24) has for all $\epsilon > 0$ a unique smooth solution $h_\epsilon(t, x)$, $t \in [0, t(\Xi)]$ and there exists $h \in \mathcal{D}'([0, t(\Xi)]) \times T$ such that $h_\epsilon \to h$ in $\mathcal{D}'([0, t(\Xi)]) \times T$. The limit $h$ is independent of the regularization function $\chi$.

Notice that we are able to construct the solution of (6.9) up to a certain time $t(\Xi)$ which depends on the noise, in particular one has that $t(\Xi) = O(L^{-2m})$ where $N_+ \geq m = m(\Xi) < \infty$ a.s. and $L > 1$ is the scale parameter. The renormalization counterterm $C_\epsilon$ has the form

$$
 C_\epsilon = m_1 \epsilon^{-1} + m_2 \log \epsilon^{-1} + m_3
$$

28
where the constants $m_1$ and $m_3$ depend on $\chi$ whereas the $m_2$ is universal, i.e. independent of $\chi$. It is worth noticing that for the simple KPZ (6.1) the logarithmically divergent term is not needed. This is due to the fact that for equation (6.1) the two divergent terms arising from the third order of the perturbative expansion happen to cancel each other. Furthermore, the same cancellation is recovered for equation (6.9) if $M_{1y}^{\alpha}$ is totally symmetric in the three indices, so that $m_2 = 0$ in (6.25).

Given the structure of equation (6.24), one has to resort to some kind of fixed point argument in order to prove theorem 6.1. However, since the limit $h$ will be a distribution, it is not clear how to set this up due to the nonlinearity $V(h)$. Here we use the Wilsonian approach to renormalization in the setting developed by Kupiainen in [40] and inspired by earlier works by Kupiainen and collaborators on KAM theory and renormalization of PDEs [10, 11].

The main idea is to try to proceed scale by scale and derive for each scale a divergence-free effective equation, while the original divergence is recovered when looking at the flow of the effective equation over all the scales. To be more concrete, let us first rescale the equation to set the small scale cutoff to unity by using the diffusive scaling under which the linear part of the equation is invariant: let $s_\varepsilon$ be the scaling operator such that

$$(s_\varepsilon h)(t,x)e^{-\frac{1}{2}h(e^2t,e\varepsilon)} =: \varphi(t,x)$$

thus

$$\varphi = G_1 \ast (\psi^{(\varepsilon)}(\varphi) + \xi) \quad (6.26)$$

where $\varphi$ is defined on $\mathbb{R} \times \varepsilon^{-1}T$ and

$$\psi^{(\varepsilon)}(\varphi) := \varepsilon^3(\partial_x^3 \varphi, M \partial_x \varphi) + \varepsilon^2 C_\varepsilon \quad (6.27)$$

and $\xi := \varepsilon^2 s_\varepsilon \Xi$ is equal in law to the white noise on $\mathbb{R} \times \varepsilon^{-1}T$.

In these dimensionless variables the small scale cutoff is unity and the strength of the nonlinearity is small, $\varepsilon^{1/2}$ i.e. the model is subcritical or super-renormalizable in the quantum field theory jargon: this implies that there is at most a finite number of diverging terms to be renormalized. Even though (6.26) has no small scale singularity, now we need to control it for times of order $\varepsilon^{-2}$ and spatial box of size $\varepsilon^{-1}$. Since it is easy to solve it on time scale $O(1)$, we can now concretely implement the idea of “splitting” the problem in different scales with this recursive procedure: after fixing a scale parameter $L > 1$ and setting $\varepsilon = L^{-N}$ (so that taking $\varepsilon \rightarrow 0$ amounts to $N \rightarrow \infty$)

1. we solve the equation up to time scale $O(L^2)$;
2. we derive an effective equation for larger scales, i.e. $\geq O(L^2)$, where the fluctuations at smaller scales are averaged out;
3. we rescale the effective equation back to get a unitary cutoff and we apply again the previous steps.

In this way we will produce of flow of effective dimensionless equations whose solutions $\varphi_n$ live on scale $\geq O(L^{2n})$ with $n = m, \ldots, N$ for some $m$, where the original equation can now be viewed as the initial condition of this flow. In this context the renormalization constant $C_\varepsilon = C_{L^{-N}}$ is nothing but a parameter to be added to the initial condition of the equations flow so that the limit in theorem 6.1 exists.

It is actually convenient to translate the flow of effective equations into the flow of effective potentials: we study the map

$$\psi^{(N)}_{n-1}(\varphi) = R^{\dagger} \psi^{(N)}_n(\varphi) \quad (6.28)$$
where $v^{(N)}_n$ is such that
\[ \varphi_n = G_1 \ast (v^{(N)}_n(\varphi_n) + \xi_n) \]  
(6.29)
with $\xi_n := L^{-2n} s^{-n} \Xi$ having the same distribution as the white noise on $\mathbb{R} \times L^n \mathbb{T}$. The map (6.28) is called Renormalization Group map and it connects effective potentials living on adjacent scales.

Then we set the initial condition of this flow to be the original potential $v^{(N)}_{N}(\varphi) = v(\varphi)$ and we iterate $\mathcal{R}$ to get $v_m = \mathcal{R}^{N-m} v_N$ which provides the solution to the dimensionless equation on the time interval $[0, L^{2(N-m)}]$, and consequently the solution for the original regularized equation on the time interval $[0, L^{-2m}]$. The final goal is then to take the limit of $v^{(N)}_m$ for $N \to \infty$ in order to remove the cutoff and prove theorem 6.1. Recall that the smallest $m$ for which the limit exists depends on the realization of the noise.

In this specific case $v$ is a function of $\partial_x \varphi$, so it is helpful to set $\phi := \partial_x \varphi$ and $w^{(N)}_n(\phi) := v^{(N)}_n(\varphi)$, hence we will now consider the flow of the polynomial potential $w^{(N)}_n(\phi)$, and it is the composition of a scaling and a translation:
\[ \mathcal{R} w(\phi) = \mathcal{I} w(\phi + \psi) \]  
(6.30)
where $\mathcal{I}$ is a scaling operator defined by $\mathcal{I} f = L^2 s^{-1} f \circ s$ and $\psi$ is a random function of $v$:
\[ \psi = \Upsilon \ast \xi + \Upsilon \ast w(\phi + \psi) \]  
(6.31)
where $\Upsilon(t,x) = \partial_x e^{L \lambda}(x,0)(\chi(t) - \chi(Lt))$. We will identify the divergent term by looking at the linearized version of $\mathcal{R}$:
\[ \mathcal{L} w(\phi) = \mathcal{I} w(\phi + \Upsilon \ast \xi). \]  
(6.32)
The operator $\mathcal{L}$ is just the composition of the scaling operator $\mathcal{I}$ with the translation with $\Upsilon \ast \xi$ and one can easily see that the local eigenfunctions of $\mathcal{I}$ are monomials of the form $\phi^k$:
\[ \mathcal{I} \phi^k = L^{\frac{k}{2}} \phi^k. \]  
(6.33)
The form of the eigenvalues implies that for $0 \leq k \leq 2$ the corresponding eigenfunction is expanding or relevant, for $k = 3$ is marginal and for $k > 3$ is contracting or irrelevant. This observation suggests to expand $w$ by extracting the possible divergent contributions as follows:
\[ w_n = \sum_{i=1}^{3} u_{n,i}(\phi) + r_n \]  
(6.34)
where $u_{n,i}$ are explicit perturbative contributions such that in a suitable norm and after a suitable renormalization
\[ \|u_{n,i}\| = O(L^{-\frac{i}{2}}), \]  
(6.35)
and
\[ \|r_n\| = O(L^{-2n}), \quad r_{n-1} = \mathcal{L} r_n + O(L^{-\frac{5}{2}n}). \]  
(6.36)
Moreover, from our analysis of $\mathcal{L}$ we also expect that via a fixed point argument one obtains
\[ \|\mathcal{L} r_n\| \leq CL^{\frac{3}{2}} \|r_n\| \leq CL^{-\frac{1}{2}} L^{-2(n-1)}, \]
so that \( \| r_n \| \) stays \( O(L^{-2n}) \) for any \( n \). The same idea works in general for subcritical problems: in fact, if the dimensionless strength of the nonlinearity is \( L^{-N\alpha} \) for some \( \alpha > 0 \) and the norm of \( \mathcal{L} \) is \( L^\beta \) for some \( \beta > 0 \), then one needs to expand perturbatively up to order \( k - 1 \) with \( k\alpha > \beta \).

From the perturbative expansion it turns out that the iteration of the RG map generates a divergent term at the first order and at the third order: in fact, at the first order we have that \( u_{n,1} \) (which is the effective potential generated by the iteration of the sole linearized RG map) is given by

\[
u_{n,1} (\phi) = L^{-\frac{n}{2}} (\phi + \delta_n, M(\phi + \delta_n)) - L^{-\frac{n}{2}} C_{L^{-N}}. \tag{6.37}
\]

where

\[
\delta_n(t) = \partial_x \int_0^t e^{(t-s)\Delta}(\chi(t-s) - \chi(L^{2(n-n)}(t-s)))\xi_n(s)ds. \tag{6.38}
\]

Since \( \mathbb{E}(\delta_n, M\delta_n) \approx m_1 L^{N-n} \) where \( m_1 \in \mathbb{R}^3 \) is the same constant appearing in (6.25), we need to compensate this divergence as \( N \to \infty \), then at the first order we must set \( C_{L^{-N}} = m_1 L^N \). The other contribution to \( C_{L^N} \) comes from the contribution to the third order term \( u_{n,3} \) which is constant in \( \phi \) : its leading part is

\[
u_{n,3}(0) = \frac{\pi}{\sqrt{3}} (\mathcal{M}_2 - \mathcal{M}_1) \log L^{-N} \tag{6.39}
\]

where

\[
(\mathcal{M}_1)_\alpha = \sum_{\beta_1, \beta_2, \beta_3, \beta_4} M_{\beta_1}^{(\alpha)} M_{\beta_2}^{(\beta_1)} M_{\beta_3}^{(\beta_2)} M_{\beta_4}^{(\beta_3)} \tag{6.40}
\]

\[
(\mathcal{M}_2)_\alpha = \sum_{\beta_1, \beta_2, \beta_3, \beta_4} M_{\beta_1}^{(\alpha)} M_{\beta_2}^{(\beta_1)} M_{\beta_3}^{(\beta_2)} M_{\beta_4}^{(\beta_3)}.
\]

In order to subtract the divergence (6.39) we need to set

\[
m_2 = \frac{\pi}{\sqrt{3}} (\mathcal{M}_2 - \mathcal{M}_1)
\]

and we easily see from (6.40) that \( m_2 = 0 \) when \( M_{\beta\gamma}^{\alpha} \) is totally symmetric in three indices. This observation explains the fact that in the simple KPZ no logarithmically divergent counterterm is needed: in fact, the coupled KPZ equation boils down to three uncoupled simple equation when \( M_{\beta\gamma}^{\alpha} = \delta_{\alpha\beta} \delta_{\beta\gamma} \) which is obviously totally symmetric.

Finally, let us comment on the functional space setting of the fixed point problem for which (6.35) and (6.36) hold. Since the random fields generated in the RG group iteration are in the Sobolev space \( H^{2^{-2}}_{loc} \) in time and space, we can take \( \omega_n \in H^{2^{-2}}_{loc} \). Moreover, noting that \( Y \) in (6.31) is a smoothing kernel, then we can take \( \phi \) to be smooth, in particular we set \( \phi_n \in C^{2,2}([0,L^{2n}] \times L^{n-T}) \) to match the regularity of \( \omega_n \).

We want to emphasize that the heuristic idea presented above and summarized by the expected behaviour (6.35) and (6.36) can hold true only if the random fields are not too large. In particular, we carry out the argument leading to (6.35) and (6.36) in the probability set \( \mathcal{S}_m \) where \( \| \xi_n \| \leq L^m \) for any \( m \leq n \leq N \) and for some \( \gamma \) > 0, where \( \xi_n \) is any of the random fields appearing in \( u_{n,i} \) and \( r_n \), like for example \( \delta_n, (\delta_n, M\delta_n) - m_1 L^N \) or \( u_{n,3}(0) - m_2 \log L^{-N} \).

After the renormalization, we are able to prove that almost surely \( \mathcal{S}_m \) holds for some \( m < \infty \). Therefore, we can work out the renormalization group procedure and show that the limit \( \lim_{N \to \infty} w_n^{(N)} \) exists for all \( n \geq m \).
As future prospects, it is natural to ask whether this method could be improved to show the global well-posedness in time of the equation, i.e. ruling out blow-up in finite time and also whether the equation for the height field \( h \) defined on the whole real line could be handled. Speaking of the latter issue, as already mentioned in section 6.1, some recent works on the other singular stochastic PDEs defined on unbounded spacial domains [31, 47] point out that one would need to set the Banach fixed point argument on some weighted space of distributions to control the behaviour of the equation at infinity in space.

Another interesting open question is whether this approach could be extended to treat also critical (marginal) systems. So far none of the recent techniques developed to study singular stochastic PDEs (regularity structures, paracontrolled distribution and RG) seem to work in these case. The only paper in this direction is the one by Hairer and Shen [34] were the authors studied the dynamical Sine-Gordon model

\[
\partial_t \phi = \frac{1}{2} \Delta \phi + c \sin(\beta \phi + \theta) + \xi.
\]  (6.41)

where \( \phi : \mathbb{R}_+ \times \mathbb{T}^2 \) and \( c, \beta, \theta \) are real constants. This model is subcritical for \( \beta^2 < 8\pi \), but it becomes critical for \( \beta^2 = 8\pi \). Hairer and Shen succeeded to construct the solution in the interval \( \beta^2 \in (0, \frac{16}{5} \pi) \), so they did not actually tackle the critical regime.

The problem of all the present methods is that all the relevant terms are treated one by one, but this is not possible in the critical case where the number of divergent terms in the perturbative expansion is not finite. Hence, new ideas are needed to design an efficient procedure capable of handling the relevant and marginal terms collectively.
References


7 Articles [I]-[V]
Wick polynomials and time-evolution of cumulants

Jani Lukkarinen, Matteo Marcozzi
Wick polynomials and time-evolution of cumulants

Jani Lukkarinen\textsuperscript{1, a}) and Matteo Marcozzi\textsuperscript{1, b})

University of Helsinki, Department of Mathematics and Statistics
P.O. Box 68, FI-00014 Helsingin yliopisto, Finland
(Dated: June 8, 2016)

We show how Wick polynomials of random variables can be defined combinatorially as the unique choice which removes all “internal contractions” from the related cumulant expansions, also in a non-Gaussian case. We discuss how an expansion in terms of the Wick polynomials can be used for derivation of a hierarchy of equations for the time-evolution of cumulants. These methods are then applied to simplify the formal derivation of the Boltzmann-Peierls equation in the kinetic scaling limit of the discrete nonlinear Schrödinger equation (DNLS) with suitable random initial data. We also present a reformulation of the standard perturbation expansion using cumulants which could simplify the problem of a rigorous derivation of the Boltzmann-Peierls equation by separating the analysis of the solutions to the Boltzmann-Peierls equation from the analysis of the corrections. This latter scheme is general and not tied to the DNLS evolution equations.

PACS numbers: 02.50.Cw, 05.50.+q, 05.20.Dd
Keywords: Wick polynomials, cumulants, Lattice dynamics, Non-Linear Schrödinger equation, Boltzmann equations

I. INTRODUCTION

Wick polynomials, also called Wick products, arose first in quantum field theory as a way of regularizing products of field operators\textsuperscript{1}. The principal goal there was to replace monomial products by polynomials with state dependent coefficients, chosen so as to remove singular terms appearing in the associated perturbation expansion.

The procedure can also be applied in more general probabilistic settings. The following definition is given in Wikipedia\textsuperscript{2} and in the Encyclopedia of Mathematics\textsuperscript{3}. Consider \( n \) (real) random variables \( y_j, j = 1, 2, \ldots, n \), on some probability space \((\Omega, B, \mu)\) and denote expectation over the probability measure \( \mu \) by \( \langle \cdot \rangle \). The Wick polynomial with powers \( k_j \geq 0, j = 1, 2, \ldots, n \), are then defined recursively in the total degree \( k_1 + k_2 + \cdots + k_n \) by the following conditions:

1. If \( k_1 = k_2 = \cdots = k_n = 0 \), set \( :y_1^{k_1} y_2^{k_2} \cdots y_n^{k_n} : = 1 \).

2. If the total degree is greater than zero, require that \( \langle :y_1^{k_1} y_2^{k_2} \cdots y_n^{k_n} : \rangle = 0 \).

3. For all \( j \), require that the (algebraic) derivatives of the Wick polynomials satisfy

\[
\partial_{y_j} :y_1^{k_1} \cdots y_j^{k_j} \cdots y_n^{k_n} : = k_j :y_1^{k_1} \cdots y_j^{k_j-1} \cdots y_n^{k_n} : .
\]  

(1.1)

These conditions have a unique solution for which \( :y_1^{k_1} y_2^{k_2} \cdots y_n^{k_n} : \) is a polynomial of total degree \( k_1 + k_2 + \cdots + k_n \) in the variables \( y_j \). (The uniqueness is algebraic, not only almost everywhere as random variables. That is, the conditions fix all coefficients of the polynomials. This can be seen by induction in the order \(|I|\): the requirement in item 3 fixes all new coefficients apart from the constant, which is then fixed by the vanishing of the expectation

\textsuperscript{a}) jani.lukkarinen@helsinki.fi
\textsuperscript{b}) matteo.marcozzi@helsinki.fi
The coefficients are polynomials of expectations of the random variables \( y_j \), and hence depend on the measure \( \mu \). The first order polynomial is obtained by simply centering the variable, \( y_1; = y_1 - \langle y_1 \rangle \), but already at second order more complex structures appear. \( y_1 y_2; = y_1 y_2 - \langle y_1 \rangle \langle y_2 \rangle \). If the random variables have joint exponential moments, i.e., if there is \( \beta > 0 \) such that \( \langle e^{\beta \sum_j |y_j|} \rangle < \infty \), the Wick polynomials can also be obtained by differentiating a fairly simple generating function. It can then be defined for \( \lambda \in \mathbb{R}^n \), such that \( |\lambda_j| < \beta \) for all \( j \), by

\[
G_w(\lambda; y_1, \ldots, y_n) = \frac{\exp \left( \sum_{i=1}^n \lambda_i y_i \right)}{\exp \left( \sum_{i=1}^n \lambda_i \langle y_i \rangle \right)},
\]

and then for all \( k_j \geq 0, j = 1, 2, \ldots, n \),

\[
y_1^{k_1} y_2^{k_2} \ldots y_n^{k_n} = \partial_{\lambda_1}^{k_1} \cdots \partial_{\lambda_n}^{k_n} G_w(\lambda; y_1, \ldots, y_n)\bigg|_{\lambda=0}.
\]

The generating function \( G_w(\lambda; y_1, \ldots, y_n) \) is also called “Wick exponential” and often denoted by \( \langle \exp (\sum_{i=1}^n \lambda_i y_i) \rangle \). For a derivation and basic properties of such Wick polynomials, see Ref. 4.

The Wick polynomials become particularly simple to use if the joint measure of \( y \) is Gaussian. Defining the covariance matrix by \( C_{y', y} := \text{Cov}(y_j, y_j') \), a Gaussian measure has \( \langle \exp (\lambda \cdot y) \rangle = \exp(\langle \lambda \cdot y \rangle - \frac{1}{2} \lambda^T C_{y} \lambda) \). Therefore, the generating function of the Wick polynomials then reads simply \( G_w(\lambda; y) = \exp(\lambda \cdot y - \frac{1}{2} \lambda^T \cdot C_{y} \cdot \lambda) \), and the resulting Wick polynomials are closely related to Hermite polynomials. This is the setting encountered in the original problem of renormalization of quantum field theories (the “unperturbed measures” concern free fields and hence are Gaussian). More discussion and details can be found for instance in Refs. 5–7.

In the Gaussian case, one can also identify the Wick polynomials as arising from an orthogonalization procedure. Wiener chaos expansion and Malliavin calculus used for stochastic differential equations can be viewed as applications of such orthogonal projection techniques.

In the non-Gaussian case, there are far fewer examples of applications of Wick polynomial techniques. The computations become then more involved. For instance, there is no explicit formula for the generating function unless the inverse of the moment generating function happens to be known explicitly. In addition, then the polynomials typically no longer form an orthogonal set in \( L^2(\mu) \).

The goals of the present contribution are two-fold. In the first part, we show that Wick polynomials have a natural combinatorial definition, closely connected to cumulants and the related cluster expansions of correlation functions. We also rederive their main properties without resorting to the generating function, hence without assuming Gaussianity or the existence of exponential moments.

In the second part, we show how Wick polynomial expansions may be used in the analysis of stochastic processes. In particular, the goal there is to apply the expansion to study the time-evolution of the cumulants, i.e., of the connected correlation functions, of the process. We will explain there why often it is cumulants, and not moments, which should be used as dynamical variables. For simplicity, we consider here only processes whose dynamics are deterministic and given by a differential equation, such as Hamiltonian evolution in classical particle systems. The randomness enters via the initial state. However, generalization to Markovian stochastic dynamics should be straightforward, for instance, if the generator of the process maps polynomials to polynomials.

In the general setup, the best one can hope for are recursion relations leading to an infinite hierarchy of equations connecting the evolution of the cumulants. We explain in Section IV what immediate constructions are available for hierarchical study of the evolution of cumulants and Wick polynomials.

We give more explicit applications in section V where we study the evolution on a lattice of particles following the discrete nonlinear Schrödinger (DNLS) equation with random initial data. In particular, our goal is to show how the Wick polynomial expansion of the dynamics
greatly simplifies the (still only formal) derivation of the related Boltzmann-Peierls equation. This case is one of the few examples of nonlinear Hamiltonian evolution where a rigorous analysis of the related perturbation expansion has been possible so far. It has been proven that if the initial measure is a stationary Gibbs measure, then the time-correlations of the field follow an evolution equation derived using a perturbation expansion analogous to the one needed for the Boltzmann-Peierls equation.

An ultimate goal of the present reformulation of the evolution problem would be to complete the rigorous derivation, and hence give a region of validity, of the Boltzmann-Peierls equation. We show how the Wick polynomial expansion could help in this goal by separating the problem of solving the effect of the Boltzmann-Peierls evolution from the estimation of the corrections arising from the wave nature of the microscopic evolution, such as constructive interference. For the DNLS evolution the Wick polynomial expansion coincides with what was called “pair truncation” in Ref. 9. In fact, the present work arose from an attempt to generalize this construction to other polynomial potentials, which we later realized to coincide with Wick polynomial expansions. It should already be apparent from the above example that in order to use the Wick polynomials some care is needed in the choice of notations to avoid being overcome by lengthy formulae and intractable combinatorial estimates. We begin by explaining our choices in detail in Section II. The first part containing the combinatorial definition and properties of Wick polynomials is given in Section III. The second part discussing the use of Wick polynomial expansions for the study of evolution of cumulants begins in Section IV. We conclude it with the specific application to DNLS dynamics in Section V. Some comments and possible further directions are discussed in Section VI.

II. SETUP AND NOTATIONS

We consider a collection \( y_j, j \in J \) where \( J \) is some fixed nonempty index set, of real or complex random variables on some probability space \((\Omega, \mathcal{B}, \mu)\). If \( y_j \) are complex, we assume that the collection is closed under complex conjugation, i.e., that to every \( j \) there is \( j' \in J \) such that \( y_{j'} = y_j^* \).

Expectation over the probability measure \( \mu \) will be denoted by \( \mathbb{E} \) or \( \langle \cdot \rangle \). In case the underlying measure needs to be identified, we denote the expectation by \( \mathbb{E}_\mu \) or \( \langle \cdot \rangle_\mu \). We use sequences of indices, \( I = (i_1, i_2, \ldots, i_n) \in J^n \), to label monomials of the above random variables, with the following shorthand notation
\[
y_I = y_{i_1}y_{i_2}\cdots y_{i_n} = \prod_{k=1}^n y_{i_k}.
\] (II.1)

We also set \( y^\emptyset := 1 \) if \( I \) is the empty sequence. Since all \( y_j \) commute with each other, we have \( y_I = y_{I'} \) for any two sequences \( I, I' \) which differ by a permutation.

We will need to operate not only with such sequences but also with their subsequences and “partitions”. This will be done by choosing a distinct label for each member of the sequence and collecting these into a set. How the labelling is done is not important, as long as one takes care when combining two “labelled” sets. We rely here on the following standard conventions: any sequence \( (i_k) \) can be uniquely identified with the function \( k \mapsto i_k \) which itself is uniquely determined by its graph, the subset \( \{(k, i_k) | k = 1, 2, \ldots, n\} \) of \( \mathbb{N} \times J \). We consider subsequences to be subsets of the graph of the sequence. Partitions of the sequence then correspond to partitions of its graph which can be understood as partitions into nonempty subsequences.

Mathematically, this leads to the following structure. Finite (sub)sequences of indices are now uniquely labelled by the collection \( \mathcal{I} \), which consists of those finite subsets \( A \subset \mathbb{N} \times J \) with the property that if \( (n, j), (n', j') \in A \) and \( (n, j) \neq (n', j') \) then \( n \neq n' \). We also allow the sequence to be empty which is identified with \( \emptyset \in \mathcal{I} \). For nonempty sets, the natural number in the first component serves as a distinct label for each member in \( A \). In addition,
we can use the order of the natural numbers to collapse any $A \in I$ back to a sequence $\hat{A}$ in $J$: Given $A \in I$ with $n > 0$ elements, there is a unique bijection $g : \{1, 2, \ldots, n\} \to A$ such that its first component is increasing, $g(k)_1 < g(k')_1$ for all $k < k'$. Using this $g$, we define $\hat{A}_k := g(k)_2 \in J$ for $k = 1, 2, \ldots, n$.

To each finite sequence $I = (i_k)$ of $n$ elements in $J$, we assign $\hat{I} := \{(k, i_k) \mid k = 1, 2, \ldots, n\}$ as the set of labels. Obviously, then $\hat{I}$ and any of its subsets belong to $I$. The following list summarizes some basic notations and definitions which will be used later without further remark.

1. If $I$ is a sequence, and a set is needed by the notation, the set is chosen to be $\hat{I}$. For instance, the notation “$A \subset I$” means $A \subset \hat{I}$.

2. The notation $P(E)$ denotes the collection of partitions of the set $E \in I$. If $I$ is a sequence, $P(I) := P(\hat{I})$.

3. If $A \in I$ and it is used in a place of a sequence, the formula always refers to the collapsed sequence $\hat{A}$ obtained via the increasing bijection $g$ above. For instance, then $y^A := y^A = \prod_{k=1}^n y_{g(k)_2}$. (Note that if $I$ is a sequence, then $y^I = y^I$ in agreement with (II.1).)

4. If $A \in I$, we denote the corresponding sequence of random variables by $y_A := (y_{A_k})_{k=1}^n$.

5. If $m \in \mathbb{N}$ and $A \in I$, the notation $\hat{A}^{(m)}$ refers to a set where any element with label $m$ is cancelled, i.e., $\hat{A}^{(m)} := \{(k, i_k) \in A \mid k \neq m\}$. Note that it is possible that $\hat{A}^{(m)} = A$.

6. Any two sequences $I$ and $I'$ can be merged into a new sequence $(i_1, \ldots, i_{|I|}, i'_1, \ldots, i'_{|I'|})$ which we denote by $I + I'$. If $A, B \in I$, we take $A + B := \hat{A} + \hat{B}$. For a merged sequence, the notation “$I \subset I + I'$” always refers to the collection of the labels of the first $|I|$ members and analogously “$I' \subset I + I''$” refers to the collection of the last $|I'|$ members. The merge operation is clearly associative, and we hence drop parentheses when it is applied iteratively; for instance, $I + I' + I''$ is a sequence of length $|I| + |I'| + |I''|$.

7. To avoid separate treatment of expressions involving empty sets and conditions, we employ here the following standard conventions: if the condition $P$ is false, we define

$$\sum_P (\cdots) := 0, \quad \prod_P (\cdots) := 1 \quad \text{(II.2)}$$

and set also $P(\emptyset) := \{\emptyset\}$.

Similarly to the moments, to any $I \in I$ we denote the corresponding cumulant by one of the following alternative notations

$$\kappa[y] = \kappa[y] = \mathbb{E}[y_{i_1}; y_{i_2}; \cdots; y_{i_n}] = \kappa(y_{i_1}, y_{i_2}, \cdots, y_{i_n}).$$

The corresponding Wick polynomial is denoted by

$$y_{i_1} y_{i_2} \cdots y_{i_n} := y^I := y^{I}_{\pm}.$$  \quad \text{(II.3)}

Note that this notation is slightly formal, since the result is not a function only of the power $y^I$ but depends on all subpowers, $y^A, A \subset I$, as well. It also requires that one carefully defines which random variables are being “Wick contracted”. We will use parentheses for this purpose, if necessary. For instance, “$(y^I)$” means $y^I = \mathbb{E}[y^I]$ which usually differs from $y^I$. 

As an application of the above definitions, let us point out that the earlier defining Wick polynomial condition (I.1) is equivalent to the requirement that for every nonempty sequence $I$ and any $j \in J$ we should have

$$\partial_{y_j} y^I := \sum_{k=1}^{|I|} \mathbb{1}(i_k = j) y^{I^{(k)}}.$$  \hspace{1cm} (II.4)

Here, and in the following, $\mathbb{1}$ denotes the generic characteristic function: $\mathbb{1}(P) = 1$ if the condition $P$ is true, and otherwise $\mathbb{1}(P) = 0$.

We recall that, if the random variables $y_j$, $j = 1, 2, \ldots, n$, have joint exponential moments, then moments, cumulants and Wick polynomials can be generated by differentiation of their respective generating functions which are

$$G_m(\lambda) := \mathbb{E}[e^{\lambda^T y}], \quad g_c(\lambda) := \ln G_m(\lambda), \quad G_w(\lambda; y) := \frac{e^{\lambda^T y} - g_c(\lambda)}{\mathbb{E}[e^{\lambda^T y}]} = e^{\lambda^T y - g_c(\lambda)}.$$  

Here $\lambda \cdot x := \sum_{i=1}^n \lambda_i x_i$ (for the sake of clarity we have denoted the integrated random variable by “$y$” instead of “$y$”) and the “generation” happens by evaluation of the $I$:th derivative at zero, i.e.,

$$\mathbb{E}[y^I] = \partial_I G_m(0), \quad \kappa^\pi[y_I] = \partial_I^\pi g_c(0), \quad y^I := \partial_I G_w(0; y).$$  \hspace{1cm} (II.5)

where $\partial_I^\pi$ is a shorthand notation for $\partial_{\lambda_1} \partial_{\lambda_2} \cdots \partial_{\lambda_n}$.

As a side remark, let us also recall that it is possible to replace the above definitions of generating functions by parametrizations which do not require the existence of any moments and hence work for arbitrary Borel probability measures $\mu$. If all $y_j$ are real, then replacing the exponential $e^{\lambda^T y}$ by $e^{\lambda^T y_j}$ yields an $L^1(\mu)$ function for all $\lambda \in \mathbb{R}^n$. If $y \in \mathbb{C}^n$, the same is achieved by using $e^{(\lambda^T y + \lambda^* y^*)/2}$ and $\lambda \in \mathbb{C}^n$: in this case, differentiation with respect to $\text{Re} \lambda_j$ generates “$\text{Re} y_j$” and with respect to $\text{Im} \lambda_j$ generates “$\text{Im} y_j$”. Naturally, without absolute integrability of the moments it is not guaranteed that any of the derivative exist. However, it might nevertheless be useful to inspect the time evolution of the generating function, in particular, if the time evolution is regularizing and improves the integrability of the moments.

### III. COMBINATORIAL DEFINITION AND PROPERTIES OF THE WICK POLYNOMIALS

Let us first recall the “moments-to-cumulants” formula which holds for any $I \in \mathcal{I}$ as long as all moments $y^A$, $A \subset I$, belong to $L^1(\mu)$:

$$\mathbb{E}[y^I] = \sum_{\pi \in \mathcal{P}(I)} \prod_{A \in \pi} \kappa^\pi[y_A].$$  \hspace{1cm} (III.1)

where $\mathcal{P}(I)$ denotes the collection of partitions of the set $I$. (Or to be precise, of $\tilde{I}$. Here it is important to assign a distinct label to each random variable in the power $y^I$ to get the combinatorics correctly.) For a partition $\pi \in \mathcal{P}(I)$, let us call the subsets $A \in \pi$ clusters or blocks. Let us also recall that the cumulants are multilinear, i.e., they are separately linear in each of the variables $y_j$, $j \in I$. These results are discussed, for instance, in Refs. 10 and 11, and also briefly in Appendix A here.

Let us point out that by the conventions adopted here, (III.1) is indeed valid also for the empty sequence $I = \emptyset$. Then the sum over partitions is not empty since it contains $\pi = \emptyset$. However, the corresponding product is empty since there is no $A$ with $A \in \pi$. Therefore, the right hand side of (III.1) evaluates to one which agrees with our definition of $\mathbb{E}[y^\emptyset]$.

We next show that it is possible to choose a subset of the indices and remove all its “internal clusters” from the moments-to-cumulants formula by replacing the corresponding power with a polynomial of the same order. This will be achieved by using the following recursive definition.
Definition III.1 Suppose that \( I_0 \in \mathcal{I} \) is such that \( \mathbb{E}[|y_I|] < \infty \) for all \( I \subset I_0 \). We define polynomials \( W[y_I] := \sum_{E \subset I} c_E[y_I] y^E \) for \( I \subset I_0 \) inductively in \( |I| \) using the following rule: set \( W[y^\emptyset] := 1 \), and for \( I \neq \emptyset \) use
\[
W[y_I] := y_I - \sum_{\emptyset \neq E \subset I} \mathbb{E}[y^E] W[y_I \setminus E]. \tag{III.2}
\]

The definition makes sense since the \( W \)-terms on the right hand side all have an order lower than \( |I| \). It also implies that indeed each \( W[y_I] \) is a polynomial of order \( |I| \), with only the term \( y_I \) being of the highest order. It is also straightforward to prove by induction that the coefficients \( c_E[y_I] \) can be chosen so that they only depend on \( \mathbb{E}[y^A] \) with \( A \subset I \). In Appendix A we explain how cumulants can also be defined via a similar recursive construction.

The following theorem shows that these polynomials indeed have the promised truncated moments-to-cumulants expansion. We also see that the polynomials are essentially uniquely defined by this property. What is perhaps surprising is that the coefficients of the polynomial can be chosen depending only on the moments of its constituent random variables. This implies that the same polynomial can be used for many different probability distributions, as long as the marginal distributions for the constituent random variables are the same.

**Theorem III.2** Assume that the measure \( \mu \) has all moments of order \( N \), i.e., suppose that \( \mathbb{E}[|y_I|] < \infty \) for all \( I \in \mathcal{I} \) with \( |I| \leq N \). Use Definition III.1 to define \( W[y_I] \) for every such \( I \).

Then replacing \( y_I \) by \( W[y_I] \) removes all terms with clusters internal to \( I \): the following truncated moments-to-cumulants formula holds for every \( I' \in \mathcal{I} \) with \( |I'| + |I| \leq N \)
\[
\mathbb{E}[W[y_I] y_{I'}] = \sum_{\pi \in \mathcal{P}(I+I')} \mathbb{I}(A \cap I' \neq \emptyset \text{ for all } A \in \pi) \prod_{A \in \pi} \kappa_{|A|}. \tag{III.3}
\]

In particular, \( \mathbb{E}[W[y_I]] = 0 \) if \( I \neq \emptyset \).

In addition, if \( I \in \mathcal{I} \) with \( |I| \leq N/2 \) and \( W' \) is a polynomial of order at most \( |I| \) such that (III.3) holds for all \( I' \) with \( |I'| \leq N - |I| \), then \( W' \) is \( \mu \)-almost surely equal to \( W[y_I] \).

**Corollary III.3** Assume that \( \mathbb{E}[|y_I|] < \infty \) for all \( I \in \mathcal{I} \). Then \( W[y_I] \) are \( \mu \)-almost surely unique polynomials of order \( |I| \) such that (III.3) holds for every \( I' \in \mathcal{I} \).

**Proof:** We make an induction in \( |I| \). By (III.1), the claim is true for \( |I| = 0 \) since then \( I = \emptyset \) and thus \( W[y^\emptyset] = 1 \).

Assume then that \( I \neq \emptyset \) and that the claim is true for sets of size less than \( |I| \). Consider an arbitrary \( I' \in \mathcal{I} \) such that \( |I'| + |I| \leq N \). For \( E \subset I \), denote \( E^c := (I + I') \setminus E \). Given a partition \( \pi \) of \( I + I' \), we can define
\[
\pi_1 := \{ A_1 \in \pi \mid A_1 \cap I' \neq \emptyset \} \quad \text{and} \quad \pi_0 := \pi \setminus \pi_1.
\]

Then \( E := \cup_{\pi_0 \subset I} \pi_0 \in \mathcal{P}(E) \), \( \pi_1 \in \mathcal{P}(E^c) \) (also whenever \( E \) or \( E^c \) happens to be empty). Once \( \pi \) is fixed, the decomposition \( \pi = \pi_0 \cup \pi_1 \) is unique and we thus find that
\[
1 = \sum_{E \subset I} \sum_{\pi_0 \in \mathcal{P}(E)} \sum_{\pi_1 \in \mathcal{P}(E^c)} \mathbb{I}(\pi = \pi_0 \cup \pi_1) \prod_{A_1 \in \pi_1} \mathbb{I}(A_1 \cap I' \neq \emptyset). \]
Using this in the
standard moments-to-cumulants formula shows that
\[
\mathbb{E}[y^I y^{I'}] = \mathbb{E}[y^{I+I'}] = \sum_{\pi \in \mathcal{P}(I+I')} \prod_{A \in \pi} \kappa[y_A]
\]
where in the last step we used the moments-to-cumulants formula and the induction hypothesis (note that \(E^c\) collapses to the sequence \((I \setminus E) + I'\)). Hence, by the definition (III.2) equation (III.3) holds for this \(I\). This completes the induction step and shows that (III.3) is valid for all \(I, I'\) with \(|I| + |I'| \leq N\). If \(I \neq \emptyset\) and \(I' = \emptyset\), we have \(I + I' \neq \emptyset\) so that for any \(\pi \in \mathcal{P}(I + I')\) there is some \(A \in \pi\) and then obviously \(A \cap I' = \emptyset\). Thus (III.3) implies that \(\mathbb{E}[W[y^I]] = 0\) for \(I \neq \emptyset\).

To prove uniqueness, suppose that \(I \in \mathcal{I}\) with \(|I| \leq N/2\) and \(W^x\) is a polynomial of order at most \(|I|\) such that (III.3) holds for all \(I'\) with \(|I'| \leq N - |I|\). Then \(P_I := W^x - W[y^I]\) is a polynomial of order at most \(|I|\) and \(\mathbb{E}[P_I y^{I'}] = 0\) for all \(I'\) with \(|I'| \leq N/2\). Since the collection of random variables is assumed to be closed under complex conjugation, this implies that also \(\mathbb{E}[P_I(y^I)^*] = 0\) whenever \(|I'| \leq N/2\). Thus we can take a linear combination of such equations and conclude that \(\mathbb{E}[|P_I|^2] = 0\). This implies that \(P_I = 0\) almost surely, i.e., that \(W^x = W[y^I]\) almost surely.

This concludes the proof of the Theorem. The Corollary is then an immediate consequence. \(\square\)

In fact, the polynomials given by Definition III.1 are equal to the standard Wick polynomials.

**Proposition III.4** Suppose that \(I_0 \in \mathcal{I}\) is such that \(\mathbb{E}[\|y^I\|] < \infty\) for all \(I \subset I_0\). Then \(W[y^I] = :y^I:\) for every \(I \subset I_0\).

**Proof:** If \(I = \emptyset\), we have \(W[y^I] = 1 = :y^I:\); and else by Theorem III.2 we have \(\mathbb{E}[W[y^I]] = 0\). Therefore, to prove \(W[y^I] = :y^I:\) it suffices to check that (II.4) holds when the Wick polynomials are replaced by \(W\)-polynomials. We do this by induction over \(|I|\). Firstly, if \(I = \emptyset\), we have \(\partial_y W[y^I] = 0\), as required. Assume then that \(I \neq \emptyset\) and that the claim is true for sets of size less than \(|I|\). For every \(j \in J\) and nonempty \(E \subset I\), the induction assumption implies that

\[
\partial_{y_j} W[y^I \setminus E] = \sum_{k=1}^{|I|} \mathbb{I}(i_k = j) \mathbb{I}((k, i_k) \notin E) W[y^I \setminus E].
\]  

(The second characteristic function allows keeping the labeling inherited from \(I\) by adding zero terms into the sum for the “missing” labels.) Since \(W[y^I]\) satisfies (III.2), we thus find
that its algebraic derivatives satisfy an equality
\[
\partial_{y_j} W[y^I] = \sum_{k=1}^{\lvert I \rvert} \mathbb{1}(i_k = j) \left[ y^I - \sum_{\emptyset \neq E \subset I} E[y^E] \mathbb{1}((k, i_k) \notin E) W[y^{I \setminus E}] \right]
\]
\[
= \sum_{k=1}^{\lvert I \rvert} \mathbb{1}(i_k = j) \left[ y^{I'} - \sum_{\emptyset \neq E \subset I'} E[y^E] W[y^{I' \setminus E}] \right]_{I' = \overline{I}(k)}
\]
\[
= \sum_{k=1}^{\lvert I \rvert} \mathbb{1}(i_k = j) W[y^{\overline{I}(k)}],
\]
where in the last equality we have applied (III.2). This proves that also \(W[y^I]\) satisfies (II.4) and hence completes the induction step. Therefore, the polynomials \(W[y^I]\) satisfy the defining properties of Wick polynomials and thus \(W[y^I] = :y^I:\).

\[\square\]

### A. Basic properties of the Wick polynomials

In this section, we assume that there is \(I_0 \in \mathcal{I}\) is such that \(E\|y^I\| < \infty\) for all \(I \subset I_0\). This guarantees the existence of the Wick polynomials \(W[y^I]\) for all \(I \subset I_0\), and allows using the results from the previous section. In particular, by Theorem III.4 these are equal to the standard Wick polynomials and from now on we will use the standard notation \(y^I\) for them.

The next Proposition collects some of the most important properties of Wick polynomials.

**Proposition III.5** The following statements hold for any \(I \subset I_0\):

1. \(y^I = \sum_{U \subset I} :y^U:\ E[y^{I \setminus U}] = \sum_{U \subset I} :y^U:\ \sum_{\pi \in \mathcal{P}(I \setminus U)} \prod_{A \in \pi} \kappa[y_A].\) \hfill (III.5)

2. Wick polynomials are permutation invariant: if \(I'\) is a permutation of \(I\), then \(y^{I'} = :y^{I'}:\.

3. \(y^I = \sum_{U \subset I} y^U \sum_{\pi \in \mathcal{P}(I \setminus U)} (-1)^{\lvert \pi \rvert} \prod_{A \in \pi} \kappa[y_A].\) \hfill (III.6)

4. If \(I' := \overline{I}(1)\) denotes the sequence obtained by cancelling the first element of \(I\), then \(y^{I'} = y_{i_1} y^{I'} - \sum_{\{1, i_1\} \subset V \subset I} \kappa[y_V] y^{I \setminus V} = y_{i_1} y^{I'} - \sum_{U \subset I'} \kappa[y_{\{i_1\} \cup U}] y^{I \setminus U}.\)

**Proof:** Item 1: The first equality in (III.5) follows directly from the definition (III.2) since then
\[
y^I = \sum_{E \subset I} E[y^E] :y^{I \setminus E} = \sum_{U \subset I} E[y^{I \setminus U}] :y^U.\]

The second equality follows then by using the moments-to-cumulants expansion.

Item 2: The permutation invariance of the Wick polynomials follows using straightforward induction in the definition (III.2) since the random variables commute and hence the powers \(y^E\) are always permutation invariant.
Items 3 and 4: Let us first define \( \tilde{W}[y^I] \) by setting it equal to the right hand side of (III.6) for any \( I \subset I_0 \). If \( I = \emptyset \), we have \( \tilde{W}[y^I] = 1 = :y^I: \). Suppose \( I \neq \emptyset \). Since \( \sum_{\pi \in \mathcal{P}(\emptyset)} (-1)^{|\pi|} \prod_{A \in \pi} \kappa[y_A] = 1 \), the definition yields a polynomial of order \(|I|\) in \( y \). Our goal is to prove that

\[
\partial_y \tilde{W}[y^I] = \sum_{k=1}^{|I|} \#(i_k = j) \tilde{W}[y^{I(k)}], \tag{III.8}
\]

\[
\tilde{W}[y^I] = y_I \tilde{W}[y^{I(1)}] - \kappa[y_I] - \sum_{(1,i) \in V \subset I} \kappa[y_V] \tilde{W}[y^{I \setminus V}]. \tag{III.9}
\]

Then the claim \( \tilde{W}[y^I] = :y^I:\); follows by straightforward induction in \(|I|\): Case \(|I| = 0\) was proven above. Suppose \( I \neq \emptyset \) and that \( \tilde{W}[y^{I'}] = :y^{I'}:\); whenever \(|I'| < |I|\). Then the induction assumption and Theorem III.2 can be used to evaluate the expectation of the right hand side of (III.9), implying \( \mathbb{E}[\tilde{W}[y^I]] = 0 \). By (III.8), the polynomial \( \tilde{W}[y^I] \) also satisfies the third defining condition of the Wick polynomials, equation (II.4). Hence, \( \tilde{W}[y^I] = :y^I:\); which completes the induction step. Then (III.9) implies the first identity in (III.7) and the second identity is found by a relabeling of the summation variable. Hence, also item 4 follows.

To prove (III.8), consider some \( I \neq \emptyset \). In the definition of \( \tilde{W}[y^I] \), we can express the derivatives of \( y^U, U \subset I \), as in (III.4). This shows that

\[
\partial_y \tilde{W}[y^I] = \sum_{U \subset I} \sum_{k=1}^{|I|} \#(i_k = j) \tilde{W}[y^{I(k)}] \sum_{\pi \in \mathcal{P}(I \setminus U)} (-1)^{|\pi|} \prod_{A \in \pi} \kappa[y_A]
\]

\[
= \sum_{k=1}^{|I|} \#(i_k = j) \sum_{V \subset I(k)} y^V \sum_{\pi \in \mathcal{P}(I(k) \setminus V)} (-1)^{|\pi|} \prod_{A \in \pi} \kappa[y_A]
\]

\[
= \sum_{k=1}^{|I|} \#(i_k = j) \tilde{W}[y^{I(k)}]. \tag{III.10}
\]

Therefore, (III.8) holds.

To prove (III.9), denote \( x := (1,i_1) \) and \( I' := \tilde{I}(1) \). We first split the definition into two parts as follows:

\[
\tilde{W}[y^I] = \sum_{U \subset I} \#(x \in U) y^U \sum_{\pi \in \mathcal{P}(I \setminus U)} (-1)^{|\pi|} \prod_{A \in \pi} \kappa[y_A]
\]

\[
+ \sum_{U \subset I} \#(x \notin U) y^U \sum_{\pi \in \mathcal{P}(I \setminus U)} (-1)^{|\pi|} \prod_{A \in \pi} \kappa[y_A]. \tag{III.11}
\]

Following a reasoning similar to (III.10), we find that the first term in the sum on the right hand side is equal to \( y_I \tilde{W}[y^{I'}] \). The second term is equal to

\[
\sum_{U \subset I} \#(x \notin U) y^U \sum_{\pi \in \mathcal{P}(I \setminus U)} (-1)^{|\pi|} \left( \prod_{x \notin A \in \pi} \kappa[y_A] \right) \kappa[y_V] |_{x \in V \subset I}.
\]

\[
= - \sum_{U \subset I} \#(x \notin U) y^U \sum_{x \in V \subset I \setminus U} \kappa[y_V] \sum_{\pi \in \mathcal{P}(I \setminus U \setminus V)} (-1)^{|\pi|} \prod_{A \in \pi} \kappa[y_A]
\]

\[
= - \sum_{x \in V \subset I} \kappa[y_V] \sum_{U \subset I \setminus V} y^U \sum_{\pi \in \mathcal{P}(I \setminus U \setminus V)} (-1)^{|\pi|} \prod_{A \in \pi} \kappa[y_A]
\]

\[
= - \sum_{x \in V \subset I} \kappa[y_V] \tilde{W}[y^{I \setminus V}]. \tag{III.12}
\]
Therefore, (III.11) implies that also (III.9) holds. This completes the proof of the Theorem. □

Example III.6 Written is terms of cumulants, the Wick polynomials of lowest order are

$$y : = y - \kappa (y) ,$$

$$y_1 y_2 : = y_1 y_2 - \kappa (y_1 , y_2) - \kappa (y_1) y_2 - \kappa (y_2) y_1 + \kappa (y_1) \kappa (y_2) ,$$

$$y_1 y_2 y_3 : = y_1 y_2 y_3 - \kappa (y_1 , y_2 , y_3) + \kappa (y_1 , y_2) \kappa (y_3) + \kappa (y_1 , y_3) \kappa (y_2) + \kappa (y_2 , y_3) \kappa (y_1)$$

$$- \kappa (y_1) \kappa (y_2) \kappa (y_3) - \kappa (y_1 , y_2) y_3 - \kappa (y_1 , y_3) y_2 - \kappa (y_2 , y_3) y_1$$

$$+ \kappa (y_1) \kappa (y_2) \kappa (y_3) + \kappa (y_1) \kappa (y_3) y_2 + \kappa (y_2) \kappa (y_3) y_1$$

$$- \kappa (y_1) y_2 y_3 - \kappa (y_2) y_1 y_3 - \kappa (y_3) y_1 y_2 .$$

(III.13)

Proposition III.7 The Wick polynomials are multilinear, i.e., if \( \alpha , \beta \) are constants such that \( y_j = \alpha y_i + \beta y_{i'} \) for some \( j , i , i' \in J \), then, whenever \( I \) and \( k \) are such that \( i_k = j \), we have

$$y : = \alpha y^\beta + \gamma ; + \beta y^\alpha + \delta : .$$

Proof: The claim follows using multilinearity of cumulants in the representation formula (III.6). □

The following result extends the earlier theorem and shows that multiple application of Wick polynomial replacements continues to simplify the moments-to-cumulants formula by removing all terms with any internal clusters.

Proposition III.8 Assume that the measure \( \mu \) has all moments of order \( N \), i.e., suppose that \( E[|y|^N] < \infty \) for all \( I \in \mathcal{I} \) with \(|I| \leq N \). Suppose \( L \geq 1 \) is given and consider a collection of \( L + 1 \) index sequences \( J' , J_\ell \in \mathcal{I} , \ell = 1 , \ldots , L \), such that \(|J'| + \sum \ell |J_\ell| \leq N \). Then for \( I : = \sum_{\ell=1}^L J_\ell + J' \) (with the implicit identification of \( J_\ell \) and \( J' \) with the set of its labels in \( I \)) we have

$$E \left[ \prod_{\ell=1}^L y_{J_\ell} \cdot y_{J'} \right] = \sum_{\pi \in P(I)} \prod_{A \in \pi} (\kappa |y_A| \mathbb{I} (A \not\subset J_\ell \ \forall \ell)) .$$

(III.14)

Proof: We proceed via a double induction: the first induction is over \( L \) and the second induction is over \(|J_L|\), i.e., the length of the last index sequence. The case \( L = 1 \) follows directly from Theorem III.2. Now we assume as induction hypothesis of the first level induction that \( L \geq 2 \) and

$$E \left[ \prod_{\ell=1}^{L-1} y_{J_\ell} \cdot y_{J'} \right] = \sum_{\pi \in P(\sum_{\ell=1}^{L-1} J_\ell + J')} \prod_{A \in \pi} (\kappa |y_A| \mathbb{I} (A \not\subset J_\ell \ \forall \ell < L)) .$$

(III.15)

Then we consider the second induction over \(|J_L| =: m\). For \( m = 0 \), we have \( y_{J_\ell} : = 1 \) and thus then the induction hypothesis (III.15) directly implies (III.14). As induction step of the second level hypothesis we take that, for fixed \( L \), equation (III.14) holds for all \(|J_L| < m\). Then, if \(|J_L| = m\), we can use (III.2), (III.15) and the second level induction hypothesis to...
justify the following argument analogous to the one used in the proof of Theorem III.2

\[
\mathbb{E} \left[ \prod_{\ell=1}^{L} y^{J_{\ell}}; y^{J} \right] = \mathbb{E} \left[ \prod_{\ell=1}^{L-1} y^{J_{\ell}}; y^{J_{L} \cup J} \right] - \sum_{\emptyset \neq E \subset J_{L}} \mathbb{E}[y^{E}] \mathbb{E} \left[ \prod_{\ell=1}^{L-1} y^{J_{\ell}}; y^{J_{L} \setminus E}; y^{J} \right] \\
= \sum_{\pi \in \mathcal{P}(\sum_{\ell=1}^{L-1} J_{\ell} + J_{L} \cup J')} \prod_{A \in \pi} (\kappa[y_{A}] \mathbb{1}(A \not\subset J_{L} \forall \ell)) - \sum_{\emptyset \neq E \subset J_{L}} \sum_{\pi_{0} \in \mathcal{P}(E)} \prod_{A \in \pi_{0}} \kappa[y_{A}] \\
\times \sum_{\pi_{1} \in \mathcal{P}(\sum_{\ell=1}^{L-1} J_{\ell} + (J_{L} \setminus E) + J')} \prod_{A \in \pi_{1}} (\kappa[y_{A}^{E}] \mathbb{1}(A \not\subset J_{L} \forall \ell)) \\
= \sum_{\pi \in \mathcal{P}(I)} \prod_{A \in \pi} (\kappa[y_{A}] \mathbb{1}(A \not\subset J_{L} \forall \ell)) \\
+ \sum_{\pi \in \mathcal{P}(I)} \mathbb{1}(\exists A \in \pi \text{ s.t. } A \subset J_{L}) \prod_{A \in \pi} (\kappa[y_{A}] \mathbb{1}(A \not\subset J_{L} \forall \ell)) \\
- \sum_{\pi \in \mathcal{P}(I)} \mathbb{1}(\cup \{A \in \pi | A \subset J_{L}\} \neq \emptyset) \prod_{A \in \pi} (\kappa[y_{A}] \mathbb{1}(A \not\subset J_{L} \forall \ell)) \\
= \sum_{\pi \in \mathcal{P}(I)} \prod_{A \in \pi} (\kappa[y_{A}] \mathbb{1}(A \not\subset J_{L} \forall \ell)) . \tag{III.16}
\]

This completes the induction step and hence also the proof. \qed

IV. CUMULANTS AND WICK POLYNOMIALS AS DYNAMICAL VARIABLES

To study the state of a random system, it is often better to use the cumulants rather than the moments of the random variables. For instance, if \( y, z \) are independent random variables we have \( \mathbb{E}[y^n z^m] = \mathbb{E}[y^n] \mathbb{E}[z^m] \), which is typically nonzero, whereas the corresponding cumulant is zero whenever both \( n, m \neq 0 \). Hence, for systems where two regions become “asymptotically independent” (for instance, for sufficiently mixing stochastic processes), it is the cumulants, not moments, which will vanish in the limit.

To have a concrete example, let us consider a random lattice field \( y_x, x \in \mathbb{Z}^d \), which is spatially sufficiently strongly mixing. Then, for instance, \( \kappa(y_0, y_1) \) to 0 if the distance of the index set \( I \subset \mathbb{Z}^d \) from the origin becomes unbounded. Often in the applications the mixing is so strong that the cumulants, in this case also called connected correlation functions, become “\( \ell_1 \)-clustering”: for any order \( n \) one requires that \( \sup_{x \in \mathbb{Z}^d} \sum_{I \in (\mathbb{Z}^d)^n} |\kappa(y_x, y_I)| < \infty \).

Naturally, such a property should then not be expected to hold for standard correlation functions \( \mathbb{E}[y^I] \), apart from some degenerate examples.

In addition to being mixing, the random fields found in the applications are often translation invariant. This means in particular that all moments \( \mathbb{E}[y^I] \) remain invariant if every index in \( I \) is translated by a fixed amount, i.e., \( \mathbb{E}[y^{I(x)}] = \mathbb{E}[y^I] \) for every \( x \in \mathbb{Z}^d \) if we set \( I(x)_{\ell} := \ell_x - x \). If the system is both \( \ell_1 \)-clustering and translation invariant, the cumulants of the Fourier transformed field \( \tilde{y}_k := \sum_{x \in \mathbb{Z}^d} e^{-2\pi ik \cdot x} y_x \), \( k \) indexed by the d-torus \( \mathbb{T}^d \), satisfy

\[
\kappa[\tilde{y}(k_1, k_2, \ldots, k_n)] = \delta\left(\sum_{\ell=1}^{n} k_{\ell}\right) \tilde{F}_n(k_1, k_2, \ldots, k_n) , \tag{IV.1}
\]

where “\( \delta \)” denotes the Dirac delta distribution and the arithmetic on \( \mathbb{T}^d \) is defined via periodic identification. Here \( \tilde{F}_n \) denotes the Fourier transform of \( F_{n}(X) := \mathbb{1}(X_1 = 0)\kappa(X_X) \), \( X \in (\mathbb{Z}^d)^n \), and for \( \ell_1 \)-clustering measures \( \tilde{F}_n \) is a uniformly bounded continuous function of \( k \). Therefore, although the cumulants are singular, their singularity structure is simple, entirely encoded in the \( \delta \)-multiplier. In contrast, by the moments-to-cumulants formula,
then for \( I := (1, 2, \ldots, n) \) and any \( k \in (\mathbb{I}^d)^n \)

\[
\mathbb{E}[\hat{y}^k_I] = \sum_{\pi \in \mathcal{P}(I)} \prod_{A \in \pi} \left( \delta \left( \sum_{j \in A} k_j \right) \hat{F}_{|A|}(k_A) \right),
\]

which has ever more complicated singularity structure as the order of the moment is increased. (The above discussion can be made mathematically rigorous by replacing the infinite lattice by a periodic \( d \)-dimensional lattice.)

Hence, for stochastic processes which lead towards a state which is mixing and translation invariant, it seems better to focus on the time-evolution of cumulants instead of the corresponding moments. However, it is not immediately clear how to avoid resorting to the moments as a middle step. It turns out that using Wick polynomials instead of “bare” monomials to define the time-evolution helps in achieving this goal. Recall that any monomial term \( y^n \) can always be expanded in terms of Wick polynomials using (III.5), albeit with state-dependent coefficients.

To have a concrete example how this could work in practice, we consider in the following the case of deterministic evolution with random initial data. Explicitly, we assume that the system is described by random variables \( y_j(t) \), where \( j \) belongs to a fixed (finite) index set \( J \), and \( t \geq 0 \) denotes time. The initial values \( y_j(0) \) are random with a joint distribution \( \mu_0 \), and for each realization of \( y(0) \) we assume that the values at later times \( t > 0 \) are determined from the solution to the differential equation

\[
\partial_t y_j(t) = \sum_{I \in \mathcal{I}_j} M_j^I(t) : y(t)^I :,
\]

where the functions \( M_j^I(t) \) are “interaction amplitudes” from the \( I \)-th Wick polynomial of \( y(t) \) to \( y_j(t) \). For each \( j \in J \), the set \( \mathcal{I}_j \) collects those \( I \in \mathcal{I} \) which have a nonzero amplitude, i.e., \( M_j^I(t) \neq 0 \) for some \( t > 0 \). For simplicity, we assume here that \( \mathcal{I}_j \) is finite and that the amplitudes \( M_j^I(t) \) are some fixed functions of time. (They typically might depend on the cumulants of \( y(t) \), but this is not relevant for the discussion below: it suffices that they are not random variables.)

We present a concrete example of such a dynamical system in Appendix B where we show how the evolution of \( N \) classical particles interacting via a polynomial interaction potential can be described by a system of this type assuming a known random distribution of initial positions and momenta. Another explicit example is given in Section V.

The usefulness of representing the dynamics in the form (IV.3) becomes apparent when looking at the evolution of cumulants. To avoid technical complications, let us suppose that the joint exponential moments of \( y(t) \) exist and are continuously differentiable and uniformly bounded functions of \( t \). For simplicity, we assume here that \( \mathcal{I}_j \) is finite and that the moments as a middle step. It turns out that using Wick polynomials instead of “bare” monomials to define the time-evolution helps in achieving this goal. Recall that any monomial term \( y^n \) can always be expanded in terms of Wick polynomials using (III.5), albeit with state-dependent coefficients.

To have a concrete example how this could work in practice, we consider in the following the case of deterministic evolution with random initial data. Explicitly, we assume that the system is described by random variables \( y_j(t) \), where \( j \) belongs to a fixed (finite) index set \( J \), and \( t \geq 0 \) denotes time. The initial values \( y_j(0) \) are random with a joint distribution \( \mu_0 \), and for each realization of \( y(0) \) we assume that the values at later times \( t > 0 \) are determined from the solution to the differential equation

\[
\partial_t y_j(t) = \sum_{I \in \mathcal{I}_j} M_j^I(t) : y(t)^I :,
\]

where the functions \( M_j^I(t) \) are “interaction amplitudes” from the \( I \)-th Wick polynomial of \( y(t) \) to \( y_j(t) \). For each \( j \in J \), the set \( \mathcal{I}_j \) collects those \( I \in \mathcal{I} \) which have a nonzero amplitude, i.e., \( M_j^I(t) \neq 0 \) for some \( t > 0 \). For simplicity, we assume here that \( \mathcal{I}_j \) is finite and that the amplitudes \( M_j^I(t) \) are some fixed functions of time. (They typically might depend on the cumulants of \( y(t) \), but this is not relevant for the discussion below: it suffices that they are not random variables.)

We present a concrete example of such a dynamical system in Appendix B where we show how the evolution of \( N \) classical particles interacting via a polynomial interaction potential can be described by a system of this type assuming a known random distribution of initial positions and momenta. Another explicit example is given in Section V.

The usefulness of representing the dynamics in the form (IV.3) becomes apparent when looking at the evolution of cumulants. To avoid technical complications, let us suppose that the joint exponential moments of \( y(t) \) exist and are continuously differentiable and uniformly bounded functions of \( t \). For simplicity, we assume here that \( \mathcal{I}_j \) is finite and that the moments as a middle step. It turns out that using Wick polynomials instead of “bare” monomials to define the time-evolution helps in achieving this goal. Recall that any monomial term \( y^n \) can always be expanded in terms of Wick polynomials using (III.5), albeit with state-dependent coefficients.

To have a concrete example how this could work in practice, we consider in the following the case of deterministic evolution with random initial data. Explicitly, we assume that the system is described by random variables \( y_j(t) \), where \( j \) belongs to a fixed (finite) index set \( J \), and \( t \geq 0 \) denotes time. The initial values \( y_j(0) \) are random with a joint distribution \( \mu_0 \), and for each realization of \( y(0) \) we assume that the values at later times \( t > 0 \) are determined from the solution to the differential equation

\[
\partial_t y_j(t) = \sum_{I \in \mathcal{I}_j} M_j^I(t) : y(t)^I :,
\]

where the functions \( M_j^I(t) \) are “interaction amplitudes” from the \( I \)-th Wick polynomial of \( y(t) \) to \( y_j(t) \). For each \( j \in J \), the set \( \mathcal{I}_j \) collects those \( I \in \mathcal{I} \) which have a nonzero amplitude, i.e., \( M_j^I(t) \neq 0 \) for some \( t > 0 \). For simplicity, we assume here that \( \mathcal{I}_j \) is finite and that the amplitudes \( M_j^I(t) \) are some fixed functions of time. (They typically might depend on the cumulants of \( y(t) \), but this is not relevant for the discussion below: it suffices that they are not random variables.)

We present a concrete example of such a dynamical system in Appendix B where we show how the evolution of \( N \) classical particles interacting via a polynomial interaction potential can be described by a system of this type assuming a known random distribution of initial positions and momenta. Another explicit example is given in Section V.

The usefulness of representing the dynamics in the form (IV.3) becomes apparent when looking at the evolution of cumulants. To avoid technical complications, let us suppose that the joint exponential moments of \( y(t) \) exist and are continuously differentiable and uniformly bounded functions of \( t \). For simplicity, we assume here that \( \mathcal{I}_j \) is finite and that the moments as a middle step. It turns out that using Wick polynomials instead of “bare” monomials to define the time-evolution helps in achieving this goal. Recall that any monomial term \( y^n \) can always be expanded in terms of Wick polynomials using (III.5), albeit with state-dependent coefficients.

To have a concrete example how this could work in practice, we consider in the following the case of deterministic evolution with random initial data. Explicitly, we assume that the system is described by random variables \( y_j(t) \), where \( j \) belongs to a fixed (finite) index set \( J \), and \( t \geq 0 \) denotes time. The initial values \( y_j(0) \) are random with a joint distribution \( \mu_0 \), and for each realization of \( y(0) \) we assume that the values at later times \( t > 0 \) are determined from the solution to the differential equation

\[
\partial_t y_j(t) = \sum_{I \in \mathcal{I}_j} M_j^I(t) : y(t)^I :,
\]

where the functions \( M_j^I(t) \) are “interaction amplitudes” from the \( I \)-th Wick polynomial of \( y(t) \) to \( y_j(t) \). For each \( j \in J \), the set \( \mathcal{I}_j \) collects those \( I \in \mathcal{I} \) which have a nonzero amplitude, i.e., \( M_j^I(t) \neq 0 \) for some \( t > 0 \). For simplicity, we assume here that \( \mathcal{I}_j \) is finite and that the amplitudes \( M_j^I(t) \) are some fixed functions of time. (They typically might depend on the cumulants of \( y(t) \), but this is not relevant for the discussion below: it suffices that they are not random variables.)

We present a concrete example of such a dynamical system in Appendix B where we show how the evolution of \( N \) classical particles interacting via a polynomial interaction potential can be described by a system of this type assuming a known random distribution of initial positions and momenta. Another explicit example is given in Section V.

The usefulness of representing the dynamics in the form (IV.3) becomes apparent when looking at the evolution of cumulants. To avoid technical complications, let us suppose that the joint exponential moments of \( y(t) \) exist and are continuously differentiable and uniformly bounded functions of \( t \). For simplicity, we assume here that \( \mathcal{I}_j \) is finite and that the moments as a middle step. It turns out that using Wick polynomials instead of “bare” monomials to define the time-evolution helps in achieving this goal. Recall that any monomial term \( y^n \) can always be expanded in terms of Wick polynomials using (III.5), albeit with state-dependent coefficients.
In this case, determining the evolution of expectation values of all multiplications of two Wick products, \(\langle y(t)\cdot y(t)\rangle\), where both \(I_1, I_2\) are non-empty, would also yield a solution to the evolution of cumulants.

We can now obtain a closed cumulant evolution hierarchy using (IV.6) and Theorems III.2 and III.8. First, note that for any \(I \neq \emptyset\) and \(j \in J\) we have

\[
\langle y(t)\cdot y_j(t)\rangle = \langle y(t)\rangle = \kappa[y(t)_{I+(j)}],
\]

since in this case there is exactly one non-internal cluster, the entire set \(I+(j)\). In addition, if \(I = \emptyset\), we clearly have \(\langle y(t)^I\cdot y(t)\rangle = 1(I' = \emptyset)\). Therefore, the first two cumulants satisfy, for arbitrary \(j, j' \in J\),

\[
\begin{align*}
\partial_t \kappa[y(t)] & = \mathbb{1}(0 \in I) \mathcal{M}_0(t), \\
\partial_t \kappa[y(t)] & = \sum_{\emptyset \neq I \in \mathcal{I}_i} \mathcal{M}^I_1(t) \kappa[y(t)_{I+(j')}] + \sum_{\emptyset \neq I \in \mathcal{I}_i} \mathcal{M}^I_1(t) \kappa[y(t)_{I+(j)}].
\end{align*}
\]

For higher order cumulants, with \(|I'| \geq 3\), the equation typically becomes nonlinear; we then have

\[
\begin{align*}
\partial_t \kappa[y(t)] & = \sum_{I' \neq \emptyset} \sum_{j \in I} \mathbb{1}(j \in I) \mathcal{M}^{I(j)}_1(t) \kappa[y(t)_{I+(j')}] \\
& \quad + \sum_{I' \neq \emptyset} \sum_{j \in I} \left[ \mathcal{M}^{I(j)}_1(t) \langle y(t)^I\cdot y(t)\rangle \right],
\end{align*}
\]

We have separated here the terms with \(|I| = 1\) to show how they operate linearly on the cumulants of order \(|I'|\) (note that \(\kappa[y(t)_{I+(I'\setminus j)}] = \kappa[y(t)_{I'}\rangle\) for the sequence \(I'\) which is obtained from \(I\) by replacing \(j\) with \(j\)). In the final sum, both \(|I|\) and \(|I' \setminus i|\) are greater than one, so it has a cumulant expansion

\[
\langle y(t)^I\cdot y(t)^{I'\setminus i}\rangle = \sum_{\pi \in \mathcal{P}(I+(I'\setminus i))} \prod_{A \in \pi} \kappa[y(t)_A] \mathbb{1}(A \cap I \neq \emptyset, A \cap (I' \setminus i) \neq \emptyset),
\]

i.e., all clusters have to contain at least one element from both sequences. In particular, it cannot contain any singlets, i.e., it does not depend on any of \(\kappa[y_j(t)]\), \(j \in J\). Let us also point out that since in these terms \(|I| + |I' \setminus i| > |I'|\), any linear term is necessarily of higher order. In particular, this means that lower order cumulants can appear only in nonlinear combinations.

Instead of studying the full cumulant hierarchy, one can also use evolution estimates for the Wick polynomials. The situation often encountered in the applications is that the properties of the initial measure are fairly well known, whereas very little a priori control exists for the time-evolved measure. In such a case, one can use the above result to obtain a perturbation expansion by applying the fundamental theorem of calculus. With the shorthand notation \(y := y(0)\) we have

\[
\begin{align*}
\kappa[y(t)] & = \kappa[y] + \int_0^t ds \sum_{i \in I'} \sum_{f \in \mathcal{I}_i} \mathcal{M}^I_f(s) \langle y(s)^I\cdot y(s)^{I'\setminus i}\rangle \\
& = \kappa[y] + \int_0^t ds \sum_{i \in I'} \sum_{f \in \mathcal{I}_i} \langle y^I\cdot y^{I'\setminus i}\rangle \int_0^t ds \mathcal{M}^I_f(s) \\
& \quad + \sum_{i \in I'} \int_0^t ds \partial_s \langle y(s)^I\cdot y(s)^{I'\setminus i}\rangle \int_0^t ds \mathcal{M}^I_f(s),
\end{align*}
\]

where we have applied Fubini’s theorem to the final integral. This type of expansion could be helpful if the coefficients \(\int_0^t ds \mathcal{M}^I_f(s)\) behave better than \(\mathcal{M}^I_f(t)\), such as in the presence of fast oscillations. Further iterations of this procedure, using either the above cumulant
hierarchy or any of the Wick polynomial hierarchies below, would then yield an expansion of $\kappa[y(t)_I]$ in terms of the expectations at time $t = 0$ and the time dependent amplitudes $M^I(t)$. This is particularly useful if all $M^I(t)$ are small, since each iteration adds one more such factor.

Let us conclude this section by deriving recursion formulae for the products of Wick polynomials. As mentioned earlier, these could then be used instead of the direct cumulant hierarchy to study the time-evolution of the cumulants. For this, it would suffice to study $\langle y(s) : y(s) \rangle$ appearing in (IV.6), but typically the products of two terms do not satisfy a closed evolution equation and a full hierarchy will be needed. Let us begin with the evolution equation for $y(t)$. For any deterministic evolution process, we can obtain a fairly compact evolution equation by treating the time-derivative $\partial_t y_I$ as a new random variable:

$$\partial_t :y(t)^I: = \partial_t \langle \partial_t G(\lambda; y(t)) \rangle \big|_{\lambda=0} = \sum_{i \in I} :\partial_t y_i(t)\rangle y(t)^{I \setminus i}:.$$  \hspace{1cm} (IV.13)

The form is analogous to the standard Leibniz rule. For products of Wick polynomials, we have thus:

$$\partial_t \prod_{k=1}^n y_k(t)^{I_k} = \sum_{k=1}^n \sum_{i \in I_k} y_k(t)^{I_k \setminus i} :\partial_t y_i(t)\rangle y(t)^{I_k \setminus i}: \prod_{k' \neq k} y(t)^{I_{k'}}.$$  \hspace{1cm} (IV.14)

Assuming (IV.3) and using multilinearity, we then obtain the following equation involving “nested Wick products”:

$$\partial_t \mathbb{E}\left[ \prod_{k=1}^n y_k(t)^{I_k} : \right] = \sum_{k=1}^n \sum_{i \in I_k} \sum_{I_{k+1}} M^I_k(t) \mathbb{E}\left[ \langle y(t)^I \rangle : y(t)^{I_k \setminus i} : y(t)^{I_{k+1}} : \right].$$ \hspace{1cm} (IV.15)

The formula (IV.15) is appealing in its simplicity but it does not directly lead to closed hierarchy of equations. This can be achieved by expanding the nested product in terms of cumulants and Wick products. For this, let us note that by (III.7) and the observation made after (IV.7), we have for any $I'$

$$:\partial_t y_i(t)\rangle y(t)^{I'} : = \partial_t y_i(t) : y(t)^{I'} : = \sum_{U \subset I'} \mathbb{E}[\partial_t y_i(t) : y(t)^{U'} : y(t)^{I \setminus U'} :].$$ \hspace{1cm} (IV.16)

Therefore, whenever (IV.3) holds, we find that

$$\partial_t \mathbb{E}\left[ \prod_{k=1}^n y_k(t)^{I_k} : \right] = \sum_{k=1}^n \sum_{i \in I_k} \sum_{I_{k+1}} M^I_k(t) \mathbb{E}\left[ \langle y(t)^I \rangle : y(t)^{I_k \setminus i} : \prod_{k' \neq k} y(t)^{I_{k'}} : \right] \mathbb{E}\left[ \langle y(t)^{I_{k+1}} \rangle : y(t)^{I_{k+1}} : \right].$$ \hspace{1cm} (IV.17)

This forms a closed hierarchy of evolution equations for the collection of all expectation values of the type $\mathbb{E}\left[ \prod_{k=1}^n y(t)^{I_k} : \right]$. A second alternative for the hierarchy follows from the observation that if $y(t)$ and $z(t)$ are two processes which start with independent, identically distributed initial data, then at any later moment they are also independent and identically distributed and hence

$$\partial_t G(\lambda; y(t)) = \mathbb{E}_z \left[ \lambda \cdot \partial_t z(t) G(\lambda; z(t)) \right] - \mathbb{E}_z \left[ \lambda \cdot \partial_t G(\lambda; z(t)) \right].$$ \hspace{1cm} (IV.18)

where in the second equality we have used $\mathbb{E}_z [G(\lambda; z(t))] = 1$. Consider then the product measure for the processes $y, z$ and let $G'$ denote the corresponding Wick polynomial
generating function. Since by Fubini’s theorem then $\mathbb{E}_{y,z}[e^{\lambda \cdot (y(t)+z(t))}] = \mathbb{E}_y[e^{\lambda \cdot y(t)}]^2$, now $G(\lambda; y(t))G(\lambda; z(t)) = G'(\lambda; y(t) + z(t))$. Hence, for dynamics satisfying (IV.3)

$$
\partial_t \mathbb{E}\left[ \prod_{k=1}^n (y(t))^{I_k} : \right] = \sum_{k=1}^n \sum_{i \in I_k} \sum_{t \in I_x} M_i^k(t) \mathbb{E}_{y,z} \left[ (y(t)^I_- : z(t)^I_- : (y(t) + z(t))^{I_k \setminus i} : \prod_{k' \neq k} (y(t)^I_{k'}) \right]. \tag{IV.19}
$$

Let us point out that the earlier expression in (IV.17) follows from the above one if we expand the power $(y(t) + z(t))^{I_k \setminus i}$ and then use the fact that the joint measure is a product measure. The formula does not yet yield a closed hierarchy but the following generalization does so: if $z_{k,\ell}(t)$ are processes such that their joint initial distribution is given, then

$$
\partial_t \mathbb{E}\left[ \prod_{k=1}^n : (\sum_{\ell} z_{k,\ell}(t))^{I_k} : \right] = \sum_{k=1}^n \sum_{i \in I_k} \sum_{t \in I_x} M_i^k(t)
\times \mathbb{E}_k' \left[ (z_{k,\ell}(t)^I_- : z_{k,\ell}'(t)^I_- : (\sum_{\ell'} (z_{k,\ell'}(t) + z_{k,\ell'}(t)))^{I_k \setminus i} : \prod_{k' \neq k} (\sum_{\ell'} z_{k,\ell'}(t))^{I_{k'}} \right], \tag{IV.20}
$$

where $\mathbb{E}_k'$ refers to a measure where $z_{k,\ell}$ for each $\ell$ has been independently duplicated in the $z_{k,\ell}'$-process.

A possible benefit of this formulation could be when $z_{k,\ell}(t)$ have mean zero and are independent for all $\ell$. Then the central limit theorem governs the behavior of $\sum_{\ell} z_{k,\ell}$ when there are many terms in the sum. Therefore, it could be of help in controlling the otherwise difficult case where one has performed many iterations starting from (IV.19).

V. KINETIC THEORY OF THE DISCRETE NLS EQUATION REVISITED

In this section we apply the previous Wick polynomial techniques to the discrete nonlinear Schrödinger equation. This example is chosen since it has a particularly simple, but nontrivial, Wick expansion of the evolution equation. In addition, we can then rely on the rigorous results and known properties from an earlier work on the model$^9$. We focus on the kinetic theory of the model on the “kinetic” time-scale which is $O(\lambda^{-2})$ for small nonlinear couplings $\lambda$. We refer to Ref. 12 for more details about such kinetic scaling limits of lattice systems.

We begin by going through the derivation of the Boltzmann-Peierls equation in the spatially homogeneous case, and we show how the task is simplified by using the Wick expanded dynamics and the cumulant hierarchy, as explained in Section IV. For an explicit comparison to a more standard derivation using the moment hierarchy, we refer to Ref. 13, Appendix A. In particular, we wish to point out the larger number of terms appearing in the moment version, most of which contain “tadpole graphs”. The tadpoles lead to cancellations in the phase factor so that some integration variables never appear in the oscillatory factor in the integrand, and hence the tadpole terms have bad decay properties in time. These terms will be absent from the expansion below, and in fact they do not occur in the corresponding computation$^{14}$ (Section 4.2.1) in the model considered in Ref. 13 either.

As in the earlier derivations, we only consider terms which would be present in the simplified case of Gaussian initial data. We give an example in Appendix C which highlights the mechanism leading to suppression of the additional effects of non-Gaussian initial data in the kinetic scaling limit.

For this particular setup, it is easy to find dynamical variables whose evolution equation does not have a linear part. This is an important simplification since it negates a term which becomes rapidly oscillating on the kinetic time-scale, having a divergent frequency
\(O(\lambda^{-1})\) in the kinetic scaling limit. The effect becomes apparent when looking at field time-correlations instead of the evolution of equal time cumulants. We discuss the issue in more detail in Section VI A.

In Ref. 9, the initial data is taken to be given by a thermal Gibbs measure which is stationary under translations both in space and in time. The Gibbs measure has also been proven\textsuperscript{15} to be \(\ell_1\)-clustering, with some additional mild assumptions on the dispersion relation. We do not assume the initial data to be time-stationary here, but the computations in this section require space translation invariance. The spatially inhomogeneous case is technically substantially more complicated, and we discuss it only briefly in Section VI A.

The results in this section are derived in the spirit of standard perturbation theory and focus solely on evolution on short kinetic time scales, \(t = \tau \lambda^{-2}\) with \(0 < \tau \ll 1\). It is however possible to apply the cumulant hierarchy differently, leading to equations which do not require taking a scaling limit. We conclude the study of the DNLS model in Section VI B by proposing a reformulation of the problem which leads to Boltzmann type evolution equations which could be accurate also for times longer than \(O(\lambda^{-2})\). The discussion is not completely mathematically rigorous, but we propose a scheme which could be used to this end under some natural conditions about the time-evolved state.

The discrete NLS equation on the lattice \(\mathbb{Z}^d\) deals with functions \(\psi : \mathbb{R} \times \mathbb{Z}^d \to \mathbb{C}\) which satisfy

\[
i\partial_t \psi_t(x) = \sum_{y \in \mathbb{Z}^d} \alpha(x - y)\psi_t(y) + \lambda|\psi_t(x)|^2\psi_t(x). \tag{V.1}
\]

Here the function \(\alpha : \mathbb{Z}^d \to \mathbb{R}\) is called the hopping amplitude and we assume that it is symmetric, \(\alpha(x) = \alpha(-x)\), and exponentially decreasing. The parameter \(\lambda > 0\) is considered to be small, and in the kinetic scaling limit we take \(\lambda \to 0\) and \(t = \tau \lambda^{-2}\) with \(\tau > 0\) fixed. The initial field \(\psi_0\) is assumed to be random, bounded on finite subsets of the lattice, and to have an \(\ell_1\)-clustering distribution. We aim at controlling the moments of the random variables \(\psi_t(x)\) and \(\psi_t(x)^*\) which we label using \(\psi_t(x, +1) := \psi_t(x)\) and \(\psi_t(x, -1) := \psi_t(x)^*\).

Since we do not assume that \(\psi_0\) is \(\ell_2\)-summable, even the (almost sure) existence of solutions to (V.1) becomes an issue. To our knowledge, it has not been proven for the above setup, and most likely, some additional assumptions about the increase of the values of the initial field at infinity are needed for proper existence theory. However, these problems can be easily avoided by replacing the infinite lattice \(\mathbb{Z}^d\) by a finite lattice with periodic boundary conditions\textsuperscript{9}. This would merely result in replacing the lattice, the Fourier space and transform, and the associated \(\delta\)-functions by their finite lattice counterparts. Since even then the final limits cannot be rigorously controlled, we opt here for some additional formality in the discussion, but with less complicated formulae to deal with.

For technical simplicity, here we also only consider initial data which are “gauge invariant”: we will always suppose \(\psi_0(x)\) has the same distribution as \(e^{i\theta}\psi_0(x)\) for any \(\theta \in [0, 2\pi]\). In fact, this transformation commutes with the time evolution, i.e., if the initial field is changed from \(\psi_0\) to \(e^{i\theta}\psi_0\), the time-evolved field will change from \(\psi_t\) to \(e^{i\theta}\psi_t\). In particular, also the field \(\psi_t\) will then be gauge invariant. The main reason for insisting on this assumption is that it will automatically force many cumulants to be zero and hence simplify the combinatorics associated with the cumulant hierarchy. Gauge invariance implies that a moment is zero unless it has the same number of \(\psi\) and \(\psi^*\) factors, even when the fields are evaluated at different times. Hence, it implies that every odd moment of the fields is zero and hence also every odd cumulant. Similarly, we see that even cumulants are also zero if they concern a different number of \(\psi^*\) and \(\psi\) variables.

For instance, we find using (III.6) that for any gauge invariant state and any \(a_j := \psi_{t_j}(x_j, \sigma_j), j = 1, 2, 3,\)

\[
a_1a_2a_3 = a_1a_2a_3 - E[a_1a_2]a_3 - E[a_1]a_2a_3 - E[a_2]a_3a_1. \tag{V.2}
\]

This is the definition of the “pairing truncation operation” \(\hat{P}\) given in Ref. 9, Lemma 3.2. Applying the truncation operation in the evolution equation was one of the key changes to
the standard perturbation theory which allowed the rigorous analysis in Ref. 9. With the
benefit of hindsight, we can now identify it as a Wick contraction of the random variables.

Under the above assumptions and using (III.5), we find that (V.1) is equivalent to the
following Wick contracted evolution equation

\[ i\partial_t \psi_t(x) = \sum_{y \in \mathbb{Z}^d} \alpha(x - y) \psi_t(y) + 2\lambda \mathbb{E}[\psi_t(x)^* \psi_t(x)] \psi_t(x) + \lambda \psi_t(x) \psi_t(x) \psi_t(x) \quad \text{.} \quad (V.3) \]

Hence, the random variables \( \psi_t(x, \sigma) \), \( \sigma = \pm 1 \), satisfy an evolution equation of a form
required in the previous section, in (IV.3),

\[ i\sigma \partial_t \psi_t(x, \sigma) = \sum_{y \in \mathbb{Z}^d} \alpha(x - y) \psi_t(y, \sigma) + \lambda R_t(x) \psi_t(x, \sigma) + \lambda \psi_t(x, -1) \psi_t(x, \sigma) \psi_t(x, 1) \quad \text{,} \quad (V.4) \]

where we have defined \( R_t(x) := 2\mathbb{E}[|\psi_t(x)|^2] \geq 0 \), which is also equal to \( 2\kappa(\psi_t(x, -1), \psi_t(x, 1)) \).

A. Translation invariant initial measures

The evolution problem simplifies significantly, if we assume that the initial data is not
only gauge, but also translation invariant. Since also spatial translations commute with
the time evolution, we can use the earlier results for the cumulants of Fourier transforms of the
random field. In particular, then for any \( t \geq 0 \) and \( x, y \in \mathbb{Z}^d \),

\[ \mathbb{E}[\psi_t(x)^* \psi_t(y)] = \mathbb{E}[\psi_t(0)^* \psi_t(y - x)] \quad (V.5) \]

This implies that \( R_t(x) = R_t(0) =: R_t \) for all \( t \), and therefore the evolution equation (V.4)
for translation and gauge invariant initial data can be written as

\[ i\sigma \partial_t \psi_t(x, \sigma) = \sum_{y \in \mathbb{Z}^d} \alpha^\lambda_t(x - y) \psi_t(y, \sigma) + \lambda \psi_t(x, -1) \psi_t(x, \sigma) \psi_t(x, 1) \quad \text{,} \quad (V.6) \]

where \( \alpha^\lambda_t(x) := \alpha(x) + \lambda \mathbb{1}(x = 0) R_t \).

Using multilinearity of the Wick polynomials, we thus find the following evolution equation
for the Fourier transformed fields \( \hat{\psi}_t(k, \sigma) := \sum_x e^{-ik \cdot x} \psi_t(x, \sigma) \),

\[ \partial_t \hat{\psi}_t(k, \sigma) = -i\sigma \omega^\lambda_t(k) \hat{\psi}_t(k, \sigma) : \]

\[ -i\lambda \int_{(\mathbb{T}^d)^3} dk_1 dk_2 dk_3 \delta(k - k_1 - k_2 - k_3) : \hat{\psi}_t(k_1, -1) \hat{\psi}_t(k_2, 1) \hat{\psi}_t(k_3, 1) \quad \text{,} \quad (V.7) \]

where

\[ \omega^\lambda_t(k) := \hat{\alpha}^\lambda_t(k) = \hat{\alpha}(k) + \lambda R_t \quad \text{.} \quad (V.8) \]

For later use, let us point out that the definitions of the random fields imply the following
rule for complex conjugation of the Fourier transformed fields: \( \hat{\psi}_t(k, \sigma)^* = \hat{\psi}_t(-k, -\sigma) \). In
addition, the assumed symmetry of \( \alpha \) implies the symmetry \( \omega^\lambda_t(-k) = \omega^\lambda_t(k) \).

We recall that in the present translation invariant setting, the \( n \)-th cumulants satisfy for
\( k \in (\mathbb{T}^d)^n \), \( \sigma \in \{-1, 1\}^n \)

\[ \kappa(\hat{\psi}_t(k_1, \sigma_1), \ldots, \hat{\psi}_t(k_n, \sigma_n)) = \delta \left( \sum_{\ell=1}^n k_\ell \right) \hat{F}_n(k, \sigma; t) \quad \text{,} \quad (V.9) \]
where \( F_{n}(x, \sigma; t) := \mathbb{I}(x_1 = 0) \kappa(\psi_t(x_1, \sigma_1), \ldots, \psi_t(x_n, \sigma_n)) \) is identically zero unless \( \sum_{\ell} \sigma_{\ell} = 0 \). We are now mainly interested in the evolution of the lowest nonzero cumulants, i.e., of \( F_2(x, (-1, 1)) \). We denote its Fourier transform by \( W \); more precisely, we set

\[
W_\lambda^\Delta(k) := \sum_{\mathbf{x} \in \mathbb{Z}^d} e^{-i2\pi k \cdot \mathbf{x}} \kappa(\psi_t(0)^*, \psi_t(\mathbf{x})) = \sum_{\mathbf{x} \in \mathbb{Z}^d} e^{-i2\pi k \cdot \mathbf{x}} \mathbb{E}[\psi_t(0)^* \psi_t(\mathbf{x})].
\]

(V.10)

It follows that \( \hat{F}_2((k_1, k_2), (-1, 1); t) = W_\lambda^\Delta(k_2) \) and \( \hat{F}_2((k_1, k_2), (1, -1); t) = W_\Delta(-k_2) \). Therefore, we have the following general rule for second order cumulants:

\[
\kappa(\hat{\psi}_t(k_1, \sigma_1), \hat{\psi}_t(k_2, \sigma_2)) = \delta(k_1 + k_2) \mathbb{I}(\sigma_1 + \sigma_2 = 0)W_\Delta(\sigma_2 k_2).
\]

(V.11)

Therefore, to study the evolution of all second moments in this systems, it suffices to consider the formula

\[
W_{\ell}(k) := \mathbb{I}(\mathbf{x} = 0) \kappa(\psi_t(\mathbf{x}_1, \sigma_1), \ldots, \psi_t(\mathbf{x}_n, \sigma_n)) \quad \text{for} \quad \ell = 0, 1, 2, \ldots
\]

(V.12)

with the collision operator

\[
\mathcal{C}(W(\cdot))(k) = 4\pi \int_{(T^\rho)^3} dk_1 dk_2 dk_3 \delta(k + k_2 - k_3)\delta(\omega + \omega_1 - \omega_2 - \omega_3) \times \left[ W(k_1)W(k_2)W(k_3) + W(k)W(k_2)W(k_3) - W(k)W(k_1)W(k_3) - W(k)W(k_1)W(k_2) \right]
\]

(V.13)

where \( \omega := \tilde{\omega}(k) \) and \( \omega := \tilde{\omega}(k) \).

In fact, a lucky accident hides the fact that our present random fields are actually ill suited for taking of the scaling limit: it is clear from the linear part in (V.7) that they are highly oscillatory, and only observables where these oscillations cancel out, can be hoped to have a (nonzero) limiting value in the kinetic scaling limit. Fortunately, there is a simple "renormalization" which cancels these fast oscillations. If we define a new random field by the formula

\[
a_t(k, \sigma) = \tilde{\psi}_t(k, \sigma) \exp \left( i\sigma \int_0^t ds \omega_s^\lambda(k) \right),
\]

(V.15)

then it clearly satisfies an equation without a linear term. Explicitly, then

\[
\partial_t a_t(k_1, \sigma) = -i\sigma \lambda \int_{(T^\rho)^3} dk_2 dk_3 dk_4 \delta(k_1 - k_2 - k_3 - k_4) \times e^{i[(\sigma_1 + \omega_2 - \sigma_3 - \omega_4):a_t(k_2, -1)a_t(k_3, \sigma)a_t(k_4, 1):}. \]

(V.16)
Note that due to the alternating signs, the time dependent terms cancel each other out in the oscillatory phase term inside the integral. In fact, the same happens in the second order cumulants, as can be checked by using (V.11) and the symmetry of $\omega_i^\lambda$: we then find that

$$\kappa(a_i(k_1, \sigma_1), a_i(k_2, \sigma_2)) = \delta(k_1 + k_2)\mathbb{I}(\sigma_1 + \sigma_2 = 0)W_i^\lambda(\sigma_2 k_2).$$

It is clear that multiplication with a nonrandom term as in (V.15) does not spoil the gauge invariance of the field so we can rely on it also when working with the cumulants of the $a$-fields.

We can now study the evolution of $W_i^\lambda(k)$ by employing the expansion given in (IV.12) to the cumulant $\kappa(a_i(k', \sigma'), a_i(k, \sigma))$. Note that the combinatorial properties of the Wick polynomials, mainly described by Proposition III.8, will allow us to easily identify the relevant contributions to the Boltzmann-Peierls evolution. We then find using any $\sigma' = -\sigma$ that

$$\delta(k' + k)(W_i^\lambda(\sigma k) - W_i^\lambda(\sigma k)) = -i\lambda\sigma \int_{\mathbb{T}^d}^t ds \int_{\mathbb{T}^d}^s d\delta(k - k_1 - k_2 - k_3)e^{i\omega(k_1 - k_2 - k_3)k_{\delta}} \kappa([a_0]_T)$$

$$- i\lambda\sigma' \int_{\mathbb{T}^d}^t ds \int_{\mathbb{T}^d}^s d\delta(k' - k_1 - k_2 - k_3)e^{i\omega(k_1 - k_2 - k_3)k_{\delta}} \kappa([a_0]_T)$$

$$- \lambda^2\sigma \sum_{\ell \in I} \int_{\mathbb{T}^d}^t ds \int_{\mathbb{T}^d}^s d\delta(k_1 + k_2 + k_3) e^{i\omega(k_1 + k_2 + k_3)k_{\delta}} \times \int_{\mathbb{T}^d}^s ds' \int_{\mathbb{T}^d}^s d\delta(k_4 + k_5 + k_6) e^{i\omega(k_4 + k_5 + k_6)k_{\delta}} \mathbb{E}[\hat{a}_i^J; \hat{a}_i^{J'}; \hat{a}_i^{J'}]$$

$$- \lambda^2\sigma' \sum_{\ell \in I'} \int_{\mathbb{T}^d}^t ds \int_{\mathbb{T}^d}^s d\delta(k_1 + k_2 + k_3) e^{i\omega(k_1 + k_2 + k_3)k_{\delta}} \times \int_{\mathbb{T}^d}^s ds' \int_{\mathbb{T}^d}^s d\delta(k_4 + k_5 + k_6) e^{i\omega(k_4 + k_5 + k_6)k_{\delta}} \mathbb{E}[\hat{a}_i^J; \hat{a}_i^{J'}; \hat{a}_i^{J'}],$$

(V.18)

where

$$I = \{(k_1, -1), (k_2, \sigma), (k_3, 1), (k', \sigma')\},$$

(V.19)

$$I' = \{(k_1, -1), (k_2, \sigma'), (k_3, 1), (k, \sigma)\},$$

(V.20)

$$J_\ell = \{(k_1, -1), (k_3, \sigma), (k_5, 1)\}.$$  

(V.21)

Following the standard perturbation recipe, we next apply the cumulant hierarchy to the terms depending on $a_{\nu}\sigma'$ in (V.18). This results in a sum of two terms: one, in which every $a_{\nu}$ has been replaced by $a_0$, plus a “correction” which we denote by $\delta(k' + k)/\mathcal{R}_3(\sigma k, t)$. Further iterations of the perturbation expansion and a careful study of the oscillations of the term by term expansion as in Ref. 9 leads us to the conjecture that $\mathcal{R}_3(\mathbf{k}, \tau \lambda^{-2})$ should converge in the kinetic scaling limit, as $\lambda \to 0$, at least for sufficiently nondegenerate dispersion relations and for large enough dimension $d$. In addition, the analysis indicates that the limit value is $O(\tau^2)$, which is negligible compared to the contribution from the other terms following from (V.18). However, the term by term analysis does not suffice to actually prove the claim since the method which was used to rigorously control the convergence of the perturbation expansion in Ref. 9 was based on time stationarity of the initial state. This assumption cannot be made here since we are interested in nontrivial time evolution effects. Instead of going into the details of the above argument, we discuss a less technically involved motivation for the claim in Section VI.B.

Next, we need to evaluate expectations of the form $\mathbb{E}[a_1 a_2 a_3; a_4 a_5 a_6]$ where each $a_i$ stands for one of the field variables. The cumulant expansion in Theorem III.8 and the
vanishing of the third order cumulants imply

\[
E[a_1 a_2 a_3; a_4 a_5 a_6] = \kappa(a_1, a_4)\kappa(a_2, a_5)\kappa(a_3, a_6) + \ldots W(-k_1)W(-k_2)W(k_3) + \text{NPC}. \quad (V.24)
\]

By changing integration variables so that \(k_1 \rightarrow -k_1\) in (V.23) and \(k_1 \rightarrow -k_3, k_2 \rightarrow -k_2\),

where the last contribution denotes a sum of the nine terms consisting of a product of a second order cumulant and a fourth order cumulant. Naturally, also some of the above terms can be zero because of the gauge invariance constraints.

To better work with the expressions arising from (V.18), let us next introduce a few shorthand notations. We denote

\[
W(k) := W_0^\delta(k),
\]

\[
\int dk_{12 \ldots n} := \int (\mathbb{T}^\omega) dk_1 dk_2 \ldots dk_n,
\]

\[
\delta(k - k_{ij\ell}) := \delta(k - k_i - k_j - k_\ell),
\]

\[
\Omega_{\pm +} := \omega(k_1) + \omega(k_2) - \omega(k_3) - \omega(k),
\]

\[
\Omega_{+ +} := \omega(k_1) - \omega(k_2) - \omega(k_3) + \omega(k). \quad (V.22)
\]

We also choose \(\sigma = 1, \sigma' = -1\) and we will only consider the pairing contractions (i.e., the Gaussian contractions) in the expansion (V.22). In fact, all terms arising from the non-pairing contractions are typically negligible in the kinetic scaling limit of the present type. As an example, in Appendix C we show how the first order terms in (V.18) vanish in the kinetic limit by assuming sufficient regularity of the dispersion relation \(\omega\) and the \(\ell_\perp\)-clustering property of the fourth order cumulants. As explained in Ref. 9, the contribution from the non-pairing terms in (V.22) can be controlled by similar techniques but we will skip this more involved analysis here.

Hence, after integrating out the variables \(k_i, i = 4, 5, 6\), the fourth term in (V.18) gives

\[
2\lambda^2 \delta(k + k') \int dk_{123} \delta(k - k_{123}) \int_0^t ds \int_0^s ds' e^{i(s-s')\Omega_{\pm +}}
\]

\[
\times [W(k)W(k_2)W(k_3) - W(k)W(-k_1)W(k_3) - W(k)W(-k_1)W(k_2)W(k_3)] + \text{NPC} \quad (V.23)
\]

where NPC stands for "non-pairing contraction terms". We proceed in the same way for the fifth term in (V.18) yielding

\[
2\lambda^2 \delta(k + k') \int dk_{123} \delta(k + k_{123}) \int_0^t ds \int_0^s ds' e^{i(s-s')\Omega_{+ +}}
\]

\[
\times [W(k)W(-k_1)W(-k_2) - W(k)W(-k_1)W(k_3) - W(k)W(-k_2)W(k_3)] + \text{NPC}. \quad (V.24)
\]

By changing integration variables so that \(k_1 \rightarrow -k_1\) in (V.23) and \(k_1 \rightarrow -k_3, k_2 \rightarrow -k_2\),
in (V.24), we obtain
\[
W_{\lambda}^1(k) - W(k) - (R_3(k,t) + NPC)
= 2\lambda^2 \int_{\mathbb{T}^3} d\mathbf{k}_{123} \delta(k + k_1 - k_2 - k_3) \int_0^t ds \int_0^s ds' e^{i(s-s')(\omega_1 - \omega_2 - \omega_3 + \omega)}
\times [W(k)W(k_2)W(k_3) - W(k)W(k_2)W(k_3)]
- 2\lambda^2 \int_{\mathbb{T}^3} d\mathbf{k}_{123} \delta(k + k_1 - k_2 - k_3) \int_0^t ds \int_0^s ds' e^{-i(s-s')(\omega_1 - \omega_2 - \omega_3 + \omega)}
\times [W(k)W(k_2)W(k_3) - W(k)W(k_2)W(k_3)]
- 2\lambda^2 \int_{\mathbb{T}^3} d\mathbf{k}_{123} \delta(k + k_1 - k_2 - k_3) \int_0^t ds \int_0^s ds' e^{i(s-s')(\omega_1 - \omega_2 - \omega_3 + \omega)}
\times [W(k_1)W(k_2)W(k_3) + W(k)W(k_2)W(k_3) - W(k)W(k_1)W(k_3) - W(k)W(k_2)W(k_3)].
\] (V.25)

Note that for any \( \Omega \in \mathbb{R} \) we have \( \lambda^2 \int_0^t ds \int_{|s| \leq s} dr e^{i\omega \Omega} = \int_{|s| \leq s} dr e^{i\omega \Omega}(\lambda^2 t - \lambda^2 |r|) \). By setting \( t = \tau \lambda^{-2} \) and taking \( \lambda \to 0 \), this expression formally converges to \( \tau \int_{-\infty}^{\infty} dr e^{i\omega \tau} = \tau 2\pi \delta(\Omega) \). Therefore, doing this in (V.25) yields the conjecture that
\[
W_\tau(k) - W(k) - O(\tau^2) = \tau 4\pi \int_{\mathbb{T}^3} d\mathbf{k}_{123} \delta(k + k_1 - k_2 - k_3) \delta(\omega + \omega_1 - \omega_2 - \omega_3)
\times [W(k_1)W(k_2)W(k_3) + W(k)W(k_2)W(k_3) - W(k)W(k_1)W(k_3) - W(k)W(k_2)W(k_3)].
\] (V.26)

Since here \( W(k) = W_0(k) \), if we divide the left hand side by \( \tau \) and then take \( \tau \to 0 \), it converges to \( \partial_\tau W_\tau(k) \) at \( \tau = 0 \). Dividing the right hand side of (V.26) by \( \tau \) yields \( C(W_0(\cdot)))(k) \), as defined in (V.14). Therefore, the Boltzmann-Peierls equation should hold at \( \tau = 0 \). Assuming that the state of the original system remains so regular that the estimates leading to the conjecture continue to hold, we thus find that the Boltzmann-Peierls equation should be valid for the limit of \( W_{\lambda \lambda^{-2}} \) also at later times \( \tau \), as was claimed in the beginning of the subsection.

2. Decay of field time-correlations

As a second example of how the standard perturbation expansion works for the cumulants, we consider a kinetic scaling limit of time-correlations. In particular, our goal is to show how the main results proven in Ref. 9 relate to the present cumulant hierarchy expansions.

The notation \( \tilde{a}_t \) was used in Ref. 9 to define the finite periodic lattice analogue of the present \( \tilde{a}_t \)-field. (One can compare the definition of \( \tilde{a}_t \) in Ref. 9, Eq. (3.9), and its evolution equation in Ref. 9, Eq. (3.10), to those given for \( a_t \) here.) Translated to the present infinite lattice setup, the main result of Ref. 9, Theorem 2.4, leads to the following conjecture about the decay of time correlations of \( a_t \): start the system from an \( \epsilon \)-clustering equilibrium Gibbs state. Then there is a continuous function \( A^\lambda_t(k) \) such that \( \mathbb{E}[a_0(k') - 1]a_t(k,1)] = \delta(k' + k) A^\lambda_t(k) \). The conjecture is that the kinetic scaling limit of \( A^\lambda_t \) exists and its decay is governed by the “loss term” of the Boltzmann-Peierls equation (V.13) evaluated at the corresponding limit equilibrium covariance function \( W^{\text{eq}}(k) = \beta^{-1}/(\omega(k) - \mu) \) where \( \beta > 0 \) and \( \mu \in \mathbb{R} \) are parameters determined by the equilibrium state. (Such functions \( W^{\text{eq}} \) are indeed stationary solutions of (V.13).) More precisely, Theorem 2.4 in Ref. 9 is consistent with the conjecture that
\[
\lim_{\lambda \to 0} A^\lambda_{t=\lambda^{-2}}(k) = W^{\text{eq}}(k)e^{-\tau(\text{eq})(\cdot)(k)}.
\] (V.27)
where
\[
\Gamma(W(\cdot))(k) = -2 \int_0^\infty \int_{(\mathbb{T}^3)^3} dk_1 dk_2 dk_3 \delta(k + k_2 - k_3) e^{i r (\omega_1 - \omega_2 - \omega_3 + \omega)} \times [W(k_2)W(k_3) - W(k_1)W(k_3) - W(k_1)W(k_2)].
\] (V.28)

Instead of assuming that the system starts from an equilibrium state, let us consider more general states which we assume to be gauge and translation invariant and \(\ell_1\)-clustering. We can immediately use the results derived in Section IV if we consider the “\(a_0\)” term to be a new field which has trivial time evolution with zero amplitudes, i.e., the corresponding \(\mathcal{Z}_f\)-set is empty. The net effect of this change is that more than half of the terms analyzed in the previous section will be absent. For instance, the expansion of the time correlation (IV.13) reads
\[
\kappa[a_0(k',-1), a_\ell(k,1)] = \kappa[a_0(k',-1), a_0(k,1)]
- i \lambda \int_0^\infty ds \int dk_1 dk_2 dk_3 \delta(k - k_2 - k_3) e^{i s (\omega_1 + \omega_2 + \omega_3)} \kappa[(a_0)_I(k',-1)]
- \lambda^2 \sum_{\ell \in I'} \int_0^\ell ds \int dk_1 dk_2 dk_3 \delta(k - k_1 - k_2 - k_3) e^{i s (\omega_1 + \omega_2 + \omega_3)} \int_0^s ds' \times \int dk_4 dk_5 dk_0 \delta(k_4 - k_1 - k_5 - k_0) e^{i s' (\sigma_1 + \omega_4 - \sigma_5 - \omega_6)} \mathbb{E} \left[ \langle a_\ell^{I'} : a_0(k',-1) a_\ell^{I'(I')} \rangle \right].
\] (V.29)

where \(I' = ((k_1, -1), (k_2, 1), (k_3, 1))\) and \(J_s = ((k_4, -1), (k_5, \sigma_1), (k_6, 1))\).

As in Section V A 1, we now assume that only pairings contribute in the kinetic scaling limit. When applying (V.22) to expand \(\mathbb{E} \left[ \langle a_\ell^{I'} : a_0(k',-1) a_\ell^{I'(I')} \rangle \right]\), we note that every pairing term results in a product containing a factor \(A_\lambda^\lambda(-k')\) and a product of two \(W_0^\lambda\)-terms. The rest of the structure is identical to the one considered earlier, some of the terms are merely missing now. We then use the perturbation expansion to the product once more. This produces a term where \(s'\) is set to 0, and a remainder which we assume to be negligible as before. The rest of the computation is essentially the same as in Section V A 1, yielding
\[
A_\tau^{\lambda - 2}(k) = A_0^\lambda(k) - \text{(terms higher order in \(\lambda\) or \(\tau\))}
= 2 \lambda^2 \int_{(\mathbb{T}^3)^3} dk_1 dk_2 dk_3 \delta(k + k_1 - k_2 - k_3) \int_0^\tau ds \int_0^s ds' e^{i r (\omega_1 - \omega_2 + \omega)} \times [A_0^\lambda(k)W_0^\lambda(k_2)W_0^\lambda(k_3) - A_0^\lambda(k)W_0^\lambda(k_1)W_0^\lambda(k_3) - A_0^\lambda(k)W_0^\lambda(k_1)W_0^\lambda(k_2)].
\] (V.30)

Hence, if we divide the equation by \(\tau\) and then take \(\lambda \to 0\), followed by \(\tau \to 0\), we find that \(A_\tau(k) := \lim_{\lambda \to 0} A_\lambda^{\lambda - 2}(k)\) should satisfy at \(\tau = 0\)
\[
\partial_\tau A_\tau(k) = - A_\tau(k) \Gamma(W_\tau(\cdot))(k),
\] (V.31)
where \(\Gamma\) has been defined in (V.28). As before, the conjecture is that this equation continues to hold for other values \(\tau > 0\), as well.

Once \(W_\tau\) is given, equation (V.31) is straightforward to solve. Since \(A_0^\lambda(k) = W_0(k)\), the solution reads
\[
A_\tau(k) = W_0(k) e^{- \int_0^\tau ds \Gamma(W_\tau(\cdot))(k)}.
\] (V.32)

If the system is started in an equilibrium state with \(W_0 = W^{\text{eq}}\), we have \(W_s = W^{\text{eq}}\) for all \(s\). Thus (V.32) implies (V.27) in this special case.
VI. DISCUSSION ABOUT FURTHER APPLICATIONS

A. Limitations of the direct renormalization procedure: inhomogeneous DNLS

The field renormalization used with the translational invariant data greatly simplified the evolution equation by removing the linear term. The renormalization procedure, given in (V.15), was a simple multiplication by a time and \( k \)-dependent function and the time-dependent first order terms had no effect in the interaction term. Unfortunately, this case is atypical: most commonly, the necessary renormalization is not a multiplication operator and the first order terms will also affect the oscillatory phase terms arising from the harmonic evolution. In fact, this happens also for the DNLS model as soon as we drop the requirement that the initial data is translation invariant. To explain the changes needed in the renormalization procedure, we discuss in this subsection the DNLS model with inhomogeneous initial data in some more detail.

Before considering the inhomogeneous case, let us begin with an example which emphasizes the importance of the field renormalization even for translation invariant initial data if one considers taking kinetic scaling limits of all field observables. We inspect the time correlation of the “bare” \( \hat{\psi} \)-fields, i.e., \( \mathbb{E}[\hat{\psi}_0(k',-1)\hat{\psi}_1(k,1)] \) assuming spatially homogeneous, gauge invariant initial data. Then there exists a function \( \Psi_t(k) \) such that \( \mathbb{E}[\hat{\psi}_0(k',-1)\hat{\psi}_1(k,1)] = \delta(k' + k)\Psi_t(k) \). By using (V.7) in (IV.6), we find the following evolution equation for \( \Psi_t \):

\[
\delta(k' + k)\Psi_t(k) = \kappa(\hat{\psi}_0(k',-1),\hat{\psi}_1(k,1)) = \delta(k' + k)\Psi_0(k) - i\delta(k' + k)\int_0^t ds \omega^\lambda_s(k)\Psi_s(k) + i\lambda\int_0^t ds \int_{\mathbb{T}^3} dk_1 dk_2 dk_3 \delta(k - k_1 - k_2 - k_3) \times \mathbb{E}[\hat{\psi}_0(k',-1)\hat{\psi}_1(k_1,1)\hat{\psi}_1(k_2,1)\hat{\psi}_1(k_3,1)].
\]

(VI.1)

The linear equation associated to (VI.1) thus has the form

\[
f_t(k) = f_0(k) - i \int_0^t ds \omega^\lambda_s(k)f_s(k),
\]

(VI.2)

which is solved by \( f_t(k) = U_t(k)f_0(k) \) where \( U_t(k) = \exp(-i\int_0^t ds \omega^\lambda_s(k)) \). We recall that \( \omega^\lambda_s(k) = \omega(k) + \lambda R_s \), and thus at a kinetic time scale, with \( t = \tau \lambda^{-2} \), we have \( \int_0^t ds \omega^\lambda_s(k) = \tau \lambda^{-2}\omega(k) + O(\lambda^{-1}) \). Therefore, \( U_t \) has unbounded oscillations in the kinetic scaling limit. Also, we find that even though the effect of the first order term proportional to \( R_s \) is subdominant, it is still rapidly oscillating on the kinetic time-scale and should not be “expanded” in any perturbative treatment of the problem.

We could solve the problem with these unbounded oscillations by considering instead of \( \hat{\psi}_1 \) the renormalized field \( a_t = U_t^{-1}\hat{\psi}_1 \). In fact, by the results of Section V A 2, and using \( a_0 = \hat{\psi}_0 \), we find that

\[
\delta(k + k')A^\lambda_t(k) = \mathbb{E}[a_0(k',-1)a_t(k,1)] = U_t(k)\mathbb{E}[\hat{\psi}_0(k',-1)\hat{\psi}_1(k,1)] = \delta(k + k')U_t(k)\Psi_t(k).
\]

(VI.3)

We have argued in Section V A 2 that the kinetic scaling limit of \( A^\lambda_t \) exists. Then \( \Psi_{\tau \lambda^{-2}}(k) \) cannot have a convergent limit as \( \lambda \rightarrow 0 \); instead, it has fast oscillations proportional to \( U_{\tau \lambda^{-2}}(k) \). Let us also once more stress that the “zeroth order renormalization”, i.e., counteracting the free evolution, does not remove all of the unbounded oscillations but still leaves those resulting from the \( R_t \) term.

However, the above renormalization procedure cannot be straightforwardly extended to more complicated cases. Consider next the DNLS model with inhomogeneous initial data.
As in section V A 1, our goal is to find the right observable which satisfies the Boltzmann equation in the kinetic scaling limit. The evolution equation for the bare field $\psi$ reads
\[
\partial_t \hat{\psi}(k, \sigma) = -i\sigma \omega(k) \hat{\psi}(k, \sigma) - 2i\sigma \lambda \int_{(\mathbb{T}^3)} dk_1 dk_2 dk_3 \delta(k - k_1 - k_2 - k_3) \kappa(\hat{\psi}_t(k_1, -1), \hat{\psi}_t(k_3, 1)) \hat{\psi}_t(k_2, \sigma)
- i\sigma \lambda \int_{(\mathbb{T}^3)} dk_1 dk_2 dk_3 \delta(k - k_1 - k_2 - k_3) : \hat{\psi}_t(k_1, -1) \hat{\psi}_t(k_2, \sigma) \hat{\psi}_t(k_3, 1) :. \tag{VI.4}
\]
By following the same strategy as for the homogeneous case, let $U_t$, $t \geq 0$, denote the family of linear operators which solves the linear part of (VI.4): we suppose that it solves the operator equation $\partial_t U_t = -iH_t U_t$ where $(H_t f)(k, \sigma) = \sigma \omega(k) f(k, \sigma) + 2\sigma \lambda \int_{\mathbb{T}^d} dk' K_t(k - k') f(k', \sigma)$. The time-dependent convolution kernel $K_t(k)$ should be equal to $\int_{\mathbb{T}^d} dk_k \kappa(\hat{\psi}_t(k_1, -1), \hat{\psi}_t(k - k_1, 1))$. This can be done either by first solving the implicit equation for the above integral over the cumulant, or by leaving $K_t$ arbitrary and fixing it by some minimization procedure at the end. If such a family $U_t$ can be found, we may define as before $\tilde{\alpha}_t(k, \sigma) = (U_t^{-1} \hat{\psi}_t)(k, \sigma)$ and, since $\partial_t U_t^{-1} = -U_t^{-1}(\partial_t U_t) U_t^{-1} = iU_t^{-1} H_t$, it then satisfies an evolution equation
\[
\partial_t \tilde{\alpha}_t(k, \sigma) = -i\sigma \lambda \int_{(\mathbb{T}^3)} d^3k \int_{(\mathbb{T}^3)} d^3k'' \int_{(\mathbb{T}^3)} d^3k' \delta(k'' - k_1' - k_2' - k_3') \times \tilde{u}_t(k, k'', \sigma) u_t(k_1', k_1, -1) u_t(k_2, k_2, \sigma) u_t(k_3', k_3, 1) : \tilde{\alpha}_t(k_1, -1) \tilde{\alpha}_t(k_2, \sigma) \tilde{\alpha}_t(k_3, 1) :. \tag{VI.5}
\]
where $u_t(k, k', \sigma)$ and $\tilde{u}_t(k, k', \sigma)$ denote the formal integral kernels of the operators $U_t$ and $U_t^{-1}$, respectively. (Note that the operators $U_t$ are diagonal in $\sigma$ but not any more in $k$.)

Apart from some special cases it seems difficult to gain sufficient control over the operators $U_t$ to consider taking a kinetic limit using the observables $\tilde{\alpha}_t$, unlike with the explicit phases factors which appeared in the spatially homogeneous case. Even though $U_t$ approach the same multiplication operator as before when $\lambda \to 0$ for a fixed $t$, it is not clear that the corrections do not contribute in the limit, since we need to consider $t = O(\lambda^{-2})$. Thus, although the cumulant expansion of $\tilde{\alpha}_t$-fields is simpler than that of $\psi_t$-fields, to control the kinetic scaling limit looks intractable. Hence, new approaches for the study of the kinetic time scales are called for.

### B. Kinetic theory beyond kinetic time-scales?

In this section we propose a new approach to the problem when a renormalization scheme with a convergent kinetic scaling limit cannot be found or controlled. The approach does not require taking $\lambda \to 0$, and, if successful, it may also yield estimates which are valid beyond the standard kinetic time scales which was $O(\lambda^{-2})$ in the above DNLS case.

The main idea can be summarized in the following simple observation. Suppose $f_t$ is a solution to the equation
\[
f_t = f_0 + R_t + \int_0^t ds F_{s,t}[f_s] \tag{VI.6}
\]
where $R_t = O(\varepsilon)$ uniformly in $t$ and $F_{s,t}$, $0 \leq s \leq t$, is an explicit, but possibly nonlinear functional of $f_s$. Suppose furthermore that there is another, “simpler”, functional $\Phi_{s,t}$ such that $F_{s,t} = \Phi_{s,t} + O(\varepsilon(1 + |t - s|)^{-p})$ with $p > 1$; by simpler we mean that the evolution problem
\[
\varphi_t = S_t + \int_0^t ds \Phi_{s,t}[\varphi_s], \tag{VI.7}
\]
for any bounded “source term” $S$, is easier to study than (VI.6). Under these assumptions, any solution to (VI.6) satisfies

$$f_t = f_0 + \rho_t + \int_0^t ds \Phi_{s,t}[f_s], \quad (VI.8)$$

where $\rho_t := R_t + \int_0^t ds (F_{s,t}[f_s] - \Phi_{s,t}[f_s])$ is $O(\varepsilon)$ uniformly in $t$. Therefore, if we could prove that (VI.7) is stable under perturbations of the source term $S$, we may conclude that the solution $\varphi_t$ to

$$\varphi_t = f_0 + \int_0^t ds \Phi_{s,t}[\varphi_s] \quad (VI.9)$$

then approximates the “true” solution $f_t$ with an error which is $O(\varepsilon)$ uniformly in time.

To have a concrete example, consider again the DNLS with gauge invariant initial data. We sketch below two conjectures about this system. The details of the conjectures should not be taken too seriously: they should be considered more as examples of what could happen in general rather than as specific conjectures about the behavior of the DNLS system. For this reason, the discussion will be kept on a very loose level; in particular, we do not wish to make any specific claims about what kind of metrics should be used for studying the uniform boundedness in time.

Let us begin with the case of inhomogeneous, gauge invariant initial data. We move to slowly varying fields by cancelling the free evolution term. This renormalization leads to equations which are almost identical to those in the homogeneous case. Namely, for the field $b_t(k, \sigma) := e^{i\sigma \omega(k) \hat{\psi}_t(k, \sigma)}$ we have

$$\frac{\partial}{\partial t} b_t(k, \sigma) = -2i\sigma \lambda \int_{(T^d)^3} dk_1 dk_2 dk_3 \delta(k - k_1 - k_2 - k_3) e^{it(\sigma \omega + \omega_1 - \sigma \omega_2 - \omega_3)} B_t(k_1, k_3) \langle b_t(k_2, \sigma), b_t(k_3, 1) \rangle \quad (VI.10)$$

where $B_t(k', k) = \kappa(b_t(k', -1), b_t(k, 1))$. By the gauge invariance, $\kappa(b_t(k', \sigma'), b_t(k, \sigma)) = 0$, unless $\sigma' + \sigma = 0$. Hence, there is only one other nonzero second order cumulant, $\kappa(b_t(k', 1), b_t(k, -1)) = B_t(k, k') = B_t(-k', -k)^*$. To study the fourth order cumulants, it suffices to concentrate on the function

$$D_t(k) := \kappa(b_t(k_1, -1), b_t(k_2, -1), b_t(k_3, 1), b_t(k_4, 1)), \quad k \in (T^d)^4.$$
Then the first two equations in the cumulant hierarchy are equivalent to
\[
\partial_t B_t(k', k) = -2i\lambda \int_{(T^d)^3} dk_1 dk_2 dk_3 \delta(k - k_1 - k_2 - k_3)e^{it(\omega_1 + \omega_2 - \omega_3)} B_t(k_1, k_3) B_t(k_2, k')
\]
\[+ 2i\lambda \int_{(T^d)^3} dk_1 dk_2 dk_3 \delta(k' - k_1 - k_2 - k_3)e^{it(-\omega_1 + \omega_2 + \omega_3)} B_t(k_1, k_3) B_t(k, k_2)
\]
\[+ \lambda \int_{(T^d)^3} dk_1 dk_2 dk_3 \delta(k' - k_1 - k_2 - k_3)e^{it(\omega_1 + \omega_2 - \omega_3)} D_t(k_1, k_2, k_3)
\]
\[+ \lambda \int_{(T^d)^3} dk_1 dk_2 dk_3 \delta(k' - k_1 - k_2 - k_3)e^{it(-\omega_1 + \omega_2 + \omega_3)} D_t(k_1, k_2, k_3, k),
\]
(VI.11)

\[
\partial_t D_t(k) = -2i\lambda \sum_{\ell=1}^4 \sigma_\ell \int_{(T^d)^3} dk'_1 dk'_2 dk'_3 \delta(k_\ell' - k_1' - k_2' - k_3')e^{it(\sigma_\ell \omega_1 + \omega'_1 - \sigma_\ell \omega'_2 - \omega'_3)}
\]
\[\times B_t(k_1', k_3') D_t(\text{“replace } k_\ell \text{ by } k_\ell'\text{”})
\]
\[+ \lambda \sum_{\ell=1}^4 \sigma_\ell \int_{(T^d)^3} dk'_1 dk'_2 dk'_3 \delta(k_\ell' - k_1' - k_2' - k_3')e^{it(\sigma_\ell \omega_1 + \omega'_1 - \sigma_\ell \omega'_2 - \omega'_3)}
\]
\[\times (\text{“}(1 \times \kappa_0) + (5 \times B_t D_t) + (6 \times B_t B_t B_t)\text{”}),
\]
(VI.12)

where in the second formula, \(\sigma := (-1, -1, 1, 1)\) and on the last line we have applied (V.22) and merely denoted how many nonzero terms each type of partition can have. Let us point out that the first two terms in (VI.11) cancel each other out if the state is spatially homogeneous. Since then \(B_t(k', k) \propto \delta(k' + k)\). For an inhomogeneous state, however, the cancellation need not be exact, and since this term is then not \(O(\lambda^2)\), it will likely prevent taking of the kinetic scaling limit of \(B_t\) directly.

If we now integrate the \(B\) equation as in (IV.12), we find that

\[
B_t(k', k) = B_0(k', k)
\]

\[\quad - i\lambda \int_0^t ds \int_{(T^d)^3} dk_1 dk_2 dk_3 \delta(k - k_1 - k_2 - k_3)e^{is(\omega_1 + \omega_2 - \omega_3)}
\]
\[\quad \times (2B_t(k_1, k_3) B_t(k_2, k') + D_0(k', k_1, k_2, k_3))
\]
\[+ i\lambda \int_0^t ds \int_{(T^d)^3} dk_1 dk_2 dk_3 \delta(k' - k_1 - k_2 - k_3)e^{is(-\omega_1 + \omega_2 + \omega_3)}
\]
\[\times (2B_t(k_1, k_3) B_t(k, k_2) + D_0(k_1, k_2, k_3, k))
\]
\[+ \lambda^2 \int_0^t ds' \int_0^{s'} ds \sum_{\ell=1}^4 \sum_{s''} \Delta_{s''}(\text{VI.13})
\]

In the final sum, each term depends on \(s\) only via the oscillatory phases. These can be collected together and they have a structure

\[
e^{i\sigma_0(\omega_0 - \omega_2)} e^{i\sigma_\ell(\omega_0 - \omega_2 - \omega')} \int_0^t ds e^{i\sigma_0(\omega_0 - \omega_2 + \omega')} \int_0^{s'} ds' e^{i\sigma_\ell(\omega_0 - \omega_2 + \omega')}.
\]
(VI.14)

where \((\sigma_0, \kappa_0)\) is equal to \((1, k)\), if the term arises from the second last term in (VI.11) and it is equal \((-1, k')\) if it arises from the last term. The pair \((\sigma_\ell, \omega')\) comes from arguments of the corresponding “\(D_t\)-term” and thus depends also on this choice. Therefore, the \(s\)-integral over the oscillatory phase can be computed explicitly and to each of the terms in the sum it will produce a factor

\[
e^{i\sigma}(\omega_0 - \omega_2 + \omega') \int_0^t ds' e^{i\sigma_0(\omega_0 - \omega_2 + \omega')}.
\]
(VI.15)
It is difficult to go further in the analysis of the oscillatory phases without resorting to graph theory, and we will not pursue it here. However, already the simple example given in Appendix C shows that, if the state is $\ell_1$-clustering, the oscillations may result in time-integrals which are absolutely convergent over $[0,\infty)$. For instance, this explicit example implies that if the initial state is homogeneous and $\ell_1$-clustering, then the two terms depending on $D_0$ in (VI.13) are $O(\lambda)$ uniformly in time. If we assume that similar bounds are valid for inhomogeneous states and every term containing $\kappa_6$, the two-component field $f_t = (B_t, D_t)$ behaves as the model considered in the beginning of the section.

Therefore, assuming uniform boundedness of $\kappa_6$ allows using the principles described in the beginning of this section and results in a conjecture about the evolution of the cumulants. Unlike in the example in Appendix C, simple $L^\infty$-norms are not expected to yield such uniform behavior for $\kappa_6$. Instead, we suppose that there is some more complicated metric—for instance, weighted $L^2$-clustering norms—for the cumulants such that the following assumptions hold for sufficiently regular dispersion relations $\omega$ and initial data:

1. Suppose that the contribution to $D_t$ from $\kappa_6(t)$ is uniformly bounded in time, with a bound $O(\lambda^{1-q})$ where $0 \leq q \leq 1$.

2. Consider the evolution equation obtained for $(B_t, D_t)$ by the standard “closure relation”, i.e., by setting $\kappa_6 \to 0$ in (VI.12). Assume that this equation is stable under all uniformly bounded time-dependent perturbations of the source term.

If both of the above hold, then the solutions to the closure evolution equation for $f_t$ remain close to $(B_t, D_t)$ uniformly in $t$. In particular, the difference in the first (covariance) component is always $O(\lambda^{2-q})$. This would validate the closure equations as good approximations even to $t \to \infty$ asymptotic behavior of the covariance function for all sufficiently small $\lambda$.

Note that no scaling limit needs to be taken; in particular, it is not claimed that the kinetic scaling limit of $B_t$ or $D_t$ would exist. However, it should be stressed that any physical quantities of a limiting steady state, such as values of conductivities when the system is endowed with boundary reservoirs, would also carry similar errors and hence the closure equation values would only be expected to become exact in the limit $\lambda \to 0$. Concerning the validity of the first assumption, let us point out a recent result\(^{(1)}\) proving that similar time-correlations at $\ell_1$-clustering equilibrium states have $\ell_2$-clustering norms which remain uniformly bounded in time.

Another application can be obtained for a one-component case with $f_t = W^\lambda_t$ as follows: Consider the spatially homogeneous case and the exact evolution equation for $W^\lambda_t$ obtained from (V.18) by “dividing out” $\delta(k' + k)$ from both sides. Take $R_t$ to include all terms which contain either $\kappa_4$ or $\kappa_6$. Then the remaining pairing terms yield an explicit definition for $F_{s,t}$ such that (VI.6) holds. (In fact, then $F_{s,t}[W]$ is equal to the right hand side of (V.25).) Next choose “$\Phi_{s,t}$” equal to the Boltzmann collision operator $\lambda^2 C$ with $C$ defined in (V.14). If $W_t$ comes from an $\ell_1$-clustering state and the free evolution is sufficiently dispersive, then $F_{s,t}[W_s] = C[W_s] + O(\lambda^2(1 + |t-s|)^{-p})$, with $p > 1$. (For instance, the estimates given in Ref. 9, Proposition 7.4 and the Appendix, prove the bound with $p = 3d/7 - 1$ for a nearest neighbor dispersion relation—the computation is essentially the same as in Appendix C below. Using the notations defined in (C.3), the result also allows to quantify the dependence on the $\ell_1$-clustering assumption: the bound is proportional to $\|\kappa_2(s)\|_1^2 \lambda^2 (1 + |t-s|)^{-p}$. Hence, then $p > 1$ at least if $d \geq 5$. However, more careful estimates or the addition of next to nearest neighbor hopping could improve the bound.)

Whenever this is the case, we obtain a second conjecture about the homogeneous DNLS equation. Suppose that there is a metric for the cumulants such that the following results hold for sufficiently regular $\omega$ and initial data:

1. Suppose that the cumulants remain uniformly bounded in this metric, with an upper bound which implies that $\|\kappa_2\|_1 = O(1)$ and that all higher order cumulants have $\ell_1$-clustering norm which is $O(\lambda^{-q})$ for some $0 \leq q < 2$.

2. Assume that the corresponding Boltzmann-Peierls equation is stable under all uniformly bounded time-dependent perturbations of the source term.
If both of the above hold, then the solutions $W_τ$ to the Boltzmann-Peierls equation with initial data $W_λ^0$ are $O(λ^{2−q})$ close to $W_λ^{3−2}$ uniformly in $τ$. In particular, any stationary limit $\lim_{t→∞} W_τ$ can differ from the limit of the solution to the Boltzmann-Peierls equation only by $O(λ^{2−q})$.

The main benefit from using the Boltzmann-Peierls equation instead of the closure hierarchy concerns the second assumption: the homogeneous Boltzmann-Peierls equations enjoy many simplifying properties and a priori estimates. For example, they typically have an entropy functional and an associated “H-theorem” which allow to classify all stationary solutions to the equation. There are also many techniques developed to control the convergence towards the stationary solution.

As a final example, let us remark that even when total uniformity in time cannot be achieved, it might be possible to go beyond the kinetic timescales using the above methods. Consider $F_{s,t}(W_I) = C[W_I] + O(λ^2(1 + |t−s|)^p)$ for some $0 < p < 1$. Then the correction is not integrable and the perturbation $ρ_t$ to the source term is $O(λ^{2(t−p)})$. This remains $O(λ^c)$ for all $t = O(λ^{−(2−c)/(1−p)})$. Hence, for instance, the earlier nearest neighbor estimate with $d = 3$ would imply that the corrections to the Boltzmann-Peierls equation remain small for $t = O(λ^{−2−3/(5+ε)})$, that is, even for times much longer than the ones implied by the kinetic scaling limit.

ACKNOWLEDGMENTS

We thank Giovanni Gallavotti, Dario Gasbarra, Antti Kupiainen, Peng Mei, Alessia Nota, Sergio Simonella, and Herbert Spohn for useful discussions on the topic. The research of J. Lukkarinen and M. Marcozzi has been supported by the Academy of Finland via the Centre of Excellence in Analysis and Dynamics Research (project 271983) and from an Academy Project (project 258302).

Appendix A: Combinatorial definition of cumulants

The cumulants are connected to moments via a formula which is very similar to the definition we used here for the Wick polynomials: if $I ≠ ∅$, for any $x ∈ I$ we have

$$E[y^I] = \sum_{E:x∈E⊂I} E[y^I,E]κ[y_E]. \quad (A.1)$$

(The formula follows straightforwardly from the identity $\partial_x G_m = G_m ∂_x g_c$.) In fact, this formula allows a definition of cumulants which does not rely on differentiation or on the existence of exponential moments. Namely, if $I_0 ∈ I$ is such that $E[|y^I|] < ∞$ for all $I ⊆ I_0$, then to each $I ⊆ I_0$, $I ≠ ∅$, we can associate a number $κ[y_I]$ by requiring that $κ[y_I] = E[y^I] − \sum_{E:x∈E⊂I} E[y^I,E]κ[y_E]$ with $x = (1,i)$. The definition is used inductively in $|I| ≥ 1$ and it has a unique solution. (Note that the empty cumulant $κ[y_0]$ never appears in the moments-to-cumulants formula, and for our purposes it can be left undefined. To be consistent with the derivatives of the generating function, we may for instance set $κ[y_0] := 0$.) Therefore, these numbers have to coincide with the standard cumulants in the case when exponential moments exist and hence (A.1) holds.

The following known properties of cumulants can then be derived directly from the above definition using induction in $|I|$ and applying techniques similar to what we have used for Wick polynomials in Section III A:

1. The cumulants are multilinear, in the same manner as was stated for Wick polynomials in Proposition III.7.

2. The moments-to-cumulants expansion (III.1) holds.
3. The cumulants are permutation invariant: if $I'$ is a permutation of $I$, then $\kappa[y_{I'}] = \kappa[y_I].$

4. If joint exponential moments exist, then $\kappa[y_I] = \partial_I^y g_c(0)$ with $g_c(\lambda) := \ln \mathbb{E}[e^{\lambda y}].$

However, let us skip the proofs here. In the text, we assume these results to be known and refer to the references for details of their proofs.

**Appendix B: Classical particle system with random initial data**

Consider the evolution of $N$ classical particles interacting via a polynomial interaction potential, with the initial data given by some random probability measure. We show here how it can be recast in the form of the evolution equation discussed in Section IV.

We consider the random variables $y_j$, indexed by $J = \{(i, n)\}_{i,n}$, where $n$ is one of the $N$ different particle labels and $i = 1, 2$ differentiates between the particle position and momentum: we define $y_{(1,n)}(t) := q_n(t)$ and $y_{(2,n)}(t) := p_n(t).$ If all particles move in $\mathbb{R}$, have the same mass, and have only pair interactions via the potential $V(q) := \sum_{n',n,n'\neq n} \lambda_{n,n'} \frac{1}{2!}(q_n - q_{n'})^2,$ $a \geq 2$ even and $\lambda_{n,n} = \lambda_{n'n'},$ then we have $\partial_t q_n(t) = p_n(t)$ and

$$\partial_t p_n(t) = -\sum_{n' \neq n} \lambda_{n,n'}(q_n(t) - q_{n'}(t))^{a-1}. \quad (B.1)$$

Here $(q_n - q_{n'})^{a-1} = \sum_{k=0}^{a-1} \frac{1}{(-1)^{a-1-k}} (\begin{array}{c} a-1-k \\ k \end{array}) q_n^k q_{n'}^{a-1-k},$ and if we define $I_{n,n',k}$ as a sequence of length $a - 1$ containing first $k$ repetitions of $(1, n)$ and then $a - 1 - k$ repetitions of $(1, n'),$ then by (III.5) we have $q_n(t)^k q_{n'}(t)^{a-1-k} = \sum V \in I_{n,n',k} y(t)^V \mathbb{E}(y(t)^{I_{n,n',k} \setminus V}).$ Define thus as the collection $I_{(2,n)}$ all such sequences $U$ which contain $k_1$ repetitions of $(1, n)$ followed by $k_2$ repetitions of $(1, n')$ where $n' \neq n,$ $k_1, k_2 \geq 0,$ and $k_1 + k_2 \leq a - 1.$ Set also for each $U \in I_{(2,n)}$

$$M^U_{(2,n)}(t) := \lambda_{n,n'}^{a-1-k} \sum_{k=k_1}^{a-1-k} \left( \frac{a-1-k}{k} \right) (-1)^{a-k} \sum_{V \subset I_{n,n'}^{a,k}} \mathbb{E}(y(t)^{I_{n,n',k} \setminus V}) \mathbb{I}(V = U). \quad (B.2)$$

Therefore, (IV.3) holds for all $j \in J$ after we also define $I_{(1,n)} := \{\emptyset, ((2,n))\}$ and set $M^\emptyset_{(1,n)}(t) := \mathbb{E}(y_{(2,n)}(t))$ and $M^U_{(1,n)}(t) := 1$ if $U = ((2,n)).$

**Appendix C: Estimation of the first order non-pairing contributions to (V.18)**

In this appendix, we show how to estimate the first order non-pairing contraction terms in (V.18). To this end we need to make an assumption on the dispersion relation $\omega(k).$ Let us consider the so called free propagator

$$p_t(x) = \int_{\mathbb{R}^d} dke^{ikx} e^{-it\omega(k)}. \quad (C.1)$$

As in the assumption “(DR2)” in Ref. 9, we now suppose that there are $C, \delta > 0$ such that for all $t \in \mathbb{R},$

$$\|p_t\|_3^3 = \sum_{x \in \mathbb{Z}^d} |p_t(x)|^3 \leq C(1 + t^2)^{-(1+\delta)/2}. \quad (C.2)$$

Furthermore, we assume also the already mentioned $\ell_1$-clustering property (see section IV) which we slightly rephrase as follows for each cumulant of order $n$: we require that

$$\|\kappa_n\|_1 := \sup_{\sigma \in \{\pm 1\}^n} \sum_{x \in (\mathbb{Z}^d)^n} \mathbb{I}(x_1 = 0) |\kappa(\psi(x_1, \sigma_1), \ldots, \psi(x_n, \sigma_n))| < \infty. \quad (C.3)$$
We recall that the physical meaning of this condition is that the cumulants decay fast enough in space so that they are summable, once the translational invariance is taken into account.

We recall that the first order non-pairing contributions in (V.18) are

\[-i\lambda\sigma \int_{0}^{t} dt \int (T^{3})^3 dk_1dk_2dk_3 \delta(k - k_1 - k_2 - k_3)e^{is(\omega_1 + \omega_2 - \omega_3)} \times \kappa[a(k_1, -1); a(k_2, \sigma); a(k_3, 1); a(k', \sigma')] \]

(C.4)

and a term which is obtained from (C.4) by swapping \((k, \sigma) \leftrightarrow (k', \sigma')\). As stated in (IV.1), by translation invariance we have

\[\kappa[a(k_1, -1); a(k_2, \sigma); a(k_3, 1); a(k', \sigma')] = \delta(k_1 + k_2 + k_3 + k')\tilde{F}(k_1, k_2, k_3, \sigma, \sigma') \]

(C.5)

where

\[F(x_1, x_2, x_3, x_4, \sigma, \sigma') = \| (x_1 = 0)\kappa[a(x_1, -1); a(x_2, \sigma); a(x_3, 1); a(x_4, \sigma')] \].

Clearly, \(\|F\|_1 \leq \|\kappa_4\|_1 < \infty\) by the assumed \(\ell_1\)-clustering.

Therefore, the term in (C.4) is bounded by

\[\lambda \left| \int_{0}^{t} dt \int (T^{3})^3 dk_1dk_2dk_3 \delta(k - k_1 - k_2 - k_3)e^{is(\omega_1 + \omega_2 - \omega_3)} \times \delta(k_1 + k_2 + k_3 + k')\tilde{F}(k_1, k_2, k_3, k', \sigma, \sigma') \right| \]

\[\leq \lambda\|k + k'\| \int_{0}^{t} dt \int (T^{3})^3 dk_1dk_2e^{is(\omega_1 - \omega_2 - \omega_3)}\tilde{F}(k_1, k_2, k_3, k', \sigma, \sigma')|k_3 = k - k_1 - k_2| \]

\[\leq \lambda\|k + k'\| \int_{0}^{t} dt \int (T^{3})^3 dk_1dk_2e^{is(\omega_1 - \omega_2 - \omega_3)}\tilde{F}(k_1, k_2, k_3, k', \sigma, \sigma')|k_3 - k_1 - k_2| \]

\[\leq \lambda\|k + k'\|\|\kappa_4\|_1 \int_{0}^{t} ds \|p_3\|_3^4 \leq \lambda C\|k + k'\|\|\kappa_4\|_1 \int_{0}^{t} ds(1 + s^2)^{-(1+\delta)/2} \]

\[\leq \lambda C'\|k + k'\|\|\kappa_4\|_1 \]

(C.6)

where \(C'\) is a constant which depends only on \(C\) and \(\delta\). We have used the inverse Fourier transform of \(\tilde{F}\) and (C.1) in the second inequality and Hölder’s inequality in the third one. Since the bound is invariant under the swap \((k, \sigma) \leftrightarrow (k', \sigma')\), it bounds also the second non-pairing contribution in (V.18). Therefore, we see that the first order contributions are \(O(\lambda)\) uniformly in \(t\).

REFERENCES


[II]

Summability of joint cumulants of nonindependent lattice fields

Jani Lukkarinen, Matteo Marcozzi, Alessia Nota
Summability of joint cumulants of nonindependent lattice fields

Jani Lukkarinen*, Matteo Marcozzi†, Alessia Nota‡

* University of Helsinki, Department of Mathematics and Statistics
P.O. Box 68, FI-00014 Helsingin yliopisto, Finland
† University of Bonn, Institute for Applied Mathematics
Endenicher Allee 60, D-53115 Bonn, Germany
‡ University of Bonn, Institute for Applied Mathematics
Endenicher Allee 60, D-53115 Bonn, Germany

Abstract

We consider two nonindependent random fields $\psi$ and $\phi$ defined on a countable set $Z$. For instance, $Z = \mathbb{Z}^d$ or $Z = \mathbb{Z}^d \times I$, where $I$ denotes a finite set of possible “internal degrees of freedom” such as spin. We prove that, if the cumulants of both $\psi$ and $\phi$ are $\ell_1$-clustering up to order $2n$, then all joint cumulants between $\psi$ and $\phi$ are $\ell_2$-summable up to order $n$, in the precise sense described in the text. We also provide explicit estimates in terms of the related $\ell_1$-clustering norms, and derive a weighted $\ell_2$-summation property of the joint cumulants if the fields are merely $\ell_2$-clustering. One immediate application of the results is given by a stochastic process $\psi_t(x)$ whose state is $\ell_1$-clustering at any time $t$: then the above estimates can be applied with $\psi = \psi_t$ and $\phi = \psi_0$ and we obtain uniform in $t$ estimates for the summability of time-correlations of the field. The above clustering assumption is obviously satisfied by any $\ell_1$-clustering stationary state of the process, and our original motivation for the control of the summability of time-correlations comes from a quest for a rigorous control of the Green-Kubo correlation function in such a system. A key role in the proof is played by the properties of non-Gaussian Wick polynomials and their connection to cumulants.

1 Introduction and physical motivation

In many problems of physical interest, the basic dynamic variable is a random field. In addition to proper stochastic processes, such as particles evolving according to Brownian motion, the random field could describe for instance a density of particles of a Hamiltonian system with random initial data or after time-averaging.

One particular instance of the second kind is the Green-Kubo formula which connects the transport coefficients, such as thermal conductivity, to integrals over equilibrium time-correlations of the current observable of the relevant conserved quantity, for instance, of the energy current. The equilibrium time-correlations are cumulants of current fields between time zero and some later time. The current fields are generated by distributing the initial data according to some fixed equilibrium measure and then solving the evolution equations: this yields a random field, even when the time-evolution itself is deterministic.

Hence, the control of correlation functions, i.e., cumulants, of random fields is a central problem for a rigorous study of transport properties. One approach, which has been used both in practical applications and in direct mathematical studies, is given by Boltzmann transport equations. It is usually derived from the microscopic system by using moment hierarchies, such as the BBGKY
hierarchy, and then ignoring higher order moments to close the hierarchy of evolution equations. Although apparently quite powerful a method, it has not been possible to give any meaningful general estimates for the accuracy or for regions of applicability of such closure approximations.

The present work arose as part of a project aiming at a rigorous derivation of a Boltzmann transport equation in the kinetic scaling limit of the weakly nonlinear discrete Schrödinger equation (DNLS). This system describes the evolution of a complex lattice field \( \psi_t(x) \), with \( x \in \mathbb{Z}^d \) and \( t \geq 0 \), by requiring that it satisfies the Hamiltonian evolution equations

\[
 i \partial_t \psi_t(x) = \sum_{y \in \mathbb{Z}^d} \alpha(x - y) \psi_t(y) + \lambda |\psi_t(x)|^2 \psi_t(x),
\]

where the function \( \alpha \) determines the “hopping amplitudes” and \( \lambda > 0 \) is a coupling constant. A kinetic scaling limit with a suitably chosen closure assumption predicts that the Green-Kubo correlation function of the energy density of this system satisfies a linearized phonon Boltzmann equation in the limit; the explicit form of the Boltzmann collision operator and discussion about the approximations involved is given in Sections 5 and 6 in [1], and we refer to [2, 3] for more details about the linearization procedure.

The method used in the derivation of the transport properties in such weak coupling limits are, naturally, based on perturbation expansions. Advances have recently been made in controlling the related oscillatory integrals (see for instance [4, 5, 6]), but for nonlinear evolution equations of the present type a major obstacle has been the lack of useful \textit{a priori} bounds for the correlation functions. For instance, Schwarz inequality estimates of moments in the “remainder terms” of finitely expanded moment hierarchies has been used for this purpose for time-stationary initial data in [5], which was inspired by the bounds from unitarity of the time-evolution of certain linear evolution equations first employed for the random Schrödinger equation in [4] and later extended to other similar models such as the Anderson model [7] and a classical harmonic lattice with random mass perturbations [8]. However, as argued in [1], using moments instead of cumulants to develop the hierarchy could lead to loss of an important decay property which is valid for cumulants but not for moments; we shall discuss this point further in Section 2.

In the present contribution we derive a generic result which allows to bound joint correlations of two random fields in terms of estimates involving only the decay properties of each of the fields separately. These estimates are immediately applicable for obtaining \textit{uniform in time a priori} bounds for time-correlation functions of time-stationary fields. In particular, they imply that if the initial state of the field is distributed according to an equilibrium measure which is \( \ell_1 \)-clustering, then all time-correlations are \( \ell_2 \)-summable. The precise assumptions are described in Section 2 and the result in Theorem 2.1 there.

If both fields are Gaussian and translation invariant, more direct estimates involving discrete Fourier-transform become available. We use this in Section 3 to give an explicit example which shows that \( \ell_1 \)-clustering of the fields does not always extend to their joint correlations, hence showing that the increase of the power from \( \ell_1 \)-clustering to \( \ell_2 \)-summability of the joint correlations in the main theorem is not superfluous.

The result is a corollary of a bound which proves summability of cumulants of any observable with finite variance with \( \ell_p \)-clustering fields, for \( p = 1 \) and \( p = 2 \). The \( p = 2 \) case is more involved than the \( p = 1 \) case, since the present bound requires taking the sum in a weighted \( \ell_2 \)-space. The precise statements and all proofs are given in Section 4. The proof is based on representation of cumulants using Wick polynomials. We rely on the results and notations of [1], and for convenience of the reader we have summarized the relevant items in Appendix A. We also present a few immediate applications of these bounds and discuss possible further applications in Section 5.

\textbf{Acknowledgements}

We thank Antti Kupiainen, Sergio Simonella, and Herbert Spohn for useful discussions on the topic. The work has been supported by the Academy of Finland via the Centre of Excellence.
in Analysis and Dynamics Research (project 271983) and from an Academy Project (project 258302). The work and the related discussions have partially occurred in workshops supported by the French Ministry of Education through the grant ANR (EDNHS) and also by the Erwin Schrödinger Institute (ESI), Vienna, Austria.

2 Notations and mathematical setting

We consider here complex lattice fields $\psi : Z \to \mathbb{C}$ where $Z$ is any nonempty countable index set. We focus on this particular setup since it is the one most directly relevant for physical applications: common examples would be $Z = \mathbb{Z}^d$ and $Z = \mathbb{Z}^d \times I$, where $I$ denotes a finite set of possible “internal degrees of freedom” such as spin. The setup can also cover more abstract index sets, such as the sequence of coefficients in the Karhunen–Loève decomposition of a stochastic process [9, 10, 11], or distribution-valued random fields evaluated at suitably chosen sequence of test-functions (details about the definition and properties of general random fields can be found for instance in [12, 14] and in other sources discussing the Bochner–Minlos theorem).

We also assume that the field is closed under complex conjugation: to every $x \in Z$ there is some $x^* \in Z$ for which $\psi(x^*) = \psi(x)$. If needed, this can always be achieved by replacing the original index set $Z$ by $Z \times \{-1, 1\}$ and defining a new field $\Psi$ by setting $\Psi(x, 1) = \psi(x)$ and $\Psi(x, -1) = \psi(x)^*$. This procedure was in fact used in [1, 5] to study the DNLS example mentioned above, resulting in the choice $Z = \mathbb{Z}^d \times \{-1, 1\}$.

A random lattice field on $Z$ is then a collection of random variables $\psi(x)$, $x \in Z$, on the probability space $(\Omega, \mathcal{M}, \mu)$, where $\Omega$ denotes the sample space, $\mathcal{M}$ the $\sigma$-algebra of measurable events, and $\mu$ the probability measure. We consider here two random fields $\psi$ and $\phi$ which are defined on the same probability space. We denote the expectation over the measure $\mu$ by $\mathbb{E}$.

The $n$:th connected correlation function $\kappa_n$ of the field $\psi$ is a map $\kappa_n : \mathbb{Z}^n \to \mathbb{C}$ which is defined as the cumulant of the $n$ random variables obtained by evaluating the field at the argument points; explicitly,

$$
\kappa_n(x) := \kappa[\psi(x_1), \ldots, \psi(x_n)], \quad x \in \mathbb{Z}^n.
$$

We employ here the notations and basic results for cumulants and the related Wick polynomials, as derived in [1]; a summary of these is also included in Appendix A.

In physics, one often encounters random fields defined on the $d$-dimensional cubic lattice, with $Z = \mathbb{Z}^d$. One could then study the decay properties of such functions as $|x| \to \infty$ by using the standard $\ell_p$-norms over $(\mathbb{Z}^d)^n$. However, this is typically too restrictive for physical applications: it would imply in particular that both the first and the second cumulant, i.e., the mean and the variance, of the random variable $\psi(x)$ decay as $|x| \to \infty$, and thus the field would be almost surely “asymptotically zero” at infinity. Instead, many stationary measures arising from physical systems are spatially translation invariant: the expectation values remain invariant if all of the fields $\psi(x)$ are replaced by $\psi(x + x_0)$ for any given $x_0 \in \mathbb{Z}^d$. Since this implies also translation invariance of all correlation functions, they cannot decay at infinity then, unless the field is almost surely zero everywhere.

To cover also such nondecaying stationary states, one uses instead of the direct $\ell_p$-norms of the function $\kappa_n$, the so-called $\ell_p$-clustering norms of the field $\psi$ defined as follows: for $1 \leq p < \infty$ and $n \in \mathbb{N}_+$ we set

$$
\|\psi\|_p^{(n)} := \sup_{x_0 \in \mathbb{Z}} \left( \sum_{x \in \mathbb{Z}^{n-1}} |\kappa[\psi(x_0), \psi(x_1), \ldots, \psi(x_{n-1})]|^p \right)^{1/p},
$$

and define analogously $\|\psi\|_\infty^{(1)} := \sup_{x_0 \in \mathbb{Z}} |\mathbb{E}[\psi(x_0)]|$ (note that $\kappa[X] = \mathbb{E}[X]$ for any random variable $X$). We shall also use the corresponding $p = \infty$ norms, which coincide with the standard sup-norms of $\kappa_n$, namely, $\|\psi\|_\infty^{(n)} = \sup_{x \in \mathbb{Z}^n} |\kappa[\psi(x_1), \psi(x_2), \ldots, \psi(x_n)]| = \|\kappa_n\|_\infty$. Since the
norms concern $L^p$-spaces over a counting measure, they are decreasing in $p$, i.e., $\|\psi\|_p^{(n)} \geq \|\psi\|_{p'}^{(n)}$ if $p \leq p'$. (This follows from the bound $|u_n(x)| \leq \|\psi\|_p^{(n)}$, valid for all $x$ and $p$.)

For a translation invariant measure with $Z = \mathbb{Z}^d$, we can translate $x_0$ to the origin in the definition (2.2), and, by a change of variables $x_n = x_0 + y_n$, obtain the simpler expression

$$\|\psi\|_p^{(n)} = \left( \sum_{y \in (\mathbb{Z}^d)^{n-1}} |\kappa[\psi(0), \psi(y_1), \ldots, \psi(y_{n-1})]|^p \right)^{1/p}. \tag{2.3}$$

The summation here goes over the displacements $y_i$ of the argument $x$, from the reference position $x_0 = 0$. The definition is tailored for random fields which become asymptotically independent for far apart regions of the lattice, i.e., when $|y_i| \to \infty$ above. For translation non-invariant measures, finiteness of the norm (2.2) yields a uniform estimate for the speed of asymptotic independence of the field. Let us use the opportunity to stress that it is crucial to use the cumulants, not moments, above: similar moments of the field would not decay as the separation grows, even if the field values would become independent (see [1] for more discussion about this point).

We now call a random field $\psi$ $\ell_p$-clustering if $\|\psi\|_p^{(n)} < \infty$ for all $n = 1, 2, \ldots$. In particular, this requires that all of the cumulants, which define the connected correlation functions $u_n$, need to exist. From the iterative definition of cumulants mentioned in the Appendix, or from the inversion formula expressing cumulants in terms of moments, it clearly suffices that $E[|\psi(x)|^n] < \infty$ for all $x \in Z$. We also say that the field $\psi$ is $\ell_p$-clustering up to order $m$ if $\|\psi\|_p^{(n)} < \infty$ for all $n \leq m$. For such a field, we use the following constants to measure its “magnitude”: we set

$$M_N(\psi; p) := \max_{1 \leq n \leq N} \left( \frac{1}{n!} \|\psi\|_p^{(n)} \right)^{1/n}. \tag{2.4}$$

Clearly, the definition yields an increasing sequence in $N$ up to the same order in which the field is $\ell_p$-clustering. We use the constants $M_N$ to control the increase of the clustering norms. It is conceivable that in special cases other choices beside (2.4) could be used with the estimates below to arrive at sharper bounds than those stated in the theorems. However, the above choice is convenient for our purposes since it leads to simple combinatorial estimates, increasing typically only factorially in the degree of the cumulant. It is possible to think of the numbers $M_N$ as measuring the range of values the field can attain. For instance, if $Z = \{0\}$ and $\psi(0)$ is a random variable which almost surely belongs to the interval $[-R, R]$ with $R > 0$, then $M_n(\psi; p)$ is independent of $p$ (since there is only one point 0) and $M_n = c_n R$ where $c_n$ remains order one, uniformly in $n$ (see, for instance, Lemma C.1 in [8]).

After these preliminaries, we are ready to state the main result:

**Theorem 2.1** Suppose $\psi$ and $\phi$ are random lattice fields which are closed under complex conjugation and defined on the same probability space. Assume that $\phi$ is $\ell_1$-clustering and $\psi$ is $\ell_\infty$-clustering, both up to order $2N$ for some $N \in \mathbb{N}_+$. Then their joint cumulants satisfy the following $\ell_2$-estimate for any $n, m \in \mathbb{N}_+$ for which $n, m \leq N$,

$$\sup_{x' \in \mathbb{Z}^m} \left[ \sum_{x \in \mathbb{Z}^n} |\kappa[\psi(x'_1), \ldots, \psi(x'_m), \phi(x_1), \ldots, \phi(x_n)]|^2 \right]^{1/2} \leq (M_{m,n} \gamma^m)^{n+m}(n + m)!, \tag{2.5}$$

where $M_{m,n} := \max(M_{2m}(\psi; \infty), M_{2n}(\phi; 1))$ and $\gamma = 2e \approx 5.44$. In particular, all of the above sums are then finite.

Loosely speaking, one can say that an $\ell_1$-clustering random field can have at worst $\ell_2$-summable joint correlations. We have stated the result in a form which assumes that the field $\psi$ is $\ell_\infty$-clustering. As mentioned above, the clustering norms are decreasing in the index; hence, the above result also holds if $\psi$ is $\ell_q$-clustering for any $1 \leq q < \infty$. One could then also replace the constants $M_{m,n}$ using the corresponding $\ell_q$-clustering norms, $\max(M_{2m}(\psi; q), M_{2n}(\phi; 1))$. However, these constants are always larger than $M_{m,n}$ and thus can only worsen the bound.
This result is a consequence of a more general covariance bound given in Theorem 4.1. There we also give a version of the estimate for fields \( \phi \) which are merely \( \ell_2 \)-clustering. The price to pay for the relaxation of the norms is an appearance of a weight factor in the \( \ell_2 \)-summation, see Theorem 4.3 for the precise statement. Before going into the details of the proofs, let us go through a special case clarifying the assumptions and the result.

3 An example: translation invariant Gaussian lattice fields

In this section, we consider real valued Gaussian random fields \( \psi \) and \( \phi \) on \( Z = \mathbb{Z} \) and assume that both fields have a zero mean and are invariant under spatial translations. Their joint measure is then determined by giving three functions \( F_1, F_2, G \in \ell_2(\mathbb{Z}, \mathbb{R}) \) for which

\[
\langle \psi(x)\psi(y) \rangle = F_1(x - y), \quad \langle \phi(x)\phi(y) \rangle = F_2(x - y), \quad \langle \psi(x)\phi(y) \rangle = G(x - y).
\]

(3.1)

The covariance operator needs to be positive semi-definite. By first using Parseval’s theorem and then computing the eigenvalues of the remaining \( 2 \times 2 \)-matrix, we find that this is guaranteed by requiring that the Fourier-transforms of the above functions, all of which belong to \( L^2(\mathbb{T}) \), satisfy almost everywhere

\[
\hat{F}_1(k) \geq 0, \quad \hat{F}_2(k) \geq 0, \quad |\hat{G}(k)|^2 \leq \hat{F}_1(k)\hat{F}_2(k).
\]

(3.2)

These three conditions hence suffice for the existence of a unique Gaussian measure on functions on \( \mathbb{Z} \) satisfying (3.1); details about such constructions are given for instance in [13, 14, 15].

The last condition restricts the magnitude of the correlations, and it implies that if each of the above fields is \( \ell_2 \)-clustering, then their correlations are \( \ell_2 \)-summable (simply because then \( \hat{G}(k) \in L^2(\mathbb{T}) \), and thus its inverse Fourier transform gives a function \( G \in \ell_2(\mathbb{Z}) \)). Hence, one might wonder if the main theorem could, in fact, be strengthened to show that \( \ell_1 \)-clustering of the fields implies \( \ell_1 \)-summability of the joint correlations. The following example shows that this is not the case.

3.1 \( \ell_1 \)-clustering fields whose joint correlations are not \( \ell_1 \)-summable

Let us consider two i.i.d. Gaussian fields \( \psi \) and \( \phi \) whose correlations are determined by the function

\[
G(x) = \frac{1}{\pi x} \sin \left( \frac{\pi}{2} x \right), \quad x \neq 0, \quad G(0) = \frac{1}{2}.
\]

(3.3)

For such i.i.d. fields \( F_1(x) = 1 (x = 0) = F_2(x) \) which is equivalent to \( \hat{F}_1(k) = 1 = \hat{F}_2(k) \) for all \( k \in \mathbb{T} \). Now for all \( x \in \mathbb{Z} \), clearly

\[
G(x) = \int_{-1/4}^{1/4} dk e^{i2\pi x k},
\]

(3.4)

and thus \( \hat{G}(k) = 1(|k| < \frac{1}{4}) \leq 1 = \sqrt{\hat{F}_1(k)\hat{F}_2(k)} \). Therefore, such \( G \) indeed defines a possible correlation between the fields \( \psi \) and \( \phi \).

For such Gaussian fields, all cumulants of order different from \( n = 2 \) are zero. We also have \( \sup_{x \in \mathbb{Z}} \sum_{y \in \mathbb{Z}} |F_1(x - y)| = 1 \), and, as \( F_2 = F_1 \), both fields are \( \ell_1 \)-clustering, with \( \|\phi\|_1^{(2)} = 1 = \|\psi\|_1^{(2)} \) and \( \|\psi\|_1^{(n)} = 0 = \|\phi\|_1^{(n)} \) for any other \( n \). However, their joint correlations satisfy for any \( x' \in \mathbb{Z} \)

\[
\sum_{x \in \mathbb{Z}} |\kappa[\psi(x'), \phi(x)]| = \sum_{y \in \mathbb{Z}} |G(y)| = \frac{1}{2} + 2 \sum_{y=1}^{\infty} \frac{1}{\pi y} \left| \sin \left( \frac{\pi}{2} y \right) \right| = \frac{1}{2} + \frac{2}{\pi} \sum_{n=0}^{\infty} \frac{1}{2n+1} = \infty.
\]

(3.5)

Thus the joint correlations are not \( \ell_1 \)-summable.

In contrast, \( \sup_{x'} \sum_x |\kappa[\psi(x'), \phi(x)]|^2 < \infty \), since it is equal to \( \sum_y |G(y)|^2 \) and \( G \in \ell^2(\mathbb{Z}) \).
4 $\ell_2$-summability of joint correlations of $\ell_p$-clustering fields

Theorem 4.1 Consider a random lattice field $\phi$ on a countable set $Z$, defined on a probability space $(\Omega, \mathcal{F}, \mu)$ and closed under complex conjugation. Suppose that $\phi$ is $\ell_p$-clustering up to order $2N$ for some $N \in \mathbb{N}_+$, and let $M_N(\psi; p)$ be defined as in (2.4). Suppose also $X \in L^2(\mu)$, i.e., $X$ is a random variable with finite variance.

1. If $p = 1$ and $n \leq N$, we have a bound

$$\left[ \sum_{x \in \mathbb{Z}^n} |\kappa[X, \phi(x_1), \ldots, \phi(x_n)]|^2 \right]^{1/2} \leq \sqrt{\text{Cov}(X^*, X)M_{2n}(\phi; 1)} e^a \sqrt{(2n)!}. \quad (4.1)$$

2. If $p = 2$ and $n \leq N$, we have a bound

$$\sup_{x' \in \mathbb{Z}^n} \left[ \sum_{x \in \mathbb{Z}^n} |\Phi_n(x', x)||\kappa[X, \phi(x_1), \ldots, \phi(x_n)]|^2 \right]^{1/2} \leq \sqrt{\text{Cov}(X^*, X)M_{2n}(\phi; 2)}^{2n} e^{2n}(2n)!,$$

where $\Phi_n(x', x) := \mathbb{E}[\phi(x'_1)^* \phi(x'_2)^* \cdots \phi(x'_n)^* : \phi(x_1)\phi(x_2) \cdots \phi(x_n);]$. \quad (4.2)

The key argument in the proof uses Wick polynomial representation of the above cumulants.

Namely, a direct consequence of the truncated moments-to-cumulants formula given in Proposition A.1 in the Appendix, is that

$$\kappa[X, \phi(x_1), \ldots, \phi(x_n)] = \mathbb{E}[X : \phi(x_1)\phi(x_2) \cdots \phi(x_n);] = \mathbb{E}[X : : \phi(x_1)\phi(x_2) \cdots \phi(x_n);], \quad (4.3)$$

The Proposition can be applied here since now $\mathbb{E}[|X| \prod_{i=1}^n |\phi(x_i)|] < \infty$ by the Schwarz inequality estimate $\mathbb{E}[|X| \prod_{i=1}^n |\phi(x_i)|]^2 \leq \mathbb{E}[|X|^2] \mathbb{E}[\prod_{i=1}^n |\phi(x_i)|]^2$ where the first factor is finite since $X \in L^2(\mu)$, and the second factor is finite since $\phi$ is assumed to be $\ell_p$-clustering up to order $2n$.

Applying Schwarz inequality in (4.3) yields a bound

$$|\kappa[X, \phi(x_1), \ldots, \phi(x_n)]|^2 \leq \mathbb{E}[|X|^2] \mathbb{E}[|\phi(x_1)\phi(x_2) \cdots \phi(x_n);] = \text{Cov}(X^*, X)\Phi_n(x, x). \quad (4.4)$$

Hence, the theorem is obviously true if $\Phi_n(x, x)$ decreases sufficiently rapidly with “increasing” $x$. However, this is typically too restrictive: since $\Phi(x, x) = \mathbb{E}[|\phi(x);|^2] = \text{Var}(\phi(x))$, this would require that the field $\phi$ becomes asymptotically deterministic. The proof below combines suitably chosen test functions with the above Schwarz estimate and results in bounds which only require summability of $\Phi_n(x', x)$ in $x$ for a fixed $x'$. Such summability is guaranteed by the $\ell_p$-clustering of the field, and the rest of the proof consists of controlling the combinatorial factors which relate these two concepts together, cf. Lemma 4.2.

Let us stress that the above result is typically not true if moments are used there instead of cumulants. The above Schwarz inequality estimates would be straightforward for moments; in fact, such a Schwarz estimate was a key method in [5] to separate time-evolved fields from their time-zero counterparts in products of these fields. However, the functions resulting from such Schwarz estimates are of the type $\mathbb{E}[\prod_{i=1}^n |\phi(x_i)|^2]$ and for these to be summable in $x$ the field not only has to become asymptotically deterministic, but it has to even vanish. Cumulants of $\ell_p$-clustering fields would, on the other hand, be summable, but there is no obvious way of generalizing the Schwarz inequality bounds for cumulants. The missing ingredient is here provided by the Wick polynomial representation (4.3).

Proof: There is a natural Hilbert space structure associated with correlations of the present type. We begin with test-functions $f : \mathbb{Z}^n \rightarrow \mathbb{C}$ which have a finite support, and define for them a (semi-)norm by the formula

$$\|f\|_{2,n}^2 := \mathbb{E}\left[ \sum_{x \in \mathbb{Z}^n} f(x)\phi(x)^{J_n}; \right]^2 = \sum_{x',x \in \mathbb{Z}^n} f(x')^* f(x) \Phi_n(x', x),$$

$$\Phi_n(x', x) := \mathbb{E}[\phi^*(x')^{J_n} \phi(x)^{J_n}], \quad (4.5)$$

6
where \( J_n = \{1, 2, \ldots, n\} = J'_n \), and thus we have \( \phi(x)^{J_n} := \phi(x_1)\phi(x_2) \cdots \phi(x_n) \), \( \phi^*(x')^{J'_n} := \phi(x'_1)^*\phi(x'_2)^* \cdots \phi(x'_n)^* \). For the definition, we do not yet need any summability properties of the field \( \phi \), it suffices that all the expectations in \( \Phi_n(x', x) \) are well-defined for all \( x', x \). By the truncated moment-to-cumulants expansion of Wick polynomials, as given in Proposition A.1 in the Appendix, we have here

\[
\Phi_n(x', x) = \sum_{\pi \in \P(J'_n + J_n)} \prod_{S \in \pi} (\kappa[\phi^*(x'), \phi(x) A] \mathbb{I}(A' \neq \emptyset, A \neq \emptyset))_{A' = S|J'_n, A = S|J_n}, \tag{4.6}
\]

where the notation \( S|J_n \) refers to the subsequence composed out of the indices belonging to \( J_n \) in the cluster \( S \) of the partition \( \pi \) of \( J'_n + J_n \). The additional restrictions \( A', A \neq \emptyset \) in the product arise from the fact that if either of them is violated, then the corresponding cluster \( S \) is contained entirely in either \( J'_n \) or \( J_n \), and vice versa. The partitions containing such a cluster are precisely those which are missing from the moments to cumulants formula by the Wick polynomial construction. Therefore, \( \Phi_n \) is finite, as soon as all cumulants up to order \( 2n \) are finite. On the other hand, this is already guaranteed by the assumed \( \ell_2 \)-clustering of the field \( \phi \). For notational simplicity, let us drop the name of the field \( \phi \) from the norm \( \| f \|_{o,n} \).

The norm can be associated with a polaroid property using the polarization identity, and we can then use it to define a Hilbert space \( \mathcal{H}_n \) by completion and dividing out the functions with zero norm, if the above formula gives only a semi-norm. The elements of \( \mathcal{H}_n \) are thus functions \( f : \mathbb{Z}^n \to \mathbb{C} \) with \( \| f \| < \infty \) (or their equivalence classes in the semi-norm case when every \( f \) and \( g \) with \( \| f - g \| = 0 \) needs to be identified). However, since we do not use these Hilbert spaces directly, let us skip the details of the construction.

We begin with joint correlations of the type \( G(x) := \mathbb{E}[Y : \phi(x)^{J_n}] \) where \( Y \in \mathcal{L}^2(\mu) \) is a random variable. Here \( G(x) \) is well defined due to the Schwarz inequality estimate \( \mathbb{E}[\| Y : \phi(x)^{J_n} \|]^2 \leq \mathbb{E}[\| Y \|^2] \Phi_n(x, x) \). If \( f : \mathbb{Z}^n \to \mathbb{C} \) has a finite support, we define

\[
\Lambda[f] := \sum_{x \in \mathbb{Z}^n} G(x)f(x) = \mathbb{E} \left[ Y \sum_{x \in \mathbb{Z}^n} \phi(x)^{J_n} : f(x) \right]. \tag{4.7}
\]

Applying the Schwarz inequality as above yields an upper bound

\[
|\Lambda[f]|^2 \leq \mathbb{E}[\| Y \|^2] \mathbb{E} \left[ \left| \sum_{x \in \mathbb{Z}^n} f(x) : \phi(x)^{J_n} \right|^2 \right] = \mathbb{E}[\| Y \|^2] \| f \|_{o,n}^2. \tag{4.8}
\]

Here, by the definition of the norm, we obtain an unweighted \( \ell_2 \)-estimate by using Hölder’s inequality as follows

\[
\| f \|_{o,n}^2 \leq \sum_{x', x \in \mathbb{Z}^n} |f(x')||f(x)||\Phi_n(x', x)| \leq \sqrt{\sum_{x', x \in \mathbb{Z}^n} |f(x')|^2|\Phi_n(x', x)|} \sqrt{\sum_{x', x \in \mathbb{Z}^n} |f(x)|^2|\Phi_n(x', x)|} \\
\leq \sum_{x \in \mathbb{Z}^n} |f(x)|^2 \sup_{x' \in \mathbb{Z}^n} \sum_{x \in \mathbb{Z}^n} |\Phi_n(x', x)|, \tag{4.9}
\]

where we have used the obvious symmetry property \( \Phi_n(x', x)^* = \Phi_n(x, x') \). As shown below, in Lemma 4.2, \( \ell_1 \)-clustering of the field \( \phi \) in fact implies that there is \( c_n < \infty \) such that \( \sup_{x' \in \mathbb{Z}^n} \sum_{x \in \mathbb{Z}^n} |\Phi_n(x', x)| \leq c_n \) (the explicit dependence of \( c_n \) on the clustering norms is given in the Lemma). Hence, we can conclude that \( |\Lambda[f]| \leq \sqrt{c_n \mathbb{E}[\| Y \|^2]} \| f \|_{\ell_2} \). Thus, thanks to the Riesz representation theorem, \( \Lambda \) can be extended into a unique functional belonging to the dual of the Hilbert space \( \ell_2(\mathbb{Z}^n) \), and hence there is a vector \( \Psi \in \ell_2(\mathbb{Z}^n) \) such that \( \Lambda[f] = \sum_{x \in \mathbb{Z}^n} \Psi(x)^* f(x) \)

\(^1\)If one has distinct labels in \( J'_n \) and \( J_n \), achievable always by relabelling of one of the sets, one can safely take here \( J'_n + J_n = J'_n \cup J_n = \mathcal{P}(J'_n + J_n) \) equal to the ordinary partitions of the set \( J'_n + J_n \), and also \( S|J_n = S \cap J_n \). However, such relabellings lead to unnecessarily clumsy notations in the present case, and we have opted to use the above notations from [1].
and \( \|\Psi\|_{\ell_2} \leq c_n \sqrt{\mathbb{E}[|Y|^2]} \). Then necessarily \( G(x) = \Psi(x)^* \) for all \( x \), and thus \( G \in \ell_2(Z^n) \) as well, with a bound
\[
\sqrt{\sum_{x \in Z^n} |G(x)|^2} \leq c_n \sqrt{\mathbb{E}[|Y|^2]}.
\tag{4.10}
\]

If \( Y = :X:\), we have \( G(x) = \mathbb{E}[X: \phi(x)^*] = \kappa[X, \phi(x_1), \ldots, \phi(x_n)] \) as explained in (4.3), and also \( \mathbb{E}[|Y|^2] = \mathbb{E}[X^*:X:] = \kappa[X^*,X] = \text{Cov}(X^*,X) \). Hence, (4.10) implies the bound stated in the first item.

For the weighted result, we apply (4.8) for specially constructed test functions \( f \). Let \( F \) be any finite subset of \( Z^n \) and choose an arbitrary point \( y \in Z^n \). Then \( f(x) = \mathbb{1}(x \in F)|\Phi_n(y,x)|G(x)^* \) has finite support and
\[
\Lambda[f] = \sum_{x \in F} |G(x)|^2|\Phi_n(y,x)| \leq \sqrt{\mathbb{E}[|Y|^2]} \|f\|_n < \infty.
\tag{4.11}
\]

On the other hand, we obtain the following estimate for \( \|f\|_n \)
\[
\|f\|_n^2 = \sum_{x',x \in F} |G(x)^*G(x')|\Phi_n(y,x')|\Phi_n(y,x)|\Phi_n(x',x)
\leq \sqrt{\sum_{x',x \in F} |G(x)|^2|\Phi_n(y,x)|\Phi_n(y,x')|\Phi_n(x',x)} \sqrt{\sum_{x',x \in F} |G(x')|^2|\Phi_n(y,x')|\Phi_n(y,x)|\Phi_n(x',x)}
\leq \sum_{x \in F} |G(x)|^2|\Phi_n(y,x)| \left( \sup_{x' \in Z^n} \sqrt{\sum_{x \in F} |\Phi_n(x',x)|^2} \right)^2,
\tag{4.12}
\]
where we have used \( \Phi_n(x',x)^* = \Phi_n(x,x') \) and the Schwarz inequality in the last estimate.

As shown below, in Lemma 4.2, \( \ell_2 \)-clustering of the field \( \phi \) implies \( \sum_{x \in Z^n} |\Phi_n(x',x)|^2 \leq c'_n < \infty \) where the explicit dependence of \( c'_n \) on the clustering norms is given in the Lemma. Therefore, \( \Lambda[f] \leq c'_n \sqrt{\mathbb{E}[|Y|^2]} \sqrt{\Lambda[f]} \). Since \( 0 \leq \Lambda[f] < \infty \) for any subset \( F \), we can conclude that the estimate \( \sqrt{\Lambda[f]} \leq c'_n \sqrt{\mathbb{E}[|Y|^2]} \) also holds. Thus by using subsets \( F = F_R \), which are constructed by choosing the first \( R \) elements from a fixed enumeration of \( Z^n \), and then taking \( R \to \infty \), we obtain that
\[
\sqrt{\sum_{x \in Z^n} |G(x)|^2|\Phi_n(y,x)|} \leq c'_n \sqrt{\mathbb{E}[|Y|^2]} < \infty,
\tag{4.13}
\]
for all \( y \in Z^n \). This implies the statement in the second item.

**Lemma 4.2** Suppose that the field \( \phi \) is closed under complex conjugation and \( \ell_p \)-clustering up to order \( 2n \), for some \( p \in [1,\infty] \) and \( n \geq 1 \). Then, for any \( x' \in Z^n \),
\[
\|\Phi_n(x',\cdot)\|_{\ell_p} \leq \sum_{\pi \in \mathcal{P}(J_{2n})} \prod_{S \in \pi} \|\phi\|_{L_p} \leq M_{2n}(\phi;p)^{2n}e^{2n}(2n)!,
\tag{4.14}
\]
where \( J_{2n} = \{1,2,\ldots,2n\} \) and \( \mathcal{P}(J_{2n}) \) denotes the collection of its partitions.

**Proof:** Let us consider some fixed \( x' \in Z^n \). We apply the Minkowski inequality to (4.6), as a function of \( x \), and conclude that
\[
\|\Phi_n(x',\cdot)\|_{\ell_p} \leq \sum_{\pi \in \mathcal{P}(J_{2n} + J_{2n})} \|F(x',\cdot;\pi)\|_{\ell_p},
\tag{4.15}
\]

where
\[
F(x', x; \pi) := \prod_{S \in \pi} (|\kappa[\phi^*(x')A', \phi(x)A]|_1 (A' \neq \emptyset, A \neq \emptyset))_{A'=S|J'_n,A=S|J_n}. \tag{4.16}
\]

Let us first consider the case \( p < \infty \). For any \( \pi \in \mathcal{P}(J'_n + J_n) \) yielding a nonzero \( F \), the restrictions of its clusters with \( J_n, A = S|J_n \) in the above form, form a partition of \( J_n \). Let us denote this partition by \( \pi_2 \). Hence, we can use this partition to reorder the summation over \( x \in Z^n \) into iterative summation over \( x_A \in Z^A \) for \( A \in \pi_2 \). Applied to (4.16) this yields
\[
\sum_{x \in Z^n} |F(x', x; \pi)|^p = \prod_{S \in \pi} \left( \|A' \neq \emptyset, A \neq \emptyset \sum_{x_A \in Z^A} |\kappa[\phi^*(x')A', \phi(x)A]|^p \right)_{A'=S|J'_n,A=S|J_n}. \tag{4.17}
\]

Since the field \( \phi \) is closed under complex conjugation, for each \( S \in \pi \) the sum over \( x_A \) is equal to \( \sum_{x \in Z^A} |\kappa[\phi(y')J_{A'}, \phi(x)A] |^p \) where \( y' = ((x')_j)_{j \in A'} \). As \( A' \neq \emptyset \), we may choose an element \( j \in A' \). We then denote \( x_0 = (x')_j \). and estimate the sum with an \( \ell_p \)-clustering norm as follows
\[
\sum_{x \in Z^A} |\kappa[\phi^*(x')A', \phi(x)A]|^p \leq \sum_{y \in Z|A'|} \sum_{x \in Z^A} |\kappa[\phi(x_0), \phi(y)J_{A'}|1, \phi(x)A]|^p \leq \left( \|\phi\|_{\ell_p}^{(|A'| + |A|)} \right)^p. \tag{4.18}
\]

Since \( |A'| + |A| = |S| \), we can conclude that, if \( p < \infty \),
\[
\|F(x', \cdot ; \pi)\|_p \leq \prod_{S \in \pi} \|\phi\|_{\ell_p}^{(|S|)}. \tag{4.19}
\]

The corresponding estimate for \( p = \infty \) is a straightforward consequence of \( |\kappa[\phi^*(x')A', \phi(x)A]| \leq \|\phi\|_{\ell_p}^{(|A'| + |A|)} \) which was discussed in Section 2 after Eq. (2.2).

Therefore, we can now conclude that the first inequality in (4.14) holds. By the definition in (2.4), we can then apply an upper bound
\[
\|\phi\|_{\ell_p}^{(m)} \leq m!M_m(\phi; p)^m \leq (\|\phi\|_{\ell_p}^{(|A'| + |A|)})^m,
\]
for any \( m \leq 2n \). If \( \pi \in \mathcal{P}(J'_n + J_n) \), we have \( |S| \leq 2n \) for any \( S \in \pi \), and thus
\[
\prod_{S \in \pi} \|\phi\|_{\ell_p}^{(|S|)} \leq M_{2n}(\phi; p)^{|\pi|} \prod_{S \in \pi} |S|! = M_{2n}(\phi; p)^{2n} \prod_{S \in \pi} |S|!. \tag{4.20}
\]

A combinatorial estimate shows that
\[
\sum_{\pi \in \mathcal{P}(J_{2n})} \prod_{S \in \pi} |S|! \leq (2n)!(2n)^{2n} \tag{4.21}
\]
(a proof of the inequality is available for instance in the proof of Lemma 7.3 in [5]). Therefore, we have proven also the second inequality in (4.14), concluding the proof of the Lemma.

The following theorem contains the already stated Theorem 2.1 in the item 1. The remarks after the Theorem at the end of Section 2 hold also in this case. In particular, it is obviously valid for any \( \ell_q \)-clustering field \( \psi \), as long as \( 1 \leq q \leq \infty \).

**Theorem 4.3** Consider two random lattice fields \( \phi(x) \) and \( \psi(x) \), \( x \in Z \) for a countable \( Z \), defined on the same probability space \((\Omega, \mathcal{A}, \mu)\) and each closed under complex conjugation. Suppose that \( \phi \) is \( \ell_q \)-clustering and \( \psi \) is \( \ell_{\infty} \)-clustering up to order \( 2N \) for some \( N \in \mathbb{N}_+ \). Let \( M_N \) be defined as in (2.4). Then their joint cumulants satisfy the following \( \ell_2 \)-estimates for any \( n, m \in \mathbb{N}_+ \) for which \( n, m \leq N \):

\[
\]
1. If \( p = 1 \), we have a bound

\[
\sup_{x^* \in \mathbb{Z}^m} \left[ \sum_{x \in \mathbb{Z}^n} \left| \kappa[\psi(x_1^*), \ldots, \psi(x_m^*), \phi(x_1), \ldots, \phi(x_n)] \right|^2 \right]^{1/2} \leq (\mathfrak{M}_{m,n} \gamma^m)^{n+m}(n+m)! \tag{4.22}
\]

where \( \mathfrak{M}_{m,n} := \max(M_{2m}(\psi; \infty), M_{2n}(\phi; 1)) \) and \( \gamma = 2e \).

2. If \( p = 2 \), we have a bound

\[
\sup_{x^* \in \mathbb{Z}^m, y \in \mathbb{Z}^n} \left[ \sum_{x \in \mathbb{Z}^n} |\Phi_n(y, x)||\kappa[\psi(x_1^*), \ldots, \psi(x_m^*), \phi(x_1), \ldots, \phi(x_n)]|^2 \right]^{1/2} \leq (\mathfrak{M}_{m,n} \gamma^m)^{2(n+m)}((n+m)!)^2 \tag{4.23}
\]

where \( \Phi_n(y, x) := \mathbb{E}[\phi(y_1^*)\phi(y_2^*)\cdots\phi(y_n^*)\cdot\phi(x_1)\phi(x_2)\cdots\phi(x_n)] \), and we set \( \mathfrak{M}_{m,n} := \max(M_{2m}(\psi; \infty), M_{2n}(\phi; 2)) \) and \( \gamma = 2e \).

**Proof:** We will proceed by induction over \( m \). Let us recall the above definition of \( \Phi_n \) and define analogously \( \Psi_m \) as \( \Psi_m(y', x') := \mathbb{E}[\psi(y_1'\psi(y_2')\psi(y_n')\cdot\psi(x_1')\psi(x_2')\cdots\psi(x_m')] \). In particular, then we can apply Theorem 4.1 with \( X = \psi(x') \). By Lemma 4.2, then \( \mathbb{E}||X||^2 = \psi_1(x_1', x_1') \leq M_2(\psi; \infty)^2e^22! \), and thus, say for \( \gamma = 2e \), both items 1 and 2 can be seen to hold for \( m = 1 \) and any \( n \leq N \) thanks to Theorem 4.1 and the estimate \((2n)! \leq ((2n)!)^2 = 2^{2n}(n!)^2\).

As an induction hypothesis, we consider some \( 1 \leq m \leq N \) and assume that the thesis holds for values up to \( m - 1 \) with any \( n \leq N \). We also give the details only for the first \( \ell_1 \)-clustering case, i.e., with \( p = 1 \).

Let us decompose the cumulant using Proposition A.1. Namely, consider

\[
P(x', x) := \mathbb{E}[\psi(x_1')\psi(x_m')\cdot\phi(x_1)\cdot\phi(x_n)],
\]

for which

\[
\kappa[\psi(x_1')\psi(x_m')\cdot\phi(x_1)\cdot\phi(x_n)] = P(x', x) - Q(x', x),
\]

with

\[
Q(x', x) := \sum_{x \in \mathcal{P}(J_m + J_n)} \prod_{A \in \pi} \mathbb{E}[\kappa\psi(x_A')\phi(x_A)] = (\prod_{A \in \pi} \mathbb{E}[\kappa\psi(x_A')\phi(x_A)])_{A = S|J_m, A = S|J_n}
\]

Then we can conclude from the Minkowski inequality that

\[
\left[ \sum_{x \in \mathbb{Z}^n} |\kappa[\psi(x_1')\psi(x_m')\cdot\phi(x_1)\cdot\phi(x_n)]|^2 \right]^{1/2} \leq \|P(x', \cdot)\|_2 + \|Q(x', \cdot)\|_2. \tag{4.24}
\]

We now estimate \( \|P(x', \cdot)\|_2 \) using item 1 in Theorem 4.1 with \( X = \psi(x_1')\cdots\psi(x_m') \). Clearly, then

\[
P(x, x') = \mathbb{E}[X\cdot\phi(x_1)\cdot\phi(x_n)] = \kappa[X, \phi(x_1)\cdots\phi(x_n)], \tag{4.25}
\]

and, by applying Theorem 4.1 and Lemma 4.2

\[
\|P(x', \cdot)\|_2 \leq \sqrt{M_2(\psi; \infty)^2e^n \sqrt{2^n}!} \leq M_{2m}(\psi; \infty)^m M_{2n}(\phi; 1)^n e^{n+m} \sqrt{2^n}!10^{-11} \tag{4.26}
\]

Note that \((2n)! \leq 2^{2n}(n!)^2\) and \( n!n! \leq (n + m)! \), so, recalling the definition of \( \mathfrak{M}_{m,n} \), we have

\[
\|P(x', \cdot)\|_2 \leq (\mathfrak{M}_{m,n} 2e)^{n+m}(n+m)! \tag{4.27}
\]
To control the second term in (4.24), we first use the Minkowski inequality to the sum over the partitions, as in the proof of Theorem 4.1, yielding

\[
\|Q(x', \cdot)\|_{\ell_2} \\
\leq \sum_{\pi \in \mathcal{P}(J_m + J_n)} \prod_{S \in \pi} \left( \sum_{x \in Z^d} |\kappa[\psi(x), \phi(x)]|^2 \right)^{1/2} \mathbb{1}(A' \neq \emptyset, A \neq \emptyset) .
\]

In the final expression all the cumulants \(\kappa[\psi(x), \phi(x)]\) are such that \(|A'| < m\) and \(|A| < n\). Therefore, the induction hypothesis can be applied to estimate their \(\ell_2\)-norms:

\[
\left[ \sum_{x \in Z^d} |\kappa[\psi(x), \phi(x)]|^2 \right]^{1/2} \leq (\mathcal{M}_{m', n'}\gamma^{m'})|S|! |m' = |S| |m' = |S| |J_m'|, n' = |S| J_n| .
\]

Note that \(\mathcal{M}_{m,n}\) is non-decreasing in \(m, m' \leq m - 1\); hence,

\[
\|Q(x', \cdot)\|_{\ell_2} \leq \sum_{\pi \in \mathcal{P}(J_m + J_n)} \prod_{S \in \pi} (\mathcal{M}_{m', n'}\gamma^{m'})|S|! |m' = |S| |m' = |S| |J_m'|, n' = |S| J_n| .
\]

Collecting the two estimates together we have proven

\[
\|P(x', \cdot)\|_{\ell_2} + \|Q(x', \cdot)\|_{\ell_2} \leq \mathcal{M}_{m,n}(n + m)![(2e)^{n+m} + (e/\gamma)^{n+m}\gamma^{m+n}] . \tag{4.28}
\]

In order to close the induction, we need to choose \(\gamma\) such that

\[
(2e)^{n+m} + (e/\gamma)^{n+m}\gamma^{m+n} \leq \gamma^{n+m} .
\]

Since \(m \geq 2\), it suffices to set, for instance, \(\gamma = 2e\).

The proof of the second item, with \(p = 2\), is essentially the same: one only needs to replace the flat \(\ell_2\)-norm by the above weighted \(\ell_2\)-norm containing the factor \(|\Phi_n(y, x)|\) (which can also be understood as integrals over the corresponding weighted counting measure over \(Z\)), and to apply item 2 in Theorem 4.1 instead of item 1 there. To reach the same combinatorial estimates, we can reduce the resulting second powers of \(|S|!\) to sums over first powers via the bound \(\prod_{S \in \pi} |S|! \leq (n + m)!\) which is valid for any partition \(\pi \in \mathcal{P}(J_m + J_n)\).

\[\square\]

5 Discussion with an application to DNLS

Suppose that for each \(t \geq 0\) there is given \(\psi_t(x, \sigma)\) which is a random field on \(\mathbb{Z}^d \times I\) for some finite index set \(I\). Suppose also that all \(\psi_t\) are identically distributed, according to an \(\ell_1\)-clustering measure; such fields arise, for instance, from stochastic processes by choosing the initial data from a stationary measure which is \(\ell_1\)-clustering. For such a system, using \(X = \psi_0(0, \sigma_0)\) in Theorem 4.1 implies that any time-correlation function of the form

\[
F_{t,n,\sigma_0}(x, \sigma) := \kappa[\psi_0(0, \sigma_0), \psi_t(x_1, \sigma_1), \ldots, \psi_t(x_{n-1}, \sigma_{n-1})] \tag{5.1}
\]

belongs to \(\ell_2((\mathbb{Z}^d \times I)^{n-1})\) and its norm is uniformly bounded in \(t\) by a constant which depends only on the initial measure.

As an explicit example, let us come back to the discrete NLS evolution and its equilibrium time-correlations, as discussed in the Introduction. At the time of writing, we are not aware of a rigorous definition of the infinite volume dynamics for an equilibrium measure of the DNLS.
However, DNLS evolution is well-defined on any finite periodic lattice, and there is a range of hopping amplitudes and equilibrium parameters for which the corresponding thermal Gibbs states are $\ell_1$-clustering, uniformly in the lattice size, as proven in [16]. Therefore, it seems likely that there are harmonic couplings for which also the DNLS evolution equations on $\mathbb{Z}^d$ with initial data distributed according to a stationary measure can be solved almost surely. In addition, it should be possible to define the stationary measure so that it is $\ell_1$-clustering and translation invariant.

For any such $\ell_1$-clustering stationary measure, we could then study the evolution of the functions $F_{t,n,s_0}$ using the above results. Referring to [1] for details, for instance $f_t(x) := \kappa[\psi_0(0, -1), \psi_t(x_1, 1)] = \mathbb{E}[\psi_0(0, -1): \psi_t(x_1, 1)]$ would then satisfy an evolution equation

$$i\partial_t f_t(x) = \sum_{y \in \mathbb{Z}^d} \alpha(x - y)f_t(y) + \lambda g_t(x),$$

(5.2)

where $g_t(x) := \mathbb{E}[\psi_0(0, -1): \psi_t(x, -1)\psi_t(x, 1)\psi_t(x, 1)]$. Applying Proposition A.1, $g_t$ can be represented in terms of the constant $\mathbb{E}[\psi_t(x, 1)]$ and the functions $F_{t,n,-1}$, with $n = 2, 3, 4$. Using the above estimates, it then follows that there is a constant $C$ such that $\|g_t\|_{\ell_2} \leq C$ for all $t$. Taking a Fourier-transform of (5.2) and solving it in Duhamel form implies that $\hat{f}_t = \mathcal{F}f_t$ satisfies

$$\hat{f}_t(k) = e^{-i\tilde{\alpha}(k)}\hat{f}_0(k) - i\lambda \int_0^t ds e^{-i(t-s)\tilde{\alpha}(k)}\hat{g}_s(k).$$

(5.3)

For stable harmonic interactions one needs to have $\tilde{\alpha}(k) \geq 0$. Therefore, for such systems we may conclude, without any complicated analysis of oscillatory integrals or graph expansion of cumulants, that the harmonic evolution dominates the behavior of $f_t$ up to times of order $\lambda^{-1}$.

More precisely, we find that the $\ell_2$-norm of the error is bounded by

$$\left\|f_t - \mathcal{F}^{-1}(e^{-i\tilde{\alpha}}\hat{f}_0)\right\|_{\ell_2} = \left\|\hat{f}_t - e^{-i\tilde{\alpha}}\hat{f}_0\right\|_{L^2(\mathbb{T}^d)} \leq Ct\lambda,$$

(5.4)

for all $t \geq 0$.

The above example perhaps does not appear very significant: after all, it is simply stating that the nonlinearity acts as a perturbation in $\ell_2$-norm with its “natural” strength, having an effect of order $\lambda t$ to the time-evolution. Let us however stress that without the present a priori bounds there seems to be no other alternative to prove this than to resort to the heavy machinery of time-dependent perturbation theory for Feynman graph classification of oscillatory integrals and careful applications of moment hierarchies, see [5] for a detailed example for DNLS.

Another important property hidden in the bound (5.4) is the fact that $\hat{f}_t(k)$ is a function, and not a distribution. This would not be true in general if instead of cumulants we would have used moments to define $f_t$: if $\mathbb{E}[\psi_0(0)] \neq 0$ and the initial state is translation invariant, already at $t = 0$ the Fourier transform of $\mathbb{E}[\psi_0(0, -1)\psi_t(x_1, 1)]$ is a distribution proportional to the Dirac delta $\delta(k)$. The fact that the cumulants produce functions, which are uniformly bounded in $\ell_2$, allows not only taking Fourier transforms but also simplifying the study of nonlinear terms in the hierarchies as products of distributions are notoriously difficult to control rigorously.

On the mathematical side, it would be of interest to study more carefully the above combinatorial bounds. We do not claim that the above constants, or their dependence on the orders $n$ and $m$, should be optimal, and there could be room for significant improvement there, possibly of importance in problems requiring the full infinite order cumulant hierarchy. Also, it is not clear what are the optimal powers and weights for the summability of the correlations as the clustering power $p$ of the field $\phi$ is varied. These questions could prove to be hard to resolve in the greatest generality, but we remain optimistic that already the present bounds suffice to control the time-evolution of cumulants in some of the above mentioned open transport problems.

### A Cumulants and Wick polynomials

We have collected in this Appendix the main results relating to Wick polynomials which were used in the text. The goal is not to give an exhaustive exposition of the topic, but rather to list the
minimal amount of definitions and properties needed for the proofs here. In the Gaussian case, one can also identify the Wick polynomials as arising from an orthogonalization procedure, with applications in Wiener chaos expansions and Malliavin calculus [13]. The present general, non-Gaussian case, is not directly connected to orthogonal polynomials, and leads to more complicated combinatorial expansions: we refer to [17] for a recent review of these from the point of view of probability theory and for more results how graphical representations may be used to facilitate the analysis involving Wick polynomials.

Let us consider a collection \( y_j, j \in J \) where \( J \) is some fixed nonempty index set, of real or complex random variables on some probability space \((\Omega, \mathcal{F}, \mu)\). Then for any sequence of indices, \( I = (i_1, i_2, \ldots, i_n) \in J^n \), we use the following shorthand notations to label monomials of the above random variables:

\[
y_I = y_{i_1} y_{i_2} \cdots y_{i_n} = \prod_{k=1}^n y_{i_k}, \quad y^\emptyset := 1 \quad \text{if} \quad I = \emptyset. \tag{A.1}
\]

As already explained in an earlier footnote, we consider sequences of indices and not sets of indices in order to avoid more cumbersome notations involving relabelling of elements whenever an element is repeated in a sequence. We will however continue to use set-like notations for subsequences and partitions of sequences. To be precise, these notations are valid only after one has added a unique label for each element of the sequence. For this, we introduce a collection \( \mathcal{I} \) which consists of those finite subsets \( A \subset \mathbb{N} \times J \) with the property that if \((n, j), (n', j') \in A \) and \((n, j) \neq (n', j') \) then \( n \neq n' \). The empty sequence is identified with \( \emptyset \in \mathcal{I} \). For nonempty sets, the natural number in the first component serves as a distinct label for each member in \( A \) and their order determines the order of the elements in the sequence.

For any \( I \in \mathcal{I} \) we denote the corresponding moment by \( E[y_I^+] \), and the related cumulant by

\[
\kappa[y_I] = \kappa_\mu[y_I] = \kappa[y_{i_1}, y_{i_2}, \cdots, y_{i_n}] . \tag{A.2}
\]

The corresponding Wick polynomial is denoted by

\[
y_I^+ = y_I^{+\mu} = :y_{i_1} y_{i_2} \cdots y_{i_n}:. \tag{A.3}
\]

Both \( \kappa[y_I^+] \) and \( y_I^+ \) can be defined recursively if \( I \in \mathcal{I} \) is such that \( E[|y_I^+|] < \infty \) for all \( E \subset I \) (see [1]). Explicitly, it suffices to require that

\[
y_I^+ = y_I - \sum_{E : x \in E \subseteq I} E[y_I^+ \setminus E] y_E^+, \tag{A.4}
\]

and, choosing some \( x \in I \),

\[
\kappa[y_I] = E[y_I^+] - \sum_{E : x \in E \subseteq I} E[y_I^+ \setminus E] \kappa[y_E] . \tag{A.5}
\]

Let us also recall that both cumulants and Wick polynomials are multilinear and permutation invariant.

If the random variables \( y_j, j = 1, 2, \ldots, n \), have joint exponential moments, then moments, cumulants and Wick polynomials can also be easily generated by differentiation of their respective generating functions which are

\[
G_m(\lambda) := E[e^{\lambda x}], \quad g_c(\lambda) := \ln G_m(\lambda) \quad \text{and} \quad G_w(\lambda; y) := \frac{e^{\lambda y}}{E[e^{\lambda x}]} = e^{\lambda y - g_c(\lambda)} . \tag{A.6}
\]

By evaluation of the \( I \)-th derivative at zero, we have

\[
E[y_I^+] = \partial^I \lambda G_m(0), \quad \kappa[y_I] = \partial^I \lambda g_c(0) \quad \text{and} \quad :y_I^+: = \partial^I \lambda G_w(0; y) , \tag{A.7}
\]

where “\( \partial^I \lambda \)” is a shorthand notation for \( \partial_{\lambda_{i_1}} \partial_{\lambda_{i_2}} \cdots \partial_{\lambda_{i_n}} \).

It is remarkable that expectations of products of Wick polynomials can be expanded in terms of cumulants, merely cancelling some terms from the standard moments-to-cumulants expansion. The following result, proven as Proposition 3.8 in [1], details the result using the above notations:

13
Proposition A.1 Assume that the measure \( \mu \) has all moments of order \( N \), i.e., suppose that 
\[ E[|y_I|] < \infty \] 
for all \( I \in \mathcal{J} \) with \(|I| \leq N \). Suppose \( L \geq 1 \) is given and consider a collection of \( L+1 \) index sequences \( J'_\ell, J_\ell \in \mathcal{J}, \ell = 1, \ldots, L \), such that \(|J'_\ell| + \sum_{\ell} |J_\ell| \leq N \). Then for \( I := \sum_{\ell=1}^L J_\ell + J' \) (with the implicit identification of \( J_\ell \) and \( J'_\ell \) with the set of its labels in \( I \)) we have 
\[ E \left[ \prod_{\ell=1}^L y_{J'_\ell} : y_{J_\ell} \right] = \sum_{\pi \in \mathcal{P}(I)} \prod_{A \in \pi} (\kappa[y_A] \mathbb{1}(A \not\subset J_\ell \forall \ell)) . \tag{A.8} \]

In words, the constraint determined by the characteristic functions on the right hand side of (A.8) amounts to removing from the standard cumulant expansion all terms which have any clusters internal to one of the sets \( J_\ell \). For instance, thanks to Proposition A.1, if we consider the expectation of the product of two second order Wick polynomials, we get 
\[ E[y_{J_1} y_{J_2} : y_{J_3} y_{J_4} :] = \kappa(y_1, y_3) \kappa(y_2, y_4) + \kappa(y_1, y_4) \kappa(y_2, y_3) + \kappa(y_1, y_2, y_3, y_4) . \]

Proposition A.1 turns out to be a powerful technical tool, used several times in the proofs of Theorems 4.1 and 4.3.

References


Harmonic chain with velocity flips: thermalization and kinetic theory

Jani Lukkarinen, Matteo Marcozzi, Alessia Nota
Harmonic chain with velocity flips: thermalization and kinetic theory

Jani Lukkarinen*, Matteo Marcozzi†, Alessia Nota‡

* University of Helsinki, Department of Mathematics and Statistics
P.O. Box 68, FI-00014 Helsinki, Finland

† University of Bonn, Institute for Applied Mathematics
Endenicher Allee 60, D-53115 Bonn, Germany

Abstract

We consider the detailed structure of correlations in harmonic chains with pinning and a bulk velocity flip noise during the heat relaxation phase which occurs on diffusive time scales, for \( t = O(L^2) \) where \( L \) is the chain length. It has been shown earlier that for non-degenerate harmonic interactions these systems thermalize, and the dominant part of the correlations is given by local thermal equilibrium determined by a temperature profile which satisfies a linear heat equation. Here we are concerned with two new aspects about the thermalization process: the first order corrections in \( 1/L \) to the local equilibrium correlations and the applicability of kinetic theory to study the relaxation process. Employing previously derived explicit uniform estimates for the temperature profile, we first derive an explicit form for the first order corrections to the particle position-momentum correlations. By suitably revising the definition of the Wigner transform and the kinetic scaling limit we derive a phonon Boltzmann equation whose predictions agree with the explicit computation. Comparing the two results, the corrections can be understood as arising from two different sources: a current-related term and a correction to the position-position correlations related to spatial changes in the phonon eigenbasis.

Contents

1 Introduction 2

2 Evolution of the first two moments in the velocity flip model 3

3 Uniform estimates in the large flip rate regime 6

3.1 The main result 6

3.2 Proof of Theorem 3.2 9

3.2.1 Review of the properties of \( T_t(x) \) 9

3.2.2 Derivation of equation (3.13) 11

4 Kinetic theory of the velocity flip model 15

4.1 Time evolution of the mean Wigner function of normal modes 15

4.2 Stationary solutions 21

4.3 Boltzmann equation for the energy density 22

4.4 Kinetic theory prediction for diffusion of energy 23

*E-mail: jani.lukkarinen@helsinki.fi
†E-mail: matteo.marcozzi@helsinki.fi
‡E-mail: nota@iam.uni-bonn.de
1 Introduction

We consider a harmonic chain with velocity flips, or the velocity flip model for short. The model dynamics consists of a classical Hamiltonian evolution of the particle positions and velocities, as determined by a quadratic Hamiltonian, intercepted with random flips of the particle velocities. This model was first considered in [1], and it is one of the simplest known particle chain models which has a finite thermal conductivity and satisfies the time-dependent Fourier’s law [2, 3].

The model is blessed with many simplifying features which make possible the usually intractable rigorous analysis of heat transport properties. For instance, it is proven in [1] that under quite general conditions every translation invariant stationary state of the infinite chain with a finite entropy density is given by a mixture of canonical Gibbs states. This indicates that temperature is the sole thermodynamic parameter in the velocity flip model with pinning. The numerical simulations of the model with boundary heat baths in [4] support these findings and provide more information about the resulting nonequilibrium states. The structure of steady state correlations and energy fluctuations are discussed in [5] with supporting numerical evidence presented in [6].

The validity of the proposed hydrodynamic limit equations (Fourier’s law) is proven rigorously in [3] (see the Remark after Theorem 1.2. for the changes needed in case the model has pinning).

The strategy for proving the hydrodynamic limit in [3] is based on the relative entropy method introduced by Yau and Varadhan; we refer to [7] for a review of the method. There one studies the relaxation of initial states which are already close to a local thermal equilibrium state and as a result one obtains estimates on how local observables, averaged over regions of size $O(L)$, evolve at diffusive time scales $O(L^2)$. The method was applied earlier to a similar model with somewhat different stochastic perturbation in [8, 9]. This model shares many features with the velocity flip model with pinning considered here. For instance, also there thermal conductivity is constant and hence temperature evolves according to a linear diffusion equation.

A different approach was chosen in [2] to study the evolution of the kinetic temperature profile, $T_t(x) = \langle p_x^2(t) \rangle$, where $p_x(t)$ is the momentum at time $t$ of the particle at the lattice site $x$. It was first observed that the temperature profile satisfies a closed renewal-type equation, and the analysis of the properties of the equation lead to a strong, pointwise, control of the errors between the temperature profile and its hydrodynamic description by Fourier’s law.

The goal of this paper is to clarify the physical meaning of the results in [2], and to explore its implications on the structure of general local correlations after local equilibrium has been reached. We consider the evolution of the full spatial covariance matrix of positions and momenta, and by defining a suitable Wigner function from the covariance matrix, we compute the first order corrections to the local thermal equilibrium. The first order correction, at diffusive time scales $t = O(L^2)$, turns out to be proportional to the temperature gradient, and hence is $O(L^{-1})$. In particular, we expect these results to be valid also for the leading covariance in a nonequilibrium steady state of the velocity flip model induced, for instance, by boundary thermostats. More precisely, we expect that the local correlations sufficiently far away from the boundary are then given by the appropriate equilibrium correlations with the leading correction given by the first order term derived here.

In the first part of the paper, Section 3, we deal with a periodic chain of length $L$ under the same assumptions as used in [2]. In particular, the stochastic flip rate $\gamma$ is assumed to be sufficiently large compared to the Hamiltonian dispersion relation. Then the estimates derived in
for the kinetic temperature profile can be applied to study the evolution of the full covariance matrix. This leads to an explicit, fairly simple, form for the first order correction, with rigorous upper bounds for the magnitude of the higher order corrections.

The simple form of the first order corrections begs for an explanation. According to the Fourier’s law the energy current is proportional to the temperature gradient, and thus necessarily $O(L^{-1})$, and this is indeed the dominant correction found in the position-momentum correlations. However, there are also other corrections of the same order, namely in the position-position correlations, while momentum-momentum correlations feature no such corrections. In the second part, Section 4, we derive the same dominant correction term from the kinetic theory of phonons. This provides a qualitative description of the correction and explains also the origin of the position-position correlations.

The kinetic theory of a similar system—merely with somewhat different, momentum conserving noise—has been derived in [10]. It is shown there that a kinetic scaling limit of a lattice Wigner function of phonon modes satisfies a linear phonon Boltzmann equation, and thus its evolution can be studied via the solutions of the Boltzmann equation. Here we diverge from the standard scheme on two accounts. Firstly, we employ a somewhat different definition of the Wigner function in which explicit real-valuedness is sacrificed for simpler analytical estimates. Several alternative definitions and basic properties of more standard Wigner functions for continuum and lattice waves are available in [11, 12] and in Appendix B of [13]. Secondly, we do not take any scaling limits explicitly but rather introduce spatial averaging into the definition of Wigner function. This allows separating phonon collisions from the large scale transport without resorting to scaling limits.

The resulting kinetic theory of the velocity flip model is determined by a phonon Boltzmann equation with a very simple collision operator and solving the equation is most standard. However, proper application of the result for spatially inhomogeneous states requires also analysis of polarization effects, in particular, of the evolution of field self-correlations. Our treatment of the kinetic theory is not fully rigorous but it is vindicated in the answer to the question about first order corrections to local equilibrium at diffusive scales: the corrections are found to be entirely consistent with the previous rigorously derived result. In particular, the somewhat unexpected position-position correlations are found to arise from changes in the phonon eigenbasis resulting from the inhomogeneities in the energy profile.

We compare the two results in more detail in Section 5. The three Appendices contain more details about some of the main computations used in the text.

Let us emphasize that we only consider models with pinning here. If the onsite potential is absent, a second locally conserved field related to the tension in the chain appears, in addition to the present temperature field. For results about the hydrodynamics of the velocity flip model without pinning, we refer to [3, 5, 6]. A more general overview about thermal transport in similar particle chains can be found in [14]. In particular, in Chapter 5 [15] the results of [10] are reviewed along with other rigorous works dealing with similar stochastic models.

## 2 Evolution of the first two moments in the velocity flip model

In this section we briefly recall the velocity flip model and the notations used in [2]. We consider a one-dimensional periodic chain (circle) of $L$ particles and we parametrize the sites on the chain by

\[
\Lambda_L := \left\{ -\frac{L-1}{2}, \ldots, \frac{L-1}{2} \right\}, \quad \text{if } L \text{ is odd,} \tag{2.1}
\]

\[
\Lambda_L := \left\{ -\frac{L}{2} + 1, \ldots, \frac{L}{2} \right\}, \quad \text{if } L \text{ is even.} \tag{2.2}
\]

Then always $|\Lambda_L| = L$ and $\Lambda_L \subset \Lambda_{L'}$ if $L \leq L'$. In addition, for odd $L$, we have $\Lambda_L = \{ n \in \mathbb{Z} \mid |n| < \frac{L}{2} \}$. We use periodic arithmetic on $\Lambda_L$, setting $x' + x := (x' + x) \mod \Lambda_L$ for
On occasion, we would like to stress the use of periodicity in the arithmetic, and we use then the somewhat heavier notation \([x' + x]_L\) for \(x' + x\).

The particles are identical with unit mass and interact via linear forces with a finite range given by the potential \(\Phi : \mathbb{Z} \to \mathbb{R}\) which is assumed to be symmetric, \(\Phi(-x) = \Phi(x)\). The range of \(\Phi\) is described by \(r_\Phi\) which we assume to be odd and chosen so that \(\Phi(x) = 0\) for all \(|x| \geq r_\Phi/2\). Moreover, the forces are assumed to be stable and pinning, i.e., the Fourier transform \(\hat{\gamma}\) is required to be strictly positive. The related dispersion relation \(\omega : \mathbb{T} \to \mathbb{R}\) is defined as \(\omega := \sqrt{\Phi}\), and it is then a smooth function on the circle \(\mathbb{T} := \mathbb{R}/\mathbb{Z}\) with \(\omega_\Phi := \min_{k \in \mathbb{T}} \omega(k) > 0\). The corresponding periodic interaction matrices \(\Phi_L \in \mathbb{R}^{\Lambda_L \times \Lambda_L}\) on \(\Lambda_L\) are defined by

\[
(\Phi_L)_{x',x} := \Phi(|x' - x|_L), \quad \text{for all } x', x \in \Lambda_L. \quad (2.3)
\]

This clearly results in a real symmetric matrix.

The discrete Fourier transform \(\mathcal{F}_L\) maps functions \(f : \Lambda_L \to \mathbb{C}\) to \(\hat{f} : \Lambda^*_L \to \mathbb{C}\), where \(\Lambda^*_L := \Lambda_L/L \subset (-\frac{1}{2}, \frac{1}{2}]\) is the dual lattice and for \(k \in \Lambda^*_L\) we set

\[
\hat{f}(k) = \sum_{x \in \Lambda_L} f(x)e^{-i2\pi k \cdot x}. \quad (2.4)
\]

The inverse transform \(\mathcal{F}_L^{-1} : g \mapsto \hat{g}\) is given by

\[
\hat{g}(x) = \int_{\Lambda^*_L} dk \, g(k)e^{i2\pi k \cdot x}, \quad (2.5)
\]

where we use the convenient shorthand notation

\[
\int_{\Lambda^*_L} dk \cdots = \frac{1}{|\Lambda_L|} \sum_{k \in \Lambda^*_L} \cdots. \quad (2.6)
\]

With the above conventions, for any \(L \geq r_\Phi\) we have

\[
(\mathcal{F}_L \Phi_L f)(k) = \omega(k)^2 \hat{f}(k), \quad \text{for all } k \in \Lambda^*_L, \quad (2.7)
\]

i.e., the functional form of the interaction in the Fourier space does not depend on \(L\).

We also use \(\delta_L\) to denote a "discrete \(\delta\)-function" on \(\Lambda^*_L\), defined by

\[
\delta_L(k) = |\Lambda_L| \mathbb{1}(k = 0), \quad \text{for } k \in \Lambda^*_L. \quad (2.8)
\]

Here, and in the following, \(\mathbb{1}\) denotes the generic characteristic function: \(\mathbb{1}(P) = 1\) if the condition \(P\) is true, and otherwise \(\mathbb{1}(P) = 0\).

The linear forces on the circle are then generated by the Hamiltonian

\[
H_L(X) := \frac{1}{2} \sum_{x \in \Lambda_L} (X^2_x)^2 + \sum_{x', x \in \Lambda_L} \frac{1}{2} X^1_{x'}X^1_x \Phi(|x' - x|_L) = \frac{1}{2} X^T \mathcal{G}_L X, \quad (2.9)
\]

\[
\mathcal{G}_L := \begin{pmatrix} \Phi_L & 0 \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{(2\Lambda_L) \times (2\Lambda_L)}, \quad (2.10)
\]

on the phase space \(X \in \Omega := \mathbb{R}^{\Lambda_L} \times \mathbb{R}^{\Lambda_L}\). The canonical pair of variables for the site \(x\) are the position \(q_x := X^1_x\), and the momentum \(p_x := X^2_x\). By adding to the Hamiltonian evolution a velocity-flip noise, the system can be identified with a Markov process \(X(t)\) and the process generates a Feller semigroup on the space of observables vanishing at infinity (see [16, 3]). For \(t > 0\) and any \(F\) in the domain of the generator \(\mathcal{L}\) of the Feller process the expectation values of \(F(X(t))\) satisfy an evolution equation

\[
\partial_t \langle F(X(t)) \rangle = \langle (\mathcal{L}F)(X(t)) \rangle, \quad (2.11)
\]
where \( \mathcal{L} := A + S \), with
\[
A := \sum_{x \in \Lambda_L} \left( X^2_x \partial_{X^1_x} - (\Phi_L X^1_x) \partial_{X^2_x} \right), \tag{2.12}
\]
\[
(SF)(X) := \gamma \sum_{x_0 \in \Lambda_L} \left( F(Sx_0 X) - F(X) \right), \tag{2.13}
\]
\[
(S_{x_0} X)^i_x := \begin{cases} -X^i_x, & \text{if } i = 2 \text{ and } x = x_0, \\ X^i_x, & \text{otherwise}. \end{cases} \tag{2.14}
\]

Let \( \tilde{q}_t = E[q_t] \) and \( \tilde{p}_t = E[p_t] \). Then
\[
\partial_t \tilde{q}_t(x) = E[\mathcal{L}\tilde{q}_t(x)] = \tilde{p}_t(x), \tag{2.15}
\]
\[
\partial_t \tilde{p}_t(x) = E[\mathcal{L}\tilde{p}_t(x)] = - (\Phi_L \tilde{q}_t)(x) - \gamma \tilde{p}_t(x). \tag{2.16}
\]

We set
\[
M_t(x, y) = \begin{pmatrix} 0 & \Phi_L(x - y) \\ -\gamma (x = y) & \gamma \delta(x = y) \end{pmatrix} \tag{2.17}
\]
so that we can rewrite (2.15) in a compact form as
\[
\partial_t \begin{pmatrix} \tilde{q}_t \\ \tilde{p}_t \end{pmatrix} = -M^\top_t \begin{pmatrix} \tilde{q}_t \\ \tilde{p}_t \end{pmatrix}. \tag{2.18}
\]

The full spacial position-momentum covariance matrix is defined by
\[
C^{ij}_t(x, y) := E[X^i_x(t) X^j_y(t)]. \tag{2.19}
\]

Strictly speaking, \( C_t \) denotes the matrix of second moments, and to get the covariance matrix we should subtract the appropriate products of the mean values \( \tilde{q}_t \) and \( \tilde{p}_t \). However, as we will prove later using the above explicit solutions, the mean values decay to zero exponentially fast on time scale \( O(\gamma^{-1}) \), and hence the difference between \( C_t \) and the covariance matrix is exponentially small in the length \( L \) for diffusive time scales \( t = O(L^2) \). Thus the distinction is not relevant in the case considered here.

**Remark 2.1** We observe that in this Section we have used mathematically nonstandard, but common in physics, notations for orders of magnitude. These notations implicitly assume that it has to be known which quantities are large and which small. Since one of the aims of the present paper is to explore the importance of scaling limits for the validity of kinetic theory, we now state more explicitly what is meant by the above notations. Time and space scales are always assumed to be “large enough”, so “\( \gamma = O(L^2) \)” actually means that there is some constant \( C > 0 \) such that \( t \geq CL^2 \). In particular, any such \( C \) must be independent from the lattice size \( L \), velocity flips and initial data. However, any succeeding bounds are allowed to depend on the choice of \( C \) and they might blow up as \( C \to 0 \) or \( C \to \infty \). If we considered the limit \( L \to \infty \), this could be written using the standard notation as \( t^{-1} = O(L^{-2}) \). We will in fact later use the notation \( O(L^{-2}) \) to denote the order of magnitude of many error terms: its precise meaning is to say that there is a constant \( C \) as above such that the term is bounded by \( CL^{-2} \) for all large enough \( L \).

According to (2.11), the entries of \( C_t \) satisfy
\[
\partial_t C^{11}_t(x, y) = C^{21}_t(x, y) + C^{12}_t(x, y),
\]
\[
\partial_t C^{12}_t(x, y) = C^{22}_t(x, y) - (C^{11}_t \Phi_L)(x, y) - \gamma C^{12}_t(x, y),
\]
\[
\partial_t C^{21}_t(x, y) = C^{12}_t(x, y) - (\Phi_L C^{11}_t)(x, y) - \gamma C^{21}_t(x, y),
\]
\[
\partial_t C^{22}_t(x, y) = - (\Phi_L C^{12}_t)(x, y) - (C^{21}_t \Phi_L)(x, y) - 2\gamma C^{22}_t(x, y) + 2\gamma \delta(x = y)T_t(x). \nonumber
\]


Here \( T_t(x) = \mathbb{E}[p_t(x)^2] \) denotes the kinetic temperature at site \( x \). Therefore, we can write the evolution equation for \( C_t \) in a more compact way:

\[
\partial_t C_t = -M_t^T C_t - C_t M_t + 2\gamma G_t, \tag{2.20}
\]

where

\[
G_t(x, y) = \left( \begin{array}{cc} 0 & 0 \\ 0 & \mathbb{1}(x = y)T_t(x) \end{array} \right).
\]

The matrix evolution equation (2.20) can be rewritten in Duhamel’s form, so that only the last \( G_t \)-term remains as a perturbation. Namely, integrating the identity

\[
\partial_s \left( e^{-(t-s)M_t^T} C_s e^{-(t-s)M_t} \right) = e^{-(t-s)M_t^T} \left( M_s^T C_s + \partial_s C_s + C_s M_s \right) e^{-(t-s)M_t}, \tag{2.21}
\]

over \( s \) from 0 to \( t \), we find that any solution to (2.20) also satisfies

\[
C_t = e^{-tM_t^T} C_0 e^{-tM_t} + 2\gamma \int_0^t ds e^{-(t-s)M_t^T} G_s e^{-(t-s)M_t}. \tag{2.22}
\]

In fact, the right hand side in (2.22) is a known function which thus determines the evolution of the covariance matrix \( C_t \) on the left hand side: the first term on the right depends only on the initial data covariance \( C_0 \), and the second term on the matrix \( G_s \). On the other hand, the matrix \( G_s \) is a function of the temperature profile \( T_s(x) \) only, and its behaviour has already been solved in [2]. As we will show next, the strong control derived for the temperature profile in [2] suffices to determine the local covariances up to order \( O(L^{-2}) \) at diffusive time scales.

3 Uniform estimates in the large flip rate regime

3.1 The main result

We first consider a regime in which the flip rate is sufficiently large. More precisely in this section we assume that \( \gamma > 2 \max_{k \in \mathbb{T}} \omega(k) \). Under this condition several analytical results become available from [2]. We recall that we want to derive a suitable approximation on the diffusive scale of the full spatial position-momentum covariance matrix. The structure of the correlations is conveniently studied by introducing the following variant of Wigner functions,

\[
U_t(x, k) := \sum_{y \in \Lambda_L} e^{-2\pi i k y} C_t(x, x + y). \tag{3.1}
\]

This corresponds to taking a Fourier transform of the covariance matrix with respect to the spatial displacement at the point \( x \). Whenever the correlations decay at least square summably, this definition results in a function of \( x, k \) instead of a distribution as can occur in other alternatives. We discuss lattice Wigner functions in more detail later together with the kinetic theory description in Section 4.1.

To get a more explicit expression for \( U_t \), we use (2.22) and write

\[
C_t(x, x + y) = \sum_{z_1, z_2 \in \Lambda_L} \left( e^{-tM_t^T} \right)_{x, z_1} C_0(z_1, z_2) \left( e^{-tM_t} \right)_{z_2, x + y} + 2\gamma \int_0^t ds \sum_{z_1, z_2 \in \Lambda_L} \left( e^{-(t-s)M_t^T} \right)_{x, z_1} G_s(z_1, z_2) \left( e^{-(t-s)M_t} \right)_{z_2, x + y}. \tag{3.2}
\]

Thanks to the translation invariance of the matrix \( M_s \), we can define a matrix \( A_t \) by the condition \( A_t(x - y) = (e^{-tM_t})_{x, y} \). As shown in [2, Appendix A], its Fourier transform is

\[
\hat{A}_t(k) = \sum_{\sigma = \pm 1} \frac{e^{-t\mu_\sigma(k)}}{\mu_\sigma(k) - \mu_{-\sigma}(k)} \begin{pmatrix} -\mu_{-\sigma}(k) & \omega(k)^2 \\ 1 & \mu_{\sigma}(k) \end{pmatrix} = \frac{e^{-\gamma t/2}}{\Omega} \begin{pmatrix} \frac{1}{2} \sinh \Omega t + \Omega \cosh \Omega t & -\omega(k)^2 \sinh \Omega t \\ \sinh \Omega t & -\frac{1}{2} \sinh \Omega t + \Omega \cosh \Omega t \end{pmatrix}. \tag{3.3}
\]
with \( \Omega = (\gamma/2) \sqrt{1 - (2\omega(k)/\gamma)^2} < \gamma/2 \) and \( \mu_\sigma(k) = \gamma/2 + \sigma \Omega(k) \). (To facilitate comparison, let us point out that the function \( \Phi \) was denoted by \( \varphi \), and only the second column of \( A \) was used in [2].) Since \( \omega(k) = \omega(-k) \), it follows that \( A_t(k) = A_t(-k) \) and thus also \( A_t(x) = A_t(-x) \) because \( A_t \) is real-valued. Let \( P^{(2)} \) denote the projection matrix to the second component, i.e., it is the diagonal 2 \( \times \) 2-matrix defined as \( P^{(2)} = \text{diag}(0, 1) \). Then, thanks to (3.1), (3.2) and (3.3) we can rewrite \( \bar{U}_t(x, k) \) as

\[
\bar{U}_t(x, k) = \sum_{y, z \in \Lambda_L} e^{-i2\pi k \cdot y} A_t^\top(z - x)U_0(z, k)A_t(x + y - z) \\
+ 2\gamma \int_0^t ds \sum_{y, z \in \Lambda_L} e^{-i2\pi k \cdot y} T_s(z) A_{t-s}^\top(z - x)P^{(2)} A_{t-s}(x + y - z).
\]

We rename the first term depending on the initial data as

\[
Z_t(x, k) := \sum_{y, z \in \Lambda_L} e^{-i2\pi k \cdot y} A_t^\top(z - x)U_0(z, k)A_t(x + y - z)
\]

and, by shifting the summation and integration variables, find

\[
\bar{U}_t(x, k) = Z_t(x, k) + 2\gamma \int_0^t ds \sum_{y, z \in \Lambda_L} e^{-i2\pi k \cdot y} T_{t-s}(x + z) A_s^\top(z)P^{(2)} A_s(y - z).
\]

We now define matrices

\[
\mathcal{A}_s(k) := \hat{A}_s^\top(k)P^{(2)} \hat{A}_s(k), \quad \mathcal{F}_s(k) := \frac{1}{2\pi}(\partial_k \hat{A}_s(k))P^{(2)} \hat{A}_s(k)
\]

and

\[
U_0(k) := 2\gamma \int_0^\infty dt \mathcal{F}_t(k) = \begin{pmatrix} \omega(k)^{-2} & -q(k) \\ 0 & 1 \end{pmatrix}, \quad U_1(k) := 2\gamma \int_0^\infty dt \hat{\mathcal{F}}_t(k).
\]

Then \( \partial_k U_0(k) = 2\pi(U_1(k) + U_1(k)\top) \), and thus we have

\[
U_1(k) = \frac{1}{2\pi} \begin{pmatrix} -\omega(k)^{-3} \partial_k \omega(k) & q(k) \\ -q(k) & 0 \end{pmatrix}
\]

where

\[
q(k) = 2\gamma \int_0^\infty dt \hat{A}_t(k)^{22} \partial_k \hat{A}_t(k)^{21}.
\]

In Appendix A we show by an explicit computation that

\[
q(k) = \frac{\partial_k \omega(k)}{\omega(k)}.
\]

The dispersion relation determines the velocity of the lattice waves with wave number \( k \), and with the present choices of normalization, the velocity is given by \( v(k) := \partial_k \omega(k)/2\pi \). Hence,

\[
U_1(k) = \frac{v(k)}{\omega(k)} \begin{pmatrix} \omega(k)^{-2} & -\gamma^{-1} \\ \gamma^{-1} & 0 \end{pmatrix}.
\]

We are interested in controlling the behaviour of \( \bar{U}_t(x, k) \) at the diffusive scale \( t = O(L^2) \). We rely on the estimates derived in [2] and, for the sake of completeness, let us begin by summarizing the necessary assumptions from [2].

**Assumption 3.1** We assume that the map \( \Phi : \mathbb{Z} \to \mathbb{R} \) and the flip rate \( \gamma \) satisfy all of the following properties where \( \omega(k) = \sqrt{\Phi(k)} \) denotes the related phonon dispersion relation:
1. (exponential decay) There are $C, \delta > 0$ such $|\Phi(x)| \leq Ce^{-\delta|x|}$ for all $x \in \mathbb{Z}$,
2. (symmetry) $\Phi(-x) = \Phi(x)$ for all $x \in \mathbb{Z}$,
3. (pinning) There is $\omega_0 > 0$ such that $\omega(k) \geq \omega_0$ for all $k \in \mathbb{T}$,
4. (noise dominates) $\gamma > 2 \max_{k \in \mathbb{T}} \omega(k)$,
5. (harmonic forces are nondegenerate) For any $\varepsilon > 0$ there is $C_\varepsilon > 0$ such that

$$
\int_0^\infty dt \int dk \left( F_1 \left(k + \frac{k_0}{2}\right) - F_1 \left(k - \frac{k_0}{2}\right) \right)^2 \geq C_\varepsilon, \quad \text{whenever } \varepsilon \leq |k_0| \leq \frac{1}{2},
$$

(3.12)

where $F_1(k) := \hat{A}_{12}^2(k)$, for $\hat{A}$ defined in (3.3).

These assumptions are satisfied for instance by the nearest neighbor interactions, for which $\omega(k) = \sqrt{\omega_0^2 + 4\sin^2(\pi k)}$, whenever $\omega_0 > 0$ and $\gamma > 2\sqrt{\omega_0^2 + 4}$. We now state the first result of this paper.

**Theorem 3.2** Suppose that Assumption 3.1 holds. Then there is $L_0 > 0$ such that for any $\mathcal{E}_0, c_0 > 0$ we can find a constant $c_1 > 0$ using which the following result holds for every $L \geq L_0$.

Assume that the initial state is such that its energy density, $\mathcal{E} := |\Lambda_L|^{-1}(H_L(X(0)))$, is bounded by $\mathcal{E}_0$, that is, suppose that $\mathcal{E} \leq \mathcal{E}_0$. Then $U_t(x, k)$ defined in (3.1) yields a finite function of $x, k$ which satisfies for every $t \geq c_0 L^2$, $x \in \Lambda_L$, $k \in \Lambda_L^*$,

$$
|U_t(x, k) - (T_t(x)U_0(k) + i\nabla_x T_t(x)U_1(k))| \leq c_1 L^{-2},
$$

(3.13)

where $\nabla_x T_t(x) := T_t(x+1) - T_t(x)$ denotes a discrete gradient.

In the above, the constant $c_1$ may thus depend on $\mathcal{E}_0$, $c_0$, and $L_0$ but it is otherwise independent of the initial data and of $L$.

The bound in (3.13) can also be written as

$$
\begin{pmatrix}
U_t^{11}(x, k) & U_t^{12}(x, k) \\
U_t^{21}(x, k) & U_t^{22}(x, k)
\end{pmatrix} = T_t(x) \begin{pmatrix}
(\omega(k))^{-2} & 0 \\
0 & (\omega(k))^{-1}
\end{pmatrix} - \frac{iv(k)\nabla_x T_t(x)}{\omega(k)} \begin{pmatrix}
(\omega(k))^{-2} & -\gamma^{-1} \\
\gamma^{-1} & 0
\end{pmatrix} + O(L^{-2}).
$$

(3.14)

Here the $(2, 2)$-component of the dominant first term on the right hand side corresponds to the diffusive temperature profile found already in [2]. Together with the other three matrix components, the first term gives the expected local thermal equilibrium correlations since the $(q, p)$-correlation matrix of the equilibrium Gibbs state at temperature $T$ is equal to

$$
T \begin{pmatrix}
\Phi^{-1} & 0 \\
0 & 1
\end{pmatrix}.
$$

The second term on the right hand side is an $O(L^{-1})$ correction to the local equilibrium correlations. Its off-diagonal components can be interpreted as “current terms” while the origin of the diagonal terms will be clarified by the kinetic theory description discussed in Section 5.

### 3.2 Proof of Theorem 3.2

All computations in this subsection are made supposing that Assumption 3.1 holds. Since then also Assumptions 4.1 and 4.3 in [2] are valid, this will allow directly applying all results derived in that reference.

Let us begin the proof of the theorem by explaining how the assumptions, in particular the boundedness of the initial energy density, immediately yield an upper bound for the function $U_t$
proving, in particular, that it is finite. First, by the discrete Plancherel theorem and using the assumed pinning property, we have for any real \( q \)

\[
\omega_0^2 \sum_{x \in \Lambda_L} q_x^2 = \omega_0^2 \int_{\Lambda_L^*} dk \, |\hat{q}(k)|^2 \leq \int_{\Lambda_L^*} dk \, \omega(k)^2 |\hat{q}(k)|^2 = \sum_{x,y \in \Lambda_L} (\Phi_L)_{x,y} q_x q_y.
\]  

(3.15)

Therefore, for any real \( X = (q,p) \) we have

\[
\sum_{x \in \Lambda_L} p_x^2 \leq 2H_L(X) \quad \text{and} \quad \sum_{x \in \Lambda_L} q_x^2 \leq 2\omega_0^{-2}H_L(X).
\]

(3.16)

Thus the assumption \( \mathcal{E} \leq \mathcal{E}_0 < \infty \) and the conservation of the total energy imply that for \( i = 1, 2 \)

\[
\sum_{x \in \Lambda_L} E[(X_i^j(t))^2] \leq 2(1 + \omega_0^{-2})E[H_L(X(t))] = 2(1 + \omega_0^{-2})E[H_L(X(0))] = 2L(1 + \omega_0^{-2})\mathcal{E} < \infty.
\]

(3.17)

Therefore, by the Schwarz inequality,

\[
|U_i^{ij}(x,k)| \leq \sum_{y \in \Lambda_L} E \left[ |X_i^j(t)||X_{i+y}^j(t)| \right] \leq (L E[|X_i^j(t)|^2])^{1/2} \left( E \left[ \sum_{x \in \Lambda_L} |X_i^j(t)|^2 \right] \right)^{1/2} \leq 2(1 + \omega_0^{-2})\mathcal{E} L^{3/2} < \infty.
\]

(3.18)

Since \( \mathcal{E} \leq \mathcal{E}_0 \), this shows that \( U_t \) is finite and \( O(L^{3/2}) \). The theorem significantly improves this a priori bound for diffusive times since it implies that then \( U_t = O(1) \).

### 3.2.1 Review of the properties of \( T_t(x) \)

In this section we collect from [2] all the necessary ingredients for the derivation of equation (3.13). We will adopt the notation \( a \lesssim b \) to indicate \( a \leq Cb \), where \( C \) is a constant which might depend on \( \gamma \) and the function \( \omega \), but not on \( L \), \( t \) or the initial data. Furthermore, we will use for matrices the elementwise sum norm defined as

\[
\|B\| := \sum_{i,j=1}^n |B_{ij}|,
\]

(3.19)

where \( B \in \mathbb{C}^{n \times n} \). All finite matrix norms are equivalent, but the above choice is convenient for our purposes, in particular, since it is sub-multiplicative, i.e., always \( \|AB\| \leq \|A\| \|B\| \).

Thanks to Lemma 4.6 in [2] we know that there exist strictly positive constants \( \gamma_2 \) and \( \delta_0 \) such that

1. The entries of the matrix \( \hat{A}_t(k) \) belong to \( C^1([0, \infty) \times \mathbb{T}) \).

2. For every \( k \in \mathbb{T} \) and \( t \geq 0 \)

\[
\|\hat{A}_t(k)\| \lesssim e^{-\delta_0 t}, \quad \|\partial_t \hat{A}_t(k)\| \lesssim e^{-\delta_0 t}, \quad \|\partial_k \hat{A}_t(k)\| \lesssim e^{-\delta_0 t/2}.
\]

(3.20)

3. For all \( x \in \mathbb{Z} \) and \( t \geq 0 \)

\[
\|A_t(x)\| \lesssim e^{-\delta_0 t/2 - \gamma_2|x|}.
\]

(3.21)

To be more precise, the bounds (3.20) and (3.21) are only proven for the absolute value of the entries \( \hat{A}_t^2(k) \) and \( A_t^2(x) \) for \( i = 1, 2 \) in Lemma 4.6 of [2]. However, as is evident from (3.3), all the entries of \( \hat{A}_t(k) \) have the same analyticity and decay properties. Thus, these derivations can
be extended directly to every matrix element, and hence also to the matrix norm (3.19). Note that from (3.20) we immediately obtain
\[ \|a_{ij}\| \lesssim e^{-\delta_0 t}, \quad \|a_{ij}\| \lesssim e^{-\delta_0 t}. \] (3.22)

A renewal equation was derived in [2] for the noise-averaged temperature profile \( T_t(x) \) and its solution was shown to satisfy a linear diffusion equation at diffusive time-scales. Indeed, the defining equation for \( T_t(x) \), equation (4.2) in [2], is equal to the \((2, x), (2, x)\)-component of the Duhamel formula in (2.22) and thus their solutions coincide. Since now Assumptions 4.1 and 4.3 in [2] hold, and we have also proven that all second moments of \( X(0) \) are finite, we can apply Theorem 4.4 in [2]. We can thus conclude that there is \( L_0 > 0 \) such that for all \( L \geq L_0, \ t > 0 \) and \( x \in A_L \), the temperature profile \( T_t(x) = \langle p^2_t(x) \rangle \) satisfies
\[ |T_t(x) - (e^{-tD}\tau)_x| \lesssim \mathcal{E}Lt^{-3/2} \] (3.23)
where the discrete diffusion operator \( D \) is defined by
\[ (D\phi)_x := \sum_{y \in A_L} K_y (2\phi_x - \phi_{x+y} - \phi_{x-y}), \]
with\(^1\)
\[ K_x := \frac{\gamma}{2} \int_0^\infty ds K_{s,x}, \quad K_{t,x} := 2\gamma((e^{-tM_t})_{22}^2), \] (3.24)
The initial data vector \( \tau \) for the discrete diffusion has an explicit, but somewhat involved, dependence on the initial data of the particle system. Namely,
\[ \tau_x := \sum_{y \in A_L} \int_{A_L} dk e^{2\pi i k \cdot (x-y)} a(k) \int_0^\infty ds g_{s,y}, \] (3.25)
where
\[ g_{t,x} = (e^{-tM_t^\top} \Gamma_x e^{-tM_t})_{22}^2 (0, 0), \quad (\Gamma_x)^{ij}(y, y') := \mathcal{C}_{ij}^0 (x + y, x + y'), \]
and \( a(k) \) are explicit constants satisfying \( 0 \leq a(k) \lesssim 1 \). It is proven in Proposition 4.8 of [2] that \( g_{t,x} \) are positive and satisfy a bound \( \sum_x g_{t,x} \lesssim \mathcal{E}Le^{-\delta_0 t} \). Thus the initial data vector \( \tau \) and its discrete Fourier transform are bounded by the total energy,
\[ |\tau_x| \lesssim \mathcal{E}L, \quad |\hat{\tau}(k)| \lesssim \mathcal{E}L. \] (3.26)
The Fourier transform of the diffusion operator is given by
\[ \hat{D}(k) = \sum_{y \in A_L} (1 - \cos(2\pi k \cdot y))2\hat{K}_y. \] (3.27)
It is bounded from both above and below, \( 0 \leq \hat{D}(k) \leq 2\gamma \), and the assumptions can be used to show that its small \( k \) behaviour is controlled by the estimates
\[ C_1 \min(|k|, \varepsilon_0)^2 \leq \hat{D}(k) \leq C_2 k^2, \] (3.28)
where \( C_1, C_2, \varepsilon_0 > 0 \) are constants of the kind mentioned in the beginning of this section. The lattice diffusion approximation, \( (e^{-tD}\tau)_x \), is equal to \( \int_{A_L} dk e^{2\pi i k \cdot x} e^{-t\hat{D}(k)}\hat{\tau}(k) \), and thus it is bounded by
\[ |(e^{-tD}\tau)_x| \lesssim \mathcal{E}L. \] (3.29)
\(^1\)To avoid possible confusion with the particle momenta, we deviate here from the notations in [2] where “\( K_{t,x} \)” and “\( \hat{K}_x \)” are denoted by “\( p_t,x \)” and “\( \hat{p}_x \)”, respectively.
Therefore, thanks to (3.23), for all \(t > 0\) and \(x \in \Lambda_L\) we can write
\[
T_t(x) = (e^{-tD}\tau)x + \delta_t(x) \tag{3.30}
\]
where for large \(t\)
\[
|\delta_t(x)| \lesssim \mathcal{E}Lt^{-3/2}. \tag{3.31}
\]
Since \(T_t(x) = \langle p(t)^2 \rangle\) is obviously bounded by the total energy, we can also conclude validity of the following \(a \ priori\) bounds
\[
|T_t(x)| \lesssim \mathcal{E}L, \quad |\delta_t(x)| \lesssim \mathcal{E}L. \tag{3.32}
\]
These trivial bounds are used later only to control small values of \(t\) for which the more accurate estimate in (3.31) becomes uninformative.

### 3.2.2 Derivation of equation (3.13)

Now we have all the necessary ingredients to find an approximate evolution equation for the observable \(U_t(x, k)\) at the diffusive scale. To this end, we now assume that \(c_0 > 0\) is fixed and the initial data satisfies \(\mathcal{E} \leq \mathcal{E}_0\), and we then consider arbitrary \(L \geq L_0\) and \(t \geq c_0L^2\).

Let us start by examining the source term in (3.5): by shifting the summation variables we find
\[
Z_t(x, k) = \sum_{y, z \in \Lambda_L} e^{-i2\pi k \cdot y} A_t^\top(z - x)U_0(z, k)A_t(x + y - z) \\
= \sum_{z \in \Lambda_L} e^{-i2\pi k \cdot z} A_t^\top(z)U_0(x + z, k)\tilde{A}_t(k).\]

As proven in (3.18), the assumptions imply that \(\|U_0(x, k)\| \lesssim \mathcal{E}L^{3/2}\). Then by (3.21) and (3.20) we get
\[
\|Z_t(x, k)\| \leq \sum_{z \in \Lambda_L} \|A_t^\top(z)\|\|U_0(x + z, k)\|\|\tilde{A}_t(k)\| \lesssim \mathcal{E}L^{3/2}e^{-3\delta_0 t/2} \sum_{z \in \Lambda_L} e^{-\gamma_2|z|} \lesssim \mathcal{E}L^{3/2}e^{-3\delta_0 t/2},
\]
which is exponentially small in \(L\) for \(t \geq c_0L^2\) and \(\mathcal{E} \leq \mathcal{E}_0\). Let us denote terms which are exponentially small in \(L\) by \(O(e^{-\delta L})\) in the following without specifying the exact value of \(\delta > 0\). In particular, the value of \(\delta\) might vary from one equation to the next.

Hence, we may now conclude that \(\|Z_t(x, k)\| = O(e^{-\delta L})\), i.e., that \(\|Z_t(x, k)\| \leq ce^{-\delta L}\) with a constant \(c\) which might depend on \(L_0\), \(c_0\), and \(\mathcal{E}_0\) but is independent from initial data, \(x\), \(t\), and \(L\).

In order to analyse the second term in (3.6), let us decompose \(T_{t-s}(x + z)\) by (3.30) as
\[
T_{t-s}(x + z) = T_t(x + z) + [T_{t-s}(x + z) - T_t(x + z)] \\
= T_t(x + z) + [e^{-(t-s)D}\tau]_{x+z} - [e^{-tD}\tau]_{x+z} + [\delta_{t-s}(x + z) - \delta_t(x + z)] \\
= T_t(x + z) - \int_{t-s}^t ds' \partial_{s'} [e^{-s'D}\tau]_{x+z} + [\delta_{t-s}(x + z) - \delta_t(x + z)] \\
= T_t(x) + \int_{t-s}^t ds' \int_{\Lambda_L^*} dq e^{2\pi i q \cdot (x + z)} \hat{D}(q)e^{-s'D}(q)\tilde{\tau}(q) \\
+ [\delta_{t-s}(x + z) - \delta_t(x + z)] + [T_t(x + z) - T_t(x)] \tag{3.34}
\]
where in the last passage we wrote \((e^{-s'D\tau})_{x+z}\) in terms of its Fourier transform. Therefore, from (3.6) and (3.34) we get

\[
U_t(x, k) = Z_t(x, k) + 2\gamma T_t(x) \int_0^t ds \sum_{y, z \in \Lambda_L} e^{-12\pi k \cdot y} A_\nu^\top(z) P^{(2)} A_\nu(y - z)
\]

\[
+ 2\gamma \int_0^t ds \sum_{y, z \in \Lambda_L} e^{-12\pi k \cdot y} (T_{t-s}(x + z) - T_t(x)) A_\nu^\top(z) P^{(2)} A_\nu(y - z)
\]

\[
= Z_t(x, k) + 2\gamma T_t(x) \int_0^\infty ds \alpha_\nu^\top(k) + I^{(1)}(k) + I^{(2)}(k) + I^{(3)}(k) + I^{(4)}(k)
\]

where

\[
I^{(1)}_t(x, k) = -2\gamma T_t(x) \int_t^\infty ds \alpha_\nu^\top(k)
\]

\[
I^{(2)}_t(x, k) = 2\gamma \int_0^t ds \sum_{y, z \in \Lambda_L} e^{-12\pi k \cdot y} A_\nu^\top(z) P^{(2)} A_\nu(y - z) \left( \int_{t-s}^t ds' \int_{\Lambda_L} dq e^{2\pi i q(x+z)} \hat{D}(q)e^{-s'D(q)\tau(q)} \right)
\]

\[
= 2\gamma \int_0^t ds \int_{\Lambda_L} dq e^{2\pi i q x} \hat{A}_\nu^\top(k - q) P^{(2)} \hat{A}_\nu^\top(k) \left( \int_{t-s}^t ds' \hat{D}(q)e^{-s'D(q)\tau(q)} \right)
\]

\[
I^{(3)}_t(x, k) = 2\gamma \int_0^t ds \sum_{z \in \Lambda_L} e^{-12\pi k \cdot z} [\delta_{t-s}(x + z) - \delta_t(x + z)] A_\nu^\top(z) P^{(2)} \hat{A}_\nu^\top(k)
\]

\[
I^{(4)}_t(x, k) = 2\gamma \int_0^t ds \sum_{z \in \Lambda_L} e^{-12\pi k \cdot z} [T_t(x + z) - T_t(x)] A_\nu^\top(z) P^{(2)} \hat{A}_\nu^\top(k)
\]

We now consider separately the terms \(I^{(1)}_t, I^{(2)}_t, I^{(3)}_t\) and \(I^{(4)}_t\). The bounds (3.20), (3.21), (3.22), (3.31) and (3.32) yield

\[
\|I^{(1)}_t(x, k)\| \leq 2\gamma |T_t(x)| \int_t^\infty ds \|\alpha_\nu^\top(k)\| \lesssim \mathcal{E} L e^{-\delta_0 t}
\]

and

\[
\|I^{(3)}_t(x, k)\| \leq 2\gamma \int_0^t ds \sum_{z \in \Lambda_L} |\delta_{t-s}(x + z) - \delta_t(x + z)| \|A_\nu^\top(z)\| \|\hat{A}_\nu^\top(k)\|
\]

\[
\lesssim \int_0^t ds \sum_{z \in \Lambda_L} |\delta_{t-s}(x + z) - \delta_t(x + z)| e^{-\delta_0 s - \gamma_2 |z|}
\]

\[
\lesssim \sum_{z \in \Lambda_L} \left[ \int_0^{t/2} ds |\delta_{t-s}(x + z) - \delta_t(x + z)| e^{-\delta_0 s - \gamma_2 |z|} \right. \right.
\]

\[
+ \left. \left. \int_{t/2}^t ds |\delta_{t-s}(x + z) - \delta_t(x + z)| e^{-\delta_0 s - \gamma_2 |z|} \right]\right]
\]

\[
\lesssim \mathcal{E} L \left[ \int_0^{t/2} ds (t - s)^{-3/2} e^{-\delta_0 s} + \int_{t/2}^t ds e^{-\delta_0 s} \right] \lesssim \mathcal{E} L t^{-3/2}
\]

The estimate for \(I^{(2)}_t(x, k)\) is slightly more complicated: thanks to (3.20), (3.21), (3.28) and (3.26)
we get
\[ \| I^{(2)}(t, k) \| \lesssim E L \int_0^t ds e^{-\delta_0 s} \int_{t-s}^t ds' \int_{\Lambda^0_s} dq e^{-s\hat{D}(q)} |\hat{D}(q)| \]
\[ \lesssim E L \left[ \int_0^{t/2} ds e^{-\delta_0 s} \int_{t-s}^t ds' \int_{|q|\leq \epsilon_0} dq e^{-c_1 s' q^2} q^2 + \int_{t/2}^t dt_0 ds e^{-\delta_0 s} \int_{t-t_0}^t ds' \int_{|q|> \epsilon_0} dq e^{-s c_1 q^2} q^2 \right] \]
\[ + E L \left[ \int_0^{t/2} ds e^{-\delta_0 s} \int_{t-s}^t ds' e^{-s c_1 q^2} q^2 + \int_{t/2}^t dt_0 ds e^{-\delta_0 s} \int_{t-t_0}^t ds' e^{-s c_1 q^2} q^2 \right] \]
\[ =: J_a + J_b + J_c + J_d. \]

We now study each \( J_i \)'s contribution separately: by Lemma 4.11 of [2], \( \int_{|q|\leq \epsilon_0} dq e^{-c_1 s' q^2} q^2 \leq 4(c_1 s')^{-3/2} \), and thus
\[ J_a \lesssim E L \int_0^{t/2} ds e^{-\delta_0 s} \int_{t-s}^t ds' (s')^{-3/2} \lesssim E L \int_0^{t/2} ds e^{-\delta_0 s} s^{-3/2} \]
\[ \lesssim E L t^{-3/2} \int_0^{t/2} ds e^{-\delta_0 s} \lesssim E L t^{-3/2} \]
\[ J_b \lesssim E L \int_0^{t/2} ds e^{-\delta_0 s} s \lesssim E L t e^{-\delta_0 t/2} \]
\[ J_c \lesssim E L \int_0^{t/2} ds e^{-(t-s)c_1 \epsilon_0 s} \lesssim E L t e^{-c_1 \epsilon_0 t/2} \]
\[ J_d \lesssim E L \int_0^{t/2} ds e^{-\delta_0 s} \lesssim E L t e^{-\delta_0 t/2}. \]

From the computations above it follows that, on the diffusive scale \( t \geq c_0 L^2 \), the sum of the first three contributions is \( O(L^{-2}) \), i.e., \( I^{(1)} + I^{(2)} + I^{(3)} = O(L^{-2}) \).

We now focus on \( I^{(4)}(t, k) \) from which the dominant correction arises. We define the discrete gradient as \((\nabla f)(x) := f(x+1) - f(x)\). If \( y \geq 0 \), by induction one can check that
\[ f(x + y) - f(x) = y(\nabla f)(x) + \sum_{z=0}^{y-1} [((\nabla f)(x+z) - (\nabla f)(x))] \]
and, if \( y < 0 \), by using (3.36), one gets
\[ f(x + y) - f(x) = y(\nabla f)(x) + y[(\nabla f)(x+y) - (\nabla f)(x)] - \sum_{z=0}^{y-1} [((\nabla f)(x+y+z) - (\nabla f)(x+y))]. \]

For any \( z \in \Lambda_L \), let us define
\[ R(f; x, y) := f(x + y) - f(x) - y(\nabla f)(x) \]
which is the correction to the first order discrete Taylor expansion of \( f(x+y) \) around \( x \). Then, given the Fourier transform \( \hat{f} = \mathcal{F} f \), by exploiting (3.36), (3.37) and the inequality \(|e^{ir} - 1 - ir| \leq r^2/2\), valid for \( r \in \mathbb{R} \), one has
\[ |R(f; x, y)| \lesssim y^2 \int_{\Lambda^0_L} dq q^2 |\hat{f}(q)|. \]
On the other hand, we also have the trivial bound
\[ |\mathcal{R}(f; x, y)| \lesssim |y| \sup_{x \in \Lambda_L} |f(x)|. \] (3.40)

Thus, by (3.38) we can split \( I_t^{(4)}(x, k) \) as follows
\[
I_t^{(4)}(x, k) = 2\gamma \int_0^t ds \sum_{z \in \Lambda_L} e^{-i2\pi k \cdot z} [T_t(x + z) - T_t(x)] A_s^\top(z) P(z) \hat{A}_s(k)
\]
\[
= 2\gamma (\nabla T_t)(x) \int_0^t ds \sum_{z \in \Lambda_L} e^{-i2\pi k \cdot z} A_s^\top(z) P(z) \hat{A}_s(k)
\]
\[
+ 2\gamma \int_0^t ds \sum_{z \in \Lambda_L} e^{-i2\pi k \cdot z} R(T_t; x, z) A_s^\top(z) P(z) \hat{A}_s(k)
\]
\[
= \frac{2i\gamma}{2\pi} \nabla T_t(x) \int_0^\infty ds (\partial_k \hat{A}_s^\top(k)) P(z) \hat{A}_s(k) - \frac{2i\gamma}{2\pi} \nabla T_t(x) \int_0^\infty ds (\partial_k \hat{A}_s^\top(k)) P(z) \hat{A}_s(k)
\]
\[
+ 2\gamma \int_0^t ds \sum_{z \in \Lambda_L} e^{-i2\pi k \cdot z} R(T_t; x, z) A_s^\top(z) P(z) \hat{A}_s(k) + O(\delta L^4). \] (3.41)

The additional exponentially small correction \( O(e^{-\delta L}) \) arises from the following mismatch between discrete Fourier transform and the Fourier series. Suppose \( f : \mathbb{T}^d \to C \) is a continuously differentiable function whose Fourier coefficients are exponentially decaying, i.e., \( \mathcal{F}(f)(e^{i2\pi n \cdot k}) = \mathcal{O}(e^{-|n|}) \). Then the Fourier series of \( F \) converges at every point to \( f \), i.e., pointwise \( f(k) = \sum_{n \in \mathbb{Z}^d} e^{-i2\pi n \cdot k} F(n) \), \( k \in \mathbb{T}^d \). Thus the discrete Fourier transform of \( f \) restricted to \( \Lambda_L^* \) is equal to \( f(x) = \int_{\Lambda_L^*} dq e^{2\pi i q \cdot x} f(q) = \sum_{m \in \mathbb{Z}^d} F(x + Lm) \), for all \( x \in \Lambda_L \). Moreover, for any \( k \in \Lambda_L^* \),
\[
\sum_{x \in \Lambda_L} x e^{-i2\pi k \cdot x} f(x) = \sum_{m \in \mathbb{Z}^d} \sum_{x \in \Lambda_L} x e^{-i2\pi k \cdot x} F(x + Lm) \]
\[
= \sum_{m \in \mathbb{Z}^d} \sum_{x \in \Lambda_L} (x + Lm - Lm) e^{-i2\pi k \cdot (x + Lm)} F(x + Lm)
\]
\[
= \sum_{y \in \mathbb{Z}^d} y e^{-i2\pi k \cdot y} F(y) - \sum_{m \in \mathbb{Z}^2, m \neq 0} Lm \sum_{x \in \Lambda_L} e^{-i2\pi k \cdot (x + Lm)} F(x + Lm)
\]
\[
= \frac{i}{2\pi} \nabla f(k) + \mathcal{O}(e^{-\delta L/4}),
\]
where \( \nabla f \) denotes the ordinary (continuum) gradient of \( f \).

We identify the first term on the right hand side of (3.41) as the claimed correction term, more precisely
\[
\mathcal{J}_t(x, k) = \frac{2\gamma}{2\pi} \nabla T_t(x) \int_0^\infty ds (\partial_k \hat{A}_s^\top(k)) P(z) \hat{A}_s(k) = 2\gamma \nabla T_t(x) \int_0^\infty ds \partial_k \hat{A}_s(k).
\] (3.43)

We are left with showing that the second and third term on the right hand side of (3.41) are \( \mathcal{O}(L^{-2}) \) at the diffusive scale. By (3.32) and (3.22) for the second one we simply have
\[
\left\| \frac{2i\gamma}{2\pi} \nabla T_t(x) \int_0^\infty ds (\partial_k \hat{A}_s^\top(k)) P(z) \hat{A}_s(k) \right\| \lesssim \mathcal{E} L e^{-\delta_0 t}. \] (3.44)
By using (3.30) we can decompose the third term on the right hand side of (3.41) as
\[
2\gamma \sum_{z \in \Lambda_L} e^{-i2\pi k \cdot z} R(T_t; x, z) \int_0^t \mathrm{d}s A_s^T(z) P^{(2)}(k)
\]
\[
= 2\gamma \sum_{z \in \Lambda_L} e^{-i2\pi k \cdot z} R(e^{-itD_T}; x, z) \int_0^t \mathrm{d}s A_s^T(z) P^{(2)}(k)
\]
\[
+ 2\gamma \sum_{z \in \Lambda_L} e^{-i2\pi k \cdot z} R(\delta_k; x, z) \int_0^t \mathrm{d}s A_s^T(z) P^{(2)}(k)
\]
\[
=: I_t^{(5)}(x, k) + I_t^{(6)}(x, k).
\]
By using (3.20), (3.21), (3.31) and (3.40),
\[
\| I_t^{(6)}(x, k) \| \lesssim \mathcal{E} L t^{-3/2} \sum_{z \in \Lambda_L} |z| e^{-\gamma z^2 |z|} \int_0^t \mathrm{d}s e^{-\delta_0 s} \lesssim \mathcal{E} L t^{-3/2},
\]
while, thanks to (3.20), (3.21), (3.28) and (3.39), for \( I_t^{(5)}(x, k) \) we have
\[
\| I_t^{(5)}(x, k) \| \lesssim \sum_{z \in \Lambda_L} e^{-\gamma z^2 |z|} \int_0^\infty \mathrm{d}q q^2 e^{-it\hat{D}(q)} |\hat{\tau}(q)| \int_0^t \mathrm{d}s e^{-\delta_0 s}
\]
\[
\lesssim \mathcal{E} L \int_{\Lambda_L} \mathrm{d}q q^2 e^{-it\hat{D}(q)} \lesssim \mathcal{E} L \left[ \int_{|q| \leq \epsilon_0} \mathrm{d}q q^2 e^{-tc_1 q^2} + \int_{|q| > \epsilon_0} \mathrm{d}q q^2 e^{-tc_2 q^2} \right]
\]
\[
\lesssim \mathcal{E} L t^{-3/2}.
\]
This guarantees that on the diffusive scale \( I_t^{(5)} + I_t^{(6)} = O(L^{-2}) \). Putting together all the terms, we finally get the anticipated equation (3.13):
\[
U_t(x, k) = 2\gamma T_t(x) \int_0^t \mathrm{d}s \phi_s(k) + i \int_0^t \mathrm{d}s \phi_s(k) + O(L^{-2})
\]
\[
= 2\gamma T_t(x) \int_0^\infty \mathrm{d}s \phi_s(k) + i2\gamma \nabla T_t(x) \int_0^\infty \mathrm{d}s \phi_s(k) + O(L^{-2})
\]
\[
= T_t(x) U_0(k) + i\nabla T_t(x) U_1(k) + O(L^{-2})
\]
where \( \int_0^t \mathrm{d}s \phi_s(k) \) is analytic on a neighbourhood of the real axis and, consequently, its inverse Fourier-transform is an exponentially decreasing function on \( \mathbb{Z} \).

4 Kinetic theory of the velocity flip model

4.1 Time evolution of the mean Wigner function of normal modes

In this second part, we are interested in the evolution of a suitably modified Wigner transform \( \Psi_t^{\sigma_1, \sigma_2}(\xi, k) \) of phonon normal modes for which we derive a phonon Boltzmann equation. As in the previous Sections, we deal with dispersion relations which have pinning since then \( \omega(k) = \hat{\Phi}(k) \) is analytic on a neighbourhood of the real axis and, consequently, its inverse Fourier-transform is an exponentially decreasing function on \( \mathbb{Z} \).

It is possible to convert the standard definition of the Wigner function to the lattice setup using distribution techniques to handle points which lie outside the original lattice \([12, 13]\). We opt here for a different approach: by sacrificing real-valuedness of the Wigner transform, we may continue to consider it as a function, by using suitable partial Fourier transforms. More precisely, we consider here
\[
\Psi_t^{\sigma_1, \sigma_2}(\xi, k) := e^{i\omega(k)\sigma_1 + \sigma_2} \sum_{x \in \Lambda_L} \varphi(\xi - x) \sum_{y \in \Lambda_L} e^{-2\pi i y \cdot k} \xi[\psi_t(x, \sigma_1)\psi_t(x + y, \sigma_2)]
\]
where $\psi_t(x, 1) = \psi_t(x)$ and $\psi_t(x, -1) = \psi^*_t(x)$ are the normal modes of the harmonic evolution obtained by setting $\gamma$ to zero. In Fourier space, they are related to the positions and momenta of the particles by

$$\hat{\psi}_t(k, \sigma) = \frac{1}{\sqrt{2}}(\omega(k)\hat{q}(k) + i\sigma\hat{p}(k)),$$  \hspace{1cm} (4.2)

which implies

$$\hat{q}_t(k) = \frac{1}{\sqrt{2}\omega(k)}\sum_{\sigma = \pm 1}\hat{\psi}_t(k, \sigma), \quad \hat{p}_t(k) = -\frac{i}{\sqrt{2}}\sum_{\sigma = \pm 1}\sigma\hat{\psi}_t(k, \sigma).$$ \hspace{1cm} (4.3)

It is possible to modify the definition of the Wigner function in (4.1) so that it would enjoy the symmetry properties of the standard Wigner function simply by replacing the factor “$\varphi(\xi - x)$” by “$\varphi(\xi - x - \frac{1}{2}y)$”: then $(\mathcal{W}_t^{\sigma_1, \sigma_2})^* = \mathcal{W}_t^{-\sigma_2, -\sigma_1}$ and thus $\mathcal{W}_t^{-1}$ would become real-valued. However, this choice would make the argument of the testfunction depend on both $x$ and $y$ which would substantially complicate the forthcoming analysis. Indeed, in what follows we will show that without the $y$-dependence in $\varphi$ the sum over $y$ can be done explicitly, resulting in fairly simple collision operator which is closed under the definition (4.1).

There does not seem to be any straightforward way of making the Wigner function real without unnecessary complications. For instance, $\text{Re} \mathcal{W}$ would not satisfy a closed evolution equation. In addition, the field self-correlation term, $\mathcal{W}_t^{\sigma, \sigma}$, needs the complex factor “$e^{i\omega_t(k)2\sigma}$” to compensate its fast oscillations, resulting in a standard transport term in the corresponding Boltzmann equation, see (4.20) below. An additional benefit from the above formulation is that it easily generalizes to higher order cumulants which will become important for evolution problems involving anharmonic potentials [14, Chapter 4].

In (4.1), the prefactor $e^{i\omega_t(k)(\sigma_1 + \sigma_2)}$ is needed to cancel out fast oscillations resulting from the free evolution for the expectations when a mode is measured against itself, i.e., when $\sigma_1 = \sigma_2$. We also employ a convolution with $\varphi$ to focus on the large scale evolution in space, and we assume that it corresponds to spatial averaging over a given scale $R > 0$. It also provides a map from the discrete values evaluated at $x \in \Lambda_L$ into a smooth function on $\mathbb{R}^d$, $d = 1$. A convenient construction of the test function $\varphi$, which is also well-adapted to the underlying $L$-periodic lattice, is obtained by taking a Schwartz function $\phi \in \mathcal{S}(\mathbb{R}^d)$, and defining

$$\varphi(\xi) = \frac{1}{R^d}\sum_{n \in \mathbb{Z}^d}\phi\left(\frac{\xi - Ln}{R}\right), \quad \xi \in \mathbb{R}^d.$$ \hspace{1cm} (4.4)

This definition guarantees that $\varphi$ is smooth, $L$-periodic, and $\nabla_x^k \varphi = O(R^{-k})$ for all $k$.

If it is additionally assumed that $\phi$ is a positive function, normalized to $\int dx \phi(x) = 1$, and that its Fourier transform $\hat{\varphi}(k)$ has a compact support, we can identify the above convolution with taking of a local average over a region whose spatial radius is given by $R$. In fact, as shown in Appendix B, as soon as $R$ is greater than the radius of the support of $\hat{\varphi}$, one has $\sum_{x \in \Lambda_L} \varphi(\xi - x) = 1$ for all $\xi \in \mathbb{R}^d$. Therefore, for such test functions $\varphi$ the averaging preserves constant densities exactly, in the sense that it maps constant lattice fields to constant continuum fields without altering the value of the constant. In the following we shall call test functions $\varphi$ with this property lattice averaging kernels.

In this setting the total Hamiltonian energy reads

$$H = \frac{1}{2}\sum_{\sigma} \int_{\Lambda_L^2} dk |\hat{\psi}(k, \sigma)|^2.$$ 

Let us point out that the normal mode fields have been normalized so that their $\ell^2$-density measures directly the phonon energy; another common choice would be obtained by dividing the fields $\hat{\psi}$ by $1/\sqrt{\omega(k)}$ in which case the field can be thought of as measuring the phonon number density at wavenumber $k$ and each phonon mode carries then an energy $\omega(k)$. 

\hspace{1cm} 16
We now define a matrix $M$ as
\[
M(x) = \begin{pmatrix}
0 & \Phi(x) \\
-\mathbb{1}(x = 0) & 0
\end{pmatrix},
\]
thus
\[
\hat{M}(k) = \begin{pmatrix}
0 & \omega(k)^2 \\
-1 & 0
\end{pmatrix}.
\] (4.5)
We also use the same notation for the translation invariant matrix defined by $M(x, y) = M(x - y)$. This $M$ is equal to the matrix $M_x$, defined earlier in (2.16), evaluated at $\gamma = 0$.

Explicitly, the evolution equation for the position-momentum correlation $C_t$ in (2.20) becomes
\[
\partial_t C_t(x, y) = -(M^T C_t(x, y) - (C_t M)(x, y) - \gamma ((J C_t)(x, y) + (C_t J)(x, y)) + 2\gamma G_t(x, y),
\]
where $J(x, y) = \text{diag}(0, 1(x = y))$ and $G_t(x, y) = \text{diag}(0, \mathbb{1}(x = y)T_t(x))$, as before. In Fourier space, for $\hat{C}_t(k_1, k_2) = \sum_{x,y} e^{-i2\pi(xk_1 + yk_2)}C_t(x, y)$, one has
\[
\partial_t \hat{C}_t(k_1, k_2) = -\hat{M}(k_1)^{\top}\hat{C}_t(k_1, k_2) - \hat{C}_t(k_1, k_2)\hat{M}(k_2)
- \gamma (P^{(2)}\hat{C}_t(k_1, k_2) + \hat{C}_t(k_1, k_2)P^{(2)}) + 2\gamma \hat{T}_t(k_1 + k_2)P^{(2)}.
\] (4.6)
where $P^{(2)} = \text{diag}(0, 1)$. Now consider
\[
\mathbb{E}[\psi_t(x, \sigma_1)\psi_t(y, \sigma_2)] = \int_{(\Lambda^2)^2} dk_1dk_2 e^{2\pi ix\cdot k_1} e^{2\pi iy\cdot k_2} \mathbb{E}[^\hat{\psi}_t(k_1, \sigma_1)\hat{\psi}_t(k_2, \sigma_2)]
\]
where
\[
\mathbb{E}[^\hat{\psi}_t(k_1, \sigma_1)\hat{\psi}_t(k_2, \sigma_2)] = \frac{1}{2}[\omega(k_1)\omega(k_2)\hat{C}_t^{11}(k_1, k_2) + i\sigma_2\omega(k_1)\hat{C}_t^{12}(k_1, k_2)
+ i\sigma_1\omega(k_2)\hat{C}_t^{21}(k_1, k_2) - \sigma_1\sigma_2\hat{C}_t^{22}(k_1, k_2)]
= \text{Tr}[O(k_1, k_2; \sigma_1, \sigma_2)\hat{C}_t(k_1, k_2)]
\]
with
\[
O(k_1, k_2; \sigma_1, \sigma_2) = \frac{1}{2}\begin{pmatrix}
\omega(k_1)\omega(k_2) & i\sigma_2\omega(k_2) \\
i\sigma_1\omega(k_1) & -\sigma_1\sigma_2
\end{pmatrix}.
\]
This implies that
\[
Y^{\sigma_1, \sigma_2}_t(x, k) := \sum_{y \in \Lambda^2} e^{-2\pi iy\cdot k} \mathbb{E}[\psi_t(x, \sigma_1)\psi_t(x + y, \sigma_2)]
= \int_{\Lambda^2} dk' e^{2\pi i(x + k')\cdot k}\mathbb{E}[^\hat{\psi}_t(k', \sigma_1)\hat{\psi}_t(k, \sigma_2)]
= \int_{\Lambda^2} dk' e^{2\pi i(x + k')\cdot k} \text{Tr}[O(k', k; \sigma_1, \sigma_2)\hat{C}_t(k', k)],
\] (4.8)
where $Y^{\sigma_1, \sigma_2}_t(x, k)$ is such that
\[
\mathcal{W}^{\sigma_1, \sigma_2}_t(\xi, k) = e^{i\omega(k)(\sigma_1 + \sigma_2)} \sum_{x \in \Lambda^2} \varphi(\xi - x)Y^{\sigma_1, \sigma_2}_t(x, k).
\] (4.9)
Then, by using (4.6), we have
\[
\partial_t \mathcal{W}^{\sigma_1, \sigma_2}_t(\xi, k) = \sum_{x \in \Lambda^2} \varphi(\xi - x)e^{i\omega(k)(\sigma_1 + \sigma_2)} \int_{\Lambda^2} dk' e^{2\pi i(x + k')\cdot k'}
\times \{i(\sigma_1 + \sigma_2)\omega(k) \text{Tr}[O\hat{C}_t(k', k)] - \text{Tr}[O\hat{M}(k)^{\top}\hat{C}_t(k', k) + \hat{M}(k)O\hat{C}_t(k', k)]\}
- \gamma \sum_{x \in \Lambda^2} \varphi(\xi - x)e^{i\omega(k)(\sigma_1 + \sigma_2)} \int_{\Lambda^2} dk' e^{2\pi i(x + k')\cdot k'}
\times \text{Tr}[OP^{(2)}\hat{C}_t(k', k) + P^{(2)}\hat{C}_t(k', k) - 2\hat{T}_t(k + k')OP^{(2)}],
\] (4.10)
where \( O = O(k', k; \sigma_1, \sigma_2) \).

We refer to the second term in (4.10) as a collision term, and denote it by
\[
\mathcal{C}[W_t(\xi, \cdot)]^{\sigma_1, \sigma_2}(k) = -\gamma \sum_{x \in \Lambda_L} \varphi(\xi - x)e^{i\omega(k)(\sigma_1 + \sigma_2)} \int_{\Lambda_L} dk' e^{2\pi i x_\perp (k' + k')} \times \Tr[OP_{(2)} \widetilde{C}_t(k', k) + P^{(2)}_t \widetilde{C}_t(k', k) - 2T_t(k + k')O^{(2)}] ,
\]
where \( O = O(k', k; \sigma_1, \sigma_2) \).

We first focus on the \( \gamma \)-independent part. By performing the explicit matrix products we get
\[
\sum_{x \in \Lambda_L} \varphi(\xi - x)e^{i\omega(k)(\sigma_1 + \sigma_2)} \int_{\Lambda_L} dk' e^{2\pi i x_\perp (k' + k')} \\
\times \left\{ i(\sigma_1 + \sigma_2)\omega(k) \Tr[O\widetilde{C}_t(k', k)] - \Tr[O\widetilde{M}(k')^\dagger \widetilde{C}_t(k', k) + \widetilde{M}(k)O\widetilde{C}_t(k', k)] \right\}.
\]

Since
\[
\omega(k') = \omega(-k'),
\]
we may express here
\[
\omega(k) - \omega(k') = \sum_{x' \in \Lambda_L} \tilde{\omega}(x') e^{-i2\pi x_\perp k} \left( 1 - e^{i2\pi x_\perp (k' + k)} \right),
\]
where \( \tilde{\omega} \) denotes the inverse discrete Fourier transform of \( \omega \) restricted to \( \Lambda_L \). Therefore,
\[
\int_{\Lambda_L} dk' e^{2\pi i x_\perp (k' + k')} (\omega(k) - \omega(k')) \Tr[O\widetilde{C}_t(k', k)] = \sum_{x' \in \Lambda_L} \tilde{\omega}(x') e^{-i2\pi x_\perp k} (Y_t^{\sigma_1, \sigma_2}(x, k) - Y_t^{\sigma_1, \sigma_2}(x' + x, k))
\]
Inserting the formula above into (4.11) and performing a change a variables in the second term, we obtain that (4.11) is equal to
\[
\sigma_1 e^{i\omega(k)(\sigma_1 + \sigma_2)} \sum_{x,x' \in \Lambda_L} \tilde{\omega}(x') e^{-i2\pi x_\perp k} Y_t^{\sigma_1, \sigma_2}(x, k) (\varphi(\xi - x) - \varphi(\xi + x' - x)).
\]

Here by Taylor expansion we find that \( \varphi(\xi - x) - \varphi(\xi + x' - x) = -x' \cdot \nabla \varphi(\xi - x) + O((x'/R)^2) \).

Then we can replace the discrete sum \( \sum_{x' \in \Lambda_L} (-ix') \tilde{\omega}(x') e^{-i2\pi x_\perp k} \) with the derivative \( \nabla \omega(k)/(2\pi) \) plus a correction which is exponentially small in \( L \) due to the exponential decay of the Fourier transform of \( \omega \) (see the argument in (3.42) for more details). Therefore,
\[
\partial_t W_t^{\sigma_1, \sigma_2}(\xi, k) = \sigma_1 \frac{\nabla \omega(k)}{2\pi} \cdot \nabla_\xi W_t^{\sigma_1, \sigma_2}(\xi, k) + \mathcal{C}[W_t(\xi, \cdot)]^{\sigma_1, \sigma_2}(k) + O(R^{-2}).
\]

Let us now come back to the collision term. Since \( OP_{(2)}(k') \omega(k), \sigma_1 \sigma_2)/2 \), we have
\[
\mathcal{C}[W_t(\xi, \cdot)]^{\sigma_1, \sigma_2}(k) = -\sum_{x \in \Lambda_L} \varphi(\xi - x)e^{i\omega(k)(\sigma_1 + \sigma_2)} \int_{\Lambda_L} dk' e^{2\pi i x_\perp (k' + k')} \Tr \left[ \left( O - \frac{1}{2} \diag(\omega(k)\omega(k'), \sigma_1 \sigma_2) \right) \widetilde{C}_t(k', k) \right] \\
- \sigma_1 \sigma_2 \gamma \sum_{x \in \Lambda_L} \varphi(\xi - x)e^{i\omega(k)(\sigma_1 + \sigma_2)} \int_{\Lambda_L} dk' e^{2\pi i x_\perp (k' + k')} \tilde{T}_t(k + k') \\
- \gamma \sum_{x \in \Lambda_L} \varphi(\xi - x)e^{i\omega(k)(\sigma_1 + \sigma_2)} \int_{\Lambda_L} dk' e^{2\pi i x_\perp (k' + k')} [\omega(k)\omega(k')C_t^{11}(k', k) + \sigma_1 \sigma_2 \tilde{C}_t^{22}(k', k)] \\
- \sigma_1 \sigma_2 \gamma \sum_{x \in \Lambda_L} \varphi(\xi - x)e^{i\omega(k)(\sigma_1 + \sigma_2)} \int_{\Lambda_L} dqq' e^{2\pi i (q + q') x} \tilde{C}_t^{22}(q, q').
\]
Since
\[ \hat{C}^{11}_t(k_1, k_2) = \frac{1}{2\omega(k_1)\omega(k_2)} \sum_{\sigma_3\sigma_4} E[\hat{\psi}_t(k_1, \sigma_3)\hat{\psi}_t(k_2, \sigma_4)], \]
\[ \hat{C}^{22}_t(k_1, k_2) = -\frac{1}{2} \sum_{\sigma_3\sigma_4} \sigma_3\sigma_4 E[\hat{\psi}_t(k_1, \sigma_3)\hat{\psi}_t(k_2, \sigma_4)], \]
we obtain using (4.8)
\[
\mathcal{C}[\mathcal{W}_t(\xi, \cdot)]^{\sigma_1, \sigma_2}(k) = -\gamma \mathcal{C}_t^{\sigma_1, \sigma_2}(\xi, k) + \frac{\gamma}{4} \sum_{x \in \Lambda_L} \varphi(\xi - x) e^{it\omega(k)(\sigma_1 + \sigma_2)} \sum_{\sigma_3\sigma_4} (1 - \sigma_1\sigma_2\sigma_3\sigma_4) \mathcal{W}_t^{\sigma_3, \sigma_4}(x, k)
+ \frac{\gamma}{4} \int_{\Lambda_L} dq \sum_{\sigma_3\sigma_4} \sigma_3\sigma_4 \mathcal{W}_t^{\sigma_3, \sigma_4}(x, q)
= -\gamma \mathcal{C}_t^{\sigma_1, \sigma_2}(\xi, k) + \frac{\gamma}{4} \sum_{x \in \Lambda_L} \varphi(\xi - x) e^{it\omega(k)(\sigma_1 + \sigma_2)} \sum_{\sigma_3\sigma_4} (1 - \sigma_1\sigma_2\sigma_3\sigma_4) \mathcal{W}_t^{\sigma_3, \sigma_4}(x, k)
+ \frac{\gamma}{2} \sum_{\sigma_3\sigma_4} \sigma_1\sigma_2\sigma_3\sigma_4 \int_{\Lambda_L} dq e^{it\omega(k)(\sigma_1 + \sigma_2)} \mathcal{W}_t^{\sigma_3, \sigma_4}(\xi, q). \]

Expanding the various sign combinations explicitly thus yields
\[
\mathcal{C}[\mathcal{W}_t(\xi, \cdot)]^{\sigma_1, \sigma_2}(k) = \mathcal{C}_t^{\sigma_1, \sigma_2}(\xi, k) + \frac{\gamma}{4} \sum_{x \in \Lambda_L} \varphi(\xi - x) e^{it\omega(k)(\sigma_1 + \sigma_2)} (e^{-2it\omega(k)} \mathcal{W}_t^{+, +}(\xi, k) + e^{2it\omega(k)} \mathcal{W}_t^{-+, -}(\xi, k))
+ \frac{\gamma}{4} (1 + \sigma_1\sigma_2) e^{it\omega(k)(\sigma_1 + \sigma_2)} (\mathcal{W}_t^{+, -}(\xi, k) + \mathcal{W}_t^{-+, -}(\xi, k))
+ \frac{\gamma}{2} \sigma_1\sigma_2 e^{it\omega(k)(\sigma_1 + \sigma_2)} \int_{\Lambda_L} dq (e^{-2it\omega(q)} \mathcal{W}_t^{+, +}(\xi, q) + e^{2it\omega(q)} \mathcal{W}_t^{-+, -}(\xi, q))
- \frac{\gamma}{2} \sigma_1\sigma_2 e^{it\omega(k)(\sigma_1 + \sigma_2)} \int_{\Lambda_L} dq (\mathcal{W}_t^{+, -}(\xi, q) + \mathcal{W}_t^{-+, -}(\xi, q)). \]

If \( \varphi \) is real-valued, as we assume here, the components of \( \mathcal{W} \) can be related to each other by complex conjugation. Namely, then
\[
\mathcal{W}_t^{\sigma_1, \sigma_2}(\xi, k) = \left( \sum_{x, y \in \Lambda_L} \varphi(\xi - x) e^{2\pi ik} e^{it\omega(k)(\sigma_1 + \sigma_2)} E[\psi_t(x, -\sigma_1)\psi_t(y, -\sigma_2)] \right)^* \quad (4.14)
= (\mathcal{W}_t^{-\sigma_1, -\sigma_2}(\xi, -k))^*. \]

In addition, from the regularity properties of the test function we can also estimate the effect of swapping the sign of \( k \) and the order of the \( \sigma \)-indices: making a change of variables from \( x \) to \( x' = x + y \) it follows that
\[
\mathcal{W}_t^{\sigma_1, \sigma_2}(\xi, k) = \sum_{x', y \in \Lambda_L} \varphi(\xi - x') e^{-2\pi ik} e^{it\omega(k)(\sigma_1 + \sigma_2)} E[\psi_t(x', \sigma_2)\psi_t(x' - y, \sigma_1)] \quad (4.15)
= \sum_{x', y \in \Lambda_L} (\varphi(\xi - x') + \varphi(\xi - x') - \varphi(\xi - x') e^{2\pi ik} e^{it\omega(k)(\sigma_1 + \sigma_2)} E[\psi_t(x', \sigma_2)\psi_t(x' + y, \sigma_1)]
= \mathcal{W}_t^{\sigma_2, \sigma_1}(\xi, -k) + O(R^{-1}). \]

This second formula, however, needs to be used with some care since the correction might not be bounded in the lattice size \( L \). This is guaranteed if the correlations decay fast enough in space so
that $\sum_{y} |y| |E[\psi_t(x', \sigma_2) \psi_t(x' + y, \sigma_1)]|$ remains bounded in $L$. Whenever this is the case, we can combine the above bounds and conclude that
\begin{equation}
W_{t}^{\sigma_1, \sigma_2}(\xi, k) = (W_{t}^{\sigma_1, \sigma_2}(\xi, k))^* + O(R^{-1}).
\end{equation}

The closest quantity to the standard Wigner function is the function $W_{t}^{\gamma, \gamma}(\xi, k)$. By (4.16), it satisfies $W_{t}^{\gamma, \gamma}(\xi, k) = (W_{t}^{\gamma, \gamma}(\xi, k))^* + O(R^{-1})$ whenever the correlations decay sufficiently rapidly. Therefore, although this function is not necessarily real, its imaginary part is typically very small, due to the spatial averaging.

By using (4.14) and the symmetry of $\omega$ we find from (4.13)
\begin{equation}
\mathcal{E}[W_{t}^{\sigma_1, \sigma_2}(k)] = -\gamma W_{t}^{\sigma_1, \sigma_2}(\xi, k) + \frac{\gamma}{4} (1 - \sigma_1 \sigma_2) e^{i\omega(k)} (e^{-i2t\omega(k)} W_{t}^{\gamma, \gamma}(\xi, -k) + e^{i2t\omega(k)} W_{t}^{\gamma, \gamma}(\xi, k))
+ \frac{\gamma}{4} (1 + \sigma_1 \sigma_2) e^{i\omega(k)} (W_{t}^{\gamma, \gamma}(\xi, -k) + W_{t}^{\gamma, \gamma}(\xi, k))
+ \gamma \sigma_1 \sigma_2 e^{i\omega(k)} (e^{-i2t\omega(k)} W_{t}^{\gamma, \gamma}(\xi, q) - W_{t}^{\gamma, \gamma}(\xi, q)).
\end{equation}

Then the equal sign term is given by
\begin{equation}
\mathcal{E}[W_{t}^{\gamma, \gamma}(\xi, \cdot)]^+ (k) = -\gamma W_{t}^{\gamma, \gamma}(\xi, k) + \frac{\gamma}{2} e^{-i2t\omega(k)} (W_{t}^{\gamma, \gamma}(\xi, -k) + W_{t}^{\gamma, \gamma}(\xi, k))
+ \gamma e^{-i2t\omega(k)} \int_{\Lambda_L} dq Re \left[ e^{i2t\omega(q)} W_{t}^{\gamma, \gamma}(\xi, q) - W_{t}^{\gamma, \gamma}(\xi, q) \right].
\end{equation}

and the opposite sign term by
\begin{equation}
\mathcal{E}[W_{t}^{\gamma, \gamma}(\xi, \cdot)]^- (k) = -\gamma W_{t}^{\gamma, \gamma}(\xi, k) + \frac{\gamma}{2} (e^{-i2t\omega(k)} W_{t}^{\gamma, \gamma}(\xi, -k) + e^{i2t\omega(k)} W_{t}^{\gamma, \gamma}(\xi, k))
+ \gamma \int_{\Lambda_L} dq Re \left[ W_{t}^{\gamma, \gamma}(\xi, q) - e^{i2t\omega(q)} W_{t}^{\gamma, \gamma}(\xi, q) \right].
\end{equation}

Hence, these two functions satisfy a closed pair of evolution equations of the form
\begin{equation}
\partial_t W_{t}^{\gamma, \gamma}(\xi, k) + v(k) \cdot \nabla_\xi W_{t}^{\gamma, \gamma}(\xi, k) + O(R^{-2}) = \mathcal{E}[W_{t}^{\gamma, \gamma}(\xi, \cdot)]^+ (k),
\partial_t W_{t}^{\gamma, \gamma}(\xi, k) + v(k) \cdot \nabla_\xi W_{t}^{\gamma, \gamma}(\xi, k) + O(R^{-2}) = \mathcal{E}[W_{t}^{\gamma, \gamma}(\xi, \cdot)]^- (k),
\end{equation}
where both “bands” have a phonon velocity $v(k) = \frac{\nabla \omega(k)}{\omega(k)}$.

Let us stress that no approximations have been made to get to the above pair of equations, and they are valid for all scale parameters $R > 0$, as long as the testfunction $\varphi$ is constructed as mentioned in the beginning of this section. Of course, to be of any use as a transport equation, one needs to make sure that the effect of the correction terms marked as “$O(R^{-2})$” above remains small. What is commonly done in mathematical derivations of kinetic equations is to scale also time by $R$ and then take $R \to \infty$ in such a way that the collision operator on the right hand side has a finite nontrivial limit. In the present case, this could be achieved by taking $t = \tau R$ and $R = \gamma^{-1}$ and then considering a weak noise limit $\gamma \to 0$ for a fixed $\tau > 0$. This would correspond to the standard kinetic scaling limit, and we refer to [10, 17] for methods of controlling the limit rigorously in similar stochastic systems.

The kinetic scaling limit however hides a difficulty whose solution begs for an explanation: the above computation shows that the correction term $O(R^{-2})$ would be present even for pure harmonic evolution. It is in fact a necessary term which captures the difference between transport by the discrete wave equation and its radiative transport approximation obtained by setting the $O(R^{-2})$ term to zero. Apparently, then the kinetic equation is accurate only up to times $t = O(R^2)$ which with above kinetic scales would mean $t = O(\gamma^{-2})$. However, kinetic theory does correctly
predict the leading contribution to thermal conductivity in a number of phonon systems, and in the rest of this section we will show that this is also the case for the present velocity flip model. In fact, in this special case, the kinetic prediction turns out to be exact, and as proven in [2, 3], diffusion of energy persists for all sufficiently large times and describes correctly the $t \to \infty$ asymptotics of the energy density.

To reconcile the apparent restriction of kinetic theory to times $t \lesssim R^2$ with the fact that it does correctly capture even $t \to \infty$ asymptotics, we deviate here from the standard kinetic limit approach to kinetic equations by introducing a new spatial scale $R$ to its definition. We do not specify the value of $R$ exactly in the following, merely assume that it is sufficiently large that certain homogenization properties hold. In particular, we will assume that $R$ is much larger than the mean free path of phonons but not so large that it washes out macroscopic effects: we assume that $\gamma^{-1} \ll R \lesssim L$.

4.2 Stationary solutions

The left-hand sides in (4.20) include a time-derivative and a transport term of order $O(\gamma)$ while the collision terms on the right hand side are $O(\gamma)$ for small $\gamma$. Qualitatively, the equation corresponds to phonons with wavenumber $k$ moving at a velocity $v(k)$ and experiencing collisions at a rate $O(\gamma)$. Thus the mean free path of phonons should have a magnitude $\gamma^{-1}|v(k)|$. Therefore, for spatial scales much larger than the mean free path, i.e., whenever $R \gg \gamma^{-1}|v(k)|$, the time evolution of the above Wigner functions is dominated by the right hand side, i.e. the collision term. In that case, it is reasonable to start by first solving the equation including only the effect of collisions. Since the collisions do not mix values with different $\xi$, this amounts to solving the equations (4.20) for spatially homogeneous initial data. For this reason, let us suppose in this subsection that the point $\xi$ is a fixed parameter which we drop from the notation.

We now want to find the stationary solution of the above system in the translation invariant case. We define the following quantities:

\begin{align}
\mathcal{H}_t(k) &:= \frac{1}{2}(\mathcal{W}_t^{-,+}(k) + \mathcal{W}_t^{-,-}(k)^*) , \\
\mathcal{I}_t(k) &:= \frac{1}{2}(\mathcal{W}_t^{-,+}(k) - \mathcal{W}_t^{-,-}(k)^*) , \\
\mathcal{P}_t(k) &:= \frac{1}{2}(e^{2i\omega(k)}\mathcal{W}_t^{-,-}(k) + e^{-2i\omega(k)}\mathcal{W}_t^{-,-}(k)^*) , \\
\mathcal{Q}_t(k) &:= \frac{1}{2}(e^{2i\omega(k)}\mathcal{W}_t^{-,-}(k) - e^{-2i\omega(k)}\mathcal{W}_t^{-,-}(k)^*) .
\end{align}

Each of these functions is either symmetric ($\mathcal{H, P}$) or antisymmetric ($\mathcal{I, Q}$) under the transform $F(k) \to F(-k)^*$. They allow writing the collision operator in a very compact form. Namely, by introducing the simplified collision operator $\mathcal{E}$, defined as

$$
\mathcal{E}[f](k) := \gamma \int_{L^2} dq [f(q) - f(k)] ,
$$

we find from (4.20) the homogeneous evolution equations

$$
\begin{align}
\partial_t \mathcal{H}_t(k) &= \mathcal{E}[\mathcal{H}_t - \mathcal{P}_t](k) , \\
\partial_t \mathcal{I}_t(k) &= -\gamma \mathcal{I}_t(k) .
\end{align}
$$

In particular, then $\mathcal{I}_t(k) \equiv \mathcal{I}_0(k)e^{-\gamma t}$ and it approaches the unique stationary solution $\mathcal{I}(k) = 0$ exponentially fast as soon as $t = O(\gamma^{-1})$.

The homogeneous equations for $\mathcal{P, Q}$ are slightly more complicated, namely

$$
\begin{align}
\partial_t \begin{pmatrix} \mathcal{P}_t(k) \\ \mathcal{Q}_t(k) \end{pmatrix} &= \mathcal{L} \begin{pmatrix} \mathcal{P}_t(k) \\ \mathcal{Q}_t(k) \end{pmatrix} + \left( \gamma \int_{L^2} dq \mathcal{P}_t(q) \right)_{0} - \left( \mathcal{E}[\mathcal{H}_t](k) \right)_{0} ,
\end{align}
$$

21
where
\[ \mathcal{L} := \begin{pmatrix} -\gamma & 2i\omega(k) \\ 2i\omega(k) & -\gamma \end{pmatrix} \]
and its eigenvalues are \( \lambda = -\gamma \pm 2i\omega(k) \). The stationary solutions then satisfy
\[
\begin{pmatrix} \mathcal{P}(k) \\ \mathcal{Q}(k) \end{pmatrix} = -\mathcal{L}^{-1} \begin{pmatrix} \gamma \int_{\Lambda^*_L} dq \mathcal{P}(q) - \tilde{\mathcal{C}}[\mathcal{H}](k) \\ 0 \end{pmatrix}.
\] (4.25)

Since
\[
\mathcal{L}^{-1} = -\frac{1}{\gamma^2 + 4\omega^2(k)} \begin{pmatrix} \gamma & 2i\omega(k) \\ 2i\omega(k) & \gamma \end{pmatrix},
\]
one gets
\[
\begin{pmatrix} \mathcal{P}(k) \\ \mathcal{Q}(k) \end{pmatrix} = \frac{1}{\gamma^2 + 4\omega^2(k)} \begin{pmatrix} \gamma & 2i\omega(k) \\ 2i\omega(k) & \gamma \end{pmatrix} \begin{pmatrix} \gamma \int_{\Lambda^*_L} dq \mathcal{P}(q) - \tilde{\mathcal{C}}[\mathcal{H}](k) \\ 0 \end{pmatrix}.
\] (4.26)

We observe that the stationary equation corresponding to (4.23) is \( \tilde{\mathcal{C}}[\mathcal{H} - \mathcal{P}] = 0 \). Then, since \( \tilde{\mathcal{C}} \) is linear, we have
\[
\tilde{\mathcal{C}}[\mathcal{H}] = \tilde{\mathcal{C}}[\mathcal{P}].
\] (4.27)

Thus from (4.22), (4.26) and (4.27) the equation for \( \mathcal{P}(k) \) becomes
\[
\mathcal{P}(k) = \frac{\gamma}{\gamma^2 + 4\omega^2(k)} \left[ \gamma \int_{\Lambda^*_L} dq \mathcal{P}(q) - \tilde{\mathcal{C}}[\mathcal{P}](k) \right] = \frac{\gamma^2}{\gamma^2 + 4\omega^2(k)} \mathcal{P}(k).
\] (4.28)

Since \( \gamma^2/(\gamma^2 + 4\omega^2(k)) \) \( < 1 \) for \( \omega(k) > 0 \), then necessarily \( \mathcal{P}(k) = 0 \). Therefore, by (4.27) we now get \( \tilde{\mathcal{C}}[\mathcal{H}] = 0 \) and consequently also \( \mathcal{Q}(k) = 0 \).

The equation \( \tilde{\mathcal{C}}[\mathcal{H}] = 0 \) is solved precisely by functions which are constant in \( k \). Therefore, we have now proven that to each stationary solution there is a constant \( E \) such that \( \mathcal{H}(k) = E \) and \( 0 = \mathcal{P}(k) = \mathcal{Q}(k) = \mathcal{I}(k) \). In addition, then clearly
\[
\mathcal{W}^{-\leftarrow}(k) = 0,
\]
\[
\mathcal{W}^{-\rightarrow}(k) = \mathcal{H}(k) = \mathcal{W}^{-\rightarrow}(-k)^*,
\]
where \( \mathcal{W}^{-\rightarrow}(k) \) and \( \mathcal{W}^{-\leftarrow}(k) \) denote the stationary counterparts of \( \mathcal{W}^{-\leftarrow}(k) \) and \( \mathcal{W}^{-\rightarrow}(k) \). The second equality implies also that \( \mathcal{H}(k) = E \) is a real constant.

### 4.3 Boltzmann equation for the energy density

As we already observed before, the definition of \( \mathcal{W}_t(\xi,k) \) indicates that \( \mathcal{W}^{-\rightarrow}_{t}(\xi,k) \) is the quantity closest to the standard Wigner transform. Thus we would expect it to be of special interest in the kinetic theory; let us denote
\[
W_t(\xi,k) := \mathcal{W}^{-\rightarrow}_{t}(\xi,k).
\] (4.29)

The relaxation of \( W \) is then governed by the phonon Boltzmann equation
\[
\partial_t W_t(\xi,k) + v(k) \cdot \nabla_{\xi} W_t(\xi,k) = \tilde{\mathcal{C}}[W_t(\xi,\cdot)](k),
\] (4.30)
where we have used the simplified collision operator \( \tilde{\mathcal{C}} \) defined in (4.22). This equation follows from (4.20) after we assume that equilibration is so fast that both \( \tilde{\mathcal{C}}[\mathcal{P}_t] \) and the difference between \( W_t \) and \( \mathcal{H}_t \) can be neglected (note that \( W_t - \mathcal{H}_t = \mathcal{I}_t \), and thus it relaxes to zero independently from the other fields).

Moreover, assuming that \( R \gg \gamma^{-1} \), for times \( t = O(R) \), i.e. after the collisions have had plenty of time to push the system towards equilibrium, we expect that to every \( \xi \) there should be a real
constant $E_t(\xi)$ such that $W_t(\xi,k) - E_t(\xi)$ is small. By construction, $W_t$ is a function which varies only at the scale $R$ in $\xi$, i.e. $\nabla_\xi W_t = O(R^{-1})$, and thus $E_t(\xi)$ should then also be similarly slowly varying.

In order to find $E_t(\xi)$, we integrate the definition (4.1), so that

$$
\int_{\Lambda_L^*} \frac{dk}{L} W_t(\xi,k) = \sum_{x \in \Lambda_L} \varphi(\xi-x)E[|\psi_t(x)|^2] = \int_{\Lambda_L^*} \frac{dk}{L} H_t(\xi,k)
$$

which is clearly nonnegative for nonnegative testfunctions $\varphi$. If $\varphi$ is one of the “lattice averaging kernels” discussed in Section 4.1, we also have

$$
\int_{-L/2}^{L/2} \frac{d\xi}{L} \sum_{x \in \Lambda_L} \varphi(\xi-x)E[|\psi_t(x)|^2] = H_L(q_t, p_t) = H_L(q_0, p_0),
$$

and thus then we may identify the constant $E_t(\xi) := \int_{\Lambda_L^*} \frac{dk}{L} W_t(\xi,k)$ as the energy in a volume of radius $R$ centered at point $\xi$, i.e. it is equal to the physical energy density at $\xi$ at the time $t$. Furthermore, this implies that $W_t(\xi,k)$ can be interpreted as the “density” $\gamma^2$ in the phonon phase space with variables $(\xi, k)$, similarly to how the standard Wigner transform achieves the goal in quantum mechanics. Let us stress that, since the definition involves taking expectation over the randomness, this refers to the energy density averaged over realizations of the velocity flips.

### 4.4 Kinetic theory prediction for diffusion of energy

The Boltzmann equation (4.30) also allows studying the relaxation towards global equilibrium. This is one of the standard uses of kinetic theory, and we merely recall here how the argument works in the present case, giving only heuristic justification for the various steps. As mentioned in the Introduction, diffusion of energy in the present velocity flip model at standard hydrodynamics scales has already been rigorously proven in [2, 3]. We hence skip any rigorous estimates, and focus on trying to understand how the known diffusion phenomena is connected to the above kinetic equation. As a byproduct, we obtain a simple integral formula for the thermal conductivity which is shown to coincide with the previous results, at least in the special case of nearest neighbour interactions for which the integral can be computed analytically in the limit $L \to \infty$.

Let us suppose that the final phase of equilibration occurs via processes which are slower than ballistic, in which case $\partial_t W_t$ is small compared to $v(k) \cdot \nabla_\xi W_t$. This would occur for instance if the relaxation is diffusive, since then densities averaged over a volume of radius $O(R)$ change at a rate $O(R^{-2})$, and thus $\partial_t W_t = O(R^{-2})$ and $v(k) \cdot \nabla_\xi W_t = O(R^{-1})$.

Therefore, combined with the earlier relaxation argument, for such systems we expect that

$$W_t(\xi,k) = E_t(\xi) + \epsilon_t(\xi,k),$$

where $\epsilon_t$ is small and by the definition of $E_t(\xi)$ we have $\int_{\Lambda_L} \frac{dk}{L} \epsilon_t(\xi,k) = 0$. Since $\partial_t W_t$ is assumed to be of lower order, the dominant part of $\epsilon_t$ can be found by solving the equation

$$v(k)\nabla_\xi W_t(\xi,k) \simeq \mathcal{E}[W_t(\xi,\cdot)](k). \quad (4.31)$$

In the general version of the argument, which can be found for instance in Sec. 14 of [18], one then proceeds by using the expansion $\mathcal{E}[W_t(\xi,\cdot)](k) = \mathcal{L}_{E_t(\xi)}[\epsilon_t(\xi,\cdot)](k) + O(\epsilon_t^2)$ where $\mathcal{L}_E$ denotes the linearization of the collision operator $\mathcal{E}$ around the stationary solution $E$. Then the dominant perturbation can be found by applying the inverse $\mathcal{L}_{E_t(\xi)}^{-1}$ to (4.31).

In the present case, the collision operator is not only linear—which always implies that the linearized operator is the same as the original collision operator—but it is in fact a very simple projection operator. The inverse is explicit and for our purposes can be found directly from the definition of $\epsilon_t$. Namely, since $E_t(\xi) = \int_{\Lambda_L} \frac{dk}{L} W_t(\xi,k)$, we have

$$\mathcal{E}[W_t(\xi,\cdot)](k) = \gamma(E_t(\xi) - W_t(\xi,k)) = -\gamma \epsilon_t(\xi,k).$$

This function is not necessarily positive, hence the quotation marks here.
On the other hand, the dominant part of \( v(k) \nabla_\xi W_t(\xi, k) \) is given by \( v(k) \nabla_\xi E_t(\xi) \), and thus (4.31) implies that

\[ \epsilon(\xi, k) \simeq -\gamma^{-1} v(k) \nabla_\xi E_t(\xi). \]  

(4.32)

This result can be connected with the energy flux by using the conservation law which is reflected in the identity \( \int_{\Lambda L} dk \mathcal{V} [W](k) = 0 \), valid for any function \( W \). Thus for any solution of (4.30) we have

\[ \partial_t E_t(\xi) = \int_{\Lambda L} dk (-v(k) \nabla_\xi W_t(\xi, k)) = -\nabla_\xi j_t(\xi), \]

where

\[ j_t(\xi) := \int_{\Lambda L} dk v(k) W_t(\xi, k) \]

can be identified as the energy current. At equilibrium, for \( W_t(\xi, k) = E_t(\xi) \), the flux vanishes, since \( v(-k) = -v(k) \) due to the symmetry of the dispersion relation \( \omega \). Therefore, we can now conclude that the energy current satisfies \( j_t(\xi) = \int_{\Lambda L} dk v(k) \epsilon_t(\xi, k) \). Together with (4.32) this implies that, under the above assumptions about the relaxation process, the dominant part of the energy flux is given by

\[ j_t(\xi) \simeq -\kappa \nabla_\xi E_t(\xi), \]

where

\[ \kappa = \kappa(L) := \gamma^{-1} \int_{\Lambda L} dk v(k)^2. \]  

(4.33)

Inserting the approximation into the continuity equation then results in the equation

\[ \partial_t E_t(\xi) \simeq \kappa \nabla^2_\xi E_t(\xi), \]

which is a linear diffusion equation with a diffusion constant \( \kappa \). This in fact implies that if the assumption about eventual slow relaxation holds, then energy density must relax diffusively, with a diffusion constant \( \kappa \).

Finally, let us point out that this formula coincides with the conductivity obtained from the nonequilibrium steady state current of the system with the same bulk dynamics but with heat baths at the two ends enforcing a steady state current through the system. As in the references, suppose that the harmonic interactions connect only the nearest neighbours and have the dispersion relation \( \omega(k) = \sqrt{\omega_0^2 + 4 \sin^2(\pi k)} \), with \( \omega_0 > 0 \). As shown in [4], the steady state covariance matrix is then identical to the one of the so called self-consistent heat bath model. The self-consistent model was studied in detail in [19], and its thermal conductivity is given in Equation (4.18) of the reference. As shown a few lines above the formula, in Equation (4.16), the conductivity may be represented by a one-dimensional integral as

\[ \kappa[\text{Ref. [19]}] = \frac{1}{\gamma} \int_0^1 dx \frac{\sin^2(\pi x)}{\omega_0^2 + 4 \sin^4(\pi x/2)}. \]  

(4.34)

Since in this case \( v(k) = \omega'(k)/(2\pi) = \sin(2\pi k)/\omega(k) \), after employing evenness of the integrand and performing a change of variables to \( x = 2k \), the result clearly coincides with the \( L \to \infty \) limit of \( \kappa(L) \) given in Eq. (4.33) above.
4.5 Kinetic theory prediction for particle correlations

To inspect the accuracy of the above discussion in more detail, let us derive a prediction about the structure of the \(\rho_1(t), \rho_2(t)\) covariance matrix and compare this to the earlier results derived using the exact solution of its evolution. To facilitate the comparison, let us next consider the Wigner function of the position-momentum correlation matrix \(\mathcal{W}(\xi, k)\) which we define analogously to \(\mathcal{W}(\xi, k)\) using the formula

\[
\mathcal{W}_t(\xi, k) = \sum_{x,y \in \Lambda_L} \varphi(\xi - x) e^{-2\pi i y k} C_t(x,x + y). \tag{4.35}
\]

It is a spatially averaged version of the matrix function \(U_t(x, k)\) defined earlier in (3.1).

The change of basis formula (4.3) then yields

\[
\mathcal{W}_t^{11}(\xi, k) = \frac{1}{2} \sum_{x \in \Lambda_L} \varphi(\xi - x) \int_{\Lambda_L} dk' \frac{e^{2\pi i x(k + k')}}{\omega(k) \omega(k')} \sum_{\sigma, \sigma'} E[\tilde{\psi}_t(k', \sigma') \tilde{\psi}_t(k, \sigma)],
\]

\[
\mathcal{W}_t^{12}(\xi, k) = -\frac{1}{2} \sum_{x \in \Lambda_L} \varphi(\xi - x) \int_{\Lambda_L} dk' \frac{e^{2\pi i x(k + k')}}{\omega(k')} \sum_{\sigma, \sigma'} E[\tilde{\psi}_t(k', \sigma') \tilde{\psi}_t(k, \sigma)],
\]

\[
\mathcal{W}_t^{21}(\xi, k) = -\frac{1}{2} \sum_{x \in \Lambda_L} \varphi(\xi - x) \int_{\Lambda_L} dk' \frac{e^{2\pi i x(k + k')}}{\omega(k)} \sum_{\sigma, \sigma'} E[\tilde{\psi}_t(k', \sigma') \tilde{\psi}_t(k, \sigma)],
\]

\[
\mathcal{W}_t^{22}(\xi, k) = \frac{1}{2} \sum_{x \in \Lambda_L} \varphi(\xi - x) \int_{\Lambda_L} dk' \frac{e^{2\pi i x(k + k')}}{\omega(k)} \sum_{\sigma, \sigma'} E[\tilde{\psi}_t(k', \sigma') \tilde{\psi}_t(k, \sigma)].
\]

For \(\mathcal{W}_t^{21}\) and \(\mathcal{W}_t^{22}\) we obtain immediately from (4.8), (4.9), and (4.14)

\[
\mathcal{W}_t^{21}(\xi, k) = \frac{1}{2\omega(k)} \left( \mathcal{W}_t^{-+}(\xi, k) - \mathcal{W}_t^{+-}(\xi, k) + e^{2i\omega(k)} \mathcal{W}_t^{+-}(\xi, k) - e^{-2i\omega(k)} \mathcal{W}_t^{-+}(\xi, k) \right)
\]

\[
= \frac{1}{\omega(k)} (\mathcal{T}_t(\xi, k) - \mathcal{Q}_t(\xi, k)), \tag{4.36}
\]

\[
\mathcal{W}_t^{22}(\xi, k) = \frac{1}{2} \left( \mathcal{W}_t^{-+}(\xi, k) + \mathcal{W}_t^{+-}(\xi, k) - e^{2i\omega(k)} \mathcal{W}_t^{-+}(\xi, k) - e^{-2i\omega(k)} \mathcal{W}_t^{+-}(\xi, k) \right)
\]

\[
= \mathcal{H}_t(\xi, k) - \mathcal{P}_t(\xi, k). \tag{4.37}
\]

where we have employed the definitions in (4.21).

For \(\mathcal{W}_t^{11}\) and \(\mathcal{W}_t^{12}\) the factor \(1/\omega(k')\) complicates rewriting the result in terms of \(\mathcal{W}\). However, it is possible to go back to the scheme used for estimating (4.11) and exploit the regularity of the smoothing function to find out the dominant contribution. We begin by rewriting the sum over \(k\)' as a convolution:

\[
\int_{\Lambda_L} dk' \frac{e^{2\pi i x(k + k')}}{\omega(k')} \mathcal{E}[\tilde{\psi}_t(k', \sigma') \tilde{\psi}_t(k, \sigma)] = \sum_{x' \in \Lambda_L} \tilde{\omega}^{-1}(x - x') e^{2\pi i(x - x') \cdot k} \chi_{t+\sigma'}(x', k) \tag{4.38}
\]

where \(\tilde{\omega}^{-1}(y) = \int_{\Lambda_L} dk' e^{2\pi i y \cdot k} \omega(k')^{-1}\) is the inverse Fourier transform of \(1/\omega\). Therefore,

\[
\sum_{x \in \Lambda_L} \varphi(\xi - x) \int_{\Lambda_L} dk' \frac{e^{2\pi i x(k + k')}}{\omega(k')} \mathcal{E}[\tilde{\psi}_t(k', \sigma') \tilde{\psi}_t(k, \sigma)]
\]

\[
= \sum_{y,x' \in \Lambda_L} \varphi(\xi - x') \omega^{-1}(y) e^{2\pi i (x - x') \cdot k} \chi_{t+\sigma'}(x', k)
\]

\[
= \frac{1}{\omega(-k)} e^{-i\omega(k)(\sigma' + \sigma)} \chi_{t+\sigma'}(\xi, k)
\]

\[
+ \sum_{x', y \in \Lambda_L} (\varphi(\xi - x') - \varphi(\xi - x')) \omega^{-1}(y) e^{2\pi i y \cdot k} \chi_{t+\sigma'}(x', k). \tag{4.39}
\]
Here we use that \( \varphi(\xi - x' - y) - \varphi(\xi - x') = -y \cdot \nabla \varphi(\xi - x') + O((y/R)^2) \) and, with the small caveat about the difference between Fourier transforms on a finite and an infinite lattice explained in Section 3.2.2 at (3.42), we obtain

\[
\sum_{x \in \Lambda_L} \varphi(\xi - x) \int \frac{dk}{k} \frac{e^{i2\pi(k+k')}}{\omega(k')} \Im[\hat{\psi}(k', \sigma') \hat{\psi}(k, \sigma)] \\
= \frac{1}{\omega(k)} e^{-i\omega(k)(\sigma + \sigma')} \Psi_t(\sigma, \sigma)(\xi, k) - i \frac{1}{\omega(k)^2} \nabla \omega(k) \cdot \nabla \xi \Psi_t(\sigma, \sigma)(\xi, k) + O(R^{-2}).
\]

(4.40)

Applied in the definitions of \( \Psi^{11} \) and \( \Psi^{12} \), we thus find

\[
\Psi^{11}_t(\xi, k) = \frac{1}{\omega(k)^2} (H_t(\xi, k) + P_t(\xi, k)) - i \frac{1}{\omega(k)^2} v(k) \cdot (\nabla \xi H_t(\xi, k) + \nabla \xi P_t(\xi, k)) + O(R^{-2}),
\]

\[
\Psi^{12}_t(\xi, k) = -i \frac{1}{\omega(k)} (I_t(\xi, k) - Q_t(\xi, k)) - \frac{1}{\omega(k)^2} v(k) \cdot (\nabla \xi I_t(\xi, k) - \nabla \xi Q_t(\xi, k)) + O(R^{-2}).
\]

(4.41)

In Section 4.2, we found that for stationary homogeneous systems \( 0 = I = P = Q \) and \( H \) is a real constant. We could now repeat the analysis by including the derivative terms, and obtain also the magnitude of \( O(R^{-1}) \) corrections for near stationary systems. For instance, then

\[
\partial_t I_t(\xi, k) + v(k) \cdot \nabla \xi H_t(\xi, k) = -\gamma I_t(\xi, k),
\]

(4.42)

thus near stationarity

\( I_t(\xi, k) \approx -\gamma^{-1} v(k) \cdot \nabla \xi H_t(\xi, k) = O(R^{-1}). \)

The analysis of the magnitude of the other terms is, however, more involved. If one concentrates on the scaling and ignores possible regularity issues, it is possible to reproduce the computations in Section 4.2, and in general one should have then \( P, Q = O(R^{-2}) \) and \( \Psi[H_t] = O(R^{-2}) \), implying \( H_t(\xi, k) = E_t(\xi) + O(R^{-2}) \). (We have sketched some details of the derivation in Appendix C.)

Whenever this is the case, the particle correlations satisfy

\[
\Psi^{11}_t(\xi, k) = \frac{1}{\omega(k)^2} E_t(\xi) - i \frac{1}{\omega(k)^2} v(k) \cdot \nabla \xi E_t(\xi) + O(R^{-2}),
\]

(4.43)

\[
\Psi^{12}_t(\xi, k) = i \gamma^{-1} \frac{1}{\omega(k)^2} v(k) \cdot \nabla \xi E_t(\xi) + O(R^{-2}),
\]

(4.44)

\[
\Psi^{21}_t(\xi, k) = -i \gamma^{-1} \frac{1}{\omega(k)} v(k) \cdot \nabla \xi E_t(\xi) + O(R^{-2}),
\]

(4.45)

\[
\Psi^{22}_t(\xi, k) = E_t(\xi) + O(R^{-2}).
\]

Thus, written in a matrix form,

\[
\Psi_t(\xi, k) = E_t(\xi) \begin{pmatrix} \omega(k)^2 & 0 \\ 0 & 1 \end{pmatrix} - i \frac{1}{\omega(k)} v(k) \cdot \nabla \xi E_t(\xi) \begin{pmatrix} \omega(k)^2 \gamma^{-1} & -\gamma^{-1} \\ \gamma^{-1} & 0 \end{pmatrix} + O(R^{-2}).
\]

(4.46)

Let us point out that this result could have been derived from the Boltzmann equation discussed in Section 4.4 by assuming that \( P, Q \) and \( \text{Im} W_t(\xi, k) \) are of lower order, namely \( O(R^{-2}) \). Since then \( W_t(\xi, k)^* = W_t(\xi, k) + O(R^{-2}) \), the definitions directly imply that \( H_t \) is equal to the part of \( W_t \) even in \( k \), and \( I_t \) is equal to the part odd in \( k \). Hence the result \( W_t(\xi, k) = E_t(\xi) - \gamma^{-1} v(k) \nabla \xi E_t(\xi) + O(R^{-2}) \) implies precisely that \( H_t(\xi, k) = E_t(\xi) + O(R^{-2}) \) and \( I_t(\xi, k) = -\gamma^{-1} v(k) \cdot \nabla \xi E_t(\xi) + O(R^{-2}) \). This, together with \( P, Q = O(R^{-2}) \), suffices to give the form (4.46) for the covariance matrix.
The kinetic theory of the velocity flip model discussed in Section 4 shows that from the point of view of phonons, the diffusive scale relaxation on the level of the second order correlations is entirely described by the simple phonon Boltzmann equation (4.30) for the polarization component $W_t = W - t$ while the self-polarization component $W_{t-}$ may be set to zero at that stage of the evolution. The dominant contribution in this picture is given by a local equilibrium term and the first order corrections are directly related to energy currents.

Transformed from phonon modes back to $(q,p)$-fields, these two terms yield the expansion (4.46) for the spatially averaged correlation matrix. A comparison with the earlier result derived using the explicit estimates, given in (3.14), shows that the two dominant terms are identical. The lattice kinetic temperature profile $T_t(x), x \in \Lambda$, in (3.14) merely needs to be replaced by its spatially averaged version $E_t(\xi), \xi \in \mathbb{R}$, in (4.46) (note that the temperature is equal to the energy density in this model at thermal equilibrium).

Also for $(q,p)$-fields the dominant correlations are determined by the local thermal equilibrium correlations. However, the first order corrections acquire a term which is not related to the current observable, namely an additional correction to the $(q,q)$-correlations. As seen from the computations in Section 4.5, this correction arises from the convolution which transforms the phonon modes back to particle variables. It is also evident that this correction is zero whenever the state is translation invariant, so we may interpret it as a correction arising from changes in the phonon eigenbasis related to the inhomogeneous energy distribution.

Moreover, the relation between the phonon modes and the original Hamiltonian variables—between $\mathcal{W}$ and $\mathcal{U}$ in Section 4.5—does not depend on how the harmonic Hamiltonian evolution is perturbed. However, the kinetic theory collision operator will greatly depend on the perturbation and hence it is not obvious that the self-correlation terms can be neglected in all models relevant to transport of phonons in crystalline structures. For instance, it would be of interest to check more carefully what happens in models related to real three-dimensional crystals where the perturbations are small nonlinearities in the potential and there can be many different dispersion relations, as well as multidimensional phonon mode eigenspaces.

To avoid complications arising from boundary effects, we have considered here only energy transport in periodic particle chains. Another commonly used setup is to use fixed boundary conditions and attach two thermostats to each end of the chain. The thermostats drive the ends towards thermal equilibrium with some predetermined temperatures, and such a system is then expected to reach a steady state with a temperature profile which can be solved from the Fourier’s law using the boundary conditions given by the thermostats. Then at the steady state in the bulk, i.e., sufficiently far away from the boundary, the system will have a temperature gradient $O(1/L)$ where $L$ is the length of the chain. As the effect of the thermostats to the dynamics is expected to remain concentrated to the boundary, the bulk dynamics should be well approximated by the dynamics of the periodic chain. Therefore, as a consequence of the above results, we can now make a precise conjecture about the structure of the above nonequilibrium steady state correlations: the dominant local correlations are determined by the value of temperature at the site but they exhibit a correction whose leading term is proportional to the temperature gradient and has the structure derived above in (3.14).

Here we have compared two different schemes to study thermal transport in the velocity flip model: the explicit estimates relying on the renewal equation and the kinetic theory from the spatially averaged Wigner function. The comparison highlights the strengths and weaknesses of both approaches. Renewal equation and the pointwise estimates are more sensitive to the local lattice dynamics and can detect for instance degeneracies which are washed out by the spatial averaging in the other approach. For instance, a chain with only next-to-nearest particle potential and an even number of particles will decouple into two non-interacting chains which thermalize independently from each other and might, for instance, reach different temperatures at equilibrium. This is one of the reasons for the somewhat complicated condition—which fails in the above degenerate case—for the dispersion relation in [2] where uniform microscopic control was the goal.
However, it is probably fair to assume that the explicit computations in [2] do not easily carry over to other models and the uniform control will remain a hard goal for most phonon systems. Although the spatial averaging can wash out relevant details from the dynamics, it is the key to the separation of scales between transport and collisions in the kinetic theory computations in Section 4. We assume there $\gamma^{-1} \ll R \lesssim L$, but some additional assumptions will likely be needed if one wishes to complete the missing details and prove rigorously that the conjectured behaviour actually occurs for the velocity flip model. However, as the best one could hope for from such a computation in the velocity flip model would be a reproduction of the existing diffusion proofs, the extra effort would likely pay off only in other, more complicated, models such as particle chains with anharmonic perturbations.

A Computation of $q(k)$ in (3.11)

In this section we show the explicit computation which proves (3.10). We look at (3.9), i.e.

$$q(k) = 2\gamma \int_0^\infty dt \hat{A}_t(k)^2 \partial_k \hat{A}_t(k)^{21}. \quad (A.1)$$

Here

$$\hat{A}_t(k)^{22} = \frac{e^{-\frac{\gamma}{2}}}{\Omega} \left( -\frac{\gamma}{2} \sinh \Omega t + \Omega \cosh \Omega t \right),$$

$$\partial_k \hat{A}_t(k)^{21} = \frac{e^{-\frac{\gamma}{2}\Omega'}}{\Omega} \left( t \cosh \Omega t - \frac{\sinh \Omega t}{\Omega} \right).$$

Therefore the integrand in (A.1), i.e., $2\gamma \hat{A}_t(k)^{22} \partial_k \hat{A}_t(k)^{21}$ reads

$$\begin{align*}
2\gamma \frac{e^{-\gamma'\Omega'}}{\Omega'^2} & \left( -\frac{\gamma}{2} \sinh \Omega t \cosh \Omega t + \frac{\gamma}{2\Omega} \sinh^2 \Omega t + \Omega t \cosh^2 \Omega t - \cosh \Omega t \sinh \Omega t \right) \\
& = 2\gamma \frac{\Omega'}{\Omega'^2} \left[ -\frac{\gamma}{8} \left( e^{-t(\gamma-2\Omega)} - e^{-t(\gamma+2\Omega)} \right) + \frac{\gamma}{8\Omega} \left( e^{-t(\gamma-2\Omega)} + e^{-t(\gamma+2\Omega)} - 2e^{-t\gamma} \right) \right] \\
& + 2\gamma \frac{\Omega'}{\Omega'^2} \left[ \frac{\Omega t}{4} \left( e^{-t(\gamma-2\Omega)} + e^{-t(\gamma+2\Omega)} + 2e^{-t\gamma} \right) - \frac{1}{4} \left( e^{-t(\gamma-2\Omega)} - e^{-t(\gamma+2\Omega)} \right) \right].
\end{align*}$$

where the “prime” denotes the derivative with respect to $k$. Once we integrate with respect to the time variable, we obtain

$$q(k) = 2\gamma \frac{\Omega'}{\Omega'^2} \left[ -\frac{\gamma}{8} \left( \frac{1}{(\gamma - 2\Omega)^2} - \frac{1}{(\gamma + 2\Omega)^2} \right) + \frac{\gamma}{8\Omega} \left( \frac{1}{(\gamma - 2\Omega)} + \frac{1}{(\gamma + 2\Omega)} - \frac{2}{\gamma} \right) \right] \\
+ 2\gamma \frac{\Omega'}{\Omega'^2} \left[ \frac{1}{4} \left( \frac{1}{(\gamma - 2\Omega)^2} + \frac{2}{(\gamma + 2\Omega)^2} + \frac{2}{\gamma^2} \right) - \frac{1}{4} \left( \frac{1}{(\gamma - 2\Omega)} - \frac{1}{(\gamma + 2\Omega)} \right) \right] \\
= 2\gamma \frac{\Omega'}{\Omega'^2} \left[ -\frac{\gamma^2}{(\gamma^2 - 4\Omega'^2)} + \frac{\Omega}{\gamma^2} \left( \frac{\gamma^4 - 2\Omega^2\gamma^2}{(\gamma^2 - 4\Omega'^2)^2} + \frac{8\Omega^4}{\gamma^2} \right) \right] \\
= -\frac{4\Omega\Omega'}{\gamma (\gamma^2 - 4\Omega'^2)}. $$

Since $\Omega = (\gamma/2)\sqrt{1 - (2\omega(k)/\gamma)^2}$ it results that $\Omega' = -\frac{\omega'}{\Omega'\gamma}$. By inserting the explicit expression for $\Omega'$ in the previous computation we get

$$q(k) = \frac{\partial_k \omega(k)}{\gamma \omega(k)}. $$

28
B Basic properties of lattice averaging kernels

In Section 4.1, we referred to “lattice averaging kernels” which were understood as convolution sums constructed using the kernel functions

\[ \varphi(\xi) = \frac{1}{R^d} \sum_{n \in \mathbb{Z}^d} \phi \left( \frac{\xi - Ln}{R} \right), \quad \xi \in \mathbb{R}^d, \tag{B.1} \]

for some given \( L, R > 0 \). These kernels are determined via the function \( \phi : \mathbb{R}^d \to \mathbb{R} \) which we assume to satisfy all of the following conditions

1. \( \phi \) is a Schwartz test function, i.e., \( \phi \in \mathcal{S}(\mathbb{R}^d) \).
2. \( \hat{\phi} \) has a compact support. Let \( \rho_\phi > 0 \) be such that \( \hat{\phi}(p) = 0 \) whenever \( |p|_\infty \geq \rho_\phi \).
3. \( \phi \geq 0 \).
4. \( \int dy \, \phi(y) = 1 \).

Since this construction could become useful in phonon models in higher dimensions, we write the results below for arbitrary \( d \geq 1 \), keeping in mind that in the text they are applied with \( d = 1 \).

The main difference comes from the fact that for \( d > 1 \), the max-norm \( |y|_\infty := \max_{1 \leq k \leq d} |y_k| \) and the Euclidean norm \( |y| := (y_1^2 + y_2^2 + \cdots + y_d^2)^{1/2} \) no longer give the same numbers. We will mainly need the max-norm for the present lattice systems.

Let us show next that these assumptions guarantee the following properties for \( \varphi \):

1. (positivity) \( \varphi \geq 0 \).
2. (L-periodicity) \( \varphi(\xi + Lm) = \varphi(\xi) \) for all \( \xi \in \mathbb{R}^d, m \in \mathbb{Z}^d \).
3. (continuum normalization) \( \int_{|\xi|_\infty \leq L/r} d\xi \varphi(\xi - \xi_0) = 1 \) for all \( \xi_0 \in \mathbb{R}^d \).
4. (slow variation) To every multi-index \( \alpha \) there is a constant \( C_\alpha \), which is independent of \( R \) and \( L \), such that

\[ |\partial^\alpha \varphi(\xi)| \leq R^{-|\alpha|} C_\alpha, \quad \text{for all} \ \xi \in \mathbb{R}^d. \tag{B.2} \]

5. (lattice normalization) If \( R \geq \rho_\phi \), we have \( \sum_{x \in \Lambda_L} \varphi(\xi + x) = 1 \) for all \( \xi \in \mathbb{R}^d \).
6. (discrete Fourier transform) If \( R \geq 2 \rho_\phi \), we have for all \( k \in \Lambda_L^* \), \( \xi \in \mathbb{R}^d \)

\[ \sum_{x \in \Lambda_L} \varphi(\xi - x) e^{-2\pi i x \cdot k} = e^{-2\pi i \xi \cdot k} \hat{\phi}(-Rk). \tag{B.3} \]

Hence, the constant \( L \) determines the periodicity of the kernel and \( R \) the scale of variation, in the sense that each derivative of \( \varphi \) will decrease the magnitude by \( R^{-1} \).

The items 1 and 2 are obvious consequences of the definition of \( \varphi \) and the assumptions on \( \phi \). Item 3 is derived by rewriting the sum over integrals as a single integral as follows:

\[ \int_{|\xi|_\infty \leq L} d\xi \varphi(\xi - \xi_0) = \frac{1}{R^d} \sum_{n \in \mathbb{Z}^d} \int_{|\xi|_\infty \leq L/r} d\xi \phi \left( \frac{\xi + Ln - \xi_0}{R} \right) = \frac{1}{R^d} \int_{R^d} dy \, \phi \left( \frac{y - \xi_0}{R} \right) = 1. \tag{B.4} \]

Item 4 follows by taking the derivative inside the sum over \( n \), and then noticing that the result can be bounded by \( R^{-|\alpha|} \) times a Riemann sum approximation of the integral \( \int dy \, |\partial^\alpha \phi(y)| \) which is finite since \( \phi \) is a Schwartz function.

The lattice normalization condition and Fourier transform in items 5 and 6 need slightly more effort. Applying the definitions of \( \varphi \) and of the finite lattice \( \Lambda_L \), we obtain for any \( k \in \Lambda_L^*, \xi \in \mathbb{R}^d \):

\[ \sum_{x \in \Lambda_L} \varphi(\xi - x) e^{-2\pi i x \cdot k} = \frac{1}{R^d} \sum_{n \in \mathbb{Z}^d} \phi \left( \frac{\xi - x - Ln}{R} \right) e^{-2\pi i (x + Ln) \cdot k} = \frac{1}{R^d} \sum_{m \in \mathbb{Z}^d} f(m). \tag{B.5} \]
where \( f(y) := \phi \left( \frac{\xi - y}{R} \right) e^{-i2\pi y \cdot k} \) is a Schwartz function. The Fourier transform of \( f \) is given by

\[
\hat{f}(p) = R^d e^{-i2\pi \xi \cdot (p+k)} \hat{\phi}(-R(p+k)).
\]

Therefore, by the Poisson summation formula,

\[
\sum_{x \in \Lambda_L} \varphi(\xi - x)e^{-i2\pi x \cdot k} = \frac{1}{R^d} \sum_{m \in \mathbb{Z}^d} \hat{f}(m) = \sum_{m \in \mathbb{Z}^d} e^{-i2\pi \xi \cdot (m+k)} \hat{\phi}(-R(m+k)). \quad (B.6)
\]

If \( m \neq 0 \), we have \(|m + k|_{\infty} \geq |m|_{\infty} - |k|_{\infty} \geq \frac{1}{2} \), and thus \(|-R(m+k)|_{\infty} \geq R/2 \). Hence, if \( R \geq 2\rho_\phi \), or \( k = 0 \) and \( R \geq \rho_\phi \), all these points lie outside the support of \( \hat{\phi} \), and thus only the “\( m = 0 \)” term may contribute to the sum. This yields

\[
\sum_{x \in \Lambda_L} \varphi(\xi - x)e^{-i2\pi x \cdot k} = e^{-i2\pi \xi \cdot k} \hat{\phi}(-Rk). \quad (B.7)
\]

In particular, if \( k = 0 \), we have \( \hat{\phi}(-Rk) = \hat{\phi}(0) = f \int dy \phi(y) = 1 \), and we obtain \( \sum_{x \in \Lambda_L} \varphi(\xi - x) = 1 \). This completes the proof of both item 5 and item 6.

### Appendix C Quasi-stationary inhomogeneous solutions

Here we want to show that \( \mathcal{P}_t = Q_t = O(R^{-2}) \), \( \mathcal{I}_t = O(R^{-1}) \) and \( \mathcal{H}_t = E_t + O(R^{-2}) \) as anticipated in Section 4.5. Using the definitions (4.21) including the \( \xi \)-dependence, as well as the antisymmetry \( v(-k) = -v(k) \), from (4.20) we deduce

\[
\partial_t \mathcal{H}_t(\xi, k) = -v(k) \nabla_\xi \mathcal{I}_t(k, \xi) + \mathcal{E}[\mathcal{H}_t - \mathcal{P}_t](\xi, k) + O(R^{-2}) \quad (C.1)
\]

\[
\partial_t \mathcal{I}_t(\xi, k) = -v(k) \nabla_\xi \mathcal{H}_t(k, \xi) - \gamma \mathcal{I}_t(\xi, k) + O(R^{-2}) \quad (C.2)
\]

\[
\partial_t \begin{pmatrix} \mathcal{P}_t(\xi, k) \\ \mathcal{Q}_t(\xi, k) \end{pmatrix} = \mathcal{L}_v \begin{pmatrix} \mathcal{P}_t(\xi, k) \\ \mathcal{Q}_t(\xi, k) \end{pmatrix} + \begin{pmatrix} \gamma \int_{\Lambda_L} dq \mathcal{P}_t(q) - \mathcal{E}[\mathcal{H}_t](\xi, k) \\ 0 \end{pmatrix} \quad (C.3)
\]

where

\[
\mathcal{L}_v = \begin{pmatrix} -\gamma & 2i\omega(k) - v(k) \nabla_\xi \\ 2i\omega(k) - v(k) \nabla_\xi & -\gamma \end{pmatrix}.
\]

Recall that \( \mathcal{H}_t, \mathcal{I}_t, \mathcal{P}_t \) and \( \mathcal{Q}_t \) are \( L \)-periodic in \( \xi \). To solve (C.1), (C.2) and (C.3) we look at the Fourier coefficients of those observables:

\[
\partial_t \hat{\mathcal{H}}_t(n, k) = -2i\pi nL^{-1} v(k) \hat{\mathcal{I}}_t(n, k) + \mathcal{E}[\hat{\mathcal{H}}_t - \hat{\mathcal{P}}_t](n, k) + O(R^{-2}) \quad (C.4)
\]

\[
\partial_t \hat{\mathcal{I}}_t(n, k) = -2i\pi nL^{-1} v(k) \hat{\mathcal{H}}_t(n, k) - \gamma \hat{\mathcal{I}}_t(n, k) + O(R^{-2}) \quad (C.5)
\]

\[
\partial_t \begin{pmatrix} \hat{\mathcal{P}}_t(n, k) \\ \hat{\mathcal{Q}}_t(n, k) \end{pmatrix} = \hat{\mathcal{L}}_v \begin{pmatrix} \hat{\mathcal{P}}_t(n, k) \\ \hat{\mathcal{Q}}_t(n, k) \end{pmatrix} + \begin{pmatrix} \gamma \int_{\Lambda_L} dq \hat{\mathcal{P}}_t(q) - \mathcal{E}[\hat{\mathcal{H}}_t](n, k) \\ 0 \end{pmatrix}, \quad (C.6)
\]

where

\[
\hat{\mathcal{L}}_v = \begin{pmatrix} -\gamma & 2i(\omega(k) - \pi nL^{-1} v(k)) \\ 2i(\omega(k) - \pi nL^{-1} v(k)) & -\gamma \end{pmatrix}
\]

and \( \hat{\mathcal{H}}_t(n, k) = L^{-1} \int_0^L d\xi e^{-2i\pi nL^{-1} \xi} \mathcal{H}_t(\xi, k) \) with \( n \in \mathbb{Z} \) and analogously for \( \hat{\mathcal{I}}_t, \hat{\mathcal{P}}_t \) and \( \hat{\mathcal{Q}}_t \).

Assuming that the time derivative yields a contribution order \( O(R^{-2}) \) we have

\[
\hat{\mathcal{I}}_t(n, k) = -\frac{2i\pi nL^{-1} v(k)}{\gamma} \hat{\mathcal{H}}_t(n, k) + O(R^{-2}), \quad (C.7)
\]

which implies that (C.4) becomes

\[
\mathcal{E}[\hat{\mathcal{H}}_t](n, k) = \mathcal{E}[\hat{\mathcal{P}}_t](n, k) + \gamma^{-1}(2\pi nL^{-1} v(k))^2 \hat{\mathcal{H}}_t(n, k) + O(R^{-2}). \quad (C.8)
\]
Moreover, for (C.6) we have
\[
\begin{pmatrix}
\hat{P}_t(n,k) \\
\hat{Q}_t(n,k)
\end{pmatrix}
= -\hat{L}^{-1}v
\begin{pmatrix}
\gamma\int_{\Lambda^*L} dq \hat{P}_t(n,q) - \bar{C}[\hat{H}_t](n,k)
\end{pmatrix} + O(R^{-2}),
\]
(C.9)
where
\[
\hat{L}^{-1} = \frac{1}{\gamma^2 + 4(\omega(k) - \pi nL^{-1}v(k))} \begin{pmatrix}
\gamma & 2i(\omega(k) - \pi nL^{-1}v(k)) \\
2i(\omega(k) - \pi nL^{-1}v(k)) & \gamma
\end{pmatrix}.
\]
Combining (C.8) and (C.9) we get
\[
\hat{P}_t(n,k) = -\frac{(\pi nL^{-1}v(k))}{(\omega(k) - \pi nL^{-1}v(k))} \hat{H}_t(n,k).
\]
(C.10)
By the definition of the test function \(\varphi\) given in (4.4), \(\hat{H}_t(n,k)\) is concentrated on values of \(n\) such that \(n/L\) is of order \(O(R^{-1})\). In fact, from the definition of \(\mathcal{H}_t(\xi, \eta)\) we get the explicit form of \(\hat{H}_t(n,k)\):
\[
\hat{H}_t(n,k) = \frac{1}{L} \int_0^L d\xi e^{-2\pi i L^{-1}n\xi} \sum_{x \in \Lambda_L} \varphi(\xi - x)V_t(x, k)
\]
(C.11)
where
\[
V_t(x, k) = \sum_{y \in \Lambda_L} e^{-2\pi i y \cdot k} E[\psi_t(x, -1)\psi_t(x + y, +1) + \psi_t(x, +1)\psi_t(x + y, -1)].
\]
Thanks to (B.7), (C.11) becomes
\[
\hat{H}_t(n,k) = \frac{1}{L} \int_{\Lambda^*_L} dk' \hat{V}_t(k', k)\hat{\phi}(Rk') \int_0^L d\xi e^{-2\pi i \xi \cdot (nL^{-1} - k')}
\]
\[
= \hat{V}_t(nL^{-1}, k)\hat{\phi}(RnL^{-1})
\]
(C.12)
where \(\hat{V}_t(k', k) := \sum_{x \in \Lambda_L} e^{-2\pi i x \cdot k'} V_t(x, k)\) and we used the fact that
\[
\frac{1}{L} \int_0^L d\xi e^{-2\pi i \xi \cdot (nL^{-1} - k')} = \mathbb{1}(k' = nL^{-1}) \text{ for any } k' \in \Lambda^*_L.
\]
Therefore, \(nL^{-1} \in \Lambda^*_L\) and, since \(\hat{\phi}\) has compact support (see assumption (2) in Appendix B), we get that \(\hat{\phi}(RnL^{-1}) = 0\) whenever \(|nL^{-1}| \leq \rho \lesssim R^{-1}\), from which the claim follows.

The fact that \(\hat{H}_t(n,k)\) vanishes for \(|nL^{-1}| \geq O(R^{-1})\) indicates that \(\mathcal{P}_t(\xi, k) = O(R^{-2})\). Then clearly \(\mathcal{Q}_t(\xi, k) = O(R^{-2})\) and \(\mathcal{E}[\mathcal{H}_t](\xi, k) = O(R^{-2})\), thus implying also \(\mathcal{H}_t = E_t + O(R^{-2})\).

**Acknowledgements**

We thank Giada Basile, Mario Pulvirenti, and Herbert Spohn for useful discussions on the topic, and the anonymous reviewers for their suggestions for improvements. The work has been supported by the Academy of Finland via the Centre of Excellence in Analysis and Dynamics Research (project 271983) and from an Academy Project (project 258302), and partially by the French Ministry of Education through the grant ANR (EDNHS). We are also grateful to the Erwin Schrödinger Institute (ESI), Vienna, Austria for organization of a workshop and providing an opportunity for the related discussions. Alessia Nota acknowledges support also through the CRC 1060 *The mathematics of emergent effects* at the University of Bonn, that is funded through the German Science Foundation (DFG).
References


Derivation of the linear Landau equation and linear Boltzmann equation from the Lorentz model with magnetic field

Matteo Marcozzi, Alessia Nota
Derivation of the Linear Landau Equation and Linear Boltzmann Equation from the Lorentz Model with Magnetic Field

M. Marcozzi\(^1\) · A. Nota\(^2\)

Received: 4 September 2015 / Accepted: 8 January 2016 / Published online: 1 February 2016

Abstract We consider a test particle moving in a random distribution of obstacles in the plane, under the action of a uniform magnetic field, orthogonal to the plane. We show that, in a weak coupling limit, the particle distribution behaves according to the linear Landau equation with a magnetic transport term. Moreover, we show that, in a low density regime, when each obstacle generates an inverse power law potential, the particle distribution behaves according to the linear Boltzmann equation with a magnetic transport term. We provide an explicit control of the error in the kinetic limit by estimating the contributions of the configurations which prevent the Markovianity. We compare these results with those ones obtained for a system of hard disks in Bobylev et al. (Phys Rev Lett 75:2, 1995), which show instead that the memory effects are not negligible in the Boltzmann-Grad limit.

Keywords Lorentz gas · Magnetic field · Linear Boltzmann equation · Linear Landau equation · Low density limit · Weak coupling limit

1 Introduction

Consider a point particle of mass \(m = 1\) in \(\mathbb{R}^d\), \(d = 2, 3\) moving in a random distribution of fixed scatterers, whose centers are denoted by \((c_1, \ldots, c_N)\).

We assume that the scatterers are distributed according to a Poisson distribution of parameter \(\mu > 0\). The equations of motion are

\[
\begin{align*}
\dot{x} &= v \\
\dot{v} &= -\sum_i \nabla \phi(|x - c_i|),
\end{align*}
\]

\[\text{(1.1)}\]
here \((x, v)\) denote position and velocity of the test particle, \(t\) the time and \(\dot{A} = \frac{dA}{dt}\) for any time dependent variable \(A\).

To outline a kinetic behavior it is usually introduced a scaling of the space-time variables and the density of the scatterer distribution. For this model, it is more physically intuitive to transfer the scaling to the background medium. More precisely, let \(\epsilon > 0\) be a parameter indicating the ratio between the macroscopic and microscopic variables, we keep time and space fixed and rescale the range of the interaction and the density of the scatterers, i.e.

\[
\phi_\epsilon(x) = \epsilon^\alpha \phi\left(\frac{x}{\epsilon}\right)
\]

\[
\mu_\epsilon = \mu \epsilon^{-(d-1+2\alpha)}
\]

(1.2)

where \(d = 2, 3\) is the dimension of the physical space and \(\alpha \in [0, \frac{1}{2}]\) is a suitable parameter. This means that the probability of finding \(N\) obstacles in a bounded measurable set \(\Lambda \subset \mathbb{R}^d\) is given by

\[
P_\epsilon(d c_N) = e^{-\mu_\epsilon \mu \epsilon N} dc_1 \cdots dc_N
\]

(1.3)

where \(c_N = c_1, \ldots, c_N\) and \(\mu \epsilon = \mu \epsilon N\). Consequently, the equation of motion (1.1) becomes

\[
\begin{aligned}
\dot{x} &= v \\
\dot{v} &= -\epsilon^{\alpha-1} \sum_i \nabla \phi\left(\frac{|x-c_i|}{\epsilon}\right).
\end{aligned}
\]

(1.4)

Now let \(T_{c_N}^t(x, v)\) be the Hamiltonian flow solution to Eq. (1.4) with initial datum \((x, v)\) in a given sample of obstacles (skipping the \(\epsilon\) dependence for notational simplicity) and, for a given probability distribution \(f_0 = f_0(x, v)\), consider the quantity

\[
f_\epsilon(x, v, t) = E_\epsilon[f_0(T_{c_N}^{-t}(x, v))]
\]

(1.5)

where \(E_\epsilon\) is the expectation with respect to the measure \(P_\epsilon\) given by (1.3).

In the limit \(\epsilon \to 0\) we expect that the probability distribution (1.5) will solve a linear kinetic equation depending on the value of \(\alpha\). If \(\alpha = 0\) the limit corresponding to such a scaling is called low-density (or Boltzmann-Grad) limit. In this case \(f_\epsilon\) converges to the solution of a linear Boltzmann equation. See [1, 8, 12, 17]. On the other hand, if \(\alpha = \frac{1}{2}\) the corresponding limit, called weak-coupling limit, yields the linear Landau equation, as proven in [7, 13, 14]. The intermediate scaling, namely \(\alpha \in (0, \frac{1}{2})\), although refers to a low-density regime, leads to the linear Landau equation again, see [9, 13].

We want to remark that in [8, 9] the authors exploit the original constructive idea due to Gallavotti (see [12]) for the Boltzmann-Grad limit. This method is based on a suitable change of variables which can be implemented outside a set \(E\) of pathological events which prevent the Markov property of the limit (such as the set of configurations yielding recollisions, i.e. when the test particle recollides with a given obstacle after having suffered collisions with other different obstacles). The probability \(P_\epsilon(E)\) is vanishing as \(\epsilon\) tends to 0. The main difference is that in [8] the range of the potential is infinite in the limit, therefore the test particle interacts with infinitely many obstacles. As for the case of the long range potential considered in [8] also in [9] there is a lack of the semi explicit form of the solution of the limit equation. This requires explicit estimates for the set of bad configurations of obstacles. For a short range potential, like in the case of the hard-sphere potential considered in [12], a simple dimensional argument is sufficient. For an explicit control of the error in the kinetic limit for the hard-sphere potential see for instance [6]. Moreover, in [9, 13] it was proven that
even if $\alpha > 0$, but sufficiently small, the recollisions are still negligible. Incidentally we note that, if $\alpha$ is close to $1/2$, this is not true anymore and it would be interesting to derive the Landau equation in this regime, by means of an explicit constructive approach.

Furthermore it has been observed that the presence of a given external field, in the two dimensional Lorentz model, strongly affects the derivation of the linear Boltzmann equation in the Boltzmann-Grad limit. Bobylev et al., in [2] and later in [3,4] (see also [15] for further readings), showed that the set of pathological configurations is no longer negligible when the test particle moves in a plane with a Poisson distribution of hard disks and a uniform and constant magnetic field perpendicular to the plane. See Fig. 1 for a pictorial representation of the light particle’s motion.

The following simple computation turns out to give a good heuristic argument explaining these results: consider the probability $\mathbb{P}_{R_L}$ of performing an entire Larmor circle without hitting any obstacle, $R_L$ being the Larmor radius. From Eq. (1.3) one easily gets

$$\mathbb{P}_{R_L} \approx e^{-\mu \varepsilon \text{Area}(A_\varepsilon)} \approx e^{-2\pi R_L \mu},$$

where $A_\varepsilon(R_L)$ is the annulus of radius $R_L$ and width $\varepsilon$. Hence, $\mathbb{P}_{R_L}$ is not vanishing in the limit $\varepsilon \to 0$ and the Markovianity of the limit system can not be attained. In fact, in [3,4], a kinetic equation with memory is derived, i.e. a generalized Boltzmann equation, taking into account those effects:

$$\frac{D}{Dt} f^G(x, v, t) = \mu \varepsilon \sum_{k=0}^{[t/T_L]} e^{-\nu k T_L} \int_{S_k} dn (v \cdot n)$$

$$\times [\chi(v \cdot n)b_n + \chi(-v \cdot n)] f^G(x, S_0^{-k}v, t - k T_L), \quad (1.6)$$

where $f(x, v, t)$ is the probability density of finding the moving particle at time $t$ at position $x$ with velocity $v$ and

$$f^G(x, v, t) = \begin{cases} f(x, v, t) & \text{if } 0 < t < T_L \\ (1 - e^{-v T_L}) f(x, v, t) & \text{if } t > T_L. \end{cases} \quad (1.7)$$
Here $\nu = 2|v|\mu_\varepsilon \varepsilon$ is the collision frequency and $T_L = 2\pi/\Omega$ is the cyclotron period where $\Omega = qB/m$ is the frequency, being $q$ the charge and $m$ the mass. Furthermore, note that

$$\frac{D}{Dt} = (\partial_t + v \cdot \nabla_x + (v \times B) \cdot \nabla_v)$$

is the generator of the free cyclotron motion with frequency $\Omega$ and $\lfloor t/T_L \rfloor$ the number of cyclotron periods $T_L$ completed before time $t$. The angular integration over the unit vector $n$ in (1.6) is over the entire unit sphere $S_1$ centered at the origin. In the gain term the operator $b_n$ is defined by

$$b_n\phi(v) = \phi(v - 2(v \cdot n)n)$$

where $\phi(v)$ is an arbitrary function of $v$. The precollisional velocity $v' = v - 2(v \cdot n)n$ becomes $v$ after the elastic collision with the hard disk. Note that $v' \cdot n < 0$. In the loss term, the precollisional velocity $v$ is also from the hemisphere $v \cdot n < 0$. Finally, the shift operator $S_{0}^{k\theta}$, when acting on $v$, rotates the velocity through the angle $-k\theta$, where $\theta$ is the scattering angle (from $v'$ to $v$).

For further readings in this direction we refer to [10,11], where the authors consider a stochastic Lorentz model with a smooth external force field $F(x,t)$ and with absorbing obstacles, i.e. the interaction between the obstacles and the test particle is such that the test particle disappears whenever it enters an obstacle. It is proved that the kinetic equation associated to this model in the Boltzmann-Grad limit is non-Markovian and that the Markovianity can be recovered by introducing an additional stochasticity in the velocity distribution of the obstacles.

In this paper we consider the case of a random distribution of scatterers in $\mathbb{R}^2$ where each obstacle generates a smooth positive and short-range potential $\phi$, with $\alpha > 0$ and sufficiently small. We show that, in this case, the solution of the microscopic dynamics converges, in the intermediate limit (when $\alpha \in (0, 1/8)$), to the solution of the linear Landau equation with an additional transport term due to the magnetic field. From the heuristic point of view, this result is suggested by the observation that in this case the probability $P_{RL}$ of performing an entire Larmor circle without hitting any obstacle is given by

$$P_{RL} \approx e^{-\mu_\varepsilon 2\pi RL} \approx e^{-2\pi RL \mu_\varepsilon^{-2\alpha}}$$

which vanishes as $\varepsilon \to 0$. This computation shows that one family of the pathological events preventing the Markovianity is negligible in this setting. We stress that this rough argument is not sufficient to conclude that we can recover the Markovianity in the limit. Indeed, to prove this, we need to show that all the other bad configurations of obstacles defining the set $E$ are negligible in the limit, as we will see in Sect. 4.1.

Furthermore, we observe that even if we consider a long range inverse power law interaction potential, truncated at distance $\varepsilon^{-\gamma}$ with $\gamma \in (0, 1)$ suitably large, in the low density regime $\alpha = 0$, we can prove that the memory is lost in the limit. More precisely, we prove that the microscopic solution converges to the solution of the uncutoffed linear Boltzmann equation with a magnetic transport term. With the same purpose of the rough argument presented above, we observe that the probability $P_{RL}$ of performing a complete Larmor circle without hitting any obstacle is approximatively given by

$$P_{RL} \approx e^{-2\pi RL \mu_\varepsilon \varepsilon^{-\gamma-1}}$$

which vanishes as $\varepsilon \to 0$ when $\gamma < 1$. Also in this case this represents only one example of bad configuration of scatterers. It is essential to prove that the contribution of the whole set
of pathological events is negligible in the limit, as we will show in Sect. 5. Moreover, from the technical point of view, we observe that the parameter \( \gamma \) has to be chosen close to 1 as dictated by the explicit control of the memory effects.

Thus, as we pointed out with the heuristic motivations above, the non-Markovian behaviour of the limit process, discussed in [2], disappears as soon as we slightly modify the microscopic model given by the two-dimensional Lorentz Gas.

The purpose of this paper is to provide a rigorous validation of the linear Landau equation and the linear Boltzmann equation respectively with magnetic field by using the constructive strategy due to Gallavotti. We remark that, as in [5, 6, 8, 9], we need explicit estimates of the error in the kinetic limit and this is the crucial part. Moreover, as a future target, it could be interesting to understand if a rigorous derivation of the generalized Boltzmann equation proposed in [2] can be achieved by using the same constructive techniques.

The plan of the paper is the following: in the next Section we establish the model and formulate the results; in Sect. 3 we present the strategy of the proofs, whereas Sects. 4 and 5 are dedicated to the nontrivial analysis and explicit estimates of the sets of bad configurations producing memory effects, which is the technical core of this paper.

2 The Model and Main Results

2.1 The Lorentz Model with Short Range Interactions

We consider the system (1.4) in the plane \((d = 2)\) under the action of a uniform, constant, magnetic field orthogonal to the plane. The equations of motion are

\[
\begin{align*}
\dot{x} &= v \\
\dot{v} &= Bv^\perp - \varepsilon^{-1} \sum_i \nabla \phi\left(\frac{|x - c_i|}{\varepsilon}\right),
\end{align*}
\]

where \(B\) is the magnitude of the magnetic field and \(v^\perp = (v_2, -v_1)\). We assume that the potential \(\phi : \mathbb{R}^+ \to \mathbb{R}^+\) is smooth and of range 1 i.e. \(\phi(r) = 0\) if \(r > 1\). Therefore the particle is influenced by the scatterer \(c_i\) if \(|x - c_i| < \varepsilon\).

Starting from the initial position \(x\) with initial velocity \(v\), the particle moves under the action of the Lorentz force \(Bv^\perp\). Suppose that the particle has unitary mass and unitary charge, namely \(m, q = 1\), hence between two consecutive scatterers, the particle moves with constant angular velocity \(\Omega = qB/m = B\) and performs an arc of circle of radius \(R_L = |v|/B\). \(R_L\) is the Larmor radius, i.e. the radius of the cyclotron orbit whose center is situated at the point

\[
x_c = x + \frac{\mathcal{R}(\frac{\pi}{2})}{\Omega} \cdot v,
\]

where the tensor \(\mathcal{R}(\varphi)\) denotes the rotation of angle \(\varphi\). Without loss of generality we assume from now on that \(|v| = 1\), therefore \(R_L = 1/B\). Moreover, we will denote by \(S_1\) the kinetic energy sphere with unitary radius.

The precise assumptions on the potential are the following:

A1) \(\phi \in C^2([0, \infty))\);
A2) \(\phi \geq 0, \phi' \leq 0\) in \((0, 1)\);
A3) \(\text{supp} \phi \subset [0, 1]\).
On $f_0$ we assume that

A4) $f_0 \in C_0(\mathbb{R}^2 \times \mathbb{R}^2)$ is a continuous, compactly supported initial probability density. Suppose also that $|D^k_x f_0| \leq C$, where $D_x$ is any partial derivative with respect to $x$ and $k = 1, 2$.

Moreover, we assume that

A5) The scatterers are distributed according to a Poisson distribution (1.3) of intensity $\mu_\varepsilon = \mu \varepsilon^{-\delta}$ with $\delta = 1 + 2\alpha$, $\alpha \in (0, \frac{1}{8})$.

Next we define the Hamiltonian flow $T^t_{c_N}(x, v)$ associated to the initial datum $(x, v)$, solution of (2.1) for a given configuration $c_N$ of scatterers, and we set

$$f_\varepsilon(x, v, t) = E_\varepsilon[f_0(T_{c_N}^{-t}(x, v))]$$

where $E_\varepsilon$ denotes the expectation with respect to the Poisson distribution.

The first result of the present paper is summarized in the following theorem.

**Theorem 2.1** Let $f_\varepsilon$ be defined in (2.2). Under assumption A1)–A5), for all $t \in [0, T]$, $T > 0$,

$$\lim_{\varepsilon \to 0} f_\varepsilon(\cdot; t) = g(\cdot; t)$$

where $g$ is the unique solution to the Landau equation with magnetic field

$$\begin{cases}
(\partial_t + v \cdot \nabla_x + B v^\perp \cdot \nabla_v) g(x, v, t) = \xi \Delta_{S_1} g(x, v, t) \\
g(x, v, 0) = f_0(x, v),
\end{cases}$$

(2.3)

where $\Delta_{S_1}$ is the Laplace-Beltrami operator on the circle $S_1$ and $\xi > 0$. The convergence is in $L^2(\mathbb{R}^2 \times S_1)$.

The constant $\xi$ is the diffusion coefficient and its explicit expression will be given below in (3.9) and in Remark 3.4.

### 2.2 The Lorentz Model with Long Range Interactions

We consider now the case in which each obstacle generates a potential of the form

$$\tilde{\psi}_\varepsilon(|x - c|) = \psi_{\varepsilon}\left(\frac{|x - c|}{\varepsilon}\right)$$

where the unrescaled potential $\psi_{\varepsilon}$ is an inverse power law potential truncated at large distances. More precisely we assume the following:

B1) $\psi_\varepsilon(x) = \begin{cases}
\frac{1}{|x|^s} & |x| < \varepsilon^{\gamma - 1} \\
\varepsilon^{-s(\gamma - 1)} & |x| \geq \varepsilon^{\gamma - 1}
\end{cases}$ with $\gamma \in (0, 1)$ and $s > 2$.

We point out that it could be challenging to consider directly the untruncated long range potential $\psi(|x|) = |x|^{-s}$. In fact, this problem presents deep additional difficulties as noted in Remark 2.3 in [8] and new ideas and techniques are necessary.

Moreover, we assume that

B2) The scatterers are distributed according to a Poisson law (1.3) of intensity $\mu_\varepsilon = \varepsilon^{-1} \mu$, $\mu > 0$. 

\(\mathord{\%} \) Springer
The equation of motion in macroscopic variables reads
\[
\begin{align*}
\dot{x} &= v \\
\dot{v} &= B v^\perp - \varepsilon^{-1} \sum_i \nabla \psi_e(\frac{|x-c_i|}{\varepsilon})
\end{align*}
\] (2.4)
with \(\psi_e\) given in Assumption B1).

Let \(T_{\varepsilon N}^t(x,v)\) be the Hamiltonian flow solution to Eq. (2.4) with initial datum \((x,v)\) in a given sample of obstacles. Let be \(f_0 = f_0(x,v)\) be the initial probability distribution. On \(f_0\) we assume
\[
B3) \quad f_0 \in L^1 \cap W^{1,\infty}(\mathbb{R}^2 \times \mathbb{R}^2), \quad f_0 \geq 0, \quad \int f_0 \, dx \, dv = 1.
\]

We consider the quantity
\[
f_\varepsilon(x,v,t) = E_\varepsilon[f_0(T_{\varepsilon N}^{-t}(x,v))]
\] (2.5)
where \(E_\varepsilon\) is the expectation with respect to the measure \(\mathbb{P}_\varepsilon\) given by (1.3). The second result of this paper is summarized in the following theorem.

**Theorem 2.2** Let \(f_\varepsilon\) be defined in (2.5). Under assumption B1) – B3) with \(\gamma \in (6/7, 1)\), for all \(t \in [0, T]\), \(T > 0\),
\[
\lim_{\varepsilon \to 0} f_\varepsilon(\cdot; t) = f(\cdot; t)
\]
where \(f\) is the unique solution to the linear Boltzmann equation with magnetic field
\[
\begin{align*}
\begin{cases}
(\partial_t + v \cdot \nabla_x + B v^\perp \cdot \nabla_v) f(t, x, v) = L_\varepsilon f(t, x, v) \\
f(x, v, 0) = f_0(x, v)
\end{cases}
\end{align*}
\] (2.6)
with
\[
L_\varepsilon f(v) = \mu \int_{\varepsilon}^{\varepsilon} d\rho [f(\mathcal{R}(\Theta) v) - f(v)]
\]
\(\Gamma(\Theta)\) is the differential cross section associated to the long range potential \(\psi(|x|) = |x|^{-s}\) and the operator \(\mathcal{R}(\Theta)\) rotates the velocity \(v\) by the angle \(\Theta\). The convergence is in \(D'(\mathbb{R}^2 \times S_1)\).

3 Proofs

3.1 Proof of Theorem 2.1

Following [5, 8, 9] we split the original problem into two parts. The first one concerns the asymptotic equivalence between \(f_\varepsilon\) defined in (1.5) and \(h_\varepsilon\), solution of the following Boltzmann equation
\[
\begin{align*}
\begin{cases}
(\partial_t + v \cdot \nabla_x + B v^\perp \cdot \nabla_v) h_\varepsilon(x, v, t) = L_\varepsilon h_\varepsilon(x, v, t) \\
h_\varepsilon(x, v, 0) = f_0(x, v)
\end{cases}
\end{align*}
\] (3.1)
where
\[
L_\varepsilon h_\varepsilon(v) = \mu \int_{-\varepsilon}^{\varepsilon} d\rho [h_\varepsilon(v') - h_\varepsilon(v)].
\] (3.2)
Here \( v' = v - 2(\omega \cdot v)\omega \) is the outgoing velocity after a scattering with incoming velocity \( v \) and impact parameter \( \rho \in [-\varepsilon, \varepsilon] \) generated by the potential \( \varepsilon^\alpha \phi(\frac{x}{\varepsilon}) \). Moreover, \( \omega = \omega(\rho) \) is the versor bisecting the angle between the incoming and outgoing velocity and \( \theta_{\varepsilon} \) is the scattering angle. The precise result is stated in the following proposition.

**Proposition 3.1** Under assumption A1) – A5), for any \( T > 0 \),

\[
\lim_{\varepsilon \to 0} \| f_\varepsilon - h_\varepsilon \|_{L^\infty([0,T]; L^1(\mathbb{R}^2 \times S_1))} = 0
\]  

(3.3)

where \( h_\varepsilon \) solves (3.1).

The proof of the above Proposition is postponed to Sect. 4.

The second step concerns the grazing collision limit. Note that the presence of the magnetic field does not affect the last step. More precisely we have the following

**Lemma 3.2** The deflection angle \( \theta_{\varepsilon}(\rho) \) of a particle colliding with impact parameter \( \rho \) with a scatterer generating a radial potential \( \varepsilon^\alpha \phi \) under the action of the Lorentz force \( B v^\perp \) satisfies

\[
|\theta_{\varepsilon}(\rho)| \leq C \varepsilon^\alpha.
\]  

(3.4)

**Proof** As established in [9] (Sect. 3), the estimate (3.4) holds when the test particle scatters with no external field. Hence, we just need to compare the dynamics of the test particle in presence of the constant magnetic field with the free dynamics. Let \((x(t), v(t))\) be the solution of the following

\[
\begin{aligned}
\dot{x} &= v \\
\dot{v} &= -\varepsilon^{\alpha - 1} \nabla \phi(\frac{|x| - \varepsilon}{\varepsilon}).
\end{aligned}
\]  

(3.5)

Let \( \tau \) be the collision time for the dynamics described by (2.1). The key observation is that the presence of the magnetic field does not modify the estimate for the collision time related to the dynamics in (3.5). Indeed also in our case \( \tau \leq C \varepsilon, \, C > 0 \), as in [9], see Appendix 1 for the detailed computations. Therefore

\[
|v(\tau) - \upsilon(\tau)| = \left| \varepsilon^{\alpha - 1} \int_0^\tau ds \left( F(x(s)/\varepsilon) - F(x(s)/\varepsilon) \right) + \int_0^\tau ds \, v^\perp B \right|
\leq \varepsilon^{\alpha - 1} \int_0^\tau ds \left| F(x(s)/\varepsilon) - F(x(s)/\varepsilon) \right| + C_1 \varepsilon
\leq \varepsilon^{\alpha - 2} C_2 \int_0^\tau ds \left| x(s) - \chi(s) \right| + C_1 \varepsilon
\leq \varepsilon^{\alpha - 2} C_2 \int_0^\tau ds \int_0^s dt \left| v(t) - \upsilon(t) \right| + C_1 \varepsilon
\leq \varepsilon^{\alpha - 2} C_2 \int_0^\tau ds \int_0^\tau dt \left| v(t) - \upsilon(t) \right| + C_1 \varepsilon,
\]

where \( F(x) := - (\nabla \phi)(x) \). By using Grönwall’s inequality we obtain

\[
|v(\tau) - \upsilon(\tau)| \leq C_1 \varepsilon \varepsilon e^{C_3 \varepsilon^{\alpha - 1} \tau} \leq C_1 \varepsilon \varepsilon e^{C_3 \varepsilon^\alpha}
\]  

(3.6)

for \( \alpha > 0 \) and \( \varepsilon \) sufficiently small. Hence, the velocities \( v \) and \( \upsilon \) are asymptotically equivalent up to an error term of order \( \varepsilon \). We now define \( v' \) and \( \upsilon' \) to be the outgoing velocities with and without magnetic field respectively, \( v \) the incoming velocity. By using (3.6) we have
\[ \sqrt{2(1 - \cos \theta_e)} = |v' - v| \leq |v' - \tilde{v}'| + |\tilde{v}' - v| \leq C \varepsilon + \sqrt{2(1 - \cos \tilde{\theta}_e)} \]  

(3.7)

where \( \tilde{\theta}_e \) is the scattering angle without magnetic field. From [9] we know that \( \tilde{\theta}_e \leq C' \varepsilon^\alpha \), so from (3.7) we obtain

\[ |\sin \frac{\theta_e}{2}| \leq C \varepsilon + |\sin \frac{\tilde{\theta}_e}{2}| \leq C \varepsilon + C' \varepsilon^\alpha \leq C'' \varepsilon^\alpha. \]

Since \( \theta_e \) is continuous as a function of the impact parameter \( \rho \), it results \( \theta_e \leq C \varepsilon^\alpha \). For further details see Proposition B.1 in Appendix 2.

The following proposition shows the asymptotic equivalence between the solution of Landau equation and the solution of the previous Boltzmann equation \( h^\varepsilon \).

**Proposition 3.3** Under the assumptions A1) − A5), \( h^\varepsilon \rightarrow g \) in \( L^\infty([0, T]; L^2(\mathbb{R}^2 \times S_1)) \) where \( g \) is the unique solution to the Landau equation with magnetic field

\[ \begin{cases} 
(\partial_t + v \cdot \nabla_x + B v^\perp \cdot \nabla_v) g(x, v, t) = \xi \Delta_{S_1} g(x, v, t) \\
 g(x, v, 0) = f_0(x, v),
\end{cases} \]

(3.8)

where

\[ \xi = \lim_{\varepsilon \to 0} \frac{\mu \varepsilon^{-2\alpha}}{2} \int_{-1}^{1} \theta_e^2(\rho) \, d\rho \]

(3.9)

is the diffusion coefficient.

**Remark 3.4** As shown in Appendix 2, \( \theta_e = \tilde{\theta}_e + O(\varepsilon) \), where \( \tilde{\theta}_e \) is the scattering angle without any magnetic field, i.e. the scattering angle studied in [9]. This implies that the explicit expression for the diffusion coefficient obtained in [9] still holds in our case:

\[ \xi = \frac{\mu}{2} \int_{-1}^{1} \left( \int_{\rho}^{1} \rho \phi^\prime(\rho) \, d\rho \sqrt{1 - u^2} \right)^2 \, d\rho, \]

(3.10)

where the integrand is an even function of the impact parameter \( \rho \).

**Remark 3.5** The linear Landau equation (3.8) propagates the regularity of the derivatives with respect to the \( x \) variable thanks to the transport operator. Moreover, the presence of the collision operator \( L := \Delta_{S_1} \) lets the solution gain regularity with respect to the transverse component of the velocity. Indeed, under the assumption A4) on \( f_0 \), the solution \( g : \mathbb{R}^2 \times S_1 \to \mathbb{R}^+ \) satisfies the bounds

\[ |D^k_x g| \leq C, \quad |D^h_v g(x, v)| \leq C \quad \forall k \leq 2, \ h \geq 0, \]

(3.11)

\( \forall t \in (0, T] \), where \( C = C(f_0, T) \) and \( D_v \) is the derivative with respect to the transverse component of the velocity. In particular, the solutions of (3.8) we are considering are classical.

**Proof** By using the invariance of the scattering angle with respect to the space scale, we rewrite the collision operator in the right hand side of (3.1) as

\[ L^\varepsilon h^\varepsilon(v) = \mu \varepsilon \int_{-1}^{1} d\rho [h^\varepsilon(v') - h^\varepsilon(v)]. \]

(3.12)
We look at the evolution of \( h_\varepsilon - g \), being \( g \) the solution of (3.8), namely
\[
(\partial_t + v \cdot \nabla_x + B v^\perp \cdot \nabla_v)(h_\varepsilon - g) = \left( L_\varepsilon h_\varepsilon - L g \right),
\]
where \( L := \xi \Delta_{S_1} \).

Note that \( g \in L^2(\mathbb{R}^2 \times S_1) \) because \( f_0 \in L^2(\mathbb{R}^2 \times \mathbb{R}^2) \) and \( h_\varepsilon \in L^2(\mathbb{R}^2 \times S_1) \). Indeed, from Proposition 3.1, we know that \( h_\varepsilon \in L^1(\mathbb{R}^2 \times S_1) \) but the hypothesis on the initial state implies that \( h_\varepsilon \in L^2(\mathbb{R}^2 \times S_1) \).

We now consider the scalar product of equation (3.13) with \((h_\varepsilon - g)\) in \( L^2(\mathbb{R}^2 \times S_1) \) and we obtain
\[
\frac{1}{2} \partial_t \| h_\varepsilon - g \|_2^2 = -(h_\varepsilon - g, -L_\varepsilon [h_\varepsilon - g]) + (h_\varepsilon - g, [L_\varepsilon - L] g).
\]

By exploiting the positivity of \(-L_\varepsilon\) and the Cauchy-Schwartz inequality we get
\[
\partial_t \| h_\varepsilon - g \|_2 \leq \| (L_\varepsilon - L) g \|_2.
\]

We now set
\[
g(v') - g(v) = (v' - v) \cdot \nabla_{|S_1} g(v)
+ \frac{1}{2} (v' - v) \otimes (v' - v) \nabla_{|S_1} \nabla_{|S_1} g(v)
+ \frac{1}{6} (v' - v) \otimes (v' - v) \otimes (v' - v) \nabla_{|S_1} \nabla_{|S_1} \nabla_{|S_1} g(v) + R_\varepsilon,
\]
with \( R_\varepsilon = O(|v - v'|^4) \). Integrating with respect to \( \rho \) and using symmetry arguments we obtain
\[
L_\varepsilon g = \mu \varepsilon^{-2\alpha} \left\{ \frac{1}{2} \Delta_{S_1} g \int_{-1}^{1} d\rho \ |v' - v|^2 + \int_{-1}^{1} d\rho \ R_\varepsilon \right\}.
\]

Observe that \( |v' - v|^2 = 4 \sin^2 \frac{\theta_\varepsilon(\rho)}{2} \), then by direct computation
\[
\lim_{\varepsilon \to 0} \frac{\mu \varepsilon^{-2\alpha}}{2} \int_{-1}^{1} d\rho \ |v' - v|^2 = \lim_{\varepsilon \to 0} \frac{\mu \varepsilon^{-2\alpha}}{2} \int_{-1}^{1} \theta_\varepsilon^2(\rho) \ d\rho =: \xi.
\]
Therefore, thanks to Lemma 3.2, we have
\[
\| (L_\varepsilon - L) g \|_{L^2} \leq \varepsilon^{2\alpha} \| \Delta_{S_1}^2 g \|_{L^2} \leq \varepsilon^{2\alpha} C,
\]
which vanishes for \( \varepsilon \to 0 \).

\[\square\]

**Remark 3.6** We avoided introducing the cross-section \( \Gamma(\theta_\varepsilon) := \frac{d\rho}{d\theta_\varepsilon} \) of the problem because the map \( \rho \to \theta_\varepsilon(\rho) \) is not monotonic in general.

Indeed if \( \phi \) is bounded and \( \varepsilon \) sufficiently small, \( \frac{1}{2} v^2 > \varepsilon^\alpha \phi(0) \) so that \( \theta = 0 \) for \( \rho = 0 \) and \( \rho = \pm 1 \). As a consequence, \( \Gamma(\theta_\varepsilon) \) is neither single valued nor bounded.

### 3.2 Proof of Theorem 2.2

The general structure of the proof follows the lines of [8] where an analogous result has been proven when the magnetic field is zero.
Proposition 3.7 Let $f_\varepsilon$ be defined in (2.5). Then, for any $T > 0,$
\[
\lim_{\varepsilon \to 0} \| f_\varepsilon - h_{E,Y} \|_{L^\infty([0,T];L^1(\mathbb{R}^2 \times S_1))} = 0
\]  
where $h_{E,Y}$ is the unique solution of the truncated linear Boltzmann equation with magnetic field
\[
\begin{cases}
(\partial_t + v \cdot \nabla_x + B v^\perp \cdot \nabla_v) h_{E,Y}(t, x, v) = L h_{E,Y}(t, x, v) \\
h_{E,Y}(x, v, 0) = f_0(x, v),
\end{cases}
\]  
with
\[
L f(v) = \mu \int_{-\pi}^{\pi} \Gamma^{(B)}_{E,Y}(\theta) \left( f(\mathcal{R}(\theta)v) - f(v) \right) d\theta.
\]  
and $\Gamma^{(B)}_{E,Y}$ is the differential cross section associated to the unrescaled potential $\psi_\varepsilon$ with magnetic field.

The proof of Proposition 3.7 is in Sect. 5.

This allows to reduce the problem of the transition from the solution of the truncated linear Boltzmann equation to the solution of the untruncated linear Boltzmann equation to a partial differential equation problem. Indeed, as in [8], we can prove the following

Proposition 3.8 Let $h_{E,Y}$ solution of (3.15). Then, for any $T > 0,$
\[
h_{E,Y} \to f \quad \text{in} \quad C([0,T]; \mathcal{D}')
\]  
where $f$ is the unique solution of (2.6).

Proof In Appendix 2, Proposition B.2, the cross section $\Gamma^{(B)}_{E,Y}(\theta)$ is shown to be bounded by $C \theta^{-1-1/s}$ and to converge to $\Gamma(\theta)$ almost everywhere as $\varepsilon \to 0,$ where $\Gamma(\theta)$ is the cross section associated to the truly long range potential $\Psi(r) = r^{-s}$ without magnetic field. Therefore, the proof of Proposition 3.8 is exactly the same as the one of Proposition A.2 in [8].

\[\square\]

4 Proof of Proposition 3.1

In this section we prove the asymptotic equivalence of $f_\varepsilon,$ defined by (2.2), and $h_\varepsilon,$ solution of the linear Boltzmann equation (3.1), that we recall here for the sake of clarity
\[
(\partial_t + v \cdot \nabla_x + B v^\perp \cdot \nabla_v) h_\varepsilon(x, v, t) = L_\varepsilon h_\varepsilon(x, v, t),
\]  
where
\[
L_\varepsilon h_\varepsilon(v) = \mu \varepsilon^{-2a} \int_{-1}^{1} d\rho [h_\varepsilon(v') - h_\varepsilon(v)].
\]  

This allows to reduce the problem to the analysis of a Markov process which is an easier task. Indeed, the series expansion defining $h_\varepsilon$ (obtained perturbing around the loss term) reads as
\[
h_\varepsilon(x, v, t) = e^{-2 \varepsilon^{-2a} \mu t} \sum_{Q \geq 0} \mu Q e^{-Q t} \cdot \int_0^t dt_2 \cdots \int_0^{t_{Q-1}} dt_1 \int_{-\varepsilon}^\varepsilon d\rho_1 \cdots \int_{-\varepsilon}^\varepsilon d\rho_Q f_0(\gamma^{-1}(x, v)).
\]  

\[\square\]
Here $\gamma^{-t}(x, v) = (\xi_\varepsilon(-t), \eta_\varepsilon(-t))$ where $\eta_\varepsilon$ is an autonomous jump process and $\xi_\varepsilon$ is an additive functional of $\eta_\varepsilon$. Equation (4.3) is an evolution equation for the probability density associated to a particle performing random jumps in the velocity space at random Markov times.

We start the proof by looking at the microscopic solution $f_\varepsilon$ defined by (1.5). For $(x, v) \in \mathbb{R}^2 \times \mathbb{R}^2$, $t > 0$, we have

$$f_\varepsilon(x, v, t) = e^{-\mu_\varepsilon |T(x, t)|} \sum_{N \geq 0} \frac{\mu_\varepsilon^N}{N!} \int_{B(x, t)^N} d\mathbf{c}_N f_0(T^{-t}_{\varepsilon, N}(x, v)), \quad (4.4)$$

where $T^{-t}_{\varepsilon, N}(x, v)$ is the Hamiltonian flow with initial datum $(x, v)$ and $B(x, t)$ is the disk centered in $x$ with radius $t$.

Given the configuration of obstacles $\mathbf{c}_N = c_1 \ldots c_N$, we shall say that $c_i$ is internal if it influences the motion up to the time $t$, i.e.

$$\inf_{0 \leq s \leq t} |x_\varepsilon(-s) - c_i| < \varepsilon, \quad (4.5)$$

while we shall call $c_i$ external if

$$\inf_{0 \leq s \leq t} |x_\varepsilon(-s) - c_i| \geq \varepsilon. \quad (4.6)$$

Here $(x_\varepsilon(-s), v_\varepsilon(-s)) = T^{-s}_{\varepsilon, N}(x, v), s \in [0, t]$.

We can perform the integration over the external obstacles and we get

$$f_\varepsilon(x, v, t) = \sum_{Q \geq 0} \frac{\mu_\varepsilon^Q}{Q!} \int_{B(x, t)^Q} d\mathbf{b}_Q e^{-\mu_\varepsilon |T(\mathbf{b}_Q)|} \chi((\mathbf{b}_Q \text{ internal})) f_0(T^{-t}_{\varepsilon, b_Q}(x, v)). \quad (4.7)$$

where $\chi(E)$ is the characteristic function of the event $E$ and $T(\mathbf{b}_Q)$ is the tube

$$T(\mathbf{b}_Q) = \{ y \in B(x, t) \text{ s.t. } \exists s \in (0, t) \text{ s.t. } |y - x_\varepsilon(-s)| \leq \varepsilon \}. \quad (4.8)$$

Note that in the previous integration we are not considering possible overlappings of obstacles. This is legitimate because we shall see that this event is negligible as $\varepsilon$ tends to 0.

Furthermore, let us restrict to the configurations such that the light particle’s trajectory does not start from inside an obstacle and does not end inside an obstacle: in formula

$$\chi_1(\mathbf{b}_Q) = \chi(\mathbf{b}_Q \text{ s.t. } b_i \notin B(x, \varepsilon) \text{ and } b_i \notin B(x_\varepsilon(-t), \varepsilon) \text{ for all } i = 1, \ldots, Q). \quad (4.9)$$

As for the overlappings, this choice is not really restrictive because the contribution related to $1 - \chi_1$ is going to vanish in the limit, as we shall see. Moreover, we will now list other events that will turn out to be negligible as $\varepsilon$ approaches 0.

i) **Complete cyclotronic orbit**

A first cyclotron orbit is completed without suffering any collisions and a repeated collision occurs with the same scatterer without any collision in the meantime.

We set

$$\chi_{\text{circ}}(\mathbf{b}_Q) = \chi(\{ \mathbf{b}_Q \text{ s.t. } i \text{ is realized} \}). \quad (4.10)$$

For a pictorial representation of the event i) see Fig. 2. We now define

$$f_\varepsilon(x, v, t) = \sum_{Q \geq 0} \frac{\mu_\varepsilon^Q}{Q!} \int_{B(x, t)^Q} d\mathbf{b}_Q e^{-\mu_\varepsilon |T(\mathbf{b}_Q)|} \chi((\mathbf{b}_Q \text{ internal})) \times (1 - \chi_{\text{circ}}(\mathbf{b}_Q)) \chi_1(\mathbf{b}_Q) f_0(T^{-t}_{\varepsilon, b_Q}(x, v)). \quad (4.11)$$
Fig. 2 On the left a cyclotron orbit completed without suffering collisions is represented. On the right there is a repeated collision with the same scatterer.

Note that $f_\varepsilon \geq \tilde{f}_\varepsilon$. For $t < T_L$ one expects that the approximation with the dynamics of the test particle in absence of the external field is true. The unexpected fact is that even for $t \gtrsim O(T_L)$ this still holds because (4.10) tends to 0 as $\varepsilon \to 0$. Hence, for a given configuration $b_Q$ such that $\chi_1[1 - \chi_{circ}(b_Q)] = 1$, we have that the measure of the tube can be estimated by

$$|T(b_Q)| \leq 2\varepsilon t.$$ (4.12)

At this point we define

$$\tilde{f}_\varepsilon(x, v, t) = e^{-2\mu \varepsilon^{-2\alpha}} \sum_{Q \geq 0} \frac{\mu^Q}{Q!} \int_{B(x, t)^Q} db_Q \chi(\{b_Q \text{ internal}\})$$

$$\times (1 - \chi_{circ}(b_Q)) \chi_1(b_Q) f_0(T_{b_Q}^{-1}(x, v)).$$ (4.13)

Thanks to (4.12) we get

$$f_\varepsilon \geq \tilde{f}_\varepsilon \geq \tilde{\tilde{f}}_\varepsilon.$$ (4.14)

According to a classical argument introduced in [12] (see also [5,8,9]), we now want to remove from $\tilde{\tilde{f}}_\varepsilon$ all the events that prevent the light particle’s trajectory to be the Markov process described by $h_\varepsilon$.

For any fixed initial condition $(x, v)$ we order the obstacles $b_1, \ldots, b_N$ according to the scattering sequence. Let $\rho_i$ and $t_i$ be the impact parameter and the backwards entrance time of the light particle in the protection disk around $b_i$, namely $B(b_i, \varepsilon)$. Then we perform the following change of variables

$$b_1, \ldots, b_N \to \rho_1, t_1, \ldots, \rho_N, t_N$$ (4.15)

with

$$0 \leq t_N < t_{N-1} < \ldots < t_1 \leq t.$$
Conversely, fixed the impact parameters \( \{\rho_i \} \) and the hitting times \( \{t_i \} \) we construct the centers of the obstacles \( b_i = b(\rho_i, t_i) \) and a trajectory \( \tilde{y}^{-s}(x, v) := (\tilde{\xi}_e(-s), \tilde{\eta}_e(-s)), \ s \in [0, t] \) inductively.

Suppose that we are able to define the obstacles \( b_1, \ldots, b_{i-1} \) and a trajectory \( \tilde{y}^{-s}(x, v) := (\tilde{\xi}_e(-s), \tilde{\eta}_e(-s)) \) up to the time \( s = t_{i-1} \). We then define the trajectory between times \( t_{i-1} \) and \( t_i \) as that of the evolution of a particle moving under the action of the Lorentz force and of the potential \( \varepsilon^a \phi(\varepsilon^{-1} | \cdot - b_{i-1}|) \) with initial datum \( (\tilde{\xi}_e(-t_{i-1}), \tilde{\eta}_e(-t_{i-1})) \). Then \( b_j \) is defined to be the only point at distance \( \varepsilon \) of \( \tilde{\xi}_e(-t_i) \) and algebraic distance \( \rho_i \) from the straight line which is tangent to the trajectory at the point \( \tilde{\xi}_e(-t_i) \).

However, \( \tilde{y}^{-s}(x, v) = (x_e(-s), v_e(-s)) \) (therefore the mapping (4.15) is one-to-one) only outside the following pathological situations (relative to the backward trajectory).

**ii) Overlapping**

If \( b_i \) and \( b_j \) are both internal then \( B(b_i, \varepsilon) \cap B(b_j, \varepsilon) \neq \emptyset \).

**iii) Recollisions**

There exists \( b_i \) such that for \( s \in (t_{j+1}, t_j), j > i, \xi_e(-s) \in B(b_i, \varepsilon) \).

**iv) Interferences**

There exists \( b_i \) such that \( \xi_e(-s) \in B(b_j, \varepsilon) \) for \( s \in (t_{i+1}, t_i), j > i \).

We simply skip such events by setting

\[
\chi_{ov} = \chi((b_Q \text{ s.t. ii) is realized})), \\
\chi_{rec} = \chi((b_Q \text{ s.t. iii) is realized})), \\
\chi_{int} = \chi((b_Q \text{ s.t. iv) is realized})),
\]

and defining

\[
\tilde{f}_e(x, v, t) = e^{-2\varepsilon^{2a} \mu t} \sum_{Q \geq 0} \mu^Q_e \int_0^t dt_1 \cdots \int_0^{t_{Q-1}} dt_Q \int_{-\varepsilon}^{\varepsilon} d\rho_1 \cdots \int_{-\varepsilon}^{\varepsilon} d\rho_Q \times \chi_1(1 - \chi_{circ})(1 - \chi_{ov})(1 - \chi_{rec})(1 - \chi_{int}) f_0(\tilde{y}^{-t}(x, v)).
\] (4.16)

Note that

\[
\tilde{f}_e \leq \tilde{f}_e \leq \tilde{f}_e \leq f_e.
\] (4.17)

Note also that in (4.16) we have used the change of variables (4.15) for which, outside the pathological sets i), ii), iii), iv) \( \tilde{y}^{-t}(x, v) = (x_e(-t), v_e(-t)) \).

Next we remove \( \chi_1(1 - \chi_{circ})(1 - \chi_{ov})(1 - \chi_{rec})(1 - \chi_{int}) \) by setting

\[
\tilde{h}_e(x, v, t) = e^{-2\varepsilon^{2a} \mu t} \sum_{Q \geq 0} \mu^Q_e \int_0^t dt_1 \cdots \int_0^{t_{Q-1}} dt_Q \times \int_{-\varepsilon}^{\varepsilon} d\rho_1 \cdots \int_{-\varepsilon}^{\varepsilon} d\rho_Q f_0(\tilde{y}^{-t}(x, v)).
\] (4.18)

We observe that

\[
1 - \chi_1(1 - \chi_{ov})(1 - \chi_{circ})(1 - \chi_{rec})(1 - \chi_{int}) \leq (1 - \chi_1) + \chi_{ov} + \chi_{circ} + \chi_{rec} + \chi_{int}.
\] (4.19)

Then by (4.16) and (4.18) we obtain

\[
|
\tilde{h}_e(t) - \tilde{f}_e(t)| \leq \varphi_1(\varepsilon, t)
\]
with
\[
\varphi_1(\varepsilon, t) = \|f_0\|_\infty e^{-2\varepsilon - 2\alpha \mu t} \sum_{Q \geq 0} \mu^Q \int_0^t dt_1 \cdots \int_0^{t_{Q-1}} dt_Q \int_{-\varepsilon}^\varepsilon d\rho_1 \cdots \int_{-\varepsilon}^\varepsilon d\rho_Q \times \left[ (1 - \chi_1) + \chi_{ov} + \chi_{cjr} + \chi_{int} \right].
\] (4.20)

We can prove that \( \varphi_1 \) is negligible in the limit. The precise statement follows.

**Proposition 4.1** Let \( \varphi_1(\varepsilon, t) \) be defined as in (4.20). For any \( t \in [0, T] \)
\[
\|\varphi_1(\varepsilon, t)\|_{L^1} \to 0
\]
as \( \varepsilon \to 0 \).

**Proof** See Sect. 4.1. \( \square \)

To complete the proof of Theorem 2.1 we still need to show the asymptotic equivalence of \( \tilde{h}_\varepsilon \) and \( h_\varepsilon \). Notice that \( h_\varepsilon \) is given by (4.3) where the trajectory \( \gamma^{-\varepsilon}(x, v) = (\xi_\varepsilon(-t), \eta_\varepsilon(-t)) \) is a jump process in the velocity space, i.e. the changes of velocity are instantaneous. We compare the trajectory \( \tilde{\gamma}^{-\varepsilon}(x, v) = (\tilde{\xi}_\varepsilon(-t), \tilde{\eta}_\varepsilon(-t)) \) with \( \gamma^{-\varepsilon}(x, v) \): being \( t_1, \cdots, t_Q \) the sequence of impact times and \( \tau \leq C \varepsilon \) the collision time, the spacial coordinates can differ only inside the interaction disk, while the velocities can differ only if \( t \in (t_1, t_1 + \tau) \).

In formulae we have
\[
|\xi_\varepsilon(-t) - \tilde{\xi}_\varepsilon(-t)| \leq C_1 Q \varepsilon,
\]
\[
|\eta_\varepsilon(-t) - \tilde{\eta}_\varepsilon(-t)| \leq C_2 \varepsilon^a \chi(t - t_1 \leq C \varepsilon). \tag{4.21}
\]

By exploiting the regularity of the initial condition \( f_0 \) we get
\[
|f_0(\xi_\varepsilon(-t), \eta_\varepsilon(-t)) - f_0(\tilde{\xi}_\varepsilon(-t), \tilde{\eta}_\varepsilon(-t))| \leq C' [Q \varepsilon + \varepsilon^a \chi(t - t_1 \leq C \varepsilon)] \tag{4.22}
\]
which implies
\[
|h_\varepsilon(x, v, t) - \tilde{h}_\varepsilon(x, v, t)| \leq C' e^{-2\mu \varepsilon e^{-2a}} \sum_{Q \geq 0} \mu^Q \int_0^t dt_1 \cdots \int_0^{t_{Q-1}} dt_Q \times \int_{-\varepsilon}^\varepsilon d\rho_1 \cdots \int_{-\varepsilon}^\varepsilon d\rho_Q [Q \varepsilon + \varepsilon^a \chi(t - t_1 \leq C \varepsilon)]
\leq C (t \varepsilon^{-2a} + \varepsilon^{1+a}). \tag{4.23}
\]

Hence we obtained \( \lim_{\varepsilon \to 0} \|h_\varepsilon - \tilde{h}_\varepsilon\|_{L^\infty([0, T] \times \mathbb{R}^2 \times S_1)} = 0 \). We observe that the monotonicity argument behind this strategy, see Eq. (4.17), the positivity of the solution \( h_\varepsilon \) of the Boltzmann equation and the conservation of mass imply that \( \tilde{f}_\varepsilon, \tilde{h}_\varepsilon \) and \( h_\varepsilon \) have the same asymptotic behavior in \( L^\infty ([0, T]; L^1(\mathbb{R}^2 \times S_1)) \) when \( \varepsilon \to 0 \).

### 4.1 Control of the Pathological Sets: Proof of Proposition 4.1

In this section we prove Proposition 4.1. This makes rigorous the claim of the heuristic argument presented in the paper’s introduction. For any measurable function \( u \) of the backward Markov process \( (\xi_\varepsilon, \eta_\varepsilon) \) we set
\[
\mathbb{E}_{x, \varepsilon}[u] = e^{-2\mu \varepsilon} \sum_{Q \geq 0} \mu^Q \int_0^t dt_1 \cdots \int_0^{t_{Q-1}} dt_Q \int_{-\varepsilon}^\varepsilon d\rho_1 \cdots \int_{-\varepsilon}^\varepsilon d\rho_Q u(\xi_\varepsilon, \eta_\varepsilon).
\]
Recalling (4.19) we have

\[ \varphi_1(\varepsilon, t) \leq \| f_0 \|_{\infty} E_{x,v}[(1 - \chi_1) + \chi_{ov} + \chi_{cir} + \chi_{rec} + \chi_{int}]. \]

We can skip the estimates of the first two contributions, i.e. \( E_{x,v}[(1 - \chi_1)] \) and \( E_{x,v}[\chi_{ov}] \), since the presence of the external field does not affect the classical arguments which can be found in [5,8,9]. However, the presence of the magnetic field and consequently the circular motion of the test particle strongly affects the explicit estimates of the pathological events ii), iii). Therefore, we need a detailed analysis for \( \chi_{cir} \), \( \chi_{rec} \) and \( \chi_{int} \).

For what concerns the pathological event due to a recollision with the same scatterer (see Fig. 3), we observe that \( \chi_{cir} = 1 \) if there exists an entrance time \( t_i \) such that \( |t_i - t_{i+1}| \geq T_L - \tau \geq T_L - C \varepsilon \) for some \( i = 0, \ldots Q - 1 \). Moreover, \( \chi_{circ} = 1 \) also when a test particle performs an entire Larmor orbit without colliding with any obstacles. As explained in the introduction, the probability of this event is bounded from above by \( C \exp(-\frac{2\pi \mu \varepsilon}{B \varepsilon^{2\alpha}}) := c_\alpha(\varepsilon) \). Therefore, it results

\[
E_{x,v}[\chi_{cir}]
\leq c_\alpha(\varepsilon) + e^{-2\mu_{\varepsilon} t} \sum_{Q \geq 1} \mu_{\varepsilon}^Q \int_0^t dt_1 \cdots \int_0^{t_{Q-1}} dt_Q \int_{-\varepsilon}^\varepsilon d\rho_1 \cdots \int_{-\varepsilon}^\varepsilon d\rho_Q \sum_{i=1}^Q \chi(t_i < t_{i-1} - T_L + C \varepsilon)
\]

\[
\leq c_\alpha(\varepsilon) + e^{-2\mu_{\varepsilon} t} \sum_{Q \geq 1} (2\mu_{\varepsilon})^Q \int_0^t dt_1 \cdots \int_0^{t_{Q-1}} dt_Q \sum_{i=1}^Q \chi(t_i < t_{i-1} - T_L + C \varepsilon)
\]

(4.24)

where \( \mu_{\varepsilon} = \mu \varepsilon^{1-2\alpha} \) with \( \alpha \in (0, 1/2) \) and \( t_0 = t \).
We set
\[ I_i := \int_0^t dt_1 \cdots \int_0^{t_{Q-1}} dt_Q \chi(t_i < t_{i-1} - T_L + C\varepsilon) \]
\[ = c_\alpha(\varepsilon) + 2e^{-2t\mu\varepsilon\varepsilon(2\mu\varepsilon\varepsilon)(t - TL + C\varepsilon)} \sum_{Q \geq 1} (2\mu\varepsilon\varepsilon)^Q \frac{Q}{(Q - 1)!} \]
then
\[ \mathbb{E}_{x,v}[\chi_{circ}] \leq c_\alpha(\varepsilon) + e^{-2\mu\varepsilon\varepsilon} \sum_{Q \geq 1} (2\mu\varepsilon\varepsilon)^Q \sum_{i=1}^Q I_i . \quad (4.25) \]

Note that
\[ I_{i+1} = \int_0^t dt_1 \cdots \int_0^{t_{i-1}} dt_i \frac{(t_i - T_L + C\varepsilon)^{Q-i}}{(Q-i)!} \]
\[ = \int_0^t dt_1 \cdots \int_0^{t_{i-2}} dt_{i-1} \frac{(t_{i-1} - T_L + C\varepsilon)^{Q-(i-1)} - (T_L + C\varepsilon)^{Q-(i-1)}}{(Q - (i-1))!} \]
\[ = I_i - \frac{(-T_L + C\varepsilon)^{Q-i+1}}{(Q-i+1)!} \frac{t^i}{(i-1)!} \]
\[ = I_i - f_i \]
where \( f_i := \frac{(-T_L + C\varepsilon)^{Q-i}}{(Q-i)!} \frac{t^i}{(i-1)!} \). We now look at
\[ \sum_{i=1}^Q I_i = QI_1 - \sum_{j=0}^{Q-2} (Q-1-j) f_j \]
\[ = \frac{1}{(Q-1)!} \left[ (t - T_L + C\varepsilon)^Q - \sum_{j=0}^{Q-2} (Q-1)!(Q-j)! \frac{(-T_L + C\varepsilon)^{Q-j}}{(Q-j)!} \frac{t^j}{j!} \right. \]
\[ + \sum_{j=0}^{Q-2} (Q-1)!(Q-j)! \frac{(-T_L + C\varepsilon)^{Q-j}}{(Q-j)!} \frac{t^j}{j!} \]
\[ \leq \frac{1}{(Q-1)!} \left[ 2(t - T_L + C\varepsilon)^Q + (T_L - C\varepsilon)(t - T_L + C\varepsilon)^{Q-1} \right] . \quad (4.26) \]

Finally we got
\[ \mathbb{E}_{x,v}[\chi_{circ}] \]
\[ \leq c_\alpha(\varepsilon) + e^{-2\mu\varepsilon\varepsilon} \sum_{Q \geq 1} (2\mu\varepsilon\varepsilon)^Q \left[ \frac{2(t - T_L + C\varepsilon)^Q}{(Q-1)!} + \frac{(T_L - C\varepsilon)(t - T_L + C\varepsilon)^{Q-1}}{(Q-1)!} \right] \]
\[ = c_\alpha(\varepsilon) + 2e^{-2\mu\varepsilon\varepsilon}(2\mu\varepsilon\varepsilon)(t - T_L + C\varepsilon) \sum_{Q \geq 1} (2\mu\varepsilon\varepsilon)^{Q-1} \frac{(t - T_L + C\varepsilon)^{Q-1}}{(Q-1)!} \]
\[ e^{-2\mu_\varepsilon(T_L + C\varepsilon)} \sum_{Q\geq 1} (2\mu_\varepsilon\varepsilon)^Q Q^{-1} \frac{(t - T_L + C\varepsilon)^{Q-1}}{(Q - 1)!} = c_\alpha(\varepsilon) + 2e^{-2(T_L - C\varepsilon)(2\mu_\varepsilon\varepsilon)}(t - T_L + C\varepsilon) + e^{-2(T_L - C\varepsilon)(2\mu_\varepsilon\varepsilon)}(T_L - C\varepsilon)(2\mu_\varepsilon\varepsilon) \]  

for \( \alpha > 0 \) and \( \varepsilon \) sufficiently small. Hence, \( \mathbb{E}_{x,v}[\chi_{\text{circ}}] \) vanishes as \( \varepsilon \to 0 \).

We now consider a generalization of \( \chi_{\text{circ}} \); let be \( \chi_{\text{arc}}^{(v)} \) the characteristic function of the event such that the light particle does not hit any obstacles in a time interval equal to \( T_L \varepsilon^v \), \( 0 < v < 1 \). More precisely \( \chi_{\text{arc}}^{(v)} = 1 \) if there exists an entrance time \( t_i \) such that \( |t_i - t_{i+1}| \geq T_L \varepsilon^v - \tau \geq T_L \varepsilon^v - C\varepsilon \) for some \( i = 0, \ldots, Q - 1 \). The same computations as for \( \mathbb{E}_{x,v}[\chi_{\text{circ}}] \) show that \( \mathbb{E}_{x,v}[\chi_{\text{arc}}^{(v)}] \) vanishes as \( \varepsilon \to 0 \) when \( v < 2\alpha \). In other words this shows that the motion of the light particle outside the obstacles covers arcs of circle and corresponding angles of order at most \( O(\varepsilon^v) \).

Next we pass to the control of the recollision event. We observe that

\[ \chi_{\text{rec}} = (1 - \chi_{\text{arc}}^{(v)})\chi_{\text{rec}} + \chi_{\text{rec}} \chi_{\text{arc}}^{(v)} \leq (1 - \chi_{\text{arc}}^{(v)})\chi_{\text{rec}} + \chi_{\text{arc}}^{(v)} \]

and this implies

\[ \mathbb{E}_{x,v}[\chi_{\text{rec}}] \leq \mathbb{E}_{x,v}[(1 - \chi_{\text{arc}}^{(v)})\chi_{\text{rec}}] + \mathbb{E}_{x,v}[\chi_{\text{arc}}^{(v)}], \]

but \( \mathbb{E}_{x,v}[\chi_{\text{arc}}^{(v)}] \) is vanishing in the limit \( \varepsilon \to 0 \) as we have seen before. Therefore, we can focus on \( \mathbb{E}_{x,v}[(1 - \chi_{\text{arc}}^{(v)})\chi_{\text{rec}}] \). Let \( t_i \) the first time the light particle hits the \( i \)-th scatterer \( b_i \), \( v_i^- \) the incoming velocity, \( v_i^+ \) the outgoing velocity (with respect to the backwards trajectory) and \( t_i^+ \) the exit time. Moreover, we fix the axis in such a way that \( v_i^+ \) is parallel to the \( x \)-axis. We have

\[ \chi_{\text{rec}} (1 - \chi_{\text{arc}}^{(v)}) \leq (1 - \chi_{\text{arc}}^{(v)}) \sum_{i=1}^{Q} \sum_{j>i} \chi_{\text{arc}}^{(v)}, \]

where \( \chi_{\text{arc}}^{(v)} = 1 \) if and only if \( b_i \) (constructed via the sequence \( t_1, \rho_1, \ldots, t_i, \rho_i \)) is recollided in the time interval \( (t_j, t_{j-1}) \). Note indeed that a recollision can occur only if the rotation angle \( |\sum_{h=i+1}^{j-1} (\theta_h + \varphi_h)| > \pi \) where \( \varphi_h \) is the angle covered outside the obstacles in the time interval \( (t_{h+1}, t_h^+) \), being \( \theta_h \) the \( h \)-th scattering angle. The constraint \((1 - \chi_{\text{arc}}^{(v)})\) implies that \( |\varphi_h| \leq C\varepsilon^v \).

Hence, since \( |\theta_h + \varphi_h| \leq C\varepsilon^\alpha + C'\varepsilon^v \leq C''\varepsilon^{v/2} \), in order to have a recollision there must be an intermediate velocity \( v_k, k = i + 1, \ldots, j - 1 \) such that

\[ |v_k^+ \cdot v_j^+| \leq C\varepsilon^{v/2}, \]

namely \( v_k^+ \) is almost orthogonal to \( v_j^+ \) (see Fig. 4).

Then

\[ \chi_{\text{rec}} (1 - \chi_{\text{arc}}^{(v)}) \leq (1 - \chi_{\text{arc}}^{(v)}) \sum_{i=1}^{Q} \sum_{j>i} \sum_{k=i+2}^{j-1} \chi_{\text{rec}}^{i,j,k}, \]

where \( \chi_{\text{arc}}^{i,j,k} = 1 \) if and only if \( \chi_{\text{arc}}^{i,j} = 1 \) and (4.30) is fulfilled. Following [5], we fix all the parameters \( \rho_1, \ldots, \rho_Q, t_1, \ldots, t_Q \) but \( t_{k+1} \). The two branches of the trajectory \( l_1, l_2 \) are rigid.
so that, when a recollision occurs, the integration domain with respect to $t_{k+1}$ is restricted to a time interval bounded by

$$\frac{2\varepsilon}{\cos C_{\varepsilon}v/2} \leq 4\varepsilon.$$  

Performing all the other integrations and summing over $i, j, k$ we obtain

$$E_{x,v} \left[ (1 - \chi_{arc}) \sum_{i=1}^{Q} \sum_{j=1}^{Q} \sum_{k=i+1}^{j-1} \chi_{\text{rec}}^{i,j,k} \right] \leq C_{\varepsilon} e^{-2\mu\varepsilon^{-2\alpha}} \sum_{Q \geq 3} (Q - 1)(Q - 2)(Q - 3) \frac{(2\mu\varepsilon^{-2\alpha}Q)!}{(Q - 1)!} t \varepsilon^{Q-1} \leq C' t^3 \varepsilon^{1-8\alpha},$$  

which tends to 0 as $\varepsilon$ goes to 0 for $\alpha < 1/8$.

Following the strategy used in [5], since a backward interference is a forward recollision, the estimate for the interference event can be handled by using the Liouville Theorem.

### 5 Proof of Proposition 3.7

Our aim is to prove the asymptotic equivalence of $f_{\varepsilon}$, defined by (2.5), and $h_{\varepsilon,\gamma}$ solution of the linear Boltzmann equation (3.15), namely
\[
\begin{aligned}
\left\{
(\partial_t + v \cdot \nabla x + B v^\perp \cdot \nabla v) h_{\varepsilon,\gamma}(t, x, v) &= \tilde{L} h_{\varepsilon,\gamma}(t, x, v), \\
h_{\varepsilon,\gamma}(x, v, 0) &= f_0(x, v),
\right.
\end{aligned}
\]

which reads
\[
h_{\varepsilon,\gamma}(x, v, t) = e^{-2\mu \varepsilon^{-1} t} \sum_{Q \geq 0} \frac{\mu^Q}{Q!} \int_0^t dt_Q \cdots \int_0^{t_Q} dt_1 \int_{-\mu \varepsilon^{-1}}^{\mu \varepsilon^{-1}} d\rho_1 \cdots \int_{-\mu \varepsilon^{-1}}^{\mu \varepsilon^{-1}} d\rho_Q f_0(\tilde{\xi}_\varepsilon(-t), \tilde{\eta}_\varepsilon(-t)).
\]

(5.1)

We recall that we can expand \( f_\varepsilon \) as follows:
\[
f_\varepsilon(x, v, t) = e^{-\mu_\varepsilon[B(x, t)]} \sum_{N \geq 0} \frac{\mu^N}{N!} \int_{B(x, t)^N} d\mathbf{c}_N \int_0^{T_{\varepsilon^{-1}}}(T_{\varepsilon^{-1}}(x, v));
\]

(5.2)

where \( T_{\varepsilon^{-1}}(x, v) \) is the Hamiltonian flow with initial datum \((x, v)\) and \( B(x, t) \) is the disk of center \( x \) and radius \( t \). We observe that the proof follows the same strategy of Proposition 3.1 (see also Sect. 3 in [8]). As before the hardest part is the estimate of the non-Markovian contribution which is summarized in the following proposition.

**Proposition 5.1** Let \( \phi_1(\varepsilon, t) \) be defined as in (4.20) with the only difference that the radius of the obstacles is now \( \varepsilon \gamma \) instead of \( \varepsilon \) and the collision time \( \tau \) is bounded by \( C \varepsilon \gamma \) instead of \( C \varepsilon \). Then for any \( t \in [0, T] \)
\[
\| \phi_1(\varepsilon, t) \|_{L^1} \to 0
\]
as \( \varepsilon \to 0 \).

**Proof** Also in this case we can skip the estimates of the contributions \( E_{x,v}[1 - \chi_1] \) and \( E_{x,v}[\chi_{ov}] \) since the presence of the external field does not affect the bounds in [8]. Moreover, as in Proposition 4.1, if we know that \( E_{x,v}[\chi_{rec}] \) is negligible, then the Liouville theorem guarantees that also \( E_{x,v}[\chi_{int}] \) can be disregarded in the limit. Hence, it suffices to focus on \( E_{x,v}[\chi_{circ}] \) and \( E_{x,v}[\chi_{rec}] \). So we look at
\[
E_{x,v}[\chi_{circ}] \leq c_\gamma(\varepsilon) + e^{-2t_\mu \varepsilon \gamma} \sum_{Q \geq 1} (2\mu \varepsilon \gamma)^Q \int_0^t dt_1 \cdots \int_0^{t_{Q-1}} dt_Q \times \sum_{i=0}^{Q-1} \chi(t_{i+1} < t_i - T_L + C \varepsilon \gamma)
\]

(5.3)

where \( c_\gamma(\varepsilon) = C \exp(-2\pi \mu \varepsilon \gamma), \mu_\varepsilon = \mu \varepsilon^{-1} \) and \( \gamma \in (0, 1) \). Following the same strategy as in Proposition 4.1, we obtain
\[
E_{x,v}[\chi_{circ}] \leq c_\gamma(\varepsilon) + e^{-2t_\mu \varepsilon \gamma} \sum_{Q \geq 1} (2\mu \varepsilon \gamma)^Q \times \left[ \frac{2(t - T_L + C \varepsilon \gamma)^Q}{(Q - 1)!} + \frac{(T_L - C \varepsilon \gamma)(t - T_L + C \varepsilon \gamma)^{Q-1}}{(Q - 1)!} \right]
\]

\[
= c_\gamma(\varepsilon) + 2e^{-2(t - T_L + C \varepsilon \gamma)\mu \varepsilon^{-1} - 2t_\mu \varepsilon \gamma} (2\mu \varepsilon^{-1}) (t - T_L + C \varepsilon \gamma) + e^{-2(T_L - C \varepsilon \gamma)\mu \varepsilon^{-1} - 2t_\mu \varepsilon \gamma} (2\mu \varepsilon^{-1}) (T_L - C \varepsilon \gamma)
\]

(5.4)

which vanishes as \( \varepsilon \to 0 \) and \( \gamma \in (0, 1) \).
To control the recollision event we can follow the strategy used in Sect. 4.1 and in [8]. More precisely, as in Sect. 4.1, we introduce \( \chi^{(M)}_{arc} \) such that \( \chi^{(M)}_{arc} = 1 \) if there exists an entrance time \( t_i \) such that \( |t_i - t_{i+1}| \geq T_L/M - \tau \geq T_L/M - C\varepsilon^{1/2} \) for some \( i = 0, \ldots Q - 1 \) where \( M \) is a finite constant and \( M > 1 \). One can easily see that \( \mathbb{E}_{x,v}[\chi^{(M)}_{arc}] \) vanishes as \( \varepsilon \to 0 \) when \( \gamma < 1 \). Furthermore,

\[
\mathbb{E}_{x,v}[\chi_{rec}] \leq \mathbb{E}_{x,v}[\left(1 - \chi^{(M)}_{arc}\right)\chi_{rec}] + \mathbb{E}_{x,v}[\chi^{(M)}_{arc}]
\]

but \( \mathbb{E}_{x,v}[\chi^{(M)}_{arc}] \) is vanishing in the limit \( \varepsilon \to 0 \) as we have seen before. Therefore, we can focus on \( \mathbb{E}_{x,v}[\left(1 - \chi^{(M)}_{arc}\right)\chi_{rec}] \). We now distinguish the collisions as

\[
(1 - \chi^{(M)}_{arc})\chi_{rec} \leq (1 - \chi^{(M)}_{arc}) \sum_{i=1}^{Q} \sum_{j>1}^{j} \chi^{i,j}_{rec} \left( \sin \alpha_{jk} \leq \frac{\varepsilon^{1/2}}{4}, \ \forall k = i, \ldots, j - 1 \right)
\]

\[
+ (1 - \chi^{(M)}_{arc}) \sum_{i=1}^{Q} \sum_{j>1}^{j} \chi^{i,j}_{rec} \left( \sin \alpha_{jk} \geq \frac{\varepsilon^{1/2}}{4}, \ \text{for some} \ k = i, \ldots, j - 1 \right) \quad (5.5)
\]

where \( \chi^{i,j}_{rec} = 1 \) if and only if \( b_i \) (constructed via the sequence \( t_1, \rho_1, \ldots, t_i, \rho_i \)) is recollided in the time interval \((t_j, t_j+1)\) outside the scatterers and \( \theta_{jk} \) (with \( i < k < j \)) is the absolute value of the sum of the angles between the outgoing velocity \( v_k^- \) from the \( k \)-th obstacle and the recolliding velocity \( v_j^- \), i.e.

\[
\alpha_{jk} = |\varphi_k| + \sum_{r=k+1}^{j} |\theta_r + \varphi_r|
\]

where \( \theta_r \) is the deflection angle due to the \( r \)-th scatterer and \( \varphi_r \) is the angle covered in the time interval \((t_{r+1}, t_{r+1}^-)\) outside the scatterers and \( \varphi_j \) is the angle covered between the \( j \)-th obstacle before recolliding with \( b_i \). Here \( \delta > 0 \) is a suitable parameter that we will fix later.

Note that, thanks to \( 1 - \chi^{(M)}_{arc} \), we have \( |\varphi_r| \leq 2\pi/M \) for any \( r \).

As noticed in [8], the constraint \( \sin \alpha_{jk} \leq \varepsilon^{1/2}/4 \) implies that \( |\theta_r + \varphi_r - \pi| < \varepsilon^{1/2} \) for some \( r = i, \ldots, j - 1 \), thus we get

\[
\mathbb{E}_{x,v} \left[ (1 - \chi^{(M)}_{arc}) \sum_{i=1}^{Q} \sum_{j>1}^{j} \chi^{i,j}_{rec} \left( \sin \alpha_{jk} \leq \frac{\varepsilon^{1/2}}{4}, \ \forall k = i, \ldots, j - 1 \right) \right]
\]

\[
\leq e^{-2t_{\mu_x^{1/2}}} \sum_{Q \geq 0} \mu_x^{Q} \sum_{i=1}^{Q} \sum_{j=i+1}^{j} \int_{t_0}^{t_1} dt_1 \int_{t_0}^{t_2} dt_2 \cdots \int_{t_0}^{t_Q-1} dt_Q
\]

\[
\times \int_{-\varepsilon^{1/2}}^{\varepsilon^{1/2}} d\rho_1 \int_{-\varepsilon^{1/2}}^{\varepsilon^{1/2}} d\rho_2 \cdots \int_{-\varepsilon^{1/2}}^{\varepsilon^{1/2}} d\rho_Q \chi(|\theta_k + \varphi_k - \pi| < \varepsilon^{1/2}). \quad (5.6)
\]

Now we note that for \( M \) big enough

\[
\int_{-\varepsilon^{1/2}}^{\varepsilon^{1/2}} d\rho_k \chi(|\theta_k + \varphi_k - \pi| < \varepsilon^{1/2}) = \int_{-\pi}^{\pi} d\theta_k \tilde{\Gamma}^{(B)}_{\varepsilon,\varepsilon^{1/2}}(\theta_k) \chi(|\theta_k + \varphi_k - \pi| < \varepsilon^{1/2})
\]

\[
= \varepsilon \int_{-\pi}^{\pi} d\theta_k \Gamma^{(B)}_{\varepsilon,\varepsilon^{1/2}}(\theta_k) \chi(|\theta_k + \varphi_k - \pi| < \varepsilon^{1/2})
\]
\[ = \varepsilon \int_{\pi(1 - \frac{3}{2})^{-\varepsilon^\delta}}^{\pi(1 - \frac{3}{2})^{+\varepsilon^\delta}} d\theta_k \Gamma^{(B)}_{\varepsilon, \gamma}(\theta_k) \]

\[ \leq C\varepsilon^{1+\delta} \] (5.7)

where \( \Gamma^{(B)}_{\varepsilon, \gamma}(\theta_k) \) is the differential cross section associated to the rescaled potential \( \tilde{\psi}_{\varepsilon} \), while \( \Gamma_{\varepsilon, \gamma}(\theta_k) \) is differential cross section associated to the unrescaled potential \( \psi_{\varepsilon} \). In the last line of (5.7) we used that \( \Gamma^{(B)}_{\varepsilon, \gamma}(\theta) \) is uniformly bounded in \( \varepsilon \) when \( \theta \) is far from 0, as shown in Appendix 2.

Then from (5.6) one gets

\[ E_{x,v} \left[ (1 - \chi^{(M)}_{\text{arc}}) \sum_{i=1}^{Q} \sum_{j>1}^{j_{\text{rec}}} \chi_{ij} \chi \left( \sin \alpha_{jk} \leq \frac{e^\delta}{4}, \forall k = i, \ldots, j - 1 \right) \right] \]

\[ \leq C e^{-2t\mu_{e'}^{\gamma'}} \sum_{Q \geq 1} \frac{(2t\mu_{e'}^{\gamma'})^Q Q^3 e^{1+\delta-\gamma}}{Q!} \leq C e^{2\gamma-2+\delta}. \] (5.8)

For what concerns the second term in (5.5) we note that, once we fix all the variables \( \{t_i\}_{i=1}^Q \) and \( \{\rho_i\}_{i=1}^Q \) except \( t_k \), by using the same geometrical argument as the one illustrated in Fig. 5 of [8], one gets that the integral over \( t_{k+1} \) is bounded by

\[ \frac{2e^{\gamma'}}{\sin \alpha_{jk}} \leq 8 e^{\gamma-\delta}. \] (5.9)

It follows that

\[ E_{x,v} \left[ (1 - \chi_{\text{arc}}) \sum_{i=1}^{Q} \sum_{j>1}^{j_{\text{rec}}} \chi_{ij} \chi \left( \sin \alpha_{jk} \geq \frac{e^\delta}{4}, \text{ for some } k = i, \ldots, j - 1 \right) \right] \]

\[ \leq e^{-2t\mu_{e'}^{\gamma'}} \sum_{Q \geq 0} \frac{(2t\mu_{e'}^{\gamma'})^Q Q^3 (Q-1)! 8 e^{\gamma-\delta}}{(Q-1)!} \]

\[ \leq C T e^{5\gamma-\delta-4}. \] (5.10)

We can now optimize the parameter \( \delta \) setting \( \delta = \frac{3\gamma-2}{2} \). From Eqs. (5.6) and (5.10) we finally end up with

\[ E_{x,v} [\chi_{\text{rec}}] \leq C e^{7\gamma-6 \gamma \frac{1}{2}} \] (5.11)

which tends to 0 as \( \varepsilon \to 0 \) if \( \gamma \in (6/7, 1) \).

**Acknowledgments** We thank Mario Pulvirenti for the proposed problem, Jens Marklof and Jani Lukkarinen for useful discussions on the topic. Finally we thank the anonymous referees for helpful suggestions to improve the manuscript. The research of M. Marcozzi and A. Nota has been supported by the Academy of Finland.

**Appendix 1: The Collision Time**

We want to estimate the time spent by a the test particle in the interaction disk associated to the central potential of finite range with a uniform magnetic field perpendicular to the plane. Let be \( e^{\alpha} \phi(r) \) with \( \alpha \in [0, 1/2] \) the central potential and \( \varepsilon B \) the modulus of the magnetic field.
The Lagrangian of the system is
\[ L(r, \dot{r}, \theta, \dot{\theta}) = \frac{1}{2} \dot{r}^2 + \frac{1}{2} r^2 \dot{\theta}^2 - \epsilon \alpha \phi(r) + \frac{B}{2} r^2 \dot{\theta} \]

We observe that the energy of the system is conserved. Moreover the Lagrangian does not depend on the variable \( \theta \), so we obtain the conservation of the conjugate momentum
\[ \frac{d}{dt} \left( r^2 \dot{\theta} + \epsilon \frac{B}{2} r^2 \right) = 0. \]

Therefore we obtain the following conserved quantities
\[ \dot{r}^2 + r^2 \dot{\theta}^2 + 2 \epsilon \alpha \phi = 2E, \]
\[ r^2 \dot{\theta} + \epsilon \frac{B}{2} r^2 = M, \]
and the equations of motion are
\[
\begin{align*}
\dot{r} &:= \frac{dr}{dt} = \sqrt{2(E - \epsilon \alpha \phi) - \frac{M^2}{r^2} - \frac{\epsilon^2 B^2}{4} r^2 + \epsilon M B} \\
\dot{\theta} &:= \frac{d\theta}{dt} = \frac{M}{r^2} - \epsilon \frac{B}{2}.
\end{align*}
\]

This implies that
\[
\begin{aligned}
\frac{dt}{dr} &= \left[ 2(E - \epsilon \alpha \phi(r)) - \frac{M^2}{r^2} - \frac{\epsilon^2 B^2}{4} r^2 + \epsilon M B \right]^{-1/2} \\
\frac{d\theta}{dr} &= \frac{M}{r^2} - \epsilon \frac{B}{2} \left/ \sqrt{2(E - \epsilon \alpha \phi(r)) - \frac{M^2}{r^2} - \frac{\epsilon^2 B^2}{4} r^2 + \epsilon M B} \right.. 
\end{aligned}
\]

We now define the effective potential
\[ \phi_{\text{eff}}(r) = \epsilon \alpha \phi(r) + \frac{M^2}{2 r^2} + \frac{\epsilon^2 B^2}{8} r^2 - \frac{\epsilon M B}{2}. \]

and we assume that the potential has a short range, i.e. \( \phi(r) : [0, 1] \rightarrow \mathbb{R} \) and \( \phi \) is continuous on \( [0, 1] \) and differentiable on \((0, 1)\).

Take the modulus of the initial velocity to be \(|v| = 1\). When the particle hits the obstacle of radius \( r = 1 \) the conserved quantities read
\[
\begin{align*}
E &= \frac{1}{2} \\
M &= \rho + \epsilon \frac{B}{2}
\end{align*}
\]

being \( \rho \in [0, 1] \) is the impact parameter. The effective potential is
\[
\phi_{\text{eff}}(r) = \epsilon \alpha \phi(r) + \left( \rho + \epsilon \frac{B}{2} \right)^2 \frac{1}{2 r^2} + \frac{\epsilon^2 B^2}{8} r^2 - \left( \rho + \epsilon \frac{B}{2} \right) \frac{\epsilon B}{2}
\]
\[
= \epsilon \alpha \phi(r) + \frac{1}{2} \left[ \rho \frac{r}{r} - \epsilon \frac{B}{2} \left( r - \frac{1}{r} \right) \right]^2.
\]

By integrating the equations of motion we obtain the collision time, namely the time spent inside the obstacle:
\[ \tau = 2 \int_{r_{\text{min}}}^{1} dr \left[ 1 - 2 \epsilon \alpha \phi(r) - \left( \frac{\rho}{r} - \epsilon \frac{B}{2} \left( r - \frac{1}{r} \right) \right)^2 \right]^{-1/2} \]
where $r_{\text{min}}$ (the minimum distance from the centre) is the unique zero of the radicand, i.e. 
$$1 = 2\phi_{\text{eff}}(r_{\text{min}}),$$
so we can reformulate (A.2) as
$$\tau = \sqrt{2} \int_{r_{\text{min}}}^{1} \frac{dr}{\sqrt{2(\phi_{\text{eff}}(r_{\text{min}}) - \phi_{\text{eff}}(r))}}$$
where $2\phi_{\text{eff}}(r) \leq 1$. The derivative of the effective potential reads
$$\phi_{\text{eff}}'(r) = \varepsilon^{\alpha} \phi'(r) - \frac{\rho^{2}}{r^{3}} - \frac{\varepsilon^{2} B^{2}}{4 r^{3}} - \frac{\varepsilon B \rho}{r^{3}} + \frac{\varepsilon^{2} B^{2}}{4 r}.$$

By the mean value theorem we get
$$|\phi_{\text{eff}}(r_{\text{min}}) - \phi_{\text{eff}}(r)| = |r - r_{\text{min}}|| - \phi_{\text{eff}}'(r^{*})|$$
$$\geq |r - r_{\text{min}}| \left( \inf_{r \in (r_{\text{min}},1)} | - \phi_{\text{eff}}'(r)| \right), \quad r^{*} \in (r_{\text{min}}, r)$$
and then
$$\tau \leq \frac{\sqrt{2}}{(\inf_{r \in (r_{\text{min}},1)} | - \phi_{\text{eff}}'(r)|)^{1/2}} \int_{r_{\text{min}}}^{1} \frac{1}{\sqrt{r - r_{\text{min}}} \, dr}.$$

Since $\phi_{\text{eff}}'(r) < 0$ for $r \in [0, 1)$, then $\inf_{r \in (r_{\text{min}},1)} | - \phi_{\text{eff}}'(r)| = \kappa > 0$ and it follows easily that
$$\tau \leq 2 \left( \frac{2(1 - r_{\text{min}})}{\kappa} \right)^{1/2} \leq 2 \left( \frac{2}{\kappa} \right)^{1/2}.$$

For the corresponding rescaled problem the effective potential reads
$$\phi_{\text{eff}}^{(\varepsilon)}(r) = \varepsilon^{\alpha} \phi(r / \varepsilon) + \frac{1}{2} \left[ \frac{\varepsilon \rho}{r} + B \left( \frac{\varepsilon^{2} B^{2}}{r^{3}} - r \right) \right]^{2} \tag{A.3}$$
with $\rho \in [0, 1)$. In this way one gets $-\phi_{\text{eff}}^{(\varepsilon)}(r) = \frac{1}{\varepsilon} F(r / \varepsilon, \varepsilon)$ where
$$F(y, \varepsilon) = -\varepsilon^{\alpha} \phi'(y) + (\rho + B) \rho y^{-3} + \frac{\varepsilon^{2} B^{2}}{4 y^{3}} (1 - y^{4}) \tag{A.4}$$
which is positive for $y < 1$ and uniformly in $\varepsilon$. The same argument as before yields the claimed estimate:
$$\tau_{\varepsilon} \leq \frac{(2\varepsilon)^{1/2}}{(\inf_{y \in (y_{0}, \varepsilon)} F(y, \varepsilon))^{1/2}} \int_{y_{0}}^{\varepsilon} \frac{dy}{\sqrt{y - y_{0}}} \leq C \varepsilon \tag{A.5}$$
where $y_{0} = y_{0}(\varepsilon)$ is such that
$$1 = 2\varepsilon^{\alpha} \phi(y_{0}) + \left( \frac{\rho}{y_{0}} - \frac{\varepsilon^{2} B y_{0}}{2} + \frac{B}{2 y_{0}} \right).$$

We consider now the long range unrescaled potential defined in Assumption B1), i.e. $\psi_{\varepsilon}(r) = r^{-s}$ for $r < \varepsilon^{\alpha-1}$ and $\psi_{\varepsilon}(r) = \varepsilon^{-s(\alpha-1)}$ for $r \geq \varepsilon^{\alpha-1}$. The same argument as for the short range case leads to the following estimate for the collision time after rescaling:
$$\tau \leq \frac{(2\varepsilon)^{1/2}}{(\inf_{y \in (y_{0}, \varepsilon^{\alpha-1})} \tilde{F}(y, \varepsilon))^{1/2}} \int_{y_{0}}^{\varepsilon^{\alpha}} \frac{dy}{\sqrt{y - y_{0}}} \leq C \varepsilon^{\alpha} \tag{A.6}$$
with
\[ \tilde{F}(y, \varepsilon) := -\psi'(y) + \varepsilon^{2(y-1)} \rho^2 y^{-3} + B^2 \varepsilon^{3y-2} \rho y^{-3} + B^2 y^{-3} \varepsilon^{4y-2} (1 - y^4 \varepsilon^{4(1-y)})/4 > 0 \]
for \( y \in (y_0, \varepsilon y) \) where \( y_0 = y_0(\varepsilon) \) is such that
\[ 1 = 2 \psi'(y_0) + \left[ \varepsilon^{y-1} \rho y_0^{-1} - \frac{B}{2} (\varepsilon^{2y-1} y_0^{-1} - \varepsilon y_0) \right]^2. \]

### Appendix 2: Cross Section

**Proposition B.1** Consider the scattering angle \( \theta(\rho, \varepsilon) \) of a particle with impact parameter \( \rho \) due to a uniform magnetic field perpendicular to the plane with modulus \( \varepsilon B \) and due to a radial potential \( \varepsilon^{-\alpha} \phi \), where \( \alpha > 0 \) and \( \phi \) satisfies assumptions A1, A2, A3. Consider also the scattering angle \( \tilde{\theta}(\rho, \varepsilon) \) associated to the same radial potential as before, but without any magnetic field. Then, for \( \varepsilon \) small enough one gets
\[ \theta(\rho, \varepsilon) = \tilde{\theta}(\rho, \varepsilon) + O(\varepsilon). \]  

**Proof** Following [9,16] we can write the exact formula for both of the scattering angles:
\[ \tilde{\theta}(\rho, \varepsilon) = \pi - 2 \arcsin \rho - 2 \int_{\rho}^{\tilde{u}_{max}(\rho, \varepsilon)} \frac{du}{\sqrt{1 - u^2 - 2\varepsilon^{-\alpha} \phi(\rho u^{-1})}} \]  
where \( \tilde{u}_{max}(\rho, \varepsilon) \) is the solution of the equation \( \tilde{u}_{max}^2 + 2\varepsilon^{-\alpha} \phi(\rho \tilde{u}_{max}^{-1}) = 1 \), while
\[ \theta(\rho, \varepsilon) = \pi - 2 \arcsin \rho - 2 \int_{\rho}^{u_{max}(\rho, \varepsilon)} \frac{du}{\sqrt{1 - 2\varepsilon^{-\alpha} \phi(\rho u^{-1}) - u^2[1 + \frac{\varepsilon B}{2\rho} (1 - \rho u^2)]^2}} \]  
where \( u_{max}(\rho, \varepsilon) \) is the solution of the equation \( 2\varepsilon^{-\alpha} \phi(\rho u_{max}^{-1}) + u_{max}^2[1 + \frac{\varepsilon B}{2\rho} (1 - \rho u^2)]^2 = 1 \).
Hence, an expansion of \( \theta(\rho, \varepsilon) \) for \( \varepsilon \) small enough yields the claimed asymptotic formula. \( \square \)

**Proposition B.2** Let \( \tilde{\theta} \) be the scattering angle associated to the long range potential \( \Psi(r) = r^{-s} \) with \( s > 2 \), \( \theta_{e, \gamma} \) the scattering angle due to a radial potential \( \psi_e \) defined in Assumption B1) and \( \theta_{e, \gamma}^{(B)} \) the scattering angle due to \( \psi_e \) and to a uniform, constant magnetic field perpendicular to the plane with modulus \( \varepsilon B \). Then one has
a) \( \theta_{e, \gamma}^{(B)} \rightarrow \tilde{\theta} \) as \( \varepsilon \rightarrow 0 \).
b) \( \Gamma_{e, \gamma}^{(B)}(\theta) \rightarrow \Gamma(\theta) \) as \( \varepsilon \rightarrow 0 \), where \( \Gamma_{e, \gamma}^{(B)}(\theta) \) is the differential cross section associated to the radial potential \( \psi_e \) and the magnetic field, while \( \Gamma(\theta) \) is the one associated to the radial potential \( \Psi \).
c) \( \Gamma_{e, \gamma}^{(B)}(\theta) \leq C \theta^{-1/s} \) uniformly in \( \varepsilon, B \).
Proof. a) Let us now consider the truncated potential $\tilde{\Psi} = r^{-s} - A^{-s}$ with $s > 2$ for $r \leq A$ and $\tilde{\Psi} = 0$ for $r > A$ with $A = \varepsilon^{1/2}$ and $\gamma \in (0, 1)$. Take the modulus of the initial velocity of the light particle to be $|v| = 1$.

We denote by $\rho$ the impact parameter (with $0 \leq \rho \leq A$) while the scattering angle (that is the angle between the ingoing and the outgoing relative velocities) is

$$\theta_{\varepsilon, \gamma}(\rho) = 2 \int_{\arcsin(\rho/A)}^{\pi/2} \left(1 - \frac{\sin \beta}{v + s \rho s^{-1}}\right) d\beta.$$

(B.4)

where $v = v(\beta)$ such that $v^2 + 2((v/\rho)^s - A^{-s}) = \sin^2 \beta$ and $v = \rho/r$ (see Appendix in [8]). Following [8, 16], we can write the formula for the scattering angle associated to the potential $\tilde{\Psi}$ and with the uniform magnetic field. Due to its invariance under rescaling, the scattering angle associated to the equations of motion 2.4 reads

$$\theta^{(B)}_{\varepsilon, \gamma}(M) = \pi - 2 \arcsin \left(\frac{M - \varepsilon^2 B^2 M^2}{A} - 2 \int_{r_a}^{A} \frac{1}{\sqrt{1 - 2\tilde{\Psi}_{\text{eff}}(r)}} \sin \beta \right)$$

(B.5)

where $\tilde{\Psi}_{\text{eff}}(r) = \frac{\tilde{\Psi}(r) + \frac{1}{2} \left(\frac{M}{r} - \frac{\varepsilon}{2} Br\right)^2}{M}$, $M$ is the value of the conserved momentum at the hitting time, i.e. $M = \rho + \varepsilon A^2 B^2 / 2$, and $r_a$ is defined as the solution of the equation $2\tilde{\Psi}_{\text{eff}}(r_a) = 1$. In the second line we made the change $r \to u \to \beta$ where $u = u(\beta, M) = \frac{M}{r}$ and $\sin^2 \beta = 2\tilde{\Psi}(M/u)$. Note that the change of variable $u \to \beta$ is well defined because $\tilde{\Psi}(M/u)$ is non-decreasing when $u \in [M/A, M/r^s]$ for $\varepsilon$ small enough.

From (B.4) and (B.5) it is clear that $\theta_{\varepsilon, \gamma}$ and $\theta^{(B)}_{\varepsilon, \gamma}$ have the same asymptotic behaviour as $\varepsilon$ approaches 0. Since $\theta_{\varepsilon, \gamma} \to \tilde{\theta}$, one gets the claim.

b) The inverse of the differential cross section associated to $\tilde{\theta}$ is

$$\left|\frac{d\theta_{\varepsilon, \gamma}}{d\rho}\right| = \frac{2}{\rho^{s+1}} \left(\frac{M}{A} - \varepsilon^2 B^2 / 2\right) \sin \beta \frac{sv^{s-1}}{v + s \rho s^{-1}} \left[s - \frac{1 + s(s - 1)v^{s-2} \rho^{-s}}{1 + s v^{s-2} \rho^{-s}}\right]$$

(B.6)

We want to study the limit of $d\theta^{(B)}_{\varepsilon, \gamma} / d\rho$. For a mere computational convenience, we prefer to look at $d\theta^{(B)}_{\varepsilon, \gamma} / dM$ which is related to $d\theta^{(B)}_{\varepsilon, \gamma} / d\rho$ via

$$\frac{d\theta^{(B)}_{\varepsilon, \gamma}}{dM} = \frac{d\theta^{(B)}_{\varepsilon, \gamma}}{d\rho} \left(\rho + \frac{\varepsilon^2 B^2}{2}\right).$$

(B.7)

From (B.5) one gets

$$\frac{d\theta^{(B)}_{\varepsilon, \gamma}}{dM} = -\frac{2}{A\sqrt{1 - (\frac{M}{A} - \frac{\varepsilon B A}{2})^2}} \left[\frac{s M^{-2} A^{2-s} + \varepsilon B A^2}{1 + s M^{-2} A^{2-s} + \frac{\varepsilon^2 B^2}{4} A^4 M^{-2}}\right]$$

$$-2 \int_{\arcsin(\frac{M}{A} - \frac{\varepsilon A}{2})}^{\pi/2} \frac{d\beta \sin \beta}{u + s \rho^{s-1} M^{-s} + \frac{s^2 B^2 M^2}{4u^s}}$$

$$\times \left[s^2 u^{s-1} M^{-s-1} - u'(1 + s(s - 1)u^{s-2} M^{-1})\right]$$
\[
- \varepsilon BM \int_{\arcsin(\frac{M}{2} - \frac{\varepsilon A}{2})}^{\pi/2} \frac{d\beta \sin \beta}{u} \left[ s(s + 1)u^{s-2}M^{-s}(u'M - 1) + 3u'M \right. \\
\left. + \frac{5\varepsilon^2 B^2 M^3 u'}{4u} + \varepsilon BM \frac{3 \varepsilon B^2 M^2}{2u^3} \right] 
\]

where

\[
u = \frac{du}{dM} = \frac{1}{M^{s+1}} \left( \frac{su^{s-1} + (1 - \varepsilon BM) \varepsilon BM^{s+1}}{1 + su^{s-2}M^{-s} - \frac{3 \varepsilon^2 B^2 M^2}{4u^4}} \right). 
\]

As for item a), one realizes that \(d\theta_{e,\gamma}^{(B)}/d\rho\) and \(d\theta_{e,\gamma}/d\rho\) are asymptotically equivalent for any \(\rho\), thus Proposition A.1 in [8] implies that \(\Gamma^{(B)}_{e,\gamma}(\theta) \rightarrow \Gamma(\theta)\) for \(\theta \in (-\pi, \pi)\) because its inverse map converges everywhere.

c) From (B.8) for \(\varepsilon\) small enough a tedious expansion gives

\[
\left| \frac{d\theta_{e,\gamma}^{(B)}}{dM} \right| \geq \frac{|d\theta_{e,\gamma}|}{d\rho} - \varepsilon |R(B, M, \varepsilon)| \geq \frac{1}{C} \frac{|d\theta_{e,\gamma}|}{d\rho} 
\]

where \(C > 1\) is a constant, \(R\) is bounded in \(\varepsilon\). The claim follows thanks to Proposition A.1 in [8].

References

[V]

*Renormalization of generalized KPZ equation*

Antti Kupiainen, Matteo Marcozzi
RENORMALIZATION OF GENERALIZED KPZ EQUATION

ANTTI KUPIAINEN AND MATTEO MARCOZZI

Abstract. We use Renormalization Group to prove local well posedness for a generalized KPZ equation introduced by H. Spohn in the context of stochastic hydrodynamics. The equation requires the addition of counter terms diverging with a cutoff $\epsilon$ as $\epsilon^{-1}$ and $\log \epsilon^{-1}$.

1. Introduction

Nonlinear stochastic PDE’s driven by a space time white noise have been under intensive study in recent years [1, 2, 3, 4, 5]. These equations are of the form

$$\partial_t u = \Delta u + V(u) + \Xi$$

where $u(t,x) \in \mathbb{R}^n$ is defined on $\Lambda \subset \mathbb{R}^d$, $V(u)$ is a function of $u$ and possibly its derivatives which can also be non-local and $\Xi$ is white noise on $\mathbb{R} \times \Lambda$, formally

$$E \Xi(t',x') \Xi(t,x) = \delta_{\alpha\beta} \delta(t' - t) \delta(x' - x).$$

In order to be defined these equations in general require renormalization. One first regularizes the equation by e.g. replacing the noise by a mollified version $\Xi(\epsilon)$ which is smooth on scales less than $\epsilon$ and then replaces $V$ by $V^{(\epsilon)} = V + W^{(\epsilon)}$ where $W^{(\epsilon)}$ is an $\epsilon$-dependent “counter term”. One attempts to choose this so that solutions converge as $\epsilon \to 0$.

The rationale of such counterterms is that although they diverge as $\epsilon \to 0$ their effect on solutions on scales much bigger than $\epsilon$ is small. They are needed to make the equation well posed in small scales but they disturb it little in large scales.

Such a phenomenon is familiar in quantum field theory. For instance in quantum electrodynamics the “bare” mass and charge of the electron have to be made cutoff dependent so as to have cutoff independent measurements at fixed scales. The modern way to do this is to use the Renormalization Group (RG) method which constructs a one parameter family of effective theories describing how the parameters of the theory vary with scale.

Such a RG method was applied to SPDE’s in [5] for the case $n = 1$, $d = 3$ and $V(u) = u^3$. In that case $W^{(\epsilon)} = (a\epsilon^{-1} + b \log \epsilon)u$ and path wise solutions were constructed recovering earlier results by [1, 2]. In the present paper we consider the equations of Stochastic Hydrodynamics recently introduced by Spohn [6]. They give rise to the problem (1) with $n = 3$, $d = 1$ and

$$V(u) = (\partial_x u, M \partial_x u)$$

where $(\cdot, \cdot)$ denotes the standard inner product in $\mathbb{R}^3$ and $M = (M^{(1)}, M^{(2)}, M^{(3)})$ with $M^{(i)}$ are symmetric matrices, so that (3) can be read component-wise as $V_i(u) = \cdots$
We construct path wise solutions in this case by taking
\[ W^{(\epsilon)} = a\epsilon^{-1} + b \log \epsilon. \]
The case \( n = 1 \) is the KPZ equation and this was constructed before by Hairer [7]. In that case \( b = 0 \). For a generic \( M_{\alpha \beta \gamma} \) in (3) \( b \neq 0 \). This counter term is third order in the nonlinearity as will be explained below. Thus in this case the simple Wick ordering of the nonlinearity does not suffice to make the equation well posed.

The content of the paper is as follows. In section 2 we define the model and state the result. The RG formalism is set up in a heuristic fashion in Section 3. Section 4 discusses the leading perturbative solution and sets up the fixed point problem for the remainder. Section 5 states the estimates for the perturbative noise contributions and in Section 6 the functional spaces for RG are defined and the fixed point problem solved. The main result is proved in Section 7. Finally in Sections 8 estimates for the covariances of the various noise contributions are proved.

2. THE REGULARIZED EQUATION AND MAIN RESULT

We consider the equation (1) with \( u(t, x) \) defined on \((t, x) \in \mathbb{R} \times T\) and nonlinearity given by (3). We study its integral form
\[ u = G \ast [(V(u) + \Xi)_{t \geq 0}] + e^{t\Delta} u_0 \] (4)
where \( G(t, x) = e^{t\Delta}(x, 0) \) and \( u_0 \) is the initial condition. In this paper we consider a random initial condition of Brownian type. Concretely we take \( u_0 \) the stationary solution to the linear problem \( V = 0 \) which is the Gaussian random field with covariance
\[ \mathbb{E} u_0(x) u_0(y) = \sum_{n \in \mathbb{Z} \setminus \{0\}} e^{2\pi i n(x-y)} \frac{1}{2(2\pi n)^2}. \]
\( \Xi \) is taken to be the white noise with vanishing spatial average i.e.
\[ \Xi(t, x) = \sum_{n \in \mathbb{Z} \setminus \{0\}} e^{2\pi i n x} \tilde{b}_n(t) \]
with \( b_n = \tilde{b}_{-n} \) independent complex Brownian motions. Thus (4) can be written in the form
\[ u = G \ast (V(u)_{t \geq 0} + \Xi) \] (5)
Instead of mollifying the noise we regularize the convolution by considering
\[ u = G_{\epsilon} \ast (V^{(\epsilon)}(u)_{t \geq 0} + \Xi) \] (6)
where
\[ G_{\epsilon}(t, x) = e^{t\Delta}(x, 0)(1 - \chi(\epsilon^{-2} t)) \] (7)
with \( \chi \geq 0 \) being a smooth bump, \( \chi(t) = 1 \) for \( t \in [0, 1] \) and \( \chi(t) = 0 \) for \( t \in [2, \infty) \) and
\[ V^{(\epsilon)}(u) = [(\partial u, M \partial u)_{t \geq 0} + C_{\epsilon} \mathbf{1}_{t \geq 0}] \] (8)
We look for \( C_{\epsilon} \) such that (6) has a unique solution \( u^{(\epsilon)} \) which converges as \( \epsilon \to 0 \) to a non trivial limit. Note that \( G_{\epsilon} \ast \Xi \) is a.s. smooth.

Our main result is
Theorem 1. There exists $C_{\epsilon}$ s.t. the following holds. For almost all realizations of the white noise $\Xi$ there exists $t(\Xi) > 0$ such that the equation (6) has for all $\epsilon > 0$ a unique smooth solution $u^{(\epsilon)}(t, x)$, $t \in [0, t(\Xi)]$ and there exists $u \in D^\prime([0, t(\Xi)] \times \mathbb{T})$ such that $u^{(\epsilon)} \rightarrow u$ in $D^\prime([0, t(\Xi)] \times \mathbb{T})$. The limit $u$ is independent of the regularization function $\chi$.

Remark 2. We will find that the renormalization parameter is given by

$$C_{\epsilon} = m_1 \epsilon^{-1} + m_2 \log \epsilon^{-1} + m_3$$

where the constants $m_1$ and $m_3$ depend on $\chi$ whereas the $m_2$ is universal i.e. independent on $\chi$. Furthermore, $m_2 = 0$ if $M^{(a)}_{\beta \gamma}$ is totally symmetric in the three indices.

3. Renormalization Group

The regularized equation (6) can be viewed as dealing with spatial scales larger than $\epsilon$. The idea of the Renormalization Group (RG) is to try to increase this small scale cutoff by deriving effective equations with larger cutoffs. This will be done inductively by going from scale $\ell$ to scale $L\ell$ with with $L$ fixed. One such step is called the RG transformation. It is useful to utilize the underlying scale invariance of the linear part of the equation and rescale at each step the small scale cutoff to unity. To do this define the space time scaling $s_\mu$ by

$$(s_\mu f)(t, x) = \mu^{-\frac{1}{2}} f(\mu^2 t, \mu x)$$

and set

$$\varphi = s_{\epsilon} u.$$ 

Note that $\varphi$ is defined on $\mathbb{R} \times \epsilon^{-1} \mathbb{T}$. By a simple change of variables in (6) we obtain

$$\varphi = G_1 * (v^{(\epsilon)}(\varphi) + \xi)$$

where

$$v^{(\epsilon)}(\varphi) := \epsilon^{\frac{1}{2}} (\partial_x \varphi, M \partial_x \varphi) + \epsilon^{\frac{3}{2}} C_{\epsilon}$$

and $\xi := c^2 s_{\epsilon} \Xi$ is equal in law with the white noise on $\mathbb{R} \times \epsilon^{-1} \mathbb{T}$ (we keep the convention that $v^{(\epsilon)}(\varphi) = 0$ for $t < 0$).

We note that in these dimensionless variables the small scale cutoff is unity and the strength of the nonlinearity is small, $\epsilon^{\frac{3}{2}}$ i.e. the model is subcritical. However, the price we pay is that we need to consider times of order $\epsilon^{-2}$ and spatial box of size $\epsilon^{-1}$.

Let us now attempt to increase the cutoff $\epsilon$. Fix $L > 1$ and decompose

$$G_1 = G_{L^2} + (G_1 - G_{L^2})$$

and

$$\varphi = \varphi_1 + \varphi_2.$$ 

Then (11) is equivalent to the pair of equations

$$\varphi_1 = G_{L^2}(v^{(\epsilon)}(\varphi_1 + \varphi_2) + \xi)$$

$$\varphi_2 = (G_1 - G_{L^2})(v^{(\epsilon)}(\varphi_1 + \varphi_2) + \xi).$$

$\varphi_1$ can be thought of living on scales $\geq L$ and $\varphi_2$ on scales $\in [1, L]$. Rescale now back to unit cutoff. Let $s := s_{L^{-1}}$ and set

$$\varphi_1 = s\varphi', \quad \varphi_2 = s\zeta.$$ 

Then

$$\varphi = s(\varphi' + \zeta)$$

(13)
with \( \varphi', \zeta \) solutions to
\[
\begin{align*}
\varphi' &= G_1 * (Sv(\epsilon)(\varphi' + \zeta) + \xi) \quad (14) \\
\zeta &= \Gamma * (Sv(\epsilon)(\varphi' + \zeta) + \xi) \quad (15)
\end{align*}
\]
where we defined the scaling operation
\[
(Sv)(\varphi) = L^2 s^{-1} v(s\varphi)
\]
and denoted
\[
\Gamma(t,x) := e^{\Delta(x,0)}(\chi(t) - \chi(L^2 t)). \quad (16)
\]
Note that \( \Gamma \) involves scales between \( L^{-1} \) and 1 so that the equation (15) turns out to be tractable: its solution \( \zeta \) is a function \( \zeta(\varphi') \) of \( \varphi' \). Plugging this into the large scale equation (14) yields
\[
\varphi' = G_1 * (\mathcal{R}v^{(\epsilon)}(\varphi') + \xi) \quad (17)
\]
where the new nonlinearity \( \mathcal{R}v^{(\epsilon)} \) is defined by
\[
\mathcal{R}v^{(\epsilon)}(\varphi') = Sv(\epsilon)(\varphi' + \zeta(\varphi')). \quad (18)
\]
\( \mathcal{R} \) is the Renormalization Group map: given a function \( v \) mapping a field \( \varphi(t,x) \) to a field \( v(\varphi)(t,x) \) we obtain a new function \( \mathcal{R}v \) by solving the small scale equation. Using (15) in (18) we may write the latter as an equation to determine \( \mathcal{R}v \):
\[
\mathcal{R}v(\varphi) = Sv(\varphi + \Gamma * (\mathcal{R}v(\varphi) + \xi)). \quad (19)
\]
We will set up the functional spaces where (19) is solved in Section 6. At this point let us see on a formal level how the solution of the original SPDE is reduced to the study of the map \( \mathcal{R} \). To do this it is convenient to take the cutoff \( \epsilon \) as
\[
\epsilon = L^{-N} \quad (20)
\]
so that we are interested in the limit \( N \to \infty \). With a slight abuse of notation, denote \( v^{(\epsilon)} \) by \( v^{(N)} \) and define inductively
\[
v^{(N)}_{n-1} := \mathcal{R}v^{(N)}_n. \quad (21)
\]
for \( n = N, N-1, \ldots \).

We call \( v^{(N)}_n \) the effective potential at scale \( L^{-n} \) starting with cutoff \( L^{-N} \). They are related to each other by the iteration
\[
v^{(N)}_n(\varphi) = Sv^{(N)}_n(\varphi + \Gamma_n * (v^{(N)}_{n-1}(\varphi) + \xi_{n-1})) \quad (22)
\]
where we denote explicitly the dependence of the noise on the scale:
\[
\xi_n := L^{-2n}s^{-n}\Xi.
\]
\( \xi_n \) equals in law the white noise in \( \mathbb{R} \times L^nT \). \( \Gamma_n \) is the operator (16) on \( \mathbb{R} \times L^nT \).

**Remark 3.** The definition of \( \mathcal{R} \) involves the scale \( n \) i.e. the size \( L^n \) of the spatial box where the heat kernel in (16) is defined. We suppress this dependence in the notation unless we want to emphasize it.

From (13) we infer that solutions to the equations \( v \) and \( v' = \mathcal{R}v \) are related by
\[
\varphi = s(\varphi' + \Gamma * (v'(\varphi') + \xi)).
\]
This leads to an iterative construction of the solution as follows. Suppose \( \varphi_n \) solves the effective equation
\[
\varphi_n = G_1 * (v^{(N)}_n(\varphi_n) + \xi_n). \quad (23)
\]
Then, the solution of the original equation (11) is given by
\[ \varphi = s^{-(N-n)} f_n^{(N)}(\varphi_n). \] (24)
where the maps \( f_n^{(N)} \) satisfy the induction
\[ f_{n-1}^{(N)}(\varphi) = L^{-2} S f_n^{(N)}(\varphi + \Gamma_n * (v_{n-1}^{(N)}(\varphi) + \xi_{n-1})) \] (25)
with the initial condition
\[ f_0^{(N)}(\varphi) = \varphi. \] (26)
Recalling (10) we conclude that the solution of the SPDE with cutoff \( \epsilon \) is given by
\[ u = s^N f_n^{(N)}(\varphi_n). \] (27)
Suppose now that (a) we can control the \( v_n^{(N)} \) and \( f_n^{(N)} \) for \( n \geq m \), (b) we can solve (23) for \( n = m \) on the time interval \([0, 1]\) (c) the solution \( \varphi_m \) is in the domain of \( f_m^{(N)} \). Then (27) yields the solution of the SPDE on the time interval \([0, L^{-2}m]\).
What determines the smallest \( m \) so that (a)-(c) hold? This is determined by the realization of the noise \( \Xi \). Indeed, the \( v_n^{(N)} \) are random objects i.e. functions of the white noise \( \Xi \). Let \( \mathcal{E}_m \) be the event such that the above holds for all \( N, n \) with \( m \leq n \leq N \). We will show that almost surely \( \mathcal{E}_m \) holds for some \( m < \infty \). For a precise statement see Section 5.

Equations (11), (22) and (25) involve the convolution operators \( \Gamma_n \) and \( G_1 \) respectively. These operators are infinitely smoothing and their kernels have fast decay in space time. In particular the noise \( \zeta = \Gamma_n * \xi_{n-1} \) entering equations (22) and (25) has a smooth covariance which has finite range in time and it has Gaussian decay in space. Hence the fixed point problem (22) turns out to be quite easy.

4. Perturbative contributions

The RG iteration we have defined is quite general: formally it holds for “arbitrary” nonlinearity \( v \) (and in any dimension as well, with appropriate scaling \( s \)). In the case at hand \( v \) is a function of \( \partial_x \varphi \) so it pays to change variables and denote
\[ \phi := \partial_x \varphi. \]
Denote also
\[ v_n^{(N)}(\varphi) = w_n^{(N)}(\phi) \]
and redefine the scaling operation as
\[ (\mathcal{S}\phi)(t, x) = L^{-\frac{1}{2}} \phi(L^{-2} t, L^{-1} x) \]
and
\[ (\mathcal{S}v)(\phi) = Ls^{-1} v(\mathcal{S}\phi) \]
so that the RG iteration (22) becomes
\[ w_{n-1}^{(N)}(\phi) = \mathcal{S} w_n^{(N)}(\phi + \Upsilon_n * (w_{n-1}^{(N)}(\phi) + \xi_{n-1})) \] (28)
where
\[ \Upsilon_n = \partial_x \Gamma_n. \]
Eq. (25) in turn becomes
\[ f_{n-1}^{(N)}(\phi) = L^{-1} \mathcal{S} f_n^{(N)}(\phi + \Upsilon_n * (w_{n-1}^{(N)}(\phi) + \xi_{n-1})) \] (29)
and we have the initial conditions
\[ w^{(N)}_N(\phi) = L^{\frac{1}{2}} N(\phi, M\phi) - L^{\frac{3}{2}} N C_{L^{-N}} \] (30)
\[ f^{(N)}_N(\phi) = \phi. \] (31)

From now on to avoid too many indices we suppress in the notation the superscript \((N)\) so that \(N\) is considered fixed and the scale \(n\) runs down from \(n = N\).

4.1. Solving the first order. It is instructive and useful to study the fixed point equation (28) to first order in \(w\). Define the map
\[ (Lw)(\phi) := Sw(\phi + \Upsilon_n * \xi_n - 1). \] (32)
Then (28) can be written as
\[ w^{n-1}_N(\phi) = (Lw_n)(\phi + \Upsilon_n * w^{n-1}_N(\phi)) \] (32)
so \(L\) is the linearization of the RG map \(R\): \(L = DR\). Its properties are crucial for understanding the flow of effective equations \(w_n\).

Let us consider the linear RG flow from scale \(N\) to scale \(n\) i.e. \(L^{N-n} w_N\). This can be computed by doing one RG step with \(L\) replaced by \(L^{N-n}\). We get
\[ L^{N-n} w_N(\phi) = S^{N-n} w_N(\phi + Y^{(N)}_n * \xi_n) \] (33)
where
\[ Y^{(N)}_n(t, x) = \partial_x H_n(t, x) \chi_{N-n}(t). \] (34)
with
\[ H_n(t, x) = \frac{1}{\sqrt{4\pi t}} \sum_{i \in \mathbb{Z}} e^{-\frac{(x + i t n)^2}{4t}} \] (35)
being the heat kernel on \(T_n\) and
\[ \chi_m(s) := \chi(s) - \chi(L^{2m} s) \] (36)
a smooth indicator of the interval \([L^{-2m}, 2]\). The field
\[ \vartheta_n := Y^{(N)}_n * \xi_n \]
is a stationary Gaussian vector-valued field with covariance given by
\[ \mathbb{E}\vartheta_{n, \alpha}(t, x)\vartheta_{n, \beta}(s, y) = \delta_{\alpha\beta} \mathcal{C}^{(N)}_n(|t - s|, x - y) \] (37)
where
\[ \mathcal{C}^{(N)}_n(t, x) = -\Delta \int_0^\infty H_n(t + 2\tau, x) \chi_{N-n}(t + \tau) \chi_{N-n}(\tau) d\tau. \] (38)
The scaling operator has eigenfunctions
\[ \mathcal{S} \phi^k = L^{\frac{1-k}{2}} \phi^k. \] (39)

From this one obtains
\[ L^{N-n} w_N(\phi) = L^{\frac{1}{2}} n(\phi + \vartheta_n, M(\phi + \vartheta_n)) - L^{\frac{3}{2}} n C_{L^{-N}}. \] (40)
We see now why the counter term \(C_{L^{-N}}\) is needed: the expectation of the random field \((\vartheta_n, M\vartheta_n)\) blows up as \(N \to \infty\) as shown in Lemma 4 and this divergence is the source of the renormalization constant \(m_1\) in (9).

Furthermore, we need to study the dependence of our constructions on the choice of the cutoff function \(\chi\) in (7). To this end, let us define
\[ \chi'_m(s) = \chi(s) - \chi'(L^{2m} s) \] (41)
where the lower cutoff in (36) has been replaced by a different bump function $\chi'$. In the following we will denote by $Y^{(N)}_n$ the kernel $Y^{(N)}_n$ where $\chi_{N-n}$ is replaced by $\chi'_{N-n}$. We also note that, by taking $\chi'(s) = \chi(L^2s)$, one gets $Y^{(N)}_n = Y^{(N+1)}_n$, so by varying $\chi'$ we can also study the dependence and convergence as $N \to \infty$.

We are now ready to state the Lemma which controls the dependence of the covariance $c^{(N)}_n$ on $N$ and $\chi$. See the Appendix for the proof.

**Lemma 4.** Define $m_1 \in \mathbb{R}^3$ by

$$m_1^{(\alpha)} := \left( \sum_{\beta=1}^{3} M_\beta \right) \frac{1}{2\pi^{3/2}} \int_0^\infty s^{-3/2}(1 - \chi(s)^2) ds$$

for $\alpha = 1, 2, 3$. Then

$$\mathbb{E}(\vartheta^{(N)}_n, M\vartheta^{(N)}_n) = L^{N-n}m_1 + \delta^{(N)}_n$$

where $\|\delta^{(N)}_n\|$ is uniformly bounded in $N$ and $n$ where $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^3$. Moreover, let $\delta^{(N)}_n$ be the analog of $\delta^{(N)}_n$, where the lower cutoff function is replaced by $\chi'$. Then

$$\|\delta^{(N)}_n - \delta^{(N)}_n\| \leq Ce^{-cL^{2N}N}\|\chi - \chi'\|_\infty.$$  

The counter term $C_{L^{-N}}$ is then given in this linear approximation as

$$C_{L^{-N}} = L^N m_1$$

and we end up with

$$L^{N-n}w_N(\phi) = u_{n,1}(\phi).$$

where

$$u_{n,1}(\phi) = L^{-\frac{n}{2}}((\phi + \vartheta_n, M(\phi + \vartheta_n)) - L^{N-n}m_1).$$

**4.2. Higher order terms.** The heuristic idea of our proof is now the following. We look for the RG flow in the form

$$w_n = \sum_{i=1}^{k-1} u_{n,i}(\phi) + r_n$$

where $u_{n,i}$ are explicit perturbative contributions and in a suitable norm

$$\|u_{n,i}\| = O(L^{-\frac{n}{2}}), \quad \|r_n\| = O(L^{-\frac{n}{2}})$$

and we expect

$$r_{n-1} = \mathcal{L}r_n + O(L^{-\frac{k+1}{2}}).$$

Moreover, from our analysis of $\mathcal{L}$ we also expect that

$$\|\mathcal{L}r_n\| \leq CL^{\frac{3}{2}}\|\mathcal{L}r_n\| \leq CL^{\frac{3}{2}}L^{-\frac{k}{2}}n = CL^{\frac{3}{2}}L^{-\frac{k}{2}}(n-1)$$

so that (49) should iterate provided we take $k = 4$. Hence, we should find the perturbative contributions to $w_n$ up to order 3.

**Remark 5.** The same heuristic idea works in general for subcritical problems. The dimensionless strength of the nonlinearity is $L^{-Na}$ for some $\alpha > 0$ and the norm of $\mathcal{L}$ is $L^\beta$ for some $\beta > 0$. Then one needs to do perturbation theory up to order $k - 1$ with $k\alpha > \beta$. 

The $u_{n,i}$ may be computed by doing one RG step with scaling factor $L^{N-n}$

$$w_n(\phi) = L^{-\frac{n}{2}}(\phi + \vartheta_n + Y_n * w_n(\phi), M(\phi + \vartheta_n + Y_n * w_n(\phi)))$$ (50)

where we dropped the superscript $N$ also in $Y_n^{(N)}$. We obtain

$$u_{n,2}(\phi) = 2L^{-n}(\phi + \vartheta_n, M(Y_n * u_{n,1}(\phi)))$$

and

$$u_{n,3}(\phi) = L^{-\frac{3}{2}n}(Y_n * u_{n,1}(\phi), M(Y_n * u_{n,1}(\phi)))$$

$$+ 2(\phi + \vartheta_n, M(Y_n * u_{n,1}(\phi))) - m_2 \log L^N - m_3$$

where $m_2$ and $m_3$ are constants to be determined. To write the recursion (49) let us denote $w_n$ by $w$ and $w_{n-1}$ by $w'$ and similarly for the other functions. Then

$$r'(\phi) = Lr(\phi + Y * w'(\phi)) + F(r'(\phi))$$ (51)

with

$$F(r'(\phi)) = u'_1(\phi + Y * w') - u'_1(\phi) - Du'_1(\phi)(Y * (u'_1 + u'_2)) - \frac{1}{2} D^2 u'_1(\phi)(Y * u'_1 + Y * u'_1)$$

$$+ L u_2(\phi + Y * w') - L u_2(\phi) - D L u_2(\phi) Y * u'_1$$

$$+ L u_3(\phi + Y * w') - L u_3(\phi)$$

$$= F_1(r'(\phi)) + F_2(r'(\phi)) + F_3(r'(\phi))$$ (52)

where $D$ is the (Frechet) derivative and on the LHS $w', u'$ are evaluated at $\phi$.

**Remark 6.** Note that $u_i$ are polynomials in $\phi$ so there is no problem in defining the derivative. In Section 6 we’ll see that $w$ is actually analytic.

5. Random fields

The perturbative terms $u_i$ are polynomials in $\phi$ with random coefficients. For the heuristic idea of the proof presented above to work these coefficients should not be too large. For $u_{n,1}$ these random coefficients are the random fields $\vartheta_n(t, x)$ and

$$u_{n,1}(0) = L^{-\frac{n}{2}}(\vartheta_n(t, x), M\vartheta_n(t, x)) - L^{N-n}m_1$$ (54)

In case of $u_{n,2}$ and $u_{n,3}$ we don’t need to consider all the coefficients. Indeed, the discussion of previous section was based on a bound $L^\frac{3}{2}$ for the linearized RG operator. This is indeed its eigenvalue on constants. The next eigenvalue is $L$ on linear functions, the one after $L^\frac{1}{2}$ etc. Thus for $u_{n,2}$ we should be worried only about the constant and linear terms in $\phi$ and for $u_{n,3}$ only about constants. All the other terms should be irrelevant i.e. they should contract under the RG. We will now isolate these relevant terms. Let us expand

$$u_{n,2}(\phi) = u_{n,2}(0) + Du_{n,2}(0)\phi + U_{n,2}(\phi).$$ (55)

We get

$$u_{n,2}(0) = 2L^{-n}(\vartheta_n, MY_n * u_{n,1}(0))$$ (56)

and

$$Du_{n,2}(0)\phi = L^{-n}(\phi, MY_n * u_{n,1}(0)) + L^{-n} \int_{-\infty}^{t} ds \int_{T_n} dy \sigma_n(t, x, s, y) \phi(s, y)$$ (57)
where
\[(\sigma_n(t, x, s, y))_{\alpha\beta} = 4Y_n(t - s, x - y) \sum_{\gamma, \delta, \lambda} \vartheta_{\gamma}(t, x) M^{(\alpha)}_{\gamma\delta} \vartheta_{\lambda}(s, y) M^{(\beta)}_{\lambda\beta}. \tag{58}\]

For the third order term we get
\[u_{n,3}(\phi) = u_{n,3}(0) + U_{n,3}(\phi). \tag{59}\]

with
\[u_{n,3}(0) = L^{-\frac{3}{2}n}[(Y_n * u_{n,1}(0)), M(Y_n * u_{n,1}(0)) \tag{60}\]
\[+(\vartheta_n, M(Y_n * u_{n,2}(0)) - m_2 \log L^N - m_3].\]

Consider now the random fields \(u_{n,i}(0), Du_{n,2}(0)\) with the scaling factor divided out, i.e.
\[3n, i := L^\frac{1}{2}n u_{n,i}(0), \quad D3n, 2 := L^n Du_{n,2}(0). \tag{61}\]

Then \(\vartheta_n, D3n, i, D3n, 2\) belong to the Wiener chaos of white noise of bounded order \((\leq 4)\) and their size and regularity are controlled by studying their covariances, as shown in the Section 8. For finite cutoff parameter \(N\) these noise fields are a.s. smooth but in the limit \(N \to \infty\) they become distribution valued. We estimate their size in suitable (negative index) Sobolev type norms which we now define.

The operator \((-\partial_t^2 + 1)^{-1}\) acts on \(L^2(\mathbb{R})\) by convolution with the function
\[K_1(t) = \frac{1}{2}e^{-|t|}. \tag{62}\]

and the operator \((-\Delta + 1)^{-1}\) on \(L^2(\mathbb{T}_n)\) is convolution with the periodization of (62)
\[K_2(x) = \sum_{i \in \mathbb{Z}} K_1(x + iL^n). \]

Let
\[K(t, x) = K_1(t)K_2(x). \]

Note that convolution with \(K\) is a positive operator in \(L^2(\mathbb{R} \times \mathbb{T}_n)\). We define \(\nu_n\) to be the completion of \(C_0^\infty(\mathbb{R}_+ \times \mathbb{T}_n)\) with the norm
\[||v||_{\nu_n} = \sup_i ||K \ast v||_{L^2(\epsilon_i)} \tag{63}\]
where \(\epsilon_i\) is the unit cube centered at \(i \in \mathbb{Z} \times (\mathbb{Z} \cap \mathbb{T}_n)\). To deal with the bi-local field as \(\sigma_n\) in (58) we define for \(\sigma(t, x, s, y)\) in \(C_0^\infty(\mathbb{R}_+ \times \mathbb{T}_n \times \mathbb{R}_+ \times \mathbb{T}_n)\)
\[||\sigma||_{\nu_n} = \sup_i \sum_j \|K \otimes K \ast \sigma\|_{L^2(\epsilon_i \times \epsilon_j)} \tag{64}\]

Now we can specify the admissible set of noise. Let \(\gamma > 0\) and define the sets of events \(\mathcal{E}_m, m > 0\) in the probability space of the space time white noise \(\Xi\) as follows. Let \(\zeta_n^{(N)}\) denote any fields \(\vartheta_n, D3n, i, D3n, 2\). The first condition on \(\mathcal{E}_m\) is that for all \(N \geq n \geq m\) the following hold:
\[\|h_n\zeta_n^{(N)}\|_{\nu_n} \leq L^{\gamma m}. \tag{65}\]

where \(h_n\) is a smooth indicator of the time interval \([0, \tau_n]\), \(\tau_n = L^{2(n-m)}\) which is introduced to localize in time the flow equation, as we will see in Section 6. More precisely, \(h\) is a smooth bump on \(\mathbb{R}\) with \(h(t) = 1\) for \(t \leq -L^{-2}\) and \(h(t) = 0\) for \(t \geq -\frac{1}{2}L^{-2}\) and set \(h_k(t) = h(t - \tau_k)\) so that \(h_k(t) = 1\) for \(t \leq \tau_k - L^{-2}\) and \(h_k(t) = 0\) for \(t \geq \tau_k - \frac{1}{2}L^{-2}\).
We need also to control the $N$ and $\chi$ dependence of the noise fields $\zeta^{(N)}$. We can study both by varying the lower cutoff in the operator $Y_n^{(N)}$ in (34). We denote by $\zeta^{(N)}_n$ any of the resulting noise fields. Our second condition on $E_m$ is that for all $N \geq n \geq m$ and all cutoff functions $\chi$, $\chi'$ with bounded $C^1$ norm
\[ ||h_n(\zeta^{(N)}_n - \zeta^{(N)}_m)||_{V_n} \leq L^{-\gamma(N-n)}L^{\gamma n}. \] (66)

The final condition concerns the fields $\Upsilon_n * \xi_{n-1}$ entering the RG iteration (28). Note that these fields are $N$ independent and smooth and we are going to impose on them a smoothness condition: for all $n > m$ we demand
\[ ||\Upsilon_n * \xi_{n-1}||_{\Phi_{n-1}} \leq L^{\gamma n}. \] (67)

where the norm is defined in next section. In Section 8 we prove

**Proposition 7.** There exists $\gamma > 0$ such that almost surely $E_m$ holds for some $m < \infty$.

In the following sections we suppose the noise is on $E_m$ and we will control the RG iteration (51) for scales $n \geq m$.

### 6. Banach space setup for the RG map

In this section we set up the RG iteration in suitable functional spaces along the same lines of [10, 11]. Let us first discuss the domain and range of the effective nonlinearities $w_n$. The range of $w_n, r_n$ is dictated by the noise, so we take it to be $V_n$.

In the argument of $w_n$ in (28) $\Upsilon_n * (w_{n-1} + \xi_{n-1})$ is smooth so we take the domain of $w_n(\phi)$ to consist of suitably smooth functions. Let $\Phi_n$ be the space of
\[ \phi : [0, \tau_n] \times T_n \to \mathbb{C} \]

which are $C^2$ in $t$ and $C^2$ in $x$ with $\partial^2_t \phi(0, x) = 0$ for $0 \leq i \leq 2$ and all $x \in T_n$. We equip $\Phi_n$ with the sup norm
\[ ||\phi||_{\Phi_n} := \sum_{i \leq 2, j \leq 2} ||\partial^2_t \partial^2_x \phi||_{\infty}. \]

The following lemma collects some elementary facts on how our spaces tie up with the operators entering the RG:

**Lemma 8.** (a) $\Upsilon_n : V_{n-1} \to \Phi_{n-1}$ and $h_{n-1} \Upsilon_n : V_{n-1} \to V_{n-1}$ are bounded operator with norm $C(L)$.

(b) $\mathfrak{s} : \Phi_{n-1} \to \Phi_n$ and $\mathfrak{s}^{-1} : V_n \to V_{n-1}$ are bounded with
\[ ||\mathfrak{s}|| \leq L^{-\frac{1}{2}}, \quad ||\mathfrak{s}^{-1}|| \leq CL^{\frac{1}{2}}. \]

(c) Let $\phi \in C^{2,2}(\mathbb{R} \times T_n)$ and $v \in V_n$. Then $\phi v \in V_n$ and $||\phi v||_{V_n} \leq C||\phi||_{C^{2,2}}||v||_{V_n}$.

**Proof.** Essentially the same as Lemma 9 in [5].

Consider now our fixed point problem
\[ w_{n-1}(\phi) = S w_n(\phi + \Upsilon_n * \xi_{n-1} + \Upsilon_n * w_{n-1}(\phi)). \] (68)

$w_n$ takes values in the distribution space $V_n \subset \mathcal{D}'(\mathbb{R}_+ \times T_n)$. We want to bound it on the time interval $[0, \tau_n]$ i.e. we need to localize (68) in time. Define
\[ \tilde{w}_n = h_n w_n \]

so that
\[ \tilde{w}_{n-1}(\phi) = h_{n-1} S w_n(\phi + \Upsilon_n * \xi_{n-1} + \Upsilon_n * w_{n-1}(\phi)) \] (69)
One can readily check that \( \Upsilon_n w_{n-1} = \Upsilon_n \tilde{w}_{n-1} \) on the time interval \([0, \tau_{n-1}]\) and that
\[
h_{n-1}S w_n = h_{n-1}S \tilde{w}_n.
\]
Thus (69) can be written as
\[
\tilde{w}_{n-1}(\phi) = h_{n-1}S \tilde{w}_n(\phi + \Upsilon_n \xi_{n-1} + \Upsilon_n \tilde{w}_{n-1}(\phi)).
\]
We will solve (70) in a space of analytic functions which we discuss next. Let \( \mathcal{H}, \mathcal{H}' \)
be Banach spaces and \( B(r) \subset \mathcal{H} \) open ball of radius \( r \). Let \( H^\infty(B(r), \mathcal{H}) \)
denotes the space of analytic functions \( f : B(r) \to \mathcal{H} \) with sup norm which we denote by \( ||f||_{B(r)} \).
We will use the following simple facts that are identical to those of analytic functions
on finite dimensional spaces (see [9]).

(a). Let \( w \in H^\infty(B(r), \mathcal{H}) \) and \( w' \in H^\infty(B(r'), \mathcal{H}'') \). If \( ||w|| < r' \) then \( w' \circ w \in H^\infty(B(r), \mathcal{H}'') \) and
\[
||w' \circ w||_{B(r)} \leq ||w'||_{B(r')}.
\]
(b). Let \( w \in H^\infty(B(r), \mathcal{H}) \) and \( \rho < r \). Then
\[
\sup_{||x|| < \rho} \| Dw(x) \|_{L(H, \mathcal{H}')} \leq (r - \rho)^{-1} ||w||_{B(r)},
\]
where \( L(H, \mathcal{H}') \) denotes the space of bounded linear operators from \( H \) to \( \mathcal{H}' \). Taking \( \rho = \frac{1}{2} r' \), we infer that if \( ||w_1||_{B(r)} \leq \frac{1}{2} \rho \) then
\[
||w' \circ w_1 - w' \circ w_2||_{B(r)} \leq \frac{2}{r'} ||w'||_{B(r')} ||w_1 - w_2||_{B(r)}.
\]
(c). Define \( \delta_k w(x) := w(x) - \sum_{\ell=0}^{k-1} \frac{1}{\ell!} D^\ell w(0)(x) \). Then
\[
||\delta_k w||_{B(ar)} \leq \frac{a^k}{1 - a} ||w||_{B(r)}
\]
for \( 0 \leq a < 1 \).

Furthermore, we infer this important corollary from Lemma 8:

**Proposition 9.** \( S \) maps \( H^\infty(B(R), \mathcal{V}_n) \) into \( H^\infty(B(L^{27} R), \mathcal{V}_{n-1}) \) with norm \( ||S|| \leq C L^{\frac{3}{2}} \). Here \( B(R) \subset \Phi_n \) and \( B(L^{27} R) \subset \Phi_{n-1} \) respectively.

Let now \( \gamma > 0 \) and set \( B_n = B(L^{27n}) \subset \Phi_n \). Then we have

**Proposition 10.** There exist \( L_0 > 0, \gamma_0 > 0 \) so that for \( L > L_0, \gamma < \gamma_0 \) and \( m > m(\gamma, L) \) if \( \Xi \in \mathcal{A}_m \) then for all \( N \leq n \geq m \) the equation (70) has a unique solution \( \tilde{w}_{n-1}^{(N)} \in H^\infty(B_{n-1}, \mathcal{V}_{n-1}) \). These solutions satisfy
\[
||\tilde{w}_{n-1}^{(N)}||_{B_n} \leq L^{-\frac{1}{2} n}
\]
and \( \tilde{w}_{n}^{(N)} \) converge in \( H^\infty(B_n, \mathcal{V}_n) \) to a limit \( \tilde{w}_n \) as \( N \to \infty \). Furthermore, \( \tilde{w}_n \) is independent on the small scale cutoff.

**Proof.** We will drop the tilde from now on so that \( w_n, r_n \) and \( u_{n,i} \) stand for \( \tilde{w}_n \) etc. Also, if no confusion arises we let \( w \) and \( w' \) stand for \( w_n \) and \( w_{n-1} \) respectively. We
start with the perturbative contributions \( u_i \). As a corollary of Lemma 8(c) and (65) we obtain for \( n \geq m \) and \( N \geq n \):

\[
\|u_{n,1}\|_{RB_n} \leq CR^2 L^{(4\gamma - \frac{1}{2})n} \tag{76}
\]

\[
\|u_{n,2}(0) + Du_{n,2}(0)\|_{RB_n} \leq CRL^{(3\gamma - 1)n} \tag{77}
\]

\[
\|u_{n,3}(0)\|_{V_n} \leq CL^{(\gamma - \frac{3}{2})n} \tag{78}
\]

for all \( R \geq 1 \). We used also \( \|h_n\|_{C^{2,2}} \leq C \).

We need to bound the remainder terms \( U_2 \) and \( U_3 \) in (55) and (59). We do this inductively in \( n \). We have

\[
u_2'(\phi) = Lu_2(\phi) + Du_1(\phi)Y_nu_1(\phi) := Lu_2(\phi) + v_2(\phi)
\]

Using Lemma 8(a), (76) and (72) we get

\[
\|v_2\|_{L^2 B_{n-1}} \leq C(L) L^{(\gamma - \frac{3}{2})n} \tag{81}
\]

Let us inductively assume

\[
\|u_{n,2}\|_{B_n} \leq C L^{(7\gamma - 1)n} \tag{79}
\]

Using Proposition 9 and (67) we get the following useful result

\[
\|\mathcal{W}W\|_{L^\frac{1}{2} B_{n-1}} \leq CL^\frac{3}{2} \|W\|_{B_n} \tag{80}
\]

for all \( W \in H^\infty(B_n, V_n) \) since \( B(L^{2\gamma(n-1)} + L^{-n}) \subset B_n \) if \( L > L(\gamma) \). Thus

\[
\|u_2'\|_{L^\frac{1}{2} B_{n-1}} \leq C L^\frac{3}{2} \|u_2\|_{B_n} + C(L) L^{(6\gamma - 1)n} \leq CL^\frac{1}{2} L^{(7\gamma - 1)n}
\]

if \( n > n(\gamma, L) \). Then by (74)

\[
\|U_2'\|_{B_{n-1}} = \|\delta_2 u_2'\|_{B_{n-1}} \leq CL^\frac{3}{2} L^{(7\gamma - 1)n}.
\]

Using (77), the bound (79) follows for \( n-1 \) provided we take \( \gamma \) so that \( \frac{1}{2} + (7\gamma - 1) < 0 \).

For \( u_{n,3} \) we have the recursion

\[
u_3'(\phi) = Lu_3(\phi) + v_3(\phi) \tag{81}
\]

with

\[
v_3(\phi) = \frac{1}{2} D^2 u_1'(\phi)(Y u_1, Y u_1) + Du_1'(\phi)Y u_2 + DL u_2(\phi)Y u_1.
\]

We readily get

\[
\|v_3\|_{L^\frac{1}{2} B_{n-1}} \leq C(L) L^{(\gamma - \frac{1}{2})n}.
\]

The inductive bound

\[
\|u_{n,3}\|_{B_n} \leq CL^{(9\gamma - \frac{3}{2})n} \tag{82}
\]

follows then in the same way as for \( u_2 \), using \( U_3 = \delta_1 u_3 \).

Now we are ready to solve equation (51) by Banach fixed point theorem. Thus consider the map

\[
\mathcal{G}(r') = Lr(\phi + Y * w') + \mathcal{F}(r') \tag{83}
\]

where \( \mathcal{F}(r') \) is given by (53).

We have

\[
\mathcal{F}_1(r')(\phi) = L^{-\frac{n-1}{2}}(2(\phi + \theta, M(\bar{Y} * (u_3' + r'))) + (\bar{Y} * (u_3' + u_3' + r'), M\bar{Y} * (u_2' + u_3' + r'))
\]

so that

\[
\|\mathcal{F}_1(r')\|_{B_{n-1}} \leq C(L) L^{((4\gamma - 2)(n-1) + (2\gamma - \frac{1}{2})(n-1)\|r'\|_{B_{n-1}}).
\]
Next we write
\[
F_2(r') = Lu_2(\phi + U \ast w') - Lu_2(\phi + U \ast u'_1) + Lu_2(\phi + U \ast u'_1) - Lu_2(\phi) - DLu_2(\phi)U \ast u'_1
\equiv F_{2,1}(r') + F_{2,2}(r')
\]  
(84)

Using (73) we obtain
\[
|||F_{2,1}(r')|||_{B_{n-1}} \leq C(L)(L(14\gamma - 2)(n-1) + L(7\gamma - 1)(n-1)|||r'|||_{B_{n-1}}
\]

To bound \(F_{2,2}(r')\) consider the function \(f(z) = Lu_2(\phi + zU \ast u'_1)\) for \(z \in \mathbb{C}\). Since \(Lu_2\) is analytic in \(L^2B_{n-1}\) and
\[
||f + zU \ast u'|||_{B_{n-1}} \leq L^{2\gamma(n-1)} + CL^{(4\gamma - \frac{1}{2})(n-1)}|z|
\]
we get that \(f\) is analytic in the ball \(|z| \leq CL^{(\frac{4\gamma}{2} - 2\gamma)(n-1)}\). Since \(F_{2,2}(r')(\phi) = f(1) - f(0) - f'(0)\) we conclude by a Cauchy estimate
\[
|||F_{2,2}(r')|||_{B_{n-1}} \leq C(L)L^{(15\gamma - 2)(n-1)}.
\]

For \(F_3\) we get using (80) and (73)
\[
|||F_{3}(r')|||_{B_{n-1}} \leq C(L)(L(13\gamma - 2)(n-1) + L(9\gamma - 2)(n-1)|||r'|||_{B_{n-1}}
\]

Consider finally the first term in (83). (80) implies
\[
|||Lr(\cdot + U \ast w')|||_{B_{n-1}} \leq CL^3 |||r'|||_{B_{n}}.
\]

We conclude that by taking \(\gamma\) small enough if
\[
|||r'|||_{B_{n}} \leq L^{-\frac{2\gamma}{n}}
\]
then \(G\) maps the ball \(|||r'|||_{B_{n-1}} \leq L^{-\frac{\gamma}{2}(n-1)}\) to itself. It is now straightforward to check that \(G\) is a contraction in this ball so that by induction in \(n\) (85) holds for all \(n \geq m\).

Let us address the convergence as \(N \to \infty\) and cutoff dependence of \(w_n = w_n^{(N)}\) which can be dealt with together by considering the difference \(w_n - \tilde{w}_n\) where \(w_n^\alpha\) equals \(w_n^{N+1}\) or \(w_n^\alpha\) with a different cutoff. We proceed as with \(w_n\), starting with the following bounds that follow from (66): for all \(n \geq m\) and \(N \geq n\)
\[
||u_{n,1} - u_{n,1}'|||_{RB_n} \leq CR^2 L^{-\gamma(N-n)} L^{(4\gamma - \frac{1}{2})n}
\]
\[
||u_{n,2}(0) + Du_{n,2}(0)\phi - (u_{n,2}'(0) + Du_{n,2}'(0)\phi)|||_{RB_n} \leq CRL^{-\gamma(N-n)} L^{(3\gamma - 1)n}
\]
\[
||u_{n,3}(0) - u_{n,3}'(0)|||_{\mathcal{V}_n} \leq CL^{-\gamma(N-n)} L^{(\gamma - \frac{3}{2})n}
\]
for all \(R \geq 1\). The induction then goes as for \(w_n^{N}\), except for the prefactor \(L^{-\gamma(N-n)}\) in all the bounds. This establishes the convergence of \(w_n^{(N)}\) to a limit that is independent on the short time cutoff.

\[\square\]

7. Proof of Theorem 1

We can now construct the solution \(\varphi^{(\epsilon)}\) of the \(\epsilon\) cutoff equation (11) and consequently \(u^{(\epsilon)}\) in (6). Recall that formally \(u^{(\epsilon)}\) is given on time interval \([0, L^{-2m}]\) by equation (27) with \(n = m\) and \(\varphi_m\) is the solution of equation (23) on time interval \([0, 1]\). Hence we first need to study the \(f\) iteration equation (25) which is equivalent to (29). We study instead of (29) the localized iteration
\[
\tilde{f}_{n-1}^{(N)}(\phi) = h_{n-1} L^{-1} S_{f_{n}^{(N)}}(\phi + U \ast (\tilde{w}_{n-1}^{(N)}(\phi) + \xi_{n-1}))
\]
(89)
for \( f_n = h_n f_n \). Then we can show the following Proposition.

**Proposition 11.** Let \( \tilde{w}_n \in H^\infty(B_n, \mathcal{V}_n) \), \( m \leq n \leq N \) be as in Proposition 10 and \( \phi \in B_n \). Then for \( m \leq n \leq N \) \( f_n = h_n f_n \) converges in \( H^\infty(B_n, \mathcal{V}_n) \) as \( N \to \infty \) to a limit \( \psi_n \), which is independent of the cutoff function.

**Proof.** Let us write
\[
\tilde{f}_n(\phi) = h_n(\phi + \phi_n(\phi)) + g_n(\phi).
\]

Then
\[
g_{n-1} \phi = h_{n-1} L^{-1} S \phi_n + (\hat{w}_{n-1}(\phi + \xi_n)) + h_{n-1} \hat{w}_n(\phi).
\]

Note that the operator \( L^{-1} S \) has norm bounded by \( CL^{1/2} \) and \( \hat{w}_n(\phi) \) has norm bounded by \( CL^{-1/4} n \). Hence we need to extract the leading “marginal” part from \( \hat{w}_n(\phi) \):
\[
g_n = h_n Y_n(\phi) + \hat{w}_n(\phi).
\]

As we will see in Section 8 (Lemma 14), uniformly in \( n \) we have
\[
\mathcal{Y}_n := \sup_{N \geq n} |Y_n| \in L^p(\mathbb{R} \times T_n), \quad \|Y_n - Y_n\|_p \leq CL^{-\lambda(N-n)} \chi' - \chi \|_{\infty}
\]
for \( p < 3/2 \) and some \( \lambda > 0 \). Then thanks to Lemma 8(c) and Young’s inequality we have
\[
\|h_n Y_n(\phi) + \hat{w}_n(\phi)\| \leq C \sum_i \|K \ast Y_n(\phi) + \hat{w}_n(\phi)\|_{L^2(\xi_i)} \leq C \sum_i \|Y_n\|_{L^1(\xi_i)} \|K \ast \hat{w}_n(\phi)\|_{L^2(\xi_i)}
\]
\[
\leq C \sum_i \|Y_n\|_{L^1(\xi_i)} \|K \ast \hat{w}_n(\phi)\|_{L^2(\xi_i)}
\]
\[
\leq C \|\hat{w}_n(\phi)\|_{\mathcal{V}_n}.
\]

Thus \( \|h_n Y_n(\phi) + \hat{w}_n(\phi)\|_{B_n} \leq CL^{1/4 - \lambda} n \) by Proposition 10. Then the iteration of \( b_n(\phi) \) gives easily \( ||b_n(\phi)||_{B_n} \leq L^{-\frac{1}{4} n} \), which implies that \( ||g_n(\phi)||_{B_n} \leq L^{-1/4 n} \).

The convergence and cutoff independence follows from that of \( \tilde{w}_n(\phi) \) proved in Proposition 10.

Moreover, we need also this technical Lemma.

**Lemma 12.** \( \partial_z G_1 \) is a bounded operator from \( \mathcal{V}_n \) to \( \Phi_n \) and \( \partial_z G_1 \ast (h_{n-1}(L^{-2})v) = \partial_z G_1 \ast v \).

**Proof.** As in [5], Lemma 14. \( \square \)

Now we can finally prove our main result: let \( \phi_n \in \Phi_n \) be defined inductively by \( \phi_m = 0 \) and for \( n > m \)
\[
\phi_n = g(\phi_n - 1 + Y_n * (\hat{w}_n(\phi_n - 1) + \xi_n)).
\]

We claim that for all \( m \leq n \leq N \) \( \phi_n \in B_n \) and
\[
\phi_n = \partial_z G_1 \ast (\hat{w}_n(\phi_n) + \xi_n).
\]

Indeed, this holds trivially for \( n = m \) since the RHS vanishes identically on \([0, 1]\). Suppose \( \phi_{n-1} \in B_{n-1} \) satisfies
\[
\phi_{n-1} = \partial_z G_1 \ast (\hat{w}_{n-1}(\phi_{n-1}) + \xi_{n-1}).
\]
Then, first by Lemma 8(b) and (90)
\[ \|\phi_n\|_{\Phi_n} \leq L^{-\frac{1}{2}} \|\phi_{n-1}\|_{\Phi_{n-1}} + C(L)L^{\gamma_n} \leq L^{2\gamma_n} \]
so that \(\phi_n \in B_n\). Second, we have by (92), (90) and Lemma 12
\[ \phi_n = \varrho((\partial_x G_1 + T) \ast (\tilde{w}_n^{\langle N\rangle}(\phi_{n-1}) + \xi_{n-1})) = \partial_x G_1 \ast (\tilde{w}_n^{\langle N\rangle}(\phi_n) + \xi_n). \]  
Equation (93)
Since \(\phi_m = 0\), from (89) we have
\[ \tilde{f}_m^{\langle N\rangle}(0) = h_m \gamma^{-1} \tilde{f}_{m+1}^{\langle N\rangle}(\phi_{m+1}) = h_m \gamma h_{m+1}(L^2) \gamma^{-2} \tilde{f}_{m+2}^{\langle N\rangle}(\phi_{m+2}) = h_m \gamma^{-2} \tilde{f}_{m+2}^{\langle N\rangle}(\phi_{m+2}), \]
then, iterating we get
\[ \tilde{f}_m^{\langle N\rangle}(0) = h_m \gamma^{-1-m} \tilde{f}_N^{\langle N\rangle}(\phi_N) = h_m \gamma h_N(L^{-2(N-m)}) \gamma^{-1-m} \phi_N = h_m \gamma^{-1-m} \phi_N \]
Equation (94)
since \(\tilde{f}_N^{\langle N\rangle}(\phi_N) = \phi_N\) by (31). Now \(\phi_N \in B_N\) solves (91) with \(\tilde{w}_N^{\langle N\rangle}(\phi) = h_N w_N^{\langle N\rangle}(\phi)\) with \(w_N^{\langle N\rangle}\) given by (30). Since \(h_N = 1\) on \([0, \tau_N, L^{-2}]\) we obtain
\[ \phi_N = \partial_x G_1 \ast (\tilde{w}_N^{\langle N\rangle}(\phi_N) + \xi_N) = \partial_x G_1 \ast (w_N^{\langle N\rangle}(\phi_N) + \Xi). \]
To take the limit \(N \to \infty\) we will use (94): defining \(\eta^{\langle N\rangle} := \gamma^{-N} \phi_N\), then we get
\[ \eta^{\langle N\rangle} = \gamma^{-m} \tilde{f}_m^{\langle N\rangle}(0) \]
on the time interval \([0, \frac{1}{2} L^{-2m}] \subset [0, 1]\).
By Proposition 11 \(\tilde{f}_m^{\langle N\rangle}(0)\) converges in \(V_m\) to a limit \(\psi_m\) which is independent of the short distance cutoff. Convergence in \(V_m\) implies convergence in \(D'(\{0, 1\} \times \mathbb{T}_m)\). The claim follows from continuity of \(\gamma^{-m} : D'([0, 1] \times \mathbb{T}_m) \to D'([0, L^{-2m}] \times \mathbb{T}_1)\) and from the fact that convergence of \(\eta = \partial_x u\) implies convergence of \(u\).

8. Proof of Proposition 7
We now need to show that for some \(\gamma > 0\) the conditions defining the set \(E_m\) hold almost surely for some \(m < \infty\). To do this, as in [5] the strategy is to control the covariances of the various fields in (61) and establish enough regularity for them.
We will deduce Proposition 7 from a covariance bound for the fields in (61). Let \(\zeta_n^{\langle N\rangle}(t, x)\) or \(\zeta_n^{\langle N\rangle}(t, x, s, y)\) be any of the fields in (61). Let
\[ K_n(t', t, x) = e^{\frac{1}{2} \text{dist}(t', I_n)} K(t' - t, x) h_n(t) \]
Equation (95)
where \(I_n = [0, L^{2(n-m)}]\) and define
\[ \rho_n^{\langle N\rangle} = K_n \zeta_n^{\langle N\rangle} \text{ or } \rho_n^{\langle N\rangle} = K_n \otimes K_n \zeta_n^{\langle N\rangle}. \]
Then
\[ \|K_n \zeta_n^{\langle N\rangle}\|_{L^2(\zeta_i)} \leq C e^{-\frac{1}{2} \text{dist}(i_0, I)} \|\rho_n^{\langle N\rangle}\|_{L^2(\zeta_i)}. \]
Equation (96)
where \(i_0\) is the time component of the center of the cube \(\zeta_i\). From now on in the random fields we will drop the superscript \((N)\) referring to the ultraviolet cutoff and we recall that \(\| \cdot \|\) indicates the euclidean norm for a three-dimensional vector and the Hilbert-Schmidt norm for \(3 \times 3\)-matrix. The following proposition proved in Section 8.1 provides bounds for the covariance of \(\rho_n\).
Proposition 13. There exist renormalization constants $m_1, m_2, m_3 \in \mathbb{R}^3$ and $\lambda > 0$ such that for all $0 \leq n \leq N < \infty$ and for some constant $0 < c < \frac{1}{2}$

$$
E\|\rho_n(t, x)\|_2^2 \leq C
$$

$$
E\|\rho'_n(t, x) - \rho_n(t, x)\|_2^2 \leq CL^{-\nu(N-n)}\|\chi - \chi'\|_{\infty}
$$

$$
E\|\rho_n(t, x, s, y)\|_2^2 \leq Ce^{-c(|t-s|+|x-y|)}
$$

$$
E\|\rho'_n(t, x, s, y) - \rho_n(t, x, s, y)\|_2^2 \leq CL^{-\lambda(N-n)}e^{-c(|t-s|+|x-y|)}\|\chi - \chi'\|_{\infty}
$$

where $\rho'_n = \hat{K}\zeta'_n$, i.e. we replace the lower cutoff function $\chi$ by a $\chi'$. 

Now we can prove Proposition 7: we recall that we want to show that there exist $0 \leq m < \infty$ such that the event $\mathcal{E}_m$ holds almost surely, where $\mathcal{E}_m$ is the event such that bounds (65), (66) and (67) hold for any $m \leq n \leq N$. By using the same strategy as in [5] based on the bounds in [12, 13], one can see that Proposition 13 implies the following bounds for the random fields $\zeta_n^{(N)}$ in (61) for all $p > 1$

$$
P(||\zeta_n^{(N)}||_{\gamma_n} \geq L^{\gamma_n}) \leq CL^{-2m}L^{(3-2\gamma)p}\nu
$$

$$
P\left(||\zeta_n^{(N)} - \zeta_n^{(N)}||_{\gamma_n} \geq L^{-1/2}\cdot\gamma(N-n)\cdot L^{\gamma_n}\right) \leq CL^{-p\nu(N-n)}L^{(3-2\gamma)p}\nu L^{-2m}
$$

Furthermore, to deal with the last condition on $\mathcal{E}_m$ in (67), we note that $\zeta := \Upsilon_n * \xi_{n-1}$ is a Gaussian field with covariance

$$
E\zeta(t', x')\zeta(t, x) = -\Delta x' \int_0^\infty H_n(t' - t + 2s, x' - x)\chi(t' - t + s)\chi(s)ds
$$

where $\chi$ is smooth with support in $[L^{-2}, 2]$. $E\zeta(t', x')\zeta(t, x)$ is smooth, compactly supported in $t' - t$ and exponentially decaying in $x' - x$. We get then by standard Gaussian estimates [12] for $0 \leq j \leq 2$ and $0 \leq j' \leq 2$ and for some $c(L) > 0$

$$
P\left(\sup_{\alpha}||\partial_t'\partial_x' \Upsilon_n * \xi_{n-1})_{\alpha}||_{L^{\infty}(\Omega_i)} > R\right) \leq Ce^{-c(L)R^2}
$$

and thus

$$
P\left(\sup_{\alpha}||\Upsilon_n * \xi_{n-1})_{\alpha}||_{\Phi_n} > L^{2\gamma_n}\right) \leq CL^{-2m}L^{3n}e^{-c(L)\gamma_n^\nu}\nu.
$$

The bounds (101), (102) and (104) implies that $P(\mathcal{E}_m^c) \leq CL^{-2m}$, then Proposition 7 follows from Borel-Cantelli Lemma.

8.1. Proof of Proposition 13. We will now study the random fields in (61), i.e.

$$
\zeta_n \in \{\vartheta_n, \mathfrak{m}_{n,i}, D_3^n, 2\}
$$

that enter the probabilistic estimates.

Consider first their expectations. Setting $z = (t, x)$ and using Lemma 4, the first one gives $E_{\delta_n,1} = \delta_n \leq C$, while for the second order fields we have

$$
E_{\delta_n,1} = \delta_n, \quad E\sigma_{n,\alpha,\beta} \Upsilon_n(z, z') = m_{n,\alpha,\beta}Y_n(z - z')\Upsilon_n(z - z')
$$

and finally for the third order field we get

$$
E_{\delta_n,3} = 8M_1 \int dz_1dz_2Y_n(z_2)Y_n(z_2 - z_2)\Upsilon_n(z_1 - z_2)\theta(t_1 - t_2)\Upsilon_n(z_1)
$$

$$
+ 2M_2 \int dz_1dz_2Y_n(z_1)Y_n(z_2)\Upsilon_n(z_1 - z_2)^2 - m_2 \log L^N - m_3
$$
where \( \theta(t) = 1_{t \geq 0} \) is the Heaviside function and

\[
m_{\alpha\beta} = \sum_{\gamma, \delta} M_{\gamma\delta}^{(\alpha)} M_{\gamma\beta}^{(\delta)}
\]

(107)

\[
(M_1)_\alpha = \sum_{\beta_1, \beta_2, \beta_3, \beta_4} M_{\beta_1, \beta_2}^{(\alpha)} M_{\beta_3, \beta_4}^{(\beta_1)} M_{\beta_1, \beta_3}^{(\beta_2)} M_{\beta_2, \beta_4}^{(\beta_4)}
\]

(108)

\[
(M_2)_\alpha = \sum_{\beta_1, \beta_2, \beta_3, \beta_4} M_{\beta_1, \beta_2}^{(\alpha)} M_{\beta_3, \beta_4}^{(\beta_1)} M_{\beta_1, \beta_3}^{(\beta_2)} M_{\beta_2, \beta_4}^{(\beta_4)}
\]

(109)

Define the random field

\[
\omega_{\alpha\beta} := \vartheta_{\alpha\beta} \vartheta - E \vartheta_{\alpha} \vartheta_{\beta}
\]

(110)

(here and below \( \vartheta = \vartheta_{(N)}^{(N)} \)). Then the local fields \( \zeta_n \) are linear combinations of their expectations and the following random fields

\[
\vartheta_{\alpha}, \ \omega_{\alpha\beta}, \ Y_n * \omega_{\alpha\beta}, \ Y_n * \omega_{\alpha\beta} Y_n * \omega_{\gamma\delta} - E Y_n * \omega_{\alpha\beta} Y_n * \omega_{\gamma\delta},
\]

(111)

\[
\vartheta_{\alpha} Y_n * \omega_{\beta\gamma}, \ \vartheta_{\alpha} Y_n * (\vartheta_{\beta} Y_n * \omega_{\gamma\delta}) - E \vartheta_{\alpha} Y_n * (\vartheta_{\beta} Y_n * \omega_{\gamma\delta})
\]

(112)

where we used \( Y_n * \delta_n = 0 \), while for the bi-local fields we need to consider

\[
Y_n(z - z') \vartheta_{\alpha}(z) \vartheta_{\beta}(z') - E Y_n(z - z') \vartheta_{\alpha}(z) \vartheta_{\beta}(z')
\]

(113)

To get the covariance estimates for the fields (109), (110), (111) claimed in Proposition 13 we need to introduce the mixed covariance \( C_n'(z) \) such that

\[
\delta_{\alpha\beta} C_n'(z) := E \vartheta_{\alpha}'(z) \vartheta_{\beta}(0)
\]

(114)

where, as before, the primed kernels and fields have the lower cutoff \( \chi' \). Furthermore, let us define

\[
C_n(z) := \sup_{N \geq n} |C_n'(z)|
\]

\[
\delta C_n(z) := |C_n'(z) - C_n(z)|
\]

\[
Y_n(z) := \sup_{N \geq n} |Y_n(z)|
\]

\[
\delta Y_n(z) := |Y_n'(z) - Y_n(z)|
\]

The regularity of these kernels is summarized in the following Lemma proven in the Appendix.

**Lemma 14.**  
(a) For \( p < 3 \) and uniformly in \( n \) one has \( C_n \in L^p(\mathbb{R} \times T_n) \) and

\[
\|\delta C_n\|_p^p \leq C L^{-\lambda_p(N-n)} \|\chi - \chi'\|_{\infty}
\]

(115)

for some \( \lambda_p > 0 \).

(b) For \( p < \frac{3}{2} \) and uniformly in \( n \) one has \( Y_n \in L^p(\mathbb{R} \times T_n) \) and

\[
\|\delta Y_n\|_p^p \leq C L^{-\lambda_p(N-n)} \|\chi - \chi'\|_{\infty}
\]

(116)

for some \( \lambda_p > 0 \).

Having these technical tools at hand, we can finally start to show the covariance estimates.
8.2. Fields (109) and (111). For \( z = (t, x) \) we will use the norm \( |z| = |t| + |x| \) and we will drop the subscript \( n \) from the random fields and kernels. From the definition of the smoothing kernel \( \tilde{K} \) we note that

\[
\tilde{K}(z, z') \leq C e^{-\frac{1}{2} |z - z'|}, \quad \partial_z \tilde{K}(z, z') \leq C e^{-\frac{1}{2} |z - z'|}, \quad \partial_{z'} \tilde{K}(z, z') \leq C e^{-\frac{1}{2} |z - z'|}.
\]

(115)

Defining

\[
X(z_1 - z_2) := \mathbb{E} \zeta(z_1) \zeta(z_2)
\]

we then get

\[
\mathbb{E} \|\rho(z)\|^2 = \int dz_1 dz_2 \tilde{K}(z, z_1) \tilde{K}(z, z_2) X(z_1 - z_2) \leq C \|X\|_1
\]

(116)

i.e. it suffices to bound the \( L^1 \)-norm of the covariance. We will use repeatedly the Young inequality in the form

\[
\|f_1 * f_2 * \cdots * f_m\|_p \leq \prod_{i=1}^m \|f_i\|_{p_i}
\]

(117)

if \( n - 1 + \frac{1}{p} = \sum \frac{1}{p_i} \) where \( 1 \leq p, p_i \leq \infty \). We consider now the fields one by one.

(i) For \( \zeta = \partial_\alpha \) we have \( \|X\|_1 \leq C \|C\|_1 \).

(ii) For \( \zeta = \omega_{\alpha\beta} \) we have \( \|X\|_1 \leq C \|C\|_2^2 \).

(iii) For \( \zeta = Y * \omega_{\alpha\beta} \) let \( Y^\dagger(z) = Y(-z) \). Then \( X = CY * C \). By Young inequality

\[
\|X\|_1 \leq C \|Y\|_1 \|C\|_2 \|C\|_1 \leq C \|Y\|_1 \|C\|_2 \|C\|_1
\]

(iv) For \( \zeta = Y * \omega_{\alpha\beta} * Y * \omega_{\gamma\delta} = \mathbb{E} Y * \omega_{\alpha\beta} Y * \omega_{\gamma\delta} \) we get

\[
\int dz \mathbb{E} \zeta(z) \zeta(0) \leq C \int dz dz_1 \ldots dz_4 \mathcal{V}(z - z_1) \mathcal{V}(z - z_2) \mathcal{V}(z - z_3) \mathcal{V}(z - z_4)
\]

\[
\times [\mathcal{C}(z_1 - z_3) \mathcal{C}(z_2 - z_4) (\mathcal{C}(z_1 - z_4) \mathcal{C}(z_2 - z_3) + \mathcal{C}(z_1 - z_2) \mathcal{C}(z_3 - z_4)) + \mathcal{C}^2(z_1 - z_3) \mathcal{C}^2(z_2 - z_4) + \mathcal{C}^2(z_1 - z_4) \mathcal{C}^2(z_2 - z_3)]
\]

Using the trivial inequality

\[
2|ab| \leq a^2 + b^2
\]

(119)

with \( a, b \in \mathbb{R} \) for the products of \( \mathcal{C} \), we obtain

\[
\int dz \mathbb{E} \zeta(z) \zeta(0) \leq C \int dz dz_1 \ldots dz_4 \mathcal{V}(z - z_1) \mathcal{V}(z - z_2) \mathcal{V}(z - z_3) \mathcal{V}(z - z_4)
\]

\[
\times [\mathcal{C}(z_1 - z_3) \mathcal{C}(z_2 - z_4) (\mathcal{C}(z_1 - z_4) \mathcal{C}(z_2 - z_3) + \mathcal{C}(z_1 - z_2) \mathcal{C}(z_3 - z_4)) + \mathcal{C}^2(z_1 - z_3) \mathcal{C}^2(z_2 - z_4) + \mathcal{C}^2(z_1 - z_4) \mathcal{C}^2(z_2 - z_3)]
\]

(120)

Note that \( \mathcal{C}_n \in L^p \) with \( p < \frac{3}{2} \) thanks to Lemma 14, so by Young inequality one can see that the first two terms in (120) are bounded by

\[
C\|\mathcal{C}(\mathcal{V} * \mathcal{C}) (\mathcal{V} * (\mathcal{C}^2(\mathcal{V} * \mathcal{V} * \mathcal{C})))\|_1 \leq C\|\mathcal{V} * \mathcal{C}\|_2 \|\mathcal{V} * \mathcal{C}\|_2 \|\mathcal{V} * \mathcal{C}\|_2 \leq C\|\mathcal{V}\|_1 \|\mathcal{C}\|_2 \|\mathcal{V}\|_2 \|\mathcal{C}\|_2 \|\mathcal{V}\|_\infty \leq C\|\mathcal{V}\|_1 \|\mathcal{C}\|_2 \|\mathcal{V}\|_2 \|\mathcal{C}\|_2 \]
while the third and fourth term in (120) are bounded by
\[ C\|\mathcal{C}(\mathcal{Y}^* \mathcal{Y})\|_1 \|C^2 \ast (\mathcal{Y}(\mathcal{Y}^* \mathcal{C}))\|_1 \leq C\|\mathcal{C}\|_2 \|\mathcal{Y}^* \mathcal{Y}\|_2 \|C^2 \|_1 \|\mathcal{Y}(\mathcal{Y}^* \mathcal{C})\|_1 \]
(121)
and the last contributions are bounded by
\[ C\|\mathcal{Y} \ast \mathcal{Y} \ast \mathcal{Y}^* \mathcal{Y} \|_2 \|C^2 \ast \mathcal{Y} \|_2 \leq C\|\mathcal{Y}\|_1^2 \|\mathcal{Y}\|_3^2 \|C^2\|_1^2. \]

(v) Next we consider the bi-local field \( \zeta(z_1, z_2) = \mathcal{Y}(z_1 - z_2)(\partial_\alpha(z_1) \partial_\beta(z_2) - \delta_{\alpha\beta} \mathcal{C}(z_1 - z_2)) \). Then we have
\[ \mathbb{E}\zeta(z_1, z_2)\zeta(z_3, z_4) \leq C\mathcal{Y}(z_1 - z_2)\mathcal{Y}(z_3 - z_4)(\mathcal{C}(z_1 - z_3)\mathcal{C}(z_2 - z_4) + \mathcal{C}(z_1 - z_4)\mathcal{C}(z_2 - z_3)) \]
so that
\[ C\|\mathcal{Y} \ast \mathcal{Y} \ast \mathcal{C} \|_1 \|\mathcal{C}\|_1 \leq C\|\mathcal{Y}\|_1^2 \|\mathcal{Y}\|_1 \|\mathcal{C}\|_1^2 \]
where \( \mathcal{Y}(z) := e^{c|z|} \mathcal{Y}(z) \) is in \( L^p \) with \( p < \frac{3}{2} \).

8.3. Fields (110) and (105). We observe that in the above covariance estimates, the Young inequality trick requires all the kernels to be at least in \( L^1(\mathbb{R} \times \mathbb{T}_n) \). Unfortunately in the fields (110) and (105) the kernel \( J_n(z) := Y_n(z)\mathcal{C}_n(z) \) will appear and it is easy to see that \( \|J_n\|_1 \) diverges logarithmically as \( N \to 0 \), so Young inequality cannot be applied as before.

The following Lemma shows some properties of \( J_n \) which are crucial to overcome this problem. Its proof can be found in the Appendix.

**Lemma 15.** (a) We have
\[ J_n(z) = \partial_x W_n(z) + j_n(z) \]
where \( W_n \) is in \( L^1(\mathbb{R} \times \mathbb{T}_n) \) uniformly in \( n, N \) and
\[ |j_n(z)| \leq Ce^{-|z|}1_{[0,2]}(t). \]

(b) The function \( Z_n := Y_n \ast J_n \) is in \( L^1(\mathbb{R} \times \mathbb{T}_n) \) uniformly in \( n, N \).

(c) \( \|W_n - W_n'\|_1 \leq C L^{-\lambda(N-n)} \) for some \( \lambda > 0 \), idem for \( j_n \) and \( Z_n \).

(d) Let be \( \epsilon = L^{-2(N-n)} \), then
\[ \|W_n(z) - \mathcal{C}_n(z)\|_1 \leq C(\epsilon^{-2}e^{-c|x|/\epsilon}1_{[\frac{1}{4}c,2 \frac{1}{2}])(t) + e^{-c|x|}1_{[\frac{1}{4}c,2 \frac{1}{2}])(t)} \]
\[ |Y_n(z) - 2\partial_x \mathcal{C}_n(z)| \leq C(\epsilon^{-2}e^{-c|x|/\epsilon}1_{[\frac{1}{4}c,2 \frac{1}{2}])(t) + e^{-c|x|}1_{[\frac{1}{4}c,2 \frac{1}{2}])(t)}. \]

In practice Lemma 15 guarantees that the nasty kernel \( J_n \) is actually a gradient of an \( L^1 \)-function, up to to a smooth correction. By an integration by parts this property will allow us to move the gradient and make it act on the smoothing kernels \( \tilde{K} \), so that we can still use the Young inequality. Moreover, we point out that the kernel \( Z_n \) will
appear in the last fields in (110). Finally, item (d) in Lemma 15 will be employed to study the divergence of $\mathbb{E}_{\tilde{\eta}_{n_3}}$ in (106).

In the following we will neglect the remainder term $j(z)$, since its contributions can be easily bounded as we have done for the fields in (109).

(vi) For $\zeta = \vartheta_\alpha Y * \omega_{\beta\gamma}$ we have

$$
\mathbb{E}\|\rho(z)\|^2 \leq C \left| \int dz_1 dz_2 \mathcal{K}(z, z_1) \mathcal{K}(z, z_2) J(z_1 - z_3) J(z_2 - z_4) \mathcal{C}(z_3 - z_4) \right| + C \int dz_1 dz_2 \mathcal{Y}(z - z_1) \mathcal{Y}(-z_2) [\mathcal{C}(z) C^2(z_1 - z_2) + \mathcal{C}(z - z_2) \mathcal{C}(z_1) \mathcal{C}(z_1 - z_2)]
$$

Using (119) we get

$$
\mathbb{E}\|\rho(z)\|^2 \leq C [[W_1||C \ast W_1| + ||C \ast \mathcal{Y}||_2^2 + \mathcal{C}(C \ast \mathcal{Y}) \ast \mathcal{Y}||_2 ||C||_2]] \leq C [[W_1||C||_1^2 + ||\mathcal{Y}||_1 ||C||_2 ||\mathcal{C}||_2 ||\mathcal{C}||_2 ||\mathcal{Y}||_4 ||\mathcal{C}||_4 ||\mathcal{C}||_4 ||\mathcal{C}||_4].
$$

For the last field in (110), i.e. $\vartheta_\alpha Y * (\vartheta_\beta Y * \omega_{\gamma\delta}) - \mathbb{E}\vartheta_\alpha Y * (\vartheta_\beta Y * \omega_{\gamma\delta})$ it is convenient to perform an expansion in terms of Wick polynomials to keep track of the several contributions involved in the covariance (see [14] for a recent review about Wick polynomials). In our case the “elementary” fields $\vartheta_n$ are Gaussian and with vanishing expectation value, so the combinatorics of the Wick expansion will be quite simple. Noting that $\omega_{\alpha\beta} = :\vartheta_\alpha \vartheta_\beta:$, the random fields turns out to be a linear combination of the following terms

$$
: \vartheta_\alpha Y * (\vartheta_\beta Y * \vartheta_\gamma \vartheta_\delta):, \quad Z * \omega_{\alpha\beta}, \quad : \vartheta_\alpha Z * \vartheta_\beta:.
$$

(vii) In $\zeta = :\vartheta_\alpha Y * (\vartheta_\beta Y * \vartheta_\gamma \vartheta_\delta)$: there is no $J$ appearing, so we can just estimate the corresponding $\|X\|_1$ which unfortunately has many terms:

$$
\|X\|_1 \leq C \int d z_1 d z_2 \cdots d z_4 \mathcal{Y}(z - z_1) \mathcal{Y}(z_1 - z_2) \mathcal{Y}(-z_3) \mathcal{Y}(z_3 - z_4)
$$

$$
\times [\mathcal{C}(z) \mathcal{C}(z_1 - z_3) C^2(z_2 - z_4) + \mathcal{C}(z - z_3) \mathcal{C}(z_1) C^2(z_2 - z_4)]
$$

$$
+ \mathcal{C}(z) \mathcal{C}(z_1 - z_3) C^2(z_2 - z_4) + \mathcal{C}(z - z_3) \mathcal{C}(z_1 - z_4) \mathcal{C}(z_2) \mathcal{C}(z_2 - z_4)
$$

$$
+ \mathcal{C}(z - z_3) \mathcal{C}(z_1) C(z_2 - z_4) \mathcal{C}(z_2 - z_4) + \mathcal{C}(z - z_4) \mathcal{C}(z_1 - z_3) \mathcal{C}(z_2) \mathcal{C}(z_2 - z_4)
$$

$$
+ \mathcal{C}(z - z_4) \mathcal{C}(z_1 - z_4) \mathcal{C}(z_2 - z_3) \mathcal{C}(z_2).
$$

Using (119) we get

$$
\|X\|_1 \leq C [[\mathcal{C}(\mathcal{Y} \ast \mathcal{Y} \ast \mathcal{C}^2) ||_4^2 ||\mathcal{C} \ast \mathcal{Y}||_4 + ||\mathcal{C} \oplus \mathcal{Y}||_4 \mathcal{Y} \ast (\mathcal{Y} \ast \mathcal{C}) (\mathcal{Y} \ast \mathcal{Y} \ast \mathcal{C}^2)||_2]
$$

$$
+ ||\mathcal{C} \ast \mathcal{Y}||_4 \mathcal{Y} \ast (\mathcal{Y} \ast \mathcal{C}) (\mathcal{Y} \ast \mathcal{C}^2)||_2
$$

$$
+ ||\mathcal{C} \ast \mathcal{C}||_2 \mathcal{Y} \ast (\mathcal{C} \ast \mathcal{Y})^2||_2 + ||\mathcal{Y}||_1 ||(\mathcal{C} \ast \mathcal{Y}) (\mathcal{Y} \ast (\mathcal{C}^2 \ast \mathcal{Y})))||_1
$$

$$
+ ||\mathcal{Y}||_1 ||(\mathcal{Y} \ast \mathcal{C}^2) (\mathcal{Y} \ast (\mathcal{C} \ast \mathcal{Y}))||_1 + ||\mathcal{C} \ast \mathcal{Y}||_2 ||\mathcal{Y} \ast \mathcal{C} \ast \mathcal{Y})||_2
$$

$$
\leq C [[\mathcal{Y}||_4^2 ||\mathcal{C}||_2^2 + ||\mathcal{Y}||_1 ||\mathcal{Y}||_4^3 ||\mathcal{C}||_2^3 ||\mathcal{C}||_2^2 ||\mathcal{C}||_2^2].
$$
(viii) For \( \zeta = Z \ast \omega_{\alpha \beta} \) we have by Lemma 15(b)
\[
\|X\| \leq C\|Z \ast C^2\|_1 \|Z\|_1 \leq C\|Z\|_1^2 \|C^2\|_1.
\] (126)

(ix) For \( \zeta = \vartheta_\alpha Z \ast \vartheta_\beta \): again by Lemma 15(b) we have
\[
\|X\| \leq C\|Z \ast C^2\|^2_2 \leq C\|Z\|_1^2 \|C\|_2^2.
\] (127)

(x) For \( \zeta = \int dz_1 dz_2 Y(z - z_1) Y(z_1 - z_2) C(z - z_2) :\vartheta_\alpha(z_1) \vartheta_\beta(z_2) \): by (119) we have
\[
\|X\|_1 \leq C \int dz_1 \cdots dz_4 Y(z - z_1) Y(z_1 - z_2) C(z - z_2) Y(z_3 - z_4) C(-z_4)
\times [C^2(z_1 - z_3) + C^2(z_2 - z_4) + C^2(z_1 - z_4) + C^2(z_2 - z_3)]
\leq C\|Y\|_4 \|C\|_2^2.
\]

(xi) We still need to bound \( \zeta(z_1 - z_2) = E\sigma_{\alpha \beta}(z_1, z_2) = m_{\alpha \beta} J(z_1 - z_2) \). Using Lemma 15(a) and the same strategy used for (v) we get
\[
e^{\|z_1 - z_2\| \rho(z_1, z_2)} \leq C\|\tilde{W}\|_1^2
\] (128)
where \( \tilde{W}(z) = e^{\|z\| W(z)} \in L^1(\mathbb{R} \times T_n) \).

We observe that the estimates (98) and (100) are obtained as the bounds (97) and (99) derived above by using Lemma 4, (113), (114) and Lemma 15(c).

8.4. Third order renormalization. So we are left with the analysis of the expectation \( \mathbb{E}_{\beta_3} \) which will allow us to determine the renormalization constants \( m_2, m_3 \):
\[
\mathbb{E}_{\beta_3} = 2M_2 \int (Y \ast C^2)(z) Y(z) dz + 8M_1 \int (Y \ast \theta J)(z) C(z) dz - m_2 \log L^N - m_3
\]
\[
= 4M_2 \int (Y \ast \theta C^2)(z) Y(z) dz + 8M_1 \int (Y \ast \theta J)(z) C(z) dz - m_2 \log L^N - m_3
\]
where \( \theta C^2(z) = \theta(t) C^2(z) \) and similarly for \( \theta J \). Let us call \( A = \partial_x C - \frac{1}{2} Y \) and \( B = W - C^2 \). Using Lemma 15 and an integration by parts in the second term we get
\[
\mathbb{E}_{\beta_{3a,3}} = 8M_1 \int [(Y \ast \theta j)(z) C(z) - (Y \ast \theta C^2)(z)] A(z) - (Y \ast \theta B)(z) C(z)] dz - m_3
+ 4(M_2 - M_1) \int (Y \ast \theta C^2)(z) Y(z) dz - m_2 \log L^N
\] (129)
For the first term we use the bounds in Lemma 15 to get
\[
\left| \int [(Y \ast \theta j)(z) C(z) - (Y \ast \theta C^2)(z)] A(z) - (Y \ast \theta B)(z) C(z)] dz \right| \leq C.
\] (130)
Let us now study the second term in (129) which is the divergent one. In Fourier space we have
\[
\int (Y \ast \theta \hat{\mathcal{C}}^2)(z)Y(z)dz
\]
\[
= \int_0^\infty dt \int dp \hat{Y}(t, -p) \int_0^t ds \hat{Y}(t - s, p) \int dq \hat{\mathcal{C}}(s, p + q) \hat{\mathcal{C}}(s, q)
\]
\[
= \int_0^\infty dt \int dp p^2 e^{-tp^2} \int_0^t ds \chi(t - s) e^{-(t-s)p^2} \int dq e^{-s(q^2 + (p+q)^2)}
\]
\times h_\epsilon(s, \sqrt{s}(p + q)) h_\epsilon(s, \sqrt{s}q)
\]
where
\[
h_\epsilon(t, p) = p^2 \int_0^\infty ds e^{-2sp^2 \epsilon} \chi((1 + \epsilon)t) \chi(\epsilon t)
\]
and \(0 \leq h_\epsilon < \frac{1}{4}\) uniformly on \(\mathbb{R}_+ \times \mathbb{R}\). Let us define \(\mu_\epsilon\) as
\[
\mu_\epsilon := \frac{1}{4} \int_0^2 dt \int dp p^2 e^{-tp^2} \int_0^t ds e^{-(t-s)p^2} \textbf{1}_{[\epsilon^2, 2]}(t - s) \int dq e^{-s(q^2 + (p+q)^2)}.
\]
We get that
\[
0 \leq \mu_\epsilon - \int (Y \ast \theta \hat{\mathcal{C}}^2)(z)Y(z)dz
\]
\[
\leq \frac{1}{4} \int_0^2 dt \int dp p^2 e^{-tp^2} \int_0^t ds e^{-(t-s)p^2} \int dq e^{-s(q^2 + (p+q)^2)}
\]
\times [\textbf{1}_{[\epsilon^2, 2]}(t)(\textbf{1}_{[\epsilon^2, 2s]}(t) + \textbf{1}_{[1, 2]}(t)) + \textbf{1}_{[\epsilon^2, 2]}(t - s)(\textbf{1}_{[\epsilon^2, 2s]}(t) + \textbf{1}_{[1, 2]}(t))] \leq C
\]
(134)

Let us also define \(\tilde{\mu}_\epsilon := \frac{\pi}{4\sqrt{3}} \log \epsilon^{-1}\), then by an explicit computation one gets
\[
\lim_{\epsilon \to 0}(\tilde{\mu}_\epsilon - \mu_\epsilon) = \mathcal{O}(1).
\]
(135)

Therefore, we can identify the universal renormalization constant \(m_2\) as
\[
m_2 = 4(\mathcal{M}_2 - \mathcal{M}_1) \frac{\tilde{\mu}_\epsilon}{\log \epsilon^{-1}} = \frac{\pi}{\sqrt{3}} (\mathcal{M}_2 - \mathcal{M}_1).
\]
(136)

Finally, for the \(\chi\)-dependent renormalization constant \(m_3\), let be \(\nu_\epsilon := \mathbb{E}_{\delta n, 3} - m_2 \log \epsilon^{-1}\); from (126), (134) and (135) we know that \(|\nu_\epsilon| \leq C\) and by bounds similar to ones in Lemma 15 comparing different cutoffs one can see that \(\nu_\epsilon\) is a Cauchy sequence. Therefore, in the end we obtain
\[
m_3 = \lim_{\epsilon \to 0}(\mathbb{E}_{\delta n, 3} - m_2 \log \epsilon^{-1}) = \mathcal{O}(1).
\]
(137)

**Remark 16** (Cancellation of the third order divergence). We observe that for some special class of vectors of symmetric matrices \(M = (M^{(1)}, M^{(2)}, M^{(3)})\) the normalization constants \(m_2\) and \(m_3\) are not needed, i.e. \(m_2 = m_3 = 0\) (for example this is the case of the ordinary KPZ equation where \(u \in \mathbb{R}\)).

In fact, if \(M^{(\alpha)}_{\beta\gamma}\) is totally symmetric with respect to three indices, i.e. it is also invariant under the swap \(\alpha \leftrightarrow \beta\), then \(\mathcal{M}_1 = \mathcal{M}_2\) in (106) and (107) and the divergent term is not present.
Appendix A. Proof of Lemma 4

From (37) one has
\[ \mathbb{E}(\vartheta^{(N)}_n(t, x), M^{(\alpha)}_n(t, x)) = \left( \sum_{\beta=1}^{3} M^{(\alpha)}_{\beta \beta} \right) \mathcal{C}^{(N)}(0, 0). \]

Let us split \( \mathcal{C}^{(N)}_n(0, 0) \) by isolating the term corresponding to \( i = 0 \) in (35):
\[ \mathcal{C}^{(N)}_n(0, 0) = \frac{1}{2^{7/2} \sqrt{\pi}} \int_0^\infty \frac{\chi(s)^2 - \chi'(L^{2(N-n)}s)^2}{s^{3/2}} ds + R. \quad (138) \]
where to stress the cutoff dependence we wrote this with the lower cutoff \( \chi' \).

The remainder is easily bounded by
\[ R \leq Ce^{-cL^2n}. \]
and its change with cutoff by
\[ |R - R'| \leq Ce^{-cL^2N} \|\chi - \chi'\|_\infty. \]

For the main term in (138) we define
\[ \rho_\chi = \int_0^\infty \frac{1 - \chi(s)^2}{s^{3/2}} ds. \]
Then
\[ \int_0^\infty \frac{\chi(s)^2 - \chi'(L^{2(N-n)}s)^2}{s^{3/2}} ds = LN^{-n} \rho_\chi - \rho_\chi. \]

Setting \( \delta^{(N)}_n = \sum_{\beta=1}^{3} M^{(\alpha)}_{\beta \beta} (R - \rho_\chi) \) the claim follows. ∎

Appendix B. Proof of Lemma 14

(a) We have:
\[ \mathcal{C}'_n(t, x) = -\Delta \int_0^\infty H_n(t + 2s, x) \chi_{N-n}(t + s) \chi'_{N-n}(s) ds \quad (139) \]
where \( \chi'_{N-n}(t) = \chi(t) - \chi(L^{2(N-n)}t) \). Therefore, since \( \chi_{N-n}(t + s) \chi'_{N-n}(s) \leq \mathbf{1}_{[0,2]}(s) \mathbf{1}_{[0,2]}(t) \),

one has
\[ |\mathcal{C}'_n(t, x)| \leq C \mathbf{1}_{[0,2]}(t) \sum_{j \in \mathbb{Z}} \ell(t, x + jL^n) \quad (140) \]

where
\[ \ell(t, x + jL^n) \leq C \int_0^2 ds(t + 2s)^{-\frac{3}{2}} e^{-\frac{x^2}{4(t + 2s)}} \left[ 1 + x^2(t + 2s)^{-1} \right] \]
\[ \leq Ce^{-c\rho_\chi}(x^2 + t)^{-\frac{3}{2}} \left[ 1 + x^2(x^2 + t)^{-1} \right] \mathbf{1}_{[0,2]}(t) + e^{-c\rho_\chi} t^{-\frac{3}{2}} \left[ 1 + x^2t^{-1} \right] \mathbf{1}_{[0,2]}(t). \quad (141) \]

Combining (140) with (141) one gets
\[ \mathcal{C}_n(t, x) \leq Ce^{-c\rho_\chi}(x^2 + t)^{-\frac{3}{2}} \left[ 1 + x^2(x^2 + t)^{-1} \right] \mathbf{1}_{[0,2]}(t) \in L^p(\mathbb{R} \times \mathbb{T}_n) \quad (142) \]
for \( p < 3 \). To show (114), note that
\[ \chi_{N-n}(t + s) |\chi_s(s) - \chi'_{N-n}(s)| \leq \mathbf{1}_{[e^2,2e^2]}(s) \mathbf{1}_{[0,2]}(t) \|\chi - \chi'\|_\infty \]
where \( \epsilon = L^{-(N-n)} \). Hence
\[
\delta C_n(t, x) \leq C \sum_{j \in \mathbb{Z}} \ell_{N-n}(t, x + jL^n) 1_{[0,2]}(t) \|\chi - \chi'\|_{\infty} \tag{143}
\]
where
\[
\ell_M(t, x) := \int_0^{2L^{-2M}} (t + 2s)^{-\frac{3}{2}} e^{-\frac{4s^2}{4(t + 2s)}} [1 + x^2(t + 2s)^{-1}] ds = L^M \ell_0(L^{2M} t, L^M x). \tag{144}
\]
Hence using (141) we have
\[
\|\ell_M(t, x) 1_{[0,2]}(t)\|_p^p = L^{-(3-p)M} \|\ell_0(t, x) 1_{[0,2L^{2M}]}(t)\|_p^p \leq CL^{-(3-p)M} \left(1 + \int_2^{2L^{2M}} t^{\frac{3}{2}(1-p)} dt\right) \leq CL^{-\lambda M}
\]
with \( \lambda > 0 \) for \( p < 3 \).

(b) The claim follows with the same strategy employed in item (a).

\[\]
with
\[ \hat{a}_\epsilon(t, p) = -p^2 \int_0^1 e^{-(1-\sigma)p^2} \chi_\epsilon((1-\sigma)t) \frac{1}{\sqrt{\sigma}} \hat{W}_\epsilon(\sigma t, \sqrt{\sigma} p) d\sigma. \]
\( \hat{a}_\epsilon \) is entire satisfying (147) and the claim follows.

(c) These claims follow from
\[ |\hat{W}_\epsilon(t, r) - \hat{W}_\epsilon^\prime(t, r)| \leq C e^{-c(\text{Re}r)} 2^{1_{[\frac{1}{2}, 2]}(t)} (148) \]

(d) Let \( B_\epsilon = \partial_x \mathcal{C}^2 = 2 \mathcal{C}_c \partial_x \mathcal{C}_c \). Then
\[ \hat{B}_\epsilon(t, p) = 2i \int_R dq(p + q) e^{-t((p+q)^2 + q^2)} \hat{h}_\epsilon(t, \sqrt{t} p) \hat{h}_\epsilon(t, \sqrt{t} q) \]
Comparing with (146) and noting that
\[ |2h_\epsilon(t, \sqrt{t} p) - \chi_\epsilon(t)| \leq C(1_{[\frac{1}{2}, 2]}(t) + 1_{[\frac{1}{2}, 2]}(t)) \]
we get
\[ |J_\epsilon(z) - \partial_x \mathcal{C}_c(z)| \leq C(e^{-3e^{-c|x|/\epsilon}} 1_{[\frac{1}{2}, 2]}(t) + e^{-c|x|} 1_{[\frac{1}{2}, 2]}(t)). \]
In the same way we get
\[ |Y_\epsilon(z) - 2\partial_x \mathcal{C}_c(z)| \leq C(e^{-2e^{-c|x|/\epsilon}} 1_{[\frac{1}{2}, 2]}(t) + e^{-c|x|} 1_{[\frac{1}{2}, 2]}(t)). \]

\[ \square \]

References