The mathematical theory of wave turbulence

Zaher Hani
joint work with
Yu Deng (USC)

University of Michigan, Ann Arbor

University of Helsinki Minicourse
November 8, 2022

Outline

Lecture 1

1. Hilbert's sixth problem for waves
2. Wave Turbulence Formalism: The physical conjecture
3. Statement of the mathematical result
4. Setup of the diagrammatic expansion: trees and couples

Lecture 2

1. Physics Derivation a.k.a. computation up to second order expansion
2. Leading diagrams: regular couples
3. Outline and strategy of the proof

Lectures 3 & 4 (Yu Deng) Deeper details of the proof

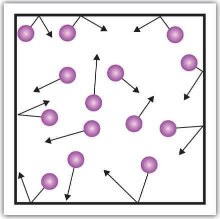
Reference

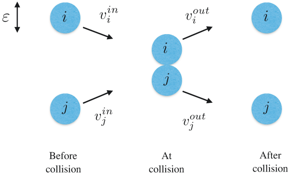
1. (Deng-H. 1), *Full derivation of the wave kinetic equation.* Preprint available at arxiv.org/abs/2104.11204.
2. (Deng-H. 2), *Propagation of chaos and higher order statistics in wave kinetic theory.* Preprint available at https://arxiv.org/abs/2110.04565
3. (Deng-H. 3), *Full Derivation of the wave kinetic equation: Full range of scaling laws.* To appear soon.
4. (Deng-H. 4), (Expository*) *Rigorous justification of the wave kinetic theory.* Preprint available at https://arxiv.org/abs/2207.08358.

Hilbert's sixth problem

- Hilbert's sixth problem asks for the axiomatic derivation of the laws of physics from first principles. A main example is justifying the laws of statistical physics starting from the laws of dynamics.
- Statistical mechanics governs *macroscopic* quantities (temperature, pressure,...), rather than *microscopic quantities* (like particle trajectories) which are described by the laws of dynamics (classical or quantum).
- The dynamical laws are *reversible* in time (Newton's laws, Hamilton's equations), whereas the statistical ones are typically *time irreversible* (e.g. second law of thermodynamics).
- The passage from the reversible to the irreversible comes as a result of *averaging* and *limiting* operations, which creates the <u>*arrow of time*</u>. This apparent paradox is adequately explained by providing a *rigorous derivation* of the laws of statistics from those of dynamics. This is the essence of Hilbert's sixth problem.
- We shall be interested today in this problem in the context of nonlinear waves.

Classical example: Boltzmann's kinetic theory





- *Boltzmann's kinetic theory* (1872): Suppose we start with a system with N particles, each of radius r undergoing elastic collisions. Assume that the initial states of the particles are random and independent, so that each particle has an initial density $f_0(x, v)$. The effective dynamics of the one-particle density function $F_N^{(1)}(t, x, v)$ is given by the **Boltzmann equation**

$$(\partial_t + v \cdot \nabla_x) f(t, x, v) = \mathcal{Q}(f, f).$$

$$\mathcal{Q}(f, f)(t, x, v) = \int_{v' + v'_* = v + v_*} dv' dv'_* \delta(|v|^2 + |v_*|^2 - |v'|^2 - |v'_*|^2) \\ (f' f'_* - f f') \tag{0.1}$$

Boltzmann Theory continued

- The justification of this approximation was a big challenge (Lanford, Cercignani-Illner-Pulvirenti, **Gallagher-Saint-Raymond-Texier**, Pulvirenti-Saffirio-Simonella).
- *Lanford's Theorem (1975):* In the limit $N \to \infty$ and $r \to 0$ under the **Boltzmann-Grad scaling law**

$$\boxed{\gamma := Nr^{d-1} \sim 1},$$

  1. *Propagation of Chaos*: The states of the particles retain their initial independence.
  2. Boltzmann's equation appears as an effective equation for $F_N^{(1)}(t, x, v)$.

- The approximation $F_N^{(1)}(t, x, v) \approx f(t, x, v)$ holds in the limit $N \to \infty$ and $r \to 0$ for times $O(\gamma^{-1})$.
- This kinetic framework has been highly informative and was extended to many other particle systems (Vlasov, Landau, etc.).

Hilbert's sixth problem for wave systems

- Naturally, a parallel theory of statistical mechanics for nonlinear wave systems followed soon after (Peierls 1929, Hasselman 1962, Zakharov, etc.)
- The microscopic dynamics is given by a nonlinear dispersive/wave PDE. Instead of a large number of particles, we have a large number of waves, represented by Fourier modes, in a box of size L with $L \to \infty$. These waves undergo "collisions" through the nonlinear interactions given by the dispersive system.
- We consider the Nonlinear Schrödinger (NLS) equation

$$\begin{cases} (i\partial_t + \Delta)v(t,x) = \alpha|v|^2 v, & x \in \mathbb{T}_L^d \\ v(t=0) = v_0(x) \end{cases} \tag{NLS}$$

where $\mathbb{T}_L^d$ is a periodic box of size L with $d \geqslant 3$ and α stands for the size of the nonlinearity.

- The NLS carries particular importance as a system for nonlinear waves due to its universality property: virtually any Hamiltonian dispersive/wave system gives NLS in an appropriate scaling limit.

Hilbert's sixth problem for wave systems

- Naturally, a parallel theory of statistical mechanics for nonlinear wave systems followed soon after (Peierls 1929, Hasselman 1962, Zakharov, etc.)
- The microscopic dynamics is given by a nonlinear dispersive/wave PDE. Instead of a large number of particles, we have a large number of waves, represented by Fourier modes, in a box of size L with $L \to \infty$. These waves undergo "collisions" through the nonlinear interactions given by the dispersive system.
- We consider the Nonlinear Schrödinger (NLS) equation

$$\begin{cases} (i\partial_t + \Delta)v(t,x) = \alpha|v|^2 v, & x \in \mathbb{T}_L^d \\ v(t=0) = v_0(x) \end{cases} \tag{NLS}$$

  where $\mathbb{T}_L^d$ is a periodic box of size L with $d \geqslant 3$ and α stands for the size of the nonlinearity.
- The NLS carries particular importance as a system for nonlinear waves due to its universality property: virtually any Hamiltonian dispersive/wave system gives NLS in an appropriate scaling limit.

Random Initial Data

- Like in the many-particle case, we start with a random distribution of initial data. This corresponds to taking random and independent Fourier modes for u_{in} as follows:

$$u_{\text{in}}(x) = L^{-d/2} \sum_k \widehat{u_{\text{in}}}(k) e^{2\pi i k \cdot x}, \qquad \widehat{u_{\text{in}}}(k) = \sqrt{n_{\text{in}}(k)} \eta_k(\omega).$$

  Here, k ranges over $\mathbb{Z}_L^d = L^{-1}\mathbb{Z}^d$, a lattice with mesh L^{-1} (which tends to continuum as $L \to \infty$), n_{in} is a non-negative Schwartz function on $\mathbb{R}^d$, and $\eta_k(\omega)$ are i.i.d. normalized random variables (Gaussian or not). Such data is called *well-prepared* since different **Fourier modes are independent** and

$$\mathbb{E}|\widehat{u_{\text{in}}}(k)|^2 = n_{\text{in}}(k).$$

- The goal of the wave kinetic theory is to understand the distribution of the Fourier modes at later times in the limit of large L and small α. This replaces the limits $N \to \infty$ and $r \to 0$ in Boltzmann's particle theory. A particularly central quantity is the variance $\mathbb{E}|\widehat{u}(t,k)|^2$.

Wave Kinetic Theory, a.k.a. Wave Turbulence Theory

- *Kinetic Conjecture:* In the limit as $L \to \infty$ and $\alpha \to 0$ (weak nonlinearity), there holds
  1. *Propagation of Chaos*: Different Fourier modes retain their independence in the limit.
  2. *Kinetic equation:* The effective dynamics of $\mathbb{E}|\widehat{u}(t,k)|^2$ is given by the **Wave Kinetic Equation** (next slide), rescaled to the "kinetic time" $T_{\text{kin}} := \alpha^{-2}$. More precisely, one expects that

$$\boxed{\mathbb{E}|\widehat{u}(t,k)|^2 = n(\frac{t}{T_{\text{kin}}}, k) + o(1), \text{ as } L \to \infty \text{ and } \alpha \to 0,} \tag{APPROX}$$

     where $n(t,\xi)$ solves the (WKE) with initial data n_{in}, and $u(t)$ solves (NLS) with the well-prepared initial data u_{in}.
  3. *Distribution of Fourier modes:* The limiting behavior of the density (law) of $\widehat{u}(t,k)$ can be described in terms of the above limiting behavior of the variance $\mathbb{E}|\widehat{u}(t,k)|^2$.

The wave kinetic equation

- The wave kinetic equation was introduced in the physics literature, and has the form

$$\begin{cases} \partial_t n &= \mathcal{K}(n,n,n) \\ n(0) &= n_{\mathrm{in}}, \end{cases} \tag{WKE}$$

where

$$\mathcal{K}(n,n,n)(\xi) = 2\int_{\substack{\xi_1,\xi_2,\xi_3\in\mathbb{R}^d \\ \xi_1-\xi_2+\xi_3=\xi}} \delta_{\mathbb{R}}(|\xi_1|^2 - |\xi_2|^2 + |\xi_3|^2 - |\xi|^2)$$
$$n(\xi_1)n(\xi_2)n(\xi_3)n(\xi)\left(\frac{1}{n(\xi_1)} - \frac{1}{n(\xi_2)} + \frac{1}{n(\xi_3)} - \frac{1}{n(\xi)}\right)$$

- WKE is a wave analog of Boltzmann's equation. Note that this kinetic approximation features a passage from the time reversible NLS equation into the time-irreversible wave kinetic equation. There is also an inhomogeneous version thereof in which n is also space-dependent and the LHS of the (WKE) has a transport term. We restrict ourselves today to the homogeneous setting.

Some History in the physics literature

- As mentioned, the emergence of wave turbulence theory was roughly in the 1920s. Since then, this kinetic formalism was generalized systematically to many nonlinear wave systems: cf. Hasselman's eq'n in the context of water waves, the Phonon Boltzmann equations for anharmonic crystals, and also WKE in plasma theory and nonlinear optics (Zakharov, Newell, Davidson, etc.). See Nazarenko's monograph for a textbook treatment
- It is highly informative in applications (oceanography, optics, crystal thermodynamics, plasma theory, etc.). Most notably, it leads to a formal framework of turbulence for nonlinear waves, hence the name *wave turbulence theory*.
- Formally, (WKE) gives conclusions similar to those made in hydrodynamic turbulence, namely *power-law cascade spectra*. These appear as special stationary solutions of the (WKE), called the *Kolmogorov-Zakharov cascade spectra*, that were discovered by Zakharov in the 1960's (see Nazarenko). Mathematically, we cite the work of [Escobedo-Velázquez, Soffer-Tran, ...] for some rigorous analysis on these equations, but our understanding of their long-time behavior is still in its infancy.

History in the mathematical literature

- Wave turbulence theory (WTT) yields fundamental mathematical implications on the generic long-time behavior of solutions of dispersive equations.
- In fact, the initial focus of the mathematical research related to (WTT) was on constructing solutions that exhibit *energy cascade* from low to high frequencies. This phenomenon is implied by the KZ-spectra mentioned above. The idea was to capture this cascade through the growth of Sobolev norms (Bourgain):

$$\|u(t)\|_{H^s}^2 := \sum_{k\in\mathbb{Z}^d} (1+|k|)^{2s} |\widehat{u}(t,k)|^2.$$

- The most influential work here is that of [CKSTT] who proved the existence of solutions on the unit torus $\mathbb{T}^d$ that exhibit *arbitrarily large but finite growth of Sobolev norms*. Such solutions are very special and far from generic. Whether this behavior is generic or whether there exists solutions that exhibit infinite growth of Sobolev norms are outstanding open problems on $\mathbb{T}^d$, despite partial results in this direction in [H.], [Guadia-Kaloshin], [H., Pausader, Tzvetkov, Visciglia], [H., Guardia, H., Haus, Maspero, Procesi].
- Of course, a rigorous understanding of the kinetic approximation (APPROX) over sufficiently long times would give deep insights into such questions.

Rigorous Derivation of the (WKE): scaling laws

- **Scaling Laws:** dictate the relative rates of the limits $L \to \infty$ and $\alpha \to 0$. We saw an example of it in Lanford's theorem, which imposes the Boltzmann-Grad scaling law $Nr^{d-1} \sim 1$.
- A priori, we may start by considering all possible scaling laws $\boxed{\alpha = L^{-\gamma}}$ where $0 \leqslant \gamma \leqslant \infty$. The case $\gamma = 0$ would correspond to taking the $L \to \infty$ limit followed by the $\alpha \to 0$ limit, and vice versa for the case $\gamma = \infty$.
- Not all scaling laws are admissible by the kinetic theory, and the admissible scaling laws can depend on the "shape" of the domain (i.e. aspect ratios of the box). We shall see that for:

  1. Rational Tori: The admissible range is $0 \leqslant \gamma < 1$ is admissible (e.g. square box).
  2. Generic Tori: i.e. with genericity (diophantine) conditions on the aspect ratios, the range widens to $0 \leqslant \gamma \leqslant d/2$. The interest in one scaling law over another seems to depend on the physical context. This is determined by the strength of the exact resonances in the nonlinearity.

The log-log plot

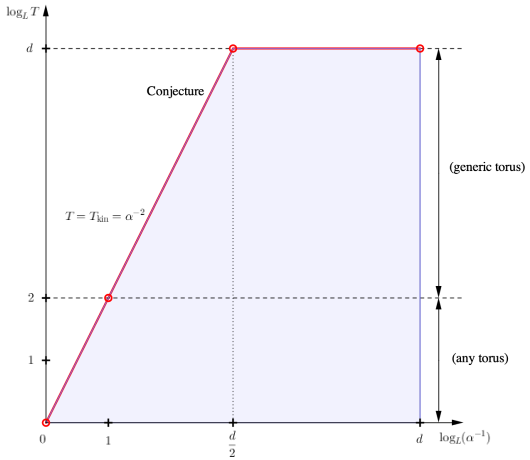


Figure: The timescales of the kinetic conjecture (the red line). The x-coordinate specifies the scaling law and the y-coordinate specifies the time T. Note the segment of the red line that is supposed to hold without any diophantine conditions on the aspect ratios of the torus.

Scaling laws (continued)

- The choice of scaling law depends on the model and the physical setting. There are a couple of particularly interesting choices for NLS:
- $T_{\text{kin}} = L^2$ **scaling law.** $(\gamma = 1)$. We can call this the "Schrodinger scaling" (space scale is L and the timescale is L^2). The importance of this scale is that the dynamics can be rescaled back to the unit torus $\mathbb{T}^d$ to give results at $O(1)$ timescales.

$$v(t,x) = L^{\frac{1}{2}} u(L^2 t, Lx), \qquad x \in \mathbb{T}^d, \qquad\qquad (i\partial_t v + \Delta)v = |v|^2 v.$$

  This ties the kinetic theory in this setting to related problems on unit torus (e.g. growth of Sobolev norms, etc.) In three dimension, it also ties the problem to questions in constructive quantum field theory, namely the invariance of the Φ^4_3-Gibbs measure (still open). Such invariance is known for the Langevin dynamics (Hairer, Gubinelli-Imkeller-Perkowski, Kupiainen) and for NLW dynamics (Bringmann-Deng-Nahmod-Yue).

Scaling laws (continued)

- **Ballistic Scaling Law.** (Spohn) $\gamma = \frac{1}{2}$, for which $T_{\text{kin}} = L$. Here the space scale and the timescale are both L. This is important in certain settings (particularly when considering inhomogeneous problems) when there is a need to match the kinetic timescale with the transport timescale of wave packets (time for a scale-1 wave packet to move across the domain).
- **To summarize:** WK theory describes a range of different phenomena at different scaling laws γ. Thus it is important to justify the theory for all such γ.
- Without imposing genericity assumption on the domain, the maximal range to expect is $\gamma \in [0, 1)$. Can we justify the theory all all such γ (at least up to endpoint)?

Previous mathematical results

- There are a several works that attempted similar or close-by questions. First is the work of Erdös, Salmhofer, Yau 2008 for the linear Schrödinger equations with random potential.
- More pioneering is the work of Lukkarinen-Spohn [Inventiones 2011] that studied the problem of time correlations of an equilibrium Gibbs measure. Analog of our problem but with stationary initial data.
- *Stochastic setting* (time-dependent random terms in the equation): Partial results by [Faou], [Dymov-Kuksin], [D-K-Maiocchi, Vladuts].
- *Deterministic setting*: The first attempt at the full nonlinear problem was in the work of Buckmaster-Germain- H.-Shatah [BGHS'19]. There, in the setting of generic irrational tori and for $d \geqslant 3$, the approximation (APPROX) was shown to hold for times $t \leqslant T_*(\alpha, L)$ where $1 \ll T_*(\alpha, L) \ll T_{\mathrm{kin}}$.
- Later works by [Deng, H.'19] and [Collot-Germain'19] improved this approximation interval considerably, all the way up to times $t \leqslant L^{-\varepsilon} T_{\mathrm{kin}}$ for arbitrarily small $\epsilon > 0$, and for some particular *scaling laws*. In [DH.'19], we did this for scaling laws $\gamma = 1$ and $\gamma = \varepsilon$, and had partial results for other scaling laws. The work of [CG'19] also gave the same result for $\gamma = 1$. See also later works [CG'20].

The log-log plot

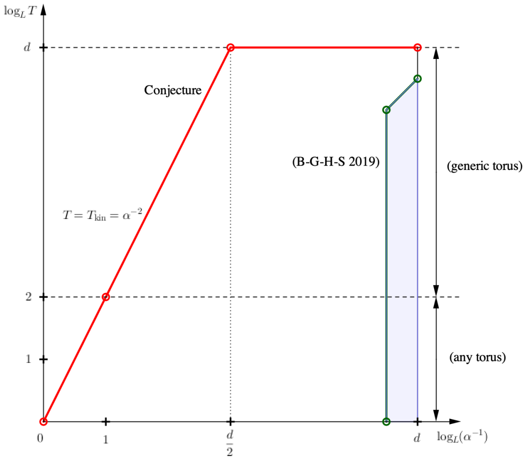


Figure: The result of Buckmaster-Germain-H.-Shatah [Inventiones '2020].

The log-log plot

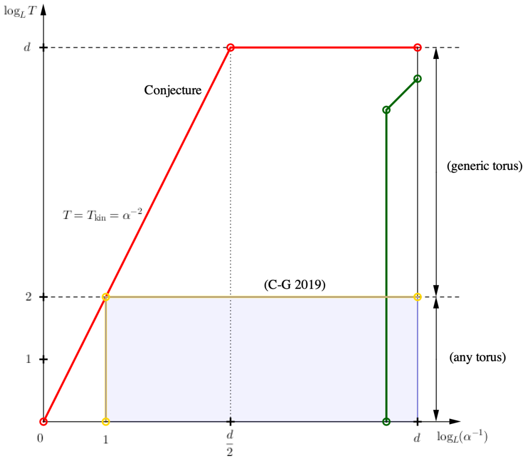


Figure: The result of Collot-Germain [CPAM 2021]

The log-log plot

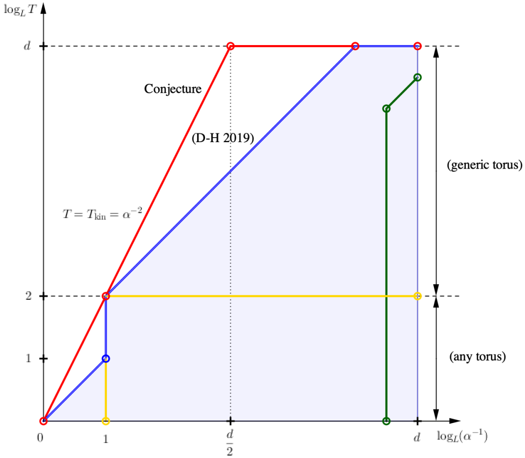


Figure: The result of Deng-Hani 2019 [Forum of Math Pi, 2021].

All these works fail to reach the kinetic timescale T_{kin}. A full justification of the equation should reach times scales $t \sim \delta \cdot T_{\text{kin}}$ where $\delta > 0$ is independent of L and α.

The Main theorem

In recent joint works with Yu Deng, we are able to give a full mathematical justification of wave turbulence theory, including a rigorous derivation of the (WKE) at the kinetic timescale T_{kin}, for some particular range of scaling laws.

Theorem (Deng-H., 2021-2022)

- *Consider (NLS) on the periodic box $\mathbb{T}_L^d$ with $d \geqslant 3$.*
- *Take $n_{\text{in}} \geqslant 0$ in $\mathcal{S}(\mathbb{R}^d)$ and u_{in} to be well-prepared, i.e. $\widehat{u}_{\text{in}}(k) = \sqrt{n_{\text{in}}(k)}\eta_k(\omega)$, and suppose that the law of $\eta_k(\omega)$ is rotationally symmetric and has exponential tails (e.g. Gaussian).*
- *Scaling laws: Let $\alpha \sim L^{-\gamma}$ for $\gamma \in (0,1]$, and recall that $T_{\text{kin}} = \alpha^{-2}$. For $\gamma = 1$, we assume suitable genericity conditions on the aspect ratios of the box.*

THEN, there exists $\delta < 1$ fixed, and an absolute constant $\nu > 0$ such that for L large enough it holds that

$$\mathbb{E}|\widehat{u}(t,k)|^2 = n(\frac{t}{T_{\text{kin}}}, k) + O(L^{-\nu})$$

uniformly in (t,k) for $t \in [0, \delta \cdot T_{\text{kin}}]$. Here $n(t,k)$ solves the wave kinetic equation with data n_{in}.

The main theorem (continued)

Theorem (Deng-H., 2021-2022(cont'd))

Moreover, suppose that $k_1, \ldots, k_r$ *are distinct, then*

1. *Propagation of Chaos: The random variables* $\widehat{u}(t, k_j)$ $(1 \leqslant j \leqslant r)$ *retain their independence in the kinetic limit* $L \to \infty$.
2. *Limiting law: The law of* $\widehat{u}(t, k)$ *converges to the density function* $\rho_k(t, v)$ *(with* $v \in \mathbb{R}^2$*) which evolves according to the linear PDE*

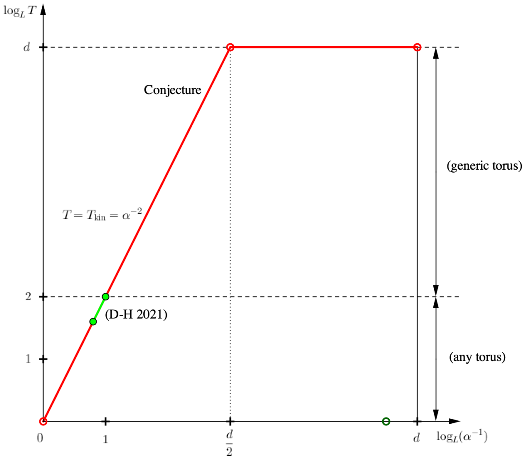
$$\partial_t \rho_k = \frac{\sigma_k(t)}{4} \Delta \rho_k - \frac{\gamma_k(t)}{2} \nabla \cdot (v \rho_k),$$

   where $\sigma_k(t) > 0$ *and* $\gamma_k(t)$ *are functions constructed from the solution* $n(t, k)$ *to the wave kinetic equation.*
3. *Propagation of Gaussianity: In particular, if* $\eta_k(\omega)$ *are Gaussian, then* $\rho_k(t, v)$ *is Gaussian with variance* $n(t, k)$ *for any* $t > 0$.

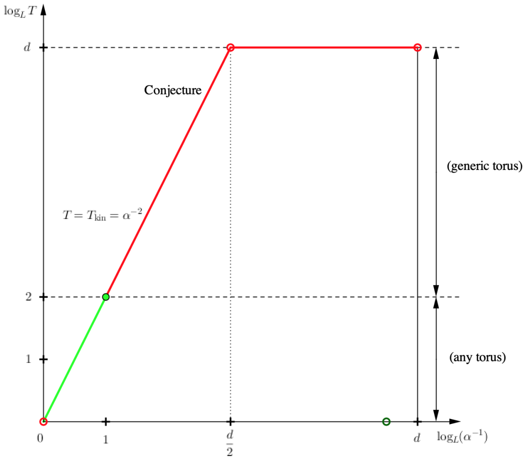
Remarks on the result

- These results provide a full mathematical foundation of wave turbulence theory, by answering all three fundamental questions: propagation of chaos, derivation of the wave kinetic equation, and the evolution of the limiting density.
- The genericity condition at $\gamma = 1$ ($T_{\text{kin}} = L^2$) is merely technical but needed (number theoretic reason). Such conditions are not needed for $\gamma < 1$.
- The case $\gamma \in [1 - 1/(20d), 1]$ was proved in Apr. 2021 (arXiv:2104.11204) for the derivation of WKE, and Oct. 2021 (arXiv:2110.04565) for the propagation of chaos and higher order statistics.
- The case $\gamma \in (0, 1 - 1/(20d))$ is proved in a third paper, coming very soon. While the main strategy (as well as a good part of the argument is the same for all γ), some very interesting new characters only show up when γ becomes $< 2/3$. Yu Deng will tell us about those on Friday.
- The equation for the density evolution equation was implicitly contained in the original work of Peierls (1929), and later rediscovered in Nazarenko (2011). Now it also has a rigorous proof!

The log-log plot (Deng-H. 2021)



Deng-Hani 2022



More recent works

- Work of Staffilani-Tran: Here, the authors consider discrete KdV-type equation, and in addition to randomizing the data, a time-dependent Stratonovich stochastic force is added to the equation (Faou). The effect of the force is to keep randomizing the phases of the Fourier modes at all times. In this setting, they give a derivation of the (WKE) at the kinetic timescale. Recent work of Hannani-Rosenzweig-Staffilani-Tran deals with the inhomogeneous case with a different well-chosen stochastic force.
- X. Ma ('22) considered the same equation without the stochastic forcing but with dissipation $\nu\Delta u$ and proved the approximation for subcritical times $T \ll L^{-\varepsilon} T_{\mathrm{kin}}$.
- Also [Dubach-Germain-Harrop-Griffiths] for a model with a random dispersion relation, [Ampatzoughlu-Collot-Germain] for a quadratic Schrodinger equation in the inhomogeneous setting (also subcritical times).

The Wave Kinetic Heirarchy

- The wave kinetic heirarchy (WKH) is a system of equations that describes the asymptotic evolution of the special moments

$$\mathbb{E}\left(|\widehat{u}_{k_1}(t)|^2|\widehat{u}_{k_2}(t)|^2 \dots |\widehat{u}_{k_r}(t)|^2\right) \to n^r(t, k_1, \dots k_r) \rightsquigarrow \text{ sol'n to the (WKH).}$$

- It can be written schematically for the infinite vector (n^r)

$$\partial_t n^r(t; k_1, \dots, k_r) = \mathcal{C}(n^{r+2}, n^{r+2}, n^{r+2}) \qquad \text{(WKH)}$$

  for some collision kernel $\mathcal{C}$ similar to the one appearing in (WKE).

- Factorization property of the (WKH)

$$\text{If} \quad n_{\text{in}}^r(k_1, \dots, k_r) = \prod_{j=1}^{r} n_{\text{in}}(k_j) \quad \text{THEN} \quad n^r(t; k_1, \dots, k_r) = \prod_{j=1}^{r} n(t, k_j)$$

- This implies: If one assumes that the initial data are independent (as we have been), then the kinetic limit of such moments is given by $\prod_{j=1}^{r} n(\frac{t}{T_{\text{kin}}}, k)$ where $n(t, \cdot)$ solves the (WKE).
- So our theorem above gives a justification of the (WKH) *for factorized solutions.* What about other solutions?

The Wave Kinetic Heirarchy (Continued)

- To access other solutions, we need to relax the independence assumption on the initial data, and allow for some correlations between the initial Fourier modes, so that

$$\mathbb{E}\left(|\widehat{u}_{k_1}^{(\mathrm{in})}|^2|\widehat{u}_{k_2}^{(\mathrm{in})}|^2\ldots|\widehat{u}_{k_r}^{(\mathrm{in})}|^2\right) = n_{\mathrm{in}}^r(k_1,\ldots k_r) \rightsquigarrow \text{ not factorized.}$$

  This is reminiscent to physics treatments describing random phases but correlated amplitudes (cf [Nazarenko]).
- Key: Such initial distributions can be written as an appropriate average of factorized distributions (a variant of the Hewitt-Savage Theorem).
- Using the linearity of the (WKH) and the uniqueness of its solutions (Rosenzweig-Staffilani), we can transfer our result for factorized data to that of unfactorized ones.
- Interestingly, this goes in the opposite direction of the paradigm followed for Boltzmann or Gross-Pitaevskii for which the heirarchy is used to justify the one-mode distribution.

Setup of the proof: the diagrammatic expansion

The equation in Fourier Space

Expanding the solution $u = L^{-d/2}\sum_{k\in\mathbb{Z}_L^d} \widehat{u}_k(t)e^{2\pi i K\cdot x}$, we obtain that

$$
\begin{aligned}
-i\partial_t \widehat{u}_k =& -2\pi|k|^2\widehat{u}_k + \frac{\alpha}{L^d}\sum_{k_1-k_2+k_3=k} \widehat{u}_{k_1}\overline{\widehat{u}_{k_2}}\widehat{u}_{k_3} \\
=& -2\pi|k|^2\widehat{u}_k + 2\frac{\alpha}{L^d}\underbrace{\left(\sum_{k_1}|\widehat{u}_{k_1}|^2\right)}_{M[u]=cst}\widehat{u}_k - \frac{\alpha}{L^d}|\widehat{u}_k|^2\widehat{u}_k + \frac{\alpha}{L^d}\sum_{\substack{k_1-k_2+k_3=k\\ k_1,k_3\neq k}} \dots \\
=& (-2\pi|k|^2 + 2\alpha L^{-d}M)\widehat{u}_k - \frac{\alpha}{L^d}|\widehat{u}_k|^2\widehat{u}_k + \frac{\alpha}{L^d}\sum_{\substack{k_1-k_2+k_3=k\\ k_1,k_3\neq k}} \widehat{u}_{k_1}\overline{\widehat{u}_{k_2}}\widehat{u}_{k_3}
\end{aligned}
$$

Let $c_k(t) = \widehat{u}_k(t)e^{i(2\pi|k|^2-2\alpha L^{-d}M)t}$, then

$$
\boxed{-i\partial_t c_k = \frac{\alpha}{L^d}\sum_{\substack{k_1-k_2+k_3=k\\ k_1,k_3\neq k}} c_{k_1}\overline{c_{k_2}}c_{k_3}e^{-2\pi i\Omega(k_1,k_2,k_3,k)t} - \frac{\alpha}{L^d}|c_k|^2c_k.}
$$

$$
\Omega(k_1,k_2,k_3,k) := |k_1|^2 - |k_2|^2 + |k_3|^2 - |k|^2 = 2\langle k_1-k, k-k_3\rangle.
$$

The equation (continued)

- It is enough to study the solution $c_k(t)$ on the interval $[0, \delta T_{\text{kin}}]$. Recall that $\alpha = L^{-\gamma}$ and $T_{\text{kin}} = \frac{1}{2\alpha^2} = \frac{L^{2\gamma}}{2}$. Setting

$$a_k(t) = c_k(\delta T_{\text{kin}} \cdot t), \qquad k \in \mathbb{Z}_L^d,$$

It is enough to study $(a_k(t))_k$ for $0 \leqslant t \leqslant 1$, and $a_k(t)$ satisfies

$$\begin{cases} \partial_t a_k = \dfrac{i\delta}{2L^{d-\gamma}} \displaystyle\sum_{\substack{k_1 - k_2 + k_3 = k \\ k_1, k_3 \neq k}} e^{\delta \pi i L^{2\gamma} \Omega(k_1, k_2, k_3, k) t} a_{k_1}(t) \overline{a_{k_2}(t)} a_{k_3}(t) \\ \qquad\qquad\qquad\qquad - \dfrac{i\delta}{2L^{d-1}} |a_k(t)|^2 a_k(t), \\ a_k(0) = (a_k)_{\text{in}} = \sqrt{n_{\text{in}}(k)} \eta_k(\omega), \end{cases}$$

- We shall expand the solution $a_k(t)$ in Duhamel iterates up to order N plus a remainder term.

The tree expansion

Now, we expand the solution into Picard iterates [Luk-Spohn, BGHS, DH'19, CG'19]

$$a_k(t) = a_k^{(0)}(t) + a_k^{(1)}(t) + a_k^{(2)}(t) + \ldots + a_k^{(N)}(t) + R_k^{(N)}(t).$$

- $a_k^{(0)}(t) = \widehat{u}_{\text{in}}(k) = \sqrt{n_{\text{in}}(k)}\eta_{k,\omega}$, and

$$a_k^{(1)} = \tfrac{i\delta}{2L^{d-\gamma}} \textstyle\sum_{k_1-k_2+k_3=k} \left(\underbrace{\int_0^t e^{\delta\pi i L^{2\gamma}\Omega(k_1,k_2,k_3,k)s} ds}_{A(t,\Omega)} \right) \textstyle\prod_{j=1}^3 \sqrt{n_{\text{in}}(k_j)}\eta_{k_j}^{\pm}$$

- $a_k^{(2)}(t)$ is a sum over (three) ternary trees with order 2 (order=number of branching nodes). $a_k^{(2)}(t) = a_k^{(T_1)}(t) + a_k^{(T_2)}(t) + a_k^{(T_3)}(t)$

$$a_k^{(T_2)}(t) = \zeta_{T_2} \left(\frac{\delta}{2L^{d-\gamma}}\right)^2 \sum_{\substack{k_1-k_2+k_3=k \\ n_1-n_2+n_3=k_2}} \underbrace{\left(\int_0^t \int_0^s e^{\delta\pi i L^{2\gamma}\Omega_1 s} e^{-\delta\pi i L^{2\gamma}\Omega_2 \tau}\, d\tau\, ds\right)}_{A(t,\Omega_1,\Omega_2)}$$

$$\times \prod_{\mathfrak{l}} \sqrt{n_{\text{in}}(k_{\mathfrak{l}})}\, \eta_{k_{\mathfrak{l}}}^{\iota_{\mathfrak{l}}}, \qquad \iota_{\mathfrak{l}} \in \{+,-\}.$$

- Here $\Omega_1 = \Omega(k_1, k_2, k_3, k)$ and $\Omega_2 = \Omega(n_1, n_2, n_3, k_2)$

The tree expansion

Now, we expand the solution into Picard iterates [Luk-Spohn, BGHS, DH'19, CG'19]

$$a_k(t) = a_k^{(0)}(t) + a_k^{(1)}(t) + a_k^{(2)}(t) + \ldots + a_k^{(N)}(t) + R_k^{(N)}(t).$$

- $a_k^{(0)}(t) = \widehat{u}_{\mathrm{in}}(k) = \sqrt{n_{\mathrm{in}}(k)}\eta_{k,\omega}$, and

$$a_k^{(1)} = \tfrac{i\delta}{2L^{d-\gamma}} \textstyle\sum_{k_1-k_2+k_3=k} \left(\underbrace{\int_0^t e^{\delta\pi i L^{2\gamma}\Omega(k_1,k_2,k_3,k)s} ds}_{A(t,\Omega)} \right) \textstyle\prod_{j=1}^3 \sqrt{n_{\mathrm{in}}(k_j)}\eta_{k_j}^{\pm}$$

- $a_k^{(2)}(t)$ is a sum over (three) ternary trees with order 2 (order=number of branching nodes). $a_k^{(2)}(t) = a_k^{(T_1)}(t) + a_k^{(T_2)}(t) + a_k^{(T_3)}(t)$

$$a_k^{(T_2)}(t) = \zeta_{T_2} \left(\frac{\delta}{2L^{d-\gamma}}\right)^2 \sum_{\substack{k_1-k_2+k_3=k \\ n_1-n_2+n_3=k_2}} \underbrace{\left(\int_0^t \int_0^s e^{\delta\pi i L^{2\gamma}\Omega_1 s} e^{-\delta\pi i L^{2\gamma}\Omega_2 \tau} \, d\tau \, ds\right)}_{A(t,\Omega_1,\Omega_2)}$$

$$\times \prod_{\mathfrak{l}} \sqrt{n_{\mathrm{in}}(k_{\mathfrak{l}})}\, \eta_{k_{\mathfrak{l}}}^{\iota_{\mathfrak{l}}}, \qquad \iota_{\mathfrak{l}} \in \{+,-\}.$$

- Here $\Omega_1 = \Omega(k_1, k_2, k_3, k)$ and $\Omega_2 = \Omega(n_1, n_2, n_3, k_2)$

Tree expansion formula

- In general, one can obtain easily by induction that $a_k^{(n)}(t) = \sum_{|\mathcal{T}|=n} a_k^{\mathcal{T}}(t)$ and

$$a_k^{\mathcal{T}}(t) = \zeta_{\mathcal{T}} \left(\frac{\delta}{2L^{d-\gamma}}\right)^n \sum_{(k_\mathfrak{n})\in\mathcal{D}} A\left(t, (\Omega_\mathfrak{n})_{\mathfrak{n}\in\mathcal{N}}\right) \prod_{\mathfrak{l}\in\mathcal{L}} \sqrt{n_{\mathrm{in}}(k_\mathfrak{l})}\eta_{k_\mathfrak{l}}^{\pm}$$

  where
  - $|\mathcal{T}|$ =order of $\mathcal{T}$, $\zeta_{\mathcal{T}}$ is the product of n factors of $\pm i$.
  - The sum over $k_\mathfrak{n} \in \mathbb{Z}_L^d$ is over decorations $\mathcal{D}$ of the tree: these are assignments of $k_\mathfrak{n} \in \mathbb{Z}_L^d$ for each $\mathfrak{n} \in \mathcal{T}$ such that $k_\mathfrak{r} = k$ and $k_\mathfrak{n} = k_{\mathfrak{n}_1} - k_{\mathfrak{n}_2} + k_{\mathfrak{n}_3}$ whenever $\mathfrak{n}$ is a branching node with children $\mathfrak{n}_1, \mathfrak{n}_2, \mathfrak{n}_3$.
  - $\Omega_\mathfrak{n} = |k_{\mathfrak{n}_1}|_\beta^2 - |k_{\mathfrak{n}_2}|_\beta^2 + |k_{\mathfrak{n}_3}|_\beta^2 - |k_\mathfrak{n}|_\beta^2$ for every $\mathfrak{n} \in \mathcal{N}$, the set of branching nodes.
  - $\mathcal{L}$ is the set of leaves, and $\eta_{k_\mathfrak{l}}^{\pm} = \eta_{k_\mathfrak{l}}$ if $\mathfrak{l}$ has $+$ sign and $\overline{\eta_{k_\mathfrak{l}}}$ if $\mathfrak{l}$ has $-$ sign.
  - $$A\left(t, (\Omega_\mathfrak{n})_{\mathfrak{n}\in\mathcal{N}}\right) = \int_{\mathscr{E}} \prod_{\mathfrak{n}\in\mathcal{N}} e^{\delta\pi i L^{2\gamma}\Omega_\mathfrak{n} t_\mathfrak{n}}\, dt_\mathfrak{n}, \quad \textit{where}$$
    $$\mathscr{E} = \{t_\mathfrak{n} \in [0,t] : \mathfrak{n} \in \mathcal{N}, \quad t_{n_f} < t_{\mathfrak{n}'} \text{ if } \mathfrak{n} \text{ is a child of } \mathfrak{n}'\}.$$

- To summarize,

$$a_k(t) = a_k^{(0)}(t) + a_k^{(1)}(t) + a_k^{(2)}(t) + \ldots + a_k^{(N)}(t) + R_k^{(N)}(t).$$

$$a_k^{(n)}(t) = \sum_{|\mathcal{T}|=n} a_k^{\mathcal{T}}(t), \qquad a_k^{\mathcal{T}}(t) \quad \text{as given above.}$$

- In computing, $\mathbb{E}|a_k(t)|^2$ we are thus led to consider

$$\mathbb{E}\, a_k^{\mathcal{T}_1}(t)\overline{a_k^{\mathcal{T}_2}(t)} = \zeta_{\mathcal{T}_1}\overline{\zeta_{\mathcal{T}_2}}\left(\frac{\delta}{2L^{d-\gamma}}\right)^{n_1+n_2} \sum_{(k_\mathfrak{n})_{\mathfrak{n}\in\mathcal{T}_1}} \sum_{(k_\mathfrak{n})_{\mathfrak{n}\in\mathcal{T}_2}} A\big(t,(\Omega_\mathfrak{n})_{\mathfrak{n}\in\mathcal{N}_1}\big)\overline{A\big(t,(\Omega_\mathfrak{n})_{\mathfrak{n}\in\mathcal{N}_2}\big)}$$

$$\times \prod_{\mathfrak{l}\in\mathcal{L}_1\cup\mathcal{L}_2} \sqrt{n_{\mathrm{in}}(k_\mathfrak{l})}\, \mathbb{E}\left(\prod_{\mathfrak{l}\in\mathcal{L}_1\cup\mathcal{L}_2} \eta_{k_\mathfrak{l}}^{\iota_\mathfrak{l}}\right).$$

- **Isserles' Theorem:** If η_k are i.i.d. complex Gaussians, then

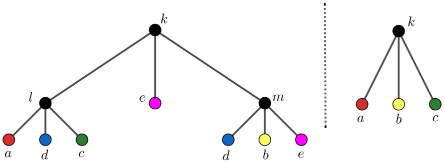
$$\mathbb{E}\left(\prod_{\mathfrak{l}\in\mathcal{L}_1\cup\mathcal{L}_2} \eta_{k_\mathfrak{l}}^{\iota_\mathfrak{l}}\right) = \sum_{\mathcal{P}} \prod_{(\mathfrak{l},\mathfrak{l}')\in\mathcal{P}} \mathbf{1}_{k_\mathfrak{l}=k_{\mathfrak{l}'}}$$

  where $\mathcal{P}$ is a partition of $\mathcal{L}_1\cup\mathcal{L}_2$ into pairs $(\mathfrak{l},\mathfrak{l}')$ such that $\mathfrak{l}$ has sign $+$ and $\mathfrak{l}'$ has sign $-$. We prove (quantitative) alternatives for this lemma in the non-Gaussian case that account for over-pairing of the leaves (see [D-H: Prop. of Chaos]).

- As such, we have

$$\mathbb{E}a_k^{\mathcal{T}_1}(t)\overline{a_k^{\mathcal{T}_2}(t)} = \zeta_{\mathcal{T}_1}\zeta_{\mathcal{T}_2}\left(\frac{\delta}{2L^{d-\gamma}}\right)^{n_1+n_2} \sum_{\mathcal{P}} \sum_{(k_{\mathfrak{n}})\in\mathcal{D}} B\left(t,(\Omega_{\mathfrak{n}})_{\mathfrak{n}\in\mathcal{N}_1\cup\mathcal{N}_2}\right) \times \prod_{\mathfrak{l}\in\mathcal{L}_1\cup\mathcal{L}_2}^{+} n_{\text{in}}(k_{\mathfrak{l}})$$

  where
  - $\mathcal{P}$ runs over all pairings of the leaves in $\mathcal{L}_1 \cup \mathcal{L}_2$ so that paired leaves have opposite signs
  - $\mathcal{D}$ is the union of decorations of the two trees $\mathcal{T}_1$ and $\mathcal{T}_2$ such that $k_{\mathfrak{l}} = k_{\mathfrak{l}'}$ if $(\mathfrak{l}, \mathfrak{l}') \in \mathcal{P}$.
  - $B\left(t,(\Omega_{\mathfrak{n}})_{\mathfrak{n}\in\mathcal{N}_1\cup\mathcal{N}_2}\right) := A\left(t,(\Omega_{\mathfrak{n}})_{\mathfrak{n}\in\mathcal{N}_1}\right)\overline{A\left(t,(\Omega_{\mathfrak{n}})_{\mathfrak{n}\in\mathcal{N}_2}\right)}$.
  - $\prod^{+}$ runs over all leaves with sign $+$.
- **Couples:** We now define the couple $\mathcal{Q}$ to be the triplet $(\mathcal{T}_1, \mathcal{T}_2, \mathcal{P})$, i.e. it is a couple of trees with their leaves paired. Also, define the order n of a couple to be $n = n_1 + n_2$ where n_j is the order of $\mathcal{T}_j$. A decoration $\mathcal{D}$ as above is now called a decoration of the couple.



- With this in hand, we can summarize

$$\mathbb{E}|a_k(t)|^2 = \sum_{\mathcal{Q}} \left(\frac{\delta}{2L^{d-\gamma}}\right)^{2n} \zeta_{\mathcal{Q}} \sum_{(k_\mathfrak{n})\in\mathcal{D}} B\left(t, (\Omega_\mathfrak{n})_{\mathfrak{n}\in\mathcal{N}_\mathcal{Q}}\right) \prod_{\mathfrak{l}\in\mathcal{L}_\mathcal{Q}}^{+} n_{\mathrm{in}}(k_\mathfrak{l}) + \text{rem. terms}$$

$$= \sum_{\mathcal{Q}} \mathcal{K}_\mathcal{Q}(t) + \text{remainder terms}$$

  - $\sum_\mathcal{Q}$ is over all couples $\mathcal{Q}$ of two trees of order $\leqslant N$.
  - $\zeta_\mathcal{Q} = \zeta_{\mathcal{T}_1}\zeta_{\mathcal{T}_2}$, $\mathcal{N}_\mathcal{Q} = \mathcal{N}_1 \cup \mathcal{N}_2$ and $\mathcal{L}_\mathcal{Q} := \mathcal{L}_1 \cup \mathcal{L}_2$.
- **Key fact:** there are C^n trees of order n, but there are $C^n n!$ couples of order $2n$.

The first iterate of (WKE)

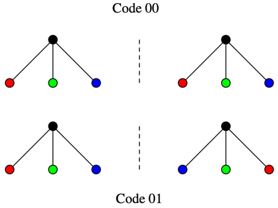
- Computing $\mathbb{E}|a_k^{(1)}(t)|^2$ (here we withhold the time rescaling so $0 \leqslant t \leqslant \delta\alpha^{-2}$)

$$(\alpha^2 t^2) \cdot \quad L^{-2d} \sum_{k_1 - k_2 + k_3 = k} n_{\text{in}}(k_1) n_{\text{in}}(k_2) n_{\text{in}}(k_3) \left(\frac{\sin(\pi t \Omega)}{\pi t \Omega} \right)^2$$

$$\sim (\alpha^2 t) \cdot \quad \int_{k_1 - k_2 + k_3 = k} n_{\text{in}}(k_1) n_{\text{in}}(k_2) n_{\text{in}}(k_3) \underbrace{t \left(\frac{\sin(\pi t \Omega)}{\pi t \Omega} \right)^2}_{\widetilde{A}(t\Omega);\ \widetilde{A} := (\sin(\pi x)/\pi x)^2 \in L^1}$$

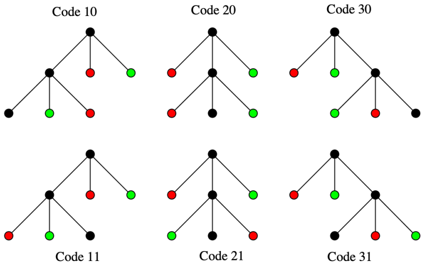
$$\to \underbrace{\alpha^2 t}_{=\delta} \int_{k_1 - k_2 + k_3 = k} n_{\text{in}}(k_1) n_{\text{in}}(k_2) n_{\text{in}}(k_3) \delta_{\mathbb{R}}(\Omega) dk_1 dk_2 dk_3.$$

which is part of the first iterate of the wave kinetic equation. This part comes from the couples



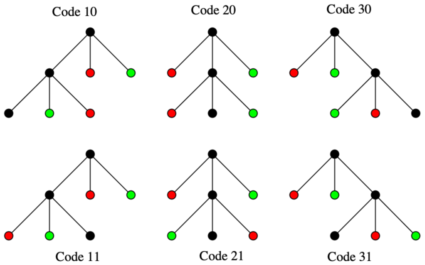
- We call those couples (1,1) minicouples.

- The rest of the first iterate comes from $\mathbb{E}\, a_k^{(0)}\overline{a_k^{(2)}}$ and $\mathbb{E}\, a_k^{(2)}\overline{a_k^{(0)}}$, which are represented by the pairing of the following trees with a trivial tree with one node.



- We call such diagrams the (2,0) minicouples. Those minicouples ((1,1) and (2,0)) converge to the first iterate of the kinetic equation.
- As such, all the remaining iterates of the (WKE) should be obtained **only** from couples constructed using the minicouples as building blocks.
- **Regular Couples** are exactly such couples. They are built by attaching the minicouples above in the natural way.

- The rest of the first iterate comes from $\mathbb{E}\, a_k^{(0)}\overline{a_k^{(2)}}$ and $\mathbb{E}\, a_k^{(2)}\overline{a_k^{(0)}}$, which are represented by the pairing of the following trees with a trivial tree with one node.



- We call such diagrams the (2,0) minicouples. Those minicouples ((1,1) and (2,0)) converge to the first iterate of the kinetic equation.
- As such, all the remaining iterates of the (WKE) should be obtained **only** from couples constructed using the minicouples as building blocks.
- **Regular Couples** are exactly such couples. They are built by attaching the minicouples above in the natural way.

Sums converging to integrals

- In the above computation, we needed to approximate a sum over the lattice $\mathbb{Z}_L^d$ by an integral over $\mathbb{R}^d$. This took the caricature form

$$L^{-2d} \sum_{(k_1,k_2)\in\mathbb{Z}_L^{2d}} W(k_1,k_2)\chi(T\Omega) \sim \int_{\mathbb{R}^{2d}} W(k_1,k_2)\chi(T\Omega)\,dk_1 dk_2,$$

  where Ω is a quadratic form like $\Omega(k_1,k_2) = \langle Ak_1, k_2\rangle$, and χ is some cutoff function (assume to be C_0^∞).
- For this inequality to hold, we need the equidistribution of the lattice points $\mathbb{Z}_L^{2d}$ in the region
$$\{(k_1,k_2)\in\mathbb{R}^{2d} : |\Omega(k_1,k_2)| \lesssim T^{-1}\}.$$
- Depending on the how large T is, and the diophantine nature of A, this can be a deep question in analytic number theory. For example, if $A = Id$ (square torus), then we don't expect this to be true if $T \geqslant L^2$ (since $\Omega \in L^{-2}\mathbb{Z}$).
- This leads to the condition $T_{\text{kin}} \ll L^2$ on the square torus, which is the range $0 \leqslant \gamma < 1$ on the scaling law.
- If the torus is generically irrational, then T can be as large as L^{d-}, which gives the bigger range $0 \leqslant \gamma < d/2^-$.

Difficulties for closing the math proof

- Regular couples of size n have the sharp estimate $\delta^{n/2}$. So the best uniform estimate we can hope to prove for $u^{(n)}$ is $\lesssim \delta^{n/2}$.
- **Probabilistic criticality.** This implies that the iterate $u^{(n+1)}$ is only better than the iterate $u^{(n)}$ by a factor of $\sqrt{\delta} = O(1)$. This is called probabilistic criticality (cf. works on stochastic PDE). In fact, this seems to be the first resolution of a (non-equilibrium) probabilistically critical problem, even in the parabolic setting.
- **Factorial divergence.** Unfortunately, there are factorially many couples of size n, so if we only use the uniform estimate $\delta^{n/2}$, we are doomed to failure. This is where the heart of the proof lies.
- **Diagrammatic Cancellations.** The regular couples are not the only large couples out there. There are two other families of couples that have maximal or almost maximal estimates, and those do not converge to the (WKE) iterates! To resolve this, one has to uncover the elaborate cancellations between those couples (at arbitrarily large size).

Thanks for your attention![1]