Séminaire Laurent Schwartz

EDP et applications

Année 2024-2025

Exposé n° XVI (18 mars 2025)

Yvain Bruned

DECORATED TREES, ARBORIFICATION FOR CANCELLATIONS IN WAVE TURBULENCE

https://doi.org/10.5802/slsedp.180

Institut des hautes études scientifiques
Le Bois-Marie • Route de Chartres
F-91440 BURES-SUR-YVETTE
http://www.ihes.fr/

Centre de mathématiques Laurent Schwartz
CMLS, CNRS, École polytechnique,
Institut Polytechnique de Paris
F-91128 PALAISEAU CEDEX
https://cmls.ip-paris.fr/

Publication membre du
Centre Mersenne pour l'édition scientifique ouverte
www.centre-mersenne.org
e-ISSN : 2266-0607

Séminaire Laurent-Schwartz — EDP et applications
Centre de mathématiques Laurent Schwartz, 2024-2025
Exposé n° XVI, 1-15

Decorated trees, arborification for cancellations in wave turbulence

Yvain Bruned

Universite de Lorraine, CNRS, IECL, F-54000 Nancy, France
Email: yvain.bruned@univ-lorraine.fr.

Abstract

In this work, we review part of the results obtained in [12] for computing cancellations for dispersive PDEs with random initial data. The idea is to get a new combinatorial perspective on the cancellations discovered by Deng and Hani (see [16]) in the context of Wave Turbulence when one wants to derive rigorously wave-kinetic equations. This new perspective is based on decorated trees developed for low regularity schemes, together with a well-chosen arborification map that rewrites these trees into linear combinations of words. With this new combinatorial basis, one develops graphical rules to compute cancellations.

Contents

1 Introduction

Perturbative expansions based on decorated trees and Feynman diagrams have become one of the main tools in the context of Wave Turbulence. In [15, 16, 17], Deng and Hani developed a rigorous justification for wave-kinetic equations. The main strategy of the proof involves summing infinitely many Feynman diagrams built out of some decorated trees. More recently, using the same ideas, a long-time derivation of the Boltzmann equation has been provided in [18]. The main combinatorics developed in these works are molecules, which are a type of Feynman diagrams. All the careful analysis for obtaining bounds on these diagrams before resummation is performed via a sophisticated cutting algorithm on these molecules. This formalism is also used for detecting cancellations in [16, Section 3.3], which they refer to as "miraculous cancellations". The authors have to come back to the iterated integrals associated with these molecules for computing them.

In [12], an alternative combinatorial formalism has been proposed for computing the cancellations between Feynmann diagrams observed in [16, Section 3.3]. The aim of the seminar was mainly to present the part of [12] that covers [16] as other cancellations could be understood via the formalism developed in the present note (see Remark 1.1), which shows the large scope of such an approach.

The main idea is to introduce a general unified framework for computing and understanding the cancellations coming from Wave Turbulence. We start by recalling the decorated tree formalism used from [11] for encoding low regularity schemes for dispersive PDEs, that are schemes that minimise the regularity on the initial data by embedding the resonance into the discretisation. This allows us to expand the k-th Fourier coefficient of the solution of a dispersive PDEs in the form of a B-series. This expansion is formed of oscillatory integrals which are multi-linear in the random Gaussian initial data. This is the subject of Section 2, which is written with the cubic non-linear Schrödinger equation as the main example for this paper. However, the formalism proposed could be applied to any dispersive PDEs. The decorated tree formalism together with the B-series expansion are extension of the classical B-series for ODEs (see [13]). Moreover, this formalism draws its inspiration from decorated trees and B-series that appeared in Regularity Structures, when one wants a systematic way to solve singular SPDEs (see [24, 8, 3]).

In Section 3 equipped with this expansion, one wants to understand the behaviour of the quantity of interest in Wave Turbulence that is $\mathbb{E}(|u_k|^2)$. We use the Wick formula for computing this expectation as the random initial data is Gaussian. We can write then an expansion using the formalism of [1] which replaces decorated trees by pairs of decorated trees where the leaves come in pairs. These pairings among the initial data come from the Wick formula. Then, the main idea of [12] is to change the perspective by using a crucial identity given in Proposition 3.1 by

$$e^{i(s-t)k^2} = \mathbb{E}(e^{-itk^2}\eta_k \overline{e^{-isk^2}\eta_k}), \tag{1.1}$$

where the η_k are i.i.d Gaussian complex random variables and it is the type of noise used for randomizing the initial data. One applies the identity (1.1) for each propagator inside the iterated integrals. One can interpret this repeated analytical transformation in a combinatorial way via a well-chosen arborification map $\mathfrak{a}$. This map allows us to move from decorated trees to words on a well-chosen alphabet. These words will also encode iterated integrals, but these integrals be simpler, as now they are on a simplex of the form $0 < t_1 < ... < t_n < t$ where the t_i are the time variables of integration. The definition of $\mathfrak{a}$ is dictated by the form of the equation and the identity (1.1). This map is inspired by the arborification used in [10] for rewriting the Poincaré-Dulac normal form proposed in [22] for dispersive equations. The idea of arborification was first introduced by Ecalle in the study of dynamical systems (see [19, 20]). Let us also stress that the idea of the arborification is to repeatedly cut some edges in a trees and to put the decorated node as letters. This is a reminiscence of the Butcher-Connes-Kreimer Hopf algebra (see [13, 14]) used for renormalising Feynman Diagrams and understanding the composition of

B-series. It is also similar in spirit to the various cutting algorithms designed for Wave Turbulence in [15] (see the proof of the Rigidity Theorem) and for the Boltzmann equation in [18].

Once the tree-based iterated integrals have been rewritten via words, one can conduct combinatorial arguments (change of the colors for some leaves, order on some letters) in order to compute the cancellation observed in [16]. We terminate Section 3 with several examples of computations with these combinatorial rules. We end this introduction with a couple of remarks that provide some perspectives.

Remark 1.1 The formalism developed in [12] can also be applied to the cancellations that appear in [4] which showed that the Φ^4_3 (Gibbs) measure is invariant under the dynamics of the three-dimensional cubic wave equation. Indeed, it is possible to derive an identity similar to (1.1) for the wave equation with random initial data and to rewrite the iterated integrals appearing in the expansion of the solution of the wave equation with the help of words. The graphical rules are a bit different, as one has to proceed with some integration by parts. This is in agreement with the idea that one cannot ignore the specificities of various dispersive equations. Therefore, some parts of the combinatorics depend crucially on the dispersive PDEs. One can still state a MetaTheorem (see [12, Metathm. 1]) saying that a well-chosen arborification and words formalism are an essential tool for computing cancellations for dispersive PDEs.

Remark 1.2 In the context of parabolic singular SPDEs, cancellations have also been computed with some graphical rules. It started in [23] where a hidden logarithmic cancellation was first observed for the KPZ equation. The approach started to become more systematic in [25, 26] as more renormalisation constants have to be computed. Therefore, graphical rules have to be introduced. These rely on the fact that the heat kernel K is non-anticipative, therefore, it loops in some oriented Feynman diagrams, which allows improved estimates. Moreover, one uses the following relation in the context of the KPZ type equations

$$(\partial_x K * \partial_x K)(z) = \frac{1}{2}(K(z) + K(-z)) \tag{1.2}$$

where $*$ is the space-time convolution. One does not use exactly this identity in the computation, but a version true up to a small error (See [26, Lemma 6.11]). Equipped with this formalism, one is able to compute and check cancellations in [9] for the generalised KPZ equation. It boils down to tedious computations on Feynman diagrams when one repeatedly applies (1.2) and some integrations by parts in order to reduce the diagram to a primitive diagram where rules cannot be applied anymore. This cancellations allows us to consider solutions that are “geometric”, meaning that they satisfy the chain rule property.

Remark 1.3 The cancellations observed for the generalsied KPZ in [9] identities have been pushed further in [21] with general integration by parts identities. They are used for quasilinear SPDEs to show the locality in the solution of the counter-terms that appear in the renormalised equation (see [7] for a general statement). These cancellations have been understood at a more conceptual level in [6], where the chain rule symmetry has been characterised as the kernel of a linear map defined on decorated trees. The dimension of this kernel and its basis are computed only for space-time white noise in [6]. The full subcritical regime is treated in a systematic way via operad theory and homological algebra in [5]. The specific case of dimension one is considered in [2] with multi-indices and elementary techniques.

Understanding cancellations for singular SPDEs via symmetries for some singular SPDEs in the full subcritical regime has been obtained only recently and requires advanced algebra. One expects to explain the cancellation obtained in [16, 4] as a consequence of symmetry coming from the equation. One can think about the symplectic nature of the equation. It is still an open question to get a more fundamental argument that justifies these cancellations.

Acknowledgements

Y. B. gratefully acknowledges funding support from the European Research Council (ERC) through the ERC Starting Grant Low Regularity Dynamics via Decorated Trees (LoRDeT), grant agreement No. 101075208. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. Y.B. also thanks the "Institut des Hautes Etudes Scientifiques" ´ (IHES) for a long research stay from 7th of January to 21st of March 2025.

2 Decorated trees for dispersive PDEs

In this section, we introduce the decorated trees and B-series formalism when one expands solutions of dispersive PDEs with iterated integrals. The formalism exposed here is coming from [11]. We focus on the cubic non-linear equation Schrödinger equation without loss of generality. This equation is given by

$$(\partial_t + i\Delta)u = i\mu^2|u|^2u, \quad u(0,x) = v(x). \tag{2.1}$$

where $x \in \mathbb{T}_L^d = [0,L]^d$. The random initial data v is given by

$$v(x) = \frac{1}{L^d}\sum_{k\in\mathbb{Z}_L^d} v_k e^{2\pi i k x}, \quad v_k = \sqrt{w_k}\eta_k$$

where $\mathbb{Z}_L^d = (L^{-1}\mathbb{Z})^d$ and $w : \mathbb{R}^d \to [0,+\infty)$ is a given Schwartz function. The η_k are i.i.d centred complex Gaussian random variables satisfying for $k, \ell \in \mathbb{Z}_L^d$

$$\mathbb{E}(|\eta_k|^2) = 1, \quad \mathbb{E}(\eta_k\eta_\ell) = 0.$$

Here, the parameter μ^2 is the strength of the non-linear interaction and L is the size of the box considered. The initial data v is said to be well-prepared in the context of Wave Turbulence. The idea is to understand the behaviour of $\mathbb{E}(|u_k|^2)$ when the size of the box L tends to infinity and μ to zero. One expects $\mathbb{E}(|u_k|^2)$ to solve a kinetic-wave equation up to a certain kinetic time. We keep these notations as the cancellations we are interested in were found in the context of Wave Turbulence in [16]. Equation (2.1) can be rewritten in Duhamel form as

$$u(t) = e^{it\Delta}v + i\mu^2 \int_0^t e^{i(t-s)\Delta}|u(s)|^2 u(s)ds.$$

In Fourier space, one has

$$u_k(t) = e^{-itk^2}v_k + i\mu^2 \sum_{k=-k_1+k_2+k_3} \int_0^t e^{-i(t-s)k^2}\bar{u}_{k_1}(s)u_{k_2}(s)u_{k_3}(s)ds \quad (2.2)$$

where pointwise product in physical space $\bar{u}u^2$ is sent to convolution product in Fourier space. Here, k_1 comes with a minus sign in $k = -k_1 + k_2 + k_3$ due to the conjugate $\bar{u}$. Moreover, $e^{it\Delta}$ is sent to e^{-itk^2} and $e^{-is\Delta}$ is sent to e^{isk^2} . One iterates (2.2) by replacing $u_{k_j}(t)$ by

$$u_{k_j}(t) = e^{-itk_j^2}v_{k_j} + \mathcal{O}(t),$$

with $j \in \{1, 2, 3\}$. We obtain the following first order approximation of the k-th Fourier coefficient $u_k(t)$:

$$u_k(t) = e^{-itk^2}v_k + i\mu^2 \sum_{k=-k_1+k_2+k_3} e^{-itk^2}$$
$$\int_0^t e^{isk^2}(e^{isk_1^2}\bar{v}_{k_1})(e^{-isk_2^2}v_{k_2})(e^{-isk_3^2}v_{k_3})ds + \mathcal{O}(t^2).$$

One can encode the previous Duhamel iterates using a decorated tree series. We denote by $U_k^r(v,t)$ the first iterated integrals of size r of the Duhamel expansion. These are integrals with r integrations in time. One has

$$|u_k(t) - U_k^r(v,t)| = \mathcal{O}(t^{r+1})$$

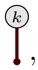
where the regularity asked on the initial data hidden in the notation $\mathcal{O}$ corresponds to the regularity needed to define the first iterated integrals up to order r. Decorated trees are used to provide a precise description of $U_k^r(v,t)$. One uses the following B-series type formula

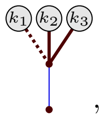
$$U_k^r(v,t) = \sum_{T\in\mathcal{T}_k^r} \frac{\Upsilon(T)(v)}{S(T)}(\Pi T)(t) \quad (2.3)$$

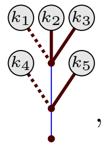
where $\mathcal{T}_k^r$ is a suitable set of decorated trees of size r, $S(T)$ is a symmetry factor and $\Upsilon(T)$ is an elementray differential associated with T depending on the initial data v. The map Π sends T to an oscillatory integral. Such a formalism is reminiscence of the B-series (named after Butcher) that describes numerical schemes for ODEs and PDEs. It has been introduced in [11] for describing low regularity schemes. It is also largely inspired from the treatment of singular stochastic partial differential equations (SPDEs) via regularity structures in [24, 8, 3] where a local ansatz for the solutions takes a similar form. One can notice that the decorated trees encode at the same time iterated integrals $(\Pi T)(t)$ and elementary differentials $\Upsilon(T)(v)$. Below, we describe the series for $r = 2$. One has

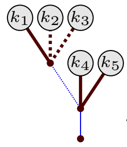
$$\mathcal{T}_k^2 = \{T_0, T_1, T_2, T_3, \;\; k_i \in \mathbb{Z}_L^d\},$$

where

$T_0 =$ 

, $T_1 =$ 

, $T_2 =$ 

, $T_3 =$ 

.

where for T_1, $k = -k_1 + k_2 + k_3$, for T_2, $k = -k_4 - k_1 + k_2 + k_3 + k_5$ and for T_3, $k = k_4 + k_1 - k_2 - k_3 + k_5$. An edge (resp.) corresponds to a factor e^{-itk^2} (resp. e^{itk^2}), while an edge (resp.) corresponds to an integral $i\mu^2 \int_0^t e^{-i(t-s)k^2} \cdots ds$ (resp. $-i\mu^2 \int_0^t e^{i(t-s)k^2} \cdots ds$). The dotted edges can be seen as taking the complex conjugate of the operator. Also, the frequencies add up to the root with a minus sign for dotted edges. Indeed, for the decorated tree T_3, one has

$$k = k_4 - \ell_3 + k_5, \quad -\ell_3 = k_1 - k_2 - k_3$$

where ℓ_3 corresponds to the node decoration of the inner nodes not connected to the root. When one interprets these decorated trees as iterated integrals, one has to order the time variables following the partial order given by the decorated tree. If two blue edges lie on the same path to the root, the edge closer to the root corresponds to a variable in time bigger than the one associated with the other blue edge. One has

$$\begin{aligned}
(\Pi T_0)(t) &= e^{-itk^2} \\
(\Pi T_1)(t) &= i\mu^2 \int_0^t e^{-i(t-s)k^2} e^{is(k_1^2 - k_2^2 - k_3^2)} ds \\
(\Pi T_2)(t) &= -\mu^4 \int_0^t e^{-i(t-s)k^2} e^{is(k_4^2 - k_5^2)} \int_0^s e^{-i(s-r)\ell_2^2} e^{ir(k_1^2 - k_2^2 - k_3^2)} dr ds \\
(\Pi T_3)(t) &= \mu^4 \int_0^t e^{-i(t-s)k^2} e^{-is(\ell_4^2 + k_5^2)} \int_0^s e^{i(s-r)\ell_3^2} e^{ir(-k_1^2 + k_2^2 + k_3^2)} dr ds.
\end{aligned} \tag{2.4}$$

The size of a decorated tree T is denoted by $|T|$ and corresponds to the number of blue edges in T. This is also the number of integrations in time inside $(\Pi T)(t)$. One has

$$(\Pi T)(t) = \mathcal{O}(t^{|T|}).$$

The symmetry factor $S(T)$ corresponds to the number of internal symmetries of the tree T taking the edge decorations into account but not the node decorations. One obtains

$$S(T_0) = 1, \quad S(T_1) = 2, \quad S(T_2) = 2, \quad S(T_2) = 4.$$

The 2 in $S(T_1)$ and $S(T_2)$ comes from the fact that one can permute the two leaves decorated by k_2 and k_3. For T_3, one can permute k_4 and k_5 in addition which gives an extra factor 2. The elementary differential $\Upsilon(T)(v)$ corresponds to a product of initial data associated with the leaves of a decorated tree. One has to take into account also a factor connected to the structure of the tree: In the case of the NLS equation, this factor is 2 for each node . One gets

$$\begin{aligned}
\Upsilon(T_0)(v) &= \eta_k \sqrt{w_k}, \quad \Upsilon(T_1)(v) = 2\bar{\eta}_{k_1}\eta_{k_2}\eta_{k_3} \prod_{j=1}^{3} \sqrt{w_{k_j}}, \\
\Upsilon(T_2)(v) &= 2\bar{\eta}_{k_1}\bar{\eta}_{k_4}\eta_{k_2}\eta_{k_3}\eta_{k_5} \prod_{j=1}^{5} \sqrt{w_{k_j}}, \\
\Upsilon(T_3)(v) &= 4\bar{\eta}_{k_2}\bar{\eta}_{k_3}\eta_{k_1}\eta_{k_4}\eta_{k_5} \prod_{j=1}^{5} \sqrt{w_{k_j}}.
\end{aligned} \tag{2.5}$$

3 Cancellations via arborification

In the previous decorated trees, we have assumed that the frequencies on the leaves, the k_j are independent. This means that one has a summation for each of the k_j. In Wave turbulence, the quantity of interest is $\mathbb{E}(|u_k|^2)$. Using the B-series for u_k and truncating to the correct order, one gets:

$$\begin{aligned}
\mathbb{E}(|u_k|^2) = \sum_{T_1, T_2 \in \mathcal{T}_k^r} & \mathbf{1}_{\{|T_1|+|T_2| \leq r\}} \mathbb{E}\big(\Upsilon(T_1)(v)\overline{\Upsilon(T_2)(v)}\big) \\
& \frac{(\Pi T_1)(t)}{S(T_1)} \frac{\overline{(\Pi T_2)(t)}}{S(T_2)} + \mathcal{O}(t^{r+1}).
\end{aligned} \tag{3.1}$$

For computing the quantity $\mathbb{E}(\Upsilon(T_1)(v)\overline{\Upsilon(T_2)(v)})$, we recall the Wick formula for computing expectations of product of random Gausssian variables. Let I be a finite set and $(X_i)_{i\in I}$ a collection of centred jointly Gaussian random variables. Then

$$\mathbb{E}\left(\prod_{i\in I} X_i\right) = \sum_{\mathfrak{p}\in\mathcal{P}(I)} \prod_{\{i,j\}\in\mathfrak{p}} \mathbb{E}(X_i X_j) \tag{3.2}$$

where $\mathcal{P}(I)$ are partitions of I with two elements of I in each block of the partition. Below, we provide an example of computation for T_0 and T_1 defined in (2.4). One has from (2.5)

$$
\begin{aligned}
\mathbb{E}(\Upsilon(T_0)(v)\overline{\Upsilon(T_1)(v)}) &= \mathbb{E}(\eta_k \eta_{k_1} \bar{\eta}_{k_2} \bar{\eta}_{k_3}) \sqrt{w_k} \prod_{j=1}^{3} \sqrt{w_{k_j}} \\
&= (\mathbb{E}(\eta_k \bar{\eta}_{k_3}) \mathbb{E}(\eta_{k_1} \bar{\eta}_{k_2}) + \mathbb{E}(\eta_k \bar{\eta}_{k_2}) \mathbb{E}(\eta_{k_1} \bar{\eta}_{k_3})) \sqrt{w_k} \prod_{j=1}^{3} \sqrt{w_{k_j}} \\
&= \big(\mathbf{1}_{\{k=k_3\}\cap\{k_1=k_2\}} + \mathbf{1}_{\{k=k_2\}\cap\{k_1=k_3\}}\big) w_k w_{k_1}.
\end{aligned}
\tag{3.3}
$$

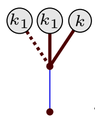
In the application of the Wick formula, we have excluded the terms of the form $\mathbb{E}(\eta_{k_i}\eta_{k_j})$ because they are zero due to the constraint on the noise η. We have also used the fact that

$$
\mathbb{E}(\eta_k \bar{\eta}_{k_3}) = \mathbf{1}_{\{k=k_3\}}.
$$

One can then rewrite (3.1) using pairs of trees that encode the pairings among the noise given by the previous indicators. One has

$$
\mathbb{E}(|u_k|^2) = \sum_{F = T_1 \cdot T_2 \in \mathcal{G}_k^r} m_F \frac{\hat{\Upsilon}(T_1)(v)}{S(T_1)} \overline{\frac{\hat{\Upsilon}(T_2)(v)}{S(T_2)}} (\Pi T_1)(t) \overline{(\Pi T_2)(t)} + \mathcal{O}(t^{r+1}).
$$

where T_1, T_2 are taken to be in $\mathcal{T}_k^r$ with $|T_1| + |T_2| \leq r$, and one assumes some pairing between the leaves of T_1 and T_2. The elementary differential $\hat{\Upsilon}$ is defined as the same as Υ but without the η_k. The symmetry factor S does not depend on the frequency decoration, therefore it is the same definition with the pairings. The factor m_F counts the number of pairings that give the same object. In the computation (3.3), one can symmetrise the result in k_2 and k_3 to see that one gets the same value twice. For the formula above, we are using the notations from [1]. Let us illustrate these new objects with an example $F = T_0 \cdot \hat{T}_1$ given by

$T_0 =$ , $\quad \hat{T}_1 =$  .

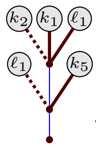
Then

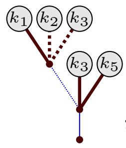
$$
\begin{aligned}
&(\Pi \hat{T}_1)(t) = it\mu^2 e^{-itk^2}, \quad S(\hat{T}_1) = S(T_1) = 2, \quad m_F = 2, \\
&\hat{\Upsilon}(T_0)(v) = \sqrt{w_k}, \quad \hat{\Upsilon}(\hat{T}_1)(v) = 2 w_{k_1} \sqrt{w_k}.
\end{aligned}
$$

In the following, we will use the shorthand notation:

$$
(\hat{\Pi} T)(v,t) := \frac{\hat{\Upsilon}(T)(v)}{S(T)} (\Pi T)(t).
$$

Now, we work in a more general case where some of the leaves come in pairs, but not necessarily all of them. We want to understand cancellations observed in [16] between the associated oscillatory integrals and to propose a new systematic way to compute them. As an example, let us consider the following decorated trees:

$\hat{T}_2 =$ 

, $\hat{T}_3 =$ 

,

where $-\ell_1 = k_1 - k_2 - k_3$ and $k = k_1 - k_2 + k_5$. The oscillatory integrals are given by

$$\begin{aligned}(\hat{\Pi}\hat{T}_2)(v,t) &= -\mu^4 w_{\ell_1}\sqrt{w_{k_1}}\sqrt{w_{k_2}}\sqrt{w_{k_5}} \\ &\qquad \int_0^t e^{-i(t-s)k^2} e^{is(k_5^2-\ell_1^2)} \int_0^s e^{-i(s-r)k_3^2} e^{ir(k_2^2-k_1^2-\ell_1^2)} dr ds \\ (\hat{\Pi}\hat{T}_3)(v,t) &= \mu^4 w_{k_3}\sqrt{w_{k_1}}\sqrt{w_{k_2}}\sqrt{w_{k_5}} \\ &\qquad \int_0^t e^{-i(t-s)k^2} e^{is(k_3^2+k_5^2)} \int_0^s e^{i(s-r)\ell_1^2} e^{ir(k_2^2-k_1^2-k_3^2)} dr ds\end{aligned}$$

If we suppose that $|k_3 - \ell_1| \leq L^{-1}$, one can make the following identification up to a small error:

$$w_{\ell_1} \approx w_{k_3}.$$

One can notice from the explicit expression of the iterated integrals described above the following cancellation:

$$\hat{\Pi}(\hat{T}_2 + \hat{T}_3)(v,t) \approx 0.$$

This is exactly the first of the three (families of) “miraculous cancellations” appearing in [16]. We want to derive a combinatorial formalism that explains this and the other cancellations. We first start with a simple observation that rewrites any internal edge of the previous tree as two edges:

Proposition 3.1 *One has*

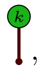
$$e^{i(s-t)k^2} = \mathbb{E}(e^{-itk^2}\eta_k \overline{e^{-isk^2}\eta_k}). \tag{3.4}$$

Proof. It is an immediate consequence of the definition of the noises η_k. Indeed, one has

$$\mathbb{E}(e^{-itk^2}\eta_k \overline{e^{-isk^2}\eta_k}) = e^{i(s-t)k^2}\mathbb{E}(\overline{\eta_k}\eta_k) = e^{i(s-t)k^2}.$$

□

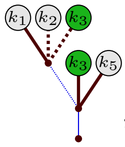
Then, the consequence of the previous proposition is that one can view the propagator $e^{i(s-t)k^2}$ as a pairing of T_0^c given by

$T_0^c =$ , $\quad (\hat{\Pi} T_0^c)(v,t) = e^{-itk^2}\eta_k.$

Here, for a leaf colored in green, we omit the $\sqrt{w_k}$ in the interpretation of the decorated tree as an oscillatory integral. Then

$$e^{i(s-t)k^2} = \mathbb{E}\Big((\hat{\Pi} T_0^c)(v,t) \overline{(\hat{\Pi} T_0^c)(v,t)} \Big).$$

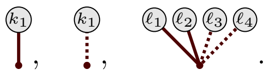
Below, we provide another example of a decorated tree with the green color on some leaves and give its oscillatory integral

$\hat{T}_3^c =$ , (3.5)

and

$$(\hat{\Pi}\hat{T}_3^c)(v,t) = \mu^4 \sqrt{w_{k_1}}\sqrt{w_{k_2}}\sqrt{w_{k_5}} \int_0^t e^{-i(t-s)k^2} e^{is(k_3^2+k_5^2)} \int_0^s e^{i(s-r)\ell_1^2} e^{ir(k_2^2-k_1^2-k_3^2)} dr ds.$$

Now, the main idea for computing cancellations is to apply Proposition 3.1 to the propagators $e^{-i(t-s)k^2}$ that correspond to the blue internal edges of a decorated tree. We introduce a new combinatorial structure based on words that will be appropriate for describing the oscillatory integrals after this transformation. We consider words on an alphabet A whose letters are given by:

 (3.6)

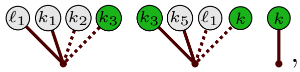
where for the last letter, one must have

$$\ell_1 + \ell_2 - \ell_3 - \ell_4 = 0. \tag{3.7}$$

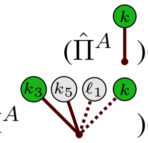
The third letter of (3.6) is close in spirit to what is happening for the hard sphere dynamic used for the long-time derivation of the Boltzmann equation (see [18]). Indeed, one can assume that the plain edges correspond to two particles that get a shock and the dotted edges to the particles after the shock. The identity (3.7) shows a conserved quantity in this dynamics. This is one of the fundamental combinatorial reasons why the algorithms developed for the non-linear Schrödinger wave-kinetic equation (see [15, 16, 17]) also work for Boltzmann.

They are also in A letters where some leaves have been colored in green. We consider the words whose rightmost letter is taken among the first two letters of

(3.6), on this alphabet that we denote by $T(A)$. Below, we provide an example of such a word:

$$W_2 =$$


with $\ell_1 + k_1 - k_2 - k_3 = 0$ and $k_3 + k_5 - \ell_1 - k = 0$. One interprets the previous words as integrals in time over a simplex via a map $\hat{\Pi}^A$. One defines the map $\hat{\Pi}^A$ inductively on the construction of the words by starting with the letters $a \in A$:



$$(\hat{\Pi}^A \;)(v,t) = e^{-itk^2},$$
$$(\hat{\Pi}^A \;)(v,t) = \sqrt{w_{k_5}}\sqrt{w_{\ell_1}}e^{-it(k_3^2+k_5^2-\ell_1^2-k^2)}.$$

One has more letters but their interpretation follows the same rules as above. Then, for a non-empty word W and a letter $a \in A$, one has

$$(\hat{\Pi}^A W a)(v,t) = \mu^2 (\hat{\Pi}^A a)(v,t) \int_0^t (\hat{\Pi}^A W)(v,s) ds.$$

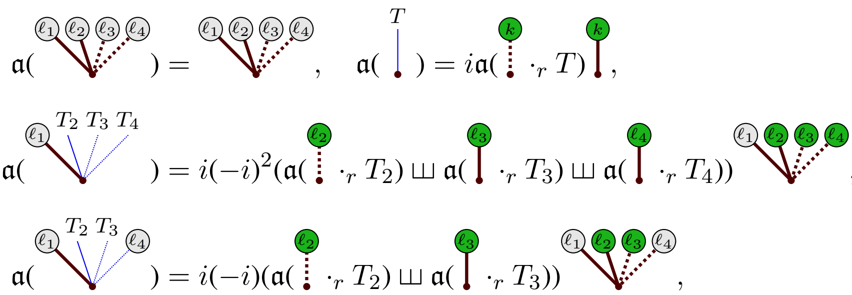
where Wa is the word that has for rightmost letters a and the rest of the word is given by W. It is the concatenation of W with a. Then

$$(\hat{\Pi}^A W_2)(v,t) = \mu^4 e^{-itk^2} \int_0^t \sqrt{w_{k_5}}\sqrt{w_{\ell_1}} e^{-is(k_3^2+k_5^2-\ell_1^2-k)}$$
$$\int_0^s \sqrt{w_{\ell_1}}\sqrt{w_{k_1}}\sqrt{w_{k_2}} e^{-ir(\ell_1^2+k_1^2-k_2^2-k_3^2)} dr ds.$$

One introduces a product $⧢$ on $T(A)$ called shuffle product. It is given inductively for two words au and bv with $a, b \in A$ by:

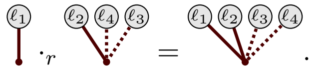
$$au ⧢ bv = a(u ⧢ bv) + b(au ⧢ v), \quad a ⧢ \mathbf{1} = \mathbf{1} ⧢ a = a. \tag{3.8}$$

Here, $\mathbf{1}$ denotes the empty word, the neutral for $⧢$. Now, we define a natural map $\mathfrak{a}$ called arborification between the decorated trees and $T(A)$. One defines $\mathfrak{a}$ as

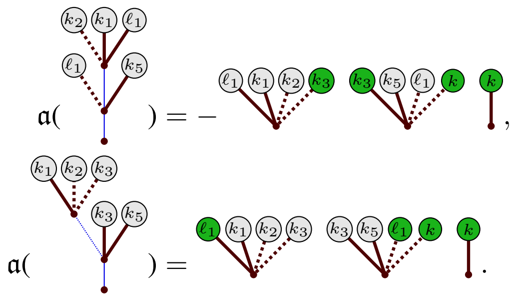


(3.9)

where we have assumed that $k, \ell_2, \ell_3, \ell_3$ are respectively the frenquencies associated with the roots of T, T_2, T_3, T_4. The product $\cdot_r$ is the merging root product by taking two decorated trees and identifying their root. One has for example



Let us comment briefly on the recursive formula (3.9). Every blue edge in the decorated tree is duplicated into two brown edges. This is a combinatorial version of a repeated application of Proposition 3.1. One has a factor i for each plain blue edge and a factor $-i$ for each dotted blue edge. The shuffle product is used for transforming integrals over a tree-shaped domain for integrals in time into an integral over a simplex. Using this transformation, one can have the following words:



A similar arborification map has been used in [10] for describing Poincaré-Dulac Normal forms. One observes that

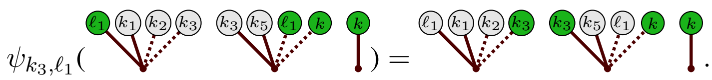
$$(\hat{\Pi}\hat{T}_2)(v,t) = (\hat{\Pi}^A \mathfrak{a}(\hat{T}_2))(v,t) = (\hat{\Pi}^A W_2)(v,t).$$

The next theorem shows that this identity is true in general and connects the two formalisms via decorated trees and words:

Theorem 3.2 *One has for every decorated tree T*

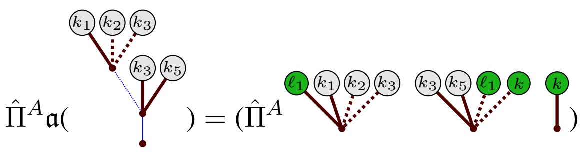
$$(\hat{\Pi}T)(v,t) = (\hat{\Pi}^A \mathfrak{a}(T))(v,t).$$

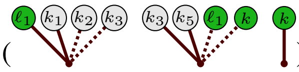
This theorem has been first stated in [12, Thm 2.3] where one can find a proof of this result. The previous theorem allows us to switch from decorated trees to words which could be easier to compute cancellations. For example, making the identification $w_{k_3} \approx w_{\ell_1}$ boils down to intervene the color of two pairs of leaves. This can be encoded via a linear map on words that we denote by ψ_{k_3,ℓ_1}. Below, we show one example of computation

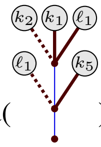


(3.10)

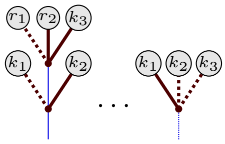
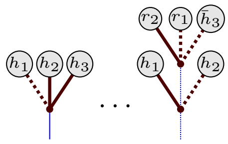
This allows us to show a combinatorial proof of the first cancellation:



$$\approx (\hat{\Pi}^A \psi_{k_3,\ell_1}($$  $$))$$

$$\approx -\hat{\Pi}^A \mathfrak{a}($$  $$).$$

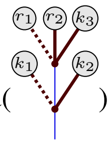
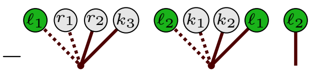
This cancellation is up to a small error coming from the approximation $w_{k_3} \approx w_{\ell_1}$. The second cancellation in [16] is between two pairs of decorated trees described below

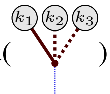
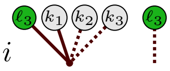
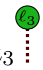
$T_1 =$  and $T_2 =$ 

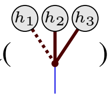
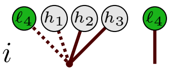
where $\ldots$ means that these two branches are connected to bigger trees. This cancellation happens under the condition that the two node decorations at the base of each of the trees are the same, i.e.

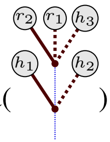
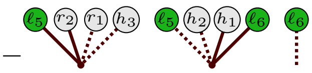
$$-k_1 + k_2 + k_3 + r_2 - r_1 = -h_1 + h_2 + h_3,$$

and that the trees containing the above subtrees are otherwise identical. We have that

$\mathfrak{a}($  $) = -$  $= -a_1 a_2$ 

$\mathfrak{a}($  $) = -i$  $= -i a_3$ 

$\mathfrak{a}($  $) = i$  $= i b_1$ 

$\mathfrak{a}($  $) = -$  $= -b_2 b_3$ 

where the condition

$$-k_1 + k_2 + k_3 + r_2 - r_1 = -h_1 + h_2 + h_3$$

can be rewritten as

$$\ell_2 = \ell_4, \quad \ell_3 = \ell_6.$$

The last branches with a green node in the computation above belong in general to a bigger letter. We now make the choice

$$h_1 = k_1, \quad h_2 = k_2, \quad h_3 = \ell_1 = k_3 + r_2 - r_1.$$

This way, we obtain that

$$a_1 = b_2, \quad a_2 = b_1, \quad a_3 = b_3.$$

From the definition of $\mathfrak{a}$ in (3.9), if T_1, T_2 are the two trees respectively, there exist u, v in the word algebra $T(A)$ such that

$$\mathfrak{a}(T_1) = i(a_1 a_2 \sqcup\!\sqcup\, a_3 \sqcup\!\sqcup\, u)v$$
$$\mathfrak{a}(T_2) = -i(b_1 \sqcup\!\sqcup\, b_2 b_3 \sqcup\!\sqcup\, u)v = -i(a_2 \sqcup\!\sqcup\, a_1 a_3 \sqcup\!\sqcup\, u)v.$$

Therefore, from the definition of the shuffle product (3.8), we have

$$\mathfrak{a}(T_1) + \mathfrak{a}(T_2) = i(a_3 a_1 a_2) \sqcup\!\sqcup\, u)v - i(a_2 a_1 a_3 \sqcup\!\sqcup\, u)v. \tag{3.11}$$

In particular, we observe that the terms with the letters a_1, a_2, a_3 in this order cancel out. It turns out that from an analytical point of view, these are all the "problematic" terms, in the sense that

$$\hat{\Pi}^A(\mathfrak{a}(T_1) + \mathfrak{a}(T_2))$$

can be estimated directly using [16, Lemma 7.1]. A third cancellation can also be computed in the same way (see [12, Sec. 2]).

References

[1] Y. Alama Bronsard, Y. Bruned, K. Schratz. *Approximations of Dispersive PDEs in the Presence of Low-Regularity Randomness.* Found. Comput. Math. **24**, (2024), 1819–1869. doi:10.1007/s10208-023-09625-8.

[2] C. Bellingeri, Y. Bruned. *Symmetries for the gKPZ equation via multi-indices* arXiv:2410.00834.

[3] Y. Bruned, A. Chandra, I. Chevyrev, M. Hairer. *Renormalising SPDEs in regularity structures.* J. Eur. Math. Soc. (JEMS), **23**, no. 3, (2021), 869-947. doi:10.4171/JEMS/1025.

[4] B. Bringmann Y. Deng, A. Nahmod, H. Yue . *Invariant Gibbs measures for the three dimensional cubic nonlinear wave equation.* Invent. Math. **236**, (2024), 1133–1411. doi:10.1007/s00222-024-01254-4.

[5] Y. Bruned, V. Dotsenko. *Chain rule symmetry for singular SPDEs* arXiv:2403.17066.

[6] Y. Bruned, F. Gabriel, M. Hairer, L. Zambotti, *Geometric stochastic heat equations.* J. Amer. Math. Soc. (JAMS) **35**, no. 1, (2022), 1-80. doi:10.1090/jams/977.

[7] Y. Bruned, M. Gerencsér, U. Nadeem *Quasi-generalised KPZ equation* arXiv:2401.13620.

[8] Y. Bruned, M. Hairer, L. Zambotti. *Algebraic renormalisation of regularity structures.* Invent. Math. **215**, no. 3, (2019), 1039–1156. doi:10.1007/s00222-018-0841-x.

[9] Y. Bruned. *Singular KPZ Type Equations.* 205 pages, PhD thesis, Université Pierre et Marie Curie - Paris VI, 2015. https://tel.archives-ouvertes.fr/tel-01306427.

[10] Y. Bruned. *Derivation of normal forms for dispersive PDEs via arborification*. arXiv:2409.03642.

[11] Y. Bruned, K. Schratz. *Resonance based schemes for dispersive equations via decorated trees*. Forum of Mathematics, Pi, 10, (2022), E2. doi:10.1017/fmp.2021.13.

[12] Y. Bruned, L. Tolomeo *Cancellations for dispersive PDEs with random initial data*. arXiv:2412.17051.

[13] J. C. Butcher. *An algebraic theory of integration methods.* Math. Comp. **26**, (1972), 79–106. doi:10.2307/2004720.

[14] A. Connes, D. Kreimer. *Hopf algebras, renormalization and noncommutative geometry.* Comm. Math. Phys. **199**, no. 1, (1998), 203–242. doi:10.1007/s002200050499.

[15] Y. Deng, Z. Hani. *Full derivation of the wave kinetic equation*. Invent. math. **233**, (2023), 543-724. doi:10.1007/s00222-023-01189-2.

[16] Y. Deng, Z. Hani. *Derivation of the wave kinetic equation: Full range of scaling laws.* arXiv:2301.07063.

[17] Y. Deng, Z. Hani. *Long time justification of wave turbulence theory.* arXiv:2408.07818.

[18] Y. Deng, Z. Hani, X. Ma. *Long time derivation of Boltzmann equation from hard sphere dynamics.* arXiv:2408.07818.

[19] J. Ecalle, B. Vallet. *The arborification-coarborification transform: analytic, combinatorial, and algebraic aspects*. Ann. Fac. Sci. Toulouse (6) **13**, no. 3, (2004), 575–657.

[20] F. Fauvet and F. Menous, *Ecalle's arborification coarborification transforms and Connes Kreimer Hopf algebra*, Annales Sc. de l'Ecole Normale Sup. **50**, no. 1, (2017), 39-83. doi:10.24033/asens.2315.

[21] M. Gerencsér. *Nondivergence form quasilinear heat equations driven by space-time white noise*. Annales de l'Institut Henri Poincaré, Analyse non linéaire **37**, no. 3. (2020), 663–682 doi:10.1016/j.anihpc.2020.01.003.

[22] Z. Guo, S. Kwon, T. Oh. *Poincaré-Dulac normal form reduction for unconditional well-posedness of the periodic cubic NLS.* Comm. Math. Phys. **322**, no. 1, (2013), 19–48. doi:10.1007/s00220-013-1755-5.

[23] M. Hairer. *Solving the KPZ equation.* Ann. Math **178**, no. 2, (2013), 559–664. doi:10.4007/annals.2013.178.2.4.

[24] M. Hairer. *A theory of regularity structures.* Invent. Math. **198**, no. 2, (2014), 269–504. doi:10.1007/s00222-014-0505-4.

[25] M. Hairer, E. Pardoux. *A Wong-Zakai theorem for stochastic PDEs.* J. Math. Soc. Japan, **67**, no. 4, (2015), 1551–1604. doi:10.2969/jmsj/06741551.

[26] M. Hairer, J. Quastel. *A class of growth models rescaling to KPZ.* Forum of Mathematics, Pi, 6, (2018), E3. doi:10.1017/fmp.2018.2.