

THE WAVE KINETIC THEORY OF THREE WAVE
AND FOUR WAVE MODELS

XIAO MA

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
MATHEMATICS
ADVISER: ALEXANDRU IONESCU

MAY 2023

© Copyright by Xiao Ma, 2023.

All Rights Reserved

Abstract

In this thesis, we introduce new techniques for studying the random series expansion of dispersive PDEs. We take a quadratic KdV type and a cubic Klein-Gordon type equation as examples to demonstrate the different techniques in three wave and four wave models. For three wave models, we introduce a counting argument to handle the degeneracy problems of the resonance surface and the loss of derivative problem. For four wave models, we introduce a novel renormalization argument and prove a renormalized Wick theorem. We provide a heuristic argument that this renormalization is able to remove all bad terms from the L^2 mass term, combining with an almost cancellation identity of the regular pairing and the Deng-Hani's Feynman diagram analysis [9], [13].

Acknowledgements

I would like to express my gratitude and appreciation to my Ph.D. advisor, Prof. Alexandru Ionescu, for his support, guidance, and mentorship throughout the journey of my Ph.D. program. I feel extremely fortunate to have had the opportunity to work under his supervision.

I am also grateful to the professors with whom I have had the privilege of interacting during the journey of my Ph.D. program, for their invaluable guidance, constructive feedback, and encouragement. My thanks go to Professors Tristan Buckmaster, Yu Deng, Zaher Hani, Hao Jia, Igor Rodnianski, Jalal Shatah, Ze Li and Lifeng Zhao for their enlightening discussions and generous assistance.

My heartfelt thanks go to my fellow math Ph.D. students, who have made this journey all the more enjoyable and memorable. I am particularly thankful for the friendship and support of Shaoyun Bai, Jiajie Chen, Shengwen Gan, Vikram Giri, Lili He, Kexin Jin, Yangyang Li, Federico Pasqualotto, Shuangping Li, Jia Shi, Zhihan Wang, Xiaoxu Wu, Lingfu Zhang, Junyan Zhang, Fan Zheng, and many others I encountered who have been there for me during the highs and lows of this journey. I would like to express my special gratitude to Zhihan Wang, Kexin Jin, and Fan Zheng. During the five years Ph.D. program, Zhihan Wang and Kexin Jin shared uncountable many invaluable and genius ideas from which I have learned a great deal. Fan Zheng, as my senior, has taught me much during my first year, and his mentorship has been instrumental in my growth as a researcher.

During my time in the Ph.D. program, I have been fortunate to be surrounded by an amazing group of friends who have not only inspired me but also enriched my graduate school experience. I would like to express my gratitude to Yixiao Chen, Yifan Chen, Zihan Ding, Hansheng Jiang, Shiji Lyu, Zhou Lu, Xiaohan Li, Qinghua Liu, Pinchen Xie, Shouda Wang, Baojun Wu, Qinxin Yan, Yucheng Yang, Bingjia Yang, Zhiyang Yuan, Xiyu Zhai and many others I encountered. I would like to

extend a special thank you to Yifan Chen, for her support during the Covid pandemic. Additionally, I also want to express my special appreciation to Yixiao Chen, Zihan Ding, Qinghua Liu, Shouda Wang, and Yucheng Yang for frequently giving me rides and spending leisure time.

Finally, I dedicate this thesis to my loving family, who has been my constant source of motivation and support. As always, I owe much more to my family than I can ever hope to give back. Thank you, mom, dad, for your unconditional love and support!

To my family and friends.

Contents

Abstract	iii
Acknowledgements	iv
1 Introduction	1
1.1 The physics of interacting wave	1
1.1.1 Dispersive PDEs	1
1.1.2 The energy cascade phenomenon	3
1.1.3 The statistical description of interacting wave	5
1.1.4 Scale of turbulence	6
1.1.5 Three-wave and four-wave models	7
1.2 The Feynman diagram expansion	7
1.3 Wave kinetic theory for three wave models	12
1.4 Wave kinetic theory for four wave models	13
1.4.1 The phase problem	13
1.4.2 The renormalization argument	14
1.4.3 From the Klein-Gordon type equation to the original Klein-Gordon equation	15
1.5 Previous research	16
2 Analysis of three wave models	18
2.1 Introduction	18

2.1.1	Statement of the results	18
2.1.2	Ideas of the proof	22
2.1.3	Notations	27
2.2	The Perturbation Expansion	28
2.2.1	The approximation series and Feynman diagrams	28
2.2.2	Estimates of the approximation solution	36
2.2.3	Bound the error of the approximation	38
2.2.4	Proof of the main theorem	40
2.3	Lattice points counting and convergence results	41
2.4	Lattice points counting and convergence results	41
2.4.1	Refined expression of coefficients	42
2.4.2	An upper bound of coefficients in expansion series	45
2.4.3	Couples and the Wick theorem	50
2.4.4	Counting lattice points	55
2.4.5	An upper bound of tree terms	81
2.4.6	Norm estimate of random matrices	89
2.4.7	Asymptotics of the main terms	110
3	Analysis of four wave models	119
3.1	Introduction	119
3.1.1	Basic set-ups	119
3.1.2	Outline of this chapter	120
3.1.3	Notations	128
3.2	The Perturbation Expansion	129
3.2.1	The renormalization argument	129
3.2.2	The approximation series and Feynman diagrams	132
3.3	Lattice points counting and convergence results	141
3.3.1	Refined expression of coefficients	142

3.3.2	Couples and the renormalized Wick theorem	148
3.3.3	Counting lattice points	161
3.3.4	A heuristic explanation of the derivation of the wave kinetic equation	193
4	Counting results in wave kinetic theory	199
4.1	Number theoretic results in three wave models	199
4.2	High Dimensional Euler-Maclaurin Formula	205
4.3	Number theoretic results in four wave models	207

Chapter 1

Introduction

1.1 The physics of interacting wave

1.1.1 Dispersive PDEs

In this thesis, we will mostly be interested in the long time behavior of dispersive PDEs on a finite domain.

$$i\partial_t\phi = \Lambda(\nabla)\phi + \lambda B(\phi, \phi) \text{ (or } \lambda^2 C(\phi, \phi, \phi)) \quad (1.1)$$

Examples of equations of the above type are listed below.

1. The nonlinear Schrodinger equations

$$i\partial_t\psi - \Delta\psi = \pm\lambda^2|\psi|^2\psi \quad (1.2)$$

2. The ZK equation

$$\partial_t\psi(t, x) + \Delta\partial_{x_1}\psi(t, x) - \nu\Delta\psi(t, x) = \lambda\partial_{x_1}(\psi^2(t, x)) \quad (1.3)$$

3. The nonlinear Klein-Gordon equation

$$\partial_{tt}u - \Delta u = \pm \lambda^2 u^3 \text{ (or } u^2) \quad (1.4)$$

4. The rotating Navier-Stokes equation (we only give the incompressible equation)

$$\begin{cases} \partial_t u + u \cdot \nabla u + e_z \times u + \nabla p = \nu \Delta u, \\ \operatorname{div} u = 0. \end{cases} \quad (1.5)$$

5. Internal wave in the stratified fluid

$$\begin{cases} \partial_t u + u \cdot \nabla u + \frac{1}{\rho} \nabla p = \frac{\nu}{\rho} \Delta u, \\ \partial_t \rho + u \cdot \nabla \rho = 0 \\ \operatorname{div} u = 0. \end{cases} \quad (1.6)$$

6. The MHD equation

$$\begin{cases} \partial_t u + u \cdot \nabla u + \nabla p = b \cdot \nabla b + \nu \Delta u, \\ \partial_t b + u \cdot \nabla b = b \cdot \nabla u + \eta \Delta b, \\ \operatorname{div} u = 0, \quad \operatorname{div} b = 0.. \end{cases} \quad (1.7)$$

7. The gravity water wave equation (we do not include the equation due to its complication).

In the rest part of this introduction, we **take the nonlinear Schrodinger equation as an example** to explain the idea of wave turbulence.

1.1.2 The energy cascade phenomenon

It is well-known that in many PDEs coming from physics, energy can be transferred from low-frequency Fourier modes to high-frequency modes. We take the nonlinear Schrödinger equation as an example to explain this.

Its Fourier coefficients satisfy

$$\dot{\psi}_k = i|k|^2\psi_k + \frac{i\lambda^2}{L^{2d}} \sum_{\substack{(k_1, k_2, k_3) \in (\mathbb{Z}_L^d)^3 \\ k_1 - k_2 + k_3 = k}} \psi_{k_1} \overline{\psi_{k_2}} \psi_{k_3} \quad (1.8)$$

ψ_{k_1} , ψ_{k_2} , ψ_{k_3} can "collide" with each other and transfer energy to ψ_k . For example, energy can transfer to ψ_6 in the "collision" of ψ_1 , ψ_2 , ψ_3 .

This can be repeated many times

$$\psi_1 \rightarrow \psi_2 \rightarrow \psi_4 \rightarrow \psi_7 \rightarrow \cdots \rightarrow \psi_{100} \rightarrow \cdots . \quad (1.9)$$

This process can be repeated many times so that finally the energy can be injected at a very high frequency. This is called the energy cascade phenomenon. Since it takes many steps to transfer energy to very high modes, the information in the low-frequency modes is lost and high-frequency modes become random and exhibit *universality*. This is very similar to the game, *Galton board*, as illustrated in Figure 1.1. Each step of energy transfer is similar to a collision of a ball with a pin. To get to the bottom (the high energy modes), a ball must collide with many pins (there must be many steps of energy transfer). Finally, the distribution of balls (energy) becomes random and universal.

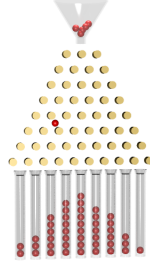


Figure 1.1: The Galton board

In the Galton board case, the final distribution is the Gaussian distribution. This distribution is universal in the sense that it is independent of the details of the input flow of balls. Although in many examples of energy cascade the final distribution is random and universal, they are usually not i.i.d Gaussian or even do not exhibit properties similar to i.i.d Gaussian.

The cascade spectrum of strong turbulence in the fluid is one example of highly non-Gaussian universality behavior. The structure functions satisfy power laws which are independent of the initial data. The second-order structure function satisfies the famous Kolmogorov 5/3 law while the high-order structure functions exhibit a significant deviation from Kolmogorov's prediction. This deviation is called *intermittency*.

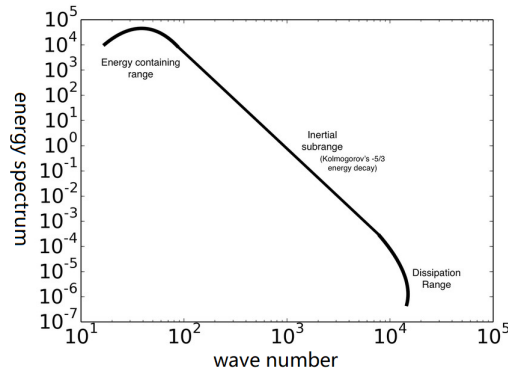


Figure 1.2: The 5/3 law

While i.i.d Gaussian assumption does not work for turbulence in fluid, it can work for turbulence generated by dispersive waves. Unlike a fluid equation, a dispersive equation can exhibit turbulent behavior even when the size of the solution is small.

In this case, the distribution can be calculated by a perturbative expansion and it is close to i.i.d Gaussian variables if it is so initially. This kind of turbulence is the so-called *wave turbulence*.

1.1.3 The statistical description of interacting wave

In the study of wave turbulence of the nonlinear Schrodinger equation, we assume that the initial data of the PDE satisfies

$$\psi_{\text{in},k} = \sqrt{n_{\text{in}}(k)}\eta_k(\omega) \quad (1.10)$$

where $\eta_k(\omega)$ are iid normal complex Gaussian variables.

Then the solution ψ is a random variable. Let ψ_k be its Fourier coefficients, then we define the energy spectral to be $n(t, k) = \mathbb{E}|\psi_k|^2$. The goal of this thesis is to study the evolution of this energy spectral.

For a more general equation, the ansatz of initial data can be imposed similarly. Here is a list of possible modifications,

1. If the solution ψ is a complex value function, then the ansatz is the same.
2. If the solution ψ is a real value function, then to ensure the real value property, we assume that $n_{\text{in}}(k) = n_{\text{in}}(-k)$, $\eta_k = \overline{\eta_{-k}}$ and η_k is independent of $\{\eta_{k'}\}_{k' \neq k, -k}$.
3. For second-order equations, we typically introduce a new variable to reduce it into a first-order equation as in section 1.4, and then impose the same ansatz as first-order equations.
4. For a system of equations with more than 1 variable, we assume that the Fourier coefficients of one component are independent of those of another component.

These modifications are not the only canonical choices. The idea of proof of other choices is similar to the above choices.

1.1.4 Scale of turbulence

The energy distribution $n(k)$ looks very different in different ranges. In general, the frequency domain can be divided into 3 parts, the *energy-containing range*, the *inertial range* and the *dissipation range*.

As in Figure 1.2, in the energy-containing range $|k| \ll 1$, the wave number k is not large and the energy distribution in this range can be influenced by the external force and initial data. There is no universality in this range.

In the inertial range $1 \lesssim |k| \ll l_d^{-1}$, the same as the Galton's board, the influence of external force and initial data starts to be lost, while the wave number is not large enough to activate the effect of dissipation. The energy distribution is universal and satisfies the *wave kinetic equation* (WKE) given below.

In the dissipation range $|k| \gtrsim l_d^{-1}$, the effect of dissipation is dominant and $n(k)$ decays to 0 very fast. In this thesis, we are mostly interested in the energy distribution in the inertial and dissipation range.

In wave turbulence, the energy distribution $n(k)$ is supposed to evolve according to the following wave kinetic equation

$$\begin{aligned} \partial_t n(t, k) &= \mathcal{K}(n(t, \cdot)), \\ \mathcal{K}(n)(k) &:= \int_{\substack{(k_1, k_2) \in \mathbb{R}^{2d} \\ k_1 + k_2 = k}} |H_{k_1 k_2}|^2 n(k_1) n(k_2) \delta(\Lambda(k_1) + \Lambda(k_2) - \Lambda(k)) dk_1 dk_2 \\ \text{or } \int_{\substack{(k_1, k_2, k_3) \in \mathbb{R}^{3d} \\ k_1 - k_2 + k_3 = k}} |H_{k_1 k_2 k_3}|^2 n(k_1) n(k_2) n(k_3) \delta(\Lambda(k_1) - \Lambda(k_2) + \Lambda(k_3) - \Lambda(k)) dk_1 dk_2 dk_3 \end{aligned} \quad (\text{WKE})$$

(WKE) admits a special solution $|k|^{-\nu}$ and the general solution of (WKE) is supposed to converge to this power law in some regimes. This power law is universal and independent of the detail of external force and initial data.

The (WKE) is a Boltzmann-type equation. Compared to the original ZK equation,

(WKE) has a monotonically decreasing entropy (see [24]) and exhibits irreversibility. The Boltzmann equation statistically describes the collision of particles and in the collision integral there are momentum and energy conservation conditions $p_1 + p_2 = p$ and $p_1^2 + p_2^2 = p^2$. (WKE) is a counterpart that describes the interaction of waves. The right-hand side of (WKE) also contains momentum and energy conservation $k_1 + k_2 = k$, $\Lambda(k_1) + \Lambda(k_2) = \Lambda(k)$. The energy conservation is the time resonance surface in the space-time resonance method.

1.1.5 Three-wave and four-wave models

For a general dispersive PDE, if the leading order term in the nonlinearity is quadratic and the solution of the resonance surface $\Lambda(k_1) + \Lambda(k_2) = \Lambda(k)$ is non-empty, then we say that this PDE is a three-wave model. If the leading order term in the nonlinearity is cubic and the solution of the resonance surface is non-empty, then we say that this PDE is a four-wave model.

In the example of dispersive PDEs in section 1.1.1, the quadratic equations (1.3), (1.5), (1.6), (1.7) are examples of the 3 wave interaction case. the cubic equations (1.2), (1.4) are examples of 4 wave interaction cases. The quadratic (1.4) equation is a four-wave model because the three wave resonance surface $\sqrt{1 + |k_1|^2} + \sqrt{1 + |k_2|^2} = \sqrt{1 + |k|^2}$ is empty in this case.

In this thesis, we choose the ZK equation as an example of 3 wave turbulence theory and the Klein-Gordon equation as an example of 4 wave turbulence theory.

1.2 The Feynman diagram expansion

Using Duhamel's formula, we can rewrite the NLS into the following form.

$$\phi = \xi + \mathcal{T}(\phi, \phi, \phi) \tag{1.11}$$

$$\mathcal{T}(\phi, \phi, \phi)_k = \frac{i\lambda^2}{L^{2d}} \sum_{\substack{(k_1, k_2, k_3) \in (\mathbb{Z}_L^d)^3 \\ S(k_1, k_2, k_3, k)=0}} \int_0^t \phi_{k_1} \overline{\phi_{k_2}} \phi_{k_3} e^{2\pi i \Omega(k_1, k_2, k_3, k)s} ds \quad (1.12)$$

where

$$\begin{aligned} S(k_1, k_2, k_3, k) &= k_1 - k_2 + k_3 - k, \\ \Omega(k_1, k_2, k_3, k) &= \Lambda(k_1) - \Lambda(k_2) + \Lambda(k_3) - \Lambda(k). \end{aligned} \quad (1.13)$$

The WKE is proved in several steps.

Step 1. Iterate the equation to obtain an approximation series $\phi = \sum_{l(T) \leq N} J_T$.

Given an equation, $\phi = \xi + C(\phi, \phi, \phi)$, we can find an approximation series by iteration,

$$\begin{aligned} \phi &= \xi + C(\phi, \phi, \phi) \\ &= \xi + C(\xi, \xi, \xi) + C(C(\xi, \xi, \xi), \xi, \xi) + C(\xi, C(\xi, \xi, \xi), \xi) + \cdots \\ &= \sum_{l(T) \leq N} J_T \end{aligned} \quad (1.14)$$

Each term is a polynomial of ξ that can be represented by a tree diagram T . The detail of the tree diagram expansion can be found in the next two chapters.

Step 2. Calculate the expectation of $|\phi_k|^2$ and compare terms in the expectation with the iteration of WKE.

Here is the detail of the second step. Integrate and iterate the WKE

$$\partial_\tau n(\tau, k) = \mathcal{K}(n, n, n), \quad (1.15)$$

then we get

$$\begin{aligned} n(\tau, k) &= n_{\text{in}} + \tau \mathcal{K}(n_{\text{in}}, n_{\text{in}}, n_{\text{in}}) + \frac{\tau^2}{2} \mathcal{K}(\mathcal{K}(n_{\text{in}}, n_{\text{in}}, n_{\text{in}}), n_{\text{in}}, n_{\text{in}}) + \cdots \\ &= \mathcal{K}^{(0)} + \mathcal{K}^{(1)} + \cdots \end{aligned} \quad (1.16)$$

By $\phi = \sum_{l(T) \leq N} J_T$, we know that

$$E|\phi_k(t)|^2 = \sum_n \sum_{l(T)+l(T')=n} E(J_{T,k} \overline{J_{T',k}}) \quad (1.17)$$

To prove WKE, we just need to show that

$$\sum_{l(T)+l(T')=n} E(J_{T,k} \overline{J_{T',k}}) \rightarrow K^{(n)} \quad (1.18)$$

All tree terms are Gaussian polynomials. (For a precise statement, see Lemma 2.4.2)

$$\mathcal{J}_{T,k} = \left(\frac{i\lambda}{L^d}\right)^{2d} \sum_{k_1, k_2, \dots, k_{2d+1}} H_{k_1 \dots k_{2d+1}}^T \xi_{k_1} \xi_{k_2} \dots \xi_{k_{2d+1}} \quad (1.19)$$

We can show that the coefficients $H_{k_1 \dots k_{2d+1}}^T$ concentrate near the resonance surface $A = \{S_{n_1}(T) = 0, \Omega_{n_1}(T) = 0, \dots, S_{n_n}(T)(T) = 0, \Omega_{n_n}(T) = 0\}$. (For a precise statement, see Lemma 2.4.4)

To calculate $E(J_{T,k} \overline{J_{T',k}})$, we just need to know how to calculate

$$\mathbb{E}\left(\xi_{k_1} \overline{\xi_{k_2}} \dots \xi_{k_{l(T)+1}} \xi_{k'_1} \overline{\xi_{k'_2}} \dots \xi_{k'_{l(T)+1}}\right) \quad (1.20)$$

This is done by applying the Wick theorem, Theorem 2.4.6.

Step 3. Analyze the size of $E(J_{T,k} \overline{J_{T',k}})$.

To prove (1.18), it suffices to get a good estimate on $E(J_{T,k} \overline{J_{T',k}})$.

The effect of applying the Wick theorem to Gaussian polynomials is the following square root cancellation principle. (Notice that in the Wick theorem, Lemma 2.4.6, paired indices are equal to each other, which means that half of the indices can be eliminated, so the size of the sum is the square root of the trivial bound)

Since the coefficients H in (1.19) concentrate near the resonance surface A , we have the following heuristic

$$\begin{aligned}
\mathcal{J}_{T,k} &= \left(\frac{i\lambda}{L^d}\right)^{2d} \sum_{k_1, k_2, \dots, k_{2d+1}} H_{k_1 \dots k_{2d+1}}^T \xi_{k_1} \xi_{k_2} \dots \xi_{k_{2d+1}} \\
&\sim \left(\frac{it\lambda}{L^d}\right)^{2d} \sum_{k_1, k_2, \dots, k_{2d+1} \in A} \xi_{k_1} \xi_{k_2} \dots \xi_{k_{2d+1}}
\end{aligned} \tag{1.21}$$

Given a sum of Gaussian monomials, its size is of order $O((\#A)^{\frac{1}{2}})$.

$$\sum_{(k_1, k_2, \dots, k_{2l+1}) \in A} \xi_{k_1} \overline{\xi_{k_2}} \dots \xi_{k_{2l+1}} \tag{1.22}$$

Therefore, to get a good estimate of $\mathcal{J}_{T,k}$, we just need to count the number of lattice points near the resonance surface. This can be done using the following volume principle. Given a set $A \subseteq (\mathbb{R}_L^d)^l$, the number of lattice points in this set is of order $\text{Vol}(A)/\text{Vol}(\text{cell})$. In our case,

$$A = \{(k_1, k_2, k_3) \in (\mathbb{Z}_L^d)^3 : k_1 - k_2 + k_3 = k, |\Omega(k_1, k_2, k_3, k)| \leq t^{-1}\}. \tag{1.23}$$

Therefore, $\text{Vol}(A) = t^{-1}$ and $\text{Vol}(\text{cell}) = L^{-2d}$. The number of solutions $\#A = L^{2d}/t$

For a precise statement of the count estimate, see Proposition 2.4.15.

As long as we get the desired counting estimate, the WKE is almost proven.

An example of applying the square root cancellation and volume principle. We can calculate the order of magnitude of the first two terms in (1.14).

The first term ξ_k is of order $O(1)$. We hope that the second term is a lower order term, which implies that $\frac{\lambda^2}{L^d} t^{\frac{1}{2}} \leq O(1)$ and $t \leq \left(\frac{L^d}{\lambda^2}\right)^2$. This is why we call $\frac{L^d}{\lambda^2}$ the

wave kinetic time.

$$\begin{aligned}
& \frac{i\lambda^2}{L^{2d}} \sum_{\substack{(k_1, k_2, k_3) \in (\mathbb{Z}_L^d)^3 \\ k_1 - k_2 + k_3 = k}} \int_0^t \xi_{k_1} \overline{\xi_{k_2}} \xi_{k_3} \underbrace{e^{2\pi i \Omega(k_1, k_2, k_3, k)s}}_{=1} ds \\
& \approx \frac{i\lambda^2 t}{L^{2d}} \sum_{\substack{(k_1, k_2, k_3) \in (\mathbb{Z}_L^d)^3 \\ k_1 - k_2 + k_3 = k \\ |\Omega(k_1, k_2, k_3, k)| \leq t^{-1}}} \xi_{k_1} \overline{\xi_{k_2}} \xi_{k_3} \\
& = \frac{\lambda^2 t}{L^{2d}} O \left(\# \left\{ \begin{array}{c} k_1 - k_2 + k_3 = k \\ |\Omega(k_1, k_2, k_3, k)| \leq t^{-1} \end{array} \right\} \right)^{\frac{1}{2}} \\
& = \frac{\lambda^2 t}{L^{2d}} (L^{2d}/t)^{\frac{1}{2}} = \frac{\lambda^2}{L^d} t^{\frac{1}{2}}
\end{aligned} \tag{1.24}$$

When executing the above steps, we will encounter several problems.

1. *The loss of derivative problem.* If the nonlinearity contains a derivative, we have the loss of derivative problem.
2. *The degeneracy problem.* In the above analysis of the second order term of the NLS equation, we actually cheated because we have the $k_1 = k_2$, $k_3 = k$ terms. A similar problem also arises in three wave models.

The first problem is solved by suitably taking advantage of the viscosity term in the equation. This argument given in section 2.4.6 can solve the loss of derivative problem for most PDEs.

The second problems in the three wave and four wave models setting are very different from each other. In the next two sections, we will discuss these two cases separately.

1.3 Wave kinetic theory for three wave models

We take the ZK equation (1.3) to explain the idea. The resonance surface for the ZK equation is

$$k_1 + k_2 = k, \quad |k_1|^2 k_{1x} + |k_2|^2 k_{2x} = |k|^2 k_x + O(T^{-1}) \quad (1.25)$$

From the volume principle, we hope the number of solutions can be bounded by $L^d T^{-1}$.

However, because of degeneracy near $k_x = 0$, the best estimate of number of solutions is $L^d T^{-1} |k_x|^{-1}$. We use the **null condition**, the fact that the multiplier k_x in the nonlinearity vanishes near k_x , to solve this problem.

WKE may not be true for equations with no null condition. Think about

$$\partial_t \psi(t, x) + \Delta \partial_{x_1} \psi(t, x) = \pm \lambda \psi^2(t, x). \quad (1.26)$$

which either blows up or decays to zero depending on the sign.

As mentioned before, for three wave systems, the resonance surfaces degenerate near $k_1 = \pm k_2 = 0, k = 0$. In [1], when studying the inhomogeneous wave turbulence of the following equation

$$i \partial_t u + \Lambda(\nabla) u = \lambda M(\nabla) (M(\nabla) \operatorname{Re} u)^2 \quad (1.27)$$

the authors propose a null condition that $M(0) = 0$ and use this property to solve the degeneracy problem.

The nonlinearity in [1] roughly equals to $\sum_{k_1+k_2=k} m(k_1)m(k_2)m(k)u_{k_1}u_{k_2}$, in which any of k_1, k_2, k equals to 0 implies $m(k_1)m(k_2)m(k)u_{k_1}u_{k_2} = 0$.

Typically, a nonlinearity from physics does not satisfy condition the condition in [1]. One example is $u|\nabla|u \sim \sum_{k_1+k_2=k} |k_2|u_{k_1}u_{k_2}$, in which case $|k_2|u_{k_1}u_{k_2} = 0$ only if $k_2 = 0$.

In chapter 2, we introduce a more complicated combinatorial argument to solve this problem.

1.4 Wave kinetic theory for four wave models

1.4.1 The phase problem

The resonance surfaces for four wave problems is

$$k_1 - k_2 + k_3 = k, \quad |k_1|^2 - |k_2|^2 + |k_3|^2 = |k|^2 + O(T^{-1}) \quad (1.28)$$

It degenerates seriously near the diagonal $k_1 = k_2, k_3 = k$. Notice that in $\sum_{k_1-k_2+k_3=k}$ of NLS, the sum over diagonal $\sum_{k_1=k_2, k_3=k}$ is

$$\sum_{k_1=k_2, k_3=k} \psi_{k_1} \overline{\psi_{k_2}} \psi_{k_3} = \left(\sum_{k_1} |\psi_{k_1}|^2 \right) \psi_k \approx L^d \quad (1.29)$$

This term is of order L^d which is much larger than

$$\sum_{\substack{k_1-k_2+k_3=k \\ |\Omega(k_1, k_2, k_3, k)| \leq t^{-1}}} \xi_{k_1} \overline{\xi_{k_2}} \xi_{k_3} \approx L^d t^{-\frac{1}{2}} \quad (1.30)$$

In NLS case, this diagonal term $\sum |\psi_{k_1}|^2$ is the conservative L^2 mass. Therefore, a change of phase argument $\psi_k \rightarrow \psi_k e^{i|k|^2 t + i t \alpha \|\psi\|_2^2}$ can remove this term completely.

$$\sum_{k_1=k_2, k_3=k} \psi_{k_1} \overline{\psi_{k_2}} \psi_{k_3} = \left(\sum_{k_1} |\psi_{k_1}|^2 \right) \psi_k \quad (1.31)$$

However, for general equations, the diagonal is of the form $\sum m_{k_1} |\psi_{k_1}|^2$ which is not a conservative quantity and cannot be removed by a naive change of phase argument.

For example, consider the following Klein-Gordon type equation.

$$i\partial_t\psi - \Lambda(\nabla)\psi = \lambda^2\Lambda(\nabla)^{-\frac{1}{2}}|\Lambda(\nabla)^{-\frac{1}{2}}\psi|^2\Lambda(\nabla)^{-\frac{1}{2}}\psi \quad (1.32)$$

The diagonal term is $\left(\sum_{k_1 \in \mathbb{Z}_L^d} \Lambda_{k_1}^{-1}|\psi_{k_1}|^2\right)\Lambda_k^{-1}\psi_k$ and $\sum_{k_1 \in \mathbb{Z}_L^d} \Lambda_{k_1}^{-1}|\psi_{k_1}|^2$ is not a conservative quantity.

1.4.2 The renormalization argument

In terms of Fourier coefficients, the equation (1.32) becomes

$$i\dot{\psi}_k = \left(\Lambda(k) + \frac{2\lambda^2}{L^d}M(t)\Lambda(k)^{-1}\right)\psi_k + \frac{\lambda^2}{L^{2d}} \sum_{S_3=0}^{\times} (\Lambda_{k_1}\Lambda_{k_2}\Lambda_{k_3}\Lambda_k)^{-\frac{1}{2}}\psi_{k_1}\bar{\psi}_{k_2}\psi_{k_3} \quad (1.33)$$

where $\sum^{\times} = \sum_{\substack{k_1, k_3 \neq k \\ S_3=0}} - \sum_{k_1=k_2=k_3=k}$ and $M(t) = \frac{1}{L^d} \sum_{k_1 \in \mathbb{Z}_L^d} \Lambda_{k_1}^{-1}|\psi_{k_1}|^2$.

The linear term $\left(\Lambda(k) + \frac{2\lambda^2}{L^d}M(t)\Lambda(k)^{-1}\right)\psi_k$ cannot be removed by changing phase.

Because $M(t)$ is a random variable.

To solve this issue, we propose a renormalization method. Instead of merging $M(t)$ into the linear part, we merge $m(t) = \mathbb{E}M(t)$

$$\begin{aligned} i\dot{\psi}_k = & \left(\Lambda(k) + \frac{2\lambda^2}{L^d}m(t)\Lambda(k)^{-1}\right)\psi_k + \frac{\lambda^2}{L^{2d}} \sum_{S_3=0}^{\times} (\Lambda_{k_1}\Lambda_{k_2}\Lambda_{k_3}\Lambda_k)^{-\frac{1}{2}}\psi_{k_1}\bar{\psi}_{k_2}\psi_{k_3} \\ & + \frac{2\lambda^2}{L^{2d}} \left(\sum_{k_1 \in \mathbb{Z}_L^d} \Lambda_{k_1}^{-1} \left(|\psi_{k_1}|^2 - \mathbb{E}|\psi_{k_1}|^2 \right) \right) \Lambda_k^{-1}\psi_k \end{aligned} \quad (1.34)$$

Now the linear term $\left(\Lambda(k) + \frac{2\lambda^2}{L^d}m(t)\Lambda(k)^{-1}\right)\psi_k$ can be removed by changing phase. The renormalized L^2 term $\left(\sum_{k_1 \in \mathbb{Z}_L^d} \Lambda_{k_1}^{-1} \left(|\psi_{k_1}|^2 - \mathbb{E}|\psi_{k_1}|^2 \right) \right) \Lambda_k^{-1}\psi_k$ do not have a phase problem due to the renormalized Wick theorem, Theorem 3.3.5.

1.4.3 From the Klein-Gordon type equation to the original Klein-Gordon equation

In this section, we explain the connection between the equation (1.32) and the Klein-Gordon equation (1.4).

Notice that in (1.4) $\partial_{tt}u - \Delta u + u = -\lambda^2 u^3$, the operator $\partial_{tt} - \Delta + 1$ can be factorized into $(-i\partial_t + \Lambda(\nabla)) \cdot (i\partial_t + \Lambda(\nabla))$, where $\Lambda(\nabla)$ is the Fourier multiplier with symbol $\Lambda(\xi) = \sqrt{1 + |\xi|^2}$. Therefore the equation can be simplified by introducing a new variable

$$\psi = \Lambda(\nabla)^{-\frac{1}{2}}(i\partial_t u + \Lambda(\nabla)u). \quad (1.35)$$

Then we have

$$(-i\partial_t + \Lambda(\nabla))\psi = (-i\partial_t + \Lambda(\nabla)) \cdot (i\partial_t u + \Lambda(\nabla))u = -\lambda^2 \Lambda(\nabla)^{-\frac{1}{2}} u^3. \quad (1.36)$$

We can solve u in terms of $\psi, \bar{\psi}$

$$u = (2\Lambda(\nabla))^{-\frac{1}{2}}(\psi + \bar{\psi}). \quad (1.37)$$

We can derive the equation of ψ by substituting (1.37) into the right hand side of (1.36)

$$i\partial_t \psi - \Lambda(\nabla)\psi = \frac{\lambda^2}{8} \Lambda(\nabla)^{-\frac{1}{2}} \left(\Lambda(\nabla)^{-\frac{1}{2}} \psi + \Lambda(\nabla)^{-\frac{1}{2}} \bar{\psi} \right)^3. \quad (1.38)$$

There are four terms in the nonlinearity,

$$\begin{aligned}
& \Lambda(\nabla)^{-\frac{1}{2}}(\Lambda(\nabla)^{-\frac{1}{2}}\psi)^3, \\
& \Lambda(\nabla)^{-\frac{1}{2}}(|\Lambda(\nabla)^{-\frac{1}{2}}\psi|^2\Lambda(\nabla)^{-\frac{1}{2}}\psi), \\
& \Lambda(\nabla)^{-\frac{1}{2}}(|\Lambda(\nabla)^{-\frac{1}{2}}\psi|^2\Lambda(\nabla)^{-\frac{1}{2}}\bar{\psi}), \\
& \Lambda(\nabla)^{-\frac{1}{2}}(\Lambda(\nabla)^{-\frac{1}{2}}\bar{\psi})^3.
\end{aligned} \tag{1.39}$$

The corresponding resonance surface of these terms are

$$\begin{aligned}
& \Lambda(k_1) + \Lambda(k_2) + \Lambda(k_3) - \Lambda(k) = 0, \\
& \Lambda(k_1) - \Lambda(k_2) + \Lambda(k_3) - \Lambda(k) = 0, \\
& \Lambda(k_1) - \Lambda(k_2) - \Lambda(k_3) - \Lambda(k) = 0, \\
& \Lambda(k_1) + \Lambda(k_2) + \Lambda(k_3) + \Lambda(k) = 0.
\end{aligned} \tag{1.40}$$

The resonance surface of the second term is the only one that contains degeneracy $k_1 = k_2, k_3 = k$.

1.5 Previous research

(1) *Early physics research.* In the physics literature, the inception of the kinetic description can be traced to the work of Peierls [34] in the study of anharmonic crystals, leading to the so-called phonon Boltzmann equation. Since then, the kinetic theory has been developed for various models and has become a systematic paradigm starting in the 1960s, with immense applications in various fields of physics and science [2], [3], [8], [26], [27], [28], [31], [33], [38], [39], [42], [45]. The name wave turbulence theory comes from the spectral energy dynamics and cascades that the wave kinetic equation predicts for nonlinear wave systems, which yields similar conclusions to Kolmogorov spectra in hydrodynamic turbulence; this connection was a major

contribution of Zakharov [44], [43].

(2) *Rigorous results about wave turbulence theory.* The WKE was first verified in the linear setting [19], [18], [37]. In the nonlinear setting, the WKE was rigorously verified for the Gibbs measure initial data by Lukkarinen and Spohn [30]. Then the basic concepts of general wave turbulence were rigorously formulated by Buckmaster, Germain, Hani, and Shatah [4] and a non-trivial result that verified WKE for a short time scale was also proved by them.

The WKE was proved for an almost sharp time scale independently by Deng and Hani [10] and Collot and Germain [6], [7] using the ideas from the study of randomly initialized PDE. The full WKE for the sharp time was proved independently by the deep works of Deng and Hani [9] and Staffilani and Tran [40] for a four-wave problem and a three-wave problem respectively. One key contribution of [9] and [40] was the classification of Feynman diagrams in the contexts of normal form expansion and Liouville equation respectively. Later in [12], [13], WKE for NLS was derived for all scaling laws by Deng and Hani.

WKE for the space-inhomogeneous case was derived by Ampatzoglou, Collot, and Germain [1] for an almost sharp time scale for a three wave problem. The result that reaching wave kinetic time was obtained by [25]. The higher-order correlation functions were studied by Deng and Hani [11]. A linearized wave kinetic equation near the Rayleigh-Jeans spectrum was derived by Faou [22]. The discrete wave turbulence was studied by Dymov and Kuksin [14], [17], [15], [16].

(3) *Rigorous results about the dynamics of WKE.* There are also many papers about the dynamics of WKE itself. For references, see [5], [20], [21], [24], [23], [36], [35], [37], [41] and the reference therein.

Chapter 2

Analysis of three wave models

2.1 Introduction

2.1.1 Statement of the results

In this chapter, we study the wave turbulence theory for the following KdV type equation

$$\begin{cases} \partial_t \psi(t, x) + \Delta \partial_{x_1} \psi(t, x) - \nu \Delta \psi(t, x) = \lambda \partial_{x_1} (\psi^2(t, x)), \\ \psi(0, x) = \psi_{\text{in}}(x), \quad x \in \mathbb{T}_L^d. \end{cases} \quad (\text{MKDV})$$

as an example of three wave system.

Here ψ is a real value function. We consider the periodic boundary condition, which implies that the spatial domain is a torus $\mathbb{T}_L^d = [0, L]^d$.

We know that the Fourier coefficients of ψ lie on the lattice $\mathbb{Z}_L^d = \{k = \frac{K}{L} : K \in \mathbb{Z}^d\}$. Let n_{in} be a known function, we assume that

$$\psi_{\text{in}}(x) = \frac{1}{L^d} \sum_{k \in \mathbb{Z}_L^d} \sqrt{n_{\text{in}}(k)} \eta_k(\omega) e^{2\pi i k x} \quad (2.1)$$

where $\eta_k(\omega)$ are mean-zero and identically distributed complex Gaussian random variables satisfying $\mathbb{E}|\eta_k|^2 = 1$. To ensure ψ_{in} to be a real value function, we assume that $n_{\text{in}}(k) = n_{\text{in}}(-k)$ and $\eta_k = \overline{\eta_{-k}}$. Finally, we assume that η_k is independent of $\{\eta_{k'}\}_{k' \neq k, -k}$.

The energy spectrum $n(t, k)$ mentioned in previous section is defined to be $\mathbb{E}|\widehat{\psi}(t, k)|^2$, where $\psi(t, k)$ are Fourier coefficients of the solution. Although the initial data is assumed to be the Gaussian random field, it is possible to develop a theory for other types of random initial data.

We define $\Lambda(k) := k_1(k_1^2 + \cdots k_d^2)$. Under this new notations the ZK equation becomes,

$$\partial_t \psi(t, x) = i\Lambda(\nabla)\psi(t, x) + \nu \Delta \psi(t, x) + \lambda \partial_{x_1}(\psi^2(t, x)).$$

In the wave turbulence, the energy distribution $n(k)$ is supposed to evolve according to the following wave kinetic equation

$$\begin{aligned} \partial_t n(t, k) &= \frac{1}{T_{\text{kin}}} \mathcal{K}(n(t, \cdot)), \\ \mathcal{K}(n)(k) &:= |k_x|^2 \int_{\substack{(k_1, k_2) \in \mathbb{R}^{2d} \\ k_1 + k_2 = k}} n(k_1) n(k_2) \delta(|k_1|^2 k_{1x} + |k_2|^2 k_{2x} - |k|^2 k_x) dk_1 dk_2 \quad (\text{WKE}) \\ &\quad - 2n(k) \int_{\mathbb{R}^d} k_x(k_x - k_{1x}) n(k_1) \delta(|k_1|^2 k_{1x} + |k|^2 k_x - |k_2|^2 k_{2x}) dk_1 \end{aligned}$$

Now we introduce the main theorem of this chapter.

Theorem 2.1.1. *Let $d \geq 3$ and L be a large number. Suppose that $n_{\text{in}} \in C_0^\infty(\mathbb{R}^d)$ is compactly supported in a domain whose diameter is bounded by D . Assume that ψ is a solution of (MKDV) with randomized initial data ψ_{in} given by (2.1). Set $\alpha = \lambda L^{-\frac{d}{2}}$ to be the strength of the nonlinearity and $T_{\text{kin}} = \frac{1}{8\pi\alpha^2}$ to be the wave kinetic time. Fix a small constant $\varepsilon > 0$, set $T_{\text{max}} = L^{-\varepsilon} \alpha^{-2} = O(L^{-\varepsilon} T_{\text{kin}})$.*

If α satisfies

$$\alpha^{-1} \leq L^{\frac{1}{2}} \tag{2.2}$$

and for some small constant c , ν satisfies

$$\nu \geq c^2 T_{\max}^{-1} \quad (2.3)$$

then for all $L^\varepsilon \leq t \leq T_{\max}$, we have the following conclusions

1. If $\sup_k n_{\text{in}}(k) \leq C_{\text{in}}$, then $\mathbb{E}|\widehat{\psi}(t, k)|^2$ is bounded by $2C_{\text{in}}$ for $t \leq T_{\max}$ and for any M , we can construct an approximation series

$$\mathbb{E}|\widehat{\psi}(t, k)|^2 = n_{\text{in}}(k) + n^{(1)}(k) + n^{(2)}(k) + \cdots + n^{(N)}(k) + O(L^{-M}) \quad (2.4)$$

where each terms $n^{(i)}(k)$ can be exactly calculated.

2. Define $l_d = (\nu T_{\max})^{\frac{1}{2}} \geq c$ and $\theta = C_1 \varepsilon$ (C_1 just depends on dimension d). We assume that $D \leq C_2 l_d^{-1}$. Then in the inertial range $|k| \leq \epsilon_1 l_d^{-1}$ and dissipation range $|k| \geq C_2 l_d^{-1}$, we have the following estimate respectively

$$n^{(1)}(k) = \begin{cases} \frac{t}{T_{\text{kin}}} \mathcal{K}(n_{\text{in}})(k) + O_{\ell_k^\infty} \left(L^{-\theta} \frac{T_{\max}}{T_{\text{kin}}} \right) + \widetilde{O}_{\ell_k^\infty} \left(\epsilon_1 \text{Err}_D(k_x) \frac{T_{\max}}{T_{\text{kin}}} \right) & \text{if } |k| \leq \epsilon_1 l_d^{-1}, \\ 0, & \text{if } |k| \geq 2C_2 l_d^{-1} \end{cases} \quad (2.5)$$

and

$$n^{(j)}(k) = O_{\ell_k^\infty} \left(L^{-\theta} \frac{t}{T_{\text{kin}}} \right), \quad j > 1 \quad (2.6)$$

where \mathcal{K} is defined in (WKE), and $O_{\ell_k^\infty}(A)$ (resp. $\widetilde{O}_{\ell_k^\infty}(A)$) is a quantity that is bounded by A in ℓ_k^∞ by some universal constant (resp. constant just depending on d). The definition of universal constant can be found in section 2.1.3. The definition of Err_D is

$$\text{Err}_D(k_x) = \begin{cases} D^{d+1}, & \text{if } |k_x| \leq D, \\ D^{d-1}(|k_x|^2 + D|k_x|), & \text{if } |k_x| \geq D. \end{cases} \quad (2.7)$$

3. By 2, $n_{\text{in}}(k) + n^{(1)}(k)$, the first two terms of (2.4), approximately equals to $n_{\text{in}}(k) + \frac{t}{T_{\text{kin}}} \mathcal{K}(n_{\text{in}})(k)$. $n_{\text{in}}(k) + \frac{t}{T_{\text{kin}}} \mathcal{K}(n_{\text{in}})(k)$ is the first two terms of the approximation series of (WKE). Therefore, $\mathbb{E}|\widehat{\psi}(t, k)|^2$ is an approximate solution of the (WKE).

Remark 2.1.2. The condition $d \geq 3$ is essential. When $d = 1$, the ZK equation becomes the KdV equation which is an integrable system whose long time behavior is quasi-periodic instead of turbulent [29]. When $d = 2$, as mentioned in Remark 4.1.3, the desire number theory result is not true and a major revision to (WKE) is required to obtain a valid wave kinetic theory.

Remark 2.1.3. $\frac{t}{T_{\text{kin}}} \mathcal{K}(n_{\text{in}})(k) = O_{\ell_k^\infty} \left(\frac{t}{T_{\text{kin}}} \right) = O_{\ell_k^\infty} \left(\frac{T_{\text{max}}}{T_{\text{kin}}} \right)$ in (2.5) is the largest term in (2.5) and (2.6). Compared to the second term in (2.5) or the first term in (2.6), $\frac{t}{T_{\text{kin}}} \mathcal{K}(n_{\text{in}})(k)$ is larger than them by a factor L^θ . It is also larger than the third term in (2.5) by a factor of ϵ_1 .

Remark 2.1.4. The restriction on α^{-1} is not optimal. The optimal result is expected to be $\alpha^{-1} \leq L$ for general torus and $\alpha^{-1} \leq L^{d/2}$ for generic torus. Here the torus is general or generic in the sense of [10]. Except for those in the the counting results in Chapter 4, section 4.1, all the arguments in this chapter work under these stronger assumptions.

Remark 2.1.5. Due to ∂_x in the nonlinearity, there is a potential risk of loss of derivative. This causes a serious difficulty in controlling the high frequency part. To resolve this difficulty, some regularization to ZK equation is required. In [40] and this paper, grid discretization and viscosity are introduced respectively. Both of them serve as canonical high frequency truncations.

2.1.2 Ideas of the proof

The basic strategy of proving the main theorem is to construct an approximation series and use probability theory and number theory to control the size and error of this approximation.

The approximate solution

The equation of Fourier coefficients is

$$\dot{\psi}_k = i\Lambda(k)\psi_k - \nu|k|^2\psi_k + \frac{i\lambda}{L^d} \sum_{\substack{(k_1, k_2) \in (\mathbb{Z}_L^d)^2 \\ k_1 + k_2 = k}} k_{x_1} \psi_{k_1} \psi_{k_2} \quad (2.8)$$

Define a new dynamical variable $\phi = e^{-it\Lambda(\nabla)}\psi$ and integrate (2.8) in time. Then (MKDV) with initial data (2.1) becomes

$$\phi_k = \xi_k + \frac{i\lambda}{L^d} \sum_{k_1 + k_2 = k} \int_0^t k_{x_1} \phi_{k_1} \phi_{k_2} e^{is\Omega(k_1, k_2, k) - \nu(t-s)|k|^2} ds. \quad (2.9)$$

Here $\Omega(k_1, k_2, k) = \Lambda(k_1) + \Lambda(k_2) - \Lambda(k)$ and ξ_k are the Fourier coefficients of the initial data of ψ defined by $\xi_k = \sqrt{n_{\text{in}}(k)} \eta_k(\omega)$.

Denote the second term of right hand side by $\mathcal{T}(\psi, \psi)_k$ and the right hand side by $\mathcal{F}(\psi)_k = \xi_k + \mathcal{T}(\psi, \psi)_k$. Then the equation becomes $\psi = \mathcal{F}(\psi)_k$. We can construct the approximation by iteration: $\psi = \mathcal{F}(\psi) = \mathcal{F}(\mathcal{F}(\psi)) = \mathcal{F}(\mathcal{F}(\mathcal{F}(\psi))) = \dots$.

Define the approximate solution by $\psi_{app} = \mathcal{F}^N(\xi)$. By recursively expanding \mathcal{F}^N , we know that ψ_{app} is a polynomial of ξ . The expansion can be described as the following,

$$\begin{aligned} \psi_{app} &= \mathcal{F}^N(\xi) = \xi + \mathcal{T}(\mathcal{F}^{N-1}(\xi), \mathcal{F}^{N-1}(\xi)) \\ &= \xi + \mathcal{T}\left(\xi + \mathcal{T}(\mathcal{F}^{N-2}(\xi), \mathcal{F}^{N-2}(\xi)), \dots\right) = \xi + \mathcal{T}(\xi, \xi) + \dots \\ &= \xi + \mathcal{T}(\xi, \xi) + \mathcal{T}(\mathcal{T}(\xi, \xi), \xi) + \mathcal{T}(\xi, \mathcal{T}(\xi, \xi)) + \dots \end{aligned}$$

In the above iteration, we recursively replace $\mathcal{F}^l(\xi)$ by $\xi + \mathcal{T}(\mathcal{F}^{l-1}(\xi), \mathcal{F}^{l-1}(\xi))$.

We need a good upper bound for each terms of ψ_{app} . To get this we introduce tree diagrams to represent terms $\xi, \mathcal{T}(\xi, \xi), \mathcal{T}(\mathcal{T}(\xi, \xi), \xi), \dots$. The basic notation of tree diagrams will be introduced in section 2.2.1.

The perturbative analysis

To prove the main theorem, we need to bound the approximation error of ψ_{app} defined by $w = \psi - \psi_{app}$. To do this, we use the follow equation of w which can be derived from (2.9):

$$w = Err(\xi) + Lw + B(w, w) \quad (2.10)$$

Here $Err(\xi)$ is a polynomial of ξ whose degree $\leq N + 1$ monomials vanish. Lw , $B(w, w)$ are linear, quadratic in w respectively.

We prove the smallness of w using the bootstrap method.

Define $\|w\|_{X^p} = \sup_k \langle k \rangle^p |w_k|$. Starting from the assumption that $\sup_t \|w\|_{X^p} \leq CL^{-M}$ ($C, M \gg 1$), we need to prove that $\sup_t \|w\|_{X^p} \leq (1 + C/2)L^{-M} < CL^{-M}$. To prove $\|w\|_{X^p} \leq (1 + C/2)L^{-M}$, we use (2.10), which gives

$$\|w\|_{X^p} \leq \|Err(\xi)\|_{X^p} + \|Lw\|_{X^p} + \|B(w, w)\|_{X^p} \quad (2.11)$$

We just need to show that

$$\|Err(\xi)\|_{X^p} \leq L^{-M}, \quad \|B(w, w)\|_{X^p} \leq C^2 L^{d+O(1)-2M}. \quad (2.12)$$

Combining with a special treatment of Lw , the above estimates imply that $\|w\|_{X^p} \leq (1 + C/2)L^{-M}$ which closes the bootstrap.

Couple diagrams, lattice points counting and $\|Err(\xi)\|_{X^p}$

In this section we explain the idea of proving upper bound of $\|Err(\xi)\|_{X^p}$.

$(Err(\xi))_k$ is a sum of terms of the form

$$\begin{aligned}\mathcal{J}_k^0(\xi) &= \xi_k, \quad \mathcal{J}_k^1(\xi) = \frac{i\lambda}{L^d} \sum_{k_1+k_2=k} H_{k_1 k_2}^1 \xi_{k_1} \xi_{k_2}, \quad \dots \\ \mathcal{J}_{T,k}^l(\xi) &= \left(\frac{i\lambda}{L^d}\right)^l \sum_{k_1+k_2+\dots+k_{l+1}=k} H_{k_1 \dots k_{l+1}}^l(T) \xi_{k_1} \xi_{k_2} \dots \xi_{k_{l+1}}, \quad \dots\end{aligned}\tag{2.13}$$

According to section 2.2.1, each terms correspond to a tree diagram and their coefficients can be calculated from these diagrams. This calculation is done in section 2.4.1. As a corollary of tree diagram representation, we know that H^l is large near a surface given by $2l$ equations $S = \{S_{n_1}(T) = 0, \Omega_{n_1}(T) = 0, \dots, S_{n_l}(T)(T) = 0, \Omega_{n_l}(T) = 0\}$.

By the large deviation principle, to obtain upper bounds of Gaussian polynomials $\mathcal{J}_{T,k}^l(\xi)$, it suffices to calculate their variance. This calculation is done in section 2.4.3 using the Wick theorem and we introduce the concept of couple diagrams to represent the final result.

As a corollary of couple diagram representation, we know that the coefficients of the variance concentrate near a surface given by n equations (n is the number of nodes in the couple) $S = \{S_{n_1}(T) = 0, \Omega_{n_1}(T) = 0, \dots, S_{n_n}(T)(T) = 0, \Omega_{n_n}(T) = 0\}$. Then in order to estimate the variance it suffices to upper bound the number of lattice points near this surface. This is done in section 2.4.4 using the edge cutting argument to reduce the size of the couple.

The method in [10] of getting number theory estimate based on tree diagram does not work in our setting. This is because the energy conservation equation $\Lambda(k_1) + \Lambda(k_2) - \Lambda(k) = 0$ of ZK equation degenerates seriously when k_x is close to 0. In (4.1), the number of solutions of the diophantine equation $\Lambda(k_1) + \Lambda(k_2) = \Lambda(k) + \sigma + O(T^{-1})$ can only be bounded by $|k_x|^{-1}$ which goes to infinity when $k_x \rightarrow 0$. This difficulty

is resolved by the fact that the multiplier k_x in the last term of (2.8) vanishes when $k_x \rightarrow 0$. Since multipliers become very complicated in higher order tree terms, we introduce the concept of norm edges to keep track of them.

In conclusion, combining the above arguments, we can show that, for any M , we can take N large enough so that $\|Err(\xi)\|_{X^p} \leq L^{-M}$.

Upper bounds for $\|B(w, w)\|_{X^p}$

$\|B(w, w)\|_{X^p}$ is a sum of terms of the form

$$\frac{i\lambda}{L^d} \int_0^t \sum_{k_1+k_2-k=0} B_{k_1 k_2}(s) w_{k_1}(\xi) w_{k_2} \quad (2.14)$$

The upper bound of $\|B(w, w)\|_{X^p}$ can be obtained by a straight forward estimate

$$\|B(w, w)\|_{X^p} \leq L^{O(1)} \|w\|_{X^p} \leq C^2 L^{O(1)-2M}, \quad (2.15)$$

Therefore, we get the desire upper bounds $\|B(w, w)\|_{X^p} \ll L^{-M}$ by taking $M > O(1)$.

A random matrix bound and Lw

To obtain a good upper bound for Lw , we need to estimate the norm of the random matrix L , following the idea in [10], [9].

In [10], they consider solutions in the Bourgain space $X^{s,b}$ and use TT^* method to get the upper bound for the operator norm of L , $\|L\|_{X^{s,b} \rightarrow X^{s,b}} \ll 1$. But we prefer to work in the simpler functional space X^p which is not a Hilbert space. Although the standard TT^* method is not useful in a non-Hilbert space, we can bypass it using a Neumann series argument.

Let us first explain how TT^* method works. Here we pretend that $\|\cdot\|_{X^p}$ is a Hilbert norm. The key idea of TT^* method is the inequality $\|L\|_{X^p \rightarrow X^p} =$

$\|(LL^*)^K\|_{X^p \rightarrow X^p}^{\frac{1}{K}} \leq (L^d \sup_{k,l} ((LL^*)^K)_{k,l})^{1/K}$. To upper bound $\|L\|_{X^p \rightarrow X^p}$, we just need to estimate $(LL^*)^K_{k,l}$ which can be calculated by couple diagrams and be estimated by the large deviation inequality. By taking K large, the loss $L^{d/K}$ could be made arbitrarily small.

Unfortunately, $\|\cdot\|_{X^p}$ is not a Hilbert norm. However, we can bypass the TT^* method using a Neumann series argument. Note that from (2.9) we have the identity

$$w - Lw = Err(\xi) + B(w, w). \quad (2.16)$$

We have good upper bounds for all of the three terms on the right hand side. By Neumann series argument we have

$$w = (1 - L)^{-1}(RHS) = (1 - L^K)^{-1}(1 + L + \cdots + L^{K-1})(RHS). \quad (2.17)$$

By calculating $(L^K)_{k,l}$ we can show that $\|L^K\|_{X^p \rightarrow X^p} \ll 1$. This implies that $\|(1 - L^K)^{-1}\|_{X^p \rightarrow X^p} \lesssim 1$. Therefore, a good upper bound of RHS gives us a good upper bound of $(1 + L + \cdots + L^{K-1})(RHS)$. Combining the above arguments, we obtain the desire estimate of w . This is done in section 2.2.3 and 2.4.6.

One additional difficulty is the unboundedness of L due to the derivative ∂_x in the nonlinearity. This is controlled by the high frequency decay coming from viscosity.

Proof of the main theorem

In summary, the above arguments in this section prove that when $t \leq \alpha^{-2}$, we have $\|w\|_{X^p} \leq L^{-M}$ with high probability ($P(false) \lesssim e^{-CL^\theta}$).

The above inequality is equivalent to $\sup_k |\langle k \rangle^s w_k| \leq CL^{-M}$. Remember that $w := \psi - \psi_{app}$, so with high probability we have the following estimate $\sup_k |\langle k \rangle^s |\psi_k - \psi_{app,k}| \leq CL^{-M}$. This implies that $\mathbb{E}|\widehat{\psi}(t, k)|^2 = \mathbb{E}|\psi_{app,k}|^2 + O(L^{-M})$. This suggests that we may get the approximation of $\mathbb{E}|\widehat{\psi}(t, k)|^2$ by calculating $\mathbb{E}|\psi_{app,k}|^2$. $\mathbb{E}|\psi_{app,k}|^2$

can be exactly calculated and the theorem can be proved by extract the main term in $\mathbb{E}|\psi_{app,k}|^2$. This is done in section 2.2.4.

2.1.3 Notations

Universal constants: In this chapter, universal constants are constants that just depend on dimension d , diameter D of the support of n_{in} and the length of the inertial range l_d^{-1} .

$O(\cdot)$, \ll , \lesssim , \sim : Throughout this chapter, we frequently use the notation, $O(\cdot)$, \ll , \lesssim . $A = O(B)$ or $A \lesssim B$ means that there exists C such that $A \lesssim CB$. $A \ll B$ means that there exists a small constant c such that $A \lesssim cB$. $A \sim B$ means that there exist two constant c, C such that $cB \lesssim A \lesssim CB$. Here the meaning of constant depends on the context. If they appear in conditions involving k, Λ, Ω , etc., like $|k| \lesssim 1$, $\iota_{\epsilon_1} k_{\epsilon_1} + \iota_{\epsilon_2} k_{\epsilon_2} + \iota_{\epsilon} k_{\epsilon} = 0$, then they are universal constants. If these constants appear in an estimate which gives upper bound of some quantity, like $\|L^K\|_{X^p \rightarrow X^p} \ll 1$ or $\sup_t \sup_k |(\mathcal{J}_T)_k| \lesssim L^{O(l(T)\theta)} \rho^{l(T)}$, then in addition to the quantities that universal constants depend, they can also depend on the quantities $\theta, \varepsilon, K, M, N, \epsilon_1$.

Order of constants: Here is the order of all constants which can appear in the exponential or superscript of L . These constants are $\theta, \varepsilon, K, M, N, \epsilon_1$.

All the constants are small compared to L in the sense they are less than L^θ for arbitrarily small $\theta > 0$.

ε can be an arbitrarily small constant less than 0.5, the reader is encouraged to assume it to be 0.01. The order of other constants can be decided by the relations $\theta \ll \varepsilon, K = O(\theta^{-1}), M \gg K, N \geq M/\theta$, here the constants in $\ll, O(\cdot)$ are universal.

$$\underline{\mathbb{Z}_L^d}: \mathbb{Z}_L^d = \{k = \frac{K}{L} : K \in \mathbb{Z}^d\}$$

k_x, k_\perp : Given any vector k , let k_x be its first component and k_\perp be the vector formed by the rest components.

$$\underline{\Lambda(k), \Lambda(\nabla)}: \Lambda(k) := k_1(k_1^2 + \cdots k_d^2) \text{ and } \Lambda(\nabla) = i|\nabla|^2 \partial_{x_1}$$

Fourier series: The spatial Fourier series of a function $u : \mathbb{T}_L^d \rightarrow \mathbb{C}$ is defined on $\mathbb{Z}_L^d := L^{-1}\mathbb{Z}^d$ by

$$u_k = \int_{\mathbb{T}_L^d} u(x) e^{-2\pi i k \cdot x}, \quad \text{so that} \quad u(x) = \frac{1}{L^d} \sum_{k \in \mathbb{Z}_L^d} u_k e^{2\pi i k \cdot x}. \quad (2.18)$$

Given any function F , let F_k or $(F)_k$ be its Fourier coefficients.

Order of L : In this chapter, L is assumed to be a constant which is much larger than all the universal constants and $\theta, \varepsilon, K, M, N, \epsilon_1$.

L -certainty: If some statement S involving ω is true with probability $\geq 1 - O_\theta(e^{-L^\theta})$, then we say this statement S is L -certain.

2.2 The Perturbation Expansion

In this section, we calculate the approximation series and introduce Feynman diagrams to represent terms in this series. Then we bound the error of this approximation by the bootstrap method, assuming several propositions about the upper bounds of higher order terms. We will prove these propositions in the rest part of the chapter.

2.2.1 The approximation series and Feynman diagrams

In this section, we derive the equation for Fourier coefficients and construct the approximate solution.

The Equation of Fourier coefficients

Let ψ_k be the Fourier coefficient of ψ . Then in term of ψ_k equation (MKDV) becomes

$$\begin{cases} \dot{\psi}_k = i\Lambda(k)\psi_k - \nu|k|^2\psi_k + \frac{i\lambda}{L^d} \sum_{\substack{(k_1, k_2) \in (\mathbb{Z}_L^d)^2 \\ k_1 + k_2 = k}} k_x \psi_{k_1} \psi_{k_2} \\ \psi_k(0) = \xi_k = \sqrt{n_{\text{in}}(k)} \eta_k(\omega) \end{cases} \quad (2.19)$$

Define the linear profile by

$$\phi_k(t) := e^{-i\Lambda(k)t} \psi_k(t) \quad (2.20)$$

Rewriting (2.19) in terms of ϕ_k gives

$$\dot{\phi}_k = -\nu|k|^2\phi_k + \frac{i\lambda}{L^d} \sum_{S(k_1, k_2, k)=0} k_x \phi_{k_1} \phi_{k_2} e^{it\Omega(k_1, k_2, k)} \quad (2.21)$$

where

$$\begin{aligned} S(k_1, k_2, k) &= k_1 + k_2 - k, \\ \Omega(k_1, k_2, k) &= \Lambda(k_1) + \Lambda(k_2) - \Lambda(k). \end{aligned} \quad (2.22)$$

We will work with (2.21) in the rest part of this chapter.

Integrating (2.21) gives

$$\phi_k = \xi_k + \underbrace{\frac{i\lambda}{L^d} \sum_{S(k_1, k_2, k)=0} \int_0^t k_x \phi_{k_1} \phi_{k_2} e^{is\Omega(k_1, k_2, k) - \nu|k|^2(t-s)} ds}_{\mathcal{T}(\phi, \phi)_k}. \quad (2.23)$$

Denote the second term on the right hand side by $\mathcal{T}(\phi, \phi)_k$. Denote the right hand

side by $\mathcal{F}(\phi)_k = \xi_k + \mathcal{T}(\phi, \phi)_k$. With these notations, (2.23) becomes $\phi = \mathcal{F}(\phi)_k$.

We construct the approximation series by iteration: $\phi = \mathcal{F}(\phi) = \mathcal{F}(\mathcal{F}(\phi)) = \mathcal{F}(\mathcal{F}(\mathcal{F}(\phi))) = \dots$. To estimate this approximation series, we need a compact graphical notation to represent the huge amount of terms generated from iteration. This is done by introducing the concept of Feynman diagrams.

Some basic definitions from graph theory

In this section, we introduce the concept of binary trees, branching nodes, leaves, subtrees, node decoration and expanding leaves.

Definition 2.2.1. 1. **Binary trees:** A binary tree T is a tree in which each node has 2 or 0 children. An example of binary trees used in this chapter is shown in Figure 2.1.

2. **Branching nodes:** A branching node in a binary tree is a node which has 2 children. The number of all branching nodes in a tree T is denoted by $l(T)$. In Figure 2.1, $l(T) = 2$.

3. **Leaves:** A leaf of a tree T is a node which has no child. In Figure 2.1, all \star nodes and \square nodes are leaves.

4. **Subtrees:** If any child of any node in a subset T' of a tree T is also contained in T' then T' also forms a tree, we call T' a subtree of T . If the root node of T' is $\mathbf{n} \in T$, we say T' is the subtree rooted at \mathbf{n} or subtree of \mathbf{n} and denote it by $T_{\mathbf{n}}$. In Figure 2.1, the tree inside the box is the subtree rooted at node \bullet .

5. **Node decoration:** In Figure 2.1, each node is associated with a symbol in $\{\bullet, \star, \square\}$. If a node \mathbf{n} has symbol \bullet (similarly \star, \square), we say \mathbf{n} is decorated by \bullet (\star, \square) or \mathbf{n} has decoration \bullet (\star, \square). In what follows, we adopt the convention that leaves always have decoration \star or \square and nodes other than leaves always have decoration \bullet .

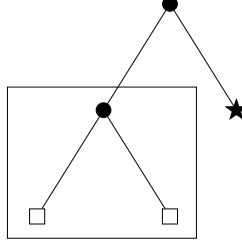


Figure 2.1: Subtrees and node decoration.

6. **Expanding and final leaves:** Leaves denoted by \square are called expanding leaves.

Other leaves denoted by \star are called final leaves. The notion of expanding leaves is useful in the construction of trees, in which the presence of \square means that the construction is not finishing and \square denotes leaves that may be replaced by a branching node later.

The concept of expanding leaves and \square is only used in this section, so the readers can safely forget it after that section.

Connection between iteration and trees

In this section, we explain non-rigorously the connection between perturbation expansion and trees. Rigorous argument can be find in the next section.

This iteration process can be described as the following,

$$\begin{aligned}
 \phi &= \mathcal{F}(\phi) = \xi + \mathcal{T}(\phi, \phi) \\
 &= \xi + \mathcal{T}\left(\xi + \mathcal{T}(\phi, \phi), \dots\right) = \xi + \mathcal{T}(\xi, \xi) + \mathcal{T}\left(\mathcal{T}(\phi, \phi), \xi\right) \dots \\
 &= \xi + \mathcal{T}(\xi, \xi) + \mathcal{T}(\mathcal{T}(\xi, \xi), \xi) + \mathcal{T}(\xi, \mathcal{T}(\xi, \xi)) + \dots
 \end{aligned}$$

In the above iteration, we recursively choose one ϕ , replace it by $\xi + \mathcal{T}(\phi, \phi)$ and use

the linearity of \mathcal{T} to expand into two terms.

$$\begin{aligned} & \mathcal{T}\left(\cdots, \mathcal{T}(\mathcal{T}(\xi, \underline{\phi}), \cdots)\right) \rightarrow \mathcal{T}\left(\cdots, \mathcal{T}(\mathcal{T}(\xi, \underline{\xi + \mathcal{T}(\phi, \phi)}), \cdots)\right) \\ &= \underbrace{\mathcal{T}\left(\cdots, \mathcal{T}(\mathcal{T}(\xi, \underline{\xi}), \cdots)\right)}_I + \underbrace{\mathcal{T}\left(\cdots, \mathcal{T}(\mathcal{T}(\xi, \underline{\mathcal{T}(\phi, \phi)}), \cdots)\right)}_{II} \end{aligned} \quad (2.24)$$

Here I and II are obtained by replacing ϕ by ξ and $\mathcal{T}(\phi, \phi)$ respectively.

In summary, all terms in the expansion can be generated by following steps

- **Step 0.** Add a term ϕ in the summation \mathcal{J} .
- **Step i ($i \geq 1$).** Assume that **Step $i - 1$** has been finished which produces a sum of terms \mathcal{J} , then choose a term in \mathcal{J} which has least number of ξ and ϕ , remove this term from \mathcal{J} and add the two terms in \mathcal{J} constructed in (2.24).

This process is very similar to the construction of binary trees, in which we recursively replace a chosen expanding node by a leaf or branching node.

- **Step 0.** Start from a expanding root node \square .
- **Step i ($i \geq 1$).** Assume that we have finish the **Step $i - 1$** which produces a collection of trees \mathcal{T} , then choose a tree in \mathcal{T} which has least number of expanding leaves \square and final leaves \star , remove this tree from \mathcal{T} and add two new trees in \mathcal{T} . In these two new trees, we replace a expanding leaf \square by a final leaf \star or a branching node \bullet with two expanding children leaves \square . This construction is illustrated by Figure 2.2.

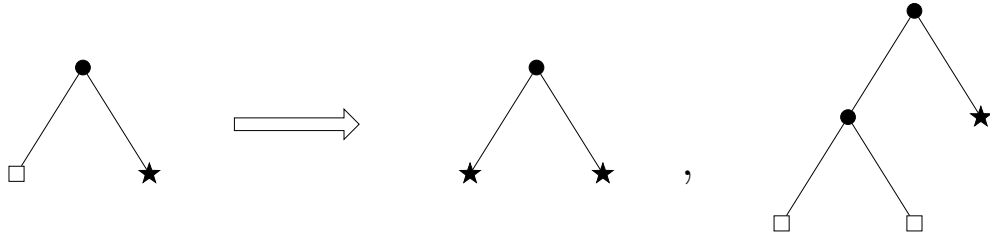


Figure 2.2: One step in the construction of binary trees

By comparing the above two process, we can make the connection between terms and trees more explicit. Each node \bullet other than leaf in the tree T corresponds to a $\mathcal{T}(\cdots, \cdots)$ in a term \mathcal{J}_T . Each final leaf \star and expanding leaf \square corresponds to ξ and ϕ respectively. The **Step** i of replacing ϕ by ξ or $\mathcal{T}(\phi, \phi)$ corresponds to replacing \square by \star or a branching node with two children \square .

We have following recursive formula for calculating a term \mathcal{J}_T from a binary tree T .

If T has only one node then $\mathcal{J}_T = \xi$. Otherwise let \bullet_1, \bullet_2 be two children of the root node \bullet , let $T_{\bullet_1}, T_{\bullet_2}$ be the subtrees of T rooted at the above nodes. If $\mathcal{J}_{T_{\bullet_1}}, \mathcal{J}_{T_{\bullet_2}}$ have been recursively calculated, then \mathcal{J}_T can be calculated by

$$\mathcal{J}_T = \mathcal{T}(\mathcal{J}_{T_{\bullet_1}}, \mathcal{J}_{T_{\bullet_2}}). \quad (2.25)$$

The formal power series obtained by iterating $\phi = \mathcal{F}(\phi)$ can be calculated from trees by $\sum_{T \in \mathcal{T}} \mathcal{J}_T$.

Let $l(T)$ be the number of branching nodes in T , then it can be shown that \mathcal{J}_T is a degree $l(T) + 1$ polynomial of ξ . We define the approximation series to be a finite degree truncation of the formal power series which equals to $\sum_{l(T) \leq N} \mathcal{J}_T$.

Feynman diagrams and construction of the approximation solution

In this section we present the rigorous argument equivalent to that in the above section.

In the construction of trees, finally all \square nodes will be replaced by \bullet, \star , so in what follows we only consider trees whose nodes are decorated by \bullet, \star .

Definition 2.2.2. Given a binary tree T whose nodes are decorated by \bullet, \star , we

inductively define the quantity \mathcal{J}_T by:

$$\mathcal{J}_T = \begin{cases} \xi, & \text{if } T \text{ has only one node } \star. \\ \mathcal{T}(\mathcal{J}_{T_{\mathbf{n}_1}}, \mathcal{J}_{T_{\mathbf{n}_2}}), & \text{otherwise.} \end{cases} \quad (2.26)$$

Here $\mathbf{n}_1, \mathbf{n}_2$ are two children of the root node \mathbf{r} and $T_{\mathbf{n}_1}, T_{\mathbf{n}_2}$ are the subtrees of T rooted at the these nodes.

Definition 2.2.3. Given a large number N , define the approximate solution ϕ_{app} by

$$\phi_{app} = \sum_{l(T) \leq N} \mathcal{J}_T \quad (2.27)$$

Section 2.2.1 explains why the approximation series should equals to (2.27), a sum of many tree terms, but if we know this fact, we can directly prove it, and forget all the motivations. The lemma below proves that ϕ_{app} defined by the above expression is an approximate solution.

Lemma 2.2.4. *Define*

$$Err = \mathcal{F}(\phi_{app}) - \phi_{app}, \quad (2.28)$$

then we have

$$Err = \sum_{T \in \mathcal{T}_{>N}^*} \mathcal{J}_T, \quad (2.29)$$

where $\mathcal{T}_{>N}^$ is defined by*

$$\begin{aligned} \mathcal{T}_{>N}^* = \{T : l(T) > N, l(T_{\mathbf{n}_1}) \leq N, l(T_{\mathbf{n}_2}) \leq N, \\ T_{\mathbf{n}_1}, T_{\mathbf{n}_2} \text{ are the subtrees defined in Definition 2.2.2}\} \end{aligned} \quad (2.30)$$

Remark 2.2.5. Notice that all terms in $\sum_{T \in \mathcal{T}_{>N}^*}$ are polynomials of ξ of degree $> N$. Therefore, the approximation error of ϕ_{app} is of very high order, which proves that ϕ_{app} is an appropriate approximation solution.

Proof. By (2.27), we get

$$\begin{aligned}
Err &= \mathcal{F}(\phi_{app}) - \phi_{app} \\
&= \xi + \mathcal{T}(\phi_{app}, \phi_{app}) - \phi_{app} \\
&= \xi + \sum_{l(T_1), l(T_2) \leq N} \mathcal{T}(\mathcal{J}_{T_1}, \mathcal{J}_{T_2}) - \sum_{l(T) \leq N} \mathcal{J}_T
\end{aligned} \tag{2.31}$$

Let T be a tree constructed by connecting the root nodes $\mathbf{n}_1, \mathbf{n}_2$ of T_1, T_2 to a new node \mathbf{r} . We define \mathbf{r} to be the root node of T .

Then by (2.26), we have

$$\mathcal{J}_T = \mathcal{T}(\mathcal{J}_{T_1}, \mathcal{J}_{T_2}) \tag{2.32}$$

and

$$\sum_{l(T_1), l(T_2) \leq N} \mathcal{T}(\mathcal{J}_{T_1}, \mathcal{J}_{T_2}) = \sum_{\substack{l(T) \geq 1 \\ l(T_1), l(T_2) \leq N}} \mathcal{J}_T \tag{2.33}$$

By (2.31), we get

$$\begin{aligned}
Err &= \xi + \sum_{\substack{l(T) \geq 1 \\ l(T_1), l(T_2) \leq N}} \mathcal{J}_T - \sum_{l(T) \leq N} \mathcal{J}_T \\
&= \sum_{\substack{T_1, T_2 \text{ are subtrees of } \mathbf{r} \\ l(T_1), l(T_2) \leq N}} \mathcal{J}_T - \sum_{\substack{T_1, T_2 \text{ are subtrees of } \mathbf{r} \\ l(T) \leq N \\ l(T_1), l(T_2) \leq N}} \mathcal{J}_T. \\
&= \sum_{T \in \mathcal{T}_{>N}^*} \mathcal{J}_T
\end{aligned} \tag{2.34}$$

Here in the second equality, we use the fact that $\sum_{l(T) \leq N} = \sum_{\substack{l(T) \leq N \\ l(T_1), l(T_2) \leq N}}$.

Therefore, we complete the proof of this lemma. \square

2.2.2 Estimates of the approximation solution

Estimates of tree terms

By (2.29), in order to control the approximation error Err , it suffices to get upper bounds of tree terms \mathcal{J}_T . We state the upper bound in the proposition below and delay its proof to section 2.4.5.

Let us introduce a definition before state the proposition.

Definition 2.2.6. Given a property A , we say A happens L -certainly if the probability that A happens satisfies $P(A) \geq 1 - Ce^{-L^\theta}$ for some $C, \theta > 0$.

Proposition 2.2.7. *We have L -certainly for any θ that*

$$\sup_t \sup_k |(\mathcal{J}_T)_k| \lesssim L^{O(l(T)\theta)} \rho^{l(T)}. \quad (2.35)$$

and $(\mathcal{J}_T)_k = 0$ if $|k| \gtrsim 1$. Here $(\mathcal{J}_T)_k$ is the Fourier coefficients of \mathcal{J}_T and

$$\rho = \alpha T_{max}^{\frac{1}{2}}. \quad (2.36)$$

Linearization around the approximation solution

Let $w = \phi - \phi_{app}$ be the deviation of ϕ to the true solution ϕ_{app} . In order to estimate w , we consider the linearized equation of

$$\phi = \mathcal{F}(\phi) = \xi + \mathcal{T}(\phi, \phi). \quad (2.37)$$

The linearized equation is a equation of w given by

$$w = Err(\xi) + Lw + B(w, w), \quad (2.38)$$

where $Err(\xi)$, Lw , $B(w, w)$ are given by

$$\begin{cases} Err = \mathcal{F}(\phi_{app}) - \phi_{app}, \\ Lw = 2\mathcal{T}(\phi_{app}, w), \\ B(w, w) = \mathcal{T}(w, w). \end{cases} \quad (2.39)$$

As explained in section 2.1.2, to control w , we need operator norm bound $\|L^K\|_{X^p}$. Notice that $\phi_{app} = \sum_{l(T) \leq N} \mathcal{J}_T$, so we know that Lw is a sum of terms like $\mathcal{T}(\mathcal{J}_T, w)$.

It suffices to get following operator norm bound for $w \rightarrow \mathcal{T}(\mathcal{J}_T, w)$ in order to upper bound L^K .

Proposition 2.2.8. *Define $\rho = \alpha T_{max}^{\frac{1}{2}}$ as in Proposition 2.2.7 and \mathcal{P}_T to be the linear operator*

$$w \rightarrow \mathcal{T}(\mathcal{J}_T, w). \quad (2.40)$$

Then for any sequence of trees $\{T_1, \dots, T_K\}$, we have L -certainly for any θ the operator bound

$$\left\| \prod_{j=1}^K \mathcal{P}_{T_j} \right\|_{L_t^\infty X^p \rightarrow L_t^\infty X^p} \leq L^{O(1+\theta \sum_{j=1}^K l(T_j))} \rho^{\sum_{j=1}^K l(T_j)}. \quad (2.41)$$

for any T_j with $l(T_j) \leq N$.

The above inequality implies that

$$\|L^K\|_{L_t^\infty X^p \rightarrow L_t^\infty X^p} \leq L^{O(1+K\theta)} \rho^K. \quad (2.42)$$

The proof of this proposition can be found in section 2.4.6.

2.2.3 Bound the error of the approximation

Define $w = v - v_{app}$ to be the approximation error. In this section we prove the following theorem that gives an upper bound of w , assuming Proposition 2.2.7 and Proposition 2.2.8 in previous section.

Theorem 2.2.9. *Let $w = \phi - \phi_{app}$. Given any $M \gg 1$, there exists N such that if ϕ_{app} is the N -th order approximate solution, then $\sup_{t \leq T_{max}} \|w(t)\|_{X^p} \lesssim L^{-M}$ L -certainly (with probability $\geq 1 - Ce^{-CL^\theta}$).*

Proof. If for some C sufficiently large we can show that

$$\sup_{t \leq T} \|w(t)\|_{X^p} \leq CL^{-M} \text{ implies that } \sup_{t \leq T} \|w(t)\|_{X^p} < CL^{-M}, \quad (2.43)$$

for all $T \leq T_{max}$, then we finish the proof of this theorem.

Here is the explanation. (2.43) implies that if we define the set $A = \{T : \sup_{t \leq T} \|w\|_{X^p} \leq CL^{-M}\}$ then the set equals to $\{T : \sup_{t \leq T} \|w\|_{H^s} < CL^{-M}\}$ which is open. The original definition $A = \{T : \sup_{t \leq T} \|w\|_{X^p} \leq CL^{-M}\}$ implies that this set is also closed. It is nonempty because $\|w(0)\|_{X^p} = 0$ implies that $0 \in A$. Therefore, A is open, closed and nonempty in $[0, T_{max}]$, so $A = [0, T_{max}]$ which implies that the theorem.

Now we prove (2.43). By (2.38),

$$w - Lw = Err(\xi) + B(w, w). \quad (2.44)$$

By Neumann series we have

$$\begin{aligned} w &= (1 - L)^{-1}(Err(\xi) + B(w, w)) \\ &= (1 - L^K)^{-1}(1 + L + \dots + L^{K-1})(Err(\xi) + B(w, w)). \end{aligned} \quad (2.45)$$

Assume that the constant in $O(1 + K\theta)$ in (2.42) is C_{norm} . Since $T_{max} \leq L^{-\varepsilon}\alpha^{-2}$,

we know that $\rho = \alpha T_{\max}^{\frac{1}{2}} \lesssim L^{-\varepsilon}$. Take $\theta \leq C_{\text{norm}}\varepsilon/2$ and $K \gg \frac{C_{\text{norm}}}{\varepsilon}$, then $\|L^K\|_{L^\infty X^p \rightarrow L^\infty X^p} \leq L^{C_{\text{norm}}(1+K\theta)} \rho^K \lesssim L^{C_{\text{norm}}(1+K\theta)} L^{-K\varepsilon} \ll 1$, so we get $\|L^K\|_{L^\infty X^p \rightarrow L^\infty X^p} \ll 1$ and thus $\|(1 - L^K)^{-1}\|_{L^\infty X^p \rightarrow L^\infty X^p} \lesssim 1$.

By (2.45), we get

$$\|w(t)\|_{X^p} \lesssim \sum_{j=1}^K \|L^j(Err(\xi))\|_{X^p} + \|(1 + L + \cdots + L^{K-1})(B(w, w))\|_{X^p} \quad (2.46)$$

By (2.29), we know that Err is a sum of tree terms $\sum_{T \in \mathcal{T}_{>N}^*} \mathcal{J}_T$ of order $\geq N$. Since $L = \sum_{1 \leq l(T) \leq N} \mathcal{P}_T$, we know that $L^j(Err(\xi))$ is a sum of terms like $\mathcal{P}_{T_1} \circ \cdots \circ \mathcal{P}_{T_j}(\mathcal{J}_T)$ which by (2.172) equals to $\mathcal{J}_{T_1 \circ \cdots \circ T_j \circ T}$. By Proposition 2.2.7, we get $\|\mathcal{J}_{T_1 \circ \cdots \circ T_j \circ T}\|_{X^p} \lesssim (L^{O(\theta)} \rho)^{l(T_1 \circ \cdots \circ T_j \circ T)} \lesssim L^{O(l(T)\theta)} \rho^{l(T)}$. Since $\rho = \alpha T_{\max}^{\frac{1}{2}} \lesssim L^{-\varepsilon}$ and $l(T) > N$ in the sum of Err , we get

$$\sum_{j=1}^K \|L^j(Err(\xi))\|_{X^p} \lesssim L^{O(N\theta)} \rho^N \lesssim L^{O(N\theta)} L^{-N\varepsilon} \ll L^{-M} \quad (2.47)$$

if we take $N \gg M/\varepsilon$ and $\theta \ll \varepsilon$.

Taking $K = 1$ in (2.41), we know that $\|L\|_{L^\infty X^p \rightarrow L^\infty X^p} \leq L^{O(1)}$, so $\|L^j\|_{L^\infty X^p \rightarrow L^\infty X^p} \leq L^{O(K)}$ if $j \leq K$. Taking $M \gg K$, by (2.43), we have $\sup_{t \leq T} \|w(t)\|_{X^p} \leq CL^{-M}$.

Therefore, we have

$$\|(1 + L + \cdots + L^{K-1})(B(w, w))\|_{X^p} \lesssim L^{O(K)} L^d \|w\|_{X^p}^2 \lesssim L^{O(K)} L^{O(1)} L^{-2M} \ll L^{-M}. \quad (2.48)$$

Combining (2.47) and (2.48), we prove $\sup_{t \leq T} \|w(t)\|_{X^p} \ll L^{-M} < CL^{-M}$ from the assumption that $\sup_{t \leq T} \|w(t)\|_{X^p} \leq CL^{-M}$. We thus complete the proof of Theorem 2.2.9. \square

2.2.4 Proof of the main theorem

In this section, we prove Theorem 2.1.1.

Proof of Theorem 2.1.1. Step 1. ($\mathbb{E}|\widehat{\psi}(t, k)|^2$ is close to $\mathbb{E}|\psi_{app,k}|^2$) By Theorem 2.2.9, we know that when $t \leq T_{\max} = L^{-\varepsilon}\alpha^{-2}$, we have $\|w\|_{X^p} \leq L^{-M}$ with L -certainly.

The above inequality is equivalent to $\sup_k |\langle k \rangle^s w_k| \leq CL^{-M}$. Remember that $w := \psi - \psi_{app}$, so L -certainly we have the following estimate

$$\sup_k \langle k \rangle^s |\psi_k - \psi_{app,k}| \leq CL^{-M} \quad (2.49)$$

Denote by A the event that the above estimate is true, then $\mathbb{E}|\widehat{\psi}(t, k)|^2 = \mathbb{E}(|\psi_k|^2 1_A) + \mathbb{E}(|\psi_k|^2 1_{A^c})$. L -certainly implies that $\mathbb{P}(A^c) \lesssim e^{-CL^\theta}$. Since $\|\psi\|_{L^2}$ is conservative and $|\psi_k|^2 \leq L^{d/2} \|\psi\|_{L^2} \leq L^{d/2}$, we know that $\mathbb{E}(|\psi_k|^2 1_{A^c}) \lesssim L^{d/2} e^{-CL^\theta} = O(L^{-M})$. Therefore, $\mathbb{E}|\widehat{\psi}(t, k)|^2 = \mathbb{E}(|\psi_k|^2 1_A) + O(L^{-M})$. Since we also have $\mathbb{E}|\psi_{app,k}|^2 = \mathbb{E}(|\psi_{app,k}|^2 1_A) + O(L^{-M})$, we conclude that

$$\mathbb{E}|\widehat{\psi}(t, k)|^2 = \mathbb{E}|\psi_{app,k}|^2 + \mathbb{E}((|\psi_k|^2 - |\psi_{app,k}|^2) 1_A) + O(L^{-M}) \quad (2.50)$$

By (2.49), $\mathbb{E}((|\psi_k|^2 - |\psi_{app,k}|^2) 1_A) = O(L^{-M})$. We may conclude that

$$\mathbb{E}|\widehat{\psi}(t, k)|^2 = \mathbb{E}|\psi_{app,k}|^2 + O(L^{-M}). \quad (2.51)$$

This suggests that we may get the approximation of $\mathbb{E}|\widehat{\psi}(t, k)|^2$ by calculating $\mathbb{E}|\psi_{app,k}|^2$.

Step 2. (Expansion of $\mathbb{E}|\psi_{app,k}|^2$) By (2.27), we know that

$$\phi_{app} = \sum_{l(T) \leq N} \mathcal{J}_T \quad (2.52)$$

Define

$$n^{(j)}(k) := \sum_{l(T)+l(T')=j} \mathbb{E} \mathcal{J}_{T,k} \overline{\mathcal{J}_{T',k}} \quad (2.53)$$

then Proposition 2.2.7 or 2.4.26 gives upper bounds of $n^{(j)}(k)$, which proves (1) and (2.6) of Theorem 2.1.1.

Step 3. (Asymptotics of $n^{(1)}(k)$) The only thing left in Theorem 2.1.1 is (2.5). This is a corollary of Proposition 2.4.38. \square

2.3 Lattice points counting and convergence results

2.4 Lattice points counting and convergence results

In this section, we prove Proposition 2.2.7 and 2.2.8 which gives upper bounds for tree terms $\mathcal{J}_{T,k}$ and the linearization operator \mathcal{P}_T . As explained before, these results are crucial in the proof of the main theorem. The proof is divided into several steps.

In section 2.4.1, we calculate the coefficients of $\mathcal{J}_{T,k}$ as polynomials of Gaussian random variables.

In section 2.4.2, we obtain upper bounds for the coefficients of these Gaussian polynomials.

Large deviation theory suggests that an upper bound of a Gaussian polynomial can be derived from an upper bound of its expectation and variance.

In section 2.4.3, we introduce the concept of couples which is a graphical method of calculating the expectation of Gaussian polynomials.

In section 2.4.4, we use couple to establish a lattice point counting result.

In section 2.4.5 and section 2.4.6, we apply the lattice points counting result

to derive upper bounds for $\mathcal{J}_{T,k}$ and \mathcal{P}_T respectively. This finishes the proof of Proposition 2.2.7 and 2.2.8 and therefore the proof of the main theorem.

2.4.1 Refined expression of coefficients

From (2.26), it is easy to show that $\mathcal{J}_{T,k}$ are polynomials of ξ . In this section, we calculate the coefficients of $\mathcal{J}_{T,k}$ using the definition (2.26) of them.

Notice that all non-leaf nodes other than the root in a tree have degree 3. For convenience, we add a new edge, called leg \mathfrak{l} , to the root node, which makes the root also of degree 3. This process is illustrated by Figure 2.3.

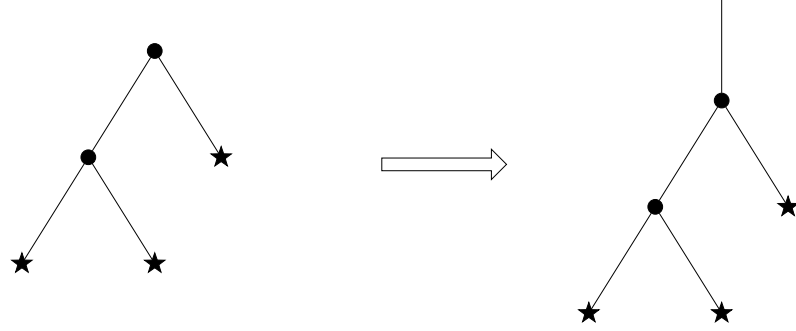


Figure 2.3: Adding a leg to a tree

We also need the following concepts about trees.

Definition 2.4.1. 1. **Nodes and children of an edge:** As in Figure 2.4, let \mathfrak{n}_u and \mathfrak{n}_l be two endpoints of an edge \mathfrak{e} and assume that \mathfrak{n}_l is a children of \mathfrak{n}_u . We define \mathfrak{n}_u (resp. \mathfrak{n}_l) to the upper node (resp. lower node) of \mathfrak{e} . Let $\mathfrak{n}_1, \mathfrak{n}_2$ be the two children of \mathfrak{n}_l and let \mathfrak{e}_1 (resp. \mathfrak{e}_2) be the edge between two nodes \mathfrak{n}_l and \mathfrak{n}_1 (resp. \mathfrak{n}_2). $\mathfrak{e}_1, \mathfrak{e}_2$ are defined to be the two children edges of \mathfrak{e} .

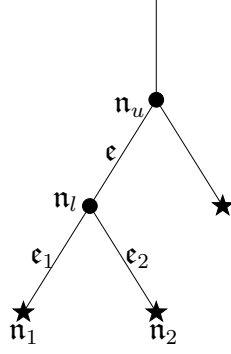


Figure 2.4: Children ϵ_1, ϵ_2 of an edge ϵ

2. **Direction of an edge:** As in Figure 2.5, each edge ϵ is assigned with a direction. This concept is mostly used to decide the value of variables $\iota_\epsilon \in \{\pm\}$ that will be defined later. Although it can be shown that the final result does not depend on the choices of the direction of each edge, for definiteness, we assign a downward direction to each edge. The orientation in Figure 2.5 is one example of this choice.

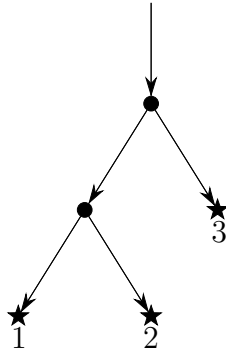


Figure 2.5: Children ϵ_1, ϵ_2 of an edge ϵ

3. **Labelling of leaves:** As in Figure 2.5, each leaf is labelled by $1, 2, \dots, l(T)+1$ from left to right. An edge pointing to a leaf \mathbf{n} is also labeled by j if \mathbf{n} is labeled by j .

Now we calculate the coefficients of $\mathcal{J}_{T,k}$.

Lemma 2.4.2. *Given a tree T of depth $l = l(T)$, denote by T_{in} the tree formed by all non-leaf nodes \mathbf{n} , then associate each node $\mathbf{n} \in T_{in}$ and edge $\mathbf{l} \in T$ with variables $t_{\mathbf{n}}$ and $k_{\mathbf{l}}$ respectively. Given a labelling of all leaves by $1, 2, \dots, l+1$, we identify $k_{\mathbf{e}}$ with k_j if \mathbf{e} is connected to a leaf labelled by j . Given a node \mathbf{n} , let $\mathbf{e}_1, \mathbf{e}_2$ and \mathbf{e} be the three edges from or pointing to \mathbf{n} (\mathbf{e} is the parent of \mathbf{e}_1 and \mathbf{e}_2), $\mathbf{n}_1, \mathbf{n}_2$ be children of \mathbf{n} and $\hat{\mathbf{n}}$ be the parent of \mathbf{n} .*

Let \mathcal{J}_T be terms defined in Definition 2.2.2, then their Fourier coefficients $\mathcal{J}_{T,k}$ are degree l polynomials of ξ given by the following formula

$$\mathcal{J}_{T,k} = \left(\frac{i\lambda}{L^d} \right)^l \sum_{k_1, k_2, \dots, k_{l+1}} H_{k_1 \dots k_{l+1}}^T \xi_{k_1} \xi_{k_2} \dots \xi_{k_{l+1}} \quad (2.54)$$

where $H_{k_1 \dots k_{l+1}}^T$ is given by

$$H_{k_1 \dots k_{l+1}}^T = \int_{\cup_{\mathbf{n} \in T_{in}} A_{\mathbf{n}}} e^{\sum_{\mathbf{n} \in T_{in}} i t_{\mathbf{n}} \Omega_{\mathbf{n}} - \nu(t_{\hat{\mathbf{n}}} - t_{\mathbf{n}}) |k_{\mathbf{e}}|^2} \prod_{\mathbf{n} \in T_{in}} dt_{\mathbf{n}} \delta_{\cap_{\mathbf{n} \in T_{in}} \{S_{\mathbf{n}}=0\}} \prod_{\mathbf{e} \in T_{in}} \iota_{\mathbf{e}} k_{\mathbf{e},x}, \quad (2.55)$$

and $\iota, A_{\mathbf{n}}, S_{\mathbf{n}}, \Omega_{\mathbf{n}}$ are defined by

$$\iota_{\mathbf{e}} = \begin{cases} +1 & \text{if } \mathbf{e} \text{ pointing inwards to } \mathbf{n} \\ -1 & \text{if } \mathbf{e} \text{ pointing outwards from } \mathbf{n} \end{cases} \quad (2.56)$$

$$A_{\mathbf{n}} = \begin{cases} \{t_{\mathbf{n}_1}, t_{\mathbf{n}_2}, t_{\mathbf{n}_3} \leq t_{\mathbf{n}}\} & \text{if } \mathbf{n} \neq \text{the root } \mathbf{r} \\ \{t_{\mathbf{r}} \leq t\} & \text{if } \mathbf{n} = \mathbf{r} \end{cases} \quad (2.57)$$

$$S_{\mathbf{n}} = \iota_{\mathbf{e}_1} k_{\mathbf{e}_1} + \iota_{\mathbf{e}_2} k_{\mathbf{e}_2} + \iota_{\mathbf{e}} k_{\mathbf{e}} \quad (2.58)$$

$$\Omega_{\mathbf{n}} = \iota_{\mathbf{e}_1} \Lambda_{k_{\mathbf{e}_1}} + \iota_{\mathbf{e}_2} \Lambda_{k_{\mathbf{e}_2}} + \iota_{\mathbf{e}} \Lambda_{k_{\mathbf{e}}} \quad (2.59)$$

For root node \mathbf{r} , we impose the constrain that $k_{\mathbf{r}} = k$ and $t_{\hat{\mathbf{r}}} = t$ (notice that \mathbf{r} does not have a parent so $\hat{\mathbf{r}}$ is not well defined).

Proof. We can check that \mathcal{J}_T is defined by (2.54) and (2.55) satisfies the recursive formula (2.26) by direct substitution, so they are the unique solution of that recursive formula, and this proves Lemma 2.4.2. \square

2.4.2 An upper bound of coefficients in expansion series

In this section, we derive an upper bound for coefficients $H_{k_1 \dots k_{l+1}}^T$.

Notice that in (2.54), $H_{k_1 \dots k_{l+1}}^T$ are integral of some oscillatory functions. An upper bound can be derived by the standard integration by parts arguments.

Associate each $\mathbf{n} \in T_{\text{in}}$ with two variables $a_{\mathbf{n}}$, $b_{\mathbf{n}}$. Then we define

$$F_T(t, \{a_{\mathbf{n}}\}_{\mathbf{n} \in T_{\text{in}}}, \{b_{\mathbf{n}}\}_{\mathbf{n} \in T_{\text{in}}}) = \int_{\cup_{\mathbf{n} \in T_{\text{in}}} A_{\mathbf{n}}} e^{\sum_{\mathbf{n} \in T_{\text{in}}} i t_{\mathbf{n}} a_{\mathbf{n}} - \nu(t_{\mathbf{n}} - t_{\mathbf{n}}) b_{\mathbf{n}}} \prod_{\mathbf{n} \in T_{\text{in}}} dt_{\mathbf{n}} \quad (2.60)$$

Lemma 2.4.3. *We have the following upper bound for $F_T(t, \{a_{\mathbf{n}}\}_{\mathbf{n}}, \{b_{\mathbf{n}}\}_{\mathbf{n}})$,*

$$\sup_{\{b_{\mathbf{n}}\}_{\mathbf{n}} \lesssim 1} |F_T(t, \{a_{\mathbf{n}}\}_{\mathbf{n}}, \{b_{\mathbf{n}}\}_{\mathbf{n}})| \lesssim \sum_{\{d_{\mathbf{n}}\}_{\mathbf{n} \in T_{\text{in}}} \in \{0,1\}^{l(T)}} \prod_{\mathbf{n} \in T_{\text{in}}} \frac{1}{|q_{\mathbf{n}}| + T_{\text{max}}^{-1}}. \quad (2.61)$$

Fix a sequence $\{d_{\mathbf{n}}\}_{\mathbf{n} \in T_{\text{in}}}$ whose elements $d_{\mathbf{n}}$ takes boolean values $\{0,1\}$. We define the two sequences $\{q_{\mathbf{n}}\}_{\mathbf{n} \in T_{\text{in}}}$, $\{r_{\mathbf{n}}\}_{\mathbf{n} \in T_{\text{in}}}$ by following recursive formula

$$q_{\mathbf{n}} = \begin{cases} a_{\mathbf{r}}, & \text{if } \mathbf{n} = \text{the root } \mathbf{r}. \\ a_{\mathbf{n}} + d_{\mathbf{n}} q_{\mathbf{n}'}, & \text{if } \mathbf{n} \neq \mathbf{r} \text{ and } \mathbf{n}' \text{ is the parent of } \mathbf{n}. \end{cases} \quad (2.62)$$

$$r_{\mathbf{n}} = \begin{cases} b_{\mathbf{r}}, & \text{if } \mathbf{n} = \text{the root } \mathbf{r}. \\ b_{\mathbf{n}} + d_{\mathbf{n}} q_{\mathbf{n}'}, & \text{if } \mathbf{n} \neq \mathbf{r} \text{ and } \mathbf{n}' \text{ is the parent of } \mathbf{n}. \end{cases} \quad (2.63)$$

Proof. The lemma is proved by induction.

For a tree T contains only one node \mathbf{r} , $F_T = 1$ and (2.61) is obviously true.

Assume that (2.61) is true for trees with $\leq n - 1$ nodes. We prove the n nodes

case.

For general T , let T_1, T_2 be the two subtrees and $\mathbf{n}_1, \mathbf{n}_2$ be the two children of the root \mathbf{r} , then by the definition of $F_T(2.60)$, we get

$$\begin{aligned}
F_T(t) &= \int_{\cup_{\mathbf{n} \in T_{\text{in}}} A_{\mathbf{n}}} e^{\sum_{\mathbf{n} \in T_{\text{in}}} it_{\mathbf{n}} a_{\mathbf{n}} - \nu(t_{\widehat{\mathbf{n}}} - t_{\mathbf{n}}) b_{\mathbf{n}}} \prod_{\mathbf{n} \in T_{\text{in}}} dt_{\mathbf{n}} \\
&= \int_{\cup_{\mathbf{n} \in T_{\text{in}}} A_{\mathbf{n}}} e^{it_{\mathbf{r}} a_{\mathbf{r}} - \nu(t - t_{\mathbf{r}}) b_{\mathbf{r}}} e^{\sum_{\mathbf{n} \in T_{\text{in},1} \cup T_{\text{in},2}} it_{\mathbf{n}} a_{\mathbf{n}} - \nu(t_{\widehat{\mathbf{n}}} - t_{\mathbf{n}}) b_{\mathbf{n}}} \left(dt_{\mathbf{r}} \prod_{j=1}^2 \prod_{\mathbf{n} \in T_{\text{in},j}} dt_{\mathbf{n}} \right) \\
&= \int_{\cup_{\mathbf{n} \in T_{\text{in}}} A_{\mathbf{n}}} e^{it_{\mathbf{r}}(a_{\mathbf{r}} + T_{\text{max}}^{-1} \text{sgn}(a_{\mathbf{r}})) - \nu(t - t_{\mathbf{r}}) b_{\mathbf{r}}} e^{-iT_{\text{max}}^{-1} t_{\mathbf{r}} \text{sgn}(a_{\mathbf{r}})} e^{\sum_{\mathbf{n} \in T_{\text{in},1} \cup T_{\text{in},2}} it_{\mathbf{n}} a_{\mathbf{n}} - \nu(t_{\widehat{\mathbf{n}}} - t_{\mathbf{n}}) b_{\mathbf{n}}} \left(dt_{\mathbf{r}} \prod_{j=1}^2 \prod_{\mathbf{n} \in T_{\text{in},j}} dt_{\mathbf{n}} \right)
\end{aligned} \tag{2.64}$$

We do integration by parts in the above integrals using the Stokes formula. Notice that for $t_{\mathbf{r}}$, there are three inequality constrains, $t_{\mathbf{r}} \leq t$ and $t_{\mathbf{r}} \geq t_{\mathbf{n}_1}, t_{\mathbf{n}_2}$.

$$\begin{aligned}
F_T(t) &= \frac{1}{ia_{\mathbf{r}} + iT_{\text{max}}^{-1} \text{sgn}(a_{\mathbf{r}}) + \nu b_{\mathbf{r}}} \int_{\cup_{\mathbf{n} \in T_{\text{in}}} A_{\mathbf{n}}} \frac{d}{dt_{\mathbf{r}}} e^{it_{\mathbf{r}}(a_{\mathbf{r}} + T_{\text{max}}^{-1} \text{sgn}(a_{\mathbf{r}})) - \nu(t - t_{\mathbf{r}}) b_{\mathbf{r}}} \\
&\quad e^{-iT_{\text{max}}^{-1} t_{\mathbf{r}} \text{sgn}(a_{\mathbf{r}})} e^{\sum_{\mathbf{n} \in T_{\text{in},1} \cup T_{\text{in},2}} it_{\mathbf{n}} a_{\mathbf{n}} - \nu(t_{\widehat{\mathbf{n}}} - t_{\mathbf{n}}) b_{\mathbf{n}}} \left(dt_{\mathbf{r}} \prod_{j=1}^2 \prod_{\mathbf{n} \in T_{\text{in},j}} dt_{\mathbf{n}} \right)
\end{aligned} \tag{2.65}$$

$$\begin{aligned}
&= \frac{1}{ia_{\mathbf{r}} + iT_{\text{max}}^{-1} \text{sgn}(a_{\mathbf{r}}) + \nu b_{\mathbf{r}}} \left(\int_{\cup_{\mathbf{n} \in T_{\text{in}}} A_{\mathbf{n}}, t_{\mathbf{r}}=t} - \int_{\cup_{\mathbf{n} \in T_{\text{in}}} A_{\mathbf{n}}, t_{\mathbf{r}}=t_{\mathbf{n}_1}} - \int_{\cup_{\mathbf{n} \in T_{\text{in}}} A_{\mathbf{n}}, t_{\mathbf{r}}=t_{\mathbf{n}_2}} \right) \\
&\quad e^{it_{\mathbf{r}}(a_{\mathbf{r}} + T_{\text{max}}^{-1} \text{sgn}(a_{\mathbf{r}})) - \nu(t - t_{\mathbf{r}}) b_{\mathbf{r}}} e^{-iT_{\text{max}}^{-1} t_{\mathbf{r}} \text{sgn}(a_{\mathbf{r}})} e^{\sum_{\mathbf{n} \in T_{\text{in},1} \cup T_{\text{in},2}} it_{\mathbf{n}} a_{\mathbf{n}} - \nu(t_{\widehat{\mathbf{n}}} - t_{\mathbf{n}}) b_{\mathbf{n}}} \left(dt_{\mathbf{r}} \prod_{j=1}^2 \prod_{\mathbf{n} \in T_{\text{in},j}} dt_{\mathbf{n}} \right)
\end{aligned}$$

$$\begin{aligned}
& -\frac{1}{ia_{\mathfrak{r}} + iT_{\max}^{-1}\text{sgn}(a_{\mathfrak{r}}) + \nu b_{\mathfrak{r}}} \int_{\cup_{\mathbf{n} \in T_{\text{in}}} A_{\mathbf{n}}} e^{it_{\mathfrak{r}}(a_{\mathfrak{r}} + T_{\max}^{-1}\text{sgn}(a_{\mathfrak{r}})) - \nu(t - t_{\mathfrak{r}})b_{\mathfrak{r}}} \\
& \quad \frac{d}{dt_{\mathfrak{r}}} (e^{-iT_{\max}^{-1}t_{\mathfrak{r}}\text{sgn}(a_{\mathfrak{r}})}) e^{\sum_{\mathbf{n} \in T_{\text{in},1} \cup T_{\text{in},2}} it_{\mathbf{n}}a_{\mathbf{n}} - \nu(t_{\mathbf{n}} - t_{\mathfrak{r}})b_{\mathbf{n}}} \left(dt_{\mathfrak{r}} \prod_{j=1}^2 \prod_{\mathbf{n} \in T_{\text{in},j}} dt_{\mathbf{n}} \right)
\end{aligned}$$

$$= \frac{1}{ia_{\mathfrak{r}} + iT_{\max}^{-1}\text{sgn}(a_{\mathfrak{r}}) + \nu b_{\mathfrak{r}}} (F_I - F_{T^{(1)}} - F_{T^{(2)}} - F_{II})$$

Here $T^{(j)}$, $j = 1, 2$ are trees that are obtained by deleting the root \mathfrak{r} , adding edges connecting \mathbf{n}_j with another node and defining \mathbf{n}_j to be the new root. For $T^{(j)}$, we can define the term $F_{T^{(j)}}$ by (2.60). It can be shown that $F_{T^{(j)}}$ defined in this way is the same as the $\int_{\cup_{\mathbf{n} \in T_{\text{in}}} A_{\mathbf{n}}, t_{\mathfrak{r}}=t_{\mathbf{n}_j}}$ term in the second equality of (2.65), so the last equality of (2.65) is true. F_I is the $\int_{\cup_{\mathbf{n} \in T_{\text{in}}} A_{\mathbf{n}}, t_{\mathfrak{r}}=t}$ term and F_{II} is the last term containing $\frac{d}{dt_{\mathfrak{r}}}$.

We can apply the induction assumption to $F_{T^{(j)}}$ and show that $\frac{1}{ia_{\mathfrak{r}} + iT_{\max}^{-1}\text{sgn}(a_{\mathfrak{r}}) + \nu b_{\mathfrak{r}}} F_{T^{(j)}}$ can be bounded by the right hand side of (2.61).

A direct calculation gives that

$$F_I(t) = e^{ita_{\mathfrak{r}}} F_{T_1}(t) F_{T_2}(t). \quad (2.66)$$

Then the induction assumption implies that $\frac{1}{ia_{\mathfrak{r}} + iT_{\max}^{-1}\text{sgn}(a_{\mathfrak{r}}) + \nu b_{\mathfrak{r}}} F_I$ can be bounded by the right hand side of (2.61).

Another direct calculation gives that

$$F_{II}(t) = \int_0^t e^{it_{\mathfrak{r}}(a_{\mathfrak{r}} + T_{\max}^{-1}\text{sgn}(a_{\mathfrak{r}})) - \nu(t - t_{\mathfrak{r}})b_{\mathfrak{r}}} \frac{d}{dt_{\mathfrak{r}}} (e^{-iT_{\max}^{-1}t_{\mathfrak{r}}\text{sgn}(a_{\mathfrak{r}})}) F_{T_1}(t_{\mathfrak{r}}) F_{T_2}(t_{\mathfrak{r}}) dt_{\mathfrak{r}}. \quad (2.67)$$

Apply the induction assumption

$$\begin{aligned}
& \left| \frac{1}{ia_{\mathfrak{r}} + iT_{\max}^{-1} \text{sgn}(a_{\mathfrak{r}}) + \nu b_{\mathfrak{r}}} F_{II}(t) \right| \\
& \leq \frac{1}{|q_{\mathfrak{r}}| + T_{\max}^{-1}} \prod_{j=1}^2 \left(\sum_{\{d_{\mathfrak{n}}\}_{\mathfrak{n} \in T_{\text{in},j}} \in \{0,1\}^{l(T_j)}} \prod_{\mathfrak{n} \in T_{\text{in},j}} \frac{1}{|q_{\mathfrak{n}}| + T_{\max}^{-1}} \right) \\
& \leq \sum_{\{d_{\mathfrak{n}}\}_{\mathfrak{n} \in T_{\text{in}}} \in \{0,1\}^{l(T)}} \prod_{\mathfrak{n} \in T_{\text{in}}} \frac{1}{|q_{\mathfrak{n}}| + T_{\max}^{-1}}.
\end{aligned} \tag{2.68}$$

Combining the bounds of F_I , $F_{T(1)}$, $F_{T(2)}$, F_{II} , we conclude that F_T can be bounded by the right hand side of (2.61) and thus complete the proof of Lemma 2.4.3. \square

A straightforward application of the above lemma gives the following upper bound of the coefficients $H_{k_1 \dots k_{l+1}}^T$.

Lemma 2.4.4. *We have the following upper bound for $H_{k_1 \dots k_{l+1}}^T$,*

$$|H_{k_1 \dots k_{l+1}}^T| \lesssim \sum_{\{d_{\mathfrak{n}}\}_{\mathfrak{n} \in T_{\text{in}}} \in \{0,1\}^{l(T)}} \prod_{\mathfrak{n} \in T_{\text{in}}} \frac{1}{|q_{\mathfrak{n}}| + T_{\max}^{-1}} \prod_{\mathfrak{e} \in T_{\text{in}}} |k_{\mathfrak{e},x}| \delta_{\cap_{\mathfrak{n} \in T_{\text{in}}} \{S_{\mathfrak{n}}=0\}}. \tag{2.69}$$

Fix a sequence $\{d_{\mathfrak{n}}\}_{\mathfrak{n} \in T_{\text{in}}}$ whose elements $d_{\mathfrak{n}}$ takes boolean values $\{0,1\}$. We define the two sequences $\{q_{\mathfrak{n}}\}_{\mathfrak{n} \in T_{\text{in}}}$, $\{r_{\mathfrak{n}}\}_{\mathfrak{n} \in T_{\text{in}}}$ by following recursive formula

$$q_{\mathfrak{n}} = \begin{cases} \Omega_{\mathfrak{r}}, & \text{if } \mathfrak{n} = \text{the root } \mathfrak{r}. \\ \Omega_{\mathfrak{n}} + d_{\mathfrak{n}} q_{\mathfrak{n}'}, & \text{if } \mathfrak{n} \neq \mathfrak{r} \text{ and } \mathfrak{n}' \text{ is the parent of } \mathfrak{n}. \end{cases} \tag{2.70}$$

$$r_{\mathfrak{n}} = \begin{cases} |k_{\mathfrak{r}}|^2, & \text{if } \mathfrak{n} = \text{the root } \mathfrak{r}. \\ |k_{\mathfrak{n}}|^2 + d_{\mathfrak{n}} q_{\mathfrak{n}'}, & \text{if } \mathfrak{n} \neq \mathfrak{r} \text{ and } \mathfrak{n}' \text{ is the parent of } \mathfrak{n}. \end{cases} \tag{2.71}$$

Proof. This is a direct corollary of Lemma 2.61 if we take $a_{\mathfrak{n}} = \Omega_{\mathfrak{n}}$, $b_{\mathfrak{n}} = |k_{\mathfrak{e}}|^2$. \square

Lemma 2.4.4 suggests that the coefficients are small when $|q_{\mathfrak{n}}| \gg T_{\max}^{-1}$. Therefore,

in order to bound $\mathcal{J}_{T,k}$, we should count the lattice points on $|q_{\mathbf{n}}| \lesssim T_{\max}^{-1}$

$$\{k_{\mathfrak{e}} \in \mathbb{Z}_L^d, |k_{\mathfrak{e}}| \lesssim 1, \forall \mathfrak{e} \in T : |q_{\mathbf{n}}| \lesssim T_{\max}^{-1}, S_{\mathbf{n}} = 0, \forall \mathbf{n} \in T. k_{\mathfrak{l}} = k\} \quad (2.72)$$

By solving (2.70), we know that $\Omega_{\mathbf{n}}$ is a linear combination of $q_{\mathbf{n}}$, so there exist constants $c_{\mathbf{n},\mathbf{n}'}$ such that $\Omega_{\mathbf{n}} = \sum_{\mathbf{n}'} c_{\mathbf{n},\mathbf{n}'} q_{\mathbf{n}'}$. Therefore, $|q_{\mathbf{n}}| \lesssim T_{\max}^{-1}$ implies that $|\Omega_{\mathbf{n}}| \leq \sum_{\mathbf{n}'} |c_{\mathbf{n},\mathbf{n}'} q_{\mathbf{n}'}| \lesssim T_{\max}^{-1}$.

$|\Omega_{\mathbf{n}}| \lesssim T_{\max}^{-1}$ implies that (2.72) is a subset of

$$\{k_{\mathfrak{e}} \in \mathbb{Z}_L^d, |k_{\mathfrak{e}}| \lesssim 1, \forall \mathfrak{e} \in T : |\Omega_{\mathbf{n}}| \lesssim T_{\max}^{-1}, S_{\mathbf{n}} = 0, \forall \mathbf{n} \in T. k_{\mathfrak{l}} = k\}. \quad (2.73)$$

To bound the number of elements of (2.72), we just need to do the same thing for (2.73).

(2.73) can be read from the tree diagrams T . As in Figure 2.6, each edge corresponds to a variable $k_{\mathfrak{e}}$. The leg \mathfrak{l} corresponds to equation $k_{\mathfrak{l}} = k$. Each node \mathbf{n} is connected with three edges $\mathfrak{e}_1, \mathfrak{e}_2, \mathfrak{e}$ whose corresponding variables $k_{\mathfrak{e}_1}, k_{\mathfrak{e}_2}, k_{\mathfrak{e}}$ satisfy the momentum conservation equation

$$\iota_{\mathfrak{e}_1} k_{\mathfrak{e}_1} + \iota_{\mathfrak{e}_2} k_{\mathfrak{e}_2} + \iota_{\mathfrak{e}} k_{\mathfrak{e}} = 0 \quad (2.74)$$

and the energy conservation equation (if the node is decorated by \bullet)

$$\iota_{\mathfrak{e}_1} \Lambda_{k_{\mathfrak{e}_1}} + \iota_{\mathfrak{e}_2} \Lambda_{k_{\mathfrak{e}_2}} + \iota_{\mathfrak{e}} \Lambda_{k_{\mathfrak{e}}} = O(T_{\max}^{-1}). \quad (2.75)$$

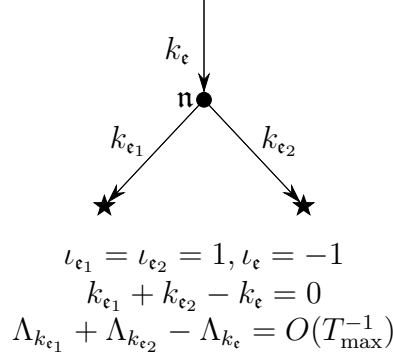


Figure 2.6: Equations of a node \mathbf{n}

The goal of the next two sections is to count the number of solutions of a modified version of the above equation (2.73).

2.4.3 Couples and the Wick theorem

In this section, we calculate $\mathbb{E}|\mathcal{J}_{T,k}|^2$ using Wick theorem. We also introduce another type of diagrams, the couple diagrams, to represent the result.

By the upper bound in the last section, the coefficients $H_{k_1 \dots k_{l+1}}^T$ concentrate near the surface $q_{\mathbf{n}} = 0, \forall \mathbf{n}$. But to get an upper bound of $\mathcal{J}_{T,k}$, we need an upper bound of their variance $\mathbb{E}|\mathcal{J}_{T,k}|^2$. The coefficients of $\mathbb{E}|\mathcal{J}_{T,k}|^2$ also concentrate near a surface whose expression is similar to (2.72).

Let's derive the expression of the coefficients of $\mathbb{E}|\mathcal{J}_{T,k}|^2$ and its concentration surface.

By Lemma 2.4.2, we know that $\mathcal{J}_{T,k}$ is a polynomial of ξ which are proportional to i.i.d Gaussians. Therefore,

$$\begin{aligned}
\mathbb{E}|\mathcal{J}_{T,k}|^2 &= \mathbb{E}(\mathcal{J}_{T,k} \overline{\mathcal{J}_{T,k}}) = \left(\frac{\lambda}{L^d}\right)^{2l(T)} \sum_{k_1, k_2, \dots, k_{l(T)+1}} \sum_{k'_1, k'_2, \dots, k'_{l(T)+1}} \\
& H_{k_1 \dots k_{l(T)+1}}^T \overline{H_{k'_1 \dots k'_{l(T)+1}}^T} \mathbb{E} \left(\xi_{k_1} \xi_{k_2} \dots \xi_{k_{l(T)+1}} \xi_{k'_1} \xi_{k'_2} \dots \xi_{k'_{l(T)+1}} \right)
\end{aligned} \tag{2.76}$$

We just need to calculate

$$\mathbb{E}\left(\xi_{k_1}\xi_{k_2}\cdots\xi_{k_{l(T)+1}}\xi_{k'_1}\xi_{k'_2}\cdots\xi_{k'_{l(T)+1}}\right). \quad (2.77)$$

Notice that $\xi_k = \sqrt{n_{\text{in}}(k)}\eta_k(\omega)$ and η_k are i.i.d Gaussians. We can apply the Wick theorem to calculate the above expectations.

To introduce the Wick theorem, we need the following definition.

Definition 2.4.5. 1. **Pairing:** Suppose that we have a set $A = \{a_1, \dots, a_{2m}\}$.

A pairing is a partition of $A = \{a_{i_1}, a_{i_2}\} \cup \dots \cup \{a_{i_{2m-1}}, a_{i_{2m}}\}$ into m disjoint subsets which have exactly two elements. Given a pairing p , elements $a_{i_k}, a_{i_{k'}}$ in the same subset of p are called paired with each other, which is denoted by $a_{i_k} \sim_p a_{i_{k'}}$.

2. $\mathcal{P}(A)$: Denote by $\mathcal{P}(A)$ the set of all pairings of A .

Lemma 2.4.6 (Wick theorem). *Let $\{\eta_k\}_{k \in \mathbb{Z}_L^d}$ be i.i.d complex Gaussian random variables with reflection symmetry (i.e. $\eta_k = \bar{\eta}_{-k}$). Let \mathcal{P} be the set of all pairings of $\{k_1, k_2, \dots, k_{2m}\}$, then*

$$\mathbb{E}(\eta_{k_1} \cdots \eta_{k_{2m}}) = \sum_{p \in \mathcal{P}} \delta_p(k_1, \dots, k_{2m}), \quad (2.78)$$

where

$$\delta_p = \begin{cases} 1 & \text{if } k_i = -k_j \text{ for all } k_i \sim_p k_j, \\ 0 & \text{otherwise.} \end{cases} \quad (2.79)$$

Proof. By Isserlis' theorem, for X_1, X_2, \dots, X_n zero-mean i.i.d Gaussian, we have

$$\mathbb{E}[X_1 X_2 \cdots X_n] = \sum_{p \in \mathcal{P}} \prod_{i \sim_p j} \mathbb{E}[X_i X_j] \quad (2.80)$$

Here \mathcal{P} is the set of pairings of $\{1, 2, \dots, n\}$.

Since $\mathbb{E}[\eta_{k_i}\eta_{k_j}] = \delta_{k_i=-k_j}$, take $X_1 = \eta_{k_1}, \dots, X_{2m} = \eta_{k_{2m}}$, then we can check that $\prod_{i \sim_p j} \mathbb{E}[X_i X_j] = \delta_p(k_1, \dots, k_{2m})$. This finishes the proof of the Wick theorem. \square

Applying Wick theorem to (2.76), we get

$$\begin{aligned} \mathbb{E}|\mathcal{J}_{T,k}|^2 &= \left(\frac{\lambda}{L^d}\right)^{2l(T)} \sum_{p \in \mathcal{P}(\{k_1, \dots, k_{l(T)+1}, k'_1, \dots, k'_{l(T)+1}\})} \\ &\quad \underbrace{\sum_{k_1, k_2, \dots, k_{l(T)+1}} \sum_{k'_1, k'_2, \dots, k'_{l(T)+1}} H_{k_1 \dots k_{l(T)+1}}^T \overline{H_{k'_1 \dots k'_{l(T)+1}}^T} \delta_p(k_1, \dots, k_{l(T)+1}, k'_1, \dots, k'_{l(T)+1}) \sqrt{n_{\text{in}}(k_1)} \dots}_{\text{Term}(T,p)_k} \end{aligned} \quad (2.81)$$

We see that the correlation of two tree terms is a sum of smaller expressions $\text{Term}(T,p)$. By (2.73), the coefficients $H_{k_1 \dots k_{l(T)+1}}^T \overline{H_{k'_1 \dots k'_{l(T)+1}}^T}$ of $\text{Term}(T,p)$ concentrate near the subset

$$\{k_{\mathbf{e}}, k_{\mathbf{e}'} \in \mathbb{Z}_L^d, |k_{\mathbf{e}}|, |k_{\mathbf{e}'}| \lesssim 1, \forall \mathbf{e}, \mathbf{e}' \in T : |\Omega_{\mathbf{n}}|, |\Omega_{\mathbf{n}'}| \lesssim T_{\max}^{-1}, S_{\mathbf{n}} = S_{\mathbf{n}'} = 0 \forall \mathbf{n}, \mathbf{n}' \in T. k_{\mathbf{l}} = -k'_{\mathbf{l}} = k\}. \quad (2.82)$$

The pairing p in Wick theorem introduces new equations $k_i = -k'_j$ (defined in (2.79)) and the coefficients $H_{k_1 \dots k_{l(T)+1}}^T \overline{H_{k'_1 \dots k'_{l(T)+1}}^T} \delta_p$ concentrate near the subset

$$\begin{aligned} &\{k_{\mathbf{e}}, k_{\mathbf{e}'} \in \mathbb{Z}_L^d, |k_{\mathbf{e}}|, |k_{\mathbf{e}'}| \lesssim 1, \forall \mathbf{e}, \mathbf{e}' \in T : |\Omega_{\mathbf{n}}|, |\Omega_{\mathbf{n}'}| \lesssim T_{\max}^{-1}, S_{\mathbf{n}} = S_{\mathbf{n}'} = 0 \forall \mathbf{n}, \mathbf{n}' \in T. \\ &k_{\mathbf{l}} = -k'_{\mathbf{l}} = k. k_i = -k'_j \text{ (and } k_i = -k_j, k'_i = -k'_j) \text{ for all } k_i \sim_p k'_j \text{ (and } k_i \sim_p k_j, k'_i \sim_p k'_j)\}. \end{aligned} \quad (2.83)$$

As in the case of (2.73), there is a graphical representation of (2.83). To explain this, we need the concept of couples.

Definition 2.4.7 (Construction of couples). Given two trees T and T' , we flip the

orientation of all edges in T' (as in the two left trees in Figure 2.7). We also label their leaves by $1, 2, \dots, l(T) + 1$ and $1, 2, \dots, l(T') + 1$ so that the corresponding variables of these leaves are $k_1, k_2, \dots, k_{l(T)+1}$ and $k_1, k_2, \dots, k_{l(T')+1}$. Assume that we have a pairing p of the set $\{k_1, k_2, \dots, k_{l(T)+1}, k_1, k_2, \dots, k_{l(T')+1}\}$, then this pairing induces a pairing between leaves (if $k_i \sim_p k_j$ then define *the i -th leaf \sim_p the j -th leaf*). Given this pairing of leaves, we define the following procedure which glues two trees T and T' into a couple $\mathcal{C}(T, T, p)$. Some examples of pairing can be found in Figure 2.7.

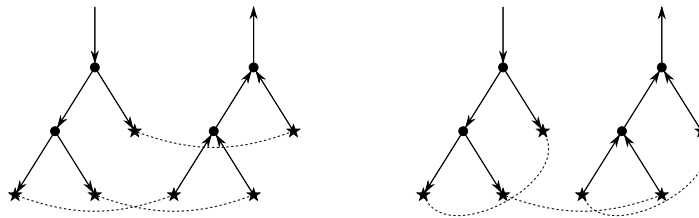


Figure 2.7: Example of pairings between trees.

1. **Merging edges connected to leaves:** Given two edges with opposite orientations connected to two paired leaves, these two edges can be merged into one edge as in Figure 2.8.

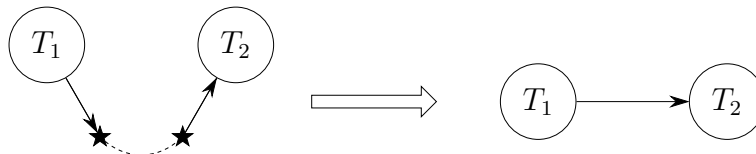


Figure 2.8: Pairing and merging of two edges

We know that two edges connected to leaves correspond to two indices k_i, k_j . Merging two such edges is a graphical interpretation that $k_i = -k_j$.

2. **Pairing of trees and couples:** Given a pairing p of the set of leaves in T, T' we merge all edges paired by p as in Figure 2.9 and the resulting combinatorial structure is called a couple.

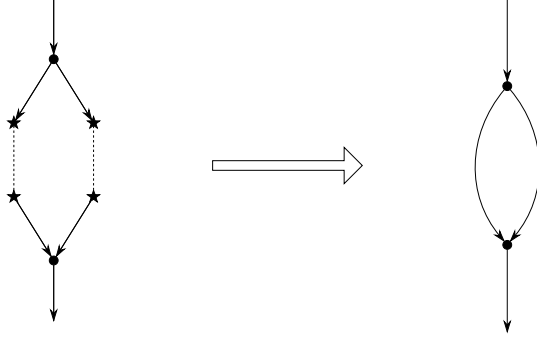


Figure 2.9: The construction of a couple

We know that each edge connected to a leaf corresponds to a variable k_i . A pairing p of $\{k_1, k_2, \dots, k_{2m}\}$ in (2.83) induces a pairing of edges connected to leaves. Merging paired edges corresponds to $k_i = -k'_j$ for all $k_i \sim_p k'_j$ in (2.83).

The following proposition introduces the graphical representation of (2.83).

Proposition 2.4.8. (2.83) can be read from a couple diagram $\mathcal{C}(T, T, p)$. Each edge corresponds to a variable k_{ϵ} . The leg \mathfrak{l} corresponds to equation $k_{\mathfrak{l}} = k$. Each node corresponds to a momentum conservation equation

$$\iota_{\epsilon_1} k_{\epsilon_1} + \iota_{\epsilon_2} k_{\epsilon_2} + \iota_{\epsilon} k_{\epsilon} = 0, \quad (2.84)$$

and an energy conservation equation

$$\iota_{\epsilon_1} \Lambda_{k_{\epsilon_1}} + \iota_{\epsilon_2} \Lambda_{k_{\epsilon_2}} + \iota_{\epsilon} \Lambda_{k_{\epsilon}} = O(T_{max}^{-1}). \quad (2.85)$$

Remark 2.4.9. In a couple diagram, we only have nodes decorated by \bullet . Nodes decorated by \star have been removed in (1), (2) of Definition 2.4.7.

Remark 2.4.10. Through the process of (1), (2) in Definition 2.4.7, a couple diagram can automatically encode the equation $k_i = -k'_j$ for all $k_i \sim_p k'_j$. Therefore, they do not appear in Proposition 2.4.8.

Proof. This directly follows from the definition of couples. \square

The calculations of this section are summarized in the following proposition.

Proposition 2.4.11. (1) Define $Term(T, p)$ in the same way as in (2.81),

$$Term(T, p)_k = \sum_{k_1, k_2, \dots, k_{l(T)+1}} \sum_{k'_1, k'_2, \dots, k'_{l(T)+1}} H_{k_1 \dots k_{l(T)+1}}^T \overline{H_{k'_1 \dots k'_{l(T)+1}}^T} \delta_p(k_1, \dots, k_{l(T)+1}, k'_1, \dots, k'_{l(T)+1}) \sqrt{n_{in}(k_1)} \dots \sqrt{n_{in}(k'_1)} \dots \quad (2.86)$$

then $\mathbb{E}|\mathcal{J}_{T,k}|^2$ is a sum of $Term(T, p)_k$ for all $p \in \mathcal{P}$, (in (2.81) the sum is over set of all possible pairing \mathcal{P})

$$\mathbb{E}|\mathcal{J}_{T,k}|^2 = \left(\frac{\lambda}{L^d}\right)^{2l(T)} \sum_{p \in \mathcal{P}(\{k_1, \dots, k_{l(T)+1}, k'_1, \dots, k'_{l(T)+1}\})} Term(T, p). \quad (2.87)$$

(2) $Term(T, p)$ concentrates near the subset (2.83) which has a simple graphical representation given by Proposition 2.4.8.

Proof. The proof of (1), (2) is easy and thus skipped. \square

2.4.4 Counting lattice points

In this section, we use the connection between couples and concentration subsets (2.83) to count the number of solutions of a generalized version of (2.83),

$$\begin{aligned} & \{k_{\mathbf{e}}, k_{\mathbf{e}'} \in \mathbb{Z}_L^d, |k_{\mathbf{e}}|, |k_{\mathbf{e}'}| \lesssim 1, \forall \mathbf{e}, \mathbf{e}' \in T : |k_{\mathbf{e}x}| \sim \kappa_{\mathbf{e}}, |k'_{\mathbf{e}x}| \sim \kappa_{\mathbf{e}'}, \forall \mathbf{e}, \mathbf{e}' \in T, \\ & |\Omega_{\mathbf{n}} - \sigma_{\mathbf{n}}|, |\Omega'_{\mathbf{n}} - \sigma'_{\mathbf{n}}| \lesssim T_{\max}^{-1}, S_{\mathbf{n}} = S_{\mathbf{n}'} = 0, \forall \mathbf{n}. k_{\mathbf{l}} = -k'_{\mathbf{l}} = k. \\ & k_i = -k'_j \text{ (and } k_i = -k_j, k'_i = -k'_j \text{) for all } k_i \sim_p k'_j \text{ (and } k_i \sim_p k_j, k'_i \sim_p k'_j \text{)}\}. \end{aligned} \quad (2.88)$$

In (2.88), $\kappa_{\mathfrak{e}} \in \{0\} \cup \mathcal{D}(\alpha, 1)$, where $\mathcal{D}(\alpha, 1) := \{2^{-K_{\mathfrak{e}}} : K_{\mathfrak{e}} \in \mathbb{Z} \cap [0, \ln \alpha^{-1}]\}$. The relation $|k_{\mathfrak{e}x}| \sim \kappa_{\mathfrak{e}}$ is defined by

$$|k_{\mathfrak{e}x}| \sim \kappa_{\mathfrak{e}} \text{ if and only if } \begin{cases} \frac{1}{2}\kappa_{\mathfrak{e}} \leq |k_{\mathfrak{e}x}| \leq 2\kappa_{\mathfrak{e}} & \text{if } \kappa_{\mathfrak{e}} \neq 0 \\ |k_{\mathfrak{e}x}| \lesssim \alpha^2, \ k_{\mathfrak{e}x} \neq 0 & \text{if } \kappa_{\mathfrak{e}} = 0 \end{cases} \quad (2.89)$$

(2.88) is obtained by replacing $\Omega_{\mathfrak{n}}, \Omega'_{\mathfrak{n}}$ by $\Omega_{\mathfrak{n}} - \sigma_{\mathfrak{n}}, \Omega'_{\mathfrak{n}} - \sigma'_{\mathfrak{n}}$ in (2.83) and adding conditions $|k_{\mathfrak{e}}| \sim \kappa_{\mathfrak{e}}$, where $\sigma_{\mathfrak{n}}, \sigma'_{\mathfrak{n}}$ and $\kappa_{\mathfrak{e}}$ are some given constants. The counterpart of Proposition 2.4.8 in this case is

Proposition 2.4.12. (2.88) *can be read from a couple diagram $\mathcal{C} = \mathcal{C}(T, T, p)$. Each edge corresponds to a variable $k_{\mathfrak{e}}$. The leg \mathfrak{l} corresponds to equation $k_{\mathfrak{l}} = k$. Each node corresponds to a momentum conservation equation*

$$\iota_{\mathfrak{e}_1} k_{\mathfrak{e}_1} + \iota_{\mathfrak{e}_2} k_{\mathfrak{e}_2} + \iota_{\mathfrak{e}} k_{\mathfrak{e}} = 0, \quad (2.90)$$

and a energy conservation equation

$$\iota_{\mathfrak{e}_1} \Lambda_{k_{\mathfrak{e}_1}} + \iota_{\mathfrak{e}_2} \Lambda_{k_{\mathfrak{e}_2}} + \iota_{\mathfrak{e}} \Lambda_{k_{\mathfrak{e}}} = \sigma_{\mathfrak{n}} + O(T_{max}^{-1}). \quad (2.91)$$

Denote the momentum and energy conservation equations by $MC_{\mathfrak{n}}$ and $EC_{\mathfrak{n}}$ respectively, then (2.88) can be rewritten as

$$(2.88) = \{k_{\mathfrak{e}} \in \mathbb{Z}_L^d, |k_{\mathfrak{e}}| \lesssim 1 \ \forall \mathfrak{e} \in \mathcal{C} : |k_{\mathfrak{e}x}| \sim \kappa_{\mathfrak{e}}, \ \forall \mathfrak{e} \in \mathcal{C}_{norm}. \ MC_{\mathfrak{n}}, \ EC_{\mathfrak{n}}, \ \forall \mathfrak{n} \in \mathcal{C}. \ k_{\mathfrak{l}} = -k'_{\mathfrak{l}} = k.\} \quad (2.92)$$

Proof. This directly follows from Proposition 2.4.8. □

To explain the counting argument in this paper, we need the following definitions related to couples.

Definition 2.4.13. 1. **Connected couples:** A couple \mathcal{C} is a connected couple if it is connected as a graph.

2. **Equations of a couple $Eq(\mathcal{C})$:** Given a couple \mathcal{C} and constants k, σ_n , let $Eq(\mathcal{C}, \{\sigma_n\}_n, k)$ (or simply $Eq(\mathcal{C})$) be the system of equation (2.92) constructed in Proposition 2.4.12. For any system of equations Eq , let $\#(Eq)$ be its number of solutions.

3. **Normal edges and leaf edges:** Remember that any couple \mathcal{C} is constructed from a pairing of two trees T, T' and therefore all edges in \mathcal{C} comes from T, T' . We define edges coming from $T_{\text{in}}, T'_{\text{in}}$ to be normal edges and edges from those connected to leaves in T, T' to be leaf edges. The set of all normal edges is denoted by $\mathcal{C}_{\text{norm}}$. A leg in \mathcal{C} which is a normal edge is called a normal leg.

The main goal of this section is to prove an upper bound of $\#Eq(\mathcal{C})$. The main idea of proving this is to decompose a large couple \mathcal{C} into smaller pieces and then prove this for the smaller piece using the induction hypothesis. To explain the idea, let us first focus on an example. Let \mathcal{C} be the left couple in the following picture. (The corresponding variables of each edge are labeled near these edges.)

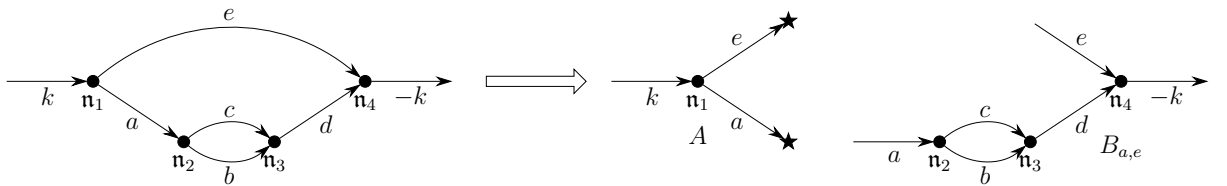


Figure 2.10: An example of decomposing a couple

By (2.92), we know that the couple \mathcal{C} corresponds to the following equations.

$$\begin{aligned}
& \{a, b, c, d, e : (|a| \text{ to } |e|) \lesssim 1, (|a_x| \text{ to } |e_x|) \sim (\kappa_a \text{ to } \kappa_e) \\
& a + e = k, \Lambda(a) + \Lambda(e) - \Lambda(k) = \sigma_1 + O(T_{\max}^{-1}) \\
& a + c = b, \Lambda(a) + \Lambda(c) - \Lambda(b) = \sigma_2 + O(T_{\max}^{-1}) \\
& b + c = d, \Lambda(b) + \Lambda(c) - \Lambda(d) = \sigma_3 + O(T_{\max}^{-1}) \\
& d + e + k = 0, \Lambda(d) + \Lambda(e) + \Lambda(k) = \sigma_4 + O(T_{\max}^{-1})\}
\end{aligned} \tag{2.93}$$

We know that (2.93) can be rewritten into the form $\bigcup_{a,e \in A} B_{a,e}$, where

$$A = \{a, e : |a|, |e| \lesssim 1, |a_x| \sim \kappa_a, |e_x| \sim \kappa_e, a + e = k, \Lambda(a) + \Lambda(e) - \Lambda(k) = \sigma_1 + O(T_{\max}^{-1})\} \tag{2.94}$$

$$\begin{aligned}
B_{a,e} = \{b, c, d : & |b|, |c|, |d| \lesssim 1, |b_x| \sim \kappa_b, |c_x| \sim \kappa_c, |d_x| \sim \kappa_d \\
& a + c = b, \Lambda(a) + \Lambda(c) - \Lambda(b) = \sigma_2 + O(T_{\max}^{-1}) \\
& b + c = d, \Lambda(b) + \Lambda(c) - \Lambda(d) = \sigma_3 + O(T_{\max}^{-1}) \\
& d + e + k = 0, \Lambda(d) + \Lambda(e) + \Lambda(k) = \sigma_4 + O(T_{\max}^{-1})\}
\end{aligned} \tag{2.95}$$

Since an upper bound of $\#Eq(\mathcal{C})$ can be derived from upper bounds of $\#A$, $\#B_{a,e}$, we just need to consider A , $B_{a,e}$ which are systems of equations of smaller size. We can reduce the size of systems of equations in this way and prove upper bounds by induction.

One problem of applying an induction argument is that A , $B_{a,e}$ cannot be represented by couple defined by Definition 2.4.7 that can contain at most two legs (an edge just connected to one node). In Definition 2.4.7, a leg is used to represent a variable that is fixed, as in the condition $k_l = -k'_l = k$ in (2.92). The definition of $\#B_{a,e}$ contains three fixed variables a , e , k which cannot be represented by just two legs. Therefore, we have to define a new type of couple that allows multiple legs.

Except for the lack of legs, we also have the problem of representing free variables.

We know that the couple representation of A should contain one node and three edges if we insist on the rule that a node corresponds to an equation and the variables in the equation correspond to edges connected to this node. All these edges are legs, but two of three edges correspond to variable a, b which are not fixed. Therefore, we have to define a type of legs that can correspond to unfixed variables.

To solve the above problems, we introduce the following definition.

Definition 2.4.14. 1. **Couples with multiple legs:** A graph in which all nodes have degree 1 or 3 is called a couples with multiple legs. The graph A and $B_{a,e}$ in Figure 2.10 are examples of this definition.

2. **Legs:** In a couple with multiple legs, an edge connected to a degree one node is called a leg. Remember that we have encountered this concept in the second paragraph of section 2.4.1 and in what follows we call the leg defined there the root leg of a tree.
3. **Free legs and fixed legs:** In a couple with multiple legs, we use two types of node decoration for degree 1 nodes as in Figure 2.11. One is \star and the other one is invisible.

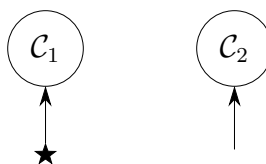


Figure 2.11: Node decoration of degree one node

An edge connected to a \star or invisible nodes is called a free leg or fixed leg respectively.

4. **Equations of a couple** $Eq(\mathcal{C}, \{c_i\}_I)$: We define the corresponding equations

for couples with multiple legs.

$$\begin{aligned}
& Eq(\mathcal{C}, \{c_{\mathfrak{l}}\}_{\mathfrak{l}}) \\
& = \{k_{\mathfrak{e}} \in \mathbb{Z}_L^d, |k_{\mathfrak{e}}| \lesssim 1 \ \forall \mathfrak{e} \in \mathcal{C} : |k_{\mathfrak{e}x}| \sim \kappa_{\mathfrak{e}}, \ \forall \mathfrak{e} \in \mathcal{C}_{\text{norm}}. MC_{\mathfrak{n}}, EC_{\mathfrak{n}}, \ \forall \mathfrak{n} \in \mathcal{C}. k_{\mathfrak{l}} = c_{\mathfrak{l}}, \ \forall \mathfrak{l}. \}
\end{aligned} \tag{2.96}$$

In this representation, the corresponding variable of a fixed leg \mathfrak{l} is fixed to be the constant $c_{\mathfrak{l}}$ and the corresponding variable of a free leg \mathfrak{l} is not fixed.

With the above definition, it's easy to show that the couple A and $B_{a,e}$ in Figure 2.10 correspond to the system of equations (2.94) and (2.95) respectively.

Using the above argument, we can prove the following proposition which gives an upper bound of a number of solutions of (2.88) (or (2.92)).

Proposition 2.4.15. *Let $\mathcal{C} = \mathcal{C}(T, T', p)$ be a connected couple with exactly one free and one fixed leg, n be the total number of nodes in \mathcal{C} and $Q = L^d T_{\max}^{-1}$. We fix $k \in \mathbb{R}$ for the legs $\mathfrak{l}, \mathfrak{l}'$ and $\sigma_{\mathfrak{n}} \in \mathbb{R}$ for each $\mathfrak{n} \in \mathcal{C}$. Assume that α satisfies (2.2). Then the number of solutions M of (2.88) (or (2.92)) is bounded by*

$$M \leq L^{O(n\theta)} Q^{\frac{n}{2}} \prod_{\mathfrak{e} \in \mathcal{C}_{\text{norm}}} \kappa_{\mathfrak{e}}^{-1}. \tag{2.97}$$

Proof. The proof is lengthy and therefore divided into several steps. The main idea of the proof is to use the operation of edge cutting to decompose the couple \mathcal{C} into smaller ones $\mathcal{C}_1, \mathcal{C}_2$, then apply Lemma 2.4.18 which relates $\#Eq(\mathcal{C})$ and $\#Eq(\mathcal{C}_i)$. The desired upper bounds of $\#Eq(\mathcal{C})$ can be obtained from that of $\#Eq(\mathcal{C}_i)$ inductively.

Step 1. In this step, we explain the cutting edge argument and prove the Lemma 2.4.18 which relates $\#Eq(\mathcal{C}_1)$, $\#Eq(\mathcal{C}_2)$ and $\#Eq(\mathcal{C})$.

Here is the formal definition of cutting

Definition 2.4.16. 1. **Cutting an edge:** Given an edge \mathfrak{e} , we can cut it into two

edges (a fixed and a free leg) as in Figure 2.12.

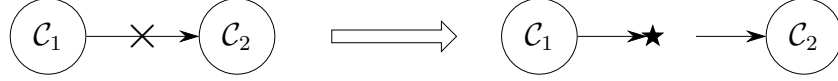


Figure 2.12: An example of cutting an edge

2. **Cut:** A cut c of a couple \mathcal{C} is a set of edges such that \mathcal{C} is disconnected after cutting all edges in c . A refined cut is a cut together with a map $rc : c \rightarrow \{\text{left}, \text{right}\}$. For each $\mathfrak{e} \in c$, if $rc(\mathfrak{e}) = \text{left}$ (resp. right), then as in Figure 2.12 the left node (resp. right node) produced by cutting \mathfrak{e} is a \star node (resp. invisible node). The map rc describes which one should be the free or fixed leg in the two legs produced by cutting an edge.
3. $c(\mathfrak{e})$, $c(\mathfrak{n})$ and $c(\mathfrak{l})$: Given an edge \mathfrak{e} that is not a leg, define $c(\mathfrak{l})$ to be the cut that contains only one edge \mathfrak{e} . Given a node $\mathfrak{n} \in \mathcal{C}$, let $\{\mathfrak{e}_i\}$ be edges that are connected to \mathfrak{n} , then define $c(\mathfrak{n})$ to be the cut that consists of edges $\{\mathfrak{e}_i\}$. Given a leg \mathfrak{l} , let \mathfrak{n} be the unique node connected to it, then define $c(\mathfrak{e})$ to be the cut $c(\mathfrak{n})$. An example of cutting $c(\mathfrak{e})$ is given by Figure 2.12. The following picture gives an example of cutting $c(\mathfrak{n})$ or $c(\mathfrak{l})$ (in this picture $\mathfrak{n} = \mathfrak{n}_1$ and \mathfrak{l} is the leg labelled by k .)

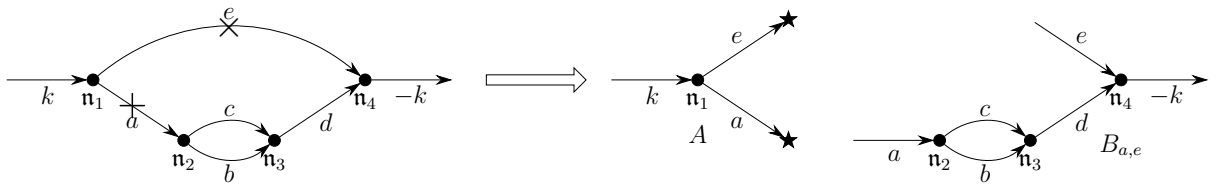


Figure 2.13: An example of cuts, $c(\mathfrak{n})$ and $c(\mathfrak{l})$

4. **Normal edges in couples with multiple legs:** In this paper, all couples with multiple legs are produced by cutting a couple defined in Definition 2.4.7. If a normal edge \mathfrak{e} is cut into \mathfrak{e}_1 and \mathfrak{e}_2 , then \mathfrak{e}_1 and \mathfrak{e}_2 are defined to be normal in the resulting couples with multiple legs.

Remark 2.4.17. Explicitly writing down the full definition of rc is often complicated, so in what follows, when defining rc , we will only describe which one should be the free or fixed leg in the two legs produced by cutting an edge.

The couples in Proposition 2.4.15 contains just 2 fixed legs, but after cutting, these couples may contain more fixed or free legs.

By Definition 2.4.14, for a couple \mathcal{C} with multiple legs, given constants c_l for each fixed leg l , the corresponding equation of \mathcal{C} is denoted by $Eq(\mathcal{C}, \{c_l\}_l)$. In $Eq(\mathcal{C}, \{c_l\}_l)$ each edge \mathfrak{e} is associated with a variable $k_{\mathfrak{e}}$ and each node \mathfrak{n} is still associated with equations $MC_{\mathfrak{n}}, EC_{\mathfrak{n}}$. The corresponding variables of free (resp. fixed) legs are free (resp. fixed to be a constant c_l).

Let us explain how does $Eq(\mathcal{C})$ and $\#Eq(\mathcal{C})$ changes after cutting. The result is summarized in the following lemma.

Lemma 2.4.18. *Let c be a cut of \mathcal{C} that consists of edges $\{\mathfrak{e}_i\}$ and $\mathcal{C}_1, \mathcal{C}_2$ be two components after cutting. Let $\mathfrak{e}_i^{(1)} \in \mathcal{C}_1, \mathfrak{e}_i^{(2)} \in \mathcal{C}_2$ be two edges obtained by cutting \mathfrak{e}_i . The rc map is defined by assigning $\{\mathfrak{e}_i^{(1)}\}$ to be free legs and $\{\mathfrak{e}_i^{(2)}\}$ to be fixed legs. Then we have*

$$Eq(\mathcal{C}, \{c_l\}_l) = \left\{ (k_{\mathfrak{e}_1}, k_{\mathfrak{e}_2}) : k_{\mathfrak{e}_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\}), k_{\mathfrak{e}_2} \in Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{\mathfrak{e}_i^{(1)}}\right\}_i\right) \right\}. \quad (2.98)$$

and

$$\sup_{\{c_l\}_l} \#Eq(\mathcal{C}, \{c_l\}_l) \leq \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_1)}} \#Eq(\mathcal{C}_1, \{c_{l_1}\}) \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}_2)}} \#Eq(\mathcal{C}_2, \{c_{l_2}\}). \quad (2.99)$$

Here $\text{leg}(\mathcal{C})$ is the set of fixed legs in \mathcal{C} (not the set of all legs!).

Proof. By definition (2.96) we have

$$\begin{aligned}
Eq(\mathcal{C}, \{c_l\}_l) &= \{k_{\mathbf{e}} \in \mathbb{Z}_L^d, |k_{\mathbf{e}}| \lesssim 1 : |k_{\mathbf{e}x}| \sim \kappa_{\mathbf{e}}, \forall \mathbf{e} \in \mathcal{C}_{\text{norm}}. MC_{\mathbf{n}}, EC_{\mathbf{n}}, \forall \mathbf{n}. k_l = c_l, \forall l \in \text{leg}(\mathcal{C}).\} \\
&= \{(k_{\mathbf{e}_1}, k_{\mathbf{e}_2}) : |k_{\mathbf{e}_1}| \lesssim 1, MC_{\mathbf{n}_1}, EC_{\mathbf{n}_1}. \forall \mathbf{e}_1, \mathbf{n}_1 \in \mathcal{C}_1. |k_{\mathbf{e}x}| \sim \kappa_{\mathbf{e}}, \forall \mathbf{e} \in \mathcal{C}_{\text{norm}}. \\
&\quad k_{l_1} = c_{l_1}, \forall l_1 \in \text{leg}(\mathcal{C}) \cap \text{leg}(\mathcal{C}_1) \\
&\quad |k_{\mathbf{e}_2}| \lesssim 1, MC_{\mathbf{n}_2}, EC_{\mathbf{n}_2}. \forall \mathbf{e}_2, \mathbf{n}_2 \in \mathcal{C}_2. |k_{\mathbf{e}x}| \sim \kappa_{\mathbf{e}}, \forall \mathbf{e} \in \mathcal{C}_{\text{norm}}. \\
&\quad k_{l_2} = c_{l_2}, \forall l_2 \in \text{leg}(\mathcal{C}) \cap \text{leg}(\mathcal{C}_2), k_{\mathbf{e}_i^{(2)}} = k_{\mathbf{e}_i^{(1)}}, \forall \mathbf{e}_i \in c\} \\
&= \left\{ (k_{\mathbf{e}_1}, k_{\mathbf{e}_2}) : k_{\mathbf{e}_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\}), k_{\mathbf{e}_2} \in Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{\mathbf{e}_i^{(1)}}\right\}_i\right) \right\}
\end{aligned} \tag{2.100}$$

Here in $Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{\mathbf{e}_i^{(1)}}\right\}_i\right)$, $k_{\mathbf{e}_i^{(1)}}$ are view as a constant value and $k_{\mathbf{e}_i^{(2)}}$ are fixed to be this constant value.

Therefore, we have the following identity of $Eq(\mathcal{C}, \{c_l\}_l)$

$$Eq(\mathcal{C}, \{c_l\}_l) = \left\{ (k_{\mathbf{e}_1}, k_{\mathbf{e}_2}) : k_{\mathbf{e}_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\}), k_{\mathbf{e}_2} \in Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{\mathbf{e}_i^{(1)}}\right\}_i\right) \right\}. \tag{2.101}$$

which proves (2.98).

We can also find the relation between $\#Eq(\mathcal{C}_1)$, $\#Eq(\mathcal{C}_2)$ and $\#Eq(\mathcal{C})$. Applying (2.98),

$$\begin{aligned}
\#Eq(\mathcal{C}, \{c_l\}_l) &= \sum_{(k_{\mathbf{e}_1}, k_{\mathbf{e}_2}) \in \#Eq(\mathcal{C}, \{c_l\}_l)} 1 \\
&= \sum_{\left\{ (k_{\mathbf{e}_1}, k_{\mathbf{e}_2}) : k_{\mathbf{e}_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\}), k_{\mathbf{e}_2} \in Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{\mathbf{e}_i^{(1)}}\right\}_i\right) \right\}} 1 \\
&= \sum_{k_{\mathbf{e}_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\})} \sum_{k_{\mathbf{e}_2} \in Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{\mathbf{e}_i^{(1)}}\right\}_i\right)} 1 \\
&= \sum_{k_{\mathbf{e}_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\})} \#Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{\mathbf{e}_i^{(1)}}\right\}_i\right)
\end{aligned} \tag{2.102}$$

Take sup in the above equation

$$\begin{aligned}
\sup_{\{c_l\}_l} \#Eq(\mathcal{C}, \{c_l\}_l) &= \sup_{\{c_l\}_l} \sum_{k_{\epsilon_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\})} \#Eq(\mathcal{C}_2, \{c_{l_2}\}, \{k_{\epsilon_i^{(1)}}\}_i) \\
&\leq \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}) \cap \text{leg}(\mathcal{C}_1)}} \sum_{k_{\epsilon_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\})} \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}) \cap \text{leg}(\mathcal{C}_2)}} \#Eq(\mathcal{C}_2, \{c_{l_2}\}, \{k_{\epsilon_i^{(1)}}\}_i) \\
&\leq \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_1)}} \sum_{k_{\epsilon_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\})} \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}_2)}} \#Eq(\mathcal{C}_2, \{c_{l_2}\}) \\
&= \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_1)}} \#Eq(\mathcal{C}_1, \{c_{l_1}\}) \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}_2)}} \#Eq(\mathcal{C}_2, \{c_{l_2}\})
\end{aligned} \tag{2.103}$$

This proves (2.99). \square

Step 2. In this step, we specify the cutting procedure.

Notice that Proposition 2.4.24, the multiple leg analog of Proposition 2.4.15, is only true for couples satisfying the "property P". When designing the cutting procedure, we must make sure that all couples generated during the execution of this procedure satisfy the "property P". The following proposition guarantees the existence of such a procedure.

Proposition 2.4.19. *There exists a recursive algorithm that repeatedly decomposes \mathcal{C} into smaller pieces and satisfies the following requirements. In the rest of this paper, we will call this algorithm "the cutting algorithm".*

(1) *The input of the 0-th step of this algorithm is $\mathcal{C}(0) = \mathcal{C}$. The inputs of other steps are the outputs of previous steps of the algorithm itself.*

(2) *In step k , $\mathcal{C}(k)$ is decomposed into 2 or 3 connected components by cutting edges and all components with more than one node are outputted. For $\mathcal{C}(1)$, $\#Eq(\mathcal{C}(1)) = \#Eq(\mathcal{C})$ and $\mathcal{C}_{norm}(1) = \mathcal{C}_{norm}$.*

(3) *One of the connected components in (2) contains exactly one node \mathbf{n} and one fixed normal leg \mathbf{l} . We call this component $\mathcal{C}(k)_l$. There are only two possibilities of $\mathcal{C}(k)_l$ as in Figure 2.14. We label them by $\mathcal{C}_I, \mathcal{C}_{II}$.*



Figure 2.14: Two possibilities of \mathcal{C}_l .

(4) The cutting algorithm satisfies the requirement that all connected components in (2) generated in each step satisfy property P , where property P is defined below

Property P of a couple $\tilde{\mathcal{C}}$: $\tilde{\mathcal{C}}$ is connected and contains exactly one free leg and at least one fixed normal leg.

Remark 2.4.20. Although the definition of cut is rather general, only three special types of cuts, $c(\mathfrak{e})$, $c(\mathfrak{n})$ and $c(\mathfrak{l})$, are used in the cutting algorithm.

Remark 2.4.21. The couple \mathcal{C} does not satisfy the property P because it does not have any free leg.

Proof of Proposition 2.4.19. Consider the following algorithm.

The cutting algorithm

Step 0. The input $\mathcal{C}(0)$ of this step is \mathcal{C} . In this step, we replace one of the two fixed legs of \mathcal{C} with a free leg to obtain a new couple $\hat{\mathcal{C}}$. By Lemma 2.4.22 (3), $\#Eq(\mathcal{C}) = \#Eq(\hat{\mathcal{C}})$. The output $\mathcal{C}(1)$ of this step is $\hat{\mathcal{C}}$. An example of step 0 can be found in the following picture.

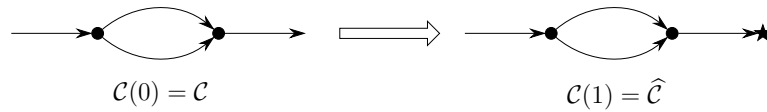


Figure 2.15: An example of step 0

Step k . Assume that the step $k - 1$ have been finished. The input $\mathcal{C}(k)$ of step k is the output of step $k - 1$. (If there are two output couples from step $k - 1$, apply to step k to these two couples separately.)

By property P, there exists a fixed normal leg in $\mathcal{C}(k)$. Choose one such leg \mathfrak{l} and define $\mathcal{C}(k)_{\mathfrak{l}}$ to be the component which contains \mathfrak{l} after cutting $c(\mathfrak{l})$ and define $\mathcal{C}(k)' = \mathcal{C}(k) \setminus \mathcal{C}(k)_{\mathfrak{l}}$. Check how many components $\mathcal{C}(k)'$ have. Jump to case 1 if the number of components is equal to 1, otherwise, jump to case 2. Examples of case 1 and case 2 can be found in the following picture.

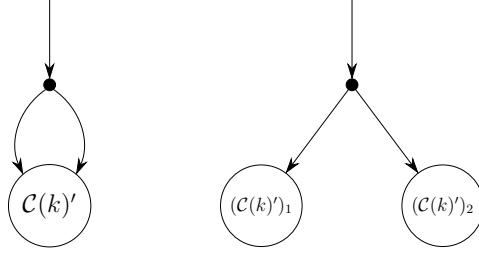


Figure 2.16: Examples of case 1 and case 2 (Left is case 1 and right is case two. $(\mathcal{C}(k)')_1$ and $(\mathcal{C}(k)')_2$ are two components of $\mathcal{C}(k)'$)

Case 1 of step k . In this case $\mathcal{C}(k)'$ has one components. We rename $\mathcal{C}(k)'$ to $\mathcal{C}(k)_1$. By property P, there exists a unique free leg \mathfrak{l}_{fr} in $\mathcal{C}(k)$. Check if \mathfrak{l}_{fr} and \mathfrak{l} are connected to the same node. If yes, jump to case 1.1, otherwise jump to case 1.2.

Case 1.1. Cut edges in $c(\mathfrak{l})$ into $\{\mathfrak{e}_i^{(1)}\}_{i=1,2}$ and $\{\mathfrak{e}_i^{(2)}\}_{i=1,2}$, then $\mathcal{C}(k)$ is decomposed into $\mathcal{C}(k)_{\mathfrak{l}}$, $\mathcal{C}(k)_1 = \mathcal{C} \setminus \mathcal{C}_{\mathfrak{l}}$. As in Lemma 2.4.18, define $\{\mathfrak{e}_i^{(1)}\} \subseteq \mathcal{C}(k)_{\mathfrak{l}}$ to be free legs and $\{\mathfrak{e}_i^{(2)}\} \subseteq \mathcal{C}(k)_1$ to be fixed legs.

If $\mathcal{C}(k)_1$ satisfies the property P, define $\mathcal{C}(k+1) = \mathcal{C}(k)_1$ to be the output of step k and apply step $k+1$ to $\mathcal{C}(k+1)$.

Otherwise, by Lemma 2.4.23 (2) there exists exactly one free normal leg and at least one fixed leg. By Lemma 2.4.22, we can define a new couple $\widehat{\mathcal{C}}$ such that the free normal leg becomes fixed and the fixed leg becomes free. Finally, define $\mathcal{C}(k+1) = \widehat{\mathcal{C}}$ to be the output.

Examples of cutting in case 1.1 can be found in the following picture.

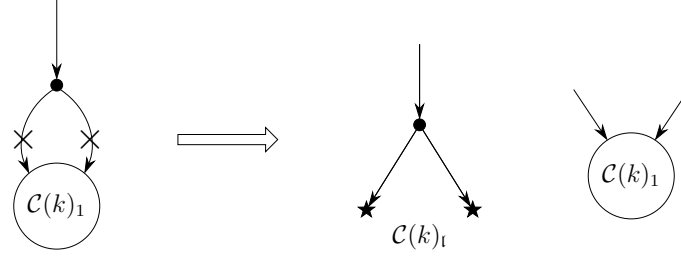


Figure 2.17: Examples of cutting in case 1.1

Case 1.2. In this case, \mathfrak{l}_{fr} and \mathfrak{l} are connected to the same node \mathfrak{n} . Let \mathfrak{e} be the edge connecting \mathfrak{n} and another interior node. Cut \mathfrak{e} into $\{\mathfrak{e}^{(1)}, \mathfrak{e}^{(2)}\}$ and then $\mathcal{C}(k)$ is decomposed into $\mathcal{C}(k)_I$, $\mathcal{C}(k)_1 = \mathcal{C}(k) \setminus \mathcal{C}(k)_I$. Define $\mathfrak{e}^{(1)} \in \mathcal{C}(k)_I$ to be fixed legs and $\mathfrak{e}^{(2)} \in \mathcal{C}(k)_1$ to be free legs.

If $\mathcal{C}(k)_1$ satisfies the property P, define $\mathcal{C}(k+1) = \mathcal{C}(k)_1$.

Otherwise by Lemma 2.4.23 (2), $\mathfrak{e}^{(2)}$ is the only normal legs. Using Lemma 2.4.22 (2), we may construct a new couple $\mathcal{C}(k+1)$ by assigning $\mathfrak{e}^{(2)}$ to be fixed and another leg to be free. Finally define $\mathcal{C}(k+1)$ to be the output of step k and apply step $k+1$ to $\mathcal{C}(k+1)$.

Examples of cutting in case 1.2 can be found in the following picture.

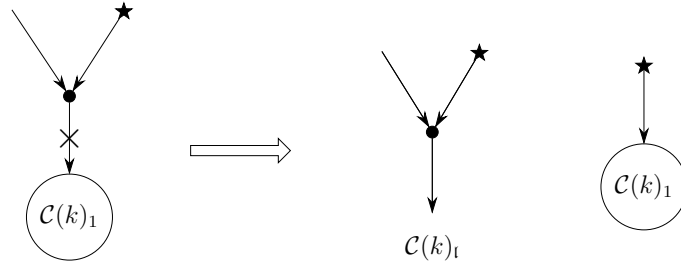


Figure 2.18: Examples of cutting in case 1.2

Case 2 of step k . Let the two connected components of $\mathcal{C}(k) \setminus \mathcal{C}(k)_I$ be $\mathcal{C}(k)_2$ and $\mathcal{C}(k)_3$. Let $\mathfrak{e}_2, \mathfrak{e}_3$ be the two edges that connect \mathfrak{l} and $\mathcal{C}(k)_2, \mathcal{C}(k)_3$ respectively. Cut $\mathfrak{e}_2, \mathfrak{e}_3$ into $\{\mathfrak{e}_2^{(1)}, \mathfrak{e}_3^{(1)}\} \subseteq \mathcal{C}(k)_I$ and $\mathfrak{e}_2^{(2)} \in \mathcal{C}(k)_2, \mathfrak{e}_3^{(2)} \in \mathcal{C}(k)_3$. By Lemma 2.4.23, $\mathcal{C}(k)_2, \mathcal{C}(k)_3$ contain at least one normal leg and two legs. By symmetry, we can

just consider $\mathcal{C}(k)_2$.

If $\mathcal{C}(k)_2$ contains free legs, define $\mathfrak{e}_2^{(2)} \in \mathcal{C}(k)_2$ to be fixed, otherwise define $\mathfrak{e}_2^{(2)}$ to be free.

In the case that $\mathcal{C}(k)_2$ contains free legs, define the output $\mathcal{C}(k+1)$ to be $\mathcal{C}(k)_2$ or $\mathcal{C}(k)_3$ and apply step $k+1$ to them separately.

In the case that $\mathcal{C}(k)_2$ contains no free legs, $\mathfrak{e}_2^{(2)}$ is the only normal leg and it is defined to be free. Use Lemma 2.4.22 (2) to construct a new couple $\widehat{\mathcal{C}}_2$ by assigning $\mathfrak{e}_2^{(2)}$ to be fixed and another leg to be free. Then define $\mathcal{C}(k+1)$ to be $\widehat{\mathcal{C}}_2$ or $\widehat{\mathcal{C}}_3$ and apply step $k+1$ to them separately.

Examples of cutting in case 2 can be found in the following picture.

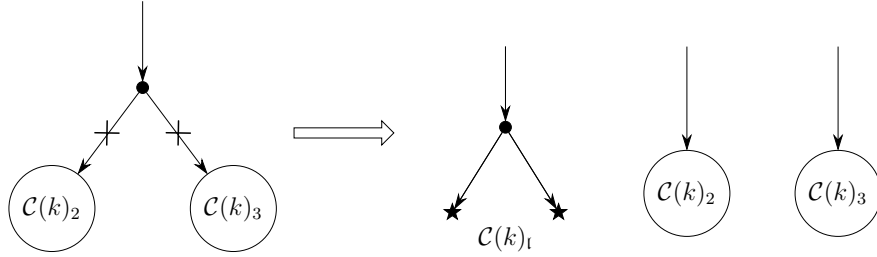


Figure 2.19: Examples of cutting in case 2

(1), (3) are true by definition. (2) is true because $\#Eq(\mathcal{C}) = \#Eq(\widehat{\mathcal{C}})$ as explained in step 0 of the algorithm. Since in step 0, we just replace a fixed leg with a free leg, $\mathcal{C}_{\text{norm}}$ should not change, so we get $\mathcal{C}_{\text{norm}}(1) = \mathcal{C}_{\text{norm}}$.

The only non-trivial part is (4), which is a corollary of Lemma 2.4.23 below.

Therefore, we complete the proof. \square

Lemma 2.4.22. *Given a connected couple \mathcal{C} with multiple legs, then we have the following conclusions.*

(1) Let $\{k_{l_i}\}_{i=1, \dots, n_{\text{leg}}}$ be the variables corresponding to legs in couple \mathcal{C} . Let $\iota_{\mathfrak{e}}$ be the same as (2.56), then $Eq(\mathcal{C})$ implies the following momentum conservation equation

$$\sum_{i=1}^{n_{\text{leg}}} \iota_{l_i} k_{l_i} = 0, \quad (2.104)$$

(2) Assume that there is exactly one free leg \mathfrak{l}_{i_0} in \mathcal{C} and all other variables $\{k_{\mathfrak{l}_i}\}_{i \neq i_0}$ corresponding to fix legs are fixed to be constants $\{c_{\mathfrak{l}_i}\}_{i \neq i_0}$. For any $i_1 = 1, \dots, n_{\text{leg}}$, we can construct a new couple $\widehat{\mathcal{C}}$ by replacing the i_0 leg by a fixed leg and i_1 leg by a free leg. If $i \neq i_0, i_1$, fix $k_{\mathfrak{l}_i}$ to be the constant $c_{\mathfrak{l}_i}$, if $i = i_0$, fix $k_{\mathfrak{l}_{i_0}}$ to be the constant $-\iota_{i_0} \sum_{i \neq i_0} \iota_{\mathfrak{l}_i} k_{\mathfrak{l}_i}$. Under the above assumptions, we have

$$Eq(\mathcal{C}, \{c_{\mathfrak{l}_i}\}_{i \neq i_0}) = Eq\left(\widehat{\mathcal{C}}, \{c_{\mathfrak{l}_i}\}_{i \neq i_0, i_1} \cup \left\{-\iota_{i_0} \sum_{i \neq i_0} \iota_{\mathfrak{l}_i} k_{\mathfrak{l}_i}\right\}\right). \quad (2.105)$$

(3) Assume that there is no free leg in \mathcal{C} and all $\{k_{\mathfrak{l}_i}\}_{i \neq i_0}$ are fixed to be constants $\{c_{\mathfrak{l}_i}\}_{i \neq i_0}$. For any $i_1 = 1, \dots, n_{\text{leg}}$, we can construct a new couple $\widehat{\mathcal{C}}$ by replacing the i_0 leg by a free leg. Then we have

$$Eq(\mathcal{C}, \{c_{\mathfrak{l}_i}\}_i) = Eq(\widehat{\mathcal{C}}, \{c_{\mathfrak{l}_i}\}_{i \neq i_0}). \quad (2.106)$$

(4) If the couple \mathcal{C} contains any leg, then it contains at least two legs.

Proof. We first prove (1). Given a node \mathbf{n} and an edge \mathfrak{e} connected to it, we define $\iota_{\mathfrak{e}}(\mathbf{n})$ by the following rule

$$\iota_{\mathfrak{e}}(\mathbf{n}) = \begin{cases} +1 & \text{if } \mathfrak{e} \text{ pointing towards } \mathbf{n} \\ -1 & \text{if } \mathfrak{e} \text{ pointing outwards from } \mathbf{n} \end{cases} \quad (2.107)$$

For a leg \mathfrak{l} , since it is connected to just one node, we may omit the (\mathbf{n}) and just write $\iota_{\mathfrak{l}}$ as in the statement of the lemma.

For each node \mathbf{n} , let $\mathfrak{e}_1(\mathbf{n})$, $\mathfrak{e}_2(\mathbf{n})$, $\mathfrak{e}(\mathbf{n})$ be the three edges connected to it. For each edge \mathfrak{e} , let $\mathbf{n}_1(\mathfrak{e})$, $\mathbf{n}_2(\mathfrak{e})$ be the two nodes connected to it. Then we know that $\iota_{\mathfrak{e}}(\mathbf{n}_1(\mathfrak{e})) + \iota_{\mathfrak{e}}(\mathbf{n}_2(\mathfrak{e}))$, since $\mathbf{n}_1(\mathfrak{e})$ and $\mathbf{n}_2(\mathfrak{e})$ have the opposite direction.

Since $k_{\mathfrak{e}}$ satisfy $Eq(\mathcal{C})$, by (2.90), we get $\iota_{\mathfrak{e}_1(\mathbf{n})}(\mathbf{n})k_{\mathfrak{e}_1(\mathbf{n})} + \iota_{\mathfrak{e}_2(\mathbf{n})}(\mathbf{n})k_{\mathfrak{e}_2(\mathbf{n})} + \iota_{\mathfrak{e}(\mathbf{n})}(\mathbf{n})k_{\mathfrak{e}(\mathbf{n})} =$

0. Summing over \mathbf{n} gives

$$\begin{aligned}
0 &= \sum_{\mathbf{n} \in \mathcal{C}} \iota_{\mathfrak{e}_1(\mathbf{n})}(\mathbf{n}) k_{\mathfrak{e}_1(\mathbf{n})} + \iota_{\mathfrak{e}_2(\mathbf{n})}(\mathbf{n}) k_{\mathfrak{e}_2(\mathbf{n})} + \iota_{\mathfrak{e}(\mathbf{n})}(\mathbf{n}) k_{\mathfrak{e}(\mathbf{n})} \\
&= \sum_{\substack{\mathfrak{e} \text{ is not a leg}}} (\iota_{\mathfrak{e}}(\mathbf{n}_1(\mathfrak{e})) + \iota_{\mathfrak{e}}(\mathbf{n}_2(\mathfrak{e}))) k_{\mathfrak{e}} + \sum_{\substack{\mathfrak{l} \text{ is a leg}}} \iota_{\mathfrak{l}} k_{\mathfrak{l}} \\
&= \sum_{i=1}^{n_{\text{leg}}} \iota_{\mathfrak{l}_i} k_{\mathfrak{l}_i}
\end{aligned} \tag{2.108}$$

This proves (2.104) and thus proves (1).

Now we prove (2). Since in $Eq(\widehat{\mathcal{C}}, \{c_{\mathfrak{l}_i}\}_{i \neq i_0, i_1} \cup \{-\iota_{\mathfrak{l}_{i_0}} \sum_{i \neq i_0} \iota_{\mathfrak{l}_i} k_{\mathfrak{l}_i}\})$, $\{k_{\mathfrak{l}_i}\}_{i \neq i_0, i_1}$ are fixed to be constants $\{c_{\mathfrak{l}_i}\}_{i \neq i_0, i_1}$ and $k_{\mathfrak{l}_{i_0}}$ is fixed to be the constant $-\iota_{\mathfrak{l}_{i_0}} \sum_{i \neq i_0} \iota_{\mathfrak{l}_i} k_{\mathfrak{l}_i}$, by (2.104), we know that

$$\iota_{\mathfrak{l}_{i_0}} \left(-\iota_{\mathfrak{l}_{i_0}} \sum_{i \neq i_0} \iota_{\mathfrak{l}_i} k_{\mathfrak{l}_i} \right) + \iota_{\mathfrak{l}_{i_1}} k_{\mathfrak{l}_{i_1}} + \sum_{i \neq i_0, i_1} \iota_{\mathfrak{l}_i} c_{\mathfrak{l}_i} = 0. \tag{2.109}$$

This implies that $k_{\mathfrak{l}_{i_1}} = c_{\mathfrak{l}_{i_1}}$ in $Eq(\widehat{\mathcal{C}})$. Therefore, equations in $Eq(\widehat{\mathcal{C}})$ automatically imply $k_{\mathfrak{l}_{i_1}} = c_{\mathfrak{l}_{i_1}}$. Notice that whether or not containing $k_{\mathfrak{l}_{i_1}} = c_{\mathfrak{l}_{i_1}}$ is the only difference between $Eq(\mathcal{C})$ and $Eq(\widehat{\mathcal{C}})$. We conclude that $Eq(\mathcal{C}) = Eq(\widehat{\mathcal{C}})$. We thus complete the proof of (2).

The proof of (3) is similar to (2). Whether or not containing $k_{\mathfrak{l}_{i_0}} = c_{\mathfrak{l}_{i_0}}$ is the only difference between $Eq(\mathcal{C})$ and $Eq(\widehat{\mathcal{C}})$. But if $\{k_{\mathfrak{l}_i}\}_{i \neq i_0}$ are fixed to be constants $\{c_{\mathfrak{l}_i}\}_{i \neq i_0}$, by momentum conservation we know that

$$k_{\mathfrak{l}_{i_0}} = -\iota_{\mathfrak{l}_{i_0}} \sum_{i \neq i_0} \iota_{\mathfrak{l}_i} c_{\mathfrak{l}_i}. \tag{2.110}$$

Therefore, $k_{\mathfrak{l}_{i_0}}$ is fixed to be the constant $-\iota_{\mathfrak{l}_{i_0}} \sum_{i \neq i_0} \iota_{\mathfrak{l}_i} c_{\mathfrak{l}_i}$ in $Eq(\widehat{\mathcal{C}})$ and we conclude that $Eq(\mathcal{C}) = Eq(\widehat{\mathcal{C}})$. We thus complete the proof of (3).

If (4) is wrong, then \mathcal{C} just has one leg \mathfrak{l} . By (2.104), $k_{\mathfrak{l}} = 0$. This contradicts

with $k_{i,x} \neq 0$ in (2.89). □

Lemma 2.4.23. (1) *The output $\mathcal{C}(k+1)$ of step k of the cutting algorithm satisfies the property P .*

(2) *All the intermediate results $\mathcal{C}(k)_1, \mathcal{C}(k)_2, \mathcal{C}(k)_3$ satisfy the weak property P : either they satisfy the property P or they contain exactly one free normal leg and at least one fixed leg. Notice that the weak property implies that these couples contain at least one normal leg and two legs.*

Proof. We prove the following stronger result by induction.

Claim. Assume that the couple $\mathcal{C} = \mathcal{C}(T, T, p)$ is the input the cutting algorithm. Then for any k , there exist a finite number of trees disjoint subtrees $T_1^{(k)}, T_2^{(k)}, \dots, T_{m(k)}^{(k)}$ of the two copies of T which satisfy the following property.

(1) Let $\text{leaf}(k)$ be the set of all leaves of these subtrees. Assume that p induce a pairing $p|_{\text{leaf}_1(k)}$ of a subset $\text{leaf}_1(k) \subseteq \text{leaf}(k)$. Apply a similar construction to Definition 2.4.7 we can construct a couple from $T_1^{(k)}, T_2^{(k)}, \dots, T_{m(k)}^{(k)}$ from p and this couple equals exactly to $\mathcal{C}(k)$.

(2) The root legs of $T_1^{(k)}, T_2^{(k)}, \dots, T_{m(k)}^{(k)}$ are exactly the normal legs of $\mathcal{C}(k)$. The edges in $\text{leaf}_2(k) = \text{leaf}(k) \setminus \text{leaf}_1(k)$ are exactly the leaf legs (legs that are leaf edges) of $\mathcal{C}(k)$.

(3) $\mathcal{C}(k)$ satisfies the property P .

(4) All the intermediate results $\mathcal{C}(k)_1, \mathcal{C}(k)_2, \mathcal{C}(k)_3$ satisfy the weak property P .

We show that $\mathcal{C}(0)$ satisfies (1) – (4) of the above claim.

Let \mathcal{C} be the input couple of step 0 obtained by pairing two copies of T . In step 0, $\mathcal{C}(0)$ is obtained by replacing a fixed leg by a free leg. Therefore, if we define $T_1^{(0)} = T$ and $T_2^{(0)}$ to be the tree obtained by replacing the fixed root leg in T by a free leg, then $\mathcal{C}(0)$ is the couple obtained by pairing $T_1^{(0)}$ and $T_2^{(0)}$. Here the pairing is p and $\text{leaf}_1(0) = \text{leaf}(0)$ and $\text{leaf}_2(0) = \emptyset$. Therefore, (1) is true for $\mathcal{C}(0)$.

The two fixed legs in \mathcal{C} are all normal edges because they come from the two root legs which belong to T_{in} . Therefore, the two legs in $\mathcal{C}(1)$ are also normal. Since the two legs in $\mathcal{C}(1)$ come from the root legs in $T_1^{(0)}$ and $T_2^{(0)}$, (2) is also true.

Since in step 0, we replace a fixed leg with a free leg, $\mathcal{C}(0)$ contains exactly one free and one fixed leg which are all normal. Therefore, the output $\mathcal{C}(0)$ of step 0 satisfies property P and (3) is proved.

In step 0, (4) does not need any proof.

Assume that $\mathcal{C}(k)$ satisfies (1) – (4) of the above claim, then we prove the same for $\mathcal{C}(k+1)$.

Remember that in step k of the algorithm, the input is $\mathcal{C}(k)$. The induction assumption implies that $\mathcal{C}(k)$ is obtained by pairing $T_1^{(k)}, T_2^{(k)}, \dots, T_{m(k)}^{(k)}$. There are several different cases in step k and we treat them separately.

In case 1.1 of the cutting algorithm, we cut $c(\mathfrak{l})$ into $\{\mathfrak{e}_i^{(1)}\}_{i=1,2}$ and $\{\mathfrak{e}_i^{(2)}\}_{i=1,2}$. By (2) \mathfrak{l} is the root leg of some subtree $T_{j_0}^{(k)}$. Assume that in $T_{j_0}^{(k)}$ the two subtrees of the root are $T_{j_0,1}^{(k)}, T_{j_0,2}^{(k)}$. After cutting $c(\mathfrak{l})$, $T_{j_0}^{(k)}$ becomes two trees $T_{j_0,1}^{(k)}, T_{j_0,2}^{(k)}$ and $T_j^{(k)}$ ($j \neq j_0$) do not change. We treat three different cases separately.

Case 1.1 (i). (Both $T_{j_0,1}^{(k)}$ and $T_{j_0,2}^{(k)}$ are one node trees.) In this case all edges in $\{\mathfrak{e}_1^{(1)}, \mathfrak{e}_2^{(1)}, \mathfrak{e}_1^{(2)}, \mathfrak{e}_2^{(2)}\}$ are leaf edges of some trees in $\{T_j^{(k)}\}$. Define $T_j^{(k+1)} = T_j^{(k)}$ for $j < j_0$ and $T_j^{(k+1)} = T_{j+1}^{(k)}$ for $j > j_0$, then $\mathcal{C}(k)_1$ can be constructed from these trees with $\text{leaf}_1(k+1) = \text{leaf}_1(k) \setminus \{\mathfrak{e}_1^{(1)}, \mathfrak{e}_2^{(1)}, \mathfrak{e}_1^{(2)}, \mathfrak{e}_2^{(2)}\}$. Remember that in case 1.1 of the algorithm, the final result $\mathcal{C}(k+1)$ is obtained by replacing (or not) some free normal leg with a fixed normal leg in $\mathcal{C}(k)_1$. Therefore, if we change some free root leg of $\{T_j^{(k)}\}$ to be fixed, then $\mathcal{C}(k+1)$ can also be obtained from pairing these trees. Therefore, (1) is true for $\mathcal{C}(k+1)$ with $\text{leaf}_1(k+1) = \text{leaf}_1(k) \setminus \{\mathfrak{e}_1^{(1)}, \mathfrak{e}_2^{(1)}, \mathfrak{e}_1^{(2)}, \mathfrak{e}_2^{(2)}\}$.

Since both the sets of root legs and normal legs are deleted by one element \mathfrak{l} and in step $k-1$ they are the same, they continue to be the same in step k . In case 1.1 (i), $\text{leaf}_2(k+1) = \text{leaf}_2(k) \cup \{\mathfrak{e}_i^{(2)}\}_{i=1,2}$ and the set of leaf legs in $\mathcal{C}(k)$ also changes in

this way, so they continue to be the same in step k . Therefore, (2) is true.

We prove (4) by contradiction. The weak property P is equivalent to that $\mathcal{C}(k)_i$ contains exactly one free leg and at least one normal leg. In case 1.1 (i), $i = 1$, assume that the weak property P is not true, then either $\mathcal{C}(k)_1$ does not contain normal leg or the number of free legs is not 1. If $\mathcal{C}(k)_1 \neq \emptyset$, then the number of $\{T_j^{(k)}\}$ is not zero. Because the root of $\{T_j^{(k)}\}$ are normal legs, there exists at least one normal leg in $\mathcal{C}(k)_1$. By our hypothesis that weak property P is wrong, the number of free legs is not 1. Since $\mathcal{C}(k)_1$ is obtained from $\mathcal{C}(k)$ by cutting a fixed leg, the number of free legs in $\mathcal{C}(k)_1$ equals to $\mathcal{C}(k)$ which is one. Therefore, we find a contradiction.

Notice that in case 1.1 of the algorithm, we change the free normal leg in $\mathcal{C}(k)_1$ to be fixed if $\mathcal{C}(k)_1$ does not satisfy the property P. The result $\mathcal{C}(k+1)$ must satisfy the property P, so we prove (3).

Case 1.1 (ii). (One of $T_{j_0,1}^{(k)}$ or $T_{j_0,2}^{(k)}$ is a one node tree.) Without loss of generality assume that $T_{j_0,2}^{(k)}$ is a one node tree. In this case, $\mathbf{e}_2^{(1)}, \mathbf{e}_2^{(2)}$ are leaf edges of some trees in $\{T_j^{(k)}\}$. Define $T_j^{(k+1)} = T_j^{(k)}$ for $j \neq j_0$ and $T_{j_0}^{(k+1)} = T_{j_0,1}^{(k)}$, then $\mathcal{C}(k)_1$ can be constructed from these trees with $\text{leaf}_1(k+1) = \text{leaf}_1(k) \setminus \{\mathbf{e}_2^{(1)}, \mathbf{e}_2^{(2)}\}$. By the same reason as in case 1.1 (i), (1) is true for $\mathcal{C}(k+1)$ with $\text{leaf}_1(k+1) = \text{leaf}_1(k) \setminus \{\mathbf{e}_2^{(1)}, \mathbf{e}_2^{(2)}\}$.

By the same reason as in case 1.1 (i), the set of root legs and the set of normal legs continue to be the same in step k . In case 1.1 (ii), $\text{leaf}_2(k+1) = \text{leaf}_2(k) \cup \{\mathbf{e}_2^{(2)}\}$ and the set of leaf legs also changes in this way, so they continue to be the same in step k . Therefore, (2) is true.

The proof of (3), and (4) in case 1.1 (ii) is the same as that in case 1.1 (i).

Case 1.1 (iii). (Neither $T_{j_0,1}^{(k)}$ or $T_{j_0,2}^{(k)}$ is an one node tree.) In this case, no edge in $\{\mathbf{e}_1^{(1)}, \mathbf{e}_2^{(1)}, \mathbf{e}_1^{(2)}, \mathbf{e}_2^{(2)}\}$ is leaf edge of $T_j^{(k)}$. Define $T_j^{(k+1)} = T_j^{(k)}$ for $j \leq j_0$, $T_j^{(k+1)} = T_{j-1}^{(k)}$ for $j \geq j_0 + 2$, $T_{j_0}^{(k+1)} = T_{j_0,1}^{(k)}$ and $T_{j_0+1}^{(k+1)} = T_{j_0,2}^{(k)}$, then $\mathcal{C}(k)_1$ can be constructed from these trees with $\text{leaf}_1(k+1) = \text{leaf}_1(k)$. By the same reason as in case 1.1 (i), (1) is true for $\mathcal{C}(k+1)$ with $\text{leaf}_1(k+1) = \text{leaf}_1(k)$.

By the same reason as in case 1.1 (i), the set of root legs and the set of normal legs continue to be the same in step k . In case 1.1 (iii), $\text{leaf}_2(k+1) = \text{leaf}_2(k)$ and the set of leaf legs also changes in this way, so they continue to be the same in step k . Therefore, (2) is true.

The proof of (3), and (4) in case 1.1 (iii) is the same as that in case 1.1 (i).

In case 1.2 of the cutting algorithm, we cut \mathfrak{e} into $\{\mathfrak{e}^{(1)}, \mathfrak{e}^{(2)}\}$. By (2), \mathfrak{l} and \mathfrak{l}_{fr} are the root leg and leaf of some subtrees $T_{j_0}^{(k)}$. Assume that in $T_{j_0}^{(k)}$ one subtree of the root is $T_{j_0,1}^{(k)}$ and the other one is a one node tree with edge \mathfrak{l}_{fr} . After cutting $c(\mathfrak{l})$, $T_{j_0}^{(k)}$ becomes $T_{j_0,1}^{(k)}$ and $T_j^{(k)}$ ($j \neq j_0$) do not change. We treat two different cases separately.

Case 1.2 (i). ($T_{j_0,1}^{(k)}$ is an one node trees.) In this case $\mathfrak{e}^{(1)}, \mathfrak{e}^{(2)}$ are leaf edges of some trees in $\{T_j^{(k)}\}$. Define $T_j^{(k+1)} = T_j^{(k)}$ for $j < j_0$ and $T_j^{(k+1)} = T_{j+1}^{(k)}$ for $j > j_0$, then $\mathcal{C}(k)_1$ can be constructed from these trees with $\text{leaf}_1(k+1) = \text{leaf}_1(k) \setminus \{\mathfrak{e}^{(1)}, \mathfrak{e}^{(2)}\}$. By the same reason as in case 1.1 (i), (1) is true for $\mathcal{C}(k+1)$ with $\text{leaf}_1(k+1) = \text{leaf}_1(k) \setminus \{\mathfrak{e}^{(1)}, \mathfrak{e}^{(2)}\}$.

By the same reason as in case 1.1 (i), the set of root legs and the set of normal legs continue to be the same in step k . In case 1.2 (i), $\text{leaf}_2(k+1) = \text{leaf}_2(k) \cup \{\mathfrak{e}^{(2)}\}$ and the set of leaf legs also changes in this way, so they continue to be the same in step k . Therefore, (2) is true.

The proof of (3), and (4) in case 1.2 (i) is the same as that in case 1.1 (i).

Case 1.2 (ii). ($T_{j_0,1}^{(k)}$ is not a one node trees.) In this case $\mathfrak{e}^{(1)}, \mathfrak{e}^{(2)}$ are not leaf edges of any trees in $\{T_j^{(k)}\}$. Define $T_j^{(k+1)} = T_j^{(k)}$ for $j \neq j_0$ and $T_{j_0}^{(k+1)} = T_{j_0,1}^{(k)}$, then $\mathcal{C}(k)_1$ can be constructed from these trees with $\text{leaf}_1(k+1) = \text{leaf}_1(k)$. By the same reason as in case 1.1 (i), (1) is true for $\mathcal{C}(k+1)$ with $\text{leaf}_1(k+1) = \text{leaf}_1(k)$.

By the same reason as in case 1.1 (i), the set of root legs and the set of normal legs continue to be the same in step k . In case 1.2 (ii), $\text{leaf}_2(k+1) = \text{leaf}_2(k)$ and the set of leaf legs also changes in this way, so they continue to be the same in step

k . Therefore, (2) is true.

The proof of (3), and (4) in case 1.2 (ii) is the same as that in case 1.1 (i).

In case 2 of the cutting algorithm, we cut $c(\mathfrak{l}) = \{\mathfrak{e}_2, \mathfrak{e}_3\}$ into $\{\mathfrak{e}_2^{(1)}, \mathfrak{e}_3^{(1)}\}$ and $\{\mathfrak{e}_2^{(2)}, \mathfrak{e}_3^{(2)}\}$.

By (2), \mathfrak{l} is the root leg of some subtree $T_{j_0}^{(k)}$. Assume that in $T_{j_0}^{(k)}$ the two subtrees of the root are $T_{j_0,1}^{(k)}, T_{j_0,2}^{(k)}$. After cutting $c(\mathfrak{l})$, $T_{j_0}^{(k)}$ becomes two trees $T_{j_0,1}^{(k)}, T_{j_0,2}^{(k)}$ and $T_j^{(k)}$ ($j \neq j_0$) do not change. Since $\mathcal{C}(k)_2$ and $\mathcal{C}(k)_3$ are disjoint, for each $j \neq j_0$, $T_j^{(k)}$ should be a subset of one of these couples. Without loss of generality we assume that $j_0 = 1$ and $T_2^{(k)}, \dots, T_{m^{(k)'} }^{(k)} \subseteq \mathcal{C}(k)_2$ and $T_{m^{(k)'}+1}^{(k)}, \dots, T_{m^{(k)}}^{(k)} \subseteq \mathcal{C}(k)_3$. Since the output, $\mathcal{C}(k+1)$ either comes from $\mathcal{C}(k)_2$ or $\mathcal{C}(k)_3$, without loss of generality we assume that the output comes from $\mathcal{C}(k)_2$. We treat two different cases separately.

Case 2 (i). ($T_{j_0,1}^{(k)}$ is a one node tree.) In this case, all edges in $\{\mathfrak{e}_2^{(1)}, \mathfrak{e}_2^{(2)}\}$ are leaf edges of some trees in $\{T_j^{(k)}\}$. Define $T_j^{(k+1)} = T_{j+1}^{(k)}$, then $\mathcal{C}(k)_2$ can be constructed from $T_2^{(k)}, \dots, T_{m^{(k)'} }^{(k)}$ with $\text{leaf}_1(k+1) = \{\text{all leaves in } T_2^{(k)}, \dots, T_{m^{(k)'} }^{(k)}\}$. Remember that in case 2 of the algorithm, the final result $\mathcal{C}(k+1)$ is obtained by replacing (or not) some free normal leg with a fixed normal leg in $\mathcal{C}(k)_2$, so by the same argument as in case 1.1 (i), (1) is true for $\mathcal{C}(k+1)$.

By the same reason as in case 1.1 (i), the set of root legs and the set of normal legs continue to be the same in step k . In case 2 (i), in the output $\mathcal{C}(k+1)$ coming from $\mathcal{C}(k)_2$, $\text{leaf}_2(k+1) = (\text{leaf}_1(k+1) \cap \mathcal{C}(k)_2) \cup \{\mathfrak{e}_2^{(2)}\}$ and the set of leaf legs also changes in this way, so they continue to be the same in step k . Therefore, (2) is true.

The proof of (3), and (4) in case 2 (i) is the same as that in case 1.1 (i).

Case 2 (ii). ($T_{j_0,1}^{(k)}$ is not a one node tree.) In this case $\{\mathfrak{e}_2^{(1)}, \mathfrak{e}_2^{(2)}\}$ are not leaf edges of any trees in $\{T_j^{(k)}\}$. Define $T_j^{(k+1)} = T_j^{(k)}$ for $j > 1$ and $T_1^{(k+1)} = T_{j_0,1}^{(k)}$, then $\mathcal{C}(k)_2$ can be constructed from $T_1^{(k+1)}, \dots, T_{m^{(k)'} }^{(k+1)}$ with $\text{leaf}_1(k+1) = \{\text{all leaves in } T_1^{(k+1)}, \dots, T_{m^{(k)'} }^{(k+1)}\}$. For the same reason as in case 1.1 (i), (1) is true for $\mathcal{C}(k+1)$.

For the same reason as in case 1.1 (i), the set of root legs and the set of normal legs continue to be the same in step k . In case 2 (ii), in the output $\mathcal{C}(k+1)$ coming

from $\mathcal{C}(k)_2$, $\text{leaf}_2(k+1) = (\text{leaf}_1(k+1) \cap \mathcal{C}(k)_2)$ and the set of leaf legs also changes in this way, so they continue to be the same in step k . Therefore, (2) is true.

The proof of (3), and (4) in case 2 (i) is the same as that in case 1.1 (i). \square

Step 3. In this step, we state Proposition 2.4.24 which is a stronger version of Proposition 2.4.15 and derive Proposition 2.4.15 from it. In the end, we prove parts (1), (2) and the 1 node case of part (3) of Proposition 2.4.24.

Proposition 2.4.24. *Let $\mathcal{C}(k)$ be a couple which is the output of step k of the cutting algorithm Proposition 2.4.19. For any couple \mathcal{C} , let $n(\mathcal{C})$ be the total number of nodes in \mathcal{C} and $n_e(\mathcal{C})$ (resp. $n_{fx}(\mathcal{C})$, $n_{fr}(\mathcal{C})$) be the total number of non-leg edges (resp. fixed legs, free legs). We fix $\sigma_{\mathbf{n}} \in \mathbb{R}$ for each $\mathbf{n} \in \mathcal{C}(k)$ and $c_{\mathbf{l}} \in \mathbb{R}$ for each fixed leg \mathbf{l} . Assume that α satisfies (2.2). Then we have*

(1) *We have following relation of $n(\mathcal{C})$, $n_e(\mathcal{C})$, $n_{fx}(\mathcal{C})$ and $n_{fr}(\mathcal{C})$*

$$2n_e(\mathcal{C}) + n_{fx}(\mathcal{C}) + n_{fr}(\mathcal{C}) = 3n(\mathcal{C}) \quad (2.111)$$

(2) *For any couple \mathcal{C} , define $\chi(\mathcal{C}) = n_e(\mathcal{C}) + n_{fr}(\mathcal{C}) - n(\mathcal{C})$. Let c , \mathcal{C}_1 , \mathcal{C}_2 be the same as in Lemma 2.4.18 and also assume that $\{\mathbf{e}_i^{(1)}\}$ are free legs and $\{\mathbf{e}_i^{(2)}\}$ are fixed legs. Then*

$$\chi(\mathcal{C}) = \chi(\mathcal{C}_1) + \chi(\mathcal{C}_2). \quad (2.112)$$

(3) *If we assume that $\mathcal{C}(k)$ satisfies the property P , then (recall that $Q = L^d T_{max}^{-1}$)*

$$\sup_{\{c_{\mathbf{l}}\}_{\mathbf{l}}} \#Eq(\mathcal{C}(k), \{c_{\mathbf{l}}\}_{\mathbf{l}}) \leq L^{O(\chi(\mathcal{C}(k))\theta)} Q^{\chi(\mathcal{C}(k))} \prod_{\mathbf{e} \in \mathcal{C}_{norm}(k)} \kappa_{\mathbf{e}}^{-1}. \quad (2.113)$$

Derivation of Proposition 2.4.15 from Proposition 2.4.24: By Proposition 2.4.19 (2),

$\mathcal{C}_{\text{norm}}(1) = \mathcal{C}_{\text{norm}}$ and $\#Eq(\mathcal{C}) = \#Eq(\mathcal{C}(1))$. Then by Proposition 2.4.24 (3),

$$\#Eq(\mathcal{C}) = \#Eq(\mathcal{C}(1)) \leq L^{O(\chi(\mathcal{C}(1))\theta)} Q^{\chi(\mathcal{C}(1))} \prod_{\mathfrak{e} \in \mathcal{C}_{\text{norm}}(1)} \kappa_{\mathfrak{e}}^{-1} \quad (2.114)$$

Since in $\mathcal{C}(1)$, $n_{fx}(\mathcal{C}(1)) = n_{fr}(\mathcal{C}(1)) = 1$, by Proposition 2.4.24 (1), $2n_e(\mathcal{C}(1)) + 2 = 3n(\mathcal{C}(1))$. Using this fact and the definition of χ , we get $\chi(\mathcal{C}(1)) = n_e(\mathcal{C}(1)) + 1 - n(\mathcal{C}(1)) = n(\mathcal{C}(1))/2$. Substituting this expression of $\chi(\mathcal{C}(1))$ into (2.114) proves the conclusion of Proposition 2.4.15.

The proof of Proposition 2.4.24 is the main goal of the rest of the proof.

Proof Proposition 2.4.24 (1): Consider the set $\mathcal{S} = \{(\mathbf{n}, \mathfrak{e}) \in \mathcal{C} : \mathbf{n} \text{ is an end point of } \mathfrak{e}\}$, then

$$\#\mathcal{S} = \sum_{\substack{(\mathbf{n}, \mathfrak{e}) \in \mathcal{C} \\ \mathbf{n} \text{ is an end point of } \mathfrak{e}}} 1 \quad (2.115)$$

First sum over \mathfrak{e} and then over \mathbf{n} , we get

$$\#\mathcal{S} = \sum_{\mathbf{n}} \sum_{\substack{\mathfrak{e} \in \mathcal{C} \\ \mathbf{n} \text{ is an end point of } \mathfrak{e}}} 1 = \sum_{\mathbf{n}} 3 = 3n(\mathcal{C}). \quad (2.116)$$

In the second equality, $\sum_{\mathfrak{e} \in \mathcal{C}: \mathbf{n} \text{ is an end point of } \mathfrak{e}} 1 = 3$ because for each node \mathbf{n} there are 3 edges connected to it.

Switch the order of summation, we get

$$\begin{aligned} \#\mathcal{S} &= \sum_{\mathfrak{e} \text{ is a non-leg edges}} \sum_{\substack{\mathbf{n} \in \mathcal{C} \\ \mathbf{n} \text{ is an endpoint of } \mathfrak{e}}} 1 + \sum_{\mathfrak{e} \text{ is a leg}} \sum_{\substack{\mathbf{n} \in \mathcal{C} \\ \mathbf{n} \text{ is an endpoint of } \mathfrak{e}}} 1 \\ &= \sum_{\mathfrak{e} \text{ is a non-leg edges}} 2 + \sum_{\mathfrak{e} \text{ is a leg}} 1 \\ &= 2n_e(\mathcal{C}) + n_{fx}(\mathcal{C}) + n_{fr}(\mathcal{C}) \end{aligned} \quad (2.117)$$

In the second equality, $\sum_{\substack{\mathbf{n} \in \mathcal{C} \\ \mathbf{n} \text{ is an endpoint of } \mathfrak{e}}} 1$ equals to 1 or 2 because for each non-leg edge (resp. leg) there are 2 (resp. 1) nodes connected to it.

Because the value of $\#\mathcal{S}$ does not depend on the order of summation, we conclude that $2n_e(\mathcal{C}) + n_{fx}(\mathcal{C}) + n_{fr}(\mathcal{C}) = 3n(\mathcal{C})$, which proves Proposition 2.4.24 (1).

Proof Proposition 2.4.24 (2): Since cutting does change the number of nodes, we have $n(\mathcal{C}) = n(\mathcal{C}_1) + n(\mathcal{C}_2)$. Let c be the cut that consists of edges $\{\mathfrak{e}_i\}$ and $n(c)$ be the number of edges in c . When cutting \mathcal{C} into \mathcal{C}_1 and \mathcal{C}_2 , $n(c)$ non-leg edges are cut into pairs of free and fixed legs, so $n_e(\mathcal{C}) = n(\mathcal{C}_1) + n(\mathcal{C}_2) + n(c)$. Because we have $n(c)$ additional free legs after cutting, so $n_{fr}(\mathcal{C}) = n_{fr}(\mathcal{C}_1) + n_{fr}(\mathcal{C}_2) - n(c)$. Therefore, we get

$$\begin{aligned}
\chi(\mathcal{C}) &= n_e(\mathcal{C}) + n_{fr}(\mathcal{C}) - n(\mathcal{C}) \\
&= n(\mathcal{C}_1) + n(\mathcal{C}_2) + n(c) + n_{fr}(\mathcal{C}_1) + n_{fr}(\mathcal{C}_2) - n(c) - (n(\mathcal{C}_1) + n(\mathcal{C}_2)) \\
&= (n(\mathcal{C}_1) + n_{fr}(\mathcal{C}_1) - n(\mathcal{C}_1)) + (n(\mathcal{C}_2) + n_{fr}(\mathcal{C}_2) - n(\mathcal{C}_2)) \\
&= \chi(\mathcal{C}_1) + \chi(\mathcal{C}_2).
\end{aligned} \tag{2.118}$$

We complete the proof of (2) of Proposition 2.4.24.

Proof of 1 node case of Proposition 2.4.24 (3): We prove the following Lemma 2.4.25 which is the 1 node case of Proposition 2.4.24 (3). Recall that \mathcal{C}_I and \mathcal{C}_{II} are the two possibilities of the 1 node couple \mathcal{C}_I in Proposition 2.4.19 (3).

Lemma 2.4.25. *$\mathcal{C}_I, \mathcal{C}_{II}$ satisfy the bound (2.113) in Proposition 2.4.24. In other words, fix c_1 (resp. c_1, c_2) for the fixed legs of \mathcal{C}_I (resp. \mathcal{C}_{II}), then we have*

$$\#Eq(\mathcal{C}_I) \leq L^\theta Q \kappa_\epsilon^{-1}, \quad \#Eq(\mathcal{C}_{II}) \leq Q^0 = 1. \tag{2.119}$$

Proof. Given c_1, c_2 , the equation of \mathcal{C}_{II} is

$$\begin{cases} k_{\epsilon_1} + k_{\epsilon_2} - k_\epsilon = 0, \quad k_{\epsilon_1} = c_1, \quad k_\epsilon = c_2, \quad |k_{\epsilon x}| \sim \kappa_\epsilon \\ \Lambda_{k_{\epsilon_1}} + \Lambda_{k_{\epsilon_2}} - \Lambda_{k_\epsilon} = \sigma_n + O(T_{\max}^{-1}) \end{cases} \tag{2.120}$$

It's obvious that there is at most one solution to this system of equations.

Given c_1 , the equation of \mathcal{C}_I is

$$\begin{cases} k_{\epsilon_1} + k_{\epsilon_2} - k_{\epsilon} = 0, & k_{\epsilon} = c_1, & |k_{\epsilon x}| \sim \kappa_{\epsilon} \\ \Lambda_{k_{\epsilon_1}} + \Lambda_{k_{\epsilon_2}} - \Lambda_{k_{\epsilon}} = \sigma_n + O(T_{\max}^{-1}) \end{cases} \quad (2.121)$$

By Theorem 4.1.1, the number of solutions of the above system of equations can be bounded by $L^{\theta} L^d T_{\max}^{-1} |k_{\epsilon x}|^{-1} \lesssim L^{\theta} Q \kappa_{\epsilon}^{-1}$

Therefore, we complete the proof of this lemma. \square

Step 4. In this step, we apply the edge cutting algorithm Proposition 2.4.19 to prove Proposition 2.4.24 (3) by induction.

If \mathcal{C} has only one node ($n = 1$), then \mathcal{C} equals to \mathcal{C}_I or \mathcal{C}_{II} and Proposition 2.4.24 (3) in this case follows from Lemma 2.4.25.

Suppose that Proposition 2.4.24 (3) holds true for couples with the number of nodes $\leq n - 1$. We prove it for couples with the number of nodes n .

Given the couple $\mathcal{C}(k)$ which is the output of the $k - 1$ -th step, apply the cutting algorithm to $\mathcal{C}(k)$. Then according to Proposition 2.4.19, $\mathcal{C}(k)$ (2) is decomposed into 2 or 3 components.

In the first case, denote by $\mathcal{C}(k)_I$ and $\mathcal{C}(k)_1$ the two components after cutting. In the second case, denote by $\mathcal{C}(k)_I$, $\mathcal{C}(k)_2$ and $\mathcal{C}(k)_3$ the three components after cutting. By Proposition 2.4.19 (4), $\mathcal{C}(k)_1$, $\mathcal{C}(k)_2$, $\mathcal{C}(k)_3$ are couples of nodes $\leq n - 1$ that satisfy the property P and thus satisfy the assumption of Proposition 2.4.24 (3). Therefore, the induction assumption is applicable and Proposition 2.4.24 (3) is true for these three couples.

Case 1. Assume that there are two components after cutting. One example of this case is the left couple in Figure 2.16.

Then applying Lemma 2.4.18 gives

$$\begin{aligned}
\sup_{\{c_l\}_l} \#Eq(\mathcal{C}(k), \{c_l\}_l) &\leq \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}(k)_l)}} \#Eq(\mathcal{C}(k)_l, \{c_{l_1}\}) \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}(k)_1)}} \#Eq(\mathcal{C}(k)_1, \{c_{l_2}\}) \\
&\lesssim L^\theta Q^{\chi(\mathcal{C}_1)} L^{O(\chi(\mathcal{C}_2)\theta)} Q^{\chi(\mathcal{C}_2)} = L^{O((\chi(\mathcal{C}_1)+\chi(\mathcal{C}_2))\theta)} Q^{\chi(\mathcal{C}_1)+\chi(\mathcal{C}_2)} \\
&= L^{O(\chi(\mathcal{C})\theta)} Q^{\chi(\mathcal{C})}.
\end{aligned} \tag{2.122}$$

Here the second inequality follows from the induction assumption and Lemma 2.4.25. The last equality follows from Proposition 2.4.24 (2).

Case 2. Assume that there are three components after cutting. One example of this case is the right couple in Figure 2.16.

Then applying Lemma 2.4.18 gives

$$\begin{aligned}
\sup_{\{c_l\}_l} \#Eq(\mathcal{C}(k), \{c_l\}_l) &\leq \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}(k) \setminus \mathcal{C}(k)_2)}} \#Eq(\mathcal{C}(k) \setminus \mathcal{C}(k)_2, \{c_{l_1}\}) \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}(k)_2)}} \#Eq(\mathcal{C}(k)_2, \{c_{l_2}\}) \\
&\lesssim L^{O(\chi(\mathcal{C}(k)_2)\theta)} Q^{\chi(\mathcal{C}(k)_2)} \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}(k) \setminus \mathcal{C}(k)_2)}} \#Eq(\mathcal{C}(k) \setminus \mathcal{C}(k)_2, \{c_{l_1}\}).
\end{aligned} \tag{2.123}$$

Here in the second inequality, we apply the induction assumption to $\mathcal{C}(k)_2$.

Applying Lemma 2.4.18 and (2.123) gives

$$\begin{aligned}
\sup_{\{c_l\}_l} \#Eq(\mathcal{C}(k), \{c_l\}_l) &\lesssim L^{O(\mathcal{C}(k)_2\theta)} Q^{\chi(\mathcal{C}(k)_2)} \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}(k) \setminus \mathcal{C}(k)_2)}} \#Eq(\mathcal{C}(k) \setminus \mathcal{C}(k)_2, \{c_{l_1}\}) \\
&\lesssim L^{O(\chi(\mathcal{C}(k)_2)\theta)} Q^{\chi(\mathcal{C}(k)_2)} \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}(k)_l)}} \#Eq(\mathcal{C}(k)_l, \{c_{l_1}\}) \sup_{\{c_{l_3}\}_{l_3 \in \text{leg}(\mathcal{C}(k)_3)}} \#Eq(\mathcal{C}(k)_3, \{c_{l_3}\}) \\
&\lesssim L^{O((\chi(\mathcal{C}(k)_2)+\chi(\mathcal{C}(k)_3))\theta)} Q^{\chi(\mathcal{C}(k)_2)+\chi(\mathcal{C}(k)_3)} \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}(k)_l)}} \#Eq(\mathcal{C}(k)_l, \{c_{l_1}\}) \\
&\lesssim (L^{O(\theta)} Q)^{\chi(\mathcal{C}(k)_2)+\chi(\mathcal{C}(k)_3)+\chi(\mathcal{C}(k)_l)} = L^{O(\chi(\mathcal{C})\theta)} Q^{\chi(\mathcal{C})}.
\end{aligned} \tag{2.124}$$

Here in the third inequality, we apply the induction assumption to $\mathcal{C}(k)_3$. The last equality follows from Proposition 2.4.24 (2).

Therefore, we complete the proof of Proposition 2.4.24 and thus the proof of Proposition 2.4.15. \square

2.4.5 An upper bound of tree terms

In this section, we first prove the following Proposition which gives an upper bound of the variance of $\mathcal{J}_{T,k}$ and then prove Proposition 2.2.7 as its corollary.

Proposition 2.4.26. *Assume that α satisfies (2.2) and $\rho = \alpha T_{\max}^{\frac{1}{2}}$. For any $\theta > 0$, we have*

$$\sup_k \mathbb{E} |(\mathcal{J}_T)_k|^2 \lesssim L^{O(l(T)\theta)} \rho^{2l(T)}. \quad (2.125)$$

and $\mathbb{E} |(\mathcal{J}_T)_k|^2 = 0$ if $|k| \gtrsim 1$.

Proof. By (2.86) and (2.87), we know that $\mathcal{J}_{T,k}$ is a linear combination of $Term(T, p)_k$.

$$\mathbb{E} |\mathcal{J}_{T,k}|^2 = \left(\frac{\lambda}{L^d} \right)^{2l(T)} \sum_{p \in \mathcal{P}(\{k_1, \dots, k_{l(T)+1}, k'_1, \dots, k'_{l(T)+1}\})} Term(T, p)_k. \quad (2.126)$$

Since $\alpha = \frac{\lambda}{L^{\frac{d}{2}}}$, $\frac{\lambda}{L^d} = \alpha L^{-\frac{d}{2}}$. Since the number of elements in \mathcal{P} can be bounded by a constant, by Lemma 2.4.27 proved below, we get

$$\mathbb{E} |\mathcal{J}_{T,k}|^2 \lesssim (\alpha L^{-\frac{d}{2}})^{2l(T)} L^{O(n\theta)} Q^{\frac{n}{2}} T_{\max}^n. \quad (2.127)$$

By definition, n is the total number of \bullet nodes in the couple constructed from tree T and pairing p . Therefore, n equals to $2l(T)$. Replacing n by $2l(T)$ and Q by

$L^d T_{\max}^{-1}$ in (2.127), we get

$$\begin{aligned}
\mathbb{E}|\mathcal{J}_{T,k}|^2 &\lesssim (\alpha L^{-\frac{d}{2}})^{2l(T)} L^{O(l(T)\theta)} (L^d T_{\max}^{-1})^{l(T)} T_{\max}^{2l(T)} \\
&= L^{O(l(T)\theta)} (\alpha^2 T_{\max})^{l(T)} \\
&= L^{O(l(T)\theta)} \rho^{2l(T)}.
\end{aligned} \tag{2.128}$$

By Lemma 2.4.27 below $Term(T, p)_k = 0$ if $|k| \gtrsim 1$, we know that the same is true for $\mathbb{E}|(\mathcal{J}_T)_k|^2 = 0$.

Therefore, we complete the proof of this proposition. \square

Lemma 2.4.27. *Let n and $Q = L^d T_{\max}^{-1}$ be the same as in Proposition 2.4.15. Assume that α satisfies (2.2) and $n_{\text{in}} \in C_0^\infty(\mathbb{R}^d)$ is compactly supported. Let \mathcal{C} be the couple constructed from tree T and pairing p . Then for any $\theta > 0$, we have*

$$\sup_k |Term(T, p)_k| \leq L^{O(n\theta)} Q^{\frac{n}{2}} T_{\max}^n. \tag{2.129}$$

and $Term(T, p)_k = 0$ if $|k| \gtrsim 1$.

Proof. By (2.86), we get

$$\begin{aligned}
Term(T, p)_k &= \sum_{k_1, k_2, \dots, k_{l(T)+1}} \sum_{k'_1, k'_2, \dots, k'_{l(T)+1}} H_{k_1 \dots k_{l(T)+1}}^T H_{k'_1 \dots k'_{l(T)+1}}^T \\
&\quad \delta_p(k_1, \dots, k_{l(T)+1}, k'_1, \dots, k'_{l(T)+1}) \sqrt{n_{\text{in}}(k_1)} \cdots \sqrt{n_{\text{in}}(k'_{l(T)+1})} \cdots
\end{aligned} \tag{2.130}$$

Since n_{in} are compactly supported and there are bounded many of them in $Term(T, p)_k$, by $k_1 + k_2 + \dots + k_{l(T)+1} = k$, we know that $Term(T, p)_k = 0$ if $|k| \gtrsim 1$.

By (2.69), we get

$$|H_{k_1 \dots k_{l+1}}^T| \lesssim \sum_{\{d_n\}_{n \in T_{\text{in}}} \in \{0,1\}^{l(T)}} \prod_{n \in T_{\text{in}}} \frac{1}{|q_n| + T_{\max}^{-1}} \prod_{\epsilon \in T_{\text{in}}} |k_{\epsilon, x}| \delta_{\cap_{n \in T_{\text{in}}} \{S_n=0\}}. \tag{2.131}$$

By (2.70), q_n is a linear combination of Ω_n , so there exist constants $c_{n,\tilde{n}}$ such that $q_n = \sum_{\tilde{n}} c_{n,\tilde{n}} \Omega_{\tilde{n}}$. Let c be the matrix $[c_{n,\tilde{n}}]$ and $\mathcal{M}(T)$ be the set of all possible such matrices, then the number of elements in $\mathcal{M}(T)$ can be bounded by a constant. Let $c(\Omega)$ be the vector $\{\sum_{\tilde{n}} c_{n,\tilde{n}} \Omega_{\tilde{n}}\}_n$ and $c(\Omega)_n = \sum_{\tilde{n}} c_{n,\tilde{n}} \Omega_{\tilde{n}}$ be the components of $c(\Omega)$.

With this notation, we know that $q_n = c(\Omega)_n$ and the right hand side of (2.131) becomes $\sum_{c \in \mathcal{M}(T)} \prod_{n \in T_{\text{in}}} \frac{1}{|c(\Omega)_n| + T_{\text{max}}^{-1}}$. Therefore, using the fact that n_{in} are compactly supported, we have

$$\begin{aligned} |Term(T, p)_k| &\lesssim \sum_{\substack{k_1, \dots, k_{l(T)+1}, k'_1, \dots, k'_{l(T)+1} \\ |k_j|, |k'_j| \lesssim 1, \forall j}} \sum_{c \in \mathcal{M}(T)} \prod_{n \in T_{\text{in}}} \frac{1}{|c(\Omega)_n| + T_{\text{max}}^{-1}} \prod_{\epsilon \in T_{\text{in}}} |k_{\epsilon, x}| \delta_{\cap_{n \in T_{\text{in}}} \{S_n=0\}} \\ &\sum_{c' \in \mathcal{M}(T)} \prod_{n' \in T_{\text{in}}} \frac{1}{|c'(\Omega)_{n'}| + T_{\text{max}}^{-1}} \prod_{\epsilon' \in T_{\text{in}}} |k_{\epsilon', x}| \delta_{\cap_{n' \in T_{\text{in}}} \{S_{n'}=0\}} \delta_p(k_1, \dots, k_{l(T)+1}, k'_1, \dots, k'_{l(T)+1}) \end{aligned} \quad (2.132)$$

Switch the order of summations and products in (2.132), then we get

$$\begin{aligned} |Term(T, p)_k| &\lesssim \sum_{\substack{k_1, \dots, k_{l(T)+1}, k'_1, \dots, k'_{l(T)+1} \\ |k_j|, |k'_j| \lesssim 1, \forall j}} \sum_{c, c' \in \mathcal{M}(T)} \prod_{n, n' \in T_{\text{in}}} \frac{1}{|c(\Omega)_n| + T_{\text{max}}^{-1}} \\ &\frac{1}{|c'(\Omega)_{n'}| + T_{\text{max}}^{-1}} \prod_{\epsilon, \epsilon' \in T_{\text{in}}} (|k_{\epsilon, x}| |k_{\epsilon', x}|) \delta_{\cap_{n, n' \in T_{\text{in}}} \{S_n=0, S_{n'}=0\}} \delta_p(k_1, \dots, k_{l(T)+1}, k'_1, \dots, k'_{l(T)+1}). \end{aligned} \quad (2.133)$$

Given a tree T and pairing p , we can construct a couple \mathcal{C} . We show that

$$\sum_{c, c' \in \mathcal{M}(T)} = \sum_{c \in \mathcal{M}(\mathcal{C})}, \quad \prod_{n, n' \in T_{\text{in}}} = \prod_{n \in \mathcal{C}}, \quad \prod_{\epsilon, \epsilon' \in T_{\text{in}}} = \prod_{\epsilon \in \mathcal{C}_{\text{norm}}}, \quad \cap_{n, n' \in T_{\text{in}}} = \cap_{n \in \mathcal{C}}. \quad (2.134)$$

Remember that \mathcal{C} is constructed by gluing two copies of T by p . In (2.133), n , ϵ , n' , ϵ' denote nodes and edges in the first or second copy respectively. Since all nodes in \mathcal{C} come from the two copies of T , we get $\prod_{n, n' \in T_{\text{in}}} = \prod_{n \in \mathcal{C}}$ and $\cap_{n, n' \in T_{\text{in}}} = \cap_{n \in \mathcal{C}}$.

Remember that T_{in} is the tree formed by all non-leaf nodes, so edges of $\mathcal{C}_{\text{norm}}$ all come from the two copies of T_{in} . Therefore $\prod_{\mathfrak{e}, \mathfrak{e}' \in T_{\text{in}}} = \prod_{\mathfrak{e} \in \mathcal{C}_{\text{norm}}}$. Given two matrices c , and c' , we can construct a new matrix $c \oplus c'$ as the following. Consider two vectors $\Omega = \{\Omega_{\mathfrak{n}}\}_{\mathfrak{n}' \in T}$, $\Omega' = \{\Omega_{\mathfrak{n}'}\}_{\mathfrak{n}' \in T}$, then define $\Omega \oplus \Omega' := \{\Omega_{\mathfrak{n}}, \Omega_{\mathfrak{n}'}\}_{\mathfrak{n}, \mathfrak{n}' \in T}$. We know that $c(\Omega) = \sum_{\tilde{\mathfrak{n}}} c_{\mathfrak{n}, \tilde{\mathfrak{n}}} \Omega_{\tilde{\mathfrak{n}}}$ and $c(\Omega') = \sum_{\tilde{\mathfrak{n}'}} c_{\mathfrak{n}', \tilde{\mathfrak{n}'}} \Omega_{\tilde{\mathfrak{n}'}}$. Define $c \oplus c'$ to be the linear map whose domain is all vector of the form $\Omega \oplus \Omega'$ and whose action is $c \oplus c'(\Omega \oplus \Omega') = \{\sum_{\tilde{\mathfrak{n}}} c_{\mathfrak{n}, \tilde{\mathfrak{n}}} \Omega_{\tilde{\mathfrak{n}}}, \sum_{\tilde{\mathfrak{n}'}} c_{\mathfrak{n}', \tilde{\mathfrak{n}'}} \Omega_{\tilde{\mathfrak{n}'}}\}_{\mathfrak{n}, \mathfrak{n}' \in T}$. Define $\mathcal{M}(\mathcal{C}) = \{c \oplus c' : c, c' \in \mathcal{M}(T)\}$, then we get $\sum_{c, c' \in \mathcal{M}(T)} = \sum_{c \in \mathcal{M}(\mathcal{C})}$.

By (2.134) and the fact that leaves corresponding to k_j, k'_j are merged in \mathcal{C} , (2.133) is equivalent to

$$|Term(T, p)_k| \lesssim \sum_{\substack{k_1, \dots, k_{l(T)+1}, k'_1, \dots, k'_{l(T)+1} \\ |k_j|, |k'_j| \lesssim 1, \forall j}} \sum_{c \in \mathcal{M}(\mathcal{C})} \prod_{\mathfrak{n} \in \mathcal{C}} \frac{1}{|c(\Omega)_{\mathfrak{n}}| + T_{\max}^{-1}} \prod_{\mathfrak{e} \in \mathcal{C}_{\text{norm}}} |k_{\mathfrak{e}, x}| \delta_{\cap_{\mathfrak{n} \in \mathcal{C}} \{S_{\mathfrak{n}}=0\}} \quad (2.135)$$

Assigning a number $\sigma_{\mathfrak{n}} \in \mathbb{Z}_{T_{\max}}$ for each node $\mathfrak{n} \in \mathcal{C}$, a number $\kappa_{\mathfrak{e}} \in \mathcal{D}(\alpha, 1) := \{2^{-K_{\mathfrak{e}}} : K_{\mathfrak{e}} \in \mathbb{Z} \cap [0, \ln \alpha^{-1}]\}$ for each edge \mathfrak{e} and a number $k \in \mathbb{Z}_L^d$ for each fixed leg, we can define the associated equation $Eq(\mathcal{C}, \{\sigma_{\mathfrak{n}}\}_{\mathfrak{n}}, \{\kappa_{\mathfrak{e}}\}_{\mathfrak{e}}, k) = Eq(\mathcal{C})$ as in (2.92). Then we have

$$\sum_{\substack{k_1, \dots, k_{l(T)+1}, k'_1, \dots, k'_{l(T)+1} \\ \cap_{\mathfrak{n} \in \mathcal{C}} \{S_{\mathfrak{n}}=0\}}} = \sum_{\kappa_{\mathfrak{e}} \in \mathcal{D}(\alpha, 1)} \sum_{\sigma_{\mathfrak{n}} \in \mathbb{Z}_{T_{\max}}} \sum_{Eq(\mathcal{C}, \{\sigma_{\mathfrak{n}}\}_{\mathfrak{n}}, \{\kappa_{\mathfrak{e}}\}_{\mathfrak{e}}, k)} , \quad (2.136)$$

which implies that

$$|Term(T, p)_k| \lesssim \sum_{\kappa_{\mathfrak{e}} \in \mathcal{D}(\alpha, 1)} \sum_{\sigma_{\mathfrak{n}} \in \mathbb{Z}_{T_{\max}}} \sum_{Eq(\mathcal{C}, \{\sigma_{\mathfrak{n}}\}_{\mathfrak{n}}, \{\kappa_{\mathfrak{e}}\}_{\mathfrak{e}}, k)} \sum_{c \in \mathcal{M}(\mathcal{C})} \prod_{\mathfrak{n} \in \mathcal{C}} \frac{1}{|c(\Omega)_{\mathfrak{n}}| + T_{\max}^{-1}} \prod_{\mathfrak{e} \in \mathcal{C}_{\text{norm}}} |k_{\mathfrak{e}, x}| \quad (2.137)$$

Remember in the definition of $Eq(\mathcal{C})$, we have conditions that $|k_{\mathfrak{e}}| \lesssim 1$. These conditions come from the fact that on the support of $\delta_{\cap_{\mathfrak{n} \in T_{\text{in}}} \{S_{\mathfrak{n}}=0\}}$, if $k_1, \dots, k_{l(T)+1}$ are

bounded, then $|k_{\mathfrak{e}}|$ are also bounded. Recall that by (2.58), $S_{\mathfrak{n}} = \iota_{\mathfrak{e}_1} k_{\mathfrak{e}_1} + \iota_{\mathfrak{e}_2} k_{\mathfrak{e}_2} + \iota_{\mathfrak{e}} k_{\mathfrak{e}}$. The conditions that $S_{\mathfrak{n}} = 0$ imply that the variables $k_{\mathfrak{e}}$ of the parent edge \mathfrak{e} are linear combinations of the variables $k_{\mathfrak{e}_1}, k_{\mathfrak{e}_2}$ of children edges $\mathfrak{e}_1, \mathfrak{e}_2$. The variables of children edge \mathfrak{e}_j is again a linear combination of the variables of their children. We may iterate this argument to show that all variables $k_{\mathfrak{e}}$ are linear combinations of variables of leaves $k_1, \dots, k_{l(T)+1}$. Because $k_1, \dots, k_{l(T)+1}$ are bounded, then their linear combinations $k_{\mathfrak{e}}$ are also bounded.

In the definition of $Eq(\mathcal{C}, \{\sigma_{\mathfrak{n}}\}_{\mathfrak{n}}, \{\kappa_{\mathfrak{e}}\}_{\mathfrak{e}}, k)$, $|\Omega_{\mathfrak{n}} - \sigma_{\mathfrak{n}}| = O(T_{\max}^{-1})$ and $|k_{\mathfrak{e}x}| \sim \kappa_{\mathfrak{e}}$. Denote the constant in $O(T_{\max}^{-1})$ by δ and then we have $|\Omega_{\mathfrak{n}} - \sigma_{\mathfrak{n}}| \leq \delta T_{\max}^{-1}$. Therefore, we have $|c(\Omega)_{\mathfrak{n}} - c(\{\sigma_{\mathfrak{n}}\})_{\mathfrak{n}}| \lesssim \delta T_{\max}^{-1}$. We have the freedom of choosing δ in the definition and we take it sufficiently small so that $|c(\Omega)_{\mathfrak{n}} - c(\{\sigma_{\mathfrak{n}}\})_{\mathfrak{n}}| \leq \frac{1}{2} T_{\max}^{-1}$. This implies that $|c(\Omega)_{\mathfrak{n}}| + T_{\max}^{-1} \gtrsim |c(\{\sigma_{\mathfrak{n}}\})_{\mathfrak{n}}| + T_{\max}^{-1}$. Since we also have $|k_{\mathfrak{e}x}| \sim \kappa_{\mathfrak{e}}$, by (2.137) we get

$$\begin{aligned}
& Term(T, p)_k \\
& \lesssim \sum_{\kappa_{\mathfrak{e}} \in \mathcal{D}(\alpha, 1)} \sum_{\sigma_{\mathfrak{n}} \in \mathbb{Z}_{T_{\max}}} \sum_{Eq(\mathcal{C}, \{\sigma_{\mathfrak{n}}\}_{\mathfrak{n}}, \{\kappa_{\mathfrak{e}}\}_{\mathfrak{e}}, k)} \sum_{c \in \mathcal{M}(\mathcal{C})} \prod_{\mathfrak{n} \in \mathcal{C}} \frac{1}{|c(\Omega)_{\mathfrak{n}}| + T_{\max}^{-1}} \prod_{\mathfrak{e} \in \mathcal{C}_{\text{norm}}} |k_{\mathfrak{e}, x}| \\
& \lesssim \sum_{c \in \mathcal{M}(\mathcal{C})} \sum_{\kappa_{\mathfrak{e}} \in \mathcal{D}(\alpha, 1)} \sum_{\substack{\sigma_{\mathfrak{n}} \in \mathbb{Z}_{T_{\max}} \\ |\sigma_{\mathfrak{n}}| \lesssim 1}} \prod_{\mathfrak{n} \in \mathcal{C}} \frac{1}{|c(\{\sigma_{\mathfrak{n}}\})_{\mathfrak{n}}| + T_{\max}^{-1}} \sum_{Eq(\mathcal{C}, \{\sigma_{\mathfrak{n}}\}_{\mathfrak{n}}, \{\kappa_{\mathfrak{e}}\}_{\mathfrak{e}}, k)} \prod_{\mathfrak{e} \in \mathcal{C}_{\text{norm}}} |k_{\mathfrak{e}, x}| \\
& \lesssim \sum_{c \in \mathcal{M}(\mathcal{C})} \sum_{\kappa_{\mathfrak{e}} \in \mathcal{D}(\alpha, 1)} \sum_{\substack{\sigma_{\mathfrak{n}} \in \mathbb{Z}_{T_{\max}} \\ |\sigma_{\mathfrak{n}}| \lesssim 1}} \prod_{\mathfrak{n} \in \mathcal{C}} \frac{1}{|c(\{\sigma_{\mathfrak{n}}\})_{\mathfrak{n}}| + T_{\max}^{-1}} \left(\prod_{\mathfrak{e} \in \mathcal{C}_{\text{norm}}} \kappa_{\mathfrak{e}} \right) \# Eq(\mathcal{C}, \{\sigma_{\mathfrak{n}}\}_{\mathfrak{n}}, \{\kappa_{\mathfrak{e}}\}_{\mathfrak{e}}, k) \\
& \lesssim \sum_{c \in \mathcal{M}(\mathcal{C})} \sum_{\kappa_{\mathfrak{e}} \in \mathcal{D}(\alpha, 1)} \sum_{\substack{\sigma_{\mathfrak{n}} \in \mathbb{Z}_{T_{\max}} \\ |\sigma_{\mathfrak{n}}| \lesssim 1}} \prod_{\mathfrak{n} \in \mathcal{C}} \frac{1}{|c(\{\sigma_{\mathfrak{n}}\})_{\mathfrak{n}}| + T_{\max}^{-1}} \left(\prod_{\mathfrak{e} \in \mathcal{C}_{\text{norm}}} \kappa_{\mathfrak{e}} \right) L^{O(n\theta)} Q^{\frac{n}{2}} \prod_{\mathfrak{e} \in \mathcal{C}_{\text{norm}}} \kappa_{\mathfrak{e}}^{-1}
\end{aligned} \tag{2.138}$$

Here in the last inequality, we applied (2.97) in Proposition 2.4.15.

After simplification, (2.138) gives us

$$\begin{aligned}
|Term(T, p)_k| &\lesssim L^{O(n\theta)} Q^{\frac{n}{2}} \sum_{c \in \mathcal{M}(\mathcal{C})} \sum_{\kappa_{\mathfrak{e}} \in \mathcal{D}(\alpha, 1)} \sum_{\substack{\sigma_{\mathbf{n}} \in \mathbb{Z}_{T_{\max}} \\ |\sigma_{\mathbf{n}}| \lesssim 1}} \prod_{\mathbf{n} \in \mathcal{C}} \frac{1}{|c(\{\sigma_{\mathbf{n}}\})_{\mathbf{n}}| + T_{\max}^{-1}} \\
&\lesssim L^{O(n\theta)} Q^{\frac{n}{2}} \left(\sum_{\kappa_{\mathfrak{e}} \in \mathcal{D}(\alpha, 1)} 1 \right) \sum_{c \in \mathcal{M}(\mathcal{C})} \sum_{\substack{\sigma_{\mathbf{n}} \in \mathbb{Z}_{T_{\max}} \\ |\sigma_{\mathbf{n}}| \lesssim 1}} \prod_{\mathbf{n} \in \mathcal{C}} \frac{1}{|c(\{\sigma_{\mathbf{n}}\})_{\mathbf{n}}| + T_{\max}^{-1}} \quad (2.139) \\
&\lesssim L^{O(n\theta)} Q^{\frac{n}{2}} \sum_{c \in \mathcal{M}(\mathcal{C})} \sum_{\substack{\sigma_{\mathbf{n}} \in \mathbb{Z}_{T_{\max}} \\ |\sigma_{\mathbf{n}}| \lesssim 1}} \prod_{\mathbf{n} \in \mathcal{C}} \frac{1}{|c(\{\sigma_{\mathbf{n}}\})_{\mathbf{n}}| + T_{\max}^{-1}}
\end{aligned}$$

Here in the last step we use the fact that $\#\mathcal{D}(\alpha, 1) \lesssim \ln(\alpha^{-1})$. $|\sigma_{\mathbf{n}}| \lesssim 1$ in the sum of the first two inequalities comes from the fact that $\#Eq(\mathcal{C}, \{\sigma_{\mathbf{n}}\}_{\mathbf{n}}, k) = 0$ if some $|\sigma_{\mathbf{n}}| \gtrsim 1$. This fact is true because in $Eq(\mathcal{C}, \{\sigma_{\mathbf{n}}\}_{\mathbf{n}}, k)$, all $|k_{\mathfrak{e}}| \lesssim 1$, which implies that $|\Omega_{\mathbf{n}}| \lesssim 1$ and therefore $|\sigma_{\mathbf{n}}| \gtrsim 1$ (notice that $|\Omega_{\mathbf{n}} - \sigma_{\mathbf{n}}| = O(T_{\max}^{-1})$).

We claim that

$$\sup_c \sum_{\substack{\sigma_{\mathbf{n}} \in \mathbb{Z}_{T_{\max}} \\ |\sigma_{\mathbf{n}}| \lesssim 1}} \prod_{\mathbf{n} \in \mathcal{C}} \frac{1}{|c(\{\sigma_{\mathbf{n}}\})_{\mathbf{n}}| + T_{\max}^{-1}} \lesssim L^{O(n\theta)} T_{\max}^n \quad (2.140)$$

Since there are only bounded many matrices in $\mathcal{M}(\mathcal{C})$. Given the above claim, we know that

$$|Term(T, p)_k| \lesssim L^{O(n\theta)} Q^{\frac{n}{2}} T_{\max}^n, \quad (2.141)$$

which proves the lemma.

Now prove the claim. Notice that there are n nodes in \mathcal{C} . We label these nodes by $h = 1, \dots, n$ and denote $\sigma_{\mathbf{n}}$ by σ_h if \mathbf{n} is labelled by h . Since $\sigma_h \in \mathbb{Z}_{T_{\max}}$, there exists $m_h \in \mathbb{Z}$ such that $\sigma_h = \frac{m_h}{T_{\max}}$. (2.140) is thus equivalent to

$$\sum_{\substack{m_h \in \mathbb{Z} \\ |m_h| \lesssim T_{\max}}} \prod_{h=1}^n \frac{T_{\max}}{|c(\{m_h\})_h| + 1} \lesssim L^{O(n\theta)} T_{\max}^n \quad (2.142)$$

Before proving (2.142), let's first look at one of its special cases. If $c = Id$, then $c(\{m_h\})_h = m_h$. The right hand side of (2.142) becomes

$$T_{\max}^n \sum_{\substack{m_h \in \mathbb{Z} \\ |m_h| \lesssim T_{\max}}} \prod_{h=1}^n \frac{1}{|m_h| + 1} = T_{\max}^n \prod_{h=1}^n \left(\sum_{\substack{m_h \in \mathbb{Z} \\ |m_h| \lesssim T_{\max}}} \frac{1}{|m_h| + 1} \right) \lesssim T_{\max}^n (\ln(\alpha^{-1}))^n = L^{O(n\theta)} T_{\max}^n. \quad (2.143)$$

Here we use the fact that $\sum_{\substack{j \in \mathbb{Z} \\ |j| \lesssim T_{\max}}} \frac{1}{|j| + 1} = O(\ln(\alpha^{-1}))$.

Now we prove (2.142). We just need to show that

$$\sum_{\substack{m_h \in \mathbb{Z} \\ |m_h| \lesssim T_{\max}}} \prod_{h=1}^n \frac{1}{|c(\{m_h\})_h| + 1} \lesssim L^{O(n\theta)} \quad (2.144)$$

By Euler-Maclaurin formula (4.27) and change of variable formula, we get

$$\begin{aligned} \sum_{\substack{m_h \in \mathbb{Z} \\ |m_h| \lesssim T_{\max}}} \prod_{h=1}^n \frac{1}{|c(\{m_h\})_h| + 1} &\leq \int_{|m_h| \lesssim T_{\max}} \prod_{h=1}^n \frac{1}{|c(\{m_h\})_h| + 1} \prod_{h=1}^n dm_h \\ &= \int_{c(\{|m_h| \lesssim T_{\max}\})} \prod_{h=1}^n \frac{1}{|m_h| + 1} |\det c| \prod_{h=1}^n dm_h \\ &\lesssim \prod_{h=1}^n \int_{|m_h| \lesssim T_{\max}} \frac{1}{|m_h| + 1} dm_h \\ &\lesssim (\ln(\alpha^{-1}))^n \lesssim L^{O(n\theta)} \end{aligned} \quad (2.145)$$

We complete the proof of the claim and thus the proof of the lemma. \square

Now we prove Proposition 2.2.7. To start with, recall the large deviation estimate for the Gaussian polynomial.

Lemma 2.4.28 (Large deviation for Gaussian polynomial). *Let $\{\eta_k(\omega)\}$ be i.i.d. complex Gaussian variables with mean 0 and variance 1. Let $F = F(\omega)$ be an degree*

n polynomial of $\{\eta_k(\omega)\}$ defined by

$$F(\omega) = \sum_{k_1, \dots, k_n} a_{k_1 \dots k_n} \prod_{j=1}^n \eta_{k_j}^{\iota_j}, \quad (2.146)$$

where $a_{k_1 \dots k_n}$ are constants, then we have

$$\mathbb{P} \left(|F(\omega)| \geq A \cdot (\mathbb{E}|F(\omega)|^2)^{\frac{1}{2}} \right) \leq C e^{-cA^{\frac{2}{n}}} \quad (2.147)$$

Proof. This is a corollary of the hypercontractivity of Ornstein-Uhlenbeck semigroup. A good reference for this topic is [32]. (2.147) is equivalent to (B.9) in [32]. \square

Proposition 2.2.7 is a corollary of the above large deviation estimate and Proposition 2.4.26.

Proof of Proposition 2.2.7. Because $(\mathcal{J}_T)_k$ are Gaussian polynomials, we can take $F(\omega) = (\mathcal{J}_T)_k$ in Lemma 2.4.28. Then we obtained

$$|(\mathcal{J}_T)_k(t)| \lesssim L^{\frac{n}{2}\theta} \sqrt{\mathbb{E}|(\mathcal{J}_T)_k|^2} \quad (2.148)$$

with probability less than $e^{-c(L^{\frac{n}{2}\theta})^{\frac{2}{n}}} = e^{-cL^\theta}$. By definition 2.2.6, the above inequality holds true L -certainly.

Since by Proposition 2.4.26, $\mathbb{E}|(\mathcal{J}_T)_k|^2 \lesssim L^{O(l(T)\theta)} \rho^{2l(T)}$, then we get

$$|(\mathcal{J}_T)_k(t)| \lesssim L^{O(l(T)\theta)} \rho^{l(T)}, \quad L\text{-certainly} \quad (2.149)$$

(2.149) is very similar to the final goal (2.35), except for the \sup_t and \sup_k in front. In what follows we apply the standard epsilon net and union bound method to remove these two sup.

Assume that $|t - t'| \lesssim \rho^{l(T)} L^{-M}$, it's not hard to show that $|(\mathcal{J}_T)_k(t) - (\mathcal{J}_T)_k(t')| \lesssim \rho^{l(T)}$. Therefore, if $\sup_i \sup_k |(\mathcal{J}_T)_k(i\rho^{l(T)} L^{-M})| \lesssim L^{O(l(T)\theta)} \rho^{l(T)}$, then $\sup_t \sup_k |(\mathcal{J}_T)_k(t)| \lesssim$

$L^{O(l(T)\theta)}\rho^{l(T)}$ and the proof is completed.

Notice that

$$\begin{aligned}
& \mathbb{P} \left(\sup_{i \in \mathbb{Z}} \sup_k |(\mathcal{J}_T)_k(i\rho^{l(T)}L^{-M})| \gtrsim L^{O(l(T)\theta)}\rho^{l(T)} \right) \\
&= \mathbb{P} \left(\bigcup_{i \in \mathbb{Z} \cap [0, T_{\max}\rho^{-l(T)}L^M]} \bigcup_{k \in \mathbb{Z}_L \cap [0, 1]} \{ |(\mathcal{J}_T)_k(i\rho^{l(T)}L^{-M})| \gtrsim L^{O(l(T)\theta)}\rho^{l(T)} \} \right) \\
&\lesssim \sum_{i \in \mathbb{Z} \cap [0, T_{\max}\rho^{-l(T)}L^M]} \sum_{k \in \mathbb{Z}_L \cap [0, 1]} \mathbb{P} (|(\mathcal{J}_T)_k(i\rho^{l(T)}L^{-M})| \gtrsim L^{O(l(T)\theta)}\rho^{l(T)}) \\
&\lesssim \sum_{i \in \mathbb{Z} \cap [0, T_{\max}\rho^{-l(T)}L^M]} \sum_{k \in \mathbb{Z}_L \cap [0, 1]} e^{-O(L^\theta)} = L^{2M}e^{-O(L^\theta)} = e^{-O(L^\theta)}
\end{aligned} \tag{2.150}$$

Here in the second inequality the two ranges $[0, T_{\max}\rho^{-l(T)}L^M]$ and $[0, 1]$ of i and k come from the fact that $t = i\rho^{l(T)}L^{-M} \lesssim T_{\max}$ and $(\mathcal{J}_T)_k = 0$ for $|k| \gtrsim 1$. In the last line we can replace the probability by $e^{-O(L^\theta)}$ because the estimate $|(\mathcal{J}_T)_k(t)| \lesssim L^{O(l(T)\theta)}\rho^{l(T)}$ holds true L -certainly.

Now we complete the proof of Proposition 2.2.7. \square

2.4.6 Norm estimate of random matrices

In this section, we prove Proposition 2.2.8.

Remember that by definition of \mathcal{P}_T and \mathcal{T} ,

$$\mathcal{P}_T(w) = \mathcal{T}(\mathcal{J}_T, w) = \frac{i\lambda}{L^d} \sum_{S(k_1, k_2, k)=0} \int_0^t k_x \mathcal{J}_{T, k_1} w_{k_2} e^{is\Omega(k_1, k_2, k) - \nu|k|^2(t-s)} ds. \tag{2.151}$$

Dyadic decomposition of \mathcal{P}_T

Remember that we have the dyadic decomposition $[0, 1] = \bigcup_{\tau=0}^{\infty} [2^{-\tau}, 2^{-\tau-1}]$.

We can then construct a dyadic decomposition $\mathcal{P}_T = \sum_{l=0}^{\infty} \mathcal{P}_T^l$. Here \mathcal{P}_T^l is defined

by the following formula.

$$\mathcal{P}_T^l(w) = \frac{i\lambda}{L^d} \sum_{S(k_1, k_2, k)=0} \int_{(t-s)/T_{\max} \in [2^{-\tau}, 2^{-\tau-1}]} k_x \mathcal{J}_{T, k_1} w_{k_2} e^{is\Omega(k_1, k_2, k) - \nu|k|^2(t-s)} ds \quad (2.152)$$

We also introduce the bilinear operator $\mathcal{T}^l(\phi, \phi)_k$

$$\mathcal{T}^l(\phi, \phi)_k = \frac{i\lambda}{L^d} \sum_{S(k_1, k_2, k)=0} \int_{(t-s)/T_{\max} \in [2^{-\tau}, 2^{-\tau-1}]} k_x \phi_{k_1} \phi_{k_2} e^{is\Omega(k_1, k_2, k) - \nu|k|^2(t-s)} ds \quad (2.153)$$

Proposition 2.2.8 is a corollary of the following proposition.

Proposition 2.4.29. *Let $\rho = \alpha T_{\max}^{\frac{1}{2}}$ and \mathcal{P}_T^l be the operator defined above. Define the space $X_{L^{2M}}^p = \{w \in X^p : w_k = 0 \text{ if } |k| \gtrsim L^{2M}\}$ and the norm $\|w\|_{X_{L^{2M}}^p} = \sup_{|k| \lesssim L^{2M}} \langle k \rangle^p |w_k|$. Then for any sequence of trees and numbers $\{T_1, \dots, T_K\}$ and $\{\tau_1, \dots, \tau_K\}$, we have L -certainly the operator bound*

$$\left\| \sum_{\tau_1, \dots, \tau_K} \prod_{j=1}^K \mathcal{P}_{T_j}^{\tau_j} \right\|_{L_t^\infty X_{L^{2M}}^p \rightarrow L_t^\infty X^p} \leq L^{O(1+\theta \sum_{j=1}^K l(T_j))} \rho^{\sum_{j=1}^K l(T_j)}. \quad (2.154)$$

for any T_j with $l(T_j) \leq N$.

Proof of Proposition 2.2.8: By Proposition 2.4.29, we have

$$\left\| \prod_{j=1}^K \mathcal{P}_{T_j} \right\|_{L_t^\infty X_{L^{2M}}^p \rightarrow L_t^\infty X^p} = \left\| \sum_{\tau_1, \dots, \tau_K} \prod_{j=1}^K \mathcal{P}_{T_j}^{\tau_j} \right\|_{L_t^\infty X_{L^{2M}}^p \rightarrow L_t^\infty X^p} \leq L^{O(1+\theta \sum_{j=1}^K l(T_j))} \rho^{\sum_{j=1}^K l(T_j)}. \quad (2.155)$$

Define $(X_{L^{2M}}^p)^\perp = \{w \in X^p : w_k = 0 \text{ if } |k| \lesssim L^{2M}\}$. To prove Proposition 2.2.8, it suffices to show that for all $w \in (X_{L^{2M}}^p)^\perp$,

$$\left\| \prod_{j=1}^K \mathcal{P}_{T_j} w \right\|_{L_t^\infty X^p} \lesssim L^{O(1+\theta \sum_{j=1}^K l(T_j))} \rho^{\sum_{j=1}^K l(T_j)} \|w\|_{L_t^\infty X^p}. \quad (2.156)$$

By (2.151), if $w \in (X_{L^{2M}}^p)^\perp$, then $\mathcal{P}_{T_j} w \in (X_{L^{2M}}^p)^\perp$ if we enlarge the constant in

$|k| \lesssim L^{2M}$ in the definition of $(X_{L^{2M}}^p)^\perp$. Therefore, to prove (2.156), it suffices to prove

$$\|\mathcal{P}_{T_j} w\|_{L_t^\infty X^p} \lesssim L^{O(1+l(T_j)\theta)} \rho^{l(T_j)} \|w\|_{L_t^\infty X^p}, \quad (2.157)$$

for all $w \in (X_{L^{2M}}^p)^\perp$.

By (2.151),

$$\begin{aligned} |\mathcal{P}_{T_j}(w)_k| &\leq \frac{\lambda}{L^d} \sum_{S(k_1, k_2, k)=0} \int_0^t |k_x| |\mathcal{J}_{T_j, k_1}| |w_{k_2}| e^{-\nu|k|^2(t-s)} ds \\ &\lesssim L^{O(l(T_j)\theta)} \rho^{l(T_j)} \frac{\lambda}{L^d} |k_x| \int_0^t e^{-\nu|k|^2(t-s)} ds \sum_{k_2: |k_2-k| \lesssim 1} \langle k_2 \rangle^{-p} \\ &\lesssim L^{O(l(T_j)\theta)} \rho^{l(T_j)} \frac{\lambda}{L^d} |k_x| \nu^{-1} \langle k \rangle^{-2} \sum_{k_2: |k_2-k| \lesssim 1} \langle k_2 \rangle^{-p} \\ &\lesssim L^{O(l(T_j)\theta)} \rho^{l(T_j)} \frac{\lambda}{L^d} \nu^{-1} \langle k \rangle^{-1} \langle k \rangle^{-p} \lesssim L^{O(l(T_j)\theta)} \rho^{l(T_j)} \frac{\lambda}{L^d} \nu^{-1} L^{-2M} \langle k \rangle^{-p} \\ &\lesssim L^{-M} \rho^{l(T_j)} \langle k \rangle^{-p} \end{aligned} \quad (2.158)$$

Here in the second inequality, we apply Proposition 2.2.7. In the third line we use the fact that $\int_0^t e^{-\nu|k|^2(t-s)} ds \leq \nu^{-1} \langle k \rangle^{-2}$. In the fourth line we use the fact that $\sum_{k_2: |k_2-k| \lesssim 1} \langle k_2 \rangle^{-p} \leq \langle k \rangle^{-p}$ and $|k| \gtrsim L^{2M}$ (since if $|k| \lesssim L^{2M}$ then $|\mathcal{P}_{T_j}(w)_k|$ vanishes and there is nothing to prove).

$$(2.158) \text{ implies that } \|\mathcal{P}_{T_j} w\|_{L_t^\infty X^p} \lesssim L^{-M} \rho^{l(T_j)} \|w\|_{L_t^\infty X^p} \lesssim L^{O(1+l(T_j)\theta)} \rho^{l(T_j)} \|w\|_{L_t^\infty X^p},$$

so we have proved (2.157) and thus (2.41).

Now we prove (2.42). Because $L = \sum_{1 \leq l(T) \leq N} \mathcal{P}_T$ and $L^K = \sum_{1 \leq l(T_1), \dots, l(T_K) \leq N} \mathcal{P}_{T_1} \cdots \mathcal{P}_{T_K}$, by (2.41)

$$\|L^K\|_{L_t^\infty X^p \rightarrow L_t^\infty X^p} \lesssim L^{O(1)} \sum_{1 \leq l(T_1), \dots, l(T_K) \leq N} (L^{O(\theta)} \rho)^{\sum_{j=1}^K l(T_j)} \leq L^{O(1+\theta \sum_{j=1}^K l(T_j))} \rho^K \quad (2.159)$$

Here in the last step, we use the fact that $l(T_j) \geq 1$ for all j and there are bounded

many trees satisfy $1 \leq l(T_1), \dots, l(T_K) \leq N$.

Therefore, we complete the proof of Proposition 2.2.8. \square

Formulas for product of random matrices $\mathcal{P}_{T_j}^l$

In order to prove Proposition 2.4.29, it is very helpful to find a good formula of $\prod_{j=1}^K \mathcal{P}_{T_j}^{\tau_j}$, which is the main goal of this section.

\mathcal{P}_T^l is almost the same as \mathcal{P}_T except for the limits in the time integral, so in the rest part of this section we do not stress their difference. Now let's find a tree representation for \mathcal{P}_T^l or \mathcal{P}_T .

By (2.26), we know that $\mathcal{J}_T = \mathcal{T}(\mathcal{J}_{T_{n_1}}, \mathcal{J}_{T_{n_2}})$ corresponds to the tree T in which the two subtrees of the root nodes are T_{n_1} and T_{n_2} . Taking T_{n_2} to be a one node tree, as the left tree in Figure 2.20, then this graph represents a term $\mathcal{T}(\mathcal{J}_{T_{n_1}}, \xi)$. The right tree in Figure 2.20 represents the term $\mathcal{T}(\mathcal{J}_{T_{n_1}}, w)$, because as in section 2.2.1, the \square node represents a function w (in section 2.2.1 the function is ϕ).



Figure 2.20: Graphical representations of $\mathcal{T}(\mathcal{J}_{T_{n_1}}, \xi)$ and $\mathcal{T}(\mathcal{J}_{T_{n_1}}, w)$.

Let's find the tree representation for $\prod_{j=1}^K \mathcal{P}_{T_j}^{\tau_j}$ or $\prod_{j=1}^K \mathcal{P}_{T_j}$. Recall the expansion process in section 2.2.1: the replacement of \square by a branching node indicates the substitution of ϕ by $\mathcal{T}(\xi, \xi)$. The action of composition $\mathcal{P}_{T_1} \circ \mathcal{P}_{T_2}(w) = \mathcal{T}(\mathcal{J}_{T_1}, \mathcal{T}(\mathcal{J}_{T_2}, w))$ is the substitution of \cdot by $\mathcal{T}(\mathcal{J}_{T_2}, w)$ in $\mathcal{T}(\mathcal{J}_{T_1}, \cdot)$. As in Figure 2.21, if $\mathcal{P}_{T_1} = \mathcal{T}(\mathcal{J}_{T_1}, \cdot)$ is represented by the left tree, then as an analog that a \square node is replaced by a branching node, the substitution of \cdot by $\mathcal{T}(\mathcal{J}_{T_2}, w)$ should correspond to the operation that the \square node in the left tree is replaced by the middle tree corresponding to

$\mathcal{T}(\mathcal{J}_{T_2}, w)$, and the final resulting right tree should correspond to $\mathcal{P}_{T_1} \circ \mathcal{P}_{T_2}(w)$.

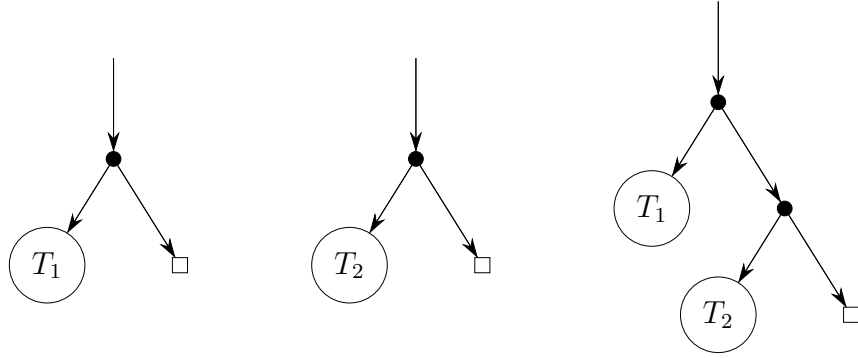


Figure 2.21: The substitution process.

In conclusion, $\mathcal{P}_{T_1} \circ \mathcal{P}_{T_2}(w)$ corresponds to the right tree in Figure 2.21 and more generally, $\prod_{j=1}^K \mathcal{P}_{T_j}(w)$ (or $\prod_{j=1}^K \mathcal{P}_{T_j}^{\tau_j}(w)$) corresponds to the tree in Figure 2.22.

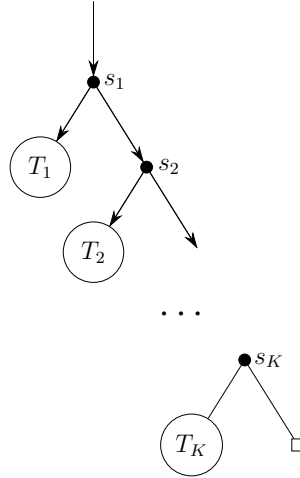


Figure 2.22: Picture of $T_1 \circ \dots \circ T_K$

We introduce the following definition.

Definition 2.4.30. 1. **Definition of $T_1 \circ \dots \circ T_K$:** We define $T_1 \circ \dots \circ T_K$ to be the tree in Figure 2.22.

2. **Substitution nodes:** A node in a tree T with \square nodes is defined to be a substitution node if it is an ancestor of some \square nodes. For a substitution node, we assign a number τ to it, called its index. In Figure 2.22, s_1, \dots, s_K are all

the substitution nodes in $T_1 \circ \cdots \circ T_K$ and we assign index τ_1, \dots, τ_K to them.

Notice that s_1 is the root \mathfrak{r} .

Because the tree in this section contains \square nodes and substitution nodes, we propose the following generalization of Definition 2.2.2 of tree terms.

Definition 2.4.31. Given a binary tree T with \square nodes and substitution nodes, we inductively define the quantity \mathcal{J}_T by:

$$\mathcal{J}_T = \begin{cases} \xi, & \text{if } T \text{ has only one node } \star. \\ w, & \text{if } T \text{ has only one node } \square. \\ \mathcal{T}^l(\mathcal{J}_{T_{\mathfrak{n}_1}}, \mathcal{J}_{T_{\mathfrak{n}_2}}), & \text{if the root of } T \text{ is a substitution node with index } \tau. \\ \mathcal{T}(\mathcal{J}_{T_{\mathfrak{n}_1}}, \mathcal{J}_{T_{\mathfrak{n}_2}}), & \text{otherwise.} \end{cases} \quad (2.160)$$

Here $\mathfrak{n}_1, \mathfrak{n}_2$ are two children of the root node \mathfrak{r} and $T_{\mathfrak{n}_1}, T_{\mathfrak{n}_2}$ are the subtrees of T rooted at above nodes.

Then we have the following formula of $\prod_{j=1}^K \mathcal{P}_{T_j}^{\tau_j}(w)$

Lemma 2.4.32. *With above definition of $T_1 \circ \cdots \circ T_K$ and \mathcal{J}_T , we have*

$$\prod_{j=1}^K \mathcal{P}_{T_j}^{\tau_j}(w) = \mathcal{J}_{T_1 \circ \cdots \circ T_K} \quad (2.161)$$

Proof. This lemma follows from the above explanation. \square

To get a good upper bound, we need a better formula of $\prod_{j=1}^K \mathcal{P}_{T_j}^{\tau_j}(w)$ which is an analog of Lemma 2.4.2.

Lemma 2.4.33. (1) *Using the same notation as Lemma 2.4.2. Given a tree T with \square nodes and substitution nodes. Assume that the root \mathfrak{r} is a substitution node of index τ_1 . Let \mathcal{J}_T be terms defined in Definition 2.4.31, then their Fourier coefficients $\mathcal{J}_{T,k}$ are degree l polynomials of ξ and w given by the following formula*

$$\mathcal{J}_{T,k} = \left(\frac{i\lambda}{L^d}\right)^l \sum_{k_1, k_2, \dots, k_{l+1}} \int_{\cup_{\mathbf{n} \in T_{in}} A_{\mathbf{n}}} e^{\sum_{\mathbf{n} \in T_{in}} i t_{\mathbf{n}} \Omega_{\mathbf{n}} - \nu(t_{\widehat{\mathbf{n}}} - t_{\mathbf{n}})|k_{\mathbf{e}}|^2} \prod_{j=1}^{l+1} (\xi|w)_{k_j} \prod_{\mathbf{n} \in T_{in}} dt_{\mathbf{n}} \delta_{\cap_{\mathbf{n} \in T_{in}} \{S_{\mathbf{n}}=0\}} \prod_{\mathbf{e} \in T_{in}} \iota_{\mathbf{e}} k_{\mathbf{e},x} \quad (2.162)$$

Here ι , $(\xi|w)_{k_j}$, $A_{\mathbf{n}}$, $S_{\mathbf{n}}$, $\Omega_{\mathbf{n}}$ are defined by

$$\iota_{\mathbf{e}} = \begin{cases} +1 & \text{if } \mathbf{e} \text{ pointing inwards to } \mathbf{n} \\ -1 & \text{if } \mathbf{e} \text{ pointing outwards from } \mathbf{n} \end{cases} \quad (2.163)$$

$$(\xi|w)_{k_j} = \begin{cases} \xi & \text{if } j\text{-th leaf node is a } \star \text{ node} \\ w & \text{if } j\text{-th leaf node is a } \square \text{ node} \end{cases} \quad (2.164)$$

$$A_{\mathbf{n}} = \begin{cases} \{t_{\mathbf{r}} \leq t, (t - t_{\mathbf{r}})/T_{max} \in [2^{-\tau_1}, 2^{-\tau_1-1}]\} & \text{if } \mathbf{n} \text{ is the root } \mathbf{r} \\ \{t_{\mathbf{n}_1}, t_{\mathbf{n}_2}, t_{\mathbf{n}_3} \leq t_{\mathbf{n}}\} & \text{if } \mathbf{n} \neq \mathbf{r} \text{ and is not a substitution node} \\ \{t_{\mathbf{n}_1}, t_{\mathbf{n}_2}, t_{\mathbf{n}_3} \leq t_{\mathbf{n}}, (t_{\widehat{\mathbf{n}}} - t_{\mathbf{n}})/T_{max} \in [2^{-\tau}, 2^{-\tau-1}]\} & \text{if } \mathbf{n} \neq \mathbf{r} \text{ is an index } \tau \text{ substitution node} \end{cases} \quad (2.165)$$

$$S_{\mathbf{n}} = \iota_{\mathbf{e}_1} k_{\mathbf{e}_1} + \iota_{\mathbf{e}_2} k_{\mathbf{e}_2} + \iota_{\mathbf{e}} k_{\mathbf{e}} \quad (2.166)$$

$$\Omega_{\mathbf{n}} = \iota_{\mathbf{e}_1} \Lambda_{k_{\mathbf{e}_1}} + \iota_{\mathbf{e}_2} \Lambda_{k_{\mathbf{e}_2}} + \iota_{\mathbf{e}} \Lambda_{k_{\mathbf{e}}} \quad (2.167)$$

For root node \mathbf{r} , we impose the constrain that $k_{\mathbf{r}} = k$. We also define $t_{\widehat{\mathbf{r}}} = t$ to fix the problem that $\widehat{\mathbf{r}}$ is not well defined because \mathbf{r} does not have a parent.

(2) Define $T = T_1 \circ \dots \circ T_K$ and by definition there are only one \square leaf. Assume that there are $l+1$ leaves in T and label all \star leaves by $1, \dots, l$, then $l = \sum_{j=1}^K l(T_j)$. As

a corollary of (1), we have the following formula for Fourier coefficients of $\prod_{j=1}^K \mathcal{P}_{T_j}^{\tau_j}$.

$$\left(\prod_{j=1}^K \mathcal{P}_{T_j}^{\tau_j}(w) \right)_k(t) = \sum_{k'} \int_0^t H_{T_{kk'}}^{\tau_1 \dots \tau_K}(t, s) w_{k'}(s) ds \quad (2.168)$$

and the kernel $H_{T_{kk'}}^{\tau_1 \dots \tau_K}$ is given by

$$H_{T_{kk'}}^{\tau_1 \dots \tau_K}(t, s) = \left(\frac{i\lambda}{L^d} \right)^l \sum_{k_1, k_2, \dots, k_l} H_{T_{k_1 \dots k_l k k'}}^{\tau_1 \dots \tau_K} \xi_{k_1} \dots \xi_{k_l} \quad (2.169)$$

and the coefficients $H_{T_{k_1 \dots k_l k k'}}^{\tau_1 \dots \tau_K}$ of the kernel is given by

$$H_{T_{k_1 \dots k_l k k'}}^{\tau_1 \dots \tau_K}(t, s) = \int_{\cup_{\mathbf{n} \in T_{in}} B_{\mathbf{n}}} e^{\sum_{\mathbf{n} \in T_{in}} i t_{\mathbf{n}} \Omega_{\mathbf{n}} - \nu(t_{\hat{\mathbf{n}}} - t_{\mathbf{n}})|k_{\mathbf{e}}|^2} \prod_{\mathbf{n} \in T_{in}} dt_{\mathbf{n}} \delta_{\cap_{\mathbf{n} \in T_{in}} \{S_{\mathbf{n}}=0\}} \prod_{\mathbf{e} \in T_{in}} \iota_{\mathbf{e}} k_{\mathbf{e}, x} \quad (2.170)$$

Here k_1, \dots, k_l are all variables corresponding to \star leaves, k' is the variable corresponding to the \square node and $B_{\mathbf{n}}$ is defined by

$$B_{\mathbf{n}} = \begin{cases} \{t_{\mathbf{r}} \leq t, (t - t_{\mathbf{r}})/T_{max} \in [2^{-\tau_1}, 2^{-\tau_1-1}]\} & \text{if } \mathbf{n} \text{ is the root } \mathbf{r} \\ \{t_{\mathbf{n}} \geq s\} & \text{if } \mathbf{n} \text{ is a parent of the } \square \text{ nodes} \\ & \text{and is not a substitution nodes} \\ \{t_{\mathbf{n}} \geq s, (t_{\hat{\mathbf{n}}} - t_{\mathbf{n}})/T_{max} \in [2^{-\tau}, 2^{-\tau-1}]\} & \text{if } \mathbf{n} \text{ is a parent of the } \square \text{ nodes} \\ & \text{and is a substitution nodes of index } \tau \\ \{t_{\mathbf{n}_1}, t_{\mathbf{n}_2}, t_{\mathbf{n}_3} \leq t_{\mathbf{n}}, (t_{\hat{\mathbf{n}}} - t_{\mathbf{n}})/T_{max} \in [2^{-\tau}, 2^{-\tau-1}]\} & \text{if not in the first three cases} \\ & \text{and } \mathbf{n} \text{ is a substitution node of index } \tau \\ \{t_{\mathbf{n}_1}, t_{\mathbf{n}_2}, t_{\mathbf{n}_3} \leq t_{\mathbf{n}}\} & \text{otherwise} \end{cases} \quad (2.171)$$

Proof. Lemma 2.4.33 (1) can be proved by the same method as Lemma 2.4.2. We can check that \mathcal{J}_T is defined by (2.54) and (2.55) satisfies the recursive formula (2.26) by direct substitution, so they are the unique solution of that recursive formula, and

this proves (1).

(2) is a corollary of (1). \square

We can also calculate $\prod_{j=1}^K \mathcal{P}_{T_j}^{\tau_j}(\mathcal{J}_T)$, replacing w by \mathcal{J}_T corresponds to replacing a \square node by a tree T , so $\prod_{j=1}^K \mathcal{P}_{T_j}^{\tau_j}(\mathcal{J}_T) = \mathcal{J}_{T_1 \circ T_2 \circ \dots \circ T_K \circ T}$. Since T does not contain \square node, so does $T_1 \circ T_2 \circ \dots \circ T_K \circ T$ and $\mathcal{J}_{T_1 \circ T_2 \circ \dots \circ T_K \circ T}$ is a polynomial of Gaussian variables ξ_k . Then we have the following lemma

Lemma 2.4.34. *We have*

$$\prod_{j=1}^K \mathcal{P}_{T_j}^{\tau_j}(\mathcal{J}_T) = \mathcal{J}_{T_1 \circ T_2 \circ \dots \circ T_K \circ T} \quad (2.172)$$

Proof. This lemma follows from the above explanation. \square

The upper bound for coefficients

In this section, we prove the Lemma 2.4.35 which gives an upper bound for $H_{T_{k_1 \dots k_l k k'}}^{\tau_1 \dots \tau_K}$.

This lemma is an analog of Lemma 2.4.4.

Lemma 2.4.35. *Assume that $|k_1|, \dots, |k_l| \lesssim 1$, then for $t \leq T_{max}$, we have the following upper bound for $H_{T_{k_1 \dots k_l k k'}}^{\tau_1 \dots \tau_K}$,*

$$|H_{T_{k_1 \dots k_l k k'}}^{\tau_1 \dots \tau_K}| \lesssim \sum_{\{d_n\}_{n \in T_{in}} \in \{0,1\}^{l(T)}} \prod_{n \in T_{in}} \frac{2^{-\frac{\tau_n}{2}}}{|q_n| + T_{max}^{-1}} \delta_{\cap_{n \in T_{in}} \{S_n=0\}} \prod_{\epsilon \in T_{in}} p_\epsilon. \quad (2.173)$$

Here τ_n is defined by

$$\tau_n = \begin{cases} 0 & \text{if } n \text{ is not a substitution node} \\ \tau_j & \text{if } n \text{ is the } j\text{-th substitution node} \end{cases} \quad (2.174)$$

Fix a sequence $\{d_n\}_{n \in T_{in}}$ whose elements d_n takes boolean values $\{0,1\}$. We define the

sequences $\{p_n\}_{n \in T_{in}}, \{q_n\}_{n \in T_{in}}, \{r_n\}_{n \in T_{in}}$ by following formulas

$$p_{\mathfrak{e}} = \frac{|k_{\mathfrak{e},x}|}{|k_{\mathfrak{e},x}| + 1} \quad (2.175)$$

$$q_n = \begin{cases} \Omega_{\mathfrak{r}}, & \text{if } n = \text{the root } \mathfrak{r}. \\ \Omega_n + d_n q_{n'}, & \text{if } n \neq \mathfrak{r} \text{ and } n' \text{ is the parent of } n. \end{cases} \quad (2.176)$$

$$r_n = \begin{cases} |k_{\mathfrak{r}}|^2, & \text{if } n = \text{the root } \mathfrak{r}. \\ |k_n|^2 + d_n q_{n'}, & \text{if } n \neq \mathfrak{r} \text{ and } n' \text{ is the parent of } n. \end{cases} \quad (2.177)$$

Proof. By definition

$$H_{T_{k_1 \dots k_l k k'}}^{\tau_1 \dots \tau_K}(t, s) = \int_{\cup_{n \in T_{in}} B_n} e^{\sum_{n \in T_{in}} i t_n \Omega_n - \nu(t_{\hat{n}} - t_n) |k_{\mathfrak{e}}|^2} \prod_{n \in T_{in}} dt_n \delta_{\cap_{n \in T_{in}} \{S_n=0\}} \prod_{\mathfrak{e} \in T_{in}} \iota_{\mathfrak{e}} k_{\mathfrak{e},x} \quad (2.178)$$

For any edge \mathfrak{e} , assume that the two end points of \mathfrak{e} are n_1 and n_2 . If neither n_1 and n_2 is a substitution node, then we claim that $|k_{\mathfrak{e}}| \lesssim 1$.

This is because if no endpoint of \mathfrak{e} is a substitution node, then the subtree $T_{\mathfrak{e}}$ rooted at the upper endpoints of \mathfrak{e} does not contain the \square node as its leaf. Therefore, the momentum conservation (Lemma 2.4.22) is applicable which implies that $k_{\mathfrak{e}}$ is a linear combination of k_1, \dots, k_l . By $|k_1|, \dots, |k_l| \lesssim 1$, we get $|k_{\mathfrak{e}}| \lesssim 1$.

Since $|k_{\mathfrak{e}}| \lesssim 1$, we get $|\prod_{\mathfrak{e} \in T_{in}} \iota_{\mathfrak{e}} k_{\mathfrak{e},x}| \lesssim \prod_{j=1}^K |k_{s_j,x}|$. Here s_1, \dots, s_K are all substitution nodes and k_{s_j} are corresponding variables of edges pointing towards s_j .

As in section 2.4.2, we define

$$F_T(t, \{a_n\}_{n \in T_{in}}, \{b_n\}_{n \in T_{in}}) = \int_{\cup_{n \in T_{in}} B_n} e^{\sum_{n \in T_{in}} i t_n a_n - \nu(t_{\hat{n}} - t_n) b_n} \prod_{n \in T_{in}} dt_n \prod_{j=1}^K |k_{s_j,x}| \quad (2.179)$$

Keeping notation the same as Lemma 2.4.3, if we can show that

$$|F_T(t, \{a_n\}_{n \in T_{\text{in}}}, \{b_n\}_{n \in T_{\text{in}}})| \lesssim \sum_{\{d_n\}_{n \in T_{\text{in}}} \in \{0,1\}^{l(T)}} \prod_{n \in T_{\text{in}}} \frac{2^{-\frac{\tau_n}{2}}}{|q_n| + T_{\text{max}}^{-1}} \prod_{\epsilon \in T_{\text{in}}} p_\epsilon, \quad (2.180)$$

then this lemma can be proved by taking $a_n = \Omega_n$, $b_n = |k_\epsilon|^2$ in (2.180).

Therefore, it suffices to prove (2.180).

We run a similar inductive integration by parts argument of Lemma 2.4.3. If the roots \mathfrak{r} is not a substitution node, then the same argument of Lemma 2.4.3 works. Therefore, we just consider the case that the roots \mathfrak{r} is a substitution node.

Using the same calculation as (2.64), we get

$$\begin{aligned} F_T(t) &= \int_{\cup_{n \in T_{\text{in}}} B_n} e^{it_{\mathfrak{r}}(a_{\mathfrak{r}} + T_{\text{max}}^{-1} \text{sgn}(a_{\mathfrak{r}})) - \nu(t - t_{\mathfrak{r}})b_{\mathfrak{r}}} e^{-iT_{\text{max}}^{-1}t_{\mathfrak{r}}\text{sgn}(a_{\mathfrak{r}})} \\ &\quad e^{\sum_{n \in T_{\text{in},1} \cup T_{\text{in},2}} it_n a_n - \nu(t_{\mathfrak{n}} - t_n)b_n} \left(dt_{\mathfrak{r}} \prod_{j=1}^2 \prod_{n \in T_{\text{in},j}} dt_n \right) |k_{s_1,x}| \prod_{j=2}^K |k_{s_j,x}| \end{aligned} \quad (2.181)$$

We do integration by parts in the above integrals using the Stokes formula. Notice that for $t_{\mathfrak{r}}$, there are four inequality constrains, $t_{\mathfrak{r}} \leq t - 2^{-\tau_1}T_{\text{max}}$, $t_{\mathfrak{r}} \geq t - 2^{-\tau_1-1}T_{\text{max}}$ and $t_{\mathfrak{r}} \geq t_{n_1}, t_{n_2}$. Notice that the first two come from $(t - t_{\mathfrak{r}})/T_{\text{max}} \in [2^{-\tau_1}, 2^{-\tau_1-1}]$.

$$\begin{aligned} F_T(t) &= \frac{|k_{s_1,x}|}{ia_{\mathfrak{r}} + iT_{\text{max}}^{-1}\text{sgn}(a_{\mathfrak{r}}) + \nu b_{\mathfrak{r}}} \int_{\cup_{n \in T_{\text{in}}} B_n} \frac{d}{dt_{\mathfrak{r}}} e^{it_{\mathfrak{r}}(a_{\mathfrak{r}} + T_{\text{max}}^{-1} \text{sgn}(a_{\mathfrak{r}})) - \nu(t - t_{\mathfrak{r}})b_{\mathfrak{r}}} \\ &\quad e^{-iT_{\text{max}}^{-1}t_{\mathfrak{r}}\text{sgn}(a_{\mathfrak{r}})} e^{\sum_{n \in T_{\text{in},1} \cup T_{\text{in},2}} it_n a_n - \nu(t_{\mathfrak{n}} - t_n)b_n} \left(dt_{\mathfrak{r}} \prod_{j=1}^2 \prod_{n \in T_{\text{in},j}} dt_n \right) \prod_{j=2}^K |k_{s_j,x}| \end{aligned} \quad (2.182)$$

$$\begin{aligned}
&= \frac{|k_{s_1,x}|}{ia_{\mathfrak{r}} + iT_{\max}^{-1} \text{sgn}(a_{\mathfrak{r}}) + \nu b_{\mathfrak{r}}} \\
&\quad \left(\int_{\cup_{\mathbf{n} \in T_{\text{in}}} B_{\mathbf{n}}, t_{\mathfrak{r}}=t-2^{-\tau_1} T_{\max}} - \int_{\cup_{\mathbf{n} \in T_{\text{in}}} B_{\mathbf{n}}, t_{\mathfrak{r}}=t-2^{-\tau_1-1} T_{\max}} - \int_{\cup_{\mathbf{n} \in T_{\text{in}}} B_{\mathbf{n}}, t_{\mathfrak{r}}=t_{n_1}} - \int_{\cup_{\mathbf{n} \in T_{\text{in}}} B_{\mathbf{n}}, t_{\mathfrak{r}}=t_{n_2}} \right) \\
&\quad e^{it_{\mathfrak{r}}(a_{\mathfrak{r}}+T_{\max}^{-1} \text{sgn}(a_{\mathfrak{r}}))-\nu(t-t_{\mathfrak{r}})b_{\mathfrak{r}}} e^{-iT_{\max}^{-1} t_{\mathfrak{r}} \text{sgn}(a_{\mathfrak{r}})} e^{\sum_{\mathbf{n} \in T_{\text{in},1} \cup T_{\text{in},2}} it_{\mathbf{n}} a_{\mathbf{n}} - \nu(t_{\widehat{\mathbf{n}}} - t_{\mathbf{n}}) b_{\mathbf{n}}} \left(dt_{\mathfrak{r}} \prod_{j=1}^2 \prod_{\mathbf{n} \in T_{\text{in},j}} dt_{\mathbf{n}} \right) \prod_{j=2}^K |k_{s_j}|
\end{aligned}$$

$$\begin{aligned}
&- \frac{|k_{s_1,x}|}{ia_{\mathfrak{r}} + iT_{\max}^{-1} \text{sgn}(a_{\mathfrak{r}}) + \nu b_{\mathfrak{r}}} \int_{\cup_{\mathbf{n} \in T_{\text{in}}} B_{\mathbf{n}}} e^{it_{\mathfrak{r}}(a_{\mathfrak{r}}+T_{\max}^{-1} \text{sgn}(a_{\mathfrak{r}}))-\nu(t-t_{\mathfrak{r}})b_{\mathfrak{r}}} \\
&\quad \frac{d}{dt_{\mathfrak{r}}} (e^{-iT_{\max}^{-1} t_{\mathfrak{r}} \text{sgn}(a_{\mathfrak{r}})}) e^{\sum_{\mathbf{n} \in T_{\text{in},1} \cup T_{\text{in},2}} it_{\mathbf{n}} a_{\mathbf{n}} - \nu(t_{\widehat{\mathbf{n}}} - t_{\mathbf{n}}) b_{\mathbf{n}}} \left(dt_{\mathfrak{r}} \prod_{j=1}^2 \prod_{\mathbf{n} \in T_{\text{in},j}} dt_{\mathbf{n}} \right) \prod_{j=2}^K |k_{s_j,x}|
\end{aligned}$$

$$= \frac{|k_{s_1,x}|}{ia_{\mathfrak{r}} + iT_{\max}^{-1} \text{sgn}(a_{\mathfrak{r}}) + \nu b_{\mathfrak{r}}} (F_I - F_{I'} - \widetilde{F}_{T^{(1)}} - \widetilde{F}_{T^{(2)}} - F_{II})$$

We now derive upper bounds for F_I , $F_{I'}$, $\widetilde{F}_{T^{(1)}}$, $\widetilde{F}_{T^{(2)}}$, F_{II} .

The argument of F_I and $F_{I'}$ is very similar, so we just consider F_I . By a direct

calculation, we know that $\frac{|k_{s_1,x}|}{ia_{\mathfrak{r}} + iT_{\max}^{-1}\text{sgn}(a_{\mathfrak{r}}) + \nu b_{\mathfrak{r}}} F_I(t)$ equals to

$$\begin{aligned}
& \frac{|k_{s_1,x}|}{ia_{\mathfrak{r}} + iT_{\max}^{-1}\text{sgn}(a_{\mathfrak{r}}) + \nu b_{\mathfrak{r}}} \int_{\cup_{\mathbf{n} \in T_{\text{in}}} B_{\mathbf{n}}, t_{\mathfrak{r}}=t-2^{-\tau_1}T_{\max}} e^{it_{\mathfrak{r}}(a_{\mathfrak{r}} + T_{\max}^{-1}\text{sgn}(a_{\mathfrak{r}})) - \nu(t-t_{\mathfrak{r}})b_{\mathfrak{r}}} e^{-iT_{\max}^{-1}t_{\mathfrak{r}}\text{sgn}(a_{\mathfrak{r}})} \\
& e^{\sum_{\mathbf{n} \in T_{\text{in},1} \cup T_{\text{in},2}} it_{\mathbf{n}}a_{\mathbf{n}} - \nu(t_{\hat{\mathbf{n}}} - t_{\mathbf{n}})b_{\mathbf{n}}} \left(dt_{\mathfrak{r}} \prod_{j=1}^2 \prod_{\mathbf{n} \in T_{\text{in},j}} dt_{\mathbf{n}} \right) \prod_{j=2}^K |k_{s_j,x}| \\
& = \frac{|k_{s_1,x}|}{ia_{\mathfrak{r}} + iT_{\max}^{-1}\text{sgn}(a_{\mathfrak{r}}) + \nu b_{\mathfrak{r}}} e^{i(t-2^{-\tau_1}T_{\max})a_{\mathfrak{r}} - \nu T_{\max}2^{-\tau_1}b_{\mathfrak{r}}} \int_{\cup_{\mathbf{n} \in T_{\text{in},1} \cup T_{\text{in},2}} B_{\mathbf{n}}, t_{\mathfrak{r}_1}, t_{\mathfrak{r}_2} \lesssim t-2^{-\tau_1}T_{\max}} \\
& e^{\sum_{\mathbf{n} \in T_{\text{in},1} \cup T_{\text{in},2}} it_{\mathbf{n}}a_{\mathbf{n}} - \nu(t_{\hat{\mathbf{n}}} - t_{\mathbf{n}})b_{\mathbf{n}}} \left(dt_{\mathfrak{r}} \prod_{j=1}^2 \prod_{\mathbf{n} \in T_{\text{in},j}} dt_{\mathbf{n}} \right) \prod_{j=2}^K |k_{s_j,x}| \\
& = O \left(\left| \frac{|k_{s_1,x}|}{ia_{\mathfrak{r}} + iT_{\max}^{-1}\text{sgn}(a_{\mathfrak{r}}) + \nu b_{\mathfrak{r}}} e^{i(t-2^{-\tau_1}T_{\max})a_{\mathfrak{r}} - \nu T_{\max}2^{-\tau_1}b_{\mathfrak{r}}} \right| |F_{T_1}(t)| |F_{T_2}(t)| \right) \\
& = O \left(\frac{(|k_{s_1,x}| + 1)e^{-2^{-\tau_1}|k_{s_1}|^2}}{|q_{\mathfrak{r}}| + T_{\max}^{-1}} |F_{T_1}(t)| |F_{T_2}(t)| \frac{|k_{s_1,x}|}{|k_{s_1,x}| + 1} \right) = O \left(\frac{2^{-\frac{\tau_1}{2}} |F_{T_1}(t)| |F_{T_2}(t)|}{|q_{\mathfrak{r}}| + T_{\max}^{-1}} p_{\mathfrak{r}} \right)
\end{aligned} \tag{2.183}$$

Here in the last line we use the fact that $b_{\mathfrak{r}} = |k_{s_1}|^2$, $\nu T_{\max} \gtrsim 1$ (2.3) and $(|k_{s_1,x}| + 1)e^{-2^{-\tau_1}|k_{s_1}|^2} \lesssim 2^{-\frac{\tau_1}{2}}$. By the above equation and the induction assumption, we know that $\frac{|k_{s_1,x}|}{ia_{\mathfrak{r}} + iT_{\max}^{-1}\text{sgn}(a_{\mathfrak{r}}) + \nu b_{\mathfrak{r}}} F_I(t)$ can be bounded by the right hand side of (2.173).

Now we find upper bounds of $\tilde{F}_{T^{(1)}}$ and $\tilde{F}_{T^{(2)}}$. Let $T^{(j)}$, $j = 1, 2$ are trees that are obtained by deleting the root \mathfrak{r} , adding edges connecting \mathbf{n}_j with another node and defining \mathbf{n}_j to be the new root. For $T^{(j)}$, we can define the term $F_{T^{(j)}}$ by (2.179). By

a direct calculation, we know that $\frac{|k_{s_1,x}|}{ia_{\mathbf{r}}+iT_{\max}^{-1}\text{sgn}(a_{\mathbf{r}})+\nu b_{\mathbf{r}}}\widetilde{F}_{T^{(1)}}(t)$ equals to

$$\begin{aligned}
& \frac{|k_{s_1,x}|}{ia_{\mathbf{r}}+iT_{\max}^{-1}\text{sgn}(a_{\mathbf{r}})+\nu b_{\mathbf{r}}}\int_{\cup_{\mathbf{n}\in T_{\text{in}}}B_{\mathbf{n}}, t_{\mathbf{r}}=t_{\mathbf{n}_j}} e^{it_{\mathbf{r}}(a_{\mathbf{r}}+T_{\max}^{-1}\text{sgn}(a_{\mathbf{r}}))-\nu(t-t_{\mathbf{r}})b_{\mathbf{r}}}e^{-iT_{\max}^{-1}t_{\mathbf{r}}\text{sgn}(a_{\mathbf{r}})} \\
& e^{\sum_{\mathbf{n}\in T_{\text{in},1}\cup T_{\text{in},2}}it_{\mathbf{n}}a_{\mathbf{n}}-\nu(t_{\widehat{\mathbf{n}}}-t_{\mathbf{n}})b_{\mathbf{n}}}\left(dt_{\mathbf{r}}\prod_{j=1}^2\prod_{\mathbf{n}\in T_{\text{in},j}}dt_{\mathbf{n}}\right)\prod_{j=2}^K|k_{s_j,x}| \\
& =\frac{|k_{s_1,x}|}{ia_{\mathbf{r}}+iT_{\max}^{-1}\text{sgn}(a_{\mathbf{r}})+\nu b_{\mathbf{r}}}e^{-\nu T_{\max}2^{-\tau_1}b_{\mathbf{r}}}\int_{\cup_{\mathbf{n}\in T_{\text{in}}}B_{\mathbf{n}}} e^{it_{\mathbf{n}_j}(a_{\mathbf{r}}+T_{\max}^{-1}\text{sgn}(a_{\mathbf{r}}))-\nu((t-T_{\max}2^{-\tau_1})-t_{\mathbf{n}_j})b_{\mathbf{r}}} \\
& e^{\sum_{\mathbf{n}\in T_{\text{in},1}\cup T_{\text{in},2}}it_{\mathbf{n}}a_{\mathbf{n}}-\nu(t_{\widehat{\mathbf{n}}}-t_{\mathbf{n}})b_{\mathbf{n}}}\left(dt_{\mathbf{r}}\prod_{j=1}^2\prod_{\mathbf{n}\in T_{\text{in},j}}dt_{\mathbf{n}}\right)\prod_{j=2}^K|k_{s_j,x}| \\
& =O\left(\frac{|k_{s_1,x}|e^{-2^{-\tau_1}|k_{s_1}|^2}}{|q_{\mathbf{r}}|+T_{\max}^{-1}}|F_{T^{(j)}}(t)|\right)=O\left(\frac{2^{-\frac{\tau_1}{2}}|F_{T^{(j)}}(t)|}{|q_{\mathbf{r}}|+T_{\max}^{-1}}p_{\mathbf{e}}\right)
\end{aligned} \tag{2.184}$$

We can apply the induction assumption to $F_{T^{(j)}}$ and show that $\frac{|k_{s_1,x}|}{ia_{\mathbf{r}}+iT_{\max}^{-1}\text{sgn}(a_{\mathbf{r}})+\nu b_{\mathbf{r}}}F_{T^{(j)}}$ can be bounded by the right hand side of (2.173).

Another direct calculation gives that

$$F_{II}(t)=\int_{(t-t_{\mathbf{r}})/T_{\max}\in[2^{-\tau_1},2^{-\tau_1-1}]}e^{it_{\mathbf{r}}(a_{\mathbf{r}}+T_{\max}^{-1}\text{sgn}(a_{\mathbf{r}}))-\nu(t-t_{\mathbf{r}})b_{\mathbf{r}}}\frac{d}{dt_{\mathbf{r}}}(e^{-iT_{\max}^{-1}t_{\mathbf{r}}\text{sgn}(a_{\mathbf{r}})})F_{T_1}(t_{\mathbf{r}})F_{T_2}(t_{\mathbf{r}})dt_{\mathbf{r}}. \tag{2.185}$$

Apply the induction assumption

$$\begin{aligned}
& \left|\frac{|k_{s_1,x}|}{ia_{\mathbf{r}}+iT_{\max}^{-1}\text{sgn}(a_{\mathbf{r}})+\nu b_{\mathbf{r}}}F_{II}(t)\right| \\
& \leq \frac{T_{\max}^{-1}|k_{s_1,x}|}{|q_{\mathbf{r}}|+T_{\max}^{-1}}\int_{(t-t_{\mathbf{r}})/T_{\max}\in[2^{-\tau_1},2^{-\tau_1-1}]}e^{-\nu(t-t_{\mathbf{r}})b_{\mathbf{r}}}|F_{T_1}(t_{\mathbf{r}})||F_{T_2}(t_{\mathbf{r}})|dt_{\mathbf{r}} \\
& \leq \frac{T_{\max}^{-1}|k_{s_1,x}|e^{-\nu T_{\max}2^{-\tau_1}|k_{s_1}|^2}}{|q_{\mathbf{r}}|+T_{\max}^{-1}}\prod_{j=1}^2\left(\sum_{\{d_{\mathbf{n}}\}_{\mathbf{n}\in T_{\text{in},j}}\in\{0,1\}^{l(T_j)}}\prod_{\mathbf{n}\in T_{\text{in},j}}\frac{2^{-\frac{\tau_{\mathbf{n}}}{2}}}{|q_{\mathbf{n}}|+T_{\max}^{-1}}\prod_{\mathbf{e}\in T_{\text{in},j}}p_{\mathbf{e}}\right) \\
& \leq \sum_{\{d_{\mathbf{n}}\}_{\mathbf{n}\in T_{\text{in}}}\in\{0,1\}^{l(T)}}\prod_{\mathbf{n}\in T_{\text{in}}}\frac{2^{-\frac{\tau_{\mathbf{n}}}{2}}}{|q_{\mathbf{n}}|+T_{\max}^{-1}}\prod_{\mathbf{e}\in T_{\text{in}}}p_{\mathbf{e}}.
\end{aligned} \tag{2.186}$$

Therefore, we get an upper bound for F_{II} .

Combining the bounds of F_I , $F_{I'}$, $\tilde{F}_{T(1)}$, $\tilde{F}_{T(2)}$, F_{II} , we conclude that F_T can be bounded by the right hand side of (2.173) and thus complete the proof of Lemma 2.4.3. \square

The upper bound for expectation of entries

In this section, we prove the Proposition 2.4.36 which gives an upper bound for $\mathbb{E}|H_{Tkk'}^{\tau_1 \dots \tau_K}|^2$. This lemma is an analog of Proposition 2.4.26.

Proposition 2.4.36. *Assume that α satisfies (2.2). For any $\theta > 0$, we have*

$$\sup_k \mathbb{E}|H_{Tkk'}^{\tau_1 \dots \tau_K}|^2 \lesssim L^{O(l(T)\theta)} 2^{-\frac{1}{2} \sum_{j=1}^K \tau_j} \rho^{2l(T)}. \quad (2.187)$$

and $\mathbb{E}|H_{Tkk'}^{\tau_1 \dots \tau_K}|^2 = 0$ if $|k - k'| \gtrsim 1$.

Proof. We first find a formula of $\mathbb{E}|H_{Tkk'}^{\tau_1 \dots \tau_K}|^2$ which is similar to (2.86) and (2.87).

A direct calculation gives

$$\mathbb{E}|H_{Tkk'}^{\tau_1 \dots \tau_K}|^2 = \mathbb{E}(H_{Tkk'}^{\tau_1 \dots \tau_K} \overline{H_{Tkk'}^{\tau_1 \dots \tau_K}}) = \left(\frac{\lambda}{L^d}\right)^{2l(T)} \sum_{k_1, k_2, \dots, k_{l(T)}} \sum_{k'_1, k'_2, \dots, k'_{l(T)}} \quad (2.188)$$

$$H_{Tk_1 \dots k_{l(T)} kk'}^{\tau_1 \dots \tau_K} \overline{H_{Tk'_1 \dots k'_{l(T)} kk'}^{\tau_1 \dots \tau_K}} \mathbb{E}(\xi_{k_1} \xi_{k_2} \dots \xi_{k_{l(T)}} \xi_{k'_1} \xi_{k'_2} \dots \xi_{k'_{l(T)}})$$

Applying Wick theorem (Lemma 2.4.6) to (2.188), we get

$$\mathbb{E}|H_{Tkk'}^{\tau_1 \dots \tau_K}|^2 = \left(\frac{\lambda}{L^d}\right)^{2l(T)} \sum_{p \in \mathcal{P}(\{k_1, \dots, k_{l(T)}, k'_1, \dots, k'_{l(T)}\})} Term(T, p)_{op, k, k'}. \quad (2.189)$$

and

$$\begin{aligned}
& Term(T, p)_{op, k, k'} \\
&= \sum_{k_1, k_2, \dots, k_{l(T)}} \sum_{k'_1, k'_2, \dots, k'_{l(T)}} H_{T k_1 \dots k_{l(T)} k k'}^{T_1 \dots T_K} \overline{H_{T k'_1 \dots k'_{l(T)} k k'}^{T_1 \dots T_K}} \delta_p(k_1, \dots, k_{l(T)}, k'_1, \dots, k'_{l(T)}) \sqrt{n_{in}(k_1)} \dots .
\end{aligned} \tag{2.190}$$

Since $\alpha = \frac{\lambda}{L^{\frac{d}{2}}}$, $\frac{\lambda}{L^d} = \alpha L^{-\frac{d}{2}}$. Since the number of elements in \mathcal{P} can be bounded by a constant, by Lemma 2.4.37 proved below, we get

$$\begin{aligned}
\mathbb{E} |H_{T k k'}^{T_1 \dots T_K}|^2 &\lesssim (\alpha L^{-\frac{d}{2}})^{2l(T)} L^{O(l(T)\theta)} 2^{-\frac{1}{2} \sum_{j=1}^K \tau_j} (L^d T_{\max}^{-1})^{l(T)} T_{\max}^{2l(T)} \\
&= L^{O(l(T)\theta)} 2^{-\frac{1}{2} \sum_{j=1}^K \tau_j} \rho^{2l(T)}
\end{aligned} \tag{2.191}$$

By Lemma 2.4.27 below $Term(T, p)_{op, k, k'} = 0$ if $|k - k'| \gtrsim 1$, we know that the same is true for $\mathbb{E} |H_{T k k'}^{T_1 \dots T_K}|^2 = 0$.

Therefore, we complete the proof of this proposition. \square

Lemma 2.4.37. *Let $Q = L^d T_{\max}^{-1}$ be the same as in Proposition 2.4.15. Assume that α satisfies (2.2) and $n_{in} \in C_0^\infty(\mathbb{R}^d)$ is compactly supported. Then for any $\theta > 0$, we have*

$$\sup_k |Term(T, p)_{op, k, k'}| \leq L^{O(l(T)\theta)} 2^{-\frac{1}{2} \sum_{j=1}^K \tau_j} Q^{l(T)} T_{\max}^{2l(T)}. \tag{2.192}$$

and $Term(T, p)_{op, k, k'} = 0$ if $|k - k'| \gtrsim 1$.

Proof. By (2.190), we get

$$\begin{aligned}
& Term(T, p)_{op, k, k'} \\
&= \sum_{k_1, k_2, \dots, k_{l(T)}} \sum_{k'_1, k'_2, \dots, k'_{l(T)}} H_{T k_1 \dots k_{l(T)} k k'}^{T_1 \dots T_K} \overline{H_{T k'_1 \dots k'_{l(T)} k k'}^{T_1 \dots T_K}} \delta_p(k_1, \dots, k_{l(T)}, k'_1, \dots, k'_{l(T)}) \sqrt{n_{in}(k_1)} \dots .
\end{aligned} \tag{2.193}$$

Since n_{in} are compactly supported and there are bounded many of them in $Term(T, p)_{op, k, k'}$,

by $k_1 + k_2 + \cdots + k_{l(T)} = k - k'$, we know that $Term(T, p)_{op, k, k'} = 0$ if $|k - k'| \gtrsim 1$.

By (2.173), we get

$$|H_{T k_1 \cdots k_l k k'}^{T_1 \cdots T_K}| \lesssim \sum_{\{d_n\}_{n \in T_{in}} \in \{0, 1\}^{l(T)}} \prod_{n \in T_{in}} \frac{2^{-\frac{\tau_n}{2}}}{|q_n| + T_{\max}^{-1}} \delta_{\cap_{n \in T_{in}} \{S_n=0\}} \prod_{\epsilon \in T_{in}} p_{\epsilon}. \quad (2.194)$$

Define $[c_{n, \tilde{n}}]$, $\mathcal{M}(T)$, $c(\Omega)$ in the same way as the proof of Lemma 2.4.27. We can apply the same derivation of (2.132) to obtain

$$\begin{aligned} |Term(T, p)_{op, k, k'}| &\lesssim \sum_{\substack{k_1, \dots, k_{l(T)}, k'_1, \dots, k'_{l(T)} \\ |k_j|, |k'_j| \lesssim 1, \forall j}} \sum_{c \in \mathcal{M}(T)} \prod_{n \in T_{in}} \frac{2^{-\frac{\tau_n}{2}}}{|c(\Omega)_n| + T_{\max}^{-1}} \delta_{\cap_{n \in T_{in}} \{S_n=0\}} \prod_{\epsilon \in T_{in}} p_{\epsilon} \\ &\sum_{c' \in \mathcal{M}(T)} \prod_{n' \in T_{in}} \frac{2^{-\frac{\tau_n}{2}}}{|c'(\Omega)_{n'}| + T_{\max}^{-1}} \prod_{\epsilon' \in T_{in}} p_{\epsilon'} \delta_{\cap_{n' \in T_{in}} \{S_{n'}=0\}} \delta_p(k_1, \dots, k_{l(T)}, k'_1, \dots, k'_{l(T)}) \end{aligned} \quad (2.195)$$

We obviously have the following inequality

$$\begin{aligned} |Term(T, p)_{op, k, k'}| &\lesssim \sum_{k'} |Term(T, p)_{op, k, k'}| \\ &\lesssim \sum_{\substack{k_1, \dots, k_{l(T)}, k'_1, \dots, k'_{l(T)}, k' \\ |k_j|, |k'_j| \lesssim 1, \forall j, \quad |k' - k| \lesssim 1}} \sum_{c \in \mathcal{M}(T)} \prod_{n \in T_{in}} \frac{2^{-\frac{\tau_n}{2}}}{|c(\Omega)_n| + T_{\max}^{-1}} \delta_{\cap_{n \in T_{in}} \{S_n=0\}} \prod_{\epsilon \in T_{in}} p_{\epsilon} \\ &\sum_{c' \in \mathcal{M}(T)} \prod_{n' \in T_{in}} \frac{2^{-\frac{\tau_n}{2}}}{|c'(\Omega)_{n'}| + T_{\max}^{-1}} \prod_{\epsilon' \in T_{in}} p_{\epsilon'} \delta_{\cap_{n' \in T_{in}} \{S_{n'}=0\}} \delta_p(k_1, \dots, k_{l(T)}, k'_1, \dots, k'_{l(T)}) \end{aligned} \quad (2.196)$$

Switch the order of summations and products in (2.132), then we get

$$\begin{aligned}
|Term(T, p)_{op, k, k'}| &\lesssim \sum_{\substack{k_1, \dots, k_{l(T)}, k'_1, \dots, k'_{l(T)}, k' \\ |k_j|, |k'_j| \lesssim 1, \forall j, |k' - k| \lesssim 1}} \sum_{c, c' \in \mathcal{M}(T)} \prod_{n, n' \in T_{in}} \frac{2^{-\frac{\tau_n}{2}}}{|c(\Omega)_n| + T_{\max}^{-1}} \frac{2^{-\frac{\tau_{n'}}{2}}}{|c'(\Omega)_{n'}| + T_{\max}^{-1}} \\
&\prod_{\epsilon, \epsilon' \in T_{in}} (p_\epsilon p_{\epsilon'}) \delta_{\cap_{n, n' \in T_{in}} \{S_n=0, S_{n'}=0\}} \delta_p(k_1, \dots, k_{l(T)}, k'_1, \dots, k'_{l(T)}).
\end{aligned} \tag{2.197}$$

Consider a tree T with a \square nodes and a pairing $p \in \mathcal{P}(\{k_1, \dots, k_{l(T)}, k'_1, \dots, k'_{l(T)}\})$. p can be viewed as a pairing of all star nodes of two copies of T . A couple \mathcal{C} can be constructed by merging all paired star nodes according to p and merging the two \square nodes. As in (2.134), we can show that

$$\sum_{c, c' \in \mathcal{M}(T)} = \sum_{c \in \mathcal{M}(\mathcal{C})}, \quad \prod_{n, n' \in T_{in}} = \prod_{n \in \mathcal{C}}, \quad \prod_{\epsilon, \epsilon' \in T_{in}} = \prod_{\epsilon \in \mathcal{C}_{norm}}, \quad \cap_{n, n' \in T_{in}} = \cap_{n \in \mathcal{C}}. \tag{2.198}$$

The analog of (2.135) is

$$\begin{aligned}
|Term(T, p)_{op, k, k'}| &\lesssim \sum_{\substack{k_1, \dots, k_{l(T)}, k'_1, \dots, k'_{l(T)}, k' \\ |k_j|, |k'_j| \lesssim 1, \forall j, |k' - k| \lesssim 1}} \sum_{c \in \mathcal{M}(\mathcal{C})} \prod_{n \in \mathcal{C}} \frac{2^{-\frac{\tau_n}{2}}}{|c(\Omega)_n| + T_{\max}^{-1}} \prod_{\epsilon \in \mathcal{C}_{norm}} p_\epsilon \delta_{\cap_{n \in \mathcal{C}} \{S_n=0\}}
\end{aligned} \tag{2.199}$$

The rest part of the proof is exactly the same as the proof of Lemma 2.4.27 after (2.135). For completeness, we include a sketch.

As (2.136), we have

$$\sum_{\substack{k_1, \dots, k_{l(T)+1}, k'_1, \dots, k'_{l(T)+1} \\ \cap_{n \in \mathcal{C}} \{S_n=0\}}} = \sum_{\kappa_\epsilon \in \mathcal{D}(\alpha, 1)} \sum_{\sigma_n \in \mathbb{Z}_{T_{\max}}} \sum_{Eq(\mathcal{C}, \{\sigma_n\}_n, \{\kappa_\epsilon\}_\epsilon, k)} , \tag{2.200}$$

which implies that

$$|Term(T, p)_{op, k, k'}| \lesssim \sum_{\kappa_{\epsilon} \in \mathcal{D}(\alpha, 1)} \sum_{\sigma_n \in \mathbb{Z}_{T_{\max}}} \sum_{Eq(\mathcal{C}, \{\sigma_n\}_n, \{\kappa_{\epsilon}\}_{\epsilon}, k)} \sum_{c \in \mathcal{M}(\mathcal{C})} \prod_{n \in \mathcal{C}} \frac{2^{-\frac{\tau_n}{2}}}{|c(\Omega)_n| + T_{\max}^{-1}} \prod_{\epsilon \in \mathcal{C}_{\text{norm}}} p_{\epsilon} \quad (2.201)$$

As (2.138), we get

$$\begin{aligned} & Term(T, p)_k \\ & \lesssim \sum_{\kappa_{\epsilon} \in \mathcal{D}(\alpha, 1)} \sum_{\sigma_n \in \mathbb{Z}_{T_{\max}}} \sum_{Eq(\mathcal{C}, \{\sigma_n\}_n, \{\kappa_{\epsilon}\}_{\epsilon}, k)} \sum_{c \in \mathcal{M}(\mathcal{C})} \prod_{n \in \mathcal{C}} \frac{2^{-\frac{\tau_n}{2}}}{|c(\Omega)_n| + T_{\max}^{-1}} \prod_{\epsilon \in \mathcal{C}_{\text{norm}}} p_{\epsilon} \\ & \lesssim \sum_{c \in \mathcal{M}(\mathcal{C})} \sum_{\kappa_{\epsilon} \in \mathcal{D}(\alpha, 1)} \sum_{\substack{\sigma_n \in \mathbb{Z}_{T_{\max}} \\ |\sigma_n| \lesssim 1}} \prod_{n \in \mathcal{C}} \frac{2^{-\frac{\tau_n}{2}}}{|c(\{\sigma_n\})_n| + T_{\max}^{-1}} \sum_{Eq(\mathcal{C}, \{\sigma_n\}_n, \{\kappa_{\epsilon}\}_{\epsilon}, k)} 1 \prod_{\epsilon \in \mathcal{C}_{\text{norm}}} \frac{\kappa_{\epsilon}}{\kappa_{\epsilon} + 1} \\ & \lesssim \sum_{c \in \mathcal{M}(\mathcal{C})} \sum_{\kappa_{\epsilon} \in \mathcal{D}(\alpha, 1)} \sum_{\substack{\sigma_n \in \mathbb{Z}_{T_{\max}} \\ |\sigma_n| \lesssim 1}} \prod_{n \in \mathcal{C}} \frac{2^{-\frac{\tau_n}{2}}}{|c(\{\sigma_n\})_n| + T_{\max}^{-1}} \#Eq(\mathcal{C}, \{\sigma_n\}_n, \{\kappa_{\epsilon}\}_{\epsilon}, k) \prod_{\epsilon \in \mathcal{C}_{\text{norm}}} \frac{\kappa_{\epsilon}}{\kappa_{\epsilon} + 1} \\ & \lesssim \sum_{c \in \mathcal{M}(\mathcal{C})} \sum_{\kappa_{\epsilon} \in \mathcal{D}(\alpha, 1)} \sum_{\substack{\sigma_n \in \mathbb{Z}_{T_{\max}} \\ |\sigma_n| \lesssim 1}} \prod_{n \in \mathcal{C}} \frac{2^{-\frac{\tau_n}{2}}}{|c(\{\sigma_n\})_n| + T_{\max}^{-1}} L^{O(n\theta)} Q^{\frac{n}{2}} \prod_{\epsilon \in \mathcal{C}_{\text{norm}}} \frac{\kappa_{\epsilon}}{\kappa_{\epsilon} + 1} \prod_{\epsilon \in \mathcal{C}_{\text{norm}}} \kappa_{\epsilon}^{-1} \end{aligned} \quad (2.202)$$

Here in the last inequality, we applied (2.97) in Proposition 2.4.15.

After simplification, (2.202) gives us

$$\begin{aligned} |Term(T, p)_{op, k, k'}| & \lesssim L^{O(n\theta)} Q^{\frac{n}{2}} \sum_{c \in \mathcal{M}(\mathcal{C})} \sum_{\kappa_{\epsilon} \in \mathcal{D}(\alpha, 1)} \sum_{\substack{\sigma_n \in \mathbb{Z}_{T_{\max}} \\ |\sigma_n| \lesssim 1}} \prod_{n \in \mathcal{C}} \frac{2^{-\frac{\tau_n}{2}}}{|c(\{\sigma_n\})_n| + T_{\max}^{-1}} \\ & \lesssim L^{O(n\theta)} Q^{\frac{n}{2}} \left(\sum_{\kappa_{\epsilon} \in \mathcal{D}(\alpha, 1)} 1 \right) \sum_{c \in \mathcal{M}(\mathcal{C})} \sum_{\substack{\sigma_n \in \mathbb{Z}_{T_{\max}} \\ |\sigma_n| \lesssim 1}} \prod_{n \in \mathcal{C}} \frac{2^{-\frac{\tau_n}{2}}}{|c(\{\sigma_n\})_n| + T_{\max}^{-1}} \\ & \lesssim L^{O(n\theta)} Q^{\frac{n}{2}} \sum_{c \in \mathcal{M}(\mathcal{C})} \sum_{\substack{\sigma_n \in \mathbb{Z}_{T_{\max}} \\ |\sigma_n| \lesssim 1}} \prod_{n \in \mathcal{C}} \frac{2^{-\frac{\tau_n}{2}}}{|c(\{\sigma_n\})_n| + T_{\max}^{-1}} \end{aligned} \quad (2.203)$$

Here in the first step we use the fact that $\prod_{\mathfrak{c} \in \mathcal{C}_{\text{norm}}} \frac{\kappa_{\mathfrak{c}}}{\kappa_{\mathfrak{c}}+1} \prod_{\mathfrak{c} \in \mathcal{C}_{\text{norm}}} \kappa_{\mathfrak{c}}^{-1} = \prod_{\mathfrak{c} \in \mathcal{C}_{\text{norm}}} \frac{1}{\kappa_{\mathfrak{c}}+1} \leq 1$. The reason for other steps can be found in the derivation of (2.139).

We claim that

$$\sup_c \sum_{\substack{\sigma_{\mathbf{n}} \in \mathbb{Z}_{T_{\max}} \\ |\sigma_{\mathbf{n}}| \lesssim 1}} \prod_{\mathbf{n} \in \mathcal{C}} \frac{2^{-\frac{\tau_{\mathbf{n}}}{2}}}{|c(\{\sigma_{\mathbf{n}}\})_{\mathbf{n}}| + T_{\max}^{-1}} \lesssim L^{O(l(T)\theta)} 2^{-\frac{1}{2} \sum_{j=1}^K \tau_j} T_{\max}^{2l(T)} \quad (2.204)$$

Since there are only bounded many matrices in $\mathcal{M}(\mathcal{C})$. Given the above claim, we know that

$$|Term(T, p)_{op, k, k'}| \lesssim L^{O(l(T)\theta)} 2^{-\frac{1}{2} \sum_{j=1}^K \tau_j} Q^{\frac{n}{2}} T_{\max}^{2l(T)}, \quad (2.205)$$

which proves the lemma since $n = 2l(T)$.

Now prove the claim. In a tree T , there are $l(T)$ branching nodes, so there are $l(T)$ nodes in T_{in} . Since all nodes of \mathcal{C} comes the two copies of T_{in} , so there are $2l(T)$ nodes in \mathcal{C} . Label these nodes by $h = 1, \dots, 2l(T)$ and denote $\sigma_{\mathbf{n}}$ by σ_h if \mathbf{n} is labelled by h . Since $\sigma_h \in \mathbb{Z}_{T_{\max}}$, there exists $m_h \in \mathbb{Z}$ such that $\sigma_h = T_{\max}^{-1} m_h$. (2.204) is thus equivalent to

$$T_{\max}^{2l(T)} \sum_{\substack{m_h \in \mathbb{Z} \\ |m_h| \lesssim T_{\max}}} \prod_{h=1}^{2l(T)} \frac{2^{-\frac{\tau_h}{2}}}{|c(\{m_h\})_h| + 1} \lesssim L^{O(l(T)\theta)} 2^{-\frac{1}{2} \sum_{j=1}^K \tau_j} T_{\max}^{2l(T)} \quad (2.206)$$

To prove (2.206). We just need to show that

$$\sum_{\substack{m_h \in \mathbb{Z} \\ |m_h| \lesssim T_{\max}}} \prod_{h=1}^{2l(T)} \frac{1}{|c(\{m_h\})_h| + 1} \lesssim L^{O(l(T)\theta)} \quad (2.207)$$

This can be proved by Euler-Maclaurin formula (4.27) as (2.144).

Now we complete the proof of the claim and thus the proof of the lemma. □

Proof of the operator norm bound

In this subsection, we finish the proof of Proposition 2.4.29.

Proof of Proposition 2.4.29. By Lemma 2.4.33, we have

$$\left(\prod_{j=1}^K \mathcal{P}_{T_j}^{\tau_j}(w) \right)_k(t) = \sum_{k'} \int_0^t H_{Tkk'}^{\tau_1 \dots \tau_K}(t, s) w_{k'}(s) ds \quad (2.208)$$

and the kernel $H_{Tkk'}^{\tau_1 \dots \tau_K}$ is a polynomial of Gaussian variables given by

$$H_{Tkk'}^{\tau_1 \dots \tau_K}(t, s) = \left(\frac{i\lambda}{L^d} \right)^l \sum_{k_1, k_2, \dots, k_l} H_{Tk_1 \dots k_l k k'}^{\tau_1 \dots \tau_K} \xi_{k_1} \dots \xi_{k_l}. \quad (2.209)$$

By Proposition 2.4.36, we have

$$\sup_k \mathbb{E} |H_{Tkk'}^{\tau_1 \dots \tau_K}|^2 \lesssim L^{O(l(T)\theta)} 2^{-\frac{1}{2} \sum_{j=1}^K \tau_j} \rho^{2l(T)}. \quad (2.210)$$

Then the large deviation estimate Lemma 2.4.28 gives

$$|H_{Tkk'}^{\tau_1 \dots \tau_K}(t, s)| \lesssim L^{\frac{n}{2}\theta} \sqrt{\mathbb{E} |H_{Tkk'}^{\tau_1 \dots \tau_K}|^2} \lesssim L^{O(l(T)\theta)} 2^{-\frac{1}{4} \sum_{j=1}^K \tau_j} \rho^{l(T)}, \quad L\text{-certainly}. \quad (2.211)$$

Summing over all τ_1, \dots, τ_K gives

$$\sum_{\tau_1, \dots, \tau_K} |H_{Tkk'}^{\tau_1 \dots \tau_K}(t, s)| \lesssim L^{O(l(T)\theta)} \rho^{l(T)}, \quad L\text{-certainly}. \quad (2.212)$$

Applying the epsilon net and union bound method as in (2.150), we obtain

$$\sup_{t, s} \sup_{|k|, |k'| \lesssim L^{2M}} \sum_{\tau_1, \dots, \tau_K} |H_{Tkk'}^{\tau_1 \dots \tau_K}(t, s)| \lesssim L^{O(l(T)\theta)} \rho^{l(T)}, \quad L\text{-certainly}. \quad (2.213)$$

For $w \in X_{L^{2M}}^p$, we have $|w_k(t)| \leq \sup_t \|w(t)\|_{X_{L^{2M}}^p} \langle k \rangle^{-p}$. Since $w_{k'} = 0$ if $|k'| \gtrsim$

L^{2M} and $H_{Tkk'}^{\tau_1 \dots \tau_K} = 0$ if $|k - k'| \gtrsim 1$, we know that

$$\begin{aligned}
\left| \left(\sum_{\tau_1, \dots, \tau_K} \prod_{j=1}^K \mathcal{P}_{T_j}^{\tau_j}(w) \right)_k(t) \right| &\leq \sum_{|k'| \lesssim L^{2M}} \int_0^t \sum_{\tau_1, \dots, \tau_K} |H_{Tkk'}^{\tau_1 \dots \tau_K}(t, s)| |w_{k'}(s)| ds \\
&\lesssim L^{O(l(T)\theta)} \rho^{l(T)} t \sup_t \|w(t)\|_{X_{L^{2M}}^p} \sum_{\substack{|k'| \lesssim L^{2M} \\ |k' - k| \lesssim 1}} \langle k' \rangle^{-p} \quad (2.214) \\
&\lesssim L^{O(1+l(T)\theta)} \rho^{l(T)} \sup_t \|w(t)\|_{X_{L^{2M}}^p} \langle k \rangle^{-p}.
\end{aligned}$$

Since $l(T) = \sum_{j=1}^K l(T_j)$, L -certainly we have

$$\left\| \sum_{\tau_1, \dots, \tau_K} \prod_{j=1}^K \mathcal{P}_{T_j}^{\tau_j} \right\|_{L_t^\infty X^p} \leq L^{O(1+\theta \sum_{j=1}^K l(T_j))} \rho^{\sum_{j=1}^K l(T_j)} \|w\|_{L_t^\infty X_{L^{2M}}^p}. \quad (2.215)$$

Therefore, we complete the proof of Proposition 2.4.29. \square

2.4.7 Asymptotics of the main terms

In this section, we prove (2.5) in Theorem 2.1.1 which characterize the asymptotic behavior of $n^{(1)}(k)$.

Proposition 2.4.38. *Using the same notation as Theorem 2.1.1 (2), then we have*

$$n^{(1)}(k) = \begin{cases} \frac{t}{T_{\text{kin}}} \mathcal{K}(n_{\text{in}})(k) + O_{\ell_k^\infty} \left(L^{-\theta} \frac{T_{\text{max}}}{T_{\text{kin}}} \right) + \tilde{O}_{\ell_k^\infty} \left(\epsilon_1 \text{Err}_D(k_x) \frac{T_{\text{max}}}{T_{\text{kin}}} \right) & \text{for any } |k| \leq \epsilon_1 l_d^{-1}, \\ 0, & \text{for any } |k| \geq 2C_2 l_d^{-1} \end{cases} \quad (2.216)$$

Here

$$\text{Err}_D(k_x) = \begin{cases} D^{d+1}, & \text{if } |k_x| \leq D, \\ D^{d-1}(|k_x|^2 + D|k_x|), & \text{if } |k_x| \geq D. \end{cases} \quad (2.217)$$

Proof. The second case of (2.5) is obvious. We divide the proof of the first case into several steps.

Step 1. (Calculation of $\psi_{app,k}$ and $n^{(1)}(k)$) The first three terms in the tree expansion (2.27) can be calculated explicitly.

$$\psi_{app,k} = \psi_{app,k}^{(0)} + \psi_{app,k}^{(1)} + \psi_{app,k}^{(2)} + \dots \quad (2.218)$$

where $\psi_{app,k}^{(0)}, \psi_{app,k}^{(1)}, \psi_{app,k}^{(2)}$ are given by

$$\psi_{app,k}^{(0)} = \xi_k \quad (2.219)$$

$$\psi_{app,k}^{(1)} = \frac{i\lambda}{L^d} \sum_{k_1+k_2=k} k_x \xi_{k_1} \xi_{k_2} \int_0^t e^{is\Omega(k_1,k_2,k)-\nu|k|^2(t-s)} ds \quad (2.220)$$

$$\begin{aligned} \psi_{app,k}^{(2)} = & -2 \left(\frac{\lambda}{L^d} \right)^2 \sum_{k_1+k_2+k_3=k} k_x (k_{2x} + k_{3x}) \xi_{k_1} \xi_{k_2} \xi_{k_3} \times \\ & \int_{0 \leq r < s \leq t} e^{is\Omega(k_1,k_2+k_3,k)-\nu|k|^2(t-s)} e^{ir\Omega(k_2,k_3,k_2+k_3)-\nu|k_2+k_3|^2(s-r)} ds dr \end{aligned} \quad (2.221)$$

By (2.53), we know that

$$n^{(1)}(k) = \mathbb{E} \left| \psi_{app,k}^{(1)} \right|^2 + 2\text{Re} \mathbb{E} \left(\psi_{app,k}^{(2)} \overline{\xi_k} \right). \quad (2.222)$$

Step 2. (Decomposition of $\psi_{app,k}$) $e^{-\nu|k|^2(t-s)}$ is supposed to be close to 1, so we have the decomposition $e^{-\nu|k|^2(t-s)} = 1 + (e^{-\nu|k|^2(t-s)} - 1) = 1 - \left(\int_0^1 e^{-a\nu|k|^2(t-s)} da \right) \nu|k|^2(t-s)$. We can also decompose $\psi_{app,k}^{(1)} = \psi_{app,k}^{(11)} + \psi_{app,k}^{(12)}$ and $\psi_{app,k}^{(2)} = \psi_{app,k}^{(21)} + \psi_{app,k}^{(22)}$ into $\psi_{app,k}^{(1)}$ accordingly.

$$\psi_{app,k}^{(11)} = \frac{i\lambda}{L^d} \sum_{k_1+k_2=k} k_x \xi_{k_1} \xi_{k_2} \int_0^t e^{is\Omega(k_1,k_2,k)} ds \quad (2.223)$$

$$\psi_{app,k}^{(12)} = -\nu|k|^2 \frac{i\lambda}{L^d} \int_0^1 \left(\sum_{k_1+k_2=k} k_x \xi_{k_1} \xi_{k_2} \int_0^t e^{is\Omega(k_1,k_2,k)-a\nu|k|^2(t-s)} (t-s) ds \right) da \quad (2.224)$$

$$\psi_{app,k}^{(21)} = -2 \left(\frac{\lambda}{L^d} \right)^2 \sum_{k_1+k_2+k_3=k} k_x(k_{2x}+k_{3x}) \xi_{k_1} \xi_{k_2} \xi_{k_3} \int_{0 \leq r < s \leq t} e^{is\Omega(k_1,k_2+k_3,k)} e^{ir\Omega(k_2,k_3,k_2+k_3)} ds dr \quad (2.225)$$

$$\begin{aligned} \psi_{app,k}^{(22)} = & 2 \left(\frac{\lambda}{L^d} \right)^2 \int_0^1 \left(\sum_{k_1+k_2+k_3=k} k_x(k_{2x}+k_{3x}) \int_{0 \leq r < s \leq t} (\nu|k|^2(t-s) + \nu|k_2+k_3|^2(s-r)) \right. \\ & \left. e^{is\Omega(k_1,k_2+k_3,k)-a\nu|k|^2(t-s)} e^{ir\Omega(k_2,k_3,k_2+k_3)-a\nu|k_2+k_3|^2(s-r)} ds dr \xi_{k_1} \xi_{k_2} \xi_{k_3} \right) da \end{aligned} \quad (2.226)$$

$n^{(1)}(k)$ is also decomposed into $n^{(1)}(k) = n^{(11)}(k) + n^{(12)}(k)$

$$n^{(11)}(k) = \mathbb{E} \left| \psi_{app,k}^{(11)} \right|^2 + 2\text{Re} \mathbb{E} \left(\psi_{app,k}^{(21)} \overline{\xi_k} \right). \quad (2.227)$$

$$n^{(12)}(k) = \mathbb{E} \left| \psi_{app,k}^{(12)} \right|^2 + 2\text{Re} \mathbb{E} \left(\psi_{app,k}^{(11)} \overline{\psi_{app,k}^{(12)}} \right) + 2\text{Re} \mathbb{E} \left(\psi_{app,k}^{(22)} \overline{\xi_k} \right). \quad (2.228)$$

Step 3. (Estimate of $n^{(12)}(k)$ for $|k| \leq \epsilon_1 l_d^{-1}$) In this step, all constants in \lesssim depend only on the dimension d . Define

$$H_{k_1 k_2 k}^{(11)} = k_x \int_0^t e^{is\Omega(k_1,k_2,k)} ds \quad (2.229)$$

$$H_{k_1 k_2 k}^{(12)} = -\nu|k|^2 k_x \int_0^t e^{is\Omega(k_1,k_2,k)-a\nu|k|^2(t-s)} (t-s) ds \quad (2.230)$$

$$H_{k_1 k_2 k_3 k}^{(21)} = 2k_x(k_{2x}+k_{3x}) \int_{0 \leq r < s \leq t} e^{is\Omega(k_1,k_2+k_3,k)} e^{ir\Omega(k_2,k_3,k_2+k_3)} ds dr \quad (2.231)$$

$$\begin{aligned} H_{k_1 k_2 k_3 k}^{(22)} = & -2k_x(k_{2x}+k_{3x}) \int_{0 \leq r < s \leq t} (\nu|k|^2(t-s) + \nu|k_2+k_3|^2(s-r)) \\ & e^{is\Omega(k_1,k_2+k_3,k)-a\nu|k|^2(t-s)} e^{ir\Omega(k_2,k_3,k_2+k_3)-a\nu|k_2+k_3|^2(s-r)} ds dr \end{aligned} \quad (2.232)$$

Then we have

$$\psi_{app,k}^{(11)} = \frac{i\lambda}{L^d} \sum_{k_1+k_2=k} H_{k_1 k_2 k}^{(11)} \xi_{k_1} \xi_{k_2} \quad (2.233)$$

$$\psi_{app,k}^{(12)} = \int_0^1 \frac{i\lambda}{L^d} \sum_{k_1+k_2=k} H_{k_1 k_2 k}^{(12)} \xi_{k_1} \xi_{k_2} da \quad (2.234)$$

$$\psi_{app,k}^{(21)} = \left(\frac{i\lambda}{L^d}\right)^2 \sum_{k_1+k_2+k_3=k} H_{k_1 k_2 k_3 k}^{(11)} \xi_{k_1} \xi_{k_2} \xi_{k_3} \quad (2.235)$$

$$\psi_{app,k}^{(22)} = \int_0^1 \left(\frac{i\lambda}{L^d}\right)^2 \sum_{k_1+k_2+k_3=k} H_{k_1 k_2 k_3 k}^{(12)} \xi_{k_1} \xi_{k_2} \xi_{k_3} da \quad (2.236)$$

To derive an upper bound for $n^{(12)}(k)$, it suffices to consider $\mathbb{E} \left| \psi_{app,k}^{(12)} \right|^2$, $\text{Re } \mathbb{E} \left(\psi_{app,k}^{(11)} \overline{\psi_{app,k}^{(12)}} \right)$ and $\text{Re } \mathbb{E} \left(\psi_{app,k}^{(22)} \overline{\xi_k} \right)$ separately.

Step 3.1. (Upper bounds of $\mathbb{E} \left| \psi_{app,k}^{(12)} \right|^2$ and $\text{Re } \mathbb{E} \left(\psi_{app,k}^{(11)} \overline{\psi_{app,k}^{(12)}} \right)$) We first derive upper bounds for $H^{(1j)}$. For $H^{(11)}$, we have

$$|H_{k_1 k_2 k}^{(11)}| \lesssim \left| \int_0^t \frac{k_x}{i(\Omega + T_{\max}^{-1} \text{sgn}(\Omega))} e^{-isT_{\max}^{-1} \text{sgn}(\Omega)} \frac{d}{ds} e^{is\Omega + is \text{sgn}(\Omega)/T_{\max}} ds \right| \quad (2.237)$$

Integration by parts we get

$$\begin{aligned} |H_{k_1 k_2 k}^{(11)}| &\lesssim \left| \int_0^t \frac{T_{\max}^{-1} k_x}{\Omega + T_{\max}^{-1} \text{sgn}(\Omega)} e^{is\Omega} ds \right| + \left| \left[\frac{k_x}{i(\Omega + T_{\max}^{-1} \text{sgn}(\Omega))} e^{is\Omega} \right]_0^t \right| \\ &\lesssim \frac{|k_x|}{|\Omega| + T_{\max}^{-1}} \end{aligned} \quad (2.238)$$

By a similar integration by parts method, we get

$$|H_{k_1 k_2 k}^{(12)}| \lesssim \frac{|k|^2 |k_x| \nu T_{\max}}{|\Omega| + T_{\max}^{-1}}. \quad (2.239)$$

By (2.234), we know that

$$\begin{aligned}
\mathbb{E} \left| \psi_{app,k}^{(12)} \right|^2 &\leq \int_0^1 \frac{\lambda^2}{L^{2d}} \mathbb{E} \left| \sum_{k_1+k_2=k} H_{k_1 k_2 k}^{(12)} \xi_{k_1} \xi_{k_2} \right|^2 da \\
&\leq \frac{2\lambda^2}{L^{2d}} \sup_{a \in [0,1]} \sum_{k_1+k_2=k} \left| H_{k_1 k_2 k}^{(12)} \right|^2 n(k_1) n(k_2) \\
&\leq \frac{2\lambda^2}{L^{2d}} (|k|^2 |k_x| \nu T_{\max})^2 \sum_{k_1} \frac{n(k_1) n(k-k_1)}{(|\Omega(k_1, k-k_1, k)| + T_{\max}^{-1})^2} \\
&\leq \frac{2\epsilon_1^2 \lambda^2}{L^{2d}} \max(|k_x|^2, D^2) T_{\max} L^d D^{d-1} = 2\epsilon_1^2 \max(|k_x|^2, D^2) T_{\max} T_{\text{kin}}^{-1} D^{d-1}
\end{aligned} \tag{2.240}$$

Here in the second inequality, we apply the Wick theorem to calculate the expectation. In the third inequality, we apply (2.239). In the last line we apply (4.3) in Theorem 4.1.4 by taking $t = T_{\max}$, $g(x) = \frac{1}{(1+|x|)^2}$ and $F(k_1) = n(k_1) n(k-k_1)$. In the last line we also use the facts that $|k| \leq \epsilon_1 l_d^{-1} = \epsilon_1 (\nu T_{\max})^{-\frac{1}{2}}$, $T_{\text{kin}} = \frac{1}{8\pi\alpha^2} = \frac{L^d}{8\pi\lambda^2}$ and $|k_x| \leq |k_{x1}| + |k_{x2}| \lesssim D$.

By (2.238) and (2.239), we know that

$$\begin{aligned}
\left| \text{Re } \mathbb{E} \left(\psi_{app,k}^{(11)} \overline{\psi_{app,k}^{(12)}} \right) \right| &\leq \int_0^1 \frac{2\lambda^2}{L^{2d}} \text{Re } \mathbb{E} \left(\sum_{k_1+k_2=k} H_{k_1 k_2 k}^{(11)} \xi_{k_1} \xi_{k_2} \sum_{k_1+k_2=k} \overline{H_{k_1 k_2 k}^{(12)} \xi_{k_1} \xi_{k_2}} \right) da \\
&\leq \frac{2\lambda^2}{L^{2d}} \sup_{a \in [0,1]} \sum_{k_1+k_2=k} \left| H_{k_1 k_2 k}^{(11)} \right| \left| H_{k_1 k_2 k}^{(12)} \right| n(k_1) n(k_2) \\
&\leq \frac{2\lambda^2}{L^{2d}} |k|^2 |k_x|^2 \nu T_{\max} \sum_{k_1} \frac{n(k_1) n(k-k_1)}{(|\Omega(k_1, k-k_1, k)| + T_{\max}^{-1})^2} D^{d-1} \\
&\leq \frac{2\epsilon_1 \lambda^2}{L^{2d}} \max(|k_x|^2, D^2) T_{\max} L^d = \epsilon_1 \max(|k_x|^2, D^2) T_{\max} T_{\text{kin}}^{-1} D^{d-1}
\end{aligned} \tag{2.241}$$

Here in the second inequality we apply the Wick theorem to calculate the expectation. In the third inequality we apply (2.238) and (2.239). In the last line we apply (4.3) in Theorem 4.1.4 by taking $t = T_{\max}$, $g(x) = \frac{1}{(1+|x|)^2}$ and $F(k_1) = n(k_1) n(k-k_1)$. In the last line we also use the facts that $|k| \leq \epsilon_1 l_d^{-1} = \epsilon_1 (\nu T_{\max})^{-\frac{1}{2}}$, $T_{\text{kin}} = \frac{1}{8\pi\alpha^2} = \frac{L^d}{8\pi\lambda^2}$

and $|k_x| \leq |k_{x1}| + |k_{x2}| \lesssim D$.

(2.240) and (2.241) give desire upper bounds of $\mathbb{E} \left| \psi_{app,k}^{(12)} \right|^2$ and $\text{Re } \mathbb{E} \left(\psi_{app,k}^{(11)} \overline{\psi_{app,k}^{(12)}} \right)$.

Step 3.2. (Upper bound of $\text{Re } \mathbb{E} \left(\psi_{app,k}^{(22)} \overline{\xi_k} \right)$) By (2.236), we have

$$\text{Re } \mathbb{E} \left(\psi_{app,k}^{(22)} \overline{\xi_k} \right) = \int_0^1 \left(\frac{i\lambda}{L^d} \right)^2 \text{Re} \sum_{k_1+k_2+k_3=k} H_{k_1 k_2 k_3 k}^{(22)} \mathbb{E} \left(\xi_{k_1} \xi_{k_2} \xi_{k_3} \overline{\xi_k} \right) da \quad (2.242)$$

By Wick theorem, $\mathbb{E} \left(\xi_{k_1} \xi_{k_2} \xi_{k_3} \overline{\xi_k} \right) = \delta_{k_1=-k_2} \delta_{k_3=k} + \delta_{k_1=-k_3} \delta_{k_2=k} + \delta_{k_1=k} \delta_{k_2=-k_3}$.

Therefore we get

$$\begin{aligned} \text{Re } \mathbb{E} \left(\psi_{app,k}^{(22)} \overline{\xi_k} \right) &= \int_0^1 \left(\frac{i\lambda}{L^d} \right)^2 \text{Re} \sum_{k_1+k_2+k_3=k} H_{k_1 k_2 k_3 k}^{(22)} (\delta_{k_1=-k_2} \delta_{k_3=k} + \delta_{k_1=-k_3} \delta_{k_2=k} + 0) da \\ &= \int_0^1 2 \left(\frac{i\lambda}{L^d} \right)^2 \sum_{k_1+k_2+k_3=k} \text{Re} \left(H_{k_1, -k_1, k, k}^{(22)} \right) da \end{aligned} \quad (2.243)$$

Here in the first equality the term corresponding to $\delta_{k_1=k} \delta_{k_2=-k_3}$ vanishes because $H_{k, k_2, -k_2, k}^{(22)} = 0$ and the two terms corresponding to $\delta_{k_1=-k_2} \delta_{k_3=k}$, $\delta_{k_1=-k_3} \delta_{k_2=k}$ are equal.

By (2.232), we get

$$\begin{aligned} H_{k_1, -k_1, k, k}^{(22)} &= -2k_x(k_x - k_{1x}) \int_{0 \leq r < s \leq t} (\nu |k|^2(t-s) + \nu |k - k_1|^2(s-r)) \\ &\quad e^{i(s-r)\Omega(k_1, k-k_1, k) - a\nu |k|^2(t-s) - a\nu |k-k_1|^2(s-r)} ds dr. \end{aligned} \quad (2.244)$$

We find upper bound of $\operatorname{Re} \left(H_{k_1, -k_1, k, k}^{(22)} \right)$ using integration by parts.

$$\begin{aligned}
\operatorname{Re} \left(H_{k_1, -k_1, k, k}^{(22)} \right) &= 2k_x(k_x - k_{1x}) \int_{0 \leq r < s \leq t} (\nu|k|^2(t-s) + \nu|k - k_1|^2(s-r)) \\
&\quad \frac{e^{irT_{\max}^{-1} \operatorname{sgn} \Omega}}{i(\Omega + T_{\max}^{-1} \operatorname{sgn} \Omega)} \frac{d}{dr} e^{i(s-r)\Omega - irT_{\max}^{-1} \operatorname{sgn} \Omega} e^{-a\nu|k|^2(t-s) - a\nu|k - k_1|^2(s-r)} ds dr \\
&= \operatorname{Re} \frac{2k_x(k_x - k_{1x})}{i(\Omega + T_{\max}^{-1} \operatorname{sgn} \Omega)} \int_{0 \leq s \leq t} \nu|k|^2(t-s) e^{-a\nu|k|^2(t-s)} ds \\
&\quad - \operatorname{Re} \frac{2k_x(k_x - k_{1x})}{i(\Omega + T_{\max}^{-1} \operatorname{sgn} \Omega)} \int_{0 \leq s \leq t} (\nu|k|^2(t-s) + \nu|k - k_1|^2 s) e^{-a\nu|k|^2(t-s) - a\nu|k - k_1|^2 s} e^{is\Omega} ds \\
&\quad - \operatorname{Re} \frac{2k_x(k_x - k_{1x})}{i(\Omega + T_{\max}^{-1} \operatorname{sgn} \Omega)} \int_{0 \leq r < s \leq t} [\nu|k - k_1|^2(\nu|k|^2(t-s) + \nu|k - k_1|^2(s-r) - 1) \\
&\quad - iT_{\max}^{-1} \operatorname{sgn} \Omega] e^{i(s-r)\Omega - a\nu|k|^2(t-s) - a\nu|k - k_1|^2(s-r)} ds dr.
\end{aligned} \tag{2.245}$$

The first term on the right hand side equals 0 after taking the real part. Using the same integration by parts argument as in (2.238), the second term can be bounded by

$$\frac{|k|^2|k_x|(|k_{1x}| + |k_x|)\nu T_{\max}}{(|\Omega(k_1, k - k_1, k)| + T_{\max}^{-1})^2}. \tag{2.246}$$

The last term can be bounded by integration by parts in the following integral

$$\begin{aligned}
&\int_{0 \leq r < s \leq t} [\nu|k - k_1|^2(\nu|k|^2(t-s) + \nu|k - k_1|^2(s-r) - 1) \\
&\quad - iT_{\max}^{-1} \operatorname{sgn} \Omega] \frac{e^{irT_{\max}^{-1} \operatorname{sgn} \Omega}}{i(\Omega + T_{\max}^{-1} \operatorname{sgn} \Omega)} \frac{d}{dr} e^{i(s-r)\Omega} e^{-a\nu|k|^2(t-s) - a\nu|k - k_1|^2(s-r)} ds dr
\end{aligned} \tag{2.247}$$

Integration by parts and bound the three resulting integral by taking the absolute value of the integrand, then we get

$$|(2.247)| \leq \frac{|k|^2}{|\Omega| + T_{\max}^{-1}} \tag{2.248}$$

Therefore, the last term in (2.245) can also be bounded by (2.246).

Then we get

$$\operatorname{Re} \left(H_{k_1, -k_1, k, k}^{(22)} \right) \lesssim \nu T_{\max} |k|^2 |k_x| \frac{|k_{1x}| + |k_x|}{(|\Omega(k_1, k - k_1, k)| + T_{\max}^{-1})^2}. \quad (2.249)$$

Substitute into (2.243), then we have

$$\begin{aligned} \operatorname{Re} \mathbb{E} \left(\psi_{app, k}^{(22)} \overline{\xi_k} \right) &\lesssim \frac{\lambda^2}{L^{2d}} |k|^2 |k_x| \nu T_{\max} n(k) \sum_{k_1} \frac{|k_{1x}| + |k_x|}{(|\Omega(k_1, k - k_1, k)| + T_{\max}^{-1})^2} n(k_1) \\ &\lesssim \frac{\epsilon_1 \lambda^2}{L^{2d}} |k_x| \sum_{k_1} \frac{D}{(|\Omega(k_1, k - k_1, k)| + T_{\max}^{-1})^2} n(k_1) \\ &\quad + \frac{\epsilon_1 \lambda^2}{L^{2d}} |k_x|^2 \sum_{k_1} \frac{1}{(|\Omega(k_1, k - k_1, k)| + T_{\max}^{-1})^2} n(k_1) \\ &\lesssim \epsilon_1 \frac{T_{\max}}{T_{\text{kin}}} \max(D^{d+1}, |k_x|^2 D^{d-1} + |k_x| D^d) \end{aligned} \quad (2.250)$$

In the second inequality, we also use the facts that $|k| \leq \epsilon_1 l_d^{-1} = \epsilon_1 (\nu T_{\max})^{-\frac{1}{2}}$ and $T_{\text{kin}} = \frac{1}{8\pi\alpha^2} = \frac{L^d}{8\pi\lambda^2}$. In the last inequality we apply (4.3) in Theorem 4.1.4 by taking $t = T_{\max}$, $g(x) = \frac{1}{(1+|x|)^2}$, $F(k_1) = n(k_1)n(k - k_1)$ and $|k_x| \leq |k_{x1}| + |k_{x2}| \lesssim D$.

Combining (2.240), (2.241) and (2.250), we get the following upper bound

$$n^{(12)}(k) \lesssim \epsilon_1 \frac{T_{\max}}{T_{\text{kin}}} \underbrace{\max(D^{d+1}, |k_x|^2 D^{d-1} + |k_x| D^d)}_{\operatorname{Err}_D(k_x)}. \quad (2.251)$$

Step 4. (Asymptotics of $n^{(11)}(k)$) By (2.227) and Wick theorem, we get

$$\begin{aligned}
n^{(11)}(k) &= \frac{2\lambda^2}{L^{2d}} |k_x|^2 \sum_{k_1+k_2=k} n(k_1)n(k_2) \left| \int_0^t e^{is\Omega(k_1,k_2,k)} ds \right|^2 \\
&\quad - \frac{8\lambda^2}{L^{2d}} \sum_{k_1} k_x(k_x - k_{1x}) n(k_1)n(k) \operatorname{Re} \left(\int_{0 \leq r < s \leq t} e^{i(s-r)\Omega(k_1,k-k_1,k)} ds dr \right) \\
&= \frac{2\lambda^2}{L^{2d}} |k_x|^2 \sum_{k_1+k_2=k} n(k_1)n(k_2) \frac{4 \sin^2 \left(\frac{t}{2} \Omega(k_1, k_2, k) \right)}{\Omega^2(k_1, k_2, k)} \\
&\quad - \frac{8\lambda^2}{L^{2d}} \sum_{k_1} k_x(k_x - k_{1x}) n(k_1)n(k) \frac{2 \sin^2 \left(\frac{t}{2} \Omega(k_1, k - k_1, k) \right)}{\Omega^2(k_1, k - k_1, k)} \\
&= 8\pi\alpha^2 t |k_x|^2 \int_{\substack{(k_1, k_2) \in \mathbb{R}^{2d} \\ k_1+k_2=k}} n(k_1)n(k_2) \delta(|k_1|^2 k_{1x} + |k_2|^2 k_{2x} - |k|^2 k_x) dk_1 dk_2 \\
&\quad - 16\pi\alpha^2 t n(k) \int_{\mathbb{R}^d} k_x(k_x - k_{1x}) n(k_1) \delta(|k_1|^2 k_{1x} + |k_2|^2 k_{2x} - |k|^2 k_x) dk_1 + O \left(L^{-\theta} \frac{T_{\max}}{T_{\text{kin}}} \right) \\
&= \frac{t}{T_{\text{kin}}} \mathcal{K}(n)(k) + O \left(L^{-\theta} \frac{T_{\max}}{T_{\text{kin}}} \right)
\end{aligned} \tag{2.252}$$

Here in the third equality we apply (4.1) in Theorem 4.1.4 by taking $t \rightarrow \frac{t}{2}$, $g(x) = \frac{\sin^2(x)}{x^2}$ and $F(k_1) = n(k_1)n(k - k_1)$ or $F(k_1) = n(k_1)$. In the third equality we also use the facts that $T_{\text{kin}} = \frac{1}{8\pi\alpha^2} = \frac{L^d}{8\pi\lambda^2}$.

Combining (2.251) and (2.252), we complete the prove of the first case in (2.5). \square

Chapter 3

Analysis of four wave models

3.1 Introduction

3.1.1 Basic set-ups

In this chapter, we study the wave turbulence theory for the following Klein-Gordon type equation

$$\begin{cases} i\partial_t \psi - \Lambda(\nabla)\psi = \lambda^2 \Lambda(\nabla)^{-\frac{1}{2}} \left(|\Lambda(\nabla)^{-\frac{1}{2}} \psi|^2 \Lambda(\nabla)^{-\frac{1}{2}} \psi \right), \\ \psi(0, x) = \psi_{\text{in}}(x), \quad x \in \mathbb{T}_{L_1 \dots L_d}^d. \end{cases} \quad (\text{NKLK})$$

as an example of general four wave systems. Here $\Lambda(\xi) := \sqrt{1 + |\xi|^2}$.

Here ψ is a complex value function. We consider the periodic boundary condition, which implies that the spatial domain is a torus $\mathbb{T}_{L_1 \dots L_d}^d = [0, L_1] \times \dots \times [0, L_d]$.

We know that the Fourier coefficients of ψ lie on the lattice $\mathbb{Z}_L^d = \{k = \frac{K}{L} : K \in \mathbb{Z}^d\}$. Let n_{in} be a known function, we assume that

$$\psi_{\text{in}}(x) = \frac{1}{L^d} \sum_{k \in \mathbb{Z}_L^d} \sqrt{n_{\text{in}}(k)} \eta_k(\omega) e^{2\pi i k x} \quad (3.1)$$

where $\eta_k(\omega)$ are mean-zero and identically distributed complex Gaussian random variables satisfying $\mathbb{E}|\eta_k|^2 = 1$.

The energy spectrum $n(t, k)$ mentioned in the previous section is defined to be $\mathbb{E}|\widehat{\psi}(t, k)|^2$, where $\psi(t, k)$ are Fourier coefficients of the solution. Although the initial data is assumed to be the Gaussian random field, it is possible to develop a theory for other types of random initial data.

In wave turbulence, the energy distribution $n(k)$ is supposed to evolve according to the following wave kinetic equation

$$\begin{aligned} \partial_t n(t, \xi) &= \mathcal{K}(n(t, \cdot)), \\ \mathcal{K}(\phi)(\xi) &:= \int_{\substack{(\xi_1, \xi_2, \xi_3) \in \mathbb{R}^{3d} \\ \xi_1 - \xi_2 + \xi_3 = \xi}} \phi \phi_1 \phi_2 \phi_3 \left(\frac{1}{\phi_1} - \frac{1}{\phi_2} + \frac{1}{\phi_3} - \frac{1}{\phi} \right) \\ &\quad (\Lambda(\xi_1)\Lambda(\xi_2)\Lambda(\xi_3)\Lambda(\xi))^{-\frac{1}{2}} \delta_{\mathbb{R}}(\Lambda(\xi_1) - \Lambda(\xi_2) + \Lambda(\xi_3) - \Lambda(\xi)) d\xi_1 d\xi_2 d\xi_3, \end{aligned} \tag{WKE}$$

Now we explain the idea of deriving the above wave kinetic equation. The derivation also uses Feynman diagram expansion. Note that in this thesis, we only consider the upper bound for the most difficult terms in the diagram expansion. Therefore, the proof in this thesis is **incomplete**. We believe that all other terms can be handled by the existing technique developed by Deng-Hani [9], [13] and we leave the full proof to be a future project.

3.1.2 Outline of this chapter

As in the previous chapter, in order to apply Deng-Hani's argument to derive the (WKE), we need to analyze the perturbative expansion series. In this section, we explain the special structure of NLS and the difficulties of Deng-Hani's argument to the general dispersive equation without this special structure.

The special structure of NLS

The NLS is given by the following equation

$$i\partial_t\psi - \Delta\psi = \lambda^2|\psi|^2\psi \quad (3.2)$$

Here we consider the focusing NLS and the sign of the nonlinear is not relevant to the argument below.

Its nonlinearity is given by $|u|^2u$. The Fourier coefficients of this nonlinearity are given by

$$\frac{\lambda^2}{L^{2d}} \sum_{\substack{(k_1, k_2, k_3) \in (\mathbb{Z}_L^d)^3 \\ k - k_1 + k_2 - k_3 = 0}} \psi_{k_1} \bar{\psi}_{k_2} \psi_{k_3} \quad (3.3)$$

In order to prove (WKE), we need to show that the nonlinearity is small compared to the linear term. To show the smallness, we have to exploit the square root cancellation given by the randomness of u_k . But in $\sum_{k_1=k_2, k_3=k}$ or $\sum_{k_1=k, k_2=k_3}$, which are equal to $(\sum_{k_1} |u_{k_1}|^2)u_k$, we do not have the desired square root cancellation.

In NLS, $\sum_{k_1} |\psi_{k_1}|^2 = \int_{\mathbb{T}_L^d} |\psi|^2$ is just the L^2 mass which is a conservative quantity. We just need to rewrite the NLS into the following form

$$i\partial_t\psi - \Delta\psi = 2\left(\lambda^2 \int_{\mathbb{T}_L^d} |\psi|^2\right)\psi + \lambda^2\left(|\psi|^2 - 2 \int_{\mathbb{T}_L^d} |\psi|^2\right)\psi \quad (3.4)$$

Due to the conservativeness and multiplier-freeness, the first term $2\lambda^2 \int_{\mathbb{T}_L^d} |\psi|^2\psi$ can be removed completely by a simple change phase argument $\psi \rightarrow e^{-2i\lambda^2 \int_{\mathbb{T}_L^d} |\psi|^2 t} \cdot \psi$.

After doing the change phase, $\lambda^2\left(|\psi|^2 - 2 \int_{\mathbb{T}_L^d} |\psi|^2\right)\psi$ has the desired square root cancellation property.

However, the Klein-Gordon equation (or many other PDEs from physics) does not have such a good structure. The Fourier coefficients of the nonlinearity $\lambda^2\Lambda(\nabla)^{-\frac{1}{2}}|\Lambda(\nabla)^{-\frac{1}{2}}\psi|^2\Lambda(\nabla)^{-\frac{1}{2}}$

are given by

$$\frac{\lambda^2}{L^{2d}} \sum_{\substack{(k_1, k_2, k_3) \in (\mathbb{Z}_L^d)^3 \\ k - k_1 + k_2 - k_3 = 0}} (\Lambda_{k_1} \Lambda_{k_2} \Lambda_{k_3})^{-\frac{1}{2}} \psi_{k_1} \bar{\psi}_{k_2} \psi_{k_3} \quad (3.5)$$

The sum $\sum_{k_1=k_2, k_3=k}$ or $\sum_{k_1=k, k_2=k_3}$ now equals to $(\sum_{k_1} \Lambda_{k_1}^{-1} |u_{k_1}|^2) \Lambda_k^{-1} u_k$. The corresponding L^2 term in this case equals to $\sum_{k_1} \Lambda_{k_1}^{-1} |u_{k_1}|^2$, which is not a conservative quantity. Even worse, this terms equals to $\left(\int_{\mathbb{T}_L^d} |\Lambda(\nabla)^{-\frac{1}{2}} \psi|^2 \right) \Lambda(\nabla)^{-1} \psi$, which contains derivative.

Therefore, the change phase argument does not work for (NKLG), since in the analogous change of variable $\psi_k(t) \rightarrow e^{\frac{2\lambda^2}{L^d} i\Lambda(k)^{-1} \int_0^t M(s) ds} \psi_k(t)$ the quantity $M(t) = \sum_{k_1} \Lambda_{k_1}^{-1} |\psi_{k_1}|^2$ is not a conservative quantity. More importantly, after a change of variable, the nonlinearity becomes

$$\frac{\lambda^2}{L^{2d}} \sum_{\substack{k - k_1 + k_2 - k_3 = 0 \\ k_1 \neq k_2, k_3 \neq k}} \psi_{k_1} \bar{\psi}_{k_2} \psi_{k_3} e^{-\frac{2i\lambda^2}{L^d} \int_0^t M(s) ds (\Lambda(k_1)^{-1} - \Lambda(k_2)^{-1} + \Lambda(k_3)^{-1} - \Lambda(k)^{-1})} \quad (3.6)$$

Due to randomness of M , the additional phase factor $e^{-\frac{2i\lambda^2}{L^d} \int_0^t M(s) ds(\dots)}$ is a random variable which cause a serious problem.

The renormalized approximation series

In order to solve the above problem we introduce a new change of variable $\psi_k(t) \rightarrow e^{\frac{2\lambda^2}{L^d} i\Lambda(k)^{-1} \int_0^t m(s) ds} \psi_k(t)$, where $m(t) = \mathbb{E}M(t)$. After doing this, the phase factor $e^{-\frac{2i\lambda^2}{L^d} \int_0^t m(s) ds(\dots)}$ is not a random variable.

Our change of variable solves the problem coming from the randomness of M . The problem of the time dependence of M will be solved later.

After the renormalization, it can be shown that $\sum_{k_1=k_2, k_3=k}$ and $\sum_{k_1=k, k_2=k_3}$ is small, so in this section, we ignore the contribution from them. For simplicity, we

also ignore the effect of renormalization. Then the equation of Fourier coefficients becomes

$$\begin{aligned} i\dot{\psi}_k = & \left(\Lambda(k) + \frac{2\lambda^2}{L^d} m(t) \Lambda(k)^{-2} \right) \psi_k + \frac{\lambda^2}{L^{2d}} \sum_{S_3=0}^{\times} \Lambda_{k_1 k_2 k_3 k}^{-1} \psi_{k_1} \bar{\psi}_{k_2} \psi_{k_3} \\ & + \frac{2\lambda^2}{L^{2d}} \left(\sum_{k_1 \in \mathbb{Z}_L^d} \Lambda_{k_1}^{-2} \left(|\psi_{k_1}|^2 - \mathbb{E} |\psi_{k_1}|^2 \right) \right) \Lambda_k^{-2} \psi_k \end{aligned} \quad (3.7)$$

Define new dynamical variable $e^{i\Lambda(k)t + \frac{2\lambda^2}{L^d} i\Lambda(k)^{-2} \int_0^t m(s) ds} \psi_k(t)$ and integrate (3.7) in time. Then (NKLG) with initial data (3.1) becomes

$$\begin{aligned} \phi_k = & \xi_k - \underbrace{\frac{i\lambda^2}{L^{2d}} \sum_{S_3=0}^{\times} \int_0^t \Lambda_{k_1 k_2 k_3 k}^{-1} \phi_{k_1} \bar{\phi}_{k_2} \phi_{k_3} e^{-i(\Omega_3 t + \tilde{\Omega}_3)} ds}_{\mathcal{T}_1(\phi, \phi, \phi)_k} \\ & - \underbrace{\frac{2i\lambda^2}{L^{2d}} \int_0^t \left(\sum_{k_1 \in \mathbb{Z}_L^d} \Lambda_{k_1}^{-2} : |\phi_{k_1}|^2 : \right) \Lambda_k^{-1} \phi_k ds}_{\mathcal{T}_2(\phi, \phi, \phi)_k}. \end{aligned} \quad (3.8)$$

Here we introduce $S_{3,k} = k_1 - k_2 + k_3 - k$, $\Omega_{3,k} = \Lambda(k_1) - \Lambda(k_2) + \Lambda(k_3) - \Lambda(k)$, $\tilde{\Omega}'_3(k_1, k_2, k_3, k) = \frac{2\lambda^2}{L^d} \int_0^t M(s) ds (\Lambda(k_1)^{-2} - \Lambda(k_2)^{-2} + \Lambda(k_3)^{-2} - \Lambda(k)^{-2})$. ξ_k are the Fourier coefficients of the initial data of ψ defined by $\xi_k = \sqrt{n_{\text{in}}(k)} \eta_k(\omega)$.

The right hand side by $\mathcal{F}(\psi)_k = \xi_k + \mathcal{T}_1(\psi, \psi, \psi)_k + \mathcal{T}_2(\psi, \psi, \psi)_k$. As in three wave models in the previous chapter, we can construct an approximate series by iteration. $\psi = \mathcal{F}(\psi) = \mathcal{F}(\mathcal{F}(\psi)) = \mathcal{F}(\mathcal{F}(\mathcal{F}(\psi))) = \dots$

Define the approximate solution by $\psi_{app} = \mathcal{F}^N(\xi)$. By recursively expanding \mathcal{F}^N , we get

$$\begin{aligned} \psi_{app} = \mathcal{F}^N(\xi) &= \xi + \mathcal{T}_1(\mathcal{F}^{N-1}(\xi), \mathcal{F}^{N-1}(\xi), \mathcal{F}^{N-1}(\xi)) + \mathcal{T}_2(\mathcal{F}^{N-1}(\xi), \mathcal{F}^{N-1}(\xi), \mathcal{F}^{N-1}(\xi)) \\ &= \xi + \mathcal{T}_1(\xi, \xi, \xi) + \mathcal{T}_2(\xi, \xi, \xi) + \mathcal{T}_1(\mathcal{T}_1(\xi, \xi, \xi), \xi, \xi) + \dots \end{aligned}$$

Since \mathcal{T}_2 contains the renormalization symbol $:\cdot:$, ψ_{app} is not a polynomial of ξ . Instead, it is a renormalized polynomial as defined by Definition 3.3.1 (3).

We need a good upper bound for each term of ψ_{app} . To get this, we analyze Feynman diagrams.

The perturbative analysis

The analysis in previous section suggests that ψ_{app} should be a good approximation of ψ . In other words, the error of this approximation $w = \psi - \psi_{app}$ is very small. The strategy for proving this is the same as in the previous chapter.

We use the follow equation of w which can be derived from (3.8):

$$w = Err(\xi) + Lw + B(w, w) + C(w, w, w) \quad (3.9)$$

Here $Err(\xi)$ is a polynomial of ξ whose degree $\leq N + 1$ monomials vanish. Lw , $B(w, w)$, $C(w, w, w)$ are linear, quadratic, cubic in w respectively.

We prove the smallness of w using the bootstrap method.

Define $\|w\|_{X^p} = \sup_k \langle k \rangle^p |w_k|$. Starting from the assumption that $\sup_t \|w\|_{X^p} \leq CL^{-M}$ ($C, M \gg 1$), in order to close the bootstrap, we need to prove that $\sup_t \|w\|_{X^p} \leq (1 + C/2)L^{-M} < CL^{-M}$. To prove $\|w\|_{X^p} \leq (1 + C/2)L^{-2}$, we use (3.9), which gives

$$\|w\|_{X^p} \leq \|Err(\xi)\|_{X^p} + \|Lw\|_{X^p} + \|B(w, w)\|_{X^p} + \|C(w, w, w)\|_{X^p} \quad (3.10)$$

In the rest part of the proof, we show that

$$\|Err(\xi)\|_{X^p} \leq L^{-M}, \quad \|B(w, w)\|_{X^p} \leq C^2 L^{O_d(1)-2M}, \quad \|C(w, w, w)\|_{X^p} \leq C^3 L^{O_d(1)-3M}. \quad (3.11)$$

Combining with a special treatment of Lw , above estimates imply that $\|w\|_{X^p} \leq (1 + C/2)L^{-M}$, which closes the bootstrap.

The proof of inequalities of $B(w, w)$ and $C(w, w)$ is easy and will be given in the later part of this section. The proof of inequalities of $Err(\xi)$ and Lw requires a complete analysis of terms in diagram expansion. As mentioned before, in this thesis, **we only consider the upper bound for those terms that are related to the issue of renormalization**, which should be the most difficult terms in the diagram expansion. We leave it a future project to adapt the existing technique developed by Deng-Hani to get full proof of the wave kinetic equation.

Lattice points counting and $||Err(\xi)||_{X^p}$

In this section we explain the idea of proving the upper bound of $||Err(\xi)||_{X^p}$.

$(Err(\xi))_k$ is a sum of terms of the form

$$\begin{aligned}\mathcal{J}_k^0(\xi) &= \xi_k, \quad \mathcal{J}_k^1(\xi) = \frac{\lambda^2}{L^{2d}} \sum_{k_1 - k_2 + k_3 - k = 0} H_{k_1 k_2 k_3}^1 [\xi_{k_1} \bar{\xi}_{k_2} \xi_{k_3}], \quad \dots \\ \mathcal{J}_k^l(\xi) &= \left(\frac{\lambda^2}{L^{2d}} \right)^l \sum_{k_1 - k_2 + \dots + k_{2l+1} - k = 0} H_{k_1 \dots k_{2l+1}}^l [\xi_{k_1} \bar{\xi}_{k_2} \dots \xi_{k_{2l+1}}], \quad \dots\end{aligned}\tag{3.12}$$

Here the $[\cdot]$ indicates that this is a renormalized polynomial whose precise definition is given in Definition 3.3.1. According to section 3.2.2, each term corresponds to a Feynman diagram, and their coefficient can be calculated from these diagrams.

Difficulties of four wave models: As in the three wave case, from the Feynman diagram representation in section 3.2.2, we know that H^l is large near a surface given by $2l$ equations $S = \{S_3(T) = 0, \Omega_3(T) = 0, \dots, \Omega_{2l+1}(T) = 0, \Omega_{2l+1}(T) = 0\}$. Then in order to estimate $\mathcal{J}_k^l(\xi)$, it suffices to upper bound the number of lattice points near this surface. Unlike the three wave case, we cannot obtain a strong enough number theory estimate without renormalization symbol $[\cdot]$.

By the large deviation principle, to obtain the upper bounds of Gaussian polynomials $\mathcal{J}_{T,k}^l(\xi)$, it suffices to calculate their variance. To calculate the variance of (3.12), we just need to calculate $\mathbb{E}[\xi_{k_1} \bar{\xi}_{k_2} \dots \bar{\xi}_{k_{2l}}]$. This calculation is done in sec-

tion 3.3.2, where we introduce a novel renormalized Wick theorem (Theorem 3.1.1 or Theorem 3.3.5).

The renormalized Wick theorem: The usual Wick theorem says that $\mathbb{E}\xi_{k_1}\bar{\xi}_{k_2}\cdots\bar{\xi}_{k_{2l}} = \sum_{p \in \mathcal{P}} \delta_p$ where \mathcal{P} is a set of pairing whose definition will be given in Definition 3.3.3. Our renormalized Wick theorem says that with the renormalization symbol $\mathbb{E}[\xi_{k_1}\bar{\xi}_{k_2}\cdots\bar{\xi}_{k_{2l}}] = \sum_{p \in \mathcal{P}_F} \delta_p$ where \mathcal{P}_F is a subset of all pairings which excludes many bad pairings.

Wick theorem solves these difficulties: In the Feynman diagram expansion, when calculating the expectation, we encounter the notion of pairing. A pairing is bad if it pairs the first two subtrees of a circle node. For each circle node with this property, the couple diagram generated by this pairing contains bad components. Other circle nodes whose first two subtrees are not paired give good components. In Proposition 3.3.22, we will show that for any couple generated by the $p \in \mathcal{P}_F$ in our renormalized Wick theorem, the number of good couples is more than the number of bad couples if there is no long chain of circle nodes. In section 3.3.4, we explain heuristically why there is no long regular chain of circle nodes due to cancellation.

Combining the heuristic argument in section 3.3.4 and a cancellation identity (3.164), we can get upper bounds for all diagram terms.

In conclusion, combining lattice points counting and renormalization argument we can show that, for any M , we can take N large enough so that $\|Err(\xi)\|_{X^p} \leq L^{-M}$.

Upper bounds for $\|B(w, w)\|_{X^p}$ and $\|C(w, w, w)\|_{X^p}$

$\|B(w, w)\|_{X^p}$ is a sum of terms of the form

$$\frac{\lambda^2}{L^{2d}} \int_0^t \sum_{k_1 - k_2 + k_3 - k = 0} B_{k_1 k_2 k_3}(s) \mathcal{J}_{k_1}^l(\xi) \bar{w}_{k_2} w_{k_3} \quad (3.13)$$

$\|C(w, w, w)\|_{X^p}$ is a cubic polynomial of w of the form

$$\frac{\lambda^2}{L^{2d}} \int_0^t \sum_{k_1 - k_2 + k_3 - k = 0} C_{k_1 k_2 k_3}(s) w_{k_1} \bar{w}_{k_2} w_{k_3} \quad (3.14)$$

By assumptions and proofs in this paper, we know that $t \leq \alpha^{-3/2} \leq L^{O_d(1)}$, $|B_{k_1 k_2 k_3}(s)|, |C_{k_1 k_2 k_3}(s)| \lesssim 1$ and $|\mathcal{J}_{k_1}^l(\xi)| \lesssim \langle k \rangle^{-p}$. By bootstrap assumption, $\sup_k \langle k \rangle^p |w_k| \leq CL^{-M}$. Therefore we have following estimate of $\|B(w, w)\|_{X^p}$ and $\|C(w, w, w)\|_{X^p}$

$$\begin{aligned} \|B(w, w)\|_{X^p} &= \sup_k \langle k \rangle^p \left| \frac{\lambda^2}{L^{2d}} \int_0^t \sum_{k_1 - k_2 + k_3 - k = 0} B_{k_1 k_2 k_3}(s) \mathcal{J}_{k_1}^l(\xi) \bar{w}_{k_2} w_{k_3} \right| \\ &\lesssim \sup_k \langle k \rangle^p \frac{\lambda^2}{L^{2d}} L^{O_d(1)} \sum_{k_1 - k_2 + k_3 - k = 0} \langle k_1 \rangle^{-p} L^{-M} \langle k_2 \rangle^{-p} L^{-M} \langle k_3 \rangle^{-p} \quad (3.15) \\ &\leq C^2 L^{O_d(1) - 2M}, \end{aligned}$$

$$\begin{aligned} \|C(w, w, w)\|_{X^p} &= \sup_k \langle k \rangle^p \left| \frac{\lambda^2}{L^{2d}} \int_0^t \sum_{k_1 - k_2 + k_3 - k = 0} C_{k_1 k_2 k_3}(s) w_{k_1} \bar{w}_{k_2} w_{k_3} \right| \\ &\lesssim \sup_k \langle k \rangle^p \frac{\lambda^2}{L^{2d}} L^{O_d(1)} \sum_{k_1 - k_2 + k_3 - k = 0} L^{-M} \langle k_1 \rangle^{-p} L^{-M} \langle k_2 \rangle^{-p} L^{-M} \langle k_3 \rangle^{-p} \\ &\leq C^3 L^{O_d(1) - 3M} \quad (3.16) \end{aligned}$$

Therefore, we get the desire upper bounds for $\|B(w, w)\|_{X^p}$ and $\|C(w, w, w)\|_{X^p}$.

A random matrix bound and Lw and proof of the main theorem

This part of the proof is the same as the three wave case if we can control all diagram terms.

Main results of this chapter

Now we summarized the main results of this chapter

Theorem 3.1.1 (The renormalized Wick theorem). *The formal statement can be found in Theorem 3.3.5.*

Theorem 3.1.2 (The counting result of closed loops). *The formal statement can be found in Proposition 3.3.22.*

Theorem 3.1.3 (The cancellation identity). *This is (3.164).*

3.1.3 Notations

Universal constants: In this paper, universal constants are constants that just depend on dimension d , diameter D of the support of n_{in} .

$O(\cdot)$, \ll , \lesssim , \sim : Throughout this paper, we frequently use the notation, $O(\cdot)$, \ll , \lesssim . $A = O(B)$ or $A \lesssim B$ means that there exists C such that $A \lesssim CB$. $A \ll B$ means that there exists a small constant c such that $A \lesssim cB$. $A \sim B$ means that there exist two constant c, C such that $cB \lesssim A \lesssim CB$. Here the meaning of constant depends on the context. If they appear in conditions involving k, Λ, Ω , etc., like $|k| \lesssim 1$, $\iota_{\epsilon_1} k_{\epsilon_1} + \iota_{\epsilon_2} k_{\epsilon_2} + \iota_{\epsilon} k_{\epsilon} = 0$, then they are universal constants. If these constants appear in an estimate which gives an upper bound of some quantity, like $\|L^K\|_{X^p \rightarrow X^p} \ll 1$ or $\sup_t \sup_k |(\mathcal{J}_T)_k| \lesssim L^{O(l(T)\theta)} \rho^{l(T)}$, then in addition to the quantities that universal constants depend, they can also depend on the quantities $\theta, \varepsilon, K, M, N, \epsilon_1$.

Order of constants: Here is the order of all constants which can appear in the exponential or superscript of L . These constants are $\theta, \varepsilon, K, M, N, \epsilon_1$.

All the constants are small compared to L in the sense they are less than L^θ for arbitrarily small $\theta > 0$.

ε can be an arbitrarily small constant less than 0.5, the reader is encouraged to assume it to be 0.01. The order of other constants can be decided by the relations $\theta \ll \varepsilon, K = O(\theta^{-1}), M \gg K, N \geq M/\theta$, here the constants in $\ll, O(\cdot)$ are universal.

$$\underline{\mathbb{Z}}_L^d: \mathbb{Z}_L^d = \{k = \frac{K}{L} : K \in \mathbb{Z}^d\}$$

k_x, k_\perp : Given any vector k , let k_x be its first component and k_\perp be the vector formed by the rest components.

$\Lambda(k), \Lambda(\nabla)$: $\Lambda(k) := k_1(k_1^2 + \cdots k_d^2)$ and $\Lambda(\nabla) = i|\nabla|^2 \partial_{x_1}$

Fourier series: The spatial Fourier series of a function $u : \mathbb{T}_L^d \rightarrow \mathbb{C}$ is defined on $\mathbb{Z}_L^d := L^{-1}\mathbb{Z}^d$ by

$$u_k = \int_{\mathbb{T}_L^d} u(x) e^{-2\pi i k \cdot x}, \quad \text{so that} \quad u(x) = \frac{1}{L^d} \sum_{k \in \mathbb{Z}_L^d} u_k e^{2\pi i k \cdot x}. \quad (3.17)$$

Given any function F , let F_k or $(F)_k$ be its Fourier coefficients.

Order of L : In this paper, L is assumed to be a constant that is much larger than all the universal constants, and $\theta, \varepsilon, K, M, N, \epsilon_1$.

3.2 The Perturbation Expansion

In this section, we calculate the renormalized approximate series and introduce Feynman diagrams to represent terms in this series. We will give incomplete proof of these propositions in the rest part of the paper.

3.2.1 The renormalization argument

In this section, we derive the equation for Fourier coefficients and specify the renormalization argument mentioned in the introduction.

Let ψ_k be the Fourier coefficient of ψ . Then in term of ψ_k equation (NKLK) becomes

$$\begin{cases} i\dot{\psi}_k = \Lambda(k)\psi_k + \frac{\lambda^2}{L^{2d}} \sum_{\substack{(k_1, k_2, k_3) \in (\mathbb{Z}_L^d)^3 \\ k - k_1 + k_2 - k_3 = 0}} \Lambda_{k_1 k_2 k_3 k}^{-1} \psi_{k_1} \bar{\psi}_{k_2} \psi_{k_3} \\ \psi_k(0) = \xi_k = \sqrt{n_{\text{in}}(k)} \eta_k(\omega), \end{cases} \quad (3.18)$$

where $\Lambda_{k_1 k_2 k_3 k} = \Lambda_{k_1} \Lambda_{k_2} \Lambda_{k_3} \Lambda_k$.

As explained in section 3.1.2, the contribution from $\sum_{k_1=k_2, k_3=k}$ or $\sum_{k_1=k, k_2=k_3}$ is very large. This suggests us split the sum in (3.18) into

$$\begin{aligned} \sum_{\substack{(k_1, k_2, k_3) \in (\mathbb{Z}_L^d)^3 \\ k-k_1+k_2-k_3=0}} &= \sum_{\substack{k_1, k_3 \neq k \\ k-k_1+k_2-k_3=0}} + \sum_{k_1=k_2, k_3=k} + \sum_{k_1=k, k_2=k_3} - \sum_{k_1=k_2=k_3=k} \\ &= 2 \sum_{k_1=k_2, k_3=k} - \sum_{k_1=k_2=k_3=k} + \sum_{\substack{k_1, k_3 \neq k \\ k-k_1+k_2-k_3=0}} \end{aligned} \quad (3.19)$$

Here in the second equality, we use symmetry to conclude that $\sum_{k_1=k_2, k_3=k} = \sum_{k_1=k, k_2=k_3}$.

Apply this splitting to (3.18),

$$\begin{aligned} i\dot{\psi}_k &= \Lambda(k)\psi_k + \frac{2\lambda^2}{L^{2d}} \left(\sum_{k_1 \in \mathbb{Z}_L^d} \Lambda_{k_1}^{-2} |\psi_{k_1}|^2 \right) \Lambda_k^{-2} \psi_k - \frac{\lambda^2}{L^{2d}} \Lambda_k^{-4} |\psi_k|^2 \psi_k + \frac{\lambda^2}{L^{2d}} \sum_{\substack{k_1, k_3 \neq k \\ S_3=0}} \Lambda_{k_1 k_2 k_3 k}^{-1} \psi_{k_1} \bar{\psi}_{k_2} \psi_{k_3} \\ &= \left(\Lambda(k) + \frac{2\lambda^2}{L^d} M(t) \Lambda(k)^{-2} \right) \psi_k + \frac{\lambda^2}{L^{2d}} \sum_{\substack{\times \\ S_3=0}} \Lambda_{k_1 k_2 k_3 k}^{-1} \psi_{k_1} \bar{\psi}_{k_2} \psi_{k_3} \end{aligned} \quad (3.20)$$

Here we have introduced the notation $\Lambda_{k_1 k_2 k_3 k} = \Lambda_{k_1} \Lambda_{k_2} \Lambda_{k_3} \Lambda_k$, $M(t) = \frac{1}{L^d} \sum_{k_1 \in \mathbb{Z}_L^d} \Lambda_{k_1}^{-2} |\psi_{k_1}|^2$ and $\sum^\times = \sum_{\substack{k_1, k_3 \neq k \\ S_3=0}} - \sum_{k_1=k_2=k_3=k}$. In what follows, a '×' on the summation symbol indicates that there are constraints of the form $k_1, k_3 \neq k$ and $k_1 = k_2 = k_3 = k$. The exact form of inequality constraints or equal summation indices depends on the context.

Change phase argument in previous paper: In [10], the L^2 phase problem was trivial and was solved by a simple change phase argument. Let us explain why this argument does not work for general four wave systems.

In [10], the authors view $\left(\Lambda(k) + \frac{2\lambda^2}{L^d} M(t) \Lambda(k)^{-2} \right) \psi_k$ as the linear part of (3.20) and remove it by rewriting the equation in term of the linear profile $\phi_k(t) := e^{i\Lambda(k)t + \frac{2\lambda^2}{L^d} i\Lambda(k)^{-2} \int_0^t M(s) ds} \psi_k$

$$\begin{cases} i\dot{\phi}_k = \frac{\lambda^2}{L^{2d}} \sum_{S_3=0}^{\times} \Lambda_{k_1 k_2 k_3 k}^{-1} \phi_{k_1} \bar{\phi}_{k_2} \phi_{k_3} e^{-i(\Omega_3 t + \tilde{\Omega}'_3)(k_1, k_2, k_3, k)} \\ \phi_k(0) = \xi_k = \sqrt{n_{\text{in}}(k)} \eta_k(\omega) \end{cases} \quad (3.21)$$

where

$$\begin{aligned} S_3(k_1, k_2, k_3, k) &= k_1 - k_2 + k_3 - k, \\ \Omega_3(k_1, k_2, k_3, k) &= \Lambda(k_1) - \Lambda(k_2) + \Lambda(k_3) - \Lambda(k), \\ \tilde{\Omega}'_3(k_1, k_2, k_3, k) &= \frac{2\lambda^2}{L^d} \int_0^t M(s) ds \left(\Lambda(k_1)^{-2} - \Lambda(k_2)^{-2} + \Lambda(k_3)^{-2} - \Lambda(k)^{-2} \right). \end{aligned} \quad (3.22)$$

Unlike [10], in the Klein-Gordon equation, M is not a conservative quantity, and the linear term of M contains a multiplier $\Lambda(k)^{-2}$. Due to these facts, we have an additional resonant phase $\tilde{\Omega}'_3$. As explained in section 3.1.2, since this resonant phase is a random variable (it contains M which is a random variable), we can't apply the Wick theorem to calculate the correlation functions in the perturbative expansion.

Our renormalization method: To avoid the above difficulties, we define $m(t) = \mathbb{E}M(t)$. Notice that a good approximation of a random variable is its expectation, so we can decompose M into $m + (M - m)$. Therefore, we can replace M with m in the definition of the linear profile. Then we remove the randomness of $\tilde{\Omega}'_3$ at the cost of introducing an error $M - m$ in the equation.

We now explain the renormalization method. Decomposing M in (3.20) gives

$$i\dot{\psi}_k = \left(\Lambda(k) + \frac{2\lambda^2}{L^d} m(t) \Lambda(k)^{-2} \right) \psi_k + \frac{\lambda^2}{L^{2d}} \sum_{S_3=0}^{\times} \Lambda_{k_1 k_2 k_3 k}^{-1} \psi_{k_1} \bar{\psi}_{k_2} \psi_{k_3} + \frac{2\lambda^2}{L^d} (M(t) - m(t)) \Lambda(k)^{-2} \psi_k \quad (3.23)$$

By the definition of M and m , the above equation is equivalent to

$$\begin{aligned}
i\dot{\psi}_k &= \left(\Lambda(k) + \frac{2\lambda^2}{L^d} m(t) \Lambda(k)^{-2} \right) \psi_k + \frac{\lambda^2}{L^{2d}} \sum_{S_3=0}^{\times} \Lambda_{k_1 k_2 k_3 k}^{-1} \psi_{k_1} \bar{\psi}_{k_2} \psi_{k_3} \\
&\quad + \frac{2\lambda^2}{L^{2d}} \left(\sum_{k_1 \in \mathbb{Z}_L^d} \Lambda_{k_1}^{-2} \left(|\psi_{k_1}|^2 - \mathbb{E} |\psi_{k_1}|^2 \right) \right) \Lambda_k^{-2} \psi_k
\end{aligned} \tag{3.24}$$

View $\left(\Lambda(k) + \frac{2\lambda^2}{L^d} M(t) \Lambda(k)^{-2} \right) \psi_k$ as the linear part and define the new linear profile by

$$\phi_k(t) := e^{i\Lambda(k)t + \frac{2\lambda^2}{L^d} i\Lambda(k)^{-2} \int_0^t m(s) ds} \psi_k(t) \tag{3.25}$$

Rewriting (3.24) in terms of ϕ_k gives

$$\begin{aligned}
i\dot{\phi}_k &= \frac{\lambda^2}{L^{2d}} \sum_{S_3=0}^{\times} \Lambda_{k_1 k_2 k_3 k}^{-1} \phi_{k_1} \bar{\phi}_{k_2} \phi_{k_3} e^{-i(\Omega_3 t + \tilde{\Omega}_3)(k_1, k_2, k_3, k)} \\
&\quad + \frac{2\lambda^2}{L^{2d}} \left(\sum_{k_1 \in \mathbb{Z}_L^d} \Lambda_{k_1}^{-2} \left(|\phi_{k_1}|^2 - \mathbb{E} |\phi_{k_1}|^2 \right) \right) \Lambda_k^{-2} \phi_k
\end{aligned} \tag{3.26}$$

where

$$\tilde{\Omega}_3(k_1, k_2, k_3, k) = \frac{2\lambda^2}{L^d} \int_0^t m(s) ds \left(\Lambda(k_1)^{-2} - \Lambda(k_2)^{-2} + \Lambda(k_3)^{-2} - \Lambda(k)^{-2} \right) \tag{3.27}$$

We will work with the renormalized equation (3.26) in the rest part of this paper.

3.2.2 The approximation series and Feynman diagrams

In this section, we construct the approximate solution.

For ease of notation, we define $:X := X - \mathbb{E}X$.

Integrating (3.26) gives

$$\begin{aligned}
\phi_k = \xi_k & - \underbrace{\frac{i\lambda^2}{L^{2d}} \sum_{S_3=0}^{\times} \int_0^t \Lambda_{k_1 k_2 k_3 k}^{-1} \phi_{k_1} \bar{\phi}_{k_2} \phi_{k_3} e^{-i(\Omega_3 t + \tilde{\Omega}_3)} ds}_{\mathcal{T}_1(\phi, \phi, \phi)_k} \\
& - \underbrace{\frac{2i\lambda^2}{L^{2d}} \int_0^t \left(\sum_{k_1 \in \mathbb{Z}_L^d} \Lambda_{k_1}^{-2} : |\phi_{k_1}|^2 : \right) \Lambda_k^{-1} \phi_k ds}_{\mathcal{T}_2(\phi, \phi, \phi)_k}.
\end{aligned} \tag{3.28}$$

Denote the second and the third term of the right hand side by $\mathcal{T}_1(\phi, \phi, \phi)_k$, $\mathcal{T}_2(\phi, \phi, \phi)_k$ respectively. Then the right hand side equals to $\mathcal{F}(\phi)_k = \xi_k + \mathcal{T}_1(\phi, \phi, \phi)_k + \mathcal{T}_2(\phi, \phi, \phi)_k$. With these notations, (3.28) becomes $\phi = \mathcal{F}(\phi)_k$.

We construct the approximation series by iteration: $\phi = \mathcal{F}(\phi) = \mathcal{F}(\mathcal{F}(\phi)) = \mathcal{F}(\mathcal{F}(\mathcal{F}(\phi))) = \dots$. To estimate this approximation series, we need a compact graphical notation to represent the huge amount of terms generated from iteration. This is done by introducing the concept of Feynman diagrams.

Some basic definitions from graph theory

In this section, we introduce the concept of ternary trees used in this paper. Some concepts below may be hard to understand, but we will provide a dictionary between these concepts and coefficients of series expansion in Table 3.2, from which the motivation of these concepts should be obvious.

Definition 3.2.1. In this paper, we use following concepts from the graph theory

1. **Ternary trees:** A ternary tree T is a tree in which each node has 3 or 0 children. An example of ternary trees used in this paper is shown in Figure 3.1.

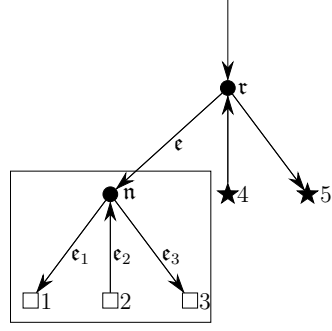


Figure 3.1: Subtrees and node decoration.

2. **Root and children:** The root of a trees and children of nodes are defined in the usual way. The three children from left to right of a node are called the primary, secondary, and tertiary child respectively. For example, in Figure 3.1, the node \mathbf{r} is the root, and nodes 1, 2, 3 are the primary, secondary, and tertiary child of \mathbf{n} respectively.
3. **Branching nodes:** A branching node in a ternary tree is a node which has 3 children. The number of all branching nodes in a tree T is denoted by $l(T)$. In Figure 3.1, \mathbf{r} and \mathbf{n} are branching nodes and $l(T) = 2$.
4. **Leaves and leaf edges:** A leaf of a tree T is a node which has no children. An edge connected to a leaf is called a leaf edge. In Figure 3.1, all \star nodes and \square nodes are leaves and the edges connected to them are leaf edges.
5. **Subtrees:** If any child of any node in a subset T' of a tree T is also contained in T' then T' also forms a tree, we call T' a subtree of T . If the root node of T' is $\mathbf{n} \in T$, we say T' is the subtree rooted at \mathbf{n} or subtree of \mathbf{n} and denote it by $T_{\mathbf{n}}$. In Figure 3.1, the tree inside the box is the subtree rooted at node \mathbf{n} .
6. **Node decoration:** In Figure 3.1, each node is associated with a symbol in $\{\bullet, \circ, \star, \square\}$. If a node \mathbf{n} has pattern \bullet (similarly \circ, \star, \square), we say \mathbf{n} is decorated by \bullet (\circ, \star, \square) or \mathbf{n} has decoration \bullet (\circ, \star, \square). In what follows, we

adopt the convention that leaves always have decoration \star or \square and nodes other than leaves always have decoration \bullet or \circ .

7. **Expanding and final leaves:** Leaves denoted by \square are called expanding leaves. Other leaves denoted by \star are called final leaves. The notion of expanding leaves is useful in the construction of trees, in which the presence of \square means that the construction is not finishing and \square denotes leaves that may be replaced by a branching node later.

The concept of expanding leaves and \square is only used in section 3.2.2, so the readers can safely forget it after that section.

8. **Leg:** The edge that is only connected to the root node is called the leg of the tree. In Figure 3.1, the edge on top of the tree is the leg.
9. **Nodes and children of an edge:** Let \mathbf{n}_u and \mathbf{n}_l be two endpoints of an edge ϵ and assume that \mathbf{n}_l is a children of \mathbf{n}_u . We define \mathbf{n}_u (resp. \mathbf{n}_l) to the upper node (resp. lower node) of ϵ . Let $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$ be the three nodes of \mathbf{n}_l and let ϵ_1 (resp. ϵ_2, ϵ_3) be the edge between two nodes \mathbf{n}_l and \mathbf{n}_1 (resp. $\mathbf{n}_2, \mathbf{n}_3$). $\epsilon_1, \epsilon_2, \epsilon_3$ are defined to be the three children edges of ϵ . We also define the primary, secondary and tertiary children in the same way as in Definition 3.2.1 (2).

For example, in Figure 3.1, if we take ϵ to be the edge between \mathbf{r} and \mathbf{n} , then $\mathbf{n}_u = \mathbf{r}$ (resp. $\mathbf{n}_l = \mathbf{n}$) are the upper node (resp. lower node of ϵ). $\epsilon_1, \epsilon_2, \epsilon_3$ are the three children edges of ϵ , which are primary, secondary and tertiary respectively.

10. **Direction of edges and conjugated tree:** In Figure 3.1, each edge is associated with a direction by the following rule.

The leg is chosen to point downwards (or upwards) to the root. The primary and tertiary children edges have the opposite direction as their parent, while

the secondary child edge has the same direction.

The tree in Figure 3.1 is an example of this rule. By default, the leg in most trees in this paper points downwards, except when we pair two trees in Definition 3.3.6. Two isomorphic trees with opposite direction association are called conjugated.

11. **Labelling of leaves:** Each leaves are labelled by $1, 2, \dots, 2l(T) + 1$ from left to right. An edge from a leaf \mathfrak{l} is also labeled by j if \mathfrak{l} is labeled by j . The tree in Figure 3.1 is an example of this labeling rule.
12. **Sign of edges and leaves:** Each edge \mathfrak{e} in a tree can be assigned with a sign $\iota_{\mathfrak{e}} \in \{\pm\}$ by following rule. If the edge points downwards (resp. upwards), then the edge is associated with a $+$ sign (a $-$ sign). The sign of a leaf equals the sign of the edge connected to it. For example, in Figure 3.1, leaves labelled by 1, 3, 5 are assigned with a $+$ sign, while leaves labelled by 2, 4 are assigned with a $-$ sign.

Connection between iteration and trees

In this section, we explain non-rigorously the connection between perturbation expansion and trees. The rigorous argument can be found in the next section.

For the ease of notation, we first assume that $\mathcal{F}(\phi)_k = \xi_k + \mathcal{T}(\phi, \phi, \phi)_k$. In this case, we do not need the decoration \circ . We will explain how to deal with two nonlinearity \mathcal{T}_1 and \mathcal{T}_2 later on. This iteration process can be described as the following,

$$\begin{aligned}
\phi &= \mathcal{F}(\phi) = \xi + \mathcal{T}(\phi, \phi, \phi) \\
&= \xi + \mathcal{T}\left(\xi + \mathcal{T}(\phi, \phi, \phi), \dots, \dots\right) = \xi + \mathcal{T}(\xi, \xi, \xi) + \mathcal{T}\left(\mathcal{T}(\phi, \phi, \phi), \xi, \xi\right) \cdots \\
&= \xi + \mathcal{T}(\xi, \xi, \xi) + \mathcal{T}(\mathcal{T}(\xi, \xi, \xi), \xi, \xi) + \mathcal{T}(\xi, \mathcal{T}(\xi, \xi, \xi), \xi) + \mathcal{T}(\xi, \xi, \mathcal{T}(\xi, \xi, \xi)) + \cdots
\end{aligned}$$

In the above iteration, we recursively choose one ϕ , replace it by $\xi + \mathcal{T}(\phi, \phi, \phi)$ and

use the linearity of \mathcal{T} to expand into two terms.

$$\begin{aligned}
& \mathcal{T}\left(\cdots, \mathcal{T}(\cdots, \mathcal{T}(\xi, \underline{\phi}, \cdots), \cdots), \cdots\right) \rightarrow \mathcal{T}\left(\cdots, \mathcal{T}(\cdots, \mathcal{T}(\xi, \underline{\xi + \mathcal{T}(\phi, \phi, \phi)}, \cdots), \cdots), \cdots\right) \\
& = \underbrace{\mathcal{T}\left(\cdots, \mathcal{T}(\cdots, \mathcal{T}(\xi, \underline{\xi}, \cdots), \cdots), \cdots\right)}_I + \underbrace{\mathcal{T}\left(\cdots, \mathcal{T}(\cdots, \mathcal{T}(\xi, \underline{\mathcal{T}(\phi, \phi, \phi)}, \cdots), \cdots), \cdots\right)}_{II}
\end{aligned} \tag{3.29}$$

Here I and II are obtained by replacing ϕ by ξ and $\mathcal{T}(\phi, \phi, \phi)$ respectively.

In summary, all terms in the expansion can be generated by following the steps

- **Step 0.** Add a term ϕ in the summation \mathcal{J} .
- **Step i ($i \geq 1$).** Assume that **Step $i - 1$** has been finished which produce a sum of terms \mathcal{J} , then choose a term in \mathcal{J} which has least number of ξ and ϕ , remove this term from \mathcal{J} and add the two terms in \mathcal{J} constructed in (3.29).

This process is very similar to the construction of ternary trees, in which we recursively replace a chosen branching node with a leaf or branch.

- **Step 0.** Start from a branching root node \square .
- **Step i ($i \geq 1$).** Assume that we have finish the **Step $i - 1$** which produce a collection of trees \mathcal{T} , then choose a tree in \mathcal{T} which has least number of expanding leaves \square and final leaves \star , remove this tree from \mathcal{T} and add two new trees in \mathcal{T} . In these two new trees, we replace an expanding leaf \square with a final leaf \star or a branched node \bullet with three branching children leaves \square . This construction is illustrated by Figure 3.2.

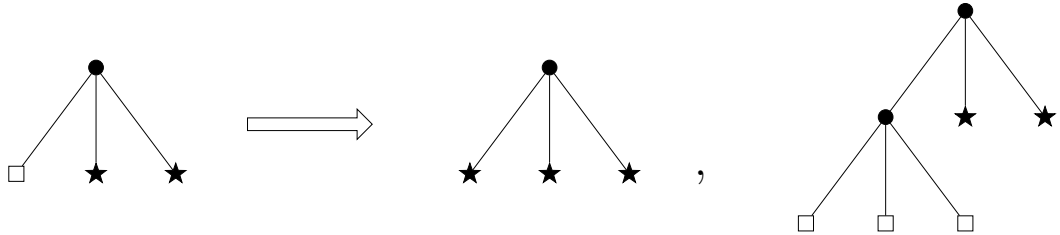


Figure 3.2: One step in the construction of binary trees

By comparing the above two processes, we can make the connection between terms and trees more explicit. Each node \bullet other than leaf in the tree T corresponds to a $\mathcal{T}(\cdots, \cdots, \cdots)$ in a term \mathcal{J}_T . Each final leaf \star and expanding leaf \square corresponds to ξ and ϕ respectively. The **Step** i of replacing ϕ by ξ or $\mathcal{T}(\phi, \phi, \phi)$ corresponds to replacing \square by \star or a branching node with three children \square .

We have the following recursive formula for calculating a term \mathcal{J}_T from a ternary tree T .

If T has only one node then $\mathcal{J}_T = \xi$. Otherwise let $\bullet_1, \bullet_2, \bullet_3$ be three children of the root node \bullet , let $T_{\bullet_1}, T_{\bullet_2}, T_{\bullet_3}$ be the subtrees of T rooted at above nodes. If $\mathcal{J}_{T_{\bullet_1}}, \mathcal{J}_{T_{\bullet_2}}, \mathcal{J}_{T_{\bullet_3}}$ have been recursively calculated, then \mathcal{J}_T can be calculated by

$$\mathcal{J}_T = \mathcal{T}(\mathcal{J}_{T_{\bullet_1}}, \mathcal{J}_{T_{\bullet_2}}, \mathcal{J}_{T_{\bullet_3}}). \quad (3.30)$$

The formal power series obtained by iterate $\phi = \mathcal{F}(\phi)$ can be calculated from trees by $\sum_{T \in \mathcal{T}} \mathcal{J}_T$.

Let $l(T)$ be the number of branches in T , then it can be shown that \mathcal{J}_T is a degree $2l(T) + 1$ polynomial of ξ . We define the approximation series to be a finite degree truncation of the formal power series which equals to $\sum_{l(T) \leq N} \mathcal{J}_T$.

Now we explain how to generalize the above ideas to the case in which two non-linearity \mathcal{T}_1 and \mathcal{T}_2 appear. Notice that in above argument, each node decorated by \bullet corresponds to a $\mathcal{T}(\cdots, \cdots, \cdots)$ in a term \mathcal{J}_T . In present case, we shall introduce two decoration \bullet, \circ correspond to \mathcal{T}_1 and \mathcal{T}_2 respectively. To adapt these changes, we should generalize the recursive formula (3.30) to (3.31) introduced in the next section.

Feynman diagrams and construction of the approximation solution

In this section we present the rigorous argument equivalent to that in the above section.

In the construction of trees, finally all \square nodes will be replaced by \bullet , \circ , \star , so in what follows, we only consider trees whose nodes are decorated by \bullet , \circ , \star .

Definition 3.2.2. Given a ternary tree T whose nodes are decorated by \bullet , \circ , \star , we inductively define the quantity \mathcal{J}_T by:

$$\mathcal{J}_T = \begin{cases} \xi, & \text{if } T \text{ has only one node } \star. \\ \mathcal{T}_1(\mathcal{J}_{T_{\mathfrak{n}_1}}, \mathcal{J}_{T_{\mathfrak{n}_2}}, \mathcal{J}_{T_{\mathfrak{n}_3}}), & \text{if the root node } \mathfrak{r} \text{ is decorated by } \bullet. \\ \mathcal{T}_2(\mathcal{J}_{T_{\mathfrak{n}_1}}, \mathcal{J}_{T_{\mathfrak{n}_2}}, \mathcal{J}_{T_{\mathfrak{n}_3}}), & \text{if the root node } \mathfrak{r} \text{ is decorated by } \circ. \end{cases} \quad (3.31)$$

Here $\mathfrak{n}_1, \mathfrak{n}_2, \mathfrak{n}_3$ are three children of the root node \mathfrak{r} and $T_{\mathfrak{n}_1}, T_{\mathfrak{n}_2}, T_{\mathfrak{n}_3}$ are the subtrees of T rooted at these nodes.

Definition 3.2.3. Given a large number N , define the approximate solution ϕ_{app} by

$$\phi_{app} = \sum_{l(T) \leq N} \mathcal{J}_T \quad (3.32)$$

Section 3.2.2 explains why the approximation series should equal to (3.32), a sum of many tree terms, but if we know this fact, we can directly prove it, and forget all the motivations. The lemma below proves that ϕ_{app} defined by the above expression is an approximate solution.

Lemma 3.2.4. *Define*

$$Err = \mathcal{F}(\phi_{app}) - \phi_{app}, \quad (3.33)$$

then we have

$$Err = \sum_{T \in \mathcal{T}_{>N}^*} \mathcal{J}_T, \quad (3.34)$$

where $\mathcal{T}_{>N}^*$ is defined by

$$\begin{aligned} \mathcal{T}_{>N}^* = \{T : l(T) > N, l(T_{\mathbf{n}_1}) \leq N, l(T_{\mathbf{n}_2}) \leq N, l(T_{\mathbf{n}_3}) \leq N, \\ T_{\mathbf{n}_1}, T_{\mathbf{n}_2}, T_{\mathbf{n}_3} \text{ are the subtrees defined in Definition 3.2.2}\} \end{aligned} \quad (3.35)$$

Remark 3.2.5. Notice that all terms in $\sum_{T \in \mathcal{T}_{>N}^*}$ are polynomials of degree $> N$. Therefore, the approximation error of ϕ_{app} is of very high order, which proves that ϕ_{app} is an appropriate approximation solution.

Proof. By (3.32), we get

$$\begin{aligned} Err &= \mathcal{F}(\phi_{app}) - \phi_{app} \\ &= \xi + \mathcal{T}_1(\phi_{app}, \phi_{app}, \phi_{app}) + \mathcal{T}_2(\phi_{app}, \phi_{app}, \phi_{app}) - \phi_{app} \\ &= \xi + \sum_{l(T_1), l(T_2), l(T_3) \leq N} \left(\mathcal{T}_1(\mathcal{J}_{T_1}, \mathcal{J}_{T_2}, \mathcal{J}_{T_3}) + \mathcal{T}_2(\mathcal{J}_{T_1}, \mathcal{J}_{T_2}, \mathcal{J}_{T_3}) \right) - \sum_{l(T) \leq N} \mathcal{J}_T \end{aligned} \quad (3.36)$$

Let T be a tree constructed by connecting the root nodes $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$ of T_1, T_2, T_3 to a new node \mathbf{r} . We define \mathbf{r} to be the root node of T .

Then by (3.31), we have

$$\mathcal{J}_T = \begin{cases} \mathcal{T}_1(\mathcal{J}_{T_1}, \mathcal{J}_{T_2}, \mathcal{J}_{T_3}), & \text{if } \mathbf{r} \text{ is decorated by } \bullet. \\ \mathcal{T}_2(\mathcal{J}_{T_1}, \mathcal{J}_{T_2}, \mathcal{J}_{T_3}), & \text{if } \mathbf{r} \text{ is decorated by } \circ. \end{cases} \quad (3.37)$$

and

$$\sum_{l(T_1), l(T_2), l(T_3) \leq N} \left(\mathcal{T}_1(\mathcal{J}_{T_1}, \mathcal{J}_{T_2}, \mathcal{J}_{T_3}) + \mathcal{T}_2(\mathcal{J}_{T_1}, \mathcal{J}_{T_2}, \mathcal{J}_{T_3}) \right) = \sum_{\substack{l(T) \geq 1 \\ l(T_1), l(T_2), l(T_3) \leq N}} \mathcal{J}_T \quad (3.38)$$

By (3.36), we get

$$\begin{aligned}
Err &= \xi + \sum_{\substack{l(T) \geq 1 \\ l(T_1), l(T_2), l(T_3) \leq N}} \mathcal{J}_T - \sum_{l(T) \leq N} \mathcal{J}_T \\
&= \sum_{\substack{T_1, T_2, T_3 \text{ are subtrees of } \mathfrak{r} \\ l(T_1), l(T_2), l(T_3) \leq N}} \mathcal{J}_T - \sum_{\substack{l(T) \leq N \\ l(T_1), l(T_2), l(T_3) \leq N}} \mathcal{J}_T. \\
&= \sum_{T \in \mathcal{T}_{>N}^*} \mathcal{J}_T
\end{aligned} \tag{3.39}$$

Here in the second equality, we use the fact that $\sum_{l(T) \leq N} = \sum_{\substack{l(T) \leq N \\ l(T_1), l(T_2), l(T_3) \leq N}}$.

Therefore, we complete the proof of this lemma. \square

3.3 Lattice points counting and convergence results

In this section, we give an incomplete proof of Proposition 3.3.27 which gives upper bounds for some tree terms $\mathcal{J}_{T,k}$. The proof is divided into several steps.

In section 3.3.1, we calculate the coefficients of $\mathcal{J}_{T,k}$ as polynomials of Gaussian random variables.

Large deviation theory suggests that an upper bound of a Gaussian polynomial can be derived from an upper bound of its expectation and variance.

In section 3.3.2, we introduce the concept of couples which is a graphical method of calculating the expectation of Gaussian polynomials. In this section, we also introduce and prove the formal version of our renormalized Wick theorem.

In section 3.3.3, we use couples to establish a lattice point counting result. This result says gives the upper bound of numbers of the lattice points near the resonance surface and this bound is strong enough to derive the wave kinetic equation up to epsilon loss, **if there is no long bad chain of circle nodes in the renormalization**

forest.

In section 3.3.4, we introduce a cancellation identity and explain **heuristically** why the Wick theorem is sufficient to exclude all bad couples.

3.3.1 Refined expression of coefficients

In this section, we calculate the coefficients of $\mathcal{J}_{T,k}$ using the definition (3.31) of them.

From (3.28) and (3.31), it is easy to show that $\mathcal{J}_{T,k}$ are polynomials of ξ if we remove the renormalization symbol $: \cdot :$. Given this renormalization symbol, we know that $\mathcal{J}_{T,k}$ should be a renormalized polynomial of ξ that will be defined in the following definition. An example of renormalized polynomial is $(: (\xi_{k_1} \bar{\xi}_{k_2} :) \xi_{k_3} \bar{\xi}_{k_4} :) \bar{\xi}_{k_5} \xi_{k_6} :$.

Definition 3.3.1. 1. **The renormalization symbol.** For any random variable X , define the renormalization symbol by $: X := X - \mathbb{E}X$.

2. **Some definitions about forests.** Given a forest F , a node \mathfrak{r} is a root if it does not have a parent. Notice that a forest may possess multiple roots. Given a node $\mathfrak{n} \in T$, its subtree $T_{\mathfrak{n}}$ is defined in the same way as the subtrees of trees. Assume that all the **non-leaf** children of \mathfrak{n} are $\mathfrak{n}_1, \dots, \mathfrak{n}_k$, then the subforest $F_{\mathfrak{n}}$ of \mathfrak{n} is defined as the union of all subtrees $T_{\mathfrak{n}_1}, \dots, T_{\mathfrak{n}_k}$. The set of leaves of $F_{\mathfrak{n}}$ (resp. $T_{\mathfrak{n}}$) is denoted by $L(F_{\mathfrak{n}})$ (resp. $L(T_{\mathfrak{n}})$). Each leaf of the forest is labeled by a number, called the label of this leaf. The example of roots, subtrees, and subforests can be found in Figure 3.3.

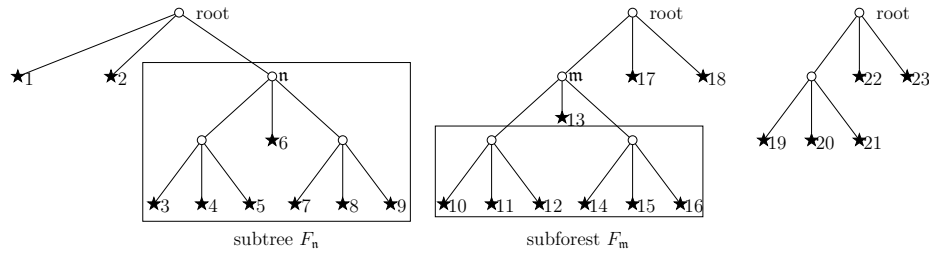


Figure 3.3: Example of roots, subtrees, and subforests

3. **Renormalization by a forest.** Given a forest F and a monomial $\xi_{k_1} \cdots \xi_{k_{2m}}$, we define $[\xi_{k_1} \cdots \xi_{k_{2m}}]_F$, the renormalization of $\xi_{k_1} \cdots \xi_{k_{2m}}$ by F , inductively. If $F = \emptyset$, we define $[\xi_{k_1} \cdots \xi_{k_{2m}}]_F = \xi_{k_1} \cdots \xi_{k_{2m}}$. If the root nodes of F are $\mathbf{r}_1, \dots, \mathbf{r}_r$ and their subforests are $F_{\mathbf{r}_1}, \dots, F_{\mathbf{r}_r}$, assume that the renormalizations by these subforests have been defined, then we define $[\xi_{k_1} \cdots \xi_{k_{2m}}]_F = \prod_{i \notin L(F)} \xi_{k_i} \prod_{j=1}^r : \left(\left[\prod_{i_j \in L(T_{\mathbf{r}_j})} \xi_{k_{i_j}} \right]_{F_{\mathbf{r}_j}} \right) :$, where the number $i_j \in L(T_{\mathbf{r}_j})$ if and only if the leaf labeled by i_j belongs to $L(T_{\mathbf{r}_j})$.

We say $[\xi_{k_1} \cdots \xi_{k_{2m}}]_F$ to be a renormalized monomial and a sum of renormalized monomial to be a renormalized polynomial.

4. **Renormalization forests.** Given a ternary tree T whose nodes are decorated by \bullet, \circ, \star , then the associated renormalization forest $R(T)$ of T is defined to be the unique forest determined by these conditions. First, it is formed by all \circ, \star nodes in T . Second, a node \mathbf{n} in $R(T)$ is a parent of another node \mathbf{n}' if and only if \mathbf{n} is the closest one in the ancestors of \mathbf{n}' that satisfies the property that the unique path from \mathbf{n} to \mathbf{n}' contains at least one primary or secondary child edge of a \circ node. (Actually, in this case, this child edge is unique and is at the beginning of this path.) Third, remove all nodes whose path to the root does not contain any primary or secondary child edge.

One example of this construction can be found in Figure 3.4.

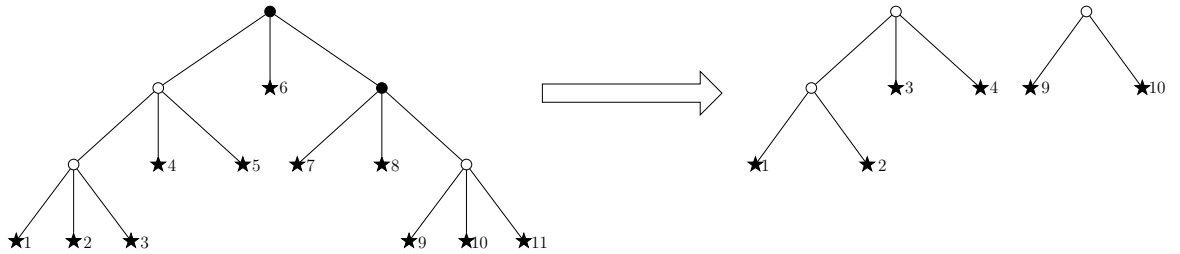


Figure 3.4: Construction of renormalization forests

Now we show that $\mathcal{J}_{T,k}$ is a renormalized polynomial and calculate the coefficients of $\mathcal{J}_{T,k}$.

Lemma 3.3.2. *Given a tree T of depth $l = l(T)$, denote by T_{in} the tree formed by all non-leaf nodes \mathbf{n} , then associate each node $\mathbf{n} \in T_{in}$ and edge $\mathbf{l} \in T$ with variables $t_{\mathbf{n}}$ and $k_{\mathbf{l}}$ respectively. Given a labelling of all leaves by $1, 2, \dots, 2l+1$, we identify $k_{\mathbf{e}}$ with k_j if \mathbf{e} connects a leaf labelled by j . Given a non-leaf node \mathbf{n} , let $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}$ be the four edges from or pointing to \mathbf{n} (\mathbf{e} is the parent of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$), $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$ be children of \mathbf{n} and $\hat{\mathbf{n}}$ be the parent of \mathbf{n} .*

Let \mathcal{J}_T be terms defined in Definition 3.2.2, then their Fourier coefficients $\mathcal{J}_{T,k}$ are degree $2l+1$ renormalized polynomials of ξ given by the following formula

$$\mathcal{J}_{T,k} = \left(\frac{-i\lambda^2}{L^{2d}} \right)^l \sum_{k_1, k_2, \dots, k_{2l+1}} H_{k_1 \dots k_{2l+1}}^T [\xi_{k_1} \bar{\xi}_{k_2} \dots \xi_{k_{2l+1}}]_{R(T)} \quad (3.40)$$

where $H_{k_1 \dots k_{2l+1}}^T$ is given by

$$H_{k_1 \dots k_{2l+1}}^T = \int_{\cup_{\mathbf{n} \in T_{in}} A_{\mathbf{n}}} e^{-i \sum_{\mathbf{n} \in T_{in}} (t_{\mathbf{n}} \Omega_{\mathbf{n}} + \tilde{\Omega}_{\mathbf{n}})} \prod_{\mathbf{n} \in T_{in}} dt_{\mathbf{n}} \delta_{\cap_{\mathbf{n} \in T_{in}} S_{\mathbf{n}}} \prod \Lambda^{-1}(k_{\mathbf{e}}), \quad (3.41)$$

and $\iota_{\mathbf{e}}$ is defined by Definition 3.2.1 (12), $A_{\mathbf{n}}, S_{\mathbf{n}}, \Omega_{\mathbf{n}}, \tilde{\Omega}_{\mathbf{n}}$ are defined by

$$A_{\mathbf{n}} = \begin{cases} \{t_{\mathbf{n}_1}, t_{\mathbf{n}_2}, t_{\mathbf{n}_3} \leq t_{\mathbf{n}}\} & \text{if } \mathbf{n} \neq \text{the root } \mathbf{r} \\ \{t_{\mathbf{r}} \leq t\} & \text{if } \mathbf{n} = \mathbf{r} \end{cases} \quad (3.42)$$

$$S_{\mathbf{n}} = \begin{cases} \{k_{\mathbf{e}_1} - k_{\mathbf{e}_2} + k_{\mathbf{e}_3} - k_{\mathbf{e}} = 0, k_{\mathbf{e}_1} \neq k_{\mathbf{e}_2} \neq k_{\mathbf{e}_3} & \text{if } \mathbf{n} \text{ is decorated by } \bullet \\ \text{or } k_{\mathbf{e}_1} = k_{\mathbf{e}_2} = k_{\mathbf{e}_3} = k\}, \\ \{k_{\mathbf{e}_1} = k_{\mathbf{e}_2}, k_{\mathbf{e}_3} = k_{\mathbf{e}}\} & \text{if } \mathbf{n} \text{ is decorated by } \circ \end{cases} \quad (3.43)$$

$$\Omega_{\mathbf{n}} = \begin{cases} \iota_{\mathbf{e}_1} \Lambda_{k_{\mathbf{e}_1}} + \iota_{\mathbf{e}_2} \Lambda_{k_{\mathbf{e}_2}} + \iota_{\mathbf{e}_3} \Lambda_{k_{\mathbf{e}_3}} + \iota_{\mathbf{e}} \Lambda_{k_{\mathbf{e}}} & \text{if } \mathbf{n} \text{ is decorated by } \bullet \\ 0 & \text{if } \mathbf{n} \text{ is decorated by } \circ \end{cases} \quad (3.44)$$

$$\tilde{\Omega}_{\mathbf{n}} = \begin{cases} 2\alpha \int_0^t m(s) ds \left(\iota_{\mathfrak{e}_1} \Lambda_{k_{\mathfrak{e}_1}}^{-1} + \iota_{\mathfrak{e}_2} \Lambda_{k_{\mathfrak{e}_2}}^{-1} + \iota_{\mathfrak{e}_3} \Lambda_{k_{\mathfrak{e}_3}}^{-1} + \iota_{\mathfrak{e}} \Lambda_{k_{\mathfrak{e}}}^{-1} \right) & \text{if } \mathbf{n} \text{ is decorated by } \bullet \\ 0 & \text{if } \mathbf{n} \text{ is decorated by } \circ \end{cases} \quad (3.45)$$

For root node \mathfrak{r} , we impose the constrain that $k_{\mathfrak{r}} = k$ and $t_{\widehat{\mathfrak{r}}} = t$ (notice that \mathfrak{r} does not have a parent so $\widehat{\mathfrak{r}}$ is not well defined).

Proof. We can check that \mathcal{J}_T defined by (3.40) and (3.41) satisfies the recursive formula (3.31) by direct substitution, so they are the unique solution of that recursive formula, and this proves Lemma 3.3.2. \square

The following table provides a dictionary between concepts in Definition 3.2.1 in the previous section and expressions in tree terms $\mathcal{J}_{T,k}$ in Lemma 3.3.2.

Concepts	Corresponding expressions
Ternary tree T	Tree term $\mathcal{J}_{T,k}$
Edge \mathfrak{e}	Summation index $k_{\mathfrak{e}}$
Leaf edges \mathfrak{l} labelled by j	Summation index k_j
\star node \mathfrak{n}	Two expressions $\widetilde{\Omega}_{\mathfrak{n}}$ and $k_{\mathfrak{e}_1} - k_{\mathfrak{e}_2} + k_{\mathfrak{e}_3} - k_{\mathfrak{e}} = 0$
\circ node \mathfrak{n}	Expression $k_{\mathfrak{e}_1} = k_{\mathfrak{e}_2}, k_{\mathfrak{e}_3} = k_{\mathfrak{e}}$
Leaf $\mathfrak{n}_{\mathfrak{l}}$ labelled by j with a sign	ξ_{k_j} or $\bar{\xi}_{k_j}$ depending on the sign
Direction of edges	$\iota_{\mathfrak{e}}$
Leg	The fixed index k
Conjugated tree \overline{T}	146 Tree term $\overline{\mathcal{J}_{T,k}}$

Lemma 3.3.2 suggests that the coefficients is small when $|\Omega_{\mathbf{n}}| \gg \omega$, ω is supposed to be $\frac{1}{T_{\text{kin}}}$. Therefore, in order to bound $\mathcal{J}_{T,k}$, we should count the lattice points on $|\Omega_{\mathbf{n}}| \lesssim \omega$

$$\{k_{\mathfrak{e}} \in \mathbb{Z}^d, |k_{\mathfrak{e}}| \lesssim 1, \forall \mathfrak{e} : |\Omega_{\mathbf{n}}| \lesssim \omega, \forall \mathbf{n}. \{k_{\mathfrak{e}}\}_{\mathfrak{e}} \in \cap_{\mathbf{n} \in T_{\text{in}}} S_{\mathbf{n}}. k_{\mathfrak{l}} = k\} \quad (3.46)$$

(3.46) can be read from the tree diagrams T . As in Figure 3.5, each edge corresponds to a variable $k_{\mathfrak{e}}$. The leg \mathfrak{l} corresponds to equation $k_{\mathfrak{l}} = k$. Each node \mathbf{n} is connected with four edges $\mathfrak{e}_1, \mathfrak{e}_2, \mathfrak{e}_3, \mathfrak{e}$ whose corresponding variables $k_{\mathfrak{e}_1}, k_{\mathfrak{e}_2}, k_{\mathfrak{e}_3}, k_{\mathfrak{e}}$ satisfy the momentum conservation equation

$$k_{\mathfrak{e}_1} - k_{\mathfrak{e}_2} + k_{\mathfrak{e}_3} - k_{\mathfrak{e}} = 0, \quad k_{\mathfrak{e}_1} \neq k_{\mathfrak{e}_2} \neq k_{\mathfrak{e}_3}, \quad \text{or} \quad k_{\mathfrak{e}_1} = k_{\mathfrak{e}_2} = k_{\mathfrak{e}_3} = k_{\mathfrak{e}} \quad (3.47)$$

and the energy conservation equation (if the node is decorated by \bullet or \circ)

$$\begin{aligned} \Lambda_{k_{\mathfrak{e}_1}} - \Lambda_{k_{\mathfrak{e}_2}} + \Lambda_{k_{\mathfrak{e}_3}} - \Lambda_{k_{\mathfrak{e}}} &= 0. \\ \text{or } k_{\mathfrak{e}_1} &= k_{\mathfrak{e}_2}, \quad k_{\mathfrak{e}_3} = k_{\mathfrak{e}} \quad (\text{depending on decoration of } \mathbf{n}) \end{aligned} \quad (3.48)$$

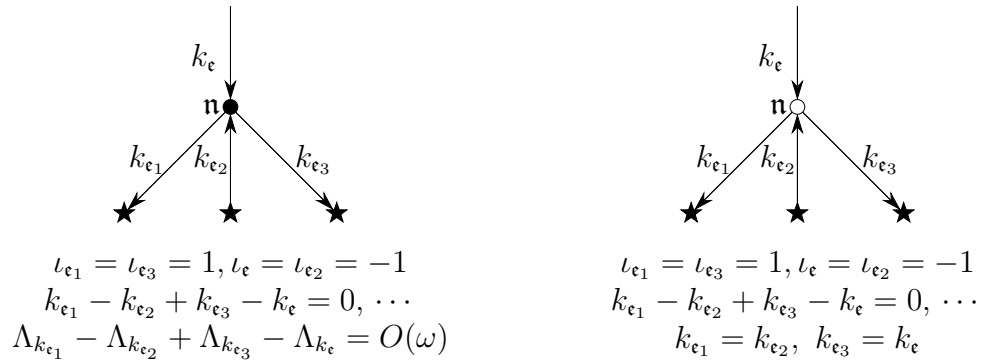


Figure 3.5: Equations of a node \mathbf{n}

3.3.2 Couples and the renormalized Wick theorem

In this section, we introduce the renormalized Wick theorem and calculate $\mathbb{E}|\mathcal{J}_{T,k}|^2$ using this theorem. We also introduce another type of diagram, the couple diagrams, to represent the result.

By the upper bound in the last section, the coefficients $H_{k_1 \dots k_{2l+1}}^T$ concentrate near the surface $\Omega_{\mathbf{n}} = 0$, $\forall \mathbf{n}$. But to get an upper bound of $\mathcal{J}_{T,k}$, we need the upper bound of their variance $\mathbb{E}|\mathcal{J}_{T,k}|^2$. The coefficients of $\mathbb{E}|\mathcal{J}_{T,k}|^2$ also concentrate near a surface whose expression is similar to (3.46).

Let's derive the expression of the coefficients of $\mathbb{E}|\mathcal{J}_{T,k}|^2$ and its concentration surface.

By Lemma 3.3.2, we know that $\mathcal{J}_{T,k}$ is a polynomial of ξ which are proportional to i.i.d Gaussians. Therefore,

$$\begin{aligned} \mathbb{E}|\mathcal{J}_{T,k}|^2 &= \mathbb{E}(\mathcal{J}_{T,k} \overline{\mathcal{J}_{T,k}}) = \left(\frac{\lambda^2}{L^{2d}} \right)^{2l(T)} \sum_{k_1, k_2, \dots, k_{2l(T)+1}} \sum_{k'_1, k'_2, \dots, k'_{2l(T)+1}} \\ &\quad H_{k_1 \dots k_{2l(T)+1}}^T \overline{H_{k'_1 \dots k'_{2l(T)+1}}^T} \mathbb{E} \left([\xi_{k_1} \bar{\xi}_{k_2} \dots \xi_{k_{2l(T)+1}}]_{R(T)} [\bar{\xi}_{k'_1} \xi_{k'_2} \dots \bar{\xi}_{k'_{2l(T)+1}}]_{R(T)} \right) \end{aligned} \quad (3.49)$$

We just need to calculate

$$\mathbb{E} \left([\xi_{k_1} \bar{\xi}_{k_2} \dots \xi_{k_{2l(T)+1}}]_{R(T)} [\bar{\xi}_{k'_1} \xi_{k'_2} \dots \bar{\xi}_{k'_{2l(T)+1}}]_{R(T)} \right). \quad (3.50)$$

Notice that $\xi_k = \sqrt{n_{\text{in}}(k)} \eta_k(\omega)$ and η_k are i.i.d Gaussians. We can apply the Wick theorem to calculate above expectations.

To introduce the renormalized Wick theorem, we need the following definition.

Definition 3.3.3. 1. **Sign map.** Given a set A , a map $\tau : A \rightarrow \{-1, 1\}$ is said

to be a sign map of A . $\tau(a) \in \{-1, 1\}$ is said to be the sign of a .

2. **Balanced set.** A set A with sign map τ is balanced if A has equal amount of elements of positive ($\tau(a) = 1$) and negative sign ($\tau(a) = -1$).
3. **Pairing.** Suppose that we have a balanced set $A = \{a_1, \dots, a_m, b_1, \dots, b_m\}$. Let a_1, \dots, a_m (resp. b_1, \dots, b_m) be elements that have a positive sign (resp. negative sign). A balanced pairing or simply balanced pairing is a partition of $A = \{a_{i_1}, b_{j_1}\} \cup \dots \cup \{a_{i_m}, b_{j_m}\}$ into m subsets which have exactly two elements of different sign. Given a pairing p , elements a_{i_k}, b_{j_k} in the same subset of p are called paired with each other, which is denoted by $a_{i_k} \sim_p b_{j_k}$.
4. **The pairing sets $\mathcal{P}(A)$ and $\mathcal{P}(B, A)$.** Denote by $\mathcal{P}(A)$ the set of all pairings of A . If B is a subset of A , let $\mathcal{P}(B, A) = \{p \in \mathcal{P}(A) : \forall a \in B, b \sim_p a \Rightarrow b \in B\}$, i.e. set of pairings that pair elements in B with elements in B .
5. **The renormalized pairing sets $\mathcal{P}_F(A)$.** Given a forest F and a set A that contains the leaf set $L(F)$ as its subset, we define the renormalized pairing set $\mathcal{P}_F(A)$ by $\mathcal{P}_F(A) = \{p \in \mathcal{P} : \forall \mathbf{n} \in F, p \notin \mathcal{P}(L(T_{\mathbf{n}}), A)\}$. $\mathcal{P}_F(A)$ is the set of pairings that do not pair all leaves of any subtree with leaves of the same subtree.
6. **Concatenation of two pairings.** Given two pairing $p_1 \in \mathcal{P}(A_1)$ and $p_2 \in \mathcal{P}(A_2)$, (see Definition 3.3.3 (4) for the definition of $\mathcal{P}(A, B)$) we can define their concatenation $p_1 \vee p_2 \in \mathcal{P}(A_1) \cup \mathcal{P}(A_2)$ as the following.
 Assume that $A_1 = \{a_1^{(1)}, \dots, a_m^{(1)}, b_1^{(1)}, \dots, b_m^{(1)}\}$, $A_2 = \{a_1^{(2)}, \dots, a_m^{(2)}, b_1^{(2)}, \dots, b_m^{(2)}\}$, p_1 is given by the partition $\{a_{i_1}^{(1)}, b_{j_1}^{(1)}\} \cup \dots \cup \{a_{i_m}^{(1)}, b_{j_m}^{(1)}\}$ and p_2 is given by the partition $\{a_{i_1}^{(2)}, b_{j_1}^{(2)}\} \cup \dots \cup \{a_{i_m}^{(2)}, b_{j_m}^{(2)}\}$, then $p_1 \vee p_2$ is given by the partition $\{a_{i_1}^{(1)}, b_{j_1}^{(1)}\} \cup \dots \cup \{a_{i_m}^{(1)}, b_{j_m}^{(1)}\} \cup \{a_{i_1}^{(2)}, b_{j_1}^{(2)}\} \cup \dots \cup \{a_{i_m}^{(2)}, b_{j_m}^{(2)}\}$.
 Similarly given pairings p_1, \dots, p_n , we can define the concatenation $p_1 \vee p_2 \vee \dots \vee p_n$

The original Wick theorem is important in the proof of the renormalized Wick theorem.

Theorem 3.3.4 (Wick theorem). *Let $\{\eta_k\}_{k \in \mathbb{Z}^d}$ be i.i.d complex Gaussian random variable. Assume that τ is a sign map of a balanced set $A = \{k_1, k_2, \dots, k_{2m}\}$. Define $\tau_j = \tau(k_j)$. Let $\eta_{k_j}^{\tau_j} = \eta_{k_j}$ if $\tau_j = 1$ and $\eta_{k_j}^{\tau_j} = \bar{\eta}_{k_j}$ if $\tau_j = -1$. Let \mathcal{P} be the set of all pairings of $\{k_1, k_2, \dots, k_{2m}\}$, then*

$$\mathbb{E}(\eta_{k_1}^{\tau_1} \cdots \eta_{k_{2m}}^{\tau_{2m}}) = \sum_{p \in \mathcal{P}} \delta_p(k_1, \dots, k_{2m}), \quad (3.51)$$

where

$$\delta_p = \begin{cases} 1 & \text{if } k_i = k_j \text{ for all } k_i \sim_p k_j, \\ 0 & \text{otherwise.} \end{cases} \quad (3.52)$$

Proof. This Wick theorem is a direct corollary of the complex Isserlis' theorem proved in Lemma A.2 of [9]. □

Theorem 3.3.5 (Renormalized Wick theorem). *In addition to assumptions in Theorem 3.3.4, define $\mathcal{P}_F = \mathcal{P}_F(A)$ (remember that $A = \{k_1, k_2, \dots, k_{2m}\}$ and $\mathcal{P}_F(A)$ is defined in Definition 3.3.3 (5)), then*

$$\mathbb{E}([\eta_{k_1}^{\tau_1} \cdots \eta_{k_{2m}}^{\tau_{2m}}]_F) = \sum_{p \in \mathcal{P}_F} \delta_p(k_1, \dots, k_{2m}) \quad (3.53)$$

Proof. Before starting the formal proof, we give proof for an example to demonstrate the idea.

Proof for a special case. Consider the following forest F in Figure 3.6.

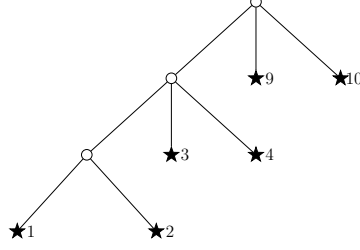


Figure 3.6: An example of forests

Let us calculate $\mathbb{E}([\xi_{k_1} \bar{\xi}_{k_2} \xi_{k_3} \bar{\xi}_{k_4} \bar{\xi}_{k_5} \xi_{k_6}])_F = \mathbb{E}((: (\xi_{k_1} \bar{\xi}_{k_2} :) \xi_{k_3} \bar{\xi}_{k_4} :) \bar{\xi}_{k_5} \xi_{k_6})$.

Expanding the renormalization symbol gives

$$\begin{aligned}
& \mathbb{E}((: (\xi_{k_1} \bar{\xi}_{k_2} :) \xi_{k_3} \bar{\xi}_{k_4} :) \bar{\xi}_{k_5} \xi_{k_6}) \\
&= \mathbb{E}(\xi_{k_1} \bar{\xi}_{k_2} \xi_{k_3} \bar{\xi}_{k_4} \bar{\xi}_{k_5} \xi_{k_6}) - \mathbb{E}(\xi_{k_1} \bar{\xi}_{k_2}) \mathbb{E}(\xi_{k_3} \bar{\xi}_{k_4} \bar{\xi}_{k_5} \xi_{k_6}) - \\
& \quad \mathbb{E}(\xi_{k_1} \bar{\xi}_{k_2} \xi_{k_3} \bar{\xi}_{k_4}) \mathbb{E}(\bar{\xi}_{k_5} \xi_{k_6}) + \mathbb{E}(\xi_{k_1} \bar{\xi}_{k_2}) \mathbb{E}(\xi_{k_3} \bar{\xi}_{k_4}) \mathbb{E}(\bar{\xi}_{k_5} \xi_{k_6}) \\
&= \sum_{\text{all pairings}} \delta_p - \sum_{A=\{p \text{ pairs } (12) \text{ and } (3456)\}} \delta_p - \sum_{B=\{p \text{ pairs } (1234) \text{ and } (56)\}} \delta_p + \sum_{A \cap B} \delta_p \\
&= \sum_{A^c \cap B^c} \delta_p = \sum_{p \in \mathcal{P}_F} \delta_p
\end{aligned} \tag{3.54}$$

Here in the second equality, we have applied the Theorem 3.3.4 and the inclusion-exclusion principle. The last equality follows from the definition of \mathcal{P}_F .

Proof for the general case. The proof for the general case also uses an analogous inclusion-exclusion argument. Since renormalization is defined by recursion in Definition 3.3.1 (3), we will prove the theorem by induction.

Step 1. (The induction assumption) For the ease of notation let $\eta_{k_j} = \eta_{k_j}^{\tau_j}$. Let $A = \{k_1, k_2, \dots, k_{2m}\}$. Then we need to show that $\mathbb{E}([\eta_{k_1} \cdots \eta_{k_{2m}}]_F) = \sum_{p \in \mathcal{P}_F} \delta_p$.

We prove the theorem by induction on $\text{Depth}(F)$.

If $\text{Depth}(F) = 0$, i.e. $F = \emptyset$, then $\mathbb{E}([\eta_{k_1} \cdots \eta_{k_{2m}}]_F) = \mathbb{E}(\eta_{k_1} \cdots \eta_{k_{2m}})$ and the theorem follows from the standard Wick theorem (Theorem 3.3.4).

Assuming that the theorem is true for $\text{Depth}(F) < l$, we show that it is true for $\text{Depth}(F) = l$ in the next steps.

Step 2. (Expanding the product) Let $\{\mathfrak{r}_1, \dots, \mathfrak{r}_r\}$ be roots of F and $F_{\mathfrak{r}_1}, \dots, F_{\mathfrak{r}_r}$ be their subforests, then by Definition 3.3.1 (3)

$$\mathbb{E}([\eta_{k_1} \cdots \eta_{k_{2m}}]_F) = \mathbb{E} \left(\prod_{k_i \notin L(F)} \eta_{k_i} \prod_{j=1}^r \left(\left[\prod_{k_{i_j} \in L(T_{\mathfrak{r}_j})} \eta_{k_{i_j}} \right]_{F_{\mathfrak{r}_j}} - \mathbb{E} \left[\prod_{k_{i_j} \in L(T_{\mathfrak{r}_j})} \eta_{k_{i_j}} \right]_{F_{\mathfrak{r}_j}} \right) \right) \quad (3.55)$$

Expanding the product $\prod_{j=1}^r$ gives

$$\begin{aligned} & \mathbb{E}([\eta_{k_1} \cdots \eta_{k_{2m}}]_F) \\ &= \mathbb{E} \left(\sum_{S \subseteq \{1, \dots, r\}} (-1)^{|S|} \prod_{k_i \notin L(F)} \eta_{k_i} \prod_{j' \notin S} \left[\prod_{k_{i_{j'}} \in L(T_{\mathfrak{r}_{j'}})} \eta_{k_{i_{j'}}} \right]_{F_{\mathfrak{r}_{j'}}} \prod_{j \in S} \mathbb{E} \left[\prod_{k_{i_j} \in L(T_{\mathfrak{r}_j})} \eta_{k_{i_j}} \right]_{F_{\mathfrak{r}_j}} \right) \\ &= \sum_{S \subseteq \{1, \dots, r\}} (-1)^{|S|} \underbrace{\mathbb{E} \left(\prod_{k_i \notin L(F)} \eta_{k_i} \prod_{j' \notin S} \left[\prod_{k_{i_{j'}} \in L(T_{\mathfrak{r}_{j'}})} \eta_{k_{i_{j'}}} \right]_{F_{\mathfrak{r}_{j'}}} \right)}_E \underbrace{\prod_{j \in S} \mathbb{E} \left[\prod_{k_{i_j} \in L(T_{\mathfrak{r}_j})} \eta_{k_{i_j}} \right]_{F_{\mathfrak{r}_j}}}_{E_{F_{\mathfrak{r}_j}}} \end{aligned} \quad (3.56)$$

Here in the second line, we have used the identity

$$\prod_{j=1}^r (a_j - b_j) = \sum_{S \subseteq \{1, \dots, r\}} (-1)^{|S|} \prod_{j \in S} a_j \prod_{j' \notin S} b_{j'} \quad (3.57)$$

Step 3. (Applying the induction assumption) Notice that E in the last line of (3.56) is the expectation of the polynomial $\prod_{k_i \notin L(F) \text{ or } k_i \in \cup_{j' \notin S} L(T_{\mathfrak{r}_{j'}})} \eta_{k_i}$ renormalized by the forest $\cup_{j' \notin S} F_{\mathfrak{r}_{j'}}$. $E_{F_{\mathfrak{r}_j}}$ is the expectation of the polynomial $\prod_{k_{i_j} \in L(T_{\mathfrak{r}_j})} \eta_{k_{i_j}}$ renormalized by the forest $F_{\mathfrak{r}_j}$.

In other words,

$$E = \mathbb{E} \left(\left[\prod_{\substack{k_i \notin L(F) \text{ or } k_i \in \cup_{j' \notin S} L(T_{\tau_j})}} \eta_{k_i} \right]_{\cup_{j' \notin S} F_{\tau_j}} \right), \quad E_{F_{\tau_j}} = \mathbb{E} \left(\left[\prod_{k_{i_j} \in L(T_{\tau_j})} \eta_{k_{i_j}} \right]_{F_{\tau_j}} \right) \quad (3.58)$$

Since $\text{Depth}(F_{\tau_j}) < l$ and $\text{Depth}(\cup_{j' \notin S} F_{\tau_j}) < l$ for any j , we can thus apply the induction assumption to calculate E and $E_{F_{\tau_j}}$.

$$\begin{aligned} E &= \sum_{p \in \mathcal{P}_{\cup_{j' \notin S} F_{\tau_{j'}}}(\{k_i \notin L(F) \text{ or } k_i \in \cup_{j' \notin S} L(T_{\tau_j})\})} \delta_p \\ &= \sum_{p \in \mathcal{P}_{\cup_{j' \notin S} F_{\tau_{j'}}}((\cup_{j \in S} L(T_{\tau_j}))^c)} \delta_p, \end{aligned} \quad (3.59)$$

and

$$E_{F_{\tau_j}} = \prod_{j \in S} \sum_{q \in \mathcal{P}_{F_{\tau_j}}(T_{\tau_j})} \delta_{q_j}, \quad (3.60)$$

where in (3.59), we have used the fact that $\{k_i \notin L(F) \text{ or } k_i \in \cup_{j' \notin S} L(T_{\tau_j})\} = (\cup_{j \in S} L(T_{\tau_j}))^c$, and in (3.59) and (3.60), we have used the notation $\mathcal{P}_F(B)$ in Definition 3.3.3 (5). In (3.59), $\mathcal{P}_{\cup_{j' \notin S} F_{\tau_{j'}}}((\cup_{j \in S} L(T_{\tau_j}))^c)$ is of the form $\mathcal{P}_F(B)$, in which $F = \cup_{j' \notin S} F_{\tau_{j'}}$ and $B = (\cup_{j \in S} L(T_{\tau_j}))^c$.

Substituting (3.59) and (3.60) into (3.56), we get

$$\mathbb{E}([\eta_{k_1} \cdots \eta_{k_{2m}}]_F) = \sum_{S \subseteq \{1, \dots, r\}} (-1)^{|S|} \left(\sum_{p \in \mathcal{P}_{\cup_{j' \notin S} F_{\tau_{j'}}}((\cup_{j \in S} L(T_{\tau_j}))^c)} \delta_p \right) \left(\prod_{j \in S} \sum_{q \in \mathcal{P}_{F_{\tau_j}}(T_{\tau_j})} \delta_{q_j} \right) \quad (3.61)$$

Step 4. (Concatenating short pairings) Expanding the product, we get

$$\begin{aligned}
\mathbb{E}([\eta_{k_1} \cdots \eta_{k_{2m}}]_F) &= \sum_{S \subseteq \{1, \dots, r\}} (-1)^{|S|} \left(\sum_{p \in \mathcal{P}_{\cup_{j' \notin S} F_{\tau_{j'}}}((\cup_{j \in S} L(T_{\tau_j}))^c)} \delta_p \right) \left(\prod_{j \in S} \sum_{q \in \mathcal{P}_{F_{\tau_j}}(T_{\tau_j})} \delta_{q_j} \right) \\
&= \sum_{S \subseteq \{1, \dots, r\}} (-1)^{|S|} \sum_{\substack{p \in \mathcal{P}_{\cup_{j' \notin S} F_{\tau_{j'}}}((\cup_{j \in S} L(T_{\tau_j}))^c) \\ q_j \in \mathcal{P}_{F_{\tau_j}}(T_{\tau_j}) \ \forall j}} \left(\delta_p \prod_{j \in S} \delta_{q_j} \right)
\end{aligned} \tag{3.62}$$

For any tree or forest T , define $\mathcal{P}[T] = \mathcal{P}(L(T), A)$ (see Definition 3.3.3 (4) for the definition of $\mathcal{P}(A, B)$), where $A = \{k_1, k_2, \dots, k_{2m}\}$. For trees or forests $\{T_j\}_{j \in S}$, define $\mathcal{P}[T_j, j \in S] = \cap_{j \in S} \mathcal{P}[L(F_j)]$.

Claim. Given $p \in \mathcal{P}_{\cup_{j' \notin S} F_{\tau_{j'}}}((\cup_{j \in S} L(T_{\tau_j}))^c)$ and $q_j \in \mathcal{P}_{F_{\tau_j}}(T_{\tau_j}) \ \forall j$, then their concatenation satisfies $p \vee (\vee_{j \in S} q_j) \in \mathcal{P}[T_j, j \in S] \cap \mathcal{P}_{\cup_{j=1}^r F_{\tau_j}}$. Conversely, all pairings in $p \vee (\vee_{j \in S} q_j) \in \mathcal{P}[T_j, j \in S] \cap \mathcal{P}_{\cup_{j=1}^r F_{\tau_j}}$ can be generated in this way.

Proof of the claim. $p \in \mathcal{P}_{\cup_{j' \notin S} F_{\tau_{j'}}}((\cup_{j \in S} L(T_{\tau_j}))^c)$ means that p does not pair all leaves of any subtree of $\cup_{j' \notin S} F_{\tau_{j'}}$ with leaves of the same subtree. $q_j \in \mathcal{P}_{F_{\tau_j}}(T_{\tau_j}) \ \forall j$ means that q_j does not pair all leaves of any subtree of F_{τ_j} with leaves of the same subtree. Therefore, $p \vee (\vee_{j \in S} q_j)$ does not pair all leaves of any subtree of $\cup_{j=1}^r F_{\tau_j}$ with leaves of the same subtree, which implies that $\mathcal{P}_{\cup_{j=1}^r F_{\tau_j}}$.

It is not difficult to show that $p \vee (\vee_{j \in S} q_j) \in \mathcal{P}[T_j, j \in S]$, so we have $p \vee (\vee_{j \in S} q_j) \in \mathcal{P}[T_j, j \in S] \cap \mathcal{P}_{\cup_{j=1}^r F_{\tau_j}}$.

Given a pairing $p' \in \mathcal{P}[T_j, j \in S] \cap \mathcal{P}_{\cup_{j=1}^r F_{\tau_j}}$, we can restrict p' to subsets and define $p = p'|_{(\cup_{j \in S} L(T_{\tau_j}))^c}$ and $q_j = p|_{\cup_{j \in S} L(T_{\tau_j})}$. It is not hard to show that $p' = p \vee (\vee_{j \in S} q_j)$, $q_j \in \mathcal{P}_{F_{\tau_j}}(T_{\tau_j}) \ \forall j$ and $p \in \mathcal{P}_{\cup_{j' \notin S} F_{\tau_{j'}}}((\cup_{j \in S} L(T_{\tau_j}))^c)$.

Therefore, we have finished the proof of the claim. \square

From above claim and (3.62), we know that

$$\mathbb{E}([\eta_{k_1} \cdots \eta_{k_{2m}}]_F) = \sum_{S \subseteq \{1, \dots, r\}} (-1)^{|S|} \sum_{p \vee (\vee_{j \in S} q_j) \in \mathcal{P}[T_j, j \in S] \cap \mathcal{P}_{\cup_{j=1}^r F_{\mathbf{r}_j}}} \delta_p \quad (3.63)$$

Step 5. (The inclusion-exclusion principle) Now we apply the following inclusion-exclusion principle.

$$\mathbb{1}_{A_1^c \cap \dots \cap A_m^c} = \sum_{S \subseteq \{1, \dots, r\}} (-1)^{|S|} \mathbb{1}_{\cap_{j \in S} A_j} \quad (3.64)$$

Since $\mathcal{P}[T_j, j \in S] = \cap_{j \in S} \mathcal{P}[T_j]$, we get

$$\begin{aligned} \mathbb{E}([\eta_{k_1} \cdots \eta_{k_{2m}}]_F) &= \sum_{S \subseteq \{1, \dots, r\}} (-1)^{|S|} \sum_p \mathbb{1}_{\cap_{j \in S} \mathcal{P}[T_j] \cap \mathcal{P}_{\cup_{j=1}^r F_{\mathbf{r}_j}}} \delta_p \\ &= \sum_p \left(\sum_{S \subseteq \{1, \dots, r\}} (-1)^{|S|} \mathbb{1}_{\cap_{j \in S} \mathcal{P}[T_j] \cap \mathcal{P}_{\cup_{j=1}^r F_{\mathbf{r}_j}}} \right) \delta_p \\ &= \sum_p \left(\sum_{S \subseteq \{1, \dots, r\}} (-1)^{|S|} \mathbb{1}_{\cap_{j \in S} \mathcal{P}[T_j]} \right) \mathbb{1}_{\mathcal{P}_{\cup_{j=1}^r F_{\mathbf{r}_j}}} \delta_p \\ &= \sum_p \mathbb{1}_{\cap_{j \in S} (\mathcal{P}[T_j])^c \cap \mathcal{P}_{\cup_{j=1}^r F_{\mathbf{r}_j}}} \delta_p = \sum_p \mathbb{1}_{\mathcal{P}_F} \delta_p \\ &= \sum_{p \in \mathcal{P}_F} \delta_p \end{aligned} \quad (3.65)$$

Here in the fourth line we have used (3.64) and the following identity

$$\mathcal{P}_F = \cap_{j \in S} (\mathcal{P}[T_j])^c \cap \mathcal{P}_{\cup_{j=1}^r F_{\mathbf{r}_j}}. \quad (3.66)$$

Now we prove the above identity. By definition (Definition 3.3.3 (5)),

$$\begin{aligned} \mathcal{P}_{\cup_{j=1}^r F_{\mathbf{r}_j}} &= \{p \in \mathcal{P} : \forall \mathbf{n} \in \cup_{j=1}^r F_{\mathbf{r}_j}, p \notin \mathcal{P}(L(T_{\mathbf{n}}), A)\} \\ &= \{p \in \mathcal{P} : \forall \mathbf{n} \in F \setminus \{\mathbf{r}_1, \dots, \mathbf{r}_r\}, p \notin \mathcal{P}(L(T_{\mathbf{n}}), A)\}, \end{aligned} \quad (3.67)$$

and

$$\mathcal{P}_F = \{p \in \mathcal{P} : \forall \mathbf{n} \in F, p \notin \mathcal{P}(L(T_{\mathbf{n}}), A)\}. \quad (3.68)$$

By definition of $\mathcal{P}[T_j]$ (it is in the paragraph above the claim),

$$(\mathcal{P}[T_j])^c = (\mathcal{P}(L(T_{\mathbf{r}_j}), A))^c = \{p \in \mathcal{P} : \mathbf{r}_j, p \notin \mathcal{P}(L(T_{\mathbf{r}_j}), A)\}. \quad (3.69)$$

Therefore, we have

$$\begin{aligned} & \cap_{j \in S} (\mathcal{P}[T_j])^c \cap \mathcal{P}_{\cup_{j=1}^r F_{\mathbf{r}_j}} \\ &= \cap_{j \in S} (\{p \in \mathcal{P} : \mathbf{r}_j, p \notin \mathcal{P}(L(T_{\mathbf{r}_j}), A)\}) \cap \{p \in \mathcal{P} : \forall \mathbf{n} \in F \setminus \{\mathbf{r}_1, \dots, \mathbf{r}_r\}, p \notin \mathcal{P}(L(T_{\mathbf{n}}), A)\} \\ &= \{p \in \mathcal{P} : \forall \mathbf{n} \in F, p \notin \mathcal{P}(L(T_{\mathbf{n}}), A)\} = \mathcal{P}_F. \end{aligned} \quad (3.70)$$

Now we finished the proof of (3.66) and thus the proof of the renormalized Wick theorem. \square

Denote by $R(T) \cup R(T)$ the union of two copies of $R(T)$, then we have

$$\begin{aligned} & \mathbb{E} \left([\xi_{k_1} \bar{\xi}_{k_2} \cdots \xi_{k_{2l(T)+1}}]_{R(T)} [\bar{\xi}_{k'_1} \xi_{k'_2} \cdots \bar{\xi}_{k'_{2l(T)+1}}]_{R(T)} \right) \\ &= \mathbb{E} \left([\xi_{k_1} \bar{\xi}_{k_2} \cdots \xi_{k_{2l(T)+1}} \bar{\xi}_{k'_1} \xi_{k'_2} \cdots \bar{\xi}_{k'_{2l(T)+1}}]_{R(T) \cup R(T)} \right) \end{aligned} \quad (3.71)$$

Applying Wick theorem to (3.49), we get

$$\begin{aligned} \mathbb{E} |\mathcal{J}_{T,k}|^2 &= \left(\frac{\lambda^2}{L^{2d}} \right)^{2l(T)} \sum_{p \in \mathcal{P}_F(\{k_1, \dots, k_{2l(T)+1}, k'_1, \dots, k'_{2l(T)+1}\})} \\ & \underbrace{\sum_{\substack{k_1, k_2, \dots, k_{2l(T)+1} \\ k'_1, k'_2, \dots, k'_{2l(T)+1}}} H_{k_1 \dots k_{2l(T)+1}}^T H_{k'_1 \dots k'_{2l(T)+1}}^T \delta_p(k_1, \dots, k_{2l(T)+1}, k'_1, \dots, k'_{2l(T)+1}) \sqrt{n_{\text{in}}(k_1)} \cdots}_{\text{Term}(T,p)} \end{aligned} \quad (3.72)$$

We see that the correlation of two tree terms is a sum of smaller expressions $Term(T, p)$. By (3.46), the coefficients $H_{k_1 \dots k_{2l(T)+1}}^T H_{k'_1 \dots k'_{2l(T)+1}}^T$ of $Term(T, p)$ concentrate near the subset

$$\{k_{\mathfrak{e}}, k'_{\mathfrak{e}} \in \mathbb{Z}^d, |k_{\mathfrak{e}}|, |k'_{\mathfrak{e}}| \lesssim 1, \forall \mathfrak{e} : |\Omega_{\mathfrak{n}}|, |\Omega'_{\mathfrak{n}}| \lesssim \omega, \forall \mathfrak{n}. \{k_{\mathfrak{e}}\}_{\mathfrak{e}}, \{k'_{\mathfrak{e}}\}_{\mathfrak{e}} \in \cap_{\mathfrak{n} \in T_{\text{in}}} S_{\mathfrak{n}}. k_l = k'_l = k\}. \quad (3.73)$$

The pairing p in Wick theorem introduces new equations $k_i = k'_j$ (defined in (3.52)) and the coefficients $H_{k_1 \dots k_{2l(T)+1}}^T H_{k'_1 \dots k'_{2l(T)+1}}^T \delta_p$ concentrate near the subset

$$\{k_{\mathfrak{e}}, k'_{\mathfrak{e}} \in \mathbb{Z}^d, |k_{\mathfrak{e}}|, |k'_{\mathfrak{e}}| \lesssim 1, \forall \mathfrak{e} : |\Omega_{\mathfrak{n}}|, |\Omega'_{\mathfrak{n}}| \lesssim \omega, \forall \mathfrak{n}. \{k_{\mathfrak{e}}\}_{\mathfrak{e}}, \{k'_{\mathfrak{e}}\}_{\mathfrak{e}} \in \cap_{\mathfrak{n} \in T_{\text{in}}} S_{\mathfrak{n}}. k_l = k'_l = k. \\ k_i = k'_j \text{ (and } k_i = k_j, k'_i = k'_j) \text{ for all } k_i \sim_p k'_j \text{ (and } k_i \sim_p k_j, k'_i \sim_p k'_j)\}. \quad (3.74)$$

As in the case of (3.46), there is a graphical representation of (3.74). To explain this, we need the concept of couples.

Definition 3.3.6 (Construction of couples). Given two trees T and T' , we flip the orientation of all edges in T' (as in the two left trees in Figure 3.7). We also label their leaves by $1, 2, \dots, 2l(T)+1$ and $1, 2, \dots, 2l(T')+1$ so that the corresponding variables of these leaves are $k_1, k_2, \dots, k_{2l(T)+1}$ and $k_1, k_2, \dots, k_{2l(T')+1}$. Assume that we have a pairing p of the set $\{k_1, k_2, \dots, k_{2l(T)+1}, k_1, k_2, \dots, k_{2l(T')+1}\}$, then this pairing induces a pairing between leaves (if $k_i \sim_p k_j$ then define *the i -th leaf \sim_p the j -th leaf*). Given this pairing of leaves, we define the following procedure which glues two trees T and T' into a couple $\mathcal{C}(T, T', p)$. Some examples of pairing can be found in Figure 3.7.

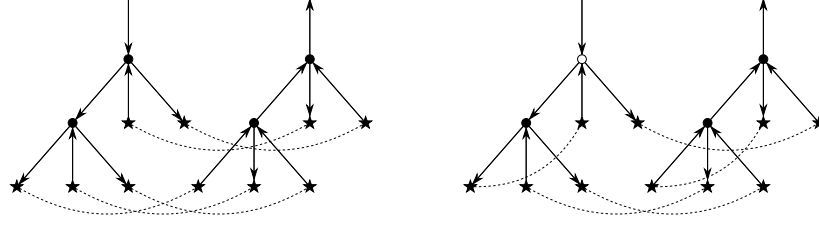


Figure 3.7: Example of pairings between trees.

1. **Merging edges connected to leaves.** Given two edges with opposite orientations connected to two paired leaves, these two edges can be merged into one edge as in Figure 3.8. Since the pairing in this paper is a balanced pairing, only edges with opposite orientations can be paired.

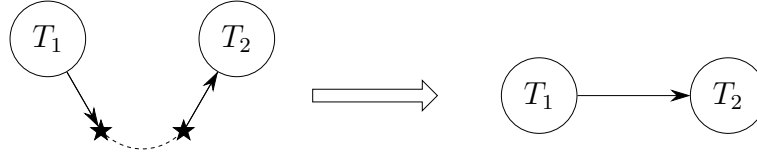


Figure 3.8: Pairing and merging of two edges

We know that two edges connected to leaves correspond to two indices k_i, k_j . Merging two such edges is a graphical interpretation that $k_i = k_j = k$.

2. **Splitting of \circ nodes:** Given a \circ node \mathbf{n} , we split it into two edges as in Figure 3.9.

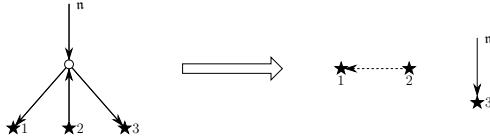


Figure 3.9: Splitting of a \circ node

One of these two edges is usual and the other one is dotted (called dotted edge). If the parent of \circ is \mathbf{n} and the children of \circ are $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$ from left to right, then we require the usual edge connects \mathbf{n} and \mathbf{n}_3 and the dotted edge connects \mathbf{n}_1 and \mathbf{n}_2 . Figure 3.10 is an example of splitting \circ in a tree

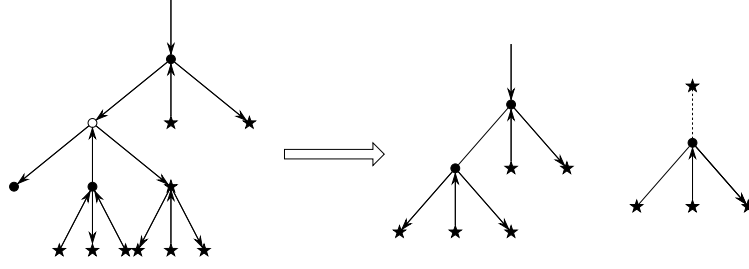


Figure 3.10: An example of splitting a \circ node

We know that the \circ node corresponds to equation $S_n = \{k_{\epsilon_1} = k_{\epsilon_2}, k_{\epsilon_3} = k_{\epsilon}\}$. The corresponding variable k_{ϵ_d} of the dotted edge ϵ_d is defined to be the common value of k_{ϵ_1} and k_{ϵ_2} . The corresponding variable k_{ϵ_n} of the usual edge ϵ_n is defined to be the common value of k_{ϵ_3} and k_{ϵ} .

3. **Pairing of trees and couples.** Given a pairing p of the set of leaves in T, T' we merge all edges paired by p as in Figure 3.11 and the resulting combinatorial structure is called a couple.

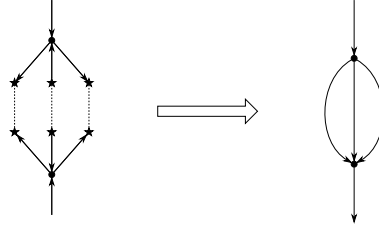


Figure 3.11: The construction of a couple

We know that each edge connected to a leaf corresponds to a variable k_i . A pairing p of $\{k_1, k_2, \dots, k_{2m}\}$ in (3.74) induces a pairing of edges connected to leaves. Merging paired edges corresponds to $k_i = k'_j$ for all $k_i \sim_p k'_j$ in (3.74).

Proposition 3.3.7. (3.74) can be read from a couple diagram $\mathcal{C}(T, T, p)$. Each edge corresponds to a variable k_{ϵ} . The leg \mathfrak{l} corresponds to equation $k_{\mathfrak{l}} = k$. Each node corresponds to a momentum conservation equation

$$k_{\epsilon_1} - k_{\epsilon_2} + k_{\epsilon_3} - k_{\epsilon} = 0, \quad k_{\epsilon_1} \neq k_{\epsilon_2} \neq k_{\epsilon_3}, \quad \text{or} \quad k_{\epsilon_1} = k_{\epsilon_2} = k_{\epsilon_3} = k, \quad (3.75)$$

and an energy conservation equation

$$\Lambda_{k_{\epsilon_1}} - \Lambda_{k_{\epsilon_2}} + \Lambda_{k_{\epsilon_3}} - \Lambda_{k_{\epsilon}} = O(\omega). \quad (3.76)$$

Remark 3.3.8. In a couple diagrams, we only have nodes decorated by \bullet . Nodes decorated by \circ and \star have been removed in (2), (3) of Definition 3.3.6.

Remark 3.3.9. Through the process of (2), (3) in Definition 3.3.6, a couple diagram can automatically encode the equation $k_i = k'_j$ for all $k_i \sim_p k'_j$ and $k_{\epsilon_1} = k_{\epsilon_2}$, $k_{\epsilon_3} = k_{\epsilon}$. Therefore, they do not appear in Proposition 3.3.7.

Proof. This directly follows from the definition of couples. \square

The calculations of this section are summarized in the following proposition.

Proposition 3.3.10. (1) Define $Term(T, p)$ in the same way as in (3.72),

$$\begin{aligned} Term(T, p) = & \sum_{k_1, k_2, \dots, k_{2l(T)+1}} \sum_{k'_1, k'_2, \dots, k'_{2l(T)+1}} \\ & H_{k_1 \dots k_{2l(T)+1}}^T H_{k'_1 \dots k'_{2l(T)+1}}^T \delta_p(k_1, \dots, k_{2l(T)+1}, k'_1, \dots, k'_{2l(T)+1}) \sqrt{n_{in}(k_1)} \cdots \sqrt{n_{in}(k'_{2l(T)+1})} \cdots \end{aligned} \quad (3.77)$$

then $\mathbb{E}|\mathcal{J}_{T,k}|^2$ is a sum of $Term(T, p)$ for all $p \in \mathcal{P}_F$, (in (3.72) the sum is over set of all possible pairing \mathcal{P})

$$\mathbb{E}|\mathcal{J}_{T,k}|^2 = \left(\frac{\lambda^2}{L^{2d}} \right)^{2l(T)} \sum_{p \in \mathcal{P}_F(\{k_1, \dots, k_{2l(T)+1}, k'_1, \dots, k'_{2l(T)+1}\})} Term(T, p). \quad (3.78)$$

(2) Since the definition (3.77) of $Term(T, p)$ do not change, it still concentrates near the subset (3.74) which has a simple graphical representation given by Proposition 3.3.7.

Proof. The proof of (1), (2) is easy and thus skipped. \square

3.3.3 Counting lattice points

In this section, we apply the connection between couple and concentration subset (3.74) to count the number of solutions of a generalized version of (3.74),

$$\begin{aligned} \{k_{\mathfrak{e}}, k'_{\mathfrak{e}} \in \mathbb{Z}^d, |k_{\mathfrak{e}}|, |k'_{\mathfrak{e}}| \lesssim 1, \forall \mathfrak{e} : |\Omega_{\mathfrak{n}} - \sigma_{\mathfrak{n}}|, |\Omega'_{\mathfrak{n}} - \sigma'_{\mathfrak{n}}| \lesssim \omega, \forall \mathfrak{n}. \{k_{\mathfrak{e}}\}_{\mathfrak{e}}, \{k'_{\mathfrak{e}}\}_{\mathfrak{e}} \in \cap_{\mathfrak{n} \in T'} S_{\mathfrak{n}}. k_{\mathfrak{l}} = k'_{\mathfrak{l}} = k. \\ k_i = k'_j \text{ (and } k_i = k_j, k'_i = k'_j) \text{ for all } k_i \sim_p k'_j \text{ (and } k_i \sim_p k_j, k'_i \sim_p k'_j)\}. \end{aligned} \quad (3.79)$$

(3.79) is obtained by replacing $\Omega_{\mathfrak{n}}, \Omega'_{\mathfrak{n}}$ by $\Omega_{\mathfrak{n}} - \sigma_{\mathfrak{n}}, \Omega'_{\mathfrak{n}} - \sigma'_{\mathfrak{n}}$ in (3.74). $\sigma_{\mathfrak{n}}$ and $\sigma'_{\mathfrak{n}}$ are some given constants. The counterpart of Proposition 3.3.7 in this case is

Proposition 3.3.11. (3.79) can be read from a couple diagram $\mathcal{C} = \mathcal{C}(T, T, p)$. Each edge corresponds to a variable $k_{\mathfrak{e}}$. The leg \mathfrak{l} corresponds to equation $k_{\mathfrak{l}} = k$. Each node corresponds to a momentum conservation equation

$$k_{\mathfrak{e}_1} - k_{\mathfrak{e}_2} + k_{\mathfrak{e}_3} - k_{\mathfrak{e}} = 0, \quad k_{\mathfrak{e}_1} \neq k_{\mathfrak{e}_2} \neq k_{\mathfrak{e}_3}, \text{ or } k_{\mathfrak{e}_1} = k_{\mathfrak{e}_2} = k_{\mathfrak{e}_3} = k, \quad (3.80)$$

and a energy conservation equation

$$\Lambda_{k_{\mathfrak{e}_1}} - \Lambda_{k_{\mathfrak{e}_2}} + \Lambda_{k_{\mathfrak{e}_3}} - \Lambda_{k_{\mathfrak{e}}} = \sigma_{\mathfrak{n}} + O(\omega). \quad (3.81)$$

Denote the momentum and energy conservation equations by $MC_{\mathfrak{n}}$ and $EC_{\mathfrak{n}}$ respectively, then (3.79) can be rewritten as

$$(3.79) = \{k_{\mathfrak{e}} \in \mathbb{Z}^d, |k_{\mathfrak{e}}| \lesssim 1, \forall \mathfrak{e} \in \mathcal{C} : MC_{\mathfrak{n}}, EC_{\mathfrak{n}}, \forall \mathfrak{n} \in \mathcal{C}. k_{\mathfrak{l}} = k'_{\mathfrak{l}} = k.\} \quad (3.82)$$

Proof. This directly follows from Proposition 3.3.7. □

Basic ideas for counting lattice points

In this section, we explain the basic idea used in this paper to count lattice points. To do this, we need the following definitions related to couples.

Definition 3.3.12. 1. **Connected couples.** A couple \mathcal{C} is a connected couple if it is connected as a graph. The connected components is also defined in the same way as the graph theory.

2. **Equations of a couple $Eq(\mathcal{C})$:** Given a couple \mathcal{C} and constants k, σ_n , let $Eq(\mathcal{C}, \{\sigma_n\}_n, k)$ (or simply $Eq(\mathcal{C})$) be the system of equation (3.82) constructed in Proposition 3.3.11. For any system of equations Eq , let $\#(Eq)$ be its number of solutions.

The main goal of this section is to prove an upper bound of $\#Eq(\mathcal{C})$. The main idea of proving this is to decompose a large couple \mathcal{C} into smaller pieces and then prove this for the smaller piece using the induction hypothesis. To explain the idea, let us first focus on an example. Let \mathcal{C} be the left couple in the following picture. (The corresponding variables of each edge are labeled near these edges.)

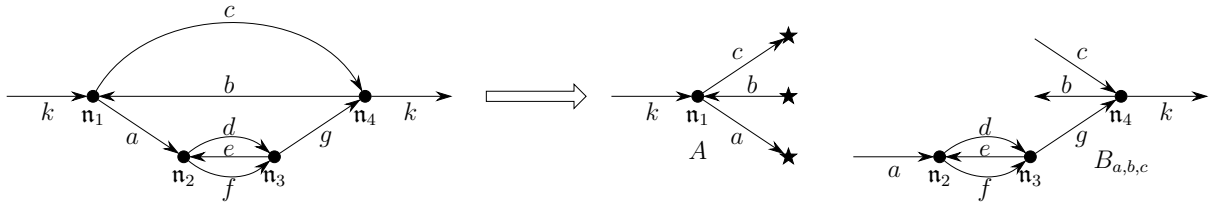


Figure 3.12: An example of decomposing a couple

By (3.82), we know that the couple \mathcal{C} corresponds to the following equations.

$$\begin{aligned}
& \{(a, b, c, d, e, f, g) : (|a| \text{ to } |g|) \lesssim 1, \\
& a - b + c = k, \Lambda(a) - \Lambda(b) + \Lambda(c) - \Lambda(k) = \sigma_1 + O(\omega) \\
& d - e + f = a, \Lambda(d) - \Lambda(e) + \Lambda(f) - \Lambda(a) = \sigma_2 + O(\omega) \\
& d - e + f = g, \Lambda(d) - \Lambda(e) + \Lambda(f) - \Lambda(g) = \sigma_3 + O(\omega) \\
& c - b + g = k, \Lambda(c) - \Lambda(b) + \Lambda(g) - \Lambda(k) = \sigma_4 + O(\omega)\}
\end{aligned} \tag{3.83}$$

We know that (3.83) can be rewritten into the form $\bigcup_{a,e \in A} B_{a,e}$, where

$$A = \{a, b, c : |a|, |b|, |c| \lesssim 1, a - b + c = k, \Lambda(a) - \Lambda(b) + \Lambda(c) - \Lambda(k) = \sigma_1 + O(\omega)\} \tag{3.84}$$

$$\begin{aligned}
B_{a,b,c} = \{d, e, f, g : & |d|, |e|, |f|, |g| \lesssim 1, \\
& d - e + f = a, \Lambda(d) - \Lambda(e) + \Lambda(f) - \Lambda(a) = \sigma_2 + O(\omega) \\
& d - e + f = g, \Lambda(d) - \Lambda(e) + \Lambda(f) - \Lambda(g) = \sigma_3 + O(\omega) \\
& c - b + g = k, \Lambda(c) - \Lambda(b) + \Lambda(g) - \Lambda(k) = \sigma_4 + O(\omega)\}
\end{aligned} \tag{3.85}$$

Since an upper bound of $\#Eq(\mathcal{C})$ can be derived from upper bounds of $\#A$, $\#B_{a,b,c}$, we just need to consider A , $B_{a,b,c}$ which are systems of equations of smaller size. We can reduce the size of systems of equations in this way and prove upper bounds by induction.

One problem of applying an induction argument is that A , $B_{a,b,c}$ cannot be represented by couple defined by Definition 3.3.6 that can contain at most two legs (an edge just connected to one node). In Definition 3.3.6, a leg is used to represent a variable that is fixed, as in the condition $k_l = k'_l = k$ in (3.82). The definition of $\#B_{a,b,c}$ contains four fixed variables a, b, c, k which cannot be represented by just two legs. Therefore, we have to define a new type of couple that allows multiple legs.

Except for the lack of legs, we also have the problem of representing free variables.

We know that the couple representation of A should contain one node and four edges if we insist on the rule that a node corresponds to an equation and the variables in the equation correspond to edges connected to this node. All these edges are legs, but three of four edges correspond to variables a, b, c which are not fixed. Therefore, we have to define a type of legs that can correspond to unfixed variables.

To solve the above problems, we introduce the following definition.

Definition 3.3.13. 1. **Couples with multiple legs:** A graph in which all nodes have degree 1 or 4 is called a couples with multiple legs. The graph A and $B_{a,b,c}$ in Figure 3.12 are examples of this definition.

2. **Legs:** In a couple with multiple legs, an edge connected to a degree one node is called a leg. Remember that we have encountered this concept in the second paragraph of section 3.3.1 and in what follows, we call the leg defined there the root leg of a tree.

3. **Free legs and fixed legs:** In a couple with multiple legs, we use two types of node decoration for degree 1 nodes as in Figure 3.13. One is \star and the other one is invisible.

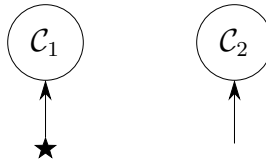


Figure 3.13: Node decoration of degree one node

An edge connected to a \star or invisible nodes is called a free leg or fixed leg respectively. They correspond to free variables or fixed variables k_{ϵ} respectively.

4. **Equations of a couple $Eq(\mathcal{C}, \{c_i\}_I)$:** We define the corresponding equations

for couples with multiple legs.

$$Eq(\mathcal{C}, \{c_l\}_l) = \{k_{\mathfrak{e}} \in \mathbb{Z}_L^d, |k_{\mathfrak{e}}| \lesssim 1 \ \forall \mathfrak{e} \in \mathcal{C} : MC_n, EC_n, \forall \mathfrak{n} \in \mathcal{C}. k_l = c_l, \forall l.\} \quad (3.86)$$

In this representation, the corresponding variable of a fixed leg l is fixed to be the constant c_l and the corresponding variable of a free leg l is not fixed.

With the above definition, it's easy to show that the couple A and $B_{a,b,c}$ in Figure 3.12 correspond to the system of equations (3.84) and (3.85) respectively.

The cutting operation and its properties

In this section, we give the formal definition of cutting and explain how $\#Eq(\mathcal{C})$ changes after cutting.

Definition 3.3.14. 1. **Cutting an edge:** Given an edge \mathfrak{e} , we can cut it into two edges (a fixed and a free leg) as in Figure 3.14.

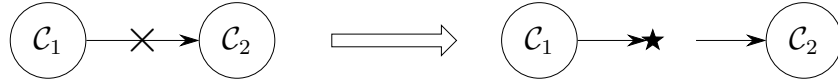


Figure 3.14: An example of cutting an edge

2. **Cut:** A cut c of a couple \mathcal{C} is a set of edges such that \mathcal{C} is disconnected after cutting all edges in c , together with a map $rc : c \rightarrow \{\text{left}, \text{right}\}$. For each $\mathfrak{e} \in c$, if $rc(\mathfrak{e}) = \text{left}$ (resp. right), then as in Figure 3.14 the left node (resp. right node) produced by cutting \mathfrak{e} is a \star node (resp. invisible node). The map rc describes which one should be the free or fixed leg in the two legs produced by cutting an edge. An admissible cut is defined to be the cut such that the new legs after cutting are all free in one component and are fixed in another component.

3. $c(\mathfrak{e})$, $c(\mathfrak{n})$ and $c(\mathfrak{l})$: Given an edge \mathfrak{e} that is not a leg, define $c(\mathfrak{e})$ to be the cut that contains only one edge \mathfrak{e} . Given a node $\mathfrak{n} \in \mathcal{C}$, let $\{\mathfrak{e}_i\}$ be edges that are connected to \mathfrak{n} , then define $c(\mathfrak{n})$ to be the cut that consists of edges $\{\mathfrak{e}_i\}$. Given a leg \mathfrak{l} , let \mathfrak{n} be the unique node connected to it, then define $c(\mathfrak{l})$ to be the cut $c(\mathfrak{n})$. An example of cutting $c(\mathfrak{e})$ is given by Figure 3.14. The following picture gives an example of cutting $c(\mathfrak{n})$ or $c(\mathfrak{l})$ (in this picture $\mathfrak{n} = \mathfrak{n}_1$ and \mathfrak{l} is the leg labelled by k .)

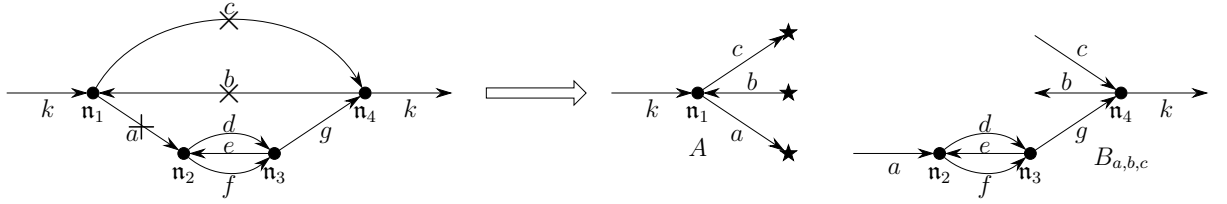


Figure 3.15: An example of cuts, $c(\mathfrak{n})$ and $c(\mathfrak{l})$

Remark 3.3.15. Explicitly writing down the full definition of rc is often complicated, so in what follows, when defining rc , we will only describe which one should be the free or fixed leg in the two legs produced by cutting an edge.

Let us explain how does $\text{Eq}(\mathcal{C})$ and $\#\text{Eq}(\mathcal{C})$ changes after cutting. The result is summarized in the following lemma.

Lemma 3.3.16. *Let c be an admissible cut of \mathcal{C} that consists of edges $\{\mathfrak{e}_i\}$ and \mathcal{C}_1 , \mathcal{C}_2 be two components after cutting. Let $\mathfrak{e}_i^{(1)} \in \mathcal{C}_1$, $\mathfrak{e}_i^{(2)} \in \mathcal{C}_2$ be two edges obtained by cutting \mathfrak{e}_i . The rc map is defined by assigning $\{\mathfrak{e}_i^{(1)}\}$ to be free legs and $\{\mathfrak{e}_i^{(2)}\}$ to be fixed legs. Then we have*

$$\text{Eq}(\mathcal{C}, \{c_i\}_l) = \left\{ (k_{\mathfrak{e}_1}, k_{\mathfrak{e}_2}) : k_{\mathfrak{e}_1} \in \text{Eq}(\mathcal{C}_1, \{c_{l_1}\}), k_{\mathfrak{e}_2} \in \text{Eq}(\mathcal{C}_2, \{c_{l_2}\}, \{k_{\mathfrak{e}_i^{(1)}}\}_i) \right\}. \quad (3.87)$$

and

$$\sup_{\{c_l\}_l} \#Eq(\mathcal{C}, \{c_l\}_l) \leq \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_1)}} \#Eq(\mathcal{C}_1, \{c_{l_1}\}) \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}_2)}} \#Eq(\mathcal{C}_2, \{c_{l_2}\}). \quad (3.88)$$

Here $\text{leg}(\mathcal{C})$ is the set of fixed legs in \mathcal{C} (not the set of all legs!).

Proof. By definition (3.86), we have

$$\begin{aligned} Eq(\mathcal{C}, \{c_l\}_l) &= \{k_{\mathbf{e}} \in \mathbb{Z}^d, |k_{\mathbf{e}}| \lesssim 1, \forall \mathbf{e} \in \mathcal{C} : MC_{\mathbf{n}}, EC_{\mathbf{n}}, \forall \mathbf{n} \in \mathcal{C}. k_l = c_l, \forall l \in \text{leg}(\mathcal{C}).\} \\ &= \{(k_{\mathbf{e}_1}, k_{\mathbf{e}_2}) : |k_{\mathbf{e}_1}| \lesssim 1, MC_{\mathbf{n}_1}, EC_{\mathbf{n}_1}, \forall \mathbf{e}_1, \mathbf{n}_1 \in \mathcal{C}_1. k_{l_1} = c_{l_1}, \forall l_1 \in \text{leg}(\mathcal{C}) \cap \text{leg}(\mathcal{C}_1) \\ &\quad |k_{\mathbf{e}_2}| \lesssim 1, MC_{\mathbf{n}_2}, EC_{\mathbf{n}_2}, \forall \mathbf{e}_2, \mathbf{n}_2 \in \mathcal{C}_2. k_{l_2} = c_{l_2}, \forall l_2 \in \text{leg}(\mathcal{C}) \cap \text{leg}(\mathcal{C}_2), k_{\mathbf{e}_i^{(2)}} = k_{\mathbf{e}_i^{(1)}}, \forall \mathbf{e}_i \in \mathcal{C}\} \\ &= \left\{ (k_{\mathbf{e}_1}, k_{\mathbf{e}_2}) : k_{\mathbf{e}_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\}), k_{\mathbf{e}_2} \in Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{\mathbf{e}_i^{(1)}}\right\}_i\right) \right\} \end{aligned} \quad (3.89)$$

Here in $Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{\mathbf{e}_i^{(1)}}\right\}_i\right)$, $k_{\mathbf{e}_i^{(1)}}$ are view as a constant value and $k_{\mathbf{e}_i^{(2)}}$ are fixed to be this constant value.

$$\begin{aligned} Eq(\mathcal{C}, \{c_l\}_l) &= \{k_{\mathbf{e}} \in \mathbb{Z}^d, |k_{\mathbf{e}}| \lesssim 1, \forall \mathbf{e} \in \mathcal{C} : MC_{\mathbf{n}}, EC_{\mathbf{n}}, \forall \mathbf{n} \in \mathcal{C}. k_l = c_l, \forall l \in \text{leg}(\mathcal{C}).\} \\ &= \{(k_{\mathbf{e}_1}, k_{\mathbf{e}_2}) : |k_{\mathbf{e}_1}| \lesssim 1, MC_{\mathbf{n}_1}, EC_{\mathbf{n}_1}, \forall \mathbf{e}_1, \mathbf{n}_1 \in \mathcal{C}_1. k_{l_1} = c_{l_1}, \forall l_1 \in \text{leg}(\mathcal{C}) \cap \text{leg}(\mathcal{C}_1) \\ &\quad |k_{\mathbf{e}_2}| \lesssim 1, MC_{\mathbf{n}_2}, EC_{\mathbf{n}_2}, \forall \mathbf{e}_2, \mathbf{n}_2 \in \mathcal{C}_2. k_{l_2} = c_{l_2}, \forall l_2 \in \text{leg}(\mathcal{C}) \cap \text{leg}(\mathcal{C}_2), k_{\mathbf{e}_i^{(2)}} = k_{\mathbf{e}_i^{(1)}}, \forall \mathbf{e}_i \in \mathcal{C}\} \\ &= \left\{ (k_{\mathbf{e}_1}, k_{\mathbf{e}_2}) : k_{\mathbf{e}_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\}), k_{\mathbf{e}_2} \in Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{\mathbf{e}_i^{(1)}}\right\}_i\right) \right\} \end{aligned} \quad (3.90)$$

Here in $Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{\mathbf{e}_i^{(1)}}\right\}_i\right)$, $k_{\mathbf{e}_i^{(1)}}$ are view as a constant value and $k_{\mathbf{e}_i^{(2)}}$ are fixed to be this constant value.

Therefore, we have the following identity of $Eq(\mathcal{C}, \{c_l\}_l)$

$$Eq(\mathcal{C}, \{c_l\}_l) = \left\{ (k_{\mathbf{e}_1}, k_{\mathbf{e}_2}) : k_{\mathbf{e}_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\}), k_{\mathbf{e}_2} \in Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{\mathbf{e}_i^{(1)}}\right\}_i\right) \right\}. \quad (3.91)$$

which proves (3.87).

We can also find the relation between $\#Eq(\mathcal{C}_1)$, $\#Eq(\mathcal{C}_2)$ and $\#Eq(\mathcal{C})$. Applying (3.87),

$$\begin{aligned}
\#Eq(\mathcal{C}, \{c_l\}_l) &= \sum_{(k_{e_1}, k_{e_2}) \in \#Eq(\mathcal{C}, \{c_l\}_l)} 1 \\
&= \sum_{\left\{ (k_{e_1}, k_{e_2}): k_{e_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\}), k_{e_2} \in Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{e_i^{(1)}}\right\}_i\right) \right\}} 1 \\
&= \sum_{k_{e_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\})} \sum_{k_{e_2} \in Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{e_i^{(1)}}\right\}_i\right)} 1 \\
&= \sum_{k_{e_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\})} \#Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{e_i^{(1)}}\right\}_i\right)
\end{aligned} \tag{3.92}$$

Take sup in the above equation

$$\begin{aligned}
\sup_{\{c_l\}_l} \#Eq(\mathcal{C}, \{c_l\}_l) &= \sup_{\{c_l\}_l} \sum_{k_{e_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\})} \#Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{e_i^{(1)}}\right\}_i\right) \\
&\leq \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}) \cap \text{leg}(\mathcal{C}_1)}} \sum_{k_{e_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\})} \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}) \cap \text{leg}(\mathcal{C}_2)}} \#Eq\left(\mathcal{C}_2, \{c_{l_2}\}, \left\{k_{e_i^{(1)}}\right\}_i\right) \\
&\leq \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_1)}} \sum_{k_{e_1} \in Eq(\mathcal{C}_1, \{c_{l_1}\})} \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}_2)}} \#Eq(\mathcal{C}_2, \{c_{l_2}\}) \\
&= \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_1)}} \#Eq(\mathcal{C}_1, \{c_{l_1}\}) \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}_2)}} \#Eq(\mathcal{C}_2, \{c_{l_2}\})
\end{aligned} \tag{3.93}$$

This proves (3.88). \square

Combinatorial properties of couples

In this section, we prove several combinatorial properties of couples. These properties are important in the counting argument.

The following lemma gives a relation between the numbers of non-leg edges, legs, and nodes.

Lemma 3.3.17. *For any couple \mathcal{C} , let $n(\mathcal{C})$ be the total number of nodes in \mathcal{C} and $n_e(\mathcal{C})$ (resp. $n_{fx}(\mathcal{C})$, $n_{fr}(\mathcal{C})$) be the total number of non-leg edges (resp. fixed legs, free legs). Then we have*

(1) $n(\mathcal{C})$, $n_e(\mathcal{C})$, $n_{fx}(\mathcal{C})$ and $n_{fr}(\mathcal{C})$ satisfies the following relation

$$2n_e(\mathcal{C}) + n_{fx}(\mathcal{C}) + n_{fr}(\mathcal{C}) = 4n(\mathcal{C}) \quad (3.94)$$

(2) For any couple \mathcal{C} , define $\chi(\mathcal{C}) = n_e(\mathcal{C}) + n_{fr}(\mathcal{C}) - n(\mathcal{C})$. Let c , \mathcal{C}_1 , \mathcal{C}_2 be the same as in Lemma 3.3.16 and also assume that $\{\mathfrak{e}_i^{(1)}\}$ are free legs and $\{\mathfrak{e}_i^{(2)}\}$ are fixed legs. Then

$$\chi(\mathcal{C}) = \chi(\mathcal{C}_1) + \chi(\mathcal{C}_2). \quad (3.95)$$

Proof. We first prove (1). Consider the set $\mathcal{S} = \{(\mathbf{n}, \mathfrak{e}) \in \mathcal{C} : \mathbf{n} \text{ is an end point of } \mathfrak{e}\}$, then

$$\#\mathcal{S} = \sum_{\substack{(\mathbf{n}, \mathfrak{e}) \in \mathcal{C} \\ \mathbf{n} \text{ is an end point of } \mathfrak{e}}} 1 \quad (3.96)$$

First sum over \mathfrak{e} and then over \mathbf{n} , we get

$$\#\mathcal{S} = \sum_{\mathbf{n}} \sum_{\substack{\mathfrak{e} \in \mathcal{C} \\ \mathbf{n} \text{ is an end point of } \mathfrak{e}}} 1 = \sum_{\mathbf{n}} 4 = 4n(\mathcal{C}). \quad (3.97)$$

In the second equality, $\sum_{\mathfrak{e} \in \mathcal{C}: \mathbf{n} \text{ is an end point of } \mathfrak{e}} 1 = 4$ because for each node \mathbf{n} there are 4 edges connected to it.

Switch the order of summation, we get

$$\begin{aligned} \#\mathcal{S} &= \sum_{\mathfrak{e} \text{ is a non-leg edges}} \sum_{\substack{\mathbf{n} \in \mathcal{C} \\ \mathbf{n} \text{ is an end point of } \mathfrak{e}}} 1 + \sum_{\mathfrak{e} \text{ is a leg}} \sum_{\substack{\mathbf{n} \in \mathcal{C} \\ \mathbf{n} \text{ is an end point of } \mathfrak{e}}} 1 \\ &= \sum_{\mathfrak{e} \text{ is a non-leg edges}} 2 + \sum_{\mathfrak{e} \text{ is a leg}} 1 \\ &= 2n_e(\mathcal{C}) + n_{fx}(\mathcal{C}) + n_{fr}(\mathcal{C}) \end{aligned} \quad (3.98)$$

In the second equality, $\sum_{\substack{\mathbf{n} \text{ is an end point of } \mathbf{e} \\ \mathbf{n} \in \mathcal{C}}} 1$ equals to 1 or 2 because for each non-leg edge (resp. leg) there are 2 (resp. 1) nodes connected to it.

Because the value of $\#\mathcal{S}$ does not depend on the order of summation, we conclude that $2n_e(\mathcal{C}) + n_{fx}(\mathcal{C}) + n_{fr}(\mathcal{C}) = 4n(\mathcal{C})$, which proves (1).

We now prove (2). Since cutting does change the number of nodes, we have $n(\mathcal{C}) = n(\mathcal{C}_1) + n(\mathcal{C}_2)$. Let c be cut that consists of edges $\{\mathbf{e}_i\}$ and $n(c)$ be the number of edges in c . When cutting \mathcal{C} into \mathcal{C}_1 and \mathcal{C}_2 , $n(c)$ non-leg edges are cut into pairs of free and fixed legs, so $n_e(\mathcal{C}) = n(\mathcal{C}_1) + n(\mathcal{C}_2) + n(c)$. Because we have $n(c)$ additional free legs after cutting, so $n_{fr}(\mathcal{C}) = n_{fr}(\mathcal{C}_1) + n_{fr}(\mathcal{C}_2) - n(c)$. Therefore, we get

$$\begin{aligned}
\chi(\mathcal{C}) &= n_e(\mathcal{C}) + n_{fr}(\mathcal{C}) - n(\mathcal{C}) \\
&= n(\mathcal{C}_1) + n(\mathcal{C}_2) + n(c) + n_{fr}(\mathcal{C}_1) + n_{fr}(\mathcal{C}_2) - n(c) - (n(\mathcal{C}_1) + n(\mathcal{C}_2)) \\
&= (n(\mathcal{C}_1) + n_{fr}(\mathcal{C}_1) - n(\mathcal{C}_1)) + (n(\mathcal{C}_2) + n_{fr}(\mathcal{C}_2) - n(\mathcal{C}_2)) \\
&= \chi(\mathcal{C}_1) + \chi(\mathcal{C}_2).
\end{aligned} \tag{3.99}$$

We complete the proof of (2). □

The following lemma gives a relation between variables of legs.

Lemma 3.3.18. *Given a connected couple \mathcal{C} with multiple legs, then we have the following conclusions.*

(1) Let $\{k_{\mathbf{l}_i}\}_{i=1, \dots, n_{leg}}$ be the variables corresponding to legs in couple \mathcal{C} . Let $\iota_{\mathbf{e}}$ be the same as Definition 3.2.1 (12), then $Eq(\mathcal{C})$ implies the following momentum conservation equation

$$\sum_{i=1}^{n_{leg}} \iota_{\mathbf{l}_i} k_{\mathbf{l}_i} = 0, \tag{3.100}$$

(2) Assume that there is exactly one free leg \mathbf{l}_{i_0} in \mathcal{C} and all other variables $\{k_{\mathbf{l}_i}\}_{i \neq i_0}$ corresponding to fix legs are fixed to be constants $\{c_{\mathbf{l}_i}\}_{i \neq i_0}$. For any $i_1 = 1, \dots, n_{leg}$, we can construct a new couple $\hat{\mathcal{C}}$ by replacing the i_0 leg by a fixed leg and i_1 leg by a

free leg. If $i \neq i_0, i_1$, fix k_{l_i} to be the constant c_{l_i} , if $i = i_0$, fix $k_{l_{i_0}}$ to be the constant $-\iota_{l_{i_0}} \sum_{i \neq i_0} \iota_{l_i} k_{l_i}$. Under the above assumptions, we have

$$Eq(\mathcal{C}, \{c_{l_i}\}_{i \neq i_0}) = Eq\left(\widehat{\mathcal{C}}, \{c_{l_i}\}_{i \neq i_0, i_1} \cup \{-\iota_{l_{i_0}} \sum_{i \neq i_0} \iota_{l_i} k_{l_i}\}\right). \quad (3.101)$$

(3) Assume that there is no free leg in \mathcal{C} and all $\{k_{l_i}\}_{i \neq i_0}$ are fixed to be constants $\{c_{l_i}\}_{i \neq i_0}$. For any $i_1 = 1, \dots, n_{leg}$, we can construct a new couple $\widehat{\mathcal{C}}$ by replacing the i_0 leg by a free leg. Then we have

$$Eq(\mathcal{C}, \{c_{l_i}\}_i) = Eq(\widehat{\mathcal{C}}, \{c_{l_i}\}_{i \neq i_0}). \quad (3.102)$$

(4) If the couple \mathcal{C} contains any leg, then it contains at least two legs.

Proof. We first prove (1). Given a node \mathbf{n} and an edge \mathbf{e} connected to it, we define $\iota_{\mathbf{e}}(\mathbf{n})$ by the following rule

$$\iota_{\mathbf{e}}(\mathbf{n}) = \begin{cases} +1 & \text{if } \mathbf{e} \text{ pointing towards } \mathbf{n} \\ -1 & \text{if } \mathbf{e} \text{ pointing outwards from } \mathbf{n} \end{cases} \quad (3.103)$$

For a leg \mathbf{l} , since it is connected to just one node, we may omit the (\mathbf{n}) and just write $\iota_{\mathbf{l}}$ as in the statement of the lemma.

For each node \mathbf{n} , let $\mathbf{e}_1(\mathbf{n})$, $\mathbf{e}_2(\mathbf{n})$, $\mathbf{e}(\mathbf{n})$ be the three edges connected to it. For each edge \mathbf{e} , let $\mathbf{n}_1(\mathbf{e})$, $\mathbf{n}_2(\mathbf{e})$ be the two nodes connected to it. Then we know that $\iota_{\mathbf{e}}(\mathbf{n}_1(\mathbf{e})) + \iota_{\mathbf{e}}(\mathbf{n}_2(\mathbf{e}))$, since $\mathbf{n}_1(\mathbf{e})$ and $\mathbf{n}_2(\mathbf{e})$ have the opposite direction.

Since $k_{\mathbf{e}}$ satisfy $Eq(\mathcal{C})$, by (3.80), we get

$$\iota_{\mathbf{e}_1(\mathbf{n})}(\mathbf{n})k_{\mathbf{e}_1(\mathbf{n})} + \iota_{\mathbf{e}_2(\mathbf{n})}(\mathbf{n})k_{\mathbf{e}_2(\mathbf{n})} + \iota_{\mathbf{e}_3(\mathbf{n})}(\mathbf{n})k_{\mathbf{e}_3(\mathbf{n})} + \iota_{\mathbf{e}(\mathbf{n})}(\mathbf{n})k_{\mathbf{e}(\mathbf{n})} = 0. \quad (3.104)$$

Summing over \mathbf{n} gives

$$\begin{aligned}
0 &= \sum_{\mathbf{n} \in \mathcal{C}} \iota_{\mathbf{e}_1(\mathbf{n})}(\mathbf{n}) k_{\mathbf{e}_1(\mathbf{n})} + \iota_{\mathbf{e}_2(\mathbf{n})}(\mathbf{n}) k_{\mathbf{e}_2(\mathbf{n})} + \iota_{\mathbf{e}_3(\mathbf{n})}(\mathbf{n}) k_{\mathbf{e}_3(\mathbf{n})} + \iota_{\mathbf{e}(\mathbf{n})}(\mathbf{n}) k_{\mathbf{e}(\mathbf{n})} \\
&= \sum_{\mathbf{e} \text{ is not a leg}} (\iota_{\mathbf{e}}(\mathbf{n}_1(\mathbf{e})) + \iota_{\mathbf{e}}(\mathbf{n}_2(\mathbf{e}))) k_{\mathbf{e}} + \sum_{\mathbf{l} \text{ is a leg}} \iota_{\mathbf{l}} k_{\mathbf{l}} \\
&= \sum_{i=1}^{n_{\text{leg}}} \iota_{\mathbf{l}_i} k_{\mathbf{l}_i}
\end{aligned} \tag{3.105}$$

This proves (3.100) and thus proves (1).

Now we prove (2). Since in $Eq(\widehat{\mathcal{C}}, \{c_{\mathbf{l}_i}\}_{i \neq i_0, i_1} \cup \{-\iota_{\mathbf{l}_{i_0}} \sum_{i \neq i_0} \iota_{\mathbf{l}_i} k_{\mathbf{l}_i}\})$, $\{k_{\mathbf{l}_i}\}_{i \neq i_0, i_1}$ are fixed to be constants $\{c_{\mathbf{l}_i}\}_{i \neq i_0, i_1}$ and $k_{\mathbf{l}_{i_0}}$ is fixed to be the constant $-\iota_{\mathbf{l}_{i_0}} \sum_{i \neq i_0} \iota_{\mathbf{l}_i} k_{\mathbf{l}_i}$, by (3.100), we know that

$$\iota_{\mathbf{l}_{i_0}} \left(-\iota_{\mathbf{l}_{i_0}} \sum_{i \neq i_0} \iota_{\mathbf{l}_i} k_{\mathbf{l}_i} \right) + \iota_{\mathbf{l}_{i_1}} k_{\mathbf{l}_{i_1}} + \sum_{i \neq i_0, i_1} \iota_{\mathbf{l}_i} c_{\mathbf{l}_i} = 0. \tag{3.106}$$

This implies that $k_{\mathbf{l}_{i_1}} = c_{\mathbf{l}_{i_1}}$ in $Eq(\widehat{\mathcal{C}})$. Therefore, equations in $Eq(\widehat{\mathcal{C}})$ automatically imply $k_{\mathbf{l}_{i_1}} = c_{\mathbf{l}_{i_1}}$. Notice that whether or not containing $k_{\mathbf{l}_{i_1}} = c_{\mathbf{l}_{i_1}}$ is the only difference between $Eq(\mathcal{C})$ and $Eq(\widehat{\mathcal{C}})$. We conclude that $Eq(\mathcal{C}) = Eq(\widehat{\mathcal{C}})$. We thus complete the proof of (2).

The proof of (3) is similar to (2). Whether or not containing $k_{\mathbf{l}_{i_0}} = c_{\mathbf{l}_{i_0}}$ is the only difference between $Eq(\mathcal{C})$ and $Eq(\widehat{\mathcal{C}})$. But if $\{k_{\mathbf{l}_i}\}_{i \neq i_0}$ are fixed to be constants $\{c_{\mathbf{l}_i}\}_{i \neq i_0}$, by momentum conservation we know that

$$k_{\mathbf{l}_{i_0}} = -\iota_{\mathbf{l}_{i_0}} \sum_{i \neq i_0} \iota_{\mathbf{l}_i} c_{\mathbf{l}_i}. \tag{3.107}$$

Therefore, $k_{\mathbf{l}_{i_0}}$ is fixed to be the constant $-\iota_{\mathbf{l}_{i_0}} \sum_{i \neq i_0} \iota_{\mathbf{l}_i} c_{\mathbf{l}_i}$ in $Eq(\widehat{\mathcal{C}})$ and we conclude that $Eq(\mathcal{C}) = Eq(\widehat{\mathcal{C}})$. We thus complete the proof of (3).

By Lemma 3.3.17, we know that the total number of legs $n_{\text{leg}} = n_{fx}(\mathcal{C}) + n_{fr}(\mathcal{C})$

is an even number. This proves (4). \square

After splitting all \circ nodes in Definition 3.3.6, a couple may have several connected components. The following properties of connected components are important in the counting argument.

Definition 3.3.19. 1. **Open and closed couples.** A connected couple is said to be open (resp. closed) if it (resp. does not) contains legs. A connected component of a couple is a open component (resp. closed component) if it is open (resp. closed).

2. **Good couples and bad couples.** A closed couple is a good couple (resp. bad couple) if it (resp. does not) contains at least two dotted edges (see Definition 3.3.6 (2)).

Figure 3.16 provides some examples of couples that are good or bad.



Figure 3.16: Example of a good couple (left) and bad couple (right)

3. **Good nodes and bad nodes in renormalization forest.** Given a pairing p and a node \mathbf{n} in a renormalization forest F , let $T_{\mathbf{n}}$ be the subtree F rooted at \mathbf{n} and $\text{Child}_{\mathbf{n}}$ be the set of all children of \mathbf{n} , then \mathbf{n} is said to be bad if the pairing $p \notin \mathcal{P}(L(T_{\mathbf{n}}) \setminus \cup_{n' \in \text{Child}_{\mathbf{n}}} L(T_{n'}), A)$. Otherwise, \mathbf{n} is said to be good. Here $A = \{k_1, k_2, \dots, k_{2l+1}\}$, $\mathcal{P}(B, A)$ is defined in Definition 3.3.3 (4)

In other word, a node is bad if p does pair all leaves in $L(T_{\mathbf{n}}) \setminus \cup_{n' \in \text{Child}_{\mathbf{n}}} L(T_{n'})$ with leaves in the this set. For example, consider the following renormalization forest (this forest is the same as the forest in Figure 3.4).

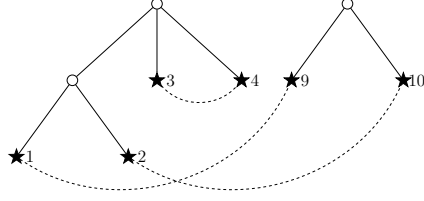


Figure 3.17: Good nodes and bad nodes

Assume that the pairing is $p = \{1, 9\} \cup \{2, 10\} \cup \{3, 4\}$, then the top left node \mathbf{n}_0 is bad because $L(T_{\mathbf{n}_0}) \setminus \cup_{n' \in \text{Child}_{\mathbf{n}_0}} L(T_{\mathbf{n}'}) = \{3, 4\}$ are paired. The other two nodes are good nodes.

4. **Chain of bad circle nodes.** Given a renormalization forest F , a chain of bad nodes is a sequence of bad nodes $\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_l$ such that this sequence have the property that for each $1 \leq i \leq l - 1$, \mathbf{n}_i only have one child \mathbf{n}_{i+1} and this sequence is one of the maximal sequences satisfies this property. We say a chain of bad nodes is long or short, if $l > 1$ or $l = 1$ respectively.

For example, consider the following renormalization forest, the pairing is $p = \{1, 7\} \cup \{2, 8\} \cup \{3, 4\} \cup \{5, 6\}$, then n_1, n_2 form a long chain of bad nodes, and n_2 or n_3 form two short chain of bad nodes

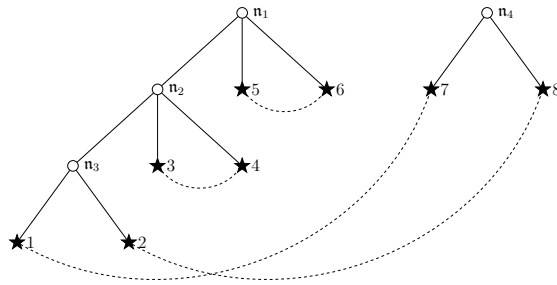


Figure 3.18: Chain of bad nodes

The counting result in the next section says that we lose a factor for each bad node but gain a factor for each good node. We have the following lemma that gives the upper bounds of the number of chains of bad nodes in terms of good nodes.

Lemma 3.3.20. *Given two trees T, T' and a pairing $p \in \mathcal{P}_F$ of leaves in the renormalized pairing set \mathcal{P}_F , let F be the forest which is the union of two forests $R(T)$ and $R(T')$. Then the couple $\mathcal{C} = \mathcal{C}(T, T', p)$ constructed from T, T', p have the following properties.*

Let $n_G(F)$ be the number of good and bad nodes of F respectively. Let $n_{BC}(F)$ be the number of chains of bad nodes of F . Then we have the following inequality.

$$n_{BC}(F) + 1 \leq 2n_G(F) \quad (3.108)$$

Remark 3.3.21. It is crucial that the couple is constructed from a pairing p from \mathcal{P}_F . For a general pairing, the lemma is not true.

Proof. We first consider the following claim.

Claim 1. If a circle node \mathbf{n} in F has no circle child, then \mathbf{n} is good.

Proof. Under this assumption, $\cup_{n' \in \text{Child}_{\mathbf{n}}} L(T_{n'}) = \emptyset$ and $L(T_{\mathbf{n}}) = L(T_{\mathbf{n}}) \setminus \cup_{n' \in \text{Child}_{\mathbf{n}}} L(T_{n'})$. By the definition of \mathcal{P}_F , Definition 3.3.3 (5), $p \notin \mathcal{P}(L(T_{\mathbf{n}}), A)$. Therefore, \mathbf{n} is a good node. \square

Claim 2. The last nodes of a chain of bad circle nodes must either have a good circle child or have at least two bad children.

Proof. If the claim is wrong, the last node has at most one circle child and this child is a bad circle node. If the last node does not have circle children, by claim 1, the last node must be good, which gives a contradiction. If the last node has exactly one circle child which is bad, then this contradicts the maximality of the chain with respect to its defining property. \square

If a chain of bad nodes has no bad ancestors, we call it top chain of bad nodes. Denote all top chain of bad nodes by $\mathbf{c}_1, \dots, \mathbf{c}_k$. Denote by $T_{\mathbf{c}_1} \cdots T_{\mathbf{c}_k}$ the subtree of

the first node of this chain. Then we know that

$$n_{BC}(F) = n_{BC}(T_{\mathbf{c}_1}) + \cdots + n_{BC}(T_{\mathbf{c}_k}) \quad (3.109)$$

$$n_G(T_{\mathbf{c}_1}) + \cdots + n_G(T_{\mathbf{c}_k}) \leq n_G(F) \quad (3.110)$$

By these two inequalities, we know that it suffices to prove (3.108) for trees T whose root is a bad circle node.

We prove (3.108) for a tree T with a bad root by induction on the total number n of bad chains.

If $n = 1$, by claim 2, it has good circle children, so $n_G(T) \geq 1$. We get $n_G(T) + 1 \geq 2 = 2n_{BC}(T)$.

Assume that the Lemma is true for $< n$. We prove the case n .

For the top chain of bad nodes, by claim 2, it either have a good child or two bad children.

The one good child case. In the first case, remove the top chain \mathbf{c} from T to get a forest, then take the subtrees $T_{\mathbf{c}'_1}, \dots, T_{\mathbf{c}'_{k'}}$ of top chains of the new forest. We have

$$n_{BC}(T) = n_{BC}(T_{\mathbf{c}'_1}) + \cdots + n_{BC}(T_{\mathbf{c}'_{k'}}) + 1 \quad (3.111)$$

$$n_G(T_{\mathbf{c}'_1}) + \cdots + n_G(T_{\mathbf{c}'_{k'}}) + 1 \leq n_G(T). \quad (3.112)$$

Here we have the $+1$ in (3.111) for n_G because in the first case, \mathbf{c} has good children.

By induction assumption, we have

$$n_{BC}(T_{\mathbf{c}'_i}) + 1 \leq 2n_G(T_{\mathbf{c}'_i}). \quad (3.113)$$

Summing over i and applying (3.111) and (3.112), we get

$$n_{BC}(T) + k' - 1 \leq 2(n_G(T) - 1). \quad (3.114)$$

Since $k' \geq 0$, this implies (3.108) in the first case.

The two bad child case. Now we consider the second case. Remove the top chain \mathbf{c} from T to get a forest, then take the subtrees $T_{\mathbf{c}'_1}, \dots, T_{\mathbf{c}'_{k'}}$ of top chains of the new forest. We have

$$n_{BC}(T) = n_{BC}(T_{\mathbf{c}'_1}) + \dots + n_{BC}(T_{\mathbf{c}'_{k'}}) + 1 \quad (3.115)$$

$$n_G(T_{\mathbf{c}'_1}) + \dots + n_G(T_{\mathbf{c}'_{k'}}) \leq n_G(T). \quad (3.116)$$

By induction assumption, we have

$$n_{BC}(T_{\mathbf{c}'_i}) + 1 \leq 2n_G(T_{\mathbf{c}'_i}). \quad (3.117)$$

Summing over i and applying (3.111) and (3.112), we get

$$n_{BC}(T) + k' - 1 \leq 2n_G(T). \quad (3.118)$$

Since $k' \geq 2$, this implies (3.108) in the second case, which completes the proof of the lemma. \square

Counting results and their proof

The following proposition gives an upper bound of the number of solutions of (3.79) (or (3.82)).

Proposition 3.3.22. *Let $\mathcal{C} = \mathcal{C}(T, T', p)$ be an open admissible couple, n be the total number of nodes in \mathcal{C} , n_c be the number of closed components of \mathcal{C} and $Q = L^d \sqrt{\omega}$.*

We fix $k \in \mathbb{R}$ for the fixed legs $\mathfrak{l}, \mathfrak{l}'$ and $\sigma_{\mathfrak{n}} \in \mathbb{R}$ for each $\mathfrak{n} \in \mathcal{C}$. Assume that $\omega \geq L^{-1}$ (or $\omega \geq L^{-2}$ in case the ratio of periods is generic in the sense of Theorem 4.3.2). Then we have the following conclusions

(1) The number of solutions M of (3.79) (or (3.82)) is bounded by

$$M \leq L^{O(\theta)} Q^n L^{dn_c}. \quad (3.119)$$

(2) We know that n equals the total number of \bullet nodes in T and T' . Let n_G and n_B be the number of good and bad nodes in T and T' respectively. Then we have

$$\begin{cases} 2n_G \geq n_{BC} + 1 \\ n_G \geq 2(n_c - n_B) \end{cases} \quad (3.120)$$

(3) Let l and l' be the total number of nodes in T and T' respectively. Assume that there is no long chain of bad circle nodes and $\omega \geq L^{-\frac{d}{3}}$. Then we also have

$$M/Q^{l+l'} \leq L^{O(\theta)}. \quad (3.121)$$

This inequality is important because the order of magnitude of $\frac{\alpha^{2l(T)}}{\rho^{l(T)} L^{2dl(T)}} \text{Term}(T, T', p)$ is supposed to be $M/Q^{l+l'}$ when $\omega = T_{kin}^{-1} = \alpha^{-2}$, see (3.156). As long as we can remove the condition that there is no long chain of bad circle nodes, up to epsilon loss the wave kinetic equation can be derived.

Proof. We first prove (2). The fact that n equals the total number of \bullet nodes in T and T' follows from the definition. The first inequality of (3.120) follows from Lemma 3.3.20.

Now we prove the second inequality of (3.120). Remember from the definition of dotted edges, Definition 3.3.6 (2), each dotted edge corresponds to a circle node. By definition of bad circle nodes, components arising from them can only contain one

dotted edge, so $n_c - n_B$ is the number of couples containing at least two dotted edges. All these couples correspond to at least two good circle nodes and thus the number of them is less than $n_G/2$. Therefore, we proved the desired inequality.

Then we prove (3) from (1) and (2). Let n_{circ} be the total number of circle nodes in T and T' . Since $l + l'$ and n are the total number of nodes and \bullet nodes in T and T' respectively, we know that $l + l' = n + n_{\text{circ}}$. By definition of n_G and n_B , we know that $n = n_G + n_B$.

By the assumption that there is no long chain of bad circle nodes, we know that $n_B = n_{BC}$. Therefore, the first inequality of (3.120) implies

$$n_G \geq \frac{1}{2}n_B. \quad (3.122)$$

Multiply this inequality by $\frac{2}{5}$, the second inequality of (3.120) $\frac{3}{5}$ and then sum them. We get the following inequality.

$$n_G \geq \frac{6}{5}n_c - n_B \quad \Leftrightarrow \quad n_c \leq \frac{5}{6}(n_G + n_B) \quad (3.123)$$

By (1), we get

$$\begin{aligned} M/Q^{l+l'} &\leq L^{O(\theta)} Q^n L^{dn_c} / Q^{n+n_G+n_B} \\ &= L^{O(\theta)} L^{dn} (\sqrt{\omega})^n L^{dn_c} / (L^{d(n+n_G+n_B)} (\sqrt{\omega})^{n+n_G+n_B}) \\ &= L^{O(\theta)} (\sqrt{\omega})^{-n_G-n_B} L^{d(n_c-n_G-n_B)} \\ &= L^{O(\theta)} (L^{-\frac{d}{6}} \sqrt{\omega}^{-1})^{n_G+n_B} \leq L^{O(\theta)}. \end{aligned} \quad (3.124)$$

Here in the third inequality, we applied (3.123) and $\omega \geq L^{-\frac{d}{3}}$ in the hypothesis of (3).

Therefore, we have finished the proof of (3)

The proof of (1) is lengthy and therefore divided into several steps. To present its proof, let us first introduce several useful notations.

Given a couple \mathcal{C} and $k, \sigma_{\mathbf{n}}$, let $Eq(\mathcal{C})$ be the system of equation (3.82) constructed in Proposition 3.3.11. For any system of equations Eq , let $\#(Eq)$ be its number of solutions.

Now we prove (1).

Step 1. In this step, we reduce Proposition 3.3.22 to the case of connected couples.

By Lemma (3.87), we know that

$$\#(Eq(\mathcal{C})) = \#(Eq(\tilde{\mathcal{C}})) \prod_{j=1}^{n_c} \#(Eq(\mathcal{C}_j)). \quad (3.125)$$

Here $\tilde{\mathcal{C}}$ is the open component of \mathcal{C} (it is not difficult to show that \mathcal{C} only has one open component), \mathcal{C}_j , $1 \leq j \leq n_c$ are the closed components of \mathcal{C} and we have $\mathcal{C} = \tilde{\mathcal{C}} \cup (\cup_{j=1}^{n_c} \mathcal{C}_j)$.

Therefore, we just need to prove the following proposition which claims that Proposition 3.3.22 is true for connected couples. Note that in the statement of the following proposition, we have an extra factor L^d for each closed component. In total, we get a factor L^{dn_c} . This how we get the L^{dn_c} in $M \leq L^{O(\theta)} Q^n L^{dn_c}$ of Proposition 3.3.22.

Proposition 3.3.23. *Let \mathcal{C} be a connected couple and n be the total number of nodes in \mathcal{C} . We fix $\sigma_{\mathbf{n}} \in \mathbb{R}$ for each $\mathbf{n} \in \mathcal{C}$ and $k \in \mathbb{R}$ if \mathcal{C} is open. Assume that $\omega \geq L^{-1}$ (or $\omega \geq L^{-2}$ in case the ratio of periods is generic in the sense of Theorem 4.3.2). Then (recall that $Q = L^d \sqrt{\omega}$)*

$$\#Eq(\mathcal{C}) \leq \begin{cases} L^{O(\theta)} Q^n, & \text{if } \mathcal{C} \text{ is open,} \\ L^{d+O(\theta)} Q^n, & \text{if } \mathcal{C} \text{ is closed.} \end{cases} \quad (3.126)$$

Steps 2-4 are devoted to the proof of Proposition 3.3.23 using a cutting edge argument.

Step 2. Although all components can have at most two legs at the beginning, after cutting, the resulting couples can have more than two legs. In this step, we introduce a stronger version of Proposition 3.3.23 that covers legs of more than two legs.

Lemma 3.3.24. *Let \mathcal{C} be a connected couple with multiple legs, n be the total number of nodes in \mathcal{C} and n_e (resp. n_{fx} , n_{fr}) be the total number of non-leg edges (resp. fixed legs, free legs). We fix $\sigma_{\mathbf{n}} \in \mathbb{R}$ for each $\mathbf{n} \in \mathcal{C}$ and $c_{\mathbf{l}} \in \mathbb{R}$ for each fixed leg \mathbf{l} . Assume that $\omega \geq L^{-1}$ (or $\omega \geq L^{-2}$ in case the ratio of periods is generic in the sense of Theorem 4.3.2). We assume that $n_{fx}, n_{fr} \neq 0$. Then*

$$\sup_{\{c_{\mathbf{l}}\}_{\mathbf{l}}} \#Eq(\mathcal{C}, \{c_{\mathbf{l}}\}_{\mathbf{l}}) \leq L^{O(\theta)} Q^{\chi(\mathcal{C})} = L^{O(\theta)} Q^{n_e + n_{fr} - n}. \quad (3.127)$$

Proposition 3.3.24 will be proved in steps 3 and 4. In the rest part of this step, we show that Proposition 3.3.23 is a corollary of Proposition 3.3.24.

Since the couples in Proposition 3.3.23 and Proposition 3.3.22 have no free leg, these two propositions are not obvious corollaries of Proposition 3.3.24.

Proof of Proposition 3.3.23. The closed couple case of Proposition 3.3.23 can be derived from the open couple case by the following argument.

Let \mathcal{C} be a closed couple, choose an edge $\mathbf{e}_* \in \mathcal{C}$ which connects $\mathbf{n}_1, \mathbf{n}_2$. Cut this edge into two legs \mathbf{l}_1 and \mathbf{l}_2 , then the couple $\tilde{\mathcal{C}}$ obtained in this way is an open couple. Therefore, we have

$$\begin{aligned}
\#Eq(\mathcal{C}) &= \#\{k_{\mathfrak{e}} \in \mathbb{Z}^d, |k_{\mathfrak{e}}| \lesssim L^+, \forall \mathfrak{e} \in \mathcal{C} : MC_{\mathfrak{n}}, EC_{\mathfrak{n}}, \forall \mathfrak{n} \in \mathcal{C}\} \\
&= \#\{k_{\mathfrak{e}} \in \mathbb{Z}^d, |k_{\mathfrak{e}}| \lesssim L^+, \forall \mathfrak{e} \in \mathcal{C} : MC_{\mathfrak{n}}, EC_{\mathfrak{n}}, \forall \mathfrak{n} \in \mathcal{C}. MC_{n_1}(k_{\mathfrak{e}_*}, \dots), MC_{n_2}(k_{\mathfrak{e}_*}, \dots)\} \\
&= \# \sum_{k_{\mathfrak{e}_*}} \{k_{\mathfrak{e}} \in \mathbb{Z}^d, |k_{\mathfrak{e}}| \lesssim L^+, \forall \mathfrak{e} \in \mathcal{C} : MC_{\mathfrak{n}}, EC_{\mathfrak{n}}, \forall \mathfrak{n} \in \mathcal{C}. MC_{n_1}(k_{l_1}, \dots), \\
&\quad MC_{n_2}(k_{l_2}, \dots), k_{l_1} = k_{l_2} = k_{\mathfrak{e}_*}\} \\
&= \sum_{|k_{\mathfrak{e}_*}| \lesssim 1} \#Eq(\tilde{\mathcal{C}}, \{k_{\mathfrak{e}_*}, k_{\mathfrak{e}_*}\})
\end{aligned} \tag{3.128}$$

If we can prove the open couple case, we have $\#Eq(\tilde{\mathcal{C}}, \{k_{\mathfrak{e}_*}, k_{\mathfrak{e}_*}\}) \leq L^{O(\theta)} Q^n$. Then by (3.128), we can prove the closed couple case:

$$\#Eq(\mathcal{C}) = \sum_{|k_{\mathfrak{e}_*}| \lesssim 1} \#Eq(\tilde{\mathcal{C}}, \{k_{\mathfrak{e}_*}, k_{\mathfrak{e}_*}\}) \lesssim L^{d+O(\theta)} Q^n. \tag{3.129}$$

The open couple case can be proved by Lemma 3.3.18.

An open couple has two fixed legs $\mathfrak{l}, \mathfrak{l}'$ with opposite sign, so by momentum conservation equation (3.100), we have $k_{\mathfrak{l}} = k_{\mathfrak{l}'}$, and $k_{\mathfrak{l}} = k_{\mathfrak{l}'} = k$ is a consequence of $k_{\mathfrak{l}} = k$. Therefore, we have

$$\begin{aligned}
Eq(\mathcal{C}) &= \{k_{\mathfrak{e}} \in \mathbb{Z}^d, |k_{\mathfrak{e}}| \lesssim L^+, \forall \mathfrak{e} \in \mathcal{C} : MC_{\mathfrak{n}}, EC_{\mathfrak{n}}, \forall \mathfrak{n} \in \mathcal{C}. k_{\mathfrak{l}} = k_{\mathfrak{l}'} = k.\} \\
&= \{k_{\mathfrak{e}} \in \mathbb{Z}^d, |k_{\mathfrak{e}}| \lesssim L^+, \forall \mathfrak{e} \in \mathcal{C} : MC_{\mathfrak{n}}, EC_{\mathfrak{n}}, \forall \mathfrak{n} \in \mathcal{C}. k_{\mathfrak{l}} = k.\} \\
&= Eq(\widehat{\mathcal{C}}, k)
\end{aligned} \tag{3.130}$$

where $\widehat{\mathcal{C}}$ is the couple obtained by replacing a fixed leg with a free leg.

In $\widehat{\mathcal{C}}$, $n_{fx} = n_{fr} = 1 \neq 0$, so by (3.94), we have $n_e + 1 = 2n$. Applying (3.127), we get

$$Eq(\mathcal{C}) = Eq(\widehat{\mathcal{C}}, k) \lesssim L^{O(\theta)} Q^{n_e + n_{fr} - n} = L^{O(\theta)} Q^n. \tag{3.131}$$

This proves Proposition 3.3.23. □

Step 3. In this step, we prove Proposition 3.3.24 for basic counting units.

To prove Proposition 3.3.24, we keep cutting nodes connected to legs from the couple \mathcal{C} . Suppose that \mathcal{C} has n nodes and we cut a node \mathfrak{n} from \mathcal{C} using cut $c(\mathfrak{n})$. Let $\mathcal{C}_{\mathfrak{n}}$ be the smaller components contains fixed leg \mathfrak{l} with 1 nodes \mathfrak{n} and $\mathcal{C}' = \mathcal{C} \setminus \mathcal{C}_{\mathfrak{n}}$ be the larger components with $n - 1$ nodes.

As in Figure 3.19, there are three possibilities of $\mathcal{C}_{\mathfrak{n}}$. We label them by \mathcal{C}_I , \mathcal{C}_{II} , \mathcal{C}_{III} .

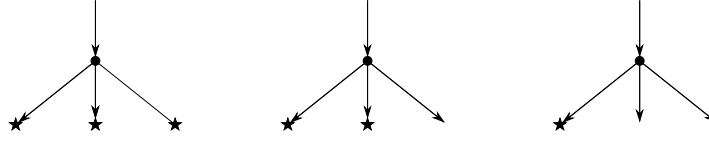


Figure 3.19: Three possibilities of $\mathcal{C}_{\mathfrak{l}}$.

Lemma 3.3.25. \mathcal{C}_I , \mathcal{C}_{II} , \mathcal{C}_{III} satisfy the bound (3.127) in Proposition 3.3.24. In other words, fix c_1 (resp. c_1, c_3 and c_1, c_2, c_3) for the legs of \mathcal{C}_I (resp. \mathcal{C}_{II} and \mathcal{C}_{III}), then we have

$$\#Eq(\mathcal{C}_I) \leq L^\theta Q^2, \quad \#Eq(\mathcal{C}_{II}) \leq L^\theta Q, \quad \#Eq(\mathcal{C}_{III}) \leq Q^0 = 1. \quad (3.132)$$

Proof. Given c_1, c_2, c_3 the equation of \mathcal{C}_{III} is

$$\begin{cases} k_{\epsilon_1} - k_{\epsilon_2} + k_{\epsilon_3} - k_{\epsilon} = 0, & k_{\epsilon_2} = c_2, & k_{\epsilon_3} = c_3, & k_{\epsilon} = c_1, \\ \Lambda_{k_{\epsilon_1}} - \Lambda_{k_{\epsilon_2}} + \Lambda_{k_{\epsilon_3}} - \Lambda_{k_{\epsilon}} = \sigma_{\mathfrak{n}} + O(\omega) \end{cases} \quad (3.133)$$

It's obvious that there is at most one solution to this system of equations.

Given c_1 , the equation of \mathcal{C}_I is

$$\begin{cases} k_{\epsilon_1} - k_{\epsilon_2} + k_{\epsilon_3} - k_{\epsilon} = 0, & k_{\epsilon} = c_1, \\ \Lambda_{k_{\epsilon_1}} - \Lambda_{k_{\epsilon_2}} + \Lambda_{k_{\epsilon_3}} - \Lambda_{k_{\epsilon}} = \sigma_{\mathbf{n}} + O(\omega) \end{cases} \quad (3.134)$$

By Theorem 4.3.1 or 4.3.2 (4.31) or (4.34), corresponding to the general or generic case, the number of solutions of the above system of equations can be bounded by $L^\theta Q^2$.

Given c_1, c_2 , the equation of \mathcal{C}_{II} is

$$\begin{cases} k_{\epsilon_1} - k_{\epsilon_2} + k_{\epsilon_3} - k_{\epsilon} = 0, & k_{\epsilon} = c_1, \quad k_{\epsilon_3} = c_3, \\ \Lambda_{k_{\epsilon_1}} - \Lambda_{k_{\epsilon_2}} + \Lambda_{k_{\epsilon_3}} - \Lambda_{k_{\epsilon}} = \sigma_{\mathbf{n}} + O(\omega) \end{cases} \quad (3.135)$$

By Theorem 4.3.1 or 4.3.2 (4.33) or (4.35), corresponding to the general or generic case, the number of solutions of the above system of equations can be bounded by $L^\theta Q$

Therefore, we complete the proof of this lemma. □

Step 4. In this step, we apply the edge cutting argument to prove Proposition 3.3.24 by induction.

If \mathcal{C} has only one node ($n = 1$), then \mathcal{C} equals to \mathcal{C}_I , \mathcal{C}_{II} or \mathcal{C}_{III} and Proposition 3.3.24 in this case follows from Lemma 3.3.25.

Suppose that Proposition 3.3.24 holds true for couples with the number of nodes $\leq n - 1$. We prove it for couples with the number of nodes n .

Since $n_{fx}, n_{fr} \neq 0$, there exists both fixed and free legs. We will cut a node $\mathbf{n}_* \in \mathcal{C}$ which is connected to a fixed leg. Denoted by \mathcal{N}_{leg} the set of all such nodes and by $\mathcal{C}_{\mathbf{n}_*}$ the components containing \mathbf{n}_* after cutting. There are several different situations in this cutting.

Case 1. Assume that for all $\mathbf{n} \in \mathcal{N}_{leg}$, after cutting \mathbf{n} , $\mathcal{C} \setminus \mathcal{C}_{\mathbf{n}}$ is still connected.

Case 1.1. Assume that not all legs are connected to one node. One example of this case is shown in Figure 3.20.

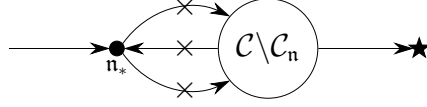


Figure 3.20: Example of case 1.1.

In this case, we have the following claim.

Claim. There exists a node $\mathbf{n}_* \in \mathcal{N}_{leg}$ such that not all free legs are connected to this node.

Proof. If there is only one node in $\mathbf{n} \in \mathcal{N}_{leg}$, then all fixed legs are connected to this node. By this, we know that not all free legs are connected to this node because, by the assumption of case 1.1, not all legs are connected to one node. We choose \mathbf{n}_* to be this node.

If there are two nodes $\mathbf{n}_1, \mathbf{n}_2 \in \mathcal{N}_{leg}$, and if all free legs are connected to one of them, then choose \mathbf{n}_* to be the other one. Otherwise, choose \mathbf{n}_* to either one of $\mathbf{n}_1, \mathbf{n}_2$ works. \square

We cut the node \mathbf{n}_* in this claim from \mathcal{C} . Let $\mathcal{C}_n, \mathcal{C}_1 = \mathcal{C} \setminus \mathcal{C}_n$ be two components after cutting. As in Lemma 3.3.16, let $\{\mathbf{e}_i^{(1)}\} \subseteq \mathcal{C}_n$ be free legs and $\{\mathbf{e}_i^{(2)}\} \subseteq \mathcal{C}_1$ be fixed legs, then applying Lemma 3.3.16 gives

$$\begin{aligned}
\sup_{\{c_l\}_l} \#Eq(\mathcal{C}, \{c_l\}_l) &\leq \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_n)}} \#Eq(\mathcal{C}_n, \{c_{l_1}\}) \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}_1)}} \#Eq(\mathcal{C}_1, \{c_{l_2}\}) \\
&\lesssim L^\theta Q^{\chi(\mathcal{C}_n)} L^{O(\theta)} Q^{\chi(\mathcal{C}_1)} = L^{O(\theta)} Q^{\chi(\mathcal{C}_n) + \chi(\mathcal{C}_1)} \\
&= L^{O(\theta)} Q^{\chi(\mathcal{C})}.
\end{aligned} \tag{3.136}$$

Here the second inequality follows from the induction assumption and Lemma 3.3.25. Because it can be verified that in both \mathcal{C}_n and \mathcal{C}_1 , n_{fx} and n_{fr} are not equal to 0, the induction assumption and Lemma 3.3.25 are applicable.

Now we provide a detailed analysis of why in both \mathcal{C}_n and \mathcal{C}_1 , n_{fx} and n_{fr} are not equal to 0.

\mathcal{C}_n contains a fixed leg because $\mathbf{n}_* \in \mathcal{N}_{leg}$ and by definition of \mathcal{N}_{leg} , \mathbf{n}_* is connected to a fixed leg. \mathcal{C}_n contains a free leg because it contains free legs $\{\mathfrak{e}_i^{(1)}\} \subseteq \mathcal{C}_n$.

By the claim, since not all free legs are connected to \mathbf{n}_* , \mathcal{C}_1 contains a free leg. \mathcal{C}_1 contains a fixed leg because it contains fixed legs $\{\mathfrak{e}_i^{(2)}\} \subseteq \mathcal{C}_1$.

Case 1.2. Assume that all legs are connected to one node \mathbf{n}_* . One example of this case is shown in Figure 3.21

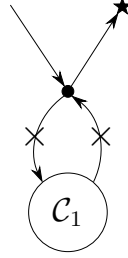


Figure 3.21: Example of case 1.2.

We cut \mathbf{n}_* from \mathcal{C} . Let $\mathcal{C}_n, \mathcal{C}_1 = \mathcal{C} \setminus \mathcal{C}_n$ be two components after cutting.

By (3.94), we know that $n_{fx} + n_{fr}$ is an even number, so it must equal to 2. Then we know that $n_{fx} = n_{fr} = 1$ and \mathcal{C} has one fixed leg \mathfrak{l} and one free leg \mathfrak{le} . The other two edges $\mathfrak{e}', \mathfrak{e}''$ from \mathbf{n}_* should be connected to \mathcal{C}_1 .

Unlike case 1.1 we cannot simply apply induction assumption to bound $\#Eq(\mathcal{C}_1, \{c_{l_2}\})$. Because in this case \mathcal{C}_1 does not have free legs ($n_{fx} = 2, n_{fr} = 0$). See Figure 3.21.

As in Lemma 3.3.16, $\mathfrak{e}', \mathfrak{e}''$ are cut into $\mathfrak{e}'^{(1)}, \mathfrak{e}''^{(1)}$ and $\mathfrak{e}'^{(2)}, \mathfrak{e}''^{(2)}$. $\mathfrak{e}'^{(1)}, \mathfrak{e}''^{(1)} \in \mathcal{C}_n$ are free legs and $\mathfrak{e}'^{(2)}, \mathfrak{e}''^{(2)} \in \mathcal{C}_1$ are fixed legs, then applying Lemma 3.3.16 (3.87) gives

$$Eq(\mathcal{C}, c_l) = \{(k_{\mathfrak{e}_1}, k_{\mathfrak{e}_2}) : k_{\mathfrak{e}_1} \in Eq(\mathcal{C}_n, c_l), k_{\mathfrak{e}_2} \in Eq(\mathcal{C}_1, \{k_{\mathfrak{e}'^{(1)}}, k_{\mathfrak{e}''^{(1)}}\}_i)\}. \quad (3.137)$$

By (3.100), we know that $k_{\epsilon'(1)} = k_{\epsilon''(1)}$. Therefore, we have

$$Eq(\mathcal{C}, c_l) = \{(k_{\epsilon_1}, k_{\epsilon_2}) : k_{\epsilon_1} \in Eq(\mathcal{C}_n, c_l) \cap \{k_{\epsilon'(1)} = k_{\epsilon''(1)}\}, k_{\epsilon_2} \in Eq(\mathcal{C}_1, \{k_{\epsilon'(1)}, k_{\epsilon''(1)}\}_i)\}. \quad (3.138)$$

By definition, we know that

$$\begin{aligned} Eq(\mathcal{C}_n, c_l) &= \{(k_l, k_{lf}, k_{\epsilon'}, k_{\epsilon''}) : |k_l|, |k_{lf}|, |k_{\epsilon'}|, |k_{\epsilon''}| \lesssim L^+, MC_{n*}, EC_{n*}, k_l = c_l\} \\ &= \{(k_l, k_{lf}, k_{\epsilon'}, k_{\epsilon''}) : |k_{lf}|, |k_{\epsilon'}|, |k_{\epsilon''}| \lesssim L^+, \\ &\quad k_{\epsilon'} - k_{\epsilon''} + k_{lf} - k_l = 0, k_{lf} \neq k_{\epsilon'} \neq k_{\epsilon''}, \text{ or } k_{\epsilon'} = k_{\epsilon''} = k_{lf} = k_l \\ &\quad \Lambda_{k_{\epsilon'}} - \Lambda_{k_{\epsilon''}} + \Lambda_{k_{lf}} - \Lambda_{k_l} = \sigma_n + O(\omega), k_l = c_l\} \end{aligned} \quad (3.139)$$

Therefore, we have

$$Eq(\mathcal{C}_n, c_l) \cap \{k_{\epsilon'(1)} = k_{\epsilon''(1)}\} = \{(k_l, k_{lf}, k_{\epsilon'}, k_{\epsilon''}) : k_{\epsilon'} = k_{\epsilon''} = k_{lf} = k_l = c_l\}. \quad (3.140)$$

By (3.138), we have

$$\#Eq(\mathcal{C}, c_l) = \#Eq(\mathcal{C}_1, \{c_l, c_l\}). \quad (3.141)$$

Let $\widehat{\mathcal{C}}_1$ be the couple obtained by replacing a fixed leg with a free leg in \mathcal{C}_1 . Using a similar argument to (3.130) and (3.131) and then applying the induction assumption, we get

$$\begin{aligned} \#Eq(\mathcal{C}, c_l) &= \#Eq(\mathcal{C}_1, \{c_l, c_l\}) = \#Eq(\widehat{\mathcal{C}}_1, \{c_l\}) \\ &\lesssim L^{O(\theta)} Q^{\chi(\widehat{\mathcal{C}}_1)} = L^{O(\theta)} Q^{\chi(\mathcal{C})-1} \leq L^{O(\theta)} Q^{\chi(\mathcal{C})} \end{aligned} \quad (3.142)$$

Here the last step is because $\chi(\mathcal{C}) = \chi(\mathcal{C}_n) + \chi(\mathcal{C}_1) = 2 + \chi(\mathcal{C}_1) = 1 + \chi(\widehat{\mathcal{C}}_1)$.

Case 2. Assume that for for some $\mathbf{n}_* \in \mathcal{N}_{leg}$, after cutting \mathbf{n}_* , $\mathcal{C} \setminus \mathcal{C}_n$ has exactly

three components. One example of this case is shown in Figure 3.22

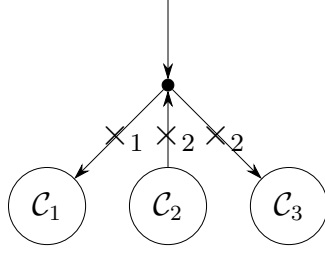


Figure 3.22: Example of case 2. \times_1 and \times_2 indicate the edges cut in the first and second time respectively.

We cut \mathbf{n}_* from \mathcal{C} . Let $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$ be the three components of $\mathcal{C} \setminus \mathcal{C}_n$ after cutting.

Since $\mathbf{n}_* \in \mathcal{N}_{leg}$, there should be a fixed leg \mathfrak{l} that is connected to \mathbf{n}_* . The other three edges $\mathfrak{e}', \mathfrak{e}'', \mathfrak{e}'''$ from \mathbf{n}_* should be connected to $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$ respectively.

Since there exists at least one free leg in \mathcal{C} ($n_{fr} \neq 0$), at least one of $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$ contain a free leg. Without loss of generality let us assume that \mathcal{C}_1 contains a free leg.

Let us first cut \mathfrak{e}' into $\mathfrak{e}'^{(1)} \in \mathcal{C} \setminus \mathcal{C}_1 = \mathcal{C}_{n_*} \cup \mathcal{C}_2 \cup \mathcal{C}_3$ and $\mathfrak{e}'^{(2)} \in \mathcal{C}_1$. Let $\mathfrak{e}'^{(1)}$ be a free leg and $\mathfrak{e}'^{(2)}$ be a fixed leg. Therefore, \mathcal{C}_1 contains both fixed and free legs. Applying Lemma 3.3.16 gives

$$\begin{aligned} \sup_{\{c_l\}_l} \#Eq(\mathcal{C}, \{c_l\}_l) &\leq \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C} \setminus \mathcal{C}_1)}} \#Eq(\mathcal{C} \setminus \mathcal{C}_1, \{c_{l_1}\}) \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}_1)}} \#Eq(\mathcal{C}_1, \{c_{l_2}\}) \\ &\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_1)} \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C} \setminus \mathcal{C}_1)}} \#Eq(\mathcal{C} \setminus \mathcal{C}_1, \{c_{l_1}\}). \end{aligned} \tag{3.143}$$

Here in the second inequality, we can apply the induction assumption to \mathcal{C}_1 because it contains both fixed and free legs.

Then we cut \mathfrak{e}'' into $\mathfrak{e}''^{(1)} \in \mathcal{C}_{n_*} \cup \mathcal{C}_3$ and $\mathfrak{e}''^{(2)} \in \mathcal{C}_2$. Let $\mathfrak{e}''^{(1)}$ be a free leg (or fix leg) and $\mathfrak{e}''^{(2)}$ be a fixed leg (or a free leg) if \mathcal{C}_2 contains a free leg (or a fixed leg). Therefore, \mathcal{C}_2 contains both fixed and free legs. Applying Lemma 3.3.16 and (3.143)

gives

$$\begin{aligned}
\sup_{\{c_l\}_l} \#Eq(\mathcal{C}, \{c_l\}_l) &\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_1)} \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C} \setminus \mathcal{C}_1)}} \#Eq(\mathcal{C} \setminus \mathcal{C}_1, \{c_{l_1}\}) \\
&\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_1)} \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_{n_*} \cup \mathcal{C}_3)}} \#Eq(\mathcal{C}_{n_*} \cup \mathcal{C}_3, \{c_{l_1}\}) \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}_2)}} \#Eq(\mathcal{C}_2, \{c_{l_2}\}) \quad (3.144) \\
&\lesssim L^\theta Q^{\chi(\mathcal{C}_1) + \chi(\mathcal{C}_2)} \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_{n_*} \cup \mathcal{C}_3)}} \#Eq(\mathcal{C}_{n_*} \cup \mathcal{C}_3, \{c_{l_1}\})
\end{aligned}$$

Here in the third inequality, we can apply the induction assumption to \mathcal{C}_2 because it contains both fixed and free legs.

The argument for \mathcal{C}_3 is similar to that of \mathcal{C}_2 . We cut \mathfrak{e}''' into $\mathfrak{e}'''^{(1)} \in \mathcal{C}_{n_*}$ and $\mathfrak{e}'''^{(2)} \in \mathcal{C}_3$. Let $\mathfrak{e}'''^{(1)}$ be a free leg (or a fixed leg) and $\mathfrak{e}'''^{(2)}$ be a fixed leg (or a free leg) if \mathcal{C}_3 contains a free leg (or a fixed leg). Therefore, \mathcal{C}_3 contains both fixed and free legs. Applying Lemma 3.3.16 and (3.144) gives

$$\begin{aligned}
\sup_{\{c_l\}_l} \#Eq(\mathcal{C}, \{c_l\}_l) &\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_1) + \chi(\mathcal{C}_2)} \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_{n_*} \cup \mathcal{C}_3)}} \#Eq(\mathcal{C}_{n_*} \cup \mathcal{C}_3, \{c_{l_1}\}) \\
&\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_1) + \chi(\mathcal{C}_2)} \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_{n_*})}} \#Eq(\mathcal{C}_{n_*}, \{c_{l_1}\}) \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}_3)}} \#Eq(\mathcal{C}_3, \{c_{l_2}\}) \\
&\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_1) + \chi(\mathcal{C}_2)} L^{O(\theta)} Q^{\chi(\mathcal{C}_{n_*})} L^\theta Q^{\chi(\mathcal{C}_3)} \quad (3.145)
\end{aligned}$$

Here in the third inequality, we can apply Lemma 3.3.25 and the induction assumption to \mathcal{C}_{n_*} and \mathcal{C}_3 because they contain both fixed and free legs. \mathcal{C}_{n_*} contains fixed legs (or free legs) because $\mathbf{n}_* \in \mathcal{N}_{leg}$ (or $\mathfrak{e}'^{(1)}$ is a free legs from \mathbf{n}_*).

Case 3. Assume that for all $\mathbf{n} \in \mathcal{N}_{leg}$, after cutting \mathbf{n} , $\mathcal{C} \setminus \mathcal{C}_n$ has exactly two components.

Choose arbitrarily $\mathbf{n}_* \in \mathcal{N}_{leg}$. We cut \mathbf{n}_* from \mathcal{C} . Let $\mathcal{C}_1, \mathcal{C}_2$ be the three components of $\mathcal{C} \setminus \mathcal{C}_n$ after cutting. Since $\mathbf{n}_* \in \mathcal{N}_{leg}$, there should be a fixed leg l that is connected to \mathbf{n}_* . There are several possibilities of the distribution of other three

edges \mathfrak{e}' , \mathfrak{e}'' , \mathfrak{e}''' from \mathfrak{n}_* .

Case 3.1. One of \mathfrak{e}' , \mathfrak{e}'' , \mathfrak{e}''' connects \mathfrak{n}_* and \mathcal{C}_1 (w.l.o.g. assume that it's \mathfrak{e}'), another one connects \mathfrak{n}_* and \mathcal{C}_2 (w.l.o.g. assume that it's \mathfrak{e}'') and the last one is a leg (w.l.o.g. assume that it's \mathfrak{e}'''). One example of this case is shown in Figure 3.23

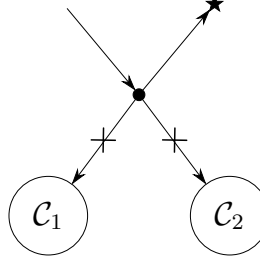


Figure 3.23: Example of case 3.1.

Case 3.1.1. Assume that \mathfrak{e}''' is an fixed leg. Then one of \mathcal{C}_1 and \mathcal{C}_2 contains a free leg (w.l.o.g. assume that it's \mathcal{C}_1).

We apply the same argument as case 2 in this case. Cut \mathfrak{e}' into $\mathfrak{e}'^{(1)} \in \mathcal{C}_{\mathfrak{n}_*} \cup \mathcal{C}_2$, and $\mathfrak{e}'^{(2)} \in \mathcal{C}_1$ and then cut \mathfrak{e}'' into $\mathfrak{e}''^{(1)} \in \mathcal{C}_{\mathfrak{n}_*}$ and $\mathfrak{e}''^{(2)} \in \mathcal{C}_2$. Let $\mathfrak{e}'^{(1)}$ be a free leg and $\mathfrak{e}'^{(2)}$ be a fixed leg. Let $\mathfrak{e}''^{(1)}$ be a free leg (or a fixed leg) and $\mathfrak{e}''^{(2)}$ be a fixed leg (or a free leg) if \mathcal{C}_2 contains a free leg (or a fixed leg). Cutting in this way, all couples contain both fixed and free legs.

As in case 2, we have

$$\begin{aligned}
\sup_{\{c_i\}_I} \#Eq(\mathcal{C}, \{c_i\}_I) &\leq \sup_{\{c_{i_1}\}_{i_1 \in \text{leg}(\mathcal{C}_{\mathfrak{n}_*} \cup \mathcal{C}_2)}} \#Eq(\mathcal{C}_{\mathfrak{n}_*} \cup \mathcal{C}_2, \{c_{i_1}\}) \sup_{\{c_{i_2}\}_{i_2 \in \text{leg}(\mathcal{C}_1)}} \#Eq(\mathcal{C}_1, \{c_{i_2}\}) \\
&\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_1)} \sup_{\{c_{i_1}\}_{i_1 \in \text{leg}(\mathcal{C}_{\mathfrak{n}_*} \cup \mathcal{C}_2)}} \#Eq(\mathcal{C}_{\mathfrak{n}_*} \cup \mathcal{C}_2, \{c_{i_1}\}) \\
&\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_1)} \sup_{\{c_{i_1}\}_{i_1 \in \text{leg}(\mathcal{C}_{\mathfrak{n}_*})}} \#Eq(\mathcal{C}_{\mathfrak{n}_*}, \{c_{i_1}\}) \sup_{\{c_{i_2}\}_{i_2 \in \text{leg}(\mathcal{C}_2)}} \#Eq(\mathcal{C}_2, \{c_{i_2}\}) \\
&\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_1) + \chi(\mathcal{C}_2) + \chi(\mathcal{C}_{\mathfrak{n}_*})} = L^{O(\theta)} Q^{\chi(\mathcal{C})}.
\end{aligned} \tag{3.146}$$

Case 3.1.2. Assume that \mathfrak{e}''' is an free leg.

We apply the same argument as case 2 and case 3.1.1 in this case. Cut \mathfrak{e}' into $\mathfrak{e}'^{(1)} \in \mathcal{C}_{n_*} \cup \mathcal{C}_2$ and $\mathfrak{e}'^{(2)} \in \mathcal{C}_1$, and then cut \mathfrak{e}'' into $\mathfrak{e}''^{(1)} \in \mathcal{C}_{n_*}$ and $\mathfrak{e}''^{(2)} \in \mathcal{C}_2$. Let $\mathfrak{e}'^{(1)}$ be a free leg (or a fixed leg) and $\mathfrak{e}'^{(2)}$ be a fixed leg (or a free leg) if \mathcal{C}_1 contains a free leg (or a fixed leg). Let $\mathfrak{e}''^{(1)}$ be a free leg (or a fixed leg) and $\mathfrak{e}''^{(2)}$ be a fixed leg (or a free leg) if \mathcal{C}_2 contains a free leg (or a fixed leg). Cutting in this way, all couples contain both fixed and free legs.

As in case 3.1.1, we have

$$\begin{aligned}
\sup_{\{c_i\}_i} \#Eq(\mathcal{C}, \{c_i\}_i) &\leq \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_{n_*} \cup \mathcal{C}_2)}} \#Eq(\mathcal{C}_{n_*} \cup \mathcal{C}_2, \{c_{l_1}\}) \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}_1)}} \#Eq(\mathcal{C}_1, \{c_{l_2}\}) \\
&\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_1)} \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_{n_*} \cup \mathcal{C}_2)}} \#Eq(\mathcal{C}_{n_*} \cup \mathcal{C}_2, \{c_{l_1}\}) \\
&\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_1)} \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_{n_*})}} \#Eq(\mathcal{C}_{n_*}, \{c_{l_1}\}) \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}_2)}} \#Eq(\mathcal{C}_2, \{c_{l_2}\}) \\
&\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_1) + \chi(\mathcal{C}_2) + \chi(\mathcal{C}_{n_*})} = L^{O(\theta)} Q^{\chi(\mathcal{C})}.
\end{aligned} \tag{3.147}$$

Case 3.2. Two of \mathfrak{e}' , \mathfrak{e}'' , \mathfrak{e}''' (w.l.o.g. assume that they are \mathfrak{e}' , \mathfrak{e}'') connects n_* and one of the two components (w.l.o.g. assume that it's \mathcal{C}_1). The last one (w.l.o.g. assume that it's \mathfrak{e}''') connects n_* and the last components (w.l.o.g. assume that it's \mathcal{C}_2). In this case \mathcal{C}_2 must contain a leg by (3.94).

One example of this case is shown in Figure 3.24

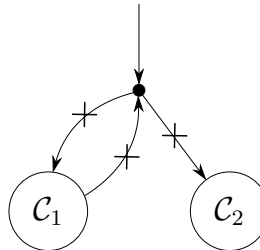


Figure 3.24: Example of case 3.2.

Case 3.2.1. Assume that \mathcal{C}_1 contains a leg.

We apply the same argument as case 2 or case 3.1 in this case. Cut \mathfrak{e}' , \mathfrak{e}'' into $\mathfrak{e}'^{(1)}, \mathfrak{e}''^{(1)} \in \mathcal{C}_{n_*} \cup \mathcal{C}_2$ and $\mathfrak{e}'^{(2)}, \mathfrak{e}''^{(2)} \in \mathcal{C}_1$, and then cut \mathfrak{e}''' into $\mathfrak{e}'''^{(1)} \in \mathcal{C}_{n_*}$ and $\mathfrak{e}'''^{(2)} \in \mathcal{C}_2$. Let $\mathfrak{e}'^{(1)}, \mathfrak{e}''^{(1)}$ be free legs (or fixed legs) and $\mathfrak{e}'^{(2)}, \mathfrak{e}''^{(2)}$ be fixed legs (or a free legs) if \mathcal{C}_1 contains a free leg (or a fixed leg). Let $\mathfrak{e}'''^{(1)}$ be a free leg (or a fixed leg) and $\mathfrak{e}'''^{(2)}$ be a fixed leg (or a free leg) if \mathcal{C}_2 contains a free leg (or a fixed leg). Cutting in this way, all couples contain both fixed and free legs.

As in case 3.1, we have

$$\begin{aligned}
\sup_{\{c_l\}_l} \#Eq(\mathcal{C}, \{c_l\}_l) &\leq \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_{n_*} \cup \mathcal{C}_2)}} \#Eq(\mathcal{C}_{n_*} \cup \mathcal{C}_2, \{c_{l_1}\}) \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}_1)}} \#Eq(\mathcal{C}_1, \{c_{l_2}\}) \\
&\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_1)} \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_{n_*} \cup \mathcal{C}_2)}} \#Eq(\mathcal{C}_{n_*} \cup \mathcal{C}_2, \{c_{l_1}\}) \\
&\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_1)} \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_{n_*})}} \#Eq(\mathcal{C}_{n_*}, \{c_{l_1}\}) \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}_2)}} \#Eq(\mathcal{C}_2, \{c_{l_2}\}) \\
&\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_1) + \chi(\mathcal{C}_2) + \chi(\mathcal{C}_{n_*})} = L^{O(\theta)} Q^{\chi(\mathcal{C})}.
\end{aligned} \tag{3.148}$$

Case 3.2.2. Assume that \mathcal{C}_1 contains no legs. Then \mathcal{C}_2 must contain free legs since \mathcal{C} contains free legs.

We apply the same argument as case 2 and case 3.1. Cut \mathfrak{e}''' into $\mathfrak{e}'''^{(1)} \in \mathcal{C}_{n_*} \cup \mathcal{C}_1$ and $\mathfrak{e}'''^{(2)} \in \mathcal{C}_2$. Let $\mathfrak{e}'''^{(1)}$ be a free leg and $\mathfrak{e}'''^{(2)}$ be a fixed leg. Cutting in this way, all couples contain both fixed and free legs.

As in case 3.1, we have

$$\begin{aligned}
\sup_{\{c_l\}_l} \#Eq(\mathcal{C}, \{c_l\}_l) &\leq \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_{n_*} \cup \mathcal{C}_1)}} \#Eq(\mathcal{C}_{n_*} \cup \mathcal{C}_1, \{c_{l_1}\}) \sup_{\{c_{l_2}\}_{l_2 \in \text{leg}(\mathcal{C}_2)}} \#Eq(\mathcal{C}_2, \{c_{l_2}\}) \\
&\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_2)} \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_{n_*} \cup \mathcal{C}_1)}} \#Eq(\mathcal{C}_{n_*} \cup \mathcal{C}_1, \{c_{l_1}\}).
\end{aligned} \tag{3.149}$$

$\mathcal{C}_{n_*} \cup \mathcal{C}_1$ has only one free leg $\mathfrak{e}'''^{(1)}$ and a fixed leg l and both of these two edges

are connected to \mathbf{n}_* . Therefore, $\mathcal{C}_{\mathbf{n}_*} \cup \mathcal{C}_1$ satisfies exactly the condition of case 1.2. Applying the conclusion there gives

$$\sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_{\mathbf{n}_*} \cup \mathcal{C}_1)}} \#Eq(\mathcal{C}_{\mathbf{n}_*} \cup \mathcal{C}_1, \{c_{l_1}\}) \lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_{\mathbf{n}_*} \cup \mathcal{C}_1)}. \quad (3.150)$$

Combine (3.149) and (3.150). We get

$$\begin{aligned} \sup_{\{c_l\}_l} \#Eq(\mathcal{C}, \{c_l\}_l) &\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_2)} \sup_{\{c_{l_1}\}_{l_1 \in \text{leg}(\mathcal{C}_{\mathbf{n}_*} \cup \mathcal{C}_1)}} \#Eq(\mathcal{C}_{\mathbf{n}_*} \cup \mathcal{C}_1, \{c_{l_1}\}) \\ &\lesssim L^{O(\theta)} Q^{\chi(\mathcal{C}_2) + \chi(\mathcal{C}_{\mathbf{n}_*} \cup \mathcal{C}_1)} = L^{O(\theta)} Q^{\chi(\mathcal{C})}. \end{aligned} \quad (3.151)$$

Therefore, we complete the proof of Proposition 3.3.24 and thus the proof of Proposition 3.3.23 and 3.3.22. □

3.3.4 A heuristic explanation of the derivation of the wave kinetic equation

In this section, we explain heuristically explain the proof of Proposition 3.3.27, which is an essential part of the derivation of the wave kinetic equation. The full derivation of the wave kinetic equation also needs an estimate on the linearized operator and an estimate that removes the ϵ -loss, which will not be discussed in this thesis.

Estimates of tree terms

By (3.34), in order to control the approximation error Err , it suffices to get upper bounds of tree terms \mathcal{J}_T .

Let us introduce a definition before stating the proposition.

Definition 3.3.26. Given a property A , we say A happens L -certainly if the proba-

bility that A happens satisfies $P(A) \geq 1 - Ke^{-L^\theta}$ for some $K, \theta > 0$.

Proposition 3.3.27. *If $\alpha \geq \max(L^{-\frac{d}{6}}, L^{-1})$, ($\alpha \geq \max(L^{-\frac{d}{6}}, L^{-2}$) in case the ratio of periods is generic in the sense of Theorem 4.3.2). Then we have L -certainly that*

$$\sup_t \sup_k |(\mathcal{J}_T)_k| \lesssim L^{O(l(T)\theta)} \rho^{l(T)}. \quad (3.152)$$

Here $(\mathcal{J}_T)_k$ is the Fourier coefficients of \mathcal{J}_T and

$$\rho = \alpha^2 t. \quad (3.153)$$

By (3.78), we know that

$$\begin{aligned} \mathbb{E}|\mathcal{J}_{T,k}|^2 &= \left(\frac{\lambda^2}{L^{2d}} \right)^{2l(T)} \sum_{p \in \mathcal{P}_F(\{k_1, \dots, k_{2l(T)+1}, k'_1, \dots, k'_{2l(T)+1}\})} Term(T, p) \\ &= \frac{\alpha^{2l(T)}}{L^{2dl(T)}} \sum_{p \in \mathcal{P}_F(\{k_1, \dots, k_{2l(T)+1}, k'_1, \dots, k'_{2l(T)+1}\})} Term(T, p). \end{aligned} \quad (3.154)$$

Therefore, to show (3.152), it suffices to show that

$$\frac{\alpha^{2l(T)}}{L^{2dl(T)}} Term(T, p) \leq L^{O(l(T)\theta)} \rho^{l(T)}. \quad (3.155)$$

Therefore, we just need to consider the order of magnitude of $Term(T, p)$.

The order of magnitude of $Term(T, p)$

In Proposition 3.3.22 (3), we claim the following inequality

$$\frac{\alpha^{2l(T)}}{\rho^{l(T)} L^{2dl(T)}} Term(T, p) \leq M/Q^{l+l'} = M/Q^{2l(T)} \quad (3.156)$$

where M is the number of solutions of (3.79) and $Q = L^d/\sqrt{t}$. After simplification, the above inequalities is equivalent to

$$Term(T, p) \leq t^{2l(T)} M \quad (3.157)$$

We heuristically prove the above inequalities.

By definition (3.72), we know that

$$Term(T, p) = \sum_{\substack{k_1, k_2, \dots, k_{2l(T)+1} \\ k'_1, k'_2, \dots, k'_{2l(T)+1}}} H_{k_1 \dots k_{2l(T)+1}}^T H_{k'_1 \dots k'_{2l(T)+1}}^T \delta_p(k_1, \dots, k_{2l(T)+1}, k'_1, \dots, k'_{2l(T)+1}) \sqrt{n_{\text{in}}(k_1)} \dots \quad (3.158)$$

We assume the following conclusion.

Essential gap 1. $H_{k_1 \dots k_{2l(T)+1}}^T H_{k'_1 \dots k'_{2l(T)+1}}^T$ concentrates near (3.79), and near (3.79) these coefficients are of order $t^{l(T)}$.

We know that

$$Term(T, p) \sim t^{2l(T)} \sum_{Eq(C)} \sqrt{n_{\text{in}}(k_1)} \dots \sim t^{2l(T)} M. \quad (3.159)$$

This proves (3.156).

To prove (3.155), we just need to apply the Proposition 3.3.22 (3) and to obtain $M/Q^{l+l'} \leq L^{O(\theta)}$ and then substitute into (3.156).

But remember that Proposition 3.3.22 (3) made the assumption that there is no long chain of bad circle nodes. This is the second essential gap.

Essential gap 2. Long chain of bad circle nodes may appear for some pairing $p \in \mathcal{P}_F$.

Removing the essential gaps

Now we discuss several proposals of remove the above two gaps.

In the three wave case, a similar conclusion to the first essential gap is proved in . The four wave case need a more detailed analysis of oscillatory integral because of the presence of $\tilde{\Omega}$.

The second essential gap can be solved because of a cancellation identity. The following picture is a schematic description of a long chain of bad nodes, in which all T_i , T'_i , T_i^e are trees. There is a pairing which pairs nodes within the groups $\{T_1, T'_1, T_2^e\}$, $\{T_2, T'_2, T_3^e\}$, \dots , $\{T_{k-1}, T'_{k-1}, T_k^e\}$, $\{T_0, T'_0, T_1^e, T_k, T'_k\}$. All nodes in group other than the last one $\{T_0, T'_0, T_1^e, T_k, T'_k\}$ are not circle nodes.

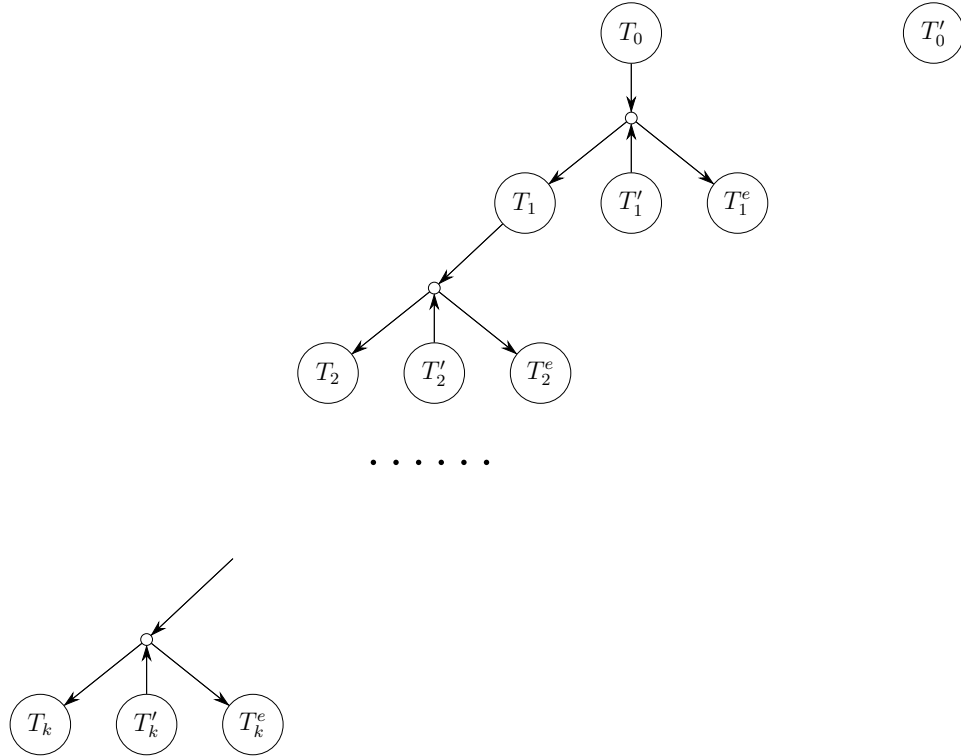


Figure 3.25: Long chain of bad circle nodes

In Proposition 3.3.22, we have the desired counting result if there is no long chain of bad nodes. For each long chain of bad nodes, we lose a factor of α^{-1} for each group $\{T_i, T'_i, T_i^e\}$ ($1 \leq i \leq k-1$). However, if the pairing within the group $\{T_i, T'_i, T_i^e\}$ is not regular (after splitting the circle node), then by the argument of Deng and Hani [9], [13], we gain a factor which can be used to compensate the loss. If the pairing is

regular, we have the following cancellation identity.

After splitting, a circle node gives a dotted edge. In a regular pairing, there are only three possibilities for the configuration of this dotted edge. As shown in the following three pictures.



Figure 3.26: The first configuration of the dotted edge and its conjugation



Figure 3.27: The second configuration of the dotted edge and its conjugation

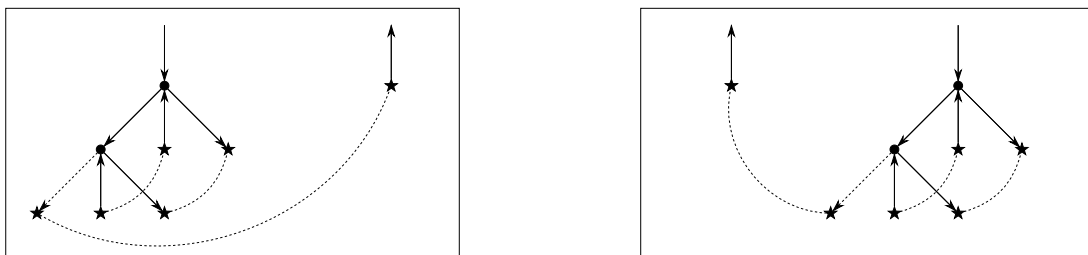


Figure 3.28: The third configuration of the dotted edge and its conjugation

We show that terms arising from conjugated configuration cancel with each other. We only consider the first configuration, because the calculation of other configurations is the same.

The H coefficients of term corresponding to the left configuration in Figure 3.26

is

$$H_{k_1 k_2 k_3 k}(t, t_3) = i \int_{\substack{0 \leq t_3 \leq t_1 \leq t \\ 0 \leq t_2 \leq t}} e^{i(t_1 - t_2)\Omega_{k_1 k_2 k_3 k}} dt_1 dt_2 \Lambda_{k_1}^{-4} \Lambda_{k_2}^{-2} \Lambda_{k_3}^{-2} \Lambda_k^{-2}. \quad (3.160)$$

The H coefficients of term corresponding to the right configuration in Figure 3.26 is

$$\bar{H}_{k_1 k_2 k_3 k}(t, t_3) = -i \int_{\substack{0 \leq t_3 \leq t_1 \leq t \\ 0 \leq t_2 \leq t}} e^{i(t_2 - t_1)\Omega_{k_1 k_2 k_3 k}} dt_1 dt_2 \Lambda_{k_1}^{-4} \Lambda_{k_2}^{-2} \Lambda_{k_3}^{-2} \Lambda_k^{-2}. \quad (3.161)$$

Sum these two terms, then we have

$$\begin{aligned} 2\text{Re } H_{k_1 k_2 k_3 k}(t, t_3) = \\ i \left(\int_{\substack{0 \leq t_3 \leq t_1 \leq t \\ 0 \leq t_2 \leq t}} e^{i(t_1 - t_2)\Omega_{k_1 k_2 k_3 k}} dt_1 dt_2 - \int_{\substack{0 \leq t_3 \leq t_1 \leq t \\ 0 \leq t_2 \leq t}} e^{i(t_2 - t_1)\Omega_{k_1 k_2 k_3 k}} dt_1 dt_2 \right) \Lambda_{k_1}^{-4} \Lambda_{k_2}^{-2} \Lambda_{k_3}^{-2} \Lambda_k^{-2}. \end{aligned} \quad (3.162)$$

Notice that

$$\int_{\substack{0 \leq t_3 \leq t_1 \leq t \\ 0 \leq t_2 \leq t}} e^{i(t_2 - t_1)\Omega_{k_1 k_2 k_3 k}} dt_1 dt_2 = \int_{\substack{0 \leq t_3 \leq t_2 \leq t \\ 0 \leq t_1 \leq t}} e^{i(t_1 - t_2)\Omega_{k_1 k_2 k_3 k}} dt_1 dt_2 \quad (3.163)$$

Then we know that

$$2\text{Re } H_{k_1 k_2 k_3 k}(t, t_3) = i \left(\int_{0 \leq t_2 \leq t_3 \leq t_1 \leq t} - \int_{0 \leq t_1 \leq t_3 \leq t_2 \leq t} \right) e^{i(t_1 - t_2)\Omega_{k_1 k_2 k_3 k}} dt_1 dt_2 \Lambda_{k_1}^{-4} \Lambda_{k_2}^{-2} \Lambda_{k_3}^{-2} \Lambda_k^{-2}. \quad (3.164)$$

When $t_3, t \rightarrow \infty$ and sum over k_1, k_2, k_3 , the two sums have the same asymptotic and thus cancel with each other.

Chapter 4

Counting results in wave kinetic theory

In this chapter, we provide several useful number theoretic results for wave turbulence theorem. Results for the three wave models can be found in section 4.1. Results for the four wave models can be found in section 3.3.3

4.1 Number theoretic results in three wave models

The mains results of this chapter is to prove Theorem 4.1.1 and Theorem 4.1.4

Theorem 4.1.1. *Let $\Lambda(k) = (\beta_x k_x^2 + \beta_2 k_2^2 + \cdots + \beta_d k_d^2)k_x$, $d \geq 3$, then for all $\beta \in [1, 2]^d$, the following number theory estimate is true*

$$\sup_{\substack{k, \sigma \in \mathbb{Z}_L^d \\ k \neq 0, T \leq L}} |k_x| T \# \left\{ \begin{array}{l} k_1, k_2 \in \mathbb{Z}_L^d : \quad k_1 + k_2 = k \\ |k_1| \lesssim 1 \quad \Lambda(k_1) + \Lambda(k_2) = \Lambda(k) + \sigma + O(T^{-1}) \end{array} \right\} \leq L^{2d}. \quad (4.1)$$

Remark 4.1.2. The restriction $T \leq L$ is not optimal. The optimal result is expected to be $T \leq L^2$ for general β and $T \leq L^d$ for generic β . It is possible to apply circle

method and probabilistic method in number theory method to prove the optimal result.

Remark 4.1.3. 4.1.1 and 4.1.4 is unlikely to be true when $d = 2$. Because in this case, the quadratic term of k_1 in (4.11) becomes $3k_x k_{1x}^2 + k_x |k_{1y}|^2 + 2k_y \cdot k_{1x} k_{1y}$ ($k_{1\perp}$ becomes $k_{1\perp} = k_y$), which is degenerate when $k_y^2 = 3k_x^2$.

Theorem 4.1.4. *Let $\Omega_k(k_1) := \Lambda(k_1) + \Lambda(k - k_1) - \Lambda(k)$ and t be a large number, then given any smooth compactly supported $F(k)$ and smooth $g(s)$ satisfying $|g|(s) + |g'(s)| \lesssim 1/(1 + s^2)$ and $\int_{\mathbb{R}} g(s) ds = c$, we have*

$$\sum_{k_1 \in \mathbb{Z}_L^d} g(t\Omega_k(k_1)) F(k_1) = cL^d t^{-1} \int F(k_1) \delta(\Omega_k(k_1)) dk_1 + O(L^{d-1}). \quad (4.2)$$

in particular

$$\sum_{k_1 \in \mathbb{Z}_L^d} g(t\Omega_k(k_1)) F(k_1) \leq 2ct^{-1} L^d D^{d-1}. \quad (4.3)$$

Proof of Theorem 4.1.1. (4.1) is a corollary of the following lemma.

Lemma 4.1.5. *For any $d \geq 4$ and β*

$$\sup_{\substack{k, \sigma \in \mathbb{Z}_L^d \\ k \neq 0}} |k_x| \# \{k_1 \in \mathbb{Z}_L^d, |k_1| \lesssim 1 : \Lambda(k_1) + \Lambda(k - k_1) - \Lambda(k) = \sigma + O(L^{-1})\} \lesssim L^{d-1}. \quad (4.4)$$

Proof. Define $\mathcal{D}_{k, \sigma} = \# \{k_1 \in \mathbb{Z}_L^d, |k_1| \lesssim 1 : \Lambda(k_1) + \Lambda(k - k_1) - \Lambda(k) = \sigma + O(L^{-1})\}$, then we just need to show that

$$\sup_{\substack{k, \sigma \in \mathbb{Z}_L^d \\ k \neq 0}} \# \mathcal{D}_{k, \sigma} \lesssim L^{d-1}. \quad (4.5)$$

We prove (4.5) using volume bound. The proof is divided into three steps.

Step 1. In this step, we show that

$$\#\mathcal{D}_{k,\sigma} \leq L^d \text{vol}(\mathcal{D}_{k,\sigma}^{\mathbb{R}}), \quad (4.6)$$

where $\mathcal{D}_{k,\sigma}^{\mathbb{R}} = \{k_1 \in \mathbb{R}^d, |k_1| \lesssim 1 : \Lambda(k_1) + \Lambda(k - k_1) - \Lambda(k) = \sigma + O(L^{-1})\}$.

(4.6) can be proved from the following claim.

Claim. If $k_1 \in \mathcal{D}_{k,\sigma}$, then $D_{1/(2L)}(k_1) \subseteq \mathcal{D}_{k,\sigma}^{\mathbb{R}}$. Here $D_r(k_1) = \{k'_1 \in \mathbb{R}^d : \sup_{i=1,\dots,d} |(k'_1)_i - (k_1)_i| \leq r\}$ ($(k_1)_i$ are the components of k_1).

We prove the claim now. $x \in \mathcal{D}_{k,\sigma}$ is equivalent to $\Lambda(k_1) + \Lambda(k - k_1) - \Lambda(k) = \sigma + O(L^{-1})$. For any $k'_1 \in D_{1/(2L)}(k_1)$, $|k'_1 - k_1| \lesssim 1/L$, because Λ is a Lifschitz function, we have $|\Lambda(k_1) - \Lambda(k'_1)| \lesssim L^{-1}$. Therefore,

$$\begin{aligned} & |\Lambda_\beta(k'_1) + \Lambda_\beta(k - k'_1) - \Lambda(k) - \sigma| \\ & \leq |\Lambda_\beta(k_1) + \Lambda_\beta(k - k_1) - \Lambda(k) - \sigma| + |\Lambda_\beta(k_1) - \Lambda_\beta(k'_1)| + |\Lambda_\beta(k - k_1) - \Lambda_\beta(k - k'_1)| \\ & \lesssim L^{-1}. \end{aligned} \quad (4.7)$$

Therefore, we have $\Lambda(k'_1) + \Lambda(k - k'_1) - \Lambda(k) = \sigma + O(L^{-1})$ and thus $k'_1 \in \mathcal{D}_{k,\sigma}^{\mathbb{R}}$. This is true for any $k'_1 \in D_{1/(2L)}(k_1)$, so $D_{1/(2L)}(k_1) \subseteq \mathcal{D}_{k,\sigma}^{\mathbb{R}}$.

Now we prove (4.6). Since for different $k_1, k'_1 \in \mathcal{D}_{k,\sigma}$, $D_{1/(2L)}(k_1) \cap D_{1/(2L)}(k'_1) = \emptyset$, we have

$$\sum_{k_1 \in \mathcal{D}_{k,\sigma}} \text{vol}(D_{1/(2L)}(k_1)) = \text{vol}\left(\bigcup_{k_1 \in \mathcal{D}_{k,\sigma}} D_{1/(2L)}(k_1)\right) \leq \text{vol}(\mathcal{D}_{k,\sigma}^{\mathbb{R}}). \quad (4.8)$$

The left hand side equals to $L^{-d} \#\mathcal{D}_{k,\sigma}$, so we get

$$L^{-d} \#\mathcal{D}_{k,\sigma} \leq \text{vol}(\mathcal{D}_{k,\sigma}^{\mathbb{R}}), \quad (4.9)$$

which implies (4.6).

Step 2. In this step, we show that

$$\text{vol}(\mathcal{D}_{k,\sigma}^{\mathbb{R}}) \leq L^{-1}|k_x|^{-1}. \quad (4.10)$$

Combining (4.6) and (4.10), we get (4.5), which proves Lemma 4.1.5.

Given a vector k , we denote the first component of k by k_x and the vector formed by other components by k_{\perp} . Then $k = (k_x, k_{\perp})$, $k_1 = (k_{1x}, k_{1\perp})$ and a simple calculation suggests that

$$\begin{aligned} & \Lambda(k_1) + \Lambda(k - k_1) - \Lambda(k) \\ &= 3k_x k_{1x}^2 + k_x |k_{1\perp}|^2 + 2k_{\perp} \cdot k_{1\perp} k_{1x} - (3k_x^2 + |k_{\perp}|^2)k_{1x} - 2k_x k_{\perp} \cdot k_{1\perp} \end{aligned} \quad (4.11)$$

If we fix k_{1x} to be a constant c in $\Lambda(k_1) + \Lambda(k - k_1) - \Lambda(k)$ and denote the resulting function by $F_{k,\sigma,c}$, then

$$F_{k,\sigma,c}(k_{1\perp}) = k_x |k_{1\perp}|^2 + 2(c - k_x)k_{\perp} \cdot k_{1\perp} + 3k_x c^2 - (3k_x^2 + |k_{\perp}|^2)c. \quad (4.12)$$

Therefore, if we define $\mathcal{D}_{k,\sigma}^{\mathbb{R}}(k_{1x} = c)$ by

$$\mathcal{D}_{k,\sigma}^{\mathbb{R}}(k_{1x} = c) = \{k_{1\perp} \in \mathbb{R}^{d-1}, |k_{1\perp}| \lesssim 1 : F_{k,\sigma,c}(k_{1\perp}) = \sigma + O(L^{-1})\} \quad (4.13)$$

Then

$$\mathcal{D}_{k,\sigma}^{\mathbb{R}} = \cup_{|c| \lesssim 1} \mathcal{D}_{k,\sigma}^{\mathbb{R}}(k_{1x} = c) \quad (4.14)$$

By Fubini theorem (or coarea formula), we get

$$\text{vol}(\mathcal{D}_{k,\sigma}^{\mathbb{R}}) = \int_{|c| \lesssim 1} \text{vol}(\mathcal{D}_{k,\sigma}^{\mathbb{R}}(k_{1x} = c)) dc. \quad (4.15)$$

To prove (4.10), it suffices to find an upper of $\text{vol}(\mathcal{D}_{k,\sigma}^{\mathbb{R}}(k_{1x} = c))$. Since $F_{k,\sigma,c}(k_{1\perp})$

is a quadratic function in $k_{1\perp}$ whose degree 2 term is $k_x|k_{1\perp}|^2$, a translation $k_{1\perp} \rightarrow k_{1\perp} - (c - k_x)k_{1\perp}/k_x$ transforms $F_{k,\sigma,c}(k_{1\perp}) = \sigma + O(L^{-1})$ to $k_x|k_{1\perp}|^2 = C_{k,\sigma,c} + O(L^{-1})$.

Therefore,

$$\begin{aligned} \text{vol}(\mathcal{D}_{k,\sigma}^{\mathbb{R}}(k_{1x} = c)) &\leq \text{vol}(\{k_{1\perp} \in \mathbb{R}^{d-1} : k_x|k_{1\perp}|^2 = C_{k,\sigma,c} + O(L^{-1})\}) \\ &= \text{vol}(\{k_{1\perp} \in \mathbb{R}^{d-1} : |k_{1\perp}|^2 = C_{k,\sigma,c}/k_x + O(L^{-1}|k_x|^{-1})\}) \quad (4.16) \\ &\lesssim L^{-1}|k_x|^{-1} \end{aligned}$$

Here the last inequality follows from the elementary fact that $\text{vol}(\{x \in \mathbb{R}^n : |x|^2 = R^2 + O(\eta)\}) \lesssim \eta$ when $n \geq 2$. Notice here the dimension of $k_{1\perp}$ is $d - 1$ which is greater than 2, so this fact is applicable.

Combining (4.15) and (4.15), we proves (4.10). Therefore, we complete the proof of Lemma 4.1.5 \square

We now return to the proof of Theorem 4.1.1. Define

$$\mathcal{D}_{k,\sigma}([-T^{-1}, T^{-1}]) = \{k_1 \in \mathbb{Z}_L^d, |k_1| \lesssim 1 : \Lambda(k_1) + \Lambda(k - k_1) - \Lambda(k) = \sigma + O(L^{-1})\} \quad (4.17)$$

Let us prove (4.1). For any $T \leq L$, let $N = [L/T] + 1$, then $\cup_{j=-N}^N [jL^{-1}, (j+1)L^{-1}]$ is a cover of $[-T^{-1}, T^{-1}]$. Therefore, $\cup_{j=-N}^N \mathcal{D}_{k,\sigma}([jL^{-1}, (j+1)L^{-1}])$ is a cover of $\mathcal{D}_{k,\sigma}([-T^{-1}, T^{-1}])$

By Lemma 4.1.5 we get

$$\#\mathcal{D}_{k,\sigma}([-T^{-1}, T^{-1}]) \lesssim \sum_{j=-N}^N \#\mathcal{D}_{k,\sigma}([jL^{-1}, (j+1)L^{-1}]) \lesssim NL^{d-1}|k_x|^{-1} \lesssim L^d T^{-1}|k_x|^{-1}. \quad (4.18)$$

This proves (4.1). \square

Proof of Theorem 4.1.4. The inequality in Theorem 4.1.4 follows from the equality because $\int F(k_1) \delta(\Omega_k(k_1)) dk_1 \leq D^{d-1}$, $t \leq L^{-\theta} \alpha^{-2} \leq L^{1-\theta}$ and $D^{d-1} \lesssim L^\theta$ if L is

large enough.

Let $k_1 = \frac{K_1}{L}$ and $k = \frac{K}{L}$. Apply the high dimensional Euler-Maclaurin formula, we know that

$$\begin{aligned} \sum_{k_1 \in \mathbb{Z}_L^d} g(t\Omega_k(k_1))F(k_1) &= \int g(t\Omega_k(K_1/L))F\left(\frac{K_1}{L}\right) dK_1 \\ &+ \sum_{|J|_\infty=1} \int_{\mathbb{R}^d} \{K_1\}^J \partial_{K_1}^J \left(g(t\Omega_k(K_1/L))F\left(\frac{K_1}{L}\right) \right) dK_1 \end{aligned} \quad (4.19)$$

We have the following estimates

$$\int g(t\Omega_k(K_1/L))F\left(\frac{K_1}{L}\right) dK_1 = L^d \int g(t\Omega_k(k_1))F(k_1) dk_1 \quad (4.20)$$

and

$$\begin{aligned} &\int_{\mathbb{R}^d} \{K_1\}^J \partial_{K_1}^J \left(g(t\Omega_k(K_1/L))F\left(\frac{K_1}{L}\right) \right) dK_1 \\ &= L^{d-|J|} \int_{\mathbb{R}^d} \{Lk_1\}^J \partial_{k_1}^J (g(t\Omega_k(k_1))F(k_1)) dk_1 \\ &= O\left(L^{d-|J|} \int_{\mathbb{R}^d} |\partial_{k_1}^J (g(t\Omega_k(k_1))F(k_1))| dk_1 \right) \\ &= O\left(\sum_{j=1}^{|J|} L^{d-j} t^j \int_{\mathbb{R}^d} g^{(j)}(t\Omega_k(k_1))F^{(j)}(k_1) dk_1 \right) \end{aligned} \quad (4.21)$$

Here in the second equality we apply the fact that $\{Lk_1\} \leq 1$. In the last line we define $g^{(j)} = \sum_{j' \leq j} \left| \frac{d^{j'}}{d^{j'} s} g(s) \right|$ and $F^{(j)}(k_1) = \sum_{|J'| \leq j} |\partial_{k_1}^{J'} F(k_1)|$ (Here J' is a multi-index) and we also use the fact that $|\partial_{k_1}^j (g(t\Omega_k(k_1)))| \leq t^j g^{(j)}(t\Omega_k(k_1))$.

Now we just need to show that

$$\int g(t\Omega_k(k_1))F(k_1) dk_1 = ct^{-1} \int F(k_1) \delta(\Omega_k(k_1)) dk_1 + O(t^{-2}) \quad (4.22)$$

and

$$\int_{\mathbb{R}^d} g^{(j)}(t\Omega_k(k_1)) F^{(j)}(k_1) dk_1 \leq t^{-1} \quad (4.23)$$

In fact, if we substitute (4.22) into (4.20), we get the main term in (4.2). If we substitute (4.23) into (4.21), we know that the error terms come from (4.21) can be bounded by $\sum_{j=1}^{|J|} L^{d-j} t^{j-1} \leq |J| L^{d-1}$. Therefore, (4.22) and (4.23) implies the lemma.

By coarea formula we know that

$$\int g(t\Omega_k(k_1)) F(k_1) dk_1 = \int g(t\omega) \underbrace{\left(\int F(k_1) \delta(\Omega_k(k_1) - \omega) dk_1 \right)}_{h(\omega)} d\omega \quad (4.24)$$

Notice that $h(\omega)$ is differentiable, then we have

$$\begin{aligned} \int g(t\omega) h(\omega) d\omega &= \int g(t\omega) (h(\omega) - h(0)) d\omega + h(0) \int g(t\omega) d\omega \\ &= ct^{-1} h(0) + O(t^{-2}). \end{aligned} \quad (4.25)$$

Therefore,

$$\begin{aligned} \int g(t\Omega_k(k_1)) F(k_1) dk_1 &= \int g(t\omega) h(\omega) d\omega \\ &= ct^{-1} h(0) + O(t^{-2}) = ct^{-1} \left(\int F(k_1) \delta(\Omega_k(k_1)) dk_1 \right) + O(t^{-2}) \end{aligned} \quad (4.26)$$

Using the same argument we can also prove (4.23), so we complete the proof of the Theorem 4.1.4. □

4.2 High Dimensional Euler-Maclaurin Formula

Theorem 4.2.1. *Assume that $J = (j_1, \dots, j_d)$ is a multi-index. Given a vector $K = (K^{(1)}, \dots, K^{(d)})$, define $K^J = (K^{(1)})^{j_1} \dots (K^{(d)})^{j_d}$. Given a number a , $\{a\} = a - [a]$ is*

its fractional part, $\{K\} := (\{K^{(1)}\}, \dots, \{K^{(d)}\})$. We also define $|J|_\infty = \sup_{1 \leq n \leq d} j_n$, $|J| = \sum_{1 \leq n \leq d} j_n$. Then we have

$$\sum_{K \in \mathbb{Z}^d} f(K) = \int_{\mathbb{R}^d} f(K) dK + \sum_{|J|_\infty=1} \int_{\mathbb{R}^d} \{K\}^J \partial_K^J f(K) dK \quad (4.27)$$

Proof. This can be proved by induction.

When $d = 1$, (4.27) becomes

$$\sum_{K \in \mathbb{Z}} f(K) = \int_{\mathbb{R}} f(K) dK + \int_{\mathbb{R}} \{K\} \partial_K f(K) dK \quad (4.28)$$

This is the standard Euler-Maclaurin formula which can be proved by integration by parts in the second integral of the right hand side.

If the formula is true in dimension d , we prove it for dimension $d+1$. Assume that $K = (\tilde{K}, K^{(d+1)})$ is a $d+1$ dimensional vector, then apply the induction assumption

$$\sum_{\tilde{K} \in \mathbb{Z}^d} f(\tilde{K}, K^{(d+1)}) = \int_{\mathbb{R}^d} f(\tilde{K}, K^{(d+1)}) d\tilde{K} + \sum_{|\tilde{J}|_\infty=1} \int_{\mathbb{R}^d} \{\tilde{K}\}^{\tilde{J}} \partial_{\tilde{K}}^{\tilde{J}} f(\tilde{K}, K^{(d+1)}) d\tilde{K}. \quad (4.29)$$

Sum over $K^{(d+1)}$ and apply $d = 1$ Euler-Maclaurin formula we get

$$\begin{aligned}
& \sum_{K^{(d+1)} \in \mathbb{Z}} \sum_{\tilde{K} \in \mathbb{Z}^d} f(\tilde{K}, K^{(d+1)}) \\
&= \int_{\mathbb{R}^d} \sum_{K^{(d+1)} \in \mathbb{Z}} f(\tilde{K}, K^{(d+1)}) d\tilde{K} + \sum_{|\tilde{J}|_\infty=1} \int_{\mathbb{R}^d} \{\tilde{K}\}^{\tilde{J}} \partial_{\tilde{K}}^{\tilde{J}} \sum_{K^{(d+1)} \in \mathbb{Z}} f(\tilde{K}, K^{(d+1)}) d\tilde{K}. \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}} f(\tilde{K}, K^{(d+1)}) dK^{(d+1)} d\tilde{K} + \int_{\mathbb{R}^d} \int_{\mathbb{R}} \{K^{(d+1)}\} \partial_{K^{(d+1)}} f(\tilde{K}, K^{(d+1)}) dK^{(d+1)} d\tilde{K} \\
&+ \sum_{|\tilde{J}|_\infty=1} \int_{\mathbb{R}^d} \int_{\mathbb{R}} \{\tilde{K}\}^{\tilde{J}} \partial_{\tilde{K}}^{\tilde{J}} f(\tilde{K}, K^{(d+1)}) dK^{(d+1)} d\tilde{K} \\
&+ \sum_{|\tilde{J}|_\infty=1} \int_{\mathbb{R}^d} \int_{\mathbb{R}} \{K^{(d+1)}\} \partial_{K^{(d+1)}} \{\tilde{K}\}^{\tilde{J}} \partial_{\tilde{K}}^{\tilde{J}} f(\tilde{K}, K^{(d+1)}) dK^{(d+1)} d\tilde{K}. \\
&= \int_{\mathbb{R}^{d+1}} f(K) dK + \sum_{|J|_\infty=1} \int_{\mathbb{R}^d} \{K\}^J \partial_K^J f(K) dK
\end{aligned} \tag{4.30}$$

In the last step, $J = (\tilde{J}, j^{(d+1)})$. In the second equality, the second term corresponds to $\tilde{J} = 0$ and $j^{(d+1)} = 1$, the third term corresponds to $\tilde{J} \neq 0$ and $j^{(d+1)} = 0$ and the fourth term corresponds to $\tilde{J} \neq 0$ and $j^{(d+1)} = 1$. \square

4.3 Number theoretic results in four wave models

The mains results of this chapter is to prove Theorem 4.3.1 and Theorem 4.3.2

Theorem 4.3.1. *Let $\Lambda(k) = \sqrt{1 + |k|^2}$, then for all $\beta \in [1, 2]^d$, the following number theory estimate is true*

$$\sup_{\substack{c, n \in \mathbb{Z}_L^d \\ T \leq L}} T \# \left\{ \begin{array}{l} k_1, k_2, k_3 \in \mathbb{Z}_L^d \\ |k_1|, |k_2|, |k_3| \lesssim 1 \end{array} : \begin{array}{l} k_1 - k_2 + k_3 = c \\ \Lambda(k_1) - \Lambda(k_2) + \Lambda(k_3) = \Lambda(c) + n + O(T^{-1}) \end{array} \right\} \leq L^{2d}. \tag{4.31}$$

$$\sup_{\substack{c_1, c_2, n \in \mathbb{Z}_L^d \\ T \leq L}} T \# \left\{ \begin{array}{l} k_1, k_2 \in \mathbb{Z}_L^d : k_1 + k_2 = c_1 + c_2 \\ |k_1|, |k_2| \lesssim 1 \quad \Lambda(k_1) + \Lambda(k_2) = \Lambda(c_1) + \Lambda(c_2) + n + O(T^{-1}) \end{array} \right\} \leq L^{2d}. \quad (4.32)$$

$$\sup_{\substack{c_1, c_2, n \in \mathbb{Z}_L^d \\ T \leq L^2}} T \# \left\{ \begin{array}{l} k_1, k_2 \in \mathbb{Z}_L^d : k_1 - k_2 = c_1 - c_2, \quad k_1 \neq k_2 \\ |k_1|, |k_2| \lesssim 1 \quad \Lambda(k_1) - \Lambda(k_2) = \Lambda(c_1) - \Lambda(c_2) + n + O(T^{-1}) \end{array} \right\} \leq L^{2d+1}. \quad (4.33)$$

More generally, the proof of this theorem works for many other choices of $\Lambda(k)$.

Theorem 4.3.2. Let $\Lambda(k) = \sqrt{1 + |k|^2}$, then for sufficiently generic $\beta \in [1, 2]^d$, the following number theory estimate is true for all $\theta \ll 1$

$$\sup_{\substack{c, n \in \mathbb{Z}_L^d \\ T \leq L^2}} T \# \left\{ \begin{array}{l} k_1, k_2, k_3 \in \mathbb{Z}_L^d : k_1 - k_2 + k_3 = c \\ |k_1|, |k_2|, |k_3| \leq L^\theta \quad \Lambda(k_1) - \Lambda(k_2) + \Lambda(k_3) = \Lambda(c) + n + O(T^{-1}) \end{array} \right\} \leq L^{2d+O(\theta)} \quad (4.34)$$

$$\sup_{\substack{c_1, c_2, n \in \mathbb{Z}_L^d \\ T \leq L^2}} T \# \left\{ \begin{array}{l} k_1, k_2 \in \mathbb{Z}_L^d : k_1 \pm k_2 = c_1 \pm c_2, \quad k_1 \neq k_2 \\ |k_1|, |k_2| \leq L^\theta \quad \Lambda(k_1) \pm \Lambda(k_2) = \Lambda(c_1) \pm \Lambda(c_2) + n + O(T^{-1}) \end{array} \right\} \leq L^{2d+O(\theta)} \quad (4.35)$$

More generally, the proof of this theorem works for many other choices of $\Lambda(k)$.

Proof of Theorem 4.3.1. (4.31) and (4.32) are corollaries of the following lemma.

Lemma 4.3.3. For any β

$$\sup_{\substack{m, n \\ m \lesssim 1}} \# \left\{ \begin{array}{l} x, y \in \mathbb{Z}_L^d : x + y = m \\ |x|, |y| \lesssim 1 \quad \Lambda_\beta(x) + \Lambda_\beta(y) = n + O(L^{-1}) \end{array} \right\} \lesssim_\beta L^{d-2}. \quad (4.36)$$

Proof. It is easy to show that

$$\begin{aligned} & \# \left\{ \begin{array}{l} x, y \in \mathbb{Z}_L^d \\ |x|, |y| \lesssim 1 \end{array} : \begin{array}{l} x + y = m \\ \Lambda_\beta(x) + \Lambda_\beta(y) = n + O(L^{-1}) \end{array} \right\} \\ &= \# \{x \in \mathbb{Z}_L^d, |x| \lesssim 1 : \Lambda_\beta(x) + \Lambda_\beta(m - x) = n + O(L^{-1})\} \end{aligned} \quad (4.37)$$

Define $\mathcal{D}_{m,n} = \{x \in \mathbb{Z}_L^d, |x| \lesssim 1 : \Lambda_\beta(x) + \Lambda_\beta(m - x) = n + O(L^{-1})\}$, then we just need to show that

$$\sup_{\substack{m,n \\ m \lesssim 1}} \# \mathcal{D}_{m,n} \lesssim_\beta L^{d-1}. \quad (4.38)$$

We prove (4.38) using volume bound. The proof is divided into two steps.

Step 1. In this step, we show that

$$\# \mathcal{D}_{m,n} \leq L^d \text{vol}(\mathcal{D}_{m,n}^\mathbb{R}), \quad (4.39)$$

where $\mathcal{D}_{m,n}^\mathbb{R} = \{x \in \mathbb{R}^d, |x| \lesssim 1 : \Lambda_\beta(x) + \Lambda_\beta(m - x) = n + O(L^{-1})\}$.

(4.39) can be proved from the following claim.

Claim. If $x \in \mathcal{D}_{m,n}$, then $D_{1/(2L)}(x) \subseteq \mathcal{D}_{m,n}^\mathbb{R}$. Here $D_r(x) = \{x' \in \mathbb{R}^d : \sup_{i=1,\dots,d} |x'_i - x_i| \leq r\}$ (x_i are the components of x).

We prove the claim now. $x \in \mathcal{D}_{m,n}$ is equivalent to $\Lambda_\beta(x) + \Lambda_\beta(m - x) = n + O(L^{-1})$. For any $x' \in D_{1/(2L)}(x)$, $|x' - x| \lesssim 1/L$, because Λ is a Lifschitz function, we have $|\Lambda_\beta(x) - \Lambda_\beta(x')| \lesssim L^{-1}$. Therefore,

$$\begin{aligned} & |\Lambda_\beta(x') + \Lambda_\beta(m - x') - n| \\ & \leq |\Lambda_\beta(x) + \Lambda_\beta(m - x) - n| + |\Lambda_\beta(x) - \Lambda_\beta(x')| + |\Lambda_\beta(m - x) - \Lambda_\beta(m - x')| \\ & \lesssim L^{-1}. \end{aligned} \quad (4.40)$$

Therefore, we have $\Lambda_\beta(x') + \Lambda_\beta(m - x') - n = O(L^{-1})$ and thus $x' \in \mathcal{D}_{m,n}^\mathbb{R}$. This is true for any $x' \in D_{1/(2L)}(x)$, so $D_{1/(2L)}(x) \subseteq \mathcal{D}_{m,n}^\mathbb{R}$.

Since for different $x_1, x_2 \in \mathcal{D}_{m,n}$, $D_{1/(2L)}(x_1) \cap D_{1/(2L)}(x_2) = \emptyset$, we have

$$\sum_{x \in \mathcal{D}_{m,n}} \text{vol}(D_{1/(2L)}(x)) = \text{vol}\left(\bigcup_{x \in \mathcal{D}_{m,n}} D_{1/(2L)}(x)\right) \leq \text{vol}(\mathcal{D}_{m,n}^{\mathbb{R}}). \quad (4.41)$$

The left hand side equals to $L^{-d} \# \mathcal{D}_{m,n}$, so we get

$$L^{-d} \# \mathcal{D}_{m,n} \leq \text{vol}(\mathcal{D}_{m,n}^{\mathbb{R}}), \quad (4.42)$$

which implies (4.39).

Step 2. In this step, we show that

$$\text{vol}(\mathcal{D}_{m,n}^{\mathbb{R}}) \leq L^{-1}. \quad (4.43)$$

Combining (4.39) and (4.43), we get (4.38), which proves Lemma 4.3.3.

By calculating the Hessian we can show that $F_{\beta,m,n} = \Lambda_{\beta}(x) + \Lambda_{\beta}(m - x) - n$ is a convex function. Let $x_{\beta,m}$ be its unique critical point. By Morse lemma, we know that there exists r_0 such that $F_{\beta,m,n} = |y|^2 - a$ after a change of variable.

In $\mathcal{D}_{m,n}^{\mathbb{R}} \setminus B_{r_0}(x_{\beta,m})$, $|\nabla F_{\beta,m,n}| \gtrsim 1$. Then by the coarea formula, we get

$$\begin{aligned} \text{vol}(\mathcal{D}_{m,n}^{\mathbb{R}} \setminus B_{r_0}(x_{\beta,m})) &= \text{vol}(\{|F_{\beta,m,n}| \lesssim L^{-1}\} \setminus B_{r_0}(x_{\beta,m})) \\ &= \int_{-CL^{-1}}^{CL^{-1}} \int_{\{F_{\beta,m,n}=s\} \setminus B_{r_0}(x_{\beta,m})} \frac{1}{|\nabla F_{\beta,m,n}|} d\sigma ds \\ &\lesssim \int_{-CL^{-1}}^{CL^{-1}} \text{Area}(\{F_{\beta,m,n}=s\} \setminus B_{r_0}(x_{\beta,m})) ds \\ &\lesssim L^{-1}. \end{aligned} \quad (4.44)$$

In $\mathcal{D}_{m,n}^{\mathbb{R}} \cap B_{r_0}(x_{\beta,m})$, $F_{\beta,m,n} = |y|^2 - a$ after a change of variable $y = y(x)$ and

$$y(B_{r_0}(x_{\beta,m})) \subseteq B_{Cr_0}(0)$$

$$\text{vol}(\mathcal{D}_{m,n}^{\mathbb{R}} \cap B_{r_0}(x_{\beta,m})) \lesssim \text{vol}(\{|y|^2 - a| \lesssim L^{-1}\} \cap B_{Cr_0}(0)) \lesssim L^{-1}. \quad (4.45)$$

Here the proof of the second inequality is elementary and is thus skipped.

Therefore, we complete the proof of Lemma 4.3.3 \square

We now return to the proof of Theorem 4.3.1. Define

$$\mathcal{D}_{m,n}^{(1)}([-T^{-1}, T^{-1}]) = \left\{ \begin{array}{l} k_1, k_2, k_3 \in \mathbb{Z}_L^d : k_1 - k_2 + k_3 = c \\ |k_1|, |k_2|, |k_3| \lesssim 1 \quad \Lambda(k_1) - \Lambda(k_2) + \Lambda(k_3) = \Lambda(c) + n + O(T^{-1}) \end{array} \right\} \quad (4.46)$$

$$\mathcal{D}_{m,n}^{(2)}([-T^{-1}, T^{-1}]) = \left\{ \begin{array}{l} k_1, k_2 \in \mathbb{Z}_L^d : k_1 + k_2 = c_1 + c_2 \\ |k_1|, |k_2| \lesssim 1 \quad \Lambda(k_1) + \Lambda(k_2) = \Lambda(c_1) + \Lambda(c_2) + n + O(T^{-1}) \end{array} \right\} \quad (4.47)$$

Let us first prove (4.32). For any $T \leq L$, let $N = [L/T] + 1$, then $\cup_{j=-N}^N [jL^{-1}, (j+1)L^{-1}]$ is a cover of $[-T^{-1}, T^{-1}]$. Therefore, $\cup_{j=-N}^N \mathcal{D}_{m,n}^{(2)}([jL^{-1}, (j+1)L^{-1}])$ is a cover of $\mathcal{D}_{m,n}^{(2)}([-T^{-1}, T^{-1}])$

By Lemma 4.3.3 we get

$$\#\mathcal{D}_{m,n}^{(2)}([-T^{-1}, T^{-1}]) \lesssim \sum_{j=-N}^N \#\mathcal{D}_{m,n}^{(2)}([jL^{-1}, (j+1)L^{-1}]) \lesssim NL^{d-1} \lesssim L^d T^{-1}. \quad (4.48)$$

This proves (4.32).

Now we prove (4.31).

$$\begin{aligned}
\#\mathcal{D}_{m,n}^{(1)}([-T^{-1}, T^{-1}]) &= \# \left\{ \begin{array}{l} k_1, k_2, k_3 \in \mathbb{Z}_L^d : k_1 - k_2 + k_3 = c \\ |k_1|, |k_2|, |k_3| \lesssim 1 \quad \Lambda(k_1) - \Lambda(k_2) + \Lambda(k_3) = \Lambda(c) + n + O(T^{-1}) \end{array} \right\} \\
&= \sum_{|c_2| \lesssim 1} \# \left\{ \begin{array}{l} k_1, k_3 \in \mathbb{Z}_L^d : k_1 + k_3 = c + c_2 \\ |k_1|, |k_3| \lesssim 1 \quad \Lambda(k_1) + \Lambda(k_3) = \Lambda(c) + \Lambda(c_2) + n + O(T^{-1}) \end{array} \right\} \\
&= \sum_{|c_2| \lesssim 1} \mathcal{D}_{m,n}^{(2)}([-T^{-1}, T^{-1}]) \lesssim \sum_{|c_2| \lesssim 1} L^d T^{-1} \\
&= L^{2d} T^{-1}.
\end{aligned} \tag{4.49}$$

This proves (4.31).

(4.33) is a simple corollary of the following lemma.

Lemma 4.3.4. *For any β*

$$\sup_{\substack{m,n \\ 0 \neq m \lesssim 1}} \# \left\{ \begin{array}{l} x, y \in \mathbb{Z}_L^d : x - y = m \neq 0 \\ |x|, |y| \lesssim 1 \quad \Lambda_\beta(x) - \Lambda_\beta(y) = n + O(L^{-2}) \end{array} \right\} \lesssim_\beta L^{d-1}. \tag{4.50}$$

Proof. The proof of this lemma is similar to that of Lemma 4.3.3.

Since $m \in \mathbb{Z}_L^d$ and $m \neq 0$, we have $|m| \geq L^{-1}$. Therefore, we have

$$\begin{aligned}
&\{x \in \mathbb{Z}_L^d, |x| \lesssim 1 : \Lambda_\beta(x) - \Lambda_\beta(x - m) - n = O(L^{-2})\} \\
&\subseteq \{x \in \mathbb{Z}_L^d, |x| \lesssim 1 : |m|^{-1}(\Lambda_\beta(x) - \Lambda_\beta(x - m) - n) = O(L^{-1})\}
\end{aligned} \tag{4.51}$$

Define $\mathcal{D}_{m,n} = \{x \in \mathbb{Z}_L^d, |x| \lesssim 1 : |m|^{-1}(\Lambda_\beta(x) - \Lambda_\beta(x - m) - n) = O(L^{-1})\}$, then we just need to show that

$$\sup_{\substack{m,n \\ m \lesssim 1}} \#\mathcal{D}_{m,n} \lesssim_\beta L^{d-1}. \tag{4.52}$$

We prove (4.52) using volume bound. The proof is divided into two steps.

Step 1. In this step, we show that

$$\#\mathcal{D}_{m,n} \leq L^d \text{vol}(\mathcal{D}_{m,n}^{\mathbb{R}}), \quad (4.53)$$

where $\{x \in \mathbb{R}^d, |x| \lesssim 1 : |m|^{-1}(\Lambda_\beta(x) - \Lambda_\beta(x - m) - n) = O(L^{-1})\}$.

Since the Lifschitz norm of the function $F_{\beta,m,n}(x) = |m|^{-1}(\Lambda_\beta(x) - \Lambda_\beta(x - m) - n)$ is bounded by a universal constant, the same argument of step 1 in the proof of Lemma 4.3.3 works and we do not repeat it.

Step 2. In this step, we show that

$$\text{vol}(\mathcal{D}_{m,n}^{\mathbb{R}}) \leq L^{-1}. \quad (4.54)$$

Combining (4.53) and (4.54), we get (4.52), which proves Lemma 4.3.4.

By calculating the Hessian we can show that $\Lambda_\beta(x)$ is a convex function, so

$$|m|^{-1} |\nabla \Lambda_\beta(x) - \nabla \Lambda_\beta(x - m)| \gtrsim 1. \quad (4.55)$$

Above equation implies that $|\nabla F_{\beta,m,n}| \gtrsim 1$ By the coarea formula, we get

$$\begin{aligned} \text{vol}(\mathcal{D}_{m,n}^{\mathbb{R}}) &= \text{vol}(\{|F_{\beta,m,n}| \lesssim L^{-1}\}) \\ &= \int_{-CL^{-1}}^{CL^{-1}} \int_{\{F_{\beta,m,n}=s\}} \frac{1}{|\nabla F_{\beta,m,n}|} d\sigma ds \\ &\lesssim \int_{-CL^{-1}}^{CL^{-1}} \text{Area}(\{F_{\beta,m,n} = s\}) ds \\ &\lesssim L^{-1}. \end{aligned} \quad (4.56)$$

Therefore, we complete the proof of Lemma 4.3.4

□

Finally we prove (4.33). Define

$$\mathcal{D}_{m,n}^{(3)}([-T^{-1}, T^{-1}]) = \left\{ \begin{array}{l} k_1, k_2 \in \mathbb{Z}_L^d : \quad k_1 - k_2 = c_1 - c_2, \quad k_1 \neq k_2 \\ |k_1|, |k_2| \lesssim 1 \quad \Lambda(k_1) - \Lambda(k_2) = \Lambda(c_1) - \Lambda(c_2) + n + O(T^{-1}) \end{array} \right\} \quad (4.57)$$

Apply the same argument of getting (4.48),

$$\#\mathcal{D}_{m,n}^{(3)}([-T^{-1}, T^{-1}]) \lesssim \sum_{j=-N}^N \#\mathcal{D}_{m,n}^{(3)}([jL^{-1}, (j+1)L^{-1}]) \lesssim NL^{d-1} \lesssim L^d T^{-1}. \quad (4.58)$$

This proves (4.33) and we complete the proof of Theorem 4.3.1. \square

Proof of Theorem 4.3.2. The $-$ case of (4.35) is equivalent to (4.33).

(4.34) and the $+$ case of (4.35) are corollaries of the following lemma.

Lemma 4.3.5. *For any θ , for almost all β*

$$\sup_{\substack{m,n \\ m \lesssim 1}} \# \left\{ \begin{array}{l} x, y \in \mathbb{Z}_L^d : \quad x + y = m \\ |x|, |y| \lesssim L^\theta \quad \Lambda_\beta(x) + \Lambda_\beta(y) = n + O(L^{-2+\theta}) \end{array} \right\} \lesssim_{\theta, \beta} L^{d-2+\theta}. \quad (4.59)$$

Proof. Let

$$\#_{m,n,L,\beta} = \# \left\{ \begin{array}{l} x, y \in \mathbb{Z}_L^d : \quad x + y = m \\ |x|, |y| \lesssim L^\theta \quad \Lambda_\beta(x) + \Lambda_\beta(y) = n + O(L^{-2+\theta}) \end{array} \right\} \quad (4.60)$$

Then the lemma is equivalent to

$$\mathbb{P}_\beta \left(\sup_{\substack{m,n \\ m \lesssim 1}} \#_{m,n,L,\beta} \lesssim_{\theta, \beta} L^{d-2+\theta} \right) = 1. \quad (4.61)$$

which is also equivalent to

$$\mathbb{P}_\beta \left(\sup_L \sup_{\substack{m,n \\ m \lesssim 1}} L^{-(d-2+\theta)} \#_{m,n,L,\beta} < \infty \right) = 1. \quad (4.62)$$

The proof of (4.62) is divided into several steps.

Step 1. In this step, we show that in the proof of (4.62), we may replace \sup_L and $\sup_{m,n}$ by supremum over discrete set \mathcal{L} and \mathcal{M}, \mathcal{N} , i.e. $\sup_{L \in \mathcal{L}}$ and $\sup_{m \in \mathcal{M}, n \in \mathcal{N}}$.

Since $x, y \in \mathbb{Z}_L^d$, we may assume that $m \in \mathbb{Z}_L^d$ and we take $\mathcal{M} = \mathbb{Z}_L^d$.

Since if $|n - n'| \lesssim L^{-2}$, then $\Lambda_\beta(x) + \Lambda_\beta(y) = n + O(L^{-2+\theta}) \Rightarrow \Lambda_\beta(x) + \Lambda_\beta(y) = n' + O(L^{-2+\theta})$, we know that (4.62) is true for all $n \in \mathbb{Z}_{L^2}^d$ implies that it is true for all $n' \in \mathbb{R}^d$. We can thus take $\mathcal{N} = \mathbb{Z}_{L^2}^d$.

Let $K = \lceil \log_2 L \rceil$ and $l = L/2^K$, then $L = l \cdot 2^K$ and $l \in [1, 2]$. Since if $|l - l'| \lesssim L^{-2}$, then $\Lambda_\beta(x) + \Lambda_\beta(y) = n + O(L^{-2+\theta}) \Rightarrow \Lambda_\beta(x) + \Lambda_\beta(y) = n' + O(L^{-2+\theta})$, we know that (4.62) is true for all $l \in [1, 2]$ implies that it is true for all $l \in [1, 2] \cap \mathbb{Z}_{L^2}$. We can thus take $\mathcal{L} = \{L = l \cdot 2^K : l \in [1, 2] \cap \mathbb{Z}_{L^2}, K \in \mathbb{N}_+\}$.

Therefore, to prove (4.62), it suffices to show that

$$\mathbb{P}_\beta \left(\sup_{L \in \mathcal{L}} \sup_{\substack{m \in \mathcal{M}, m \lesssim 1 \\ n \in \mathcal{N}}} L^{-(d-2+\theta)} \#_{m,n,L,\beta} < \infty \right) = 1. \quad (4.63)$$

In what follows, we show that

$$\mathbb{P}_\beta \left(\sup_{L \in \mathcal{L}} L^{-(d-2+\theta)} \#_{m,n,L,\beta} < \infty \right) = 1. \quad (4.64)$$

(4.64) implies that $\sup_{L \in \mathcal{L}} L^{-(d-2+\theta)} \#_{m,n,L,\beta} = \infty$ is a null set. Since a countable union of null sets is still a null set, (4.63) is a corollary of (4.64).

Step 2. In this step, we derive (4.64) from the following expectation bound.

$$\mathbb{E}(\#_{m,n,L,\beta}) \lesssim L^{d-2+\theta/2}. \quad (4.65)$$

(4.65) implies that

$$\mathbb{E}(\#_{m,n,l \cdot 2^K, \beta}) \lesssim 2^{(d-2)K + \theta K/2}. \quad (4.66)$$

Therefore, we have

$$\mathbb{E} \left(\sum_K 2^{-(d-2+\theta)K} \#_{m,n,l \cdot 2^K, \beta} \right) \lesssim \sum_K 2^{-\theta K/2} \leq C_\theta < \infty. \quad (4.67)$$

This implies that almost surely

$$\sum_K 2^{-(d-2+\theta)K} \#_{m,n,l \cdot 2^K, \beta} < \infty. \quad (4.68)$$

Let $C_{\beta, \theta} = \sum_K 2^{-(d-2+\theta)K} \#_{m,n,l \cdot 2^K, \beta}$ which is finite almost surely. We get

$$\#_{m,n,l \cdot 2^K, \beta} \leq C_{\beta, \theta} 2^{(d-2+\theta)K}, \quad (4.69)$$

which is equivalent to (4.64).

Step 3. In this step, we prove the expectation bound (4.65).

Let's calculate the expectation.

$$\begin{aligned} \mathbb{E}(\#_{m,n,L,\beta}) &= \mathbb{E} \left(\sum_{\substack{x,y \in \mathbb{Z}_L^d, x+y=m \\ |x|, |y| \lesssim L^\theta}} \chi_{\Lambda_\beta(x) + \Lambda_\beta(y) = n + O(L^{-2})} \right) \\ &= \sum_{\substack{x,y \in \mathbb{Z}_L^d, x+y=m \\ |x|, |y| \lesssim L^\theta}} \mathbb{P}_\beta(|\Lambda_\beta(x) + \Lambda_\beta(y) - n| \lesssim L^{-2}) \\ &= \sum_{\substack{x,y \in \mathbb{Z}_L^d, x+y=m \\ |x|, |y| \lesssim L^\theta}} \mathbb{P}_\beta(|F_{x,y}(\beta)| \lesssim L^{-2}) \end{aligned} \quad (4.70)$$

Here we define $F_{x,y}(\beta) = \Lambda_\beta(x) + \Lambda_\beta(y) - n$.

We estimate $\mathbb{P}_\beta(|F_{x,y}(\beta)| \lesssim L^{-2})$ using coarea formula.

$$\begin{aligned}
\mathbb{P}_\beta(|F_{x,y}(\beta)| \lesssim L^{-2}) &= \text{Area}(|F_{x,y}(\beta)| \lesssim L^{-2}) \\
&= \int_{-CL^{-2}}^{CL^{-2}} \int_{\{F_{x,y}=s\}} \frac{1}{|\nabla F_{x,y}|} d\sigma ds \\
&\lesssim L^{-2} \sup |\nabla F_{x,y}|^{-1} \text{Area}(F_{x,y} = s) \\
&\lesssim L^{-2} \sup |\nabla F_{x,y}|^{-1}
\end{aligned} \tag{4.71}$$

We thus need to get lower bound of $|\nabla F_{x,y}|$.

A simple calculation gives

$$\nabla_\beta F_{x,y}(\beta) = \frac{1}{2\Lambda_\beta(x)} \begin{bmatrix} x_1^2 \\ \vdots \\ x_d^2 \end{bmatrix} + \frac{1}{2\Lambda_\beta(y)} \begin{bmatrix} y_1^2 \\ \vdots \\ y_d^2 \end{bmatrix}, \tag{4.72}$$

which implies that

$$\begin{aligned}
|\nabla_\beta F_{x,y}(\beta)|^2 &= \frac{1}{4\Lambda_\beta^2(x)} (x_1^4 + \cdots x_d^4) + \frac{1}{4\Lambda_\beta^2(y)} (y_1^4 + \cdots y_d^4) + \frac{1}{2\Lambda_\beta(x)\Lambda_\beta(y)} (x_1^2 y_1^2 + \cdots x_d^2 y_d^2) \\
&\gtrsim L^{-} (|x|^2 + |y|^2)^2.
\end{aligned} \tag{4.73}$$

Therefore, we have

$$|\nabla_\beta F_{x,y}(\beta)| \gtrsim L^{-} (|x|^2 + |y|^2). \tag{4.74}$$

By (4.71), we get for $(x, y) \neq (0, 0)$

$$\mathbb{P}_\beta(|F_{x,y}(\beta)| \lesssim L^{-2}) \lesssim L^{-2+} (|x|^2 + |y|^2)^{-1}. \tag{4.75}$$

and for $x = y = 0$

$$\mathbb{P}_\beta(|F_{0,0}(\beta)| \leq 1). \tag{4.76}$$

By (4.70),

$$\begin{aligned}
\mathbb{E}(\#_{m,n,L,\beta}) &\lesssim 1 + L^{-2+} \sum_{\substack{x,y \in \mathbb{Z}_L^d, x+y=m \\ |x|,|y| \lesssim L^\theta}} (|x|^2 + |y|^2)^{-1} \\
&\lesssim L^{d-2+\theta/2}.
\end{aligned} \tag{4.77}$$

Therefore, we complete the proof of the expectation bound (4.65) and thus the proof of Lemma 4.3.5 \square

The derivation of (4.34) and the $+$ case of (4.35) from Lemma 4.3.5 can be done by applying a similar argument of (4.48) and (4.49). \square

Bibliography

- [1] I. Ampatzoglou, C. Collot, and P. Germain. Derivation of the kinetic wave equation for quadratic dispersive problems in the inhomogeneous setting. *arXiv preprint arXiv:2107.11819*, 2021.
- [2] D. Benney and A. C. Newell. Random wave closures. *Studies in Applied Mathematics*, 48(1):29–53, 1969.
- [3] D. Benney and P. G. Saffman. Nonlinear interactions of random waves in a dispersive medium. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 289(1418):301–320, 1966.
- [4] T. Buckmaster, P. Germain, Z. Hani, and J. Shatah. Onset of the wave turbulence description of the longtime behavior of the nonlinear Schrödinger equation. *Inventiones mathematicae*, 225(3):787–855, 2021.
- [5] C. Collot, H. Dietert, and P. Germain. Stability and cascades for the kolmogorov-zakharov spectrum of wave turbulence. *arXiv preprint arXiv:2208.00947*, 2022.
- [6] C. Collot and P. Germain. On the derivation of the homogeneous kinetic wave equation. *arXiv preprint arXiv:1912.10368*, 2019.
- [7] C. Collot and P. Germain. Derivation of the homogeneous kinetic wave equation: longer time scales. *arXiv preprint arXiv:2007.03508*, 2020.
- [8] R. Davidson. *Methods in nonlinear plasma theory*. Elsevier, 2012.

- [9] Y. Deng and Z. Hani. Full derivation of the wave kinetic equation. *arXiv preprint arXiv:2104.11204*, 2021.
- [10] Y. Deng and Z. Hani. On the derivation of the wave kinetic equation for NLS. In *Forum of Mathematics, Pi*, volume 9. Cambridge University Press, 2021.
- [11] Y. Deng and Z. Hani. Propagation of chaos and the higher order statistics in the wave kinetic theory. *arXiv preprint arXiv:2110.04565*, 2021.
- [12] Y. Deng and Z. Hani. Rigorous justification of the wave kinetic theory. *arXiv preprint arXiv:2207.08358*, 2022.
- [13] Y. Deng and Z. Hani. Derivation of the wave kinetic equation: full range of scaling laws. *arXiv preprint arXiv:2301.07063*, 2023.
- [14] A. Dymov and S. Kuksin. Formal expansions in stochastic model for wave turbulence 1: kinetic limit. *Communications in Mathematical Physics*, 382:951–1014, 2021.
- [15] A. Dymov and S. Kuksin. Formal expansions in stochastic model for wave turbulence 2: method of diagram decomposition. *Journal of Statistical Physics*, 190(1):3, 2023.
- [16] A. Dymov, S. Kuksin, A. Maiocchi, and S. Vladuts. The large-period limit for equations of discrete turbulence. *arXiv preprint arXiv:2104.11967*, 2021.
- [17] A. V. Dymov and S. B. Kuksin. On the zakharov–lvov stochastic model for wave turbulence. In *Doklady Mathematics*, volume 101, pages 102–109. Springer, 2020.
- [18] L. Erdős, M. Salmhofer, and H.-T. Yau. Quantum diffusion of the random schrödinger evolution in the scaling limit. 2008.

- [19] L. Erdős and H.-T. Yau. Linear boltzmann equation as the weak coupling limit of a random schrödinger equation. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 53(6):667–735, 2000.
- [20] M. Escobedo and J. Velázquez. Finite time blow-up and condensation for the bosonic nordheim equation. *Inventiones mathematicae*, 200:761–847, 2015.
- [21] M. Escobedo and J. J. Velázquez. *On the theory of weak turbulence for the nonlinear Schrödinger equation*. American Mathematical Soc., 2015.
- [22] E. Faou. Linearized wave turbulence convergence results for three-wave systems. *Communications in Mathematical Physics*, 378:807–849, 2020.
- [23] I. M. Gamba, L. M. Smith, and M.-B. Tran. On the wave turbulence theory for stratified flows in the ocean. *Mathematical Models and Methods in Applied Sciences*, 30(01):105–137, 2020.
- [24] P. Germain, A. D. Ionescu, and M.-B. Tran. Optimal local well-posedness theory for the kinetic wave equation. *Journal of Functional Analysis*, 279(4):108570, 2020.
- [25] A. Hannani, M. Rosenzweig, G. Staffilani, and M.-B. Tran. On the wave turbulence theory for a stochastic kdv type equation—generalization for the inhomogeneous kinetic limit. *arXiv preprint arXiv:2210.17445*, 2022.
- [26] K. Hasselmann. On the non-linear energy transfer in a gravity-wave spectrum part 1. general theory. *Journal of Fluid Mechanics*, 12(4):481–500, 1962.
- [27] K. Hasselmann. On the non-linear energy transfer in a gravity wave spectrum part 2. conservation theorems; wave-particle analogy; irrevesibility. *Journal of Fluid Mechanics*, 15(2):273–281, 1963.

- [28] P. A. Janssen. Progress in ocean wave forecasting. *Journal of Computational Physics*, 227(7):3572–3594, 2008.
- [29] K. Jin and X. Ma. Cancellations of resonances and long time dynamics of cubic Schrödinger equation on \mathbb{T} . *Communications in Mathematical Physics*, 381(3):1309–1368, 2021.
- [30] J. Lukkarinen and H. Spohn. Weakly nonlinear Schrödinger equation with random initial data. *Inventiones mathematicae*, 183(1):79–188, 2011.
- [31] S. Nazarenko. *Wave turbulence*, volume 825. Springer Science & Business Media, 2011.
- [32] T. Oh. Ornstein-uhlenbeck process on abstract wiener spaces, hypercontractivity, and logarithmic sobolev inequality. *Unpublished notes*, 2015.
- [33] W. M. Organization. *Guide to wave analysis and forecasting*. Secretariat of the World Meteorological Organization, 1998.
- [34] R. Peierls. Zur kinetischen theorie der wärmeleitung in kristallen. *Annalen der Physik*, 395(8):1055–1101, 1929.
- [35] B. Rumpf, A. Soffer, and M.-B. Tran. On the wave turbulence theory: ergodicity for the elastic beam wave equation. *arXiv preprint arXiv:2108.13223*, 2021.
- [36] A. Soffer and M.-B. Tran. On the dynamics of finite temperature trapped bose gases. *Advances in Mathematics*, 325:533–607, 2018.
- [37] A. Soffer and M.-B. Tran. On the energy cascade of 3-wave kinetic equations: beyond kolmogorov–zakharov solutions. *Communications in Mathematical Physics*, 376(3):2229–2276, 2020.
- [38] H. Spohn. The phonon boltzmann equation, properties and link to weakly anharmonic lattice dynamics. *Journal of statistical physics*, 124:1041–1104, 2006.

- [39] H. Spohn. On the boltzmann equation for weakly nonlinear wave equations. *Boltzmann's legacy, ESI Lect. Math. Phys*, pages 145–159, 2008.
- [40] G. Staffilani and M.-B. Tran. On the wave turbulence theory for stochastic and random multidimensional KdV type equations. *arXiv preprint arXiv:2106.09819*, 2021.
- [41] G. Staffilani and M.-B. Tran. On the wave turbulence theory for the beam wave equation. *arXiv preprint arXiv:2301.07903*, 2023.
- [42] A. Vedenov. Theory of a weakly turbulent plasma. *Reviews of Plasma Physics: Volume 3*, pages 229–276, 1967.
- [43] V. E. Zakharov. Weak turbulence in media with a decay spectrum. *Journal of Applied Mechanics and Technical Physics*, 6(4):22–24, 1965.
- [44] V. E. Zakharov, V. S. L'vov, and G. Falkovich. *Kolmogorov spectra of turbulence I: Wave turbulence*. Springer Science & Business Media, 2012.
- [45] G. Zaslavskii and R. Sagdeev. Limits of statistical description of a nonlinear wave field. *Soviet physics JETP*, 25:718–724, 1967.