HAL open science

On the deviations of ergodic averages on the torus and their applications

Hao Wu

▶ To cite this version:

Hao Wu. On the deviations of ergodic averages on the torus and their applications. Dynamical Systems [math.DS]. Université Paris Cité, 2022. English. NNT : 2022UNIP7296 . tel-04486279

HAL Id: tel-04486279

https://theses.hal.science/tel-04486279

Submitted on 1 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Paris Cité

École doctorale de sciences mathématiques de Paris centre (386)

Institut de Mathématiques de Jussieu-Paris Rive Gauche

On the deviations of ergodic averages on the torus and their applications.

Par

Hao WU

Thèse de doctorat

Discipline : Mathématiques

Dirigée par Bassam FAYAD

Présentée et soutenue le 17 novembre 2022 devant le jury composé de :

Mme Valérie BERTHE	DR	Université Paris Cité	examinatrice
M. Adrien BOYER	MCF	Université Paris Cité	examinateur
M. Nicolas CHEVALLIER	PU	Université de Haute Alsace	examinateur
M. Bassam FAYAD	DR	IMJ-PRG	directeur
M. François MAUCOURANT	MCF-HDR	Université Rennes I	rapporteur
M. Andreas STRÖMBERGSSON	PU	Uppsala University	rapporteur

Institut de mathématiques de Jussieu-
Paris Rive gauche. UMR 7586.
Boîte courrier 247
4 place Jussieu
75 252 Paris Cedex 05

Université Paris Cité.
École doctorale de sciences
mathématiques de Paris centre.
Boîte courrier 290
4 place Jussieu
75 252 Paris Cedex 05

Remerciements

First and foremost I would like to express my greatest gratitude to my supervisor, Bassam Fayad, for leading me into the fantastic world of dynamics and number theory, for his constant new ideas when we ran into difficulties, and for his generous help both in academic and in life. During the past several years, I have benefited enormously from the discussions with him. This thesis would not have existed without his help.

I also want to thank Fatna Abdedou, my academic sister, for her support and encouragement during our collaboration.

I would like to thank François Maucourant and Andreas Strömbergsson for accepting to be the reviewers of my thesis. I am immensely grateful for their careful reading and valuable comments.

I would like to thank, Valérie Berthé, Adrien Boyer, and Nicolas Chevallier for agreeing to be on the jury of my defense.

The past years have given me many precious memories, which owe a lot to my friends : Oussama Bensaid, Sion Chan-Lang, Jiaming Chen, Ratko Darda, Yanbo Fang, Yue Huang, Rodolfo Joaquín Gutiérrez Romo, Minsung Kim, Mingkun Liu, Long Meng, Rahman Mohammadpour, Davi Obata, Yi Pan, Mauricio Poletti, Grégoire Sergeant-Perthuis, Jieao Song, Frank Trujillo, Ioann Vasilyev, Chen Yan, Min Yu, Rui Yuan.

Lastly I want to thank my family for their constant support, which enables me to pursue my passion in math. And I want to thank my then girlfriend, Yangyang Cheng, who is now my wife, for her companion during the years, which makes my life in Paris much more enjoyable.

Résumé

Nous prouvons plusieurs résultats sur les distributions limites et les bornes absolues des fonctions de discrépances pour des actions de translation sur le tore. Nous montrons l'existence de lois limites, après normalisation adéquate, pour les fonctions de discrépance des actions $\mathbb{Z}^2$ par translations sur le tore relativement à des formes symétriques strictement convexes. Nous obtenons également des bornes supérieures presque optimales pour les fonctions de discrépance dans le cas des translations sur le tore relativement aux polygones et dans le cas des formes linéaires relativement aux intervalles sur le cercle, qui étendent le résultat de Beck des bornes absolues pour les translations sur le tore relativement aux boîtes droites vers des cas plus généraux. En nous appuyant sur l'étude des discrépances, nous donnons des exemples de marches aléatoires ergodiques sur $\mathbb{R}^2$ guidés par les retours dans une boîte d'une translation irrationnelle sur le tore.

Mots-clés

Discrépances ergodiques, Dynamiques homogènes, Approximations diophantiennes, Marches aléatoires ergodiques

On the deviations of ergodic averages on the torus and their applications.

Abstract We prove several results on limiting distributions and absolute bounds of discrepancy functions for translation actions on the torus. We show the existence of limit laws, after adequate normalizations, for the discrepancy functions of $\mathbb{Z}^2$-actions by translations on the torus relative to symmetric, strictly convex bodies. We also obtain almost optimal upper bounds for the maximal discrepancy functions in the case of toral translations relative to polygons and in the case of linear forms relative to intervals on the circle, which extend Beck's result of absolute bounds on toral translations for straight boxes to more general settings. Building on the study of discrepancy functions, we give examples of ergodic random walks on $\mathbb{R}^2$ driven by the returns to a box of an irrational toral translation.

Keywords

Ergodic discrepancies, Homogeneous dynamics, Diophantine approximations, Ergodic random walks

Résumé substantiel en français

Les principaux objets de recherche de ce travail sont les comportements asymptotiques des écarts ergodiques dans les problèmes de comptage sur un tore de dimension d. Étant donné une transformation ergodique $T : X \to X$ sur un espace mesuré (X, Σ, μ), le théorème de Birkhoff énonce que pour une fonction μ-intégrable f, sa moyenne temporelle partant de un point générique x, $\sum_{n=0}^{N-1} f(T^n x)/N$, converge vers sa moyenne spatiale $\int f d\mu/\mu(X)$. Si nous prenons la fonction f comme fonction caractéristique d'un ensemble mesurable $\mathcal{C} \subset X$, nous obtenons un problème de comptage, et le théorème de Birkhoff stipule que la proportion de "visites" de l'orbite de x à l'ensemble $\mathcal{C}$ converge vers son volume $\mathrm{Vol}(\mathcal{C})$.

Dans le cadre classique de la translation sur le tore, l'espace X est choisi comme étant un tore d-dimensionnel $\mathbb{T}^d$ équipé de la mesure de Haar, la transformation ergodique est la translation irrationnelle $T_{\boldsymbol{\alpha}} : x \mapsto x + \boldsymbol{\alpha}$, $x \in \mathbb{T}^d$, où $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_d) \in \mathbb{R}^d$ et $1, \alpha_1, \ldots, \alpha_d$ sont linéairement indépendants sur $\mathbb{Q}$; la transformation est dans le sens de prendre modulo 1 pour chaque coordonnée. Notre objet d'intérêt est l'écart ergodique, défini comme la différence entre le nombre de visites réelles et sa valeur attendue :

$$D_{\mathcal{C}}(x, \boldsymbol{\alpha}\,; N) = \sum_{n=0}^{N-1} \chi_{\mathcal{C}}(T_{\boldsymbol{\alpha}}^n(x)) - N\ \mathrm{Vol}(\mathcal{C}), \tag{1}$$

et deux questions naturelles sont les suivantes :

Q1. Distributions limites. Si on laisse le point de départ x et la translation $T_{\boldsymbol{\alpha}}$ aléatoires, x, $\boldsymbol{\alpha} \in \mathbb{T}^d$ aléatoires, existe-t-il une distribution limite, après normalisation adéquate, pour la fonction d'écart $D_{\mathcal{C}}$ comme $N \to \infty$?

Q2. Bornes absolues. Si nous fixons une transformation $T_{\boldsymbol{\alpha}}$, quelle est la borne supérieure (absolue) optimale pour la fonction d'écart $D_{\mathcal{C}}$ comme $N \to \infty$?

Ces deux questions forment les axes de ce travail, et les chapitres suivants tentent de répondre aux questions ci-dessus dans différents contextes des actions $T_{\boldsymbol{\alpha}}$ et des formes de l'ensemble $\mathcal{C}$.

Les distributions limites

Dans les années 1960, Kesten[Kes60, Kes62] a étudié le problème de la distribution limite dans le cas des rotations sur le cercle en dimension 1. Il a prouvé que la fonction d'écart, relativement à un intervalle, converge vers une distribution de Cauchy après avoir été normalisée par $\ln N$. Sa preuve s'appuie fortement sur l'algorithme de la fraction continue, ce qui rend difficile les généralisations à des dimensions supérieures. Dans les années 2000, Marklof[Mar07] a adopté une approche dynamique du problème de distribution, montrant que de telles fonctions d'écart normalisées sont souvent intimement liées à certaines orbites sous les flots géodésiques et unipotents sur des espaces homogènes.

Suivant l'approche dynamique, Dolgopyat et Fayad ont étendu les résultats de Kesten aux tores de dimension supérieure lorsque $d \geq 2$. Ils ont trouvé deux comportements distincts pour $\mathcal{C}$ étant une boule et $\mathcal{C}$ étant une boîte[DF12, DF14], qui sont tous deux des généralisations naturelles de l'intervalle en dimension 1. Ils ont montré que les boîtes à d-dimensions se comportent de la même manière que les intervalles à une dimension, et que la fonction d'écart converge également vers une distribution de Cauchy après avoir été normalisée par $(\ln N)^d$. Quant aux boules, ils ont montré que le normalisateur devrait être $N^{(d-1)/2d}$, et que la fonction d'écart converge vers une fonction de distribution définie sur l'espace produit des tores infinis et l'espace de Réseaux $(d+1)$-dimensionnel.

Dans les deux cas, le premier ingrédient de la preuve est l'analyse harmonique de la série de Fourier de l'écart, qui met en évidence la contribution principale de l'écart et aide à trouver le bon normalisateur. Le deuxième ingrédient est le principe de correspondance de Dani qui transfère l'étude des petits diviseurs qui apparaît dans la contribution principale à l'étude de la dynamique des actions diagonales et unipotentes sur l'espace homogène des réseaux. Dans le cas des boîtes, la contribution principale provient de tous les modes de Fourier dont les normes sont *inférieures* à l'ordre N, ce qui correspond aux excursions vers le cusp *jusqu'au* temps $\ln N$ des actions diagonales qui codent les petits diviseurs, ce qui conduit à une distribution de Cauchy similaire au cas unidimensionnel. Dans le cas des boules, la contribution principale ne vient que des modes de Fourier dont les normes sont d'ordre *de* N, ce qui correspond aux excursions vers le cusp à la seule échelle de temps $\ln N$ des actions diagonales qui codent les petits diviseurs, ce qui conduit à une distribution limite qui peut être calculée comme les ensembles de niveau de la fonction d'écart exprimée dans le langage des reseaux.

Dans le Chapitre 2, suivant une approche similaire à celle de Dolgopyat et Fayad, nous considérons la séquence d'actions $\mathbb{Z}^2$, $\{(n_1\alpha_1, n_2\alpha_2) \in \mathbb{T}^2\}_{1\leq n_1,n_2\leq N}$ par rapport à un disque centré $\mathcal{C}$ dans $\mathbb{T}^2$, ce qui correspond au positionnement du réseau translaté $\mathbb{Z}\alpha_1 \oplus \mathbb{Z}\alpha_2 + x$ par rapport au reseau $\mathbb{Z}^2$ dans la norme euclidienne en $\mathbb{R}^2$, rappelons que x est le point de départ de l'action. Plus précisément, la fonction d'écart mesure la fréquence à laquelle les deux réseaux deviennent $\mathcal{C}$-proches l'un de l'autre dans un rectangle croissant,

il est défini par :

$$D_{\mathcal{C}}(r,x,\boldsymbol{\alpha};N) = \sum_{\substack{0\le n_1\le N-1\\ 0\le n_2\le N-1}} \chi_{\mathcal{C}_r}(x_1+n_1\alpha_1, x_2+n_2\alpha_2) - N^2\mathrm{Vol}(\mathcal{C}_r). \tag{2}$$

Le paramètre r est la remise à l'échelle de $\mathcal{C}$ pour supprimer une éventuelle dépendance irrégulière de la distribution limite sur $\mathcal{C}$. Nous allons montrer qu'il existe deux comportements très différents pour différentes parties de la série de Fourier de $D_{\mathcal{C}}$. Après avoir choisi des normalisateurs appropriés, la partie *avec* 0-coordonnées se comporte comme la somme ergodique des observables lisses, et la partie *sans* 0-coordonnées se comporte de manière similaire aux déviations ergodiques des translations sur le tore pour les corps convexes.

Theorem 0.1. *Soit $\mathcal{C}$ un corps analytique symétrique strictement convexe qui rentre dans le cube unitaire de $\mathbb{R}^2$, et $D_{\mathcal{C}}$ défini comme dans* (2)*, on peut décomposer $D_{\mathcal{C}}$ en deux parties, notées $D_{\mathcal{C},1}$ et $D_{\mathcal{C},2}$, alors on a*

(a) Pour $D_{\mathcal{C},1}$, soit $(x,\boldsymbol{\alpha})$ uniformément distribué dans $\mathbb{T}^2\times\mathbb{T}^2$, pour tout $rfixe$, il existe une distribution limite pour $D_{\mathcal{C},1}/N$, quand $N\to\infty$.

(b) Pour $D_{\mathcal{C},2}$, soit $(r,x,\boldsymbol{\alpha})$ uniformément distribué dans $[a,b]\times\mathbb{T}^2\times\mathbb{T}^2$, alors il existe une distribution limite pour $D_{\mathcal{C},2}/(\sqrt{rN})$, quand $N\to\infty$.

Ce résultat montre que la distribution de la fonction d'écart pour l'action $\mathbb{Z}^2$ sur $\mathbb{T}^2$ par rapport à un ensemble convexe peut être comprise de la même manière qu'un développement de Taylor. Géométriquement, la partie principale avec le normalisateur N représente la résonance qui se produit naturellement puisque les deux réseaux $\mathbb{Z}\alpha_1\oplus\mathbb{Z}\alpha_2$ et $\mathbb{Z}^2$ sont parallèles. La partie restante avec le normalisateur $N^{\frac{1}{2}}$ représente l'erreur de résidu, c'est-à-dire l'interaction entre la frontière convexe lisse de $\mathcal{C}$ et l'action $\mathbb{Z}^2$.

Bien que le résultat soit indiqué en 2 dimensions, en choisissant des normalisateurs appropriés pour différentes parties de la série de Fourier, des résultats similaires peuvent être obtenus pour les dimensions *pair* de manière simple. Alors que pour les dimensions impaires, la "somme moyenne" des modes de Fourier composés de $\frac{d+1}{2}$ coordonnées nulles a des comportements fondamentalement différents, et nous n'avons pas encore de description complète pour cela.

Dans le cas général des actions $\mathbb{Z}^2$, où l'on considère le reseaux $\mathbb{Z}\boldsymbol{\alpha}_1\oplus\mathbb{Z}\boldsymbol{\alpha}_2$ pour les vecteurs bidimensionnels $\boldsymbol{\alpha}_1$ et $\boldsymbol{\alpha}_2$, la partie de résonance avec le normalisateur N n'apparaît pas, et nous pensons qu'une loi limite similaire à la partie résiduelle dans le cas parallèle est valable dans le cas général de l'action $\mathbb{Z}^2$ avec un normalisateur $N^{\frac{1}{2}}$, bien que pour le moment, nous soyons confrontés à des difficultés techniques causées par le facteur supplémentaire $\ln N$ qui découle du calcul de la mesure pour les cusps bidimensionnelles.

Bornes absolues pour les translations sur le tore

Soit $\mathcal{B}$ l'ensemble de toutes les boîtes droites dans $\mathbb{T}^d$, puis en prenant la borne supérieure absolue de la fonction d'écart (1) pour $\mathcal{C} \in \mathcal{B}$ et tous les points de départ possibles $x \in \mathbb{T}^d$, nous arrivons à la fonction d'écart maximal suivante :

$$\Delta(\boldsymbol{\alpha}; N) = \max_{\substack{\mathcal{C}\in\mathcal{B},\\ x\in\mathbb{T}^d}} \left| \sum_{1\le n\le N} \chi_{\mathcal{C}}(x + n\boldsymbol{\alpha}) - n\mathrm{Vol}(\mathcal{C}) \right|. \tag{3}$$

En 1923, Khintchine[Khi23] a prouvé les bornes précises suivantes pour $d = 1$:

$$\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} < \infty \Longleftrightarrow \frac{\Delta(\boldsymbol{\alpha}\ ; N)}{(\ln N)^d \varphi(\ln\ln N)} \text{ est borné pour presque chaque } \boldsymbol{\alpha} \in \mathbb{T}^d, \tag{4}$$

où $\varphi(n)$ est une fonction croissante positive arbitraire de n. L'outil majeur de sa preuve est la théorie métrique des irrationnels qui repose principalement sur les fractions continues. En 1964, Schmidt[Sch64] a obtenu une borne supérieure de $(\ln N)^{d+1+\epsilon}$ pour $d > 1$ en utilisant l'inégalité d'Erdős–Turán. Mais du fait de l'absence d'algorithme de fraction continue en dimension supérieure, le cas $d > 1$ n'a été complètement résolu qu'en 1994 par Beck[Bec94]. Le résultat de Beck a étendu la loi zéro-un précise de Khintchine dans (4) au cas des boîtes droites de dimensions supérieures. Beck a observé que, comme dans le cas unidimensionnel, la série de Fourier de l'écart maximal $\Delta(\boldsymbol{\alpha}; N)$ est en fait une somme de termes presque *indépendants par paire* de $\boldsymbol{\alpha}$, et l'étude de $\Delta(\boldsymbol{\alpha}; N)$ réduite à l'analyse des petits diviseurs des formes linéaires $\|n_1\alpha_1 + \cdots + n_k\alpha_k\|$. En employant " la méthode du second moment " (voir [Sch60, Sch64]), Beck a montré que l'algorithme de fraction continue utilisé dans le cas 1-d peut être efficacement remplacé par des estimations précises de la cardinalité des modes de Fourier pour dont les diviseurs sont limités à de petits intervalles. L'amélioration du facteur critique de $\ln N$ est obtenue par les annulations des parties positives avec les parties négatives pour les termes à diviseurs presque identiques. Pour une brillante discussion sur la façon dont les "irrégularités globales" des fonctions de discordance proviennent des "irrégularités locales" des propriétés diophantiennes du vecteur de translation $\boldsymbol{\alpha}$, nous nous référons à l'introduction de [Bec94].

Dans le chapitre 3, en allant dans une direction différente en dimension 1, au lieu de rotations irrationnelles, nous pouvons considérer la séquence de formes d-linéaires liées à $\boldsymbol{\alpha}$: $\left\{\sum_{1\le i\le d} k_i\alpha_i\right\}_{\substack{1\le k_i\le N\\ 1\le i\le d}}$, dont la fonction d'écart relativement à un intervalle $[0, x)$ est défini comme

$$D(\boldsymbol{\alpha}, x; N) = \sum_{\substack{1\le k_i\le N\\ 1\le i\le d}} \chi_{[0,x)} \left(\sum_{1\le i\le d} k_i\alpha_i \right) - N^d x,$$

et la fonction d'écart maximale est définie de la même manière :

$$\Delta(\boldsymbol{\alpha}; N) = \max_{0<x\leq 1} |D(\boldsymbol{\alpha}, x; N)| .$$

En généralisant les idées de Beck aux petits diviseurs $n\prod_{i=1}^{d} \|n\alpha_i\|$, on obtient une borne supérieure *presque optimale* pour l'écart maximal des formes linéaires :

Theorem 0.2. *Pour tout $\epsilon > 0$, il existe un ensemble complet de mesures de $\boldsymbol{\alpha} \in \mathbb{R}^d$, tel que*

$$\Delta(\boldsymbol{\alpha}\ ; N) \leq C(\boldsymbol{\alpha}, \epsilon)(\ln N)^d \cdot (\ln\ln N)^{\max\{3,d\}+\epsilon}, \tag{5}$$

pour une constante $C(\boldsymbol{\alpha}, \epsilon) > 0$, et

$$\Delta(\boldsymbol{\alpha}\ ; N) > (\ln N)^d \cdot \ln\ln N \tag{6}$$

pour une infinité de $N \in \mathbb{N}$.

Par rapport aux translations sur le tore, le facteur supplémentaire $(\ln\ln N)^{\max\{d-1,2\}}$ dans le cas des $d-$formes linéaires dans notre preuve est nécessaire pour contrôler les petits diviseurs $n\prod_{i=1}^{d} \|n\alpha_i\|$. Une observation intéressante est que dans la preuve, il existe une *dualité* entre le cas des translations sur le tore relativement aux boîtes droites d-dimensionnelles et le cas des formes d-linéaires relativement aux intervalles 1-dimensionnels. Les petits diviseurs apparaissant dans le cas des traductions torales sont des formes d-linéaires, et les petits diviseurs dans le cas des formes linéaires sont des produits de coordonnées de translation sur le tore. En utilisant cette dualité, nous pourrions simplifier la preuve pour un cas en appliquant l'inégalité d'Erdős–Turán pour obtenir une version plus faible du résultat de son cas dual, ce qui permet d'obtenir une estimation plus rapide de la cardinalité des ensembles dyadiques pour le cas cible, ce qui est une étape importante dans les deux preuves.

Dans le Chapitre 4, suite aux travaux de Beck sur les boîtes droites, nous considérons des translations torales relativement à des triangles rectangles à jambes droites (parallèles aux côtés du tore) dans $\mathbb{T}^2$. En remplaçant $\mathcal{B}$ dans (3) par l'ensemble des triangles droits d'hypoténuse de pente τ, on peut définir de façon similaire la fonction d'écart maximal correspondante $\Delta(\boldsymbol{\alpha}, \tau; N)$ pour les triangles droits. Ici la principale difficulté vient de l'hypoténuse du triangle rectangle. L'inclinaison τ déplace les diviseurs de la série de Fourier de la fonction d'écart de $n_1n_2(n_1\alpha_1+n_2\alpha_2-n_3)$ à $n_1(n_1\tau-n_2)(n_1\alpha+n_2\alpha_2-n_3)$, ce qui ajoute une possibilité supplémentaire que le facteur $|n_1\tau-n_2|$ soit petit. En généralisant les critères locaux de Beck aux nouveaux diviseurs, on obtient une borne supérieure *presque optimale* et la même borne inférieure pour les écarts maximaux dans le cas des triangles :

Theorem 0.3. *Fixons tout $\epsilon > 0$, pour presque chaque $\boldsymbol{\alpha} \in \mathbb{R}^2$ et presque chaque $\tau \in \mathbb{R}^+$, l'écart maximal $\Delta(\boldsymbol{\alpha}, \tau; N)$ est essentiellement délimité par le haut par $(\ln N)^2(\ln\ln N)^{2+\epsilon}$,*

mais pas par $(\ln N)^2 \ln\ln N$. *Plus précisément, il existe une constante* $C(\boldsymbol{\alpha}, \tau, \epsilon) > 0$, *telle que*

$$\Delta(\boldsymbol{\alpha}, \tau; N) \leq C(\boldsymbol{\alpha}, \tau, \epsilon)(\ln N)^2 \cdot (\ln\ln N)^{2+\epsilon}, \tag{7}$$

et il existe une infinité de $N \in \mathbb{N}$, *tels que*

$$\Delta(\boldsymbol{\alpha}, \tau; N) > (\ln N)^2 \cdot \ln\ln N. \tag{8}$$

Il est facile de voir que les mêmes limites sont valables pour presque tous les polygones de $\mathbb{T}^2$, simplement en découpant les polygones en triangles rectangles. Malheureusement, l'énoncé ne donne pas une loi zéro-un comme les résultats de Khintchine et Beck, le facteur supplémentaire $\ln\ln N$ est nécessaire pour contrôler le petit facteur $n_1\tau - n_2$ qui résulte de l'hypoténuse de le triangle, qui ne se produit pas dans le cas des rectangles droits. Notez que lorsque la pente τ ou le vecteur de translation $\boldsymbol{\alpha}$ est Liouville, les diviseurs $n_1(n_1\tau - n_2)(n_1\alpha + n_2\alpha_2 - n_3)$ peuvent être très petits, ce qui brise facilement la borne supérieure logarithmique dans le cas diophantien. Par conséquent, la restriction à "presque tous les τ et presque tous les $\boldsymbol{\alpha}$" est nécessaire. Les polyèdres de dimension supérieure devraient suivre les mêmes estimations, mais des calculs plus complexes sont nécessaires.

Le cas des disques est nettement plus difficile et semble être un problème dual au problème du cercle de Gauss. Si nous pouvions obtenir une borne supérieure *uniforme* pour les triangles d'une classe spécifique de pentes, nous pouvons approximer le disque par des triangles et obtenir la borne supérieure $N^{\frac{1}{3}+\epsilon}$, qui correspond à la borne supérieure classique $R^{\frac{2}{3}+\epsilon}$ du problème du cercle de Gauss. Alors que la borne optimale pour le terme d'erreur dans le problème du cercle de Gauss est supposée être comprise entre $R^{\frac{1}{2}}$ et $R^{\frac{1}{2}+\epsilon}$, la borne optimale dans notre cas, les translations sur le tore relatives aux boules, est supposée être comprise entre $N^{\frac{1}{4}}$ et $N^{\frac{1}{4}+\epsilon}$, puisque le résultat de Dolgopyat et Fayad[DF14] montre déjà que la magnitude *moyenne* de la fonction d'écart est de $N^{\frac{1}{4}}$.

Cascades cylindriques

Un cadre intéressant dans lequel les fonctions d'écart apparaissent naturellement est l'étude des cascades cylindriques. Étant donné une fonction $A : \mathbb{T}^d \to \mathbb{R}^r$, une cascade cylindrique au-dessus de la translation sur le tore $T_{\boldsymbol{\alpha}}$ est définie comme $W_{\boldsymbol{\alpha},A} : \mathbb{T}^d \times \mathbb{R}^r \to \mathbb{T}^d \times \mathbb{R}^r$, donné par $W_{\boldsymbol{\alpha},A}(x, y) = (x + \boldsymbol{\alpha}, y + A(x))$. Lorsque $\boldsymbol{\alpha}$ est diophantienne et A est lisse, l'équation cohomologique linéaire $A(x) - \int_{\mathbb{T}^d} A(u)du = -B(x + \boldsymbol{\alpha}) + B(x)$ a une solution lisse B, donc $W_{\boldsymbol{\alpha},A}$ est conjugué lissement à la translation $W_{\boldsymbol{\alpha},\int_{\mathbb{T}^d} A}$. Un cas intéressant pour A non lisse est lorsque les composants de A sont des fonctions indicatrices de moyenne nulle de formes typiques dans $\mathbb{T}^d$, c'est-à-dire $A = (A_1, \ldots, A_d)$, où $A_i = \chi_{\mathcal{C}_i} - \mathrm{Vol}(\mathcal{C}_i)$, et chaque $\mathcal{C}_i$ est une forme typique telle qu'une boîte ou un polytope. Ici les sommes de Birkhoff de A_i sont précisément les déviations ergodiques dont nous avons discuté précédemment.

La cascade cylindrique $W_{\boldsymbol{\alpha},A}$ peut être vue comme une marche aléatoire déterministe dans $\mathbb{R}^r$ pilotée par la translation sur le tore sur la base $\mathbb{T}^d$, ce qui rend naturel l'étude de sa récurrence et ergodicité. Pour $r = 1$, la condition *moyenne nulle* $\int_{\mathbb{T}^d} A = 0$ est à la fois nécessaire et suffisante pour que la marche $W_{\boldsymbol{\alpha},A}$ soit récurrente lorsque α est irrationnel. Pour $d = 1$, on sait que lorsque A est de moyenne nulle à variation bornée, la marche est récurrente du fait de *l'inégalité de Denjoy-Koksma.* Nous renvoyons à [DF15], Section 8 pour une introduction plus complète des résultats en 1-d.

En dimension supérieure, seuls les exemples récurrents et transitoires sont connus lorsque $d = 2$ et $\mathcal{C}_i$ sont des polytopes[CC09]. Chevallier et Conze[CC09] ont montré que pour que la marche de tout espace de base soit récurrente, une condition suffisante est que les sommes de Birkhoff de A_i croissent plus lentement que $O(n^{1/r})$ le long d'une sous-séquence N_n pour la plupart des points de départ. En utilisant l'estimation L^2 des déviations ergodiques de A, ils ont montré que pour presque chaque vecteur de translation α, la marche $W_{\boldsymbol{\alpha},A}$ est récurrente pour les polytopes. En fait, nous pouvons voir qu'à partir de nos bornes absolues de la fonction d'écart pour les triangles, la croissance des sommes de Birkhoff de A_i au temps N est au plus d'ordre $(\ln N)^{2+\epsilon}$, ce qui conduit aussi à la récurrence par le critère de Chevallier et Conze.

L'ergodicité de la marche $W_{\alpha,A}$ peut être établie si les sommes de Birkhoff de A deviennent de mieux en mieux réparties sur $\mathbb{R}^r$ le long d'une sous-suite N_n, de sorte que la traduction $T_\alpha^{N_n}$ reste proche de l'identité. Cette idée donne naissance à la notion de *valeurs essentielles*[Sch77], qui est l'outil majeur pour notre construction de marches ergodiques pour les rectangles. Dans le chapitre 5, en collaboration avec Fatna Abdedou, nous construisons des exemples ergodiques des cascades cylindriques sur $\mathbb{T}^2 \times \mathbb{R}$ en utilisant des vecteurs de translation de Liouville. Plus précisément, nous avons le suivant :

Theorem 0.4 (Abdedou-W). *Pour presque tout rectangle droit dans $\mathbb{T}^2$ avec un coin à l'origine, soit A la fonction indicatrice centrée du rectangle, alors il existe un ensemble G_δ-dense de $\boldsymbol{\alpha}$, telle que la cascade cylindrique $W_{\boldsymbol{\alpha},A}$ soit ergodique sur $\mathbb{T}^2 \times \mathbb{R}$.*

L'énoncé est pour un ensemble G_δ-dense de $\boldsymbol{\alpha}$ puisque les conditions de Liouville doivent être imposées au vecteur de translation pour atteindre l'ergodicité dans notre construction. Les contreparties de dimension supérieure dans $\mathbb{T}^d \times \mathbb{R}^r$ peuvent être obtenues de manière similaire par la propriété de sous-groupe des valeurs essentielles, mais encore une fois notre méthode exige que les longueurs des côtés supplémentaires des boîtes de dimension supérieure soient Liouville quand $d > 2$ ou $r > 1$. Exemples ergodiques utilisant des polytopes dans $\mathbb{T}^d$, $d \geq 2$, sont plus délicates en raison de leurs côtés inclinés, et leurs constructions sont toujours en cours. Bien que les conditions de Liouville semblent inévitables avec nos techniques actuelles, en raison du taux de divergence lent de la somme de Birkhoff de A, c'est-à-dire la borne supérieure logarithmique absolue des écarts maximaux dans le cas des polygones dans Théorème 0.3, l'ergodicité de la marche est attendue pour les polytopes typiques et les vecteurs de translation typiques.

Contents

1 Introduction **17**

1.1 Limiting distributions 18

1.2 Absolute bounds for toral translations 19

1.3 Cylindrical cascades 22

2 Multidimensional Degenerate Actions - Convex Bodies **25**

2.1 Introduction 25

2.2 Limiting distributions 27

2.2.1 Limiting Distribution for the case $\bar{d} = 1$ 27

2.2.2 Limiting Distribution for the case $\bar{d} = 2$ 27

2.3 Fourier series and proof of Proposition 2.1 29

2.3.1 Fourier series for convex bodies. 30

2.3.2 Proof for the limiting distribution when $\bar{d} = 1$ 30

2.4 Non-resonant terms 32

2.5 Geometry of the space of lattices 37

2.6 Oscillating terms 41

3 Absolute Bounds - Linear Form Sequences **45**

3.1 Introduction 45

3.2 Poisson's summation formula 47

3.3 Estimating the "tail" of the discrepancy function 51

3.3.1 Estimation for the sum when $|n_1|$ is large. 51

3.3.2 Estimation for the sum when $|n_1\alpha_i - n_{i+1}|$ is larger than $1/2$ for one of $1 \leq i \leq d$. 54

3.3.3 Estimation for the sum when $|n_1| \prod_{i=1}^{d} \|n_1\alpha_i\|$ is small. 56

3.3.4 Control the sum when n_1 is between $N^d/4$ and $N^d(\ln N)^d$. 60

3.4 Cancellation of the main terms. 62

3.5 Estimation of small exponentials. 65

4 Absolute Bounds - Irrational translations - Triangles **73**
4.1 Introduction 73
4.2 Poisson's summation formula 75
4.3 Local lemmas 79
4.4 Estimating the "tail" of the discrepancy function 82
4.4.1 Estimation for the sum when $|n_1|$ or $|n_2|$ is large 83
4.4.2 Estimation for the sum when $|\mathbf{n}\boldsymbol{\alpha} - n_3|$ is larger than $1/3$. 85
4.4.3 Estimation for the sum when $|n_1\tau - n_2| = \|n_1\tau\|$. 86
4.4.4 Estimation for the sum when $|n_1||n_1\tau - n_2|\|\mathbf{n}\boldsymbol{\alpha}\|$ is small. 88
4.4.5 Estimation for the sum when n_1 or n_2 is between $N^2/4$ and $N^2(\ln N)^2$. 88
4.5 Estimation of the exponential sums. 90

5 Ergodic Examples of Cylindrical Cascades **99**
5.1 Introduction 99
5.2 Arithmetic Notations 100
5.3 Main results and Lemmata 100
5.4 Proof of the Theorems using the Lemmata 104
5.5 Proofs of Lemmata 5.4 and 5.5 108

Chapter 1

Introduction

The main objects of research in this work are the asymptotic behaviors of ergodic discrepancies in counting problems on a d-dimensional torus. Given an ergodic transformation $T : X \to X$ on a measure space (X, Σ, μ), Birkhoff theorem states that for a μ-integrable function f, its time average starting from a generic point x, $\sum_{n=0}^{N-1} f(T^n x)/N$, converges to its space average $\int f d\mu/\mu(X)$. If we take the function f to be the characteristic function of a measurable set $\mathcal{C} \subset X$, we obtain a counting problem, and Birkhoff theorem states that the proportion of "visits" of the orbit of x inside the set $\mathcal{C}$ converges to its volume $\mathrm{Vol}(\mathcal{C})$.

In the classical setting of toral translation, the space X is chosen to be a d-dimensional torus $\mathbb{T}^d$ equipped with Haar measure, the ergodic transformation is the irrational translation $T_{\boldsymbol{\alpha}} : x \mapsto x + \boldsymbol{\alpha}$, where $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_d) \in \mathbb{R}^d$ and $1, \alpha_1, \ldots, \alpha_d$ are linearly independent over $\mathbb{Q}$; the map is in the sense of taking modulo 1 for each coordinate. Our object of interest is the ergodic discrepancy, defined as the difference between the number of actual visits in $\mathcal{C}$ and its expected value:

$$D_{\mathcal{C}}(x, \boldsymbol{\alpha}; N) = \sum_{n=0}^{N-1} \chi_{\mathcal{C}}(T_{\boldsymbol{\alpha}}^n(x)) - N\mathrm{Vol}(\mathcal{C}), \tag{1.1}$$

and two natural questions are the following:

Q1. Limit distributions. If we let the starting point x and the translation $T_{\boldsymbol{\alpha}}$ to be random, i.e. x, $\boldsymbol{\alpha} \in \mathbb{T}^d$ random, does there exist a limiting distribution, after adequate normalization, for the discrepancy function $D_{\mathcal{C}}$ as $N \to \infty$?

Q2. Absolute bounds. If we fix a transformation $T_{\boldsymbol{\alpha}}$, what is the optimal (absolute) upper bound for the discrepancy function $D_{\mathcal{C}}$ as $N \to \infty$?

These two questions form the axes of this work, and the following chapters are attempts to answer the above questions in different settings of the actions $T_{\boldsymbol{\alpha}}$ and the shapes of the set $\mathcal{C}$.

1.1 Limiting distributions

In the 1960s, Kesten[Kes60, Kes62] studied the limiting distribution problem in the case of 1-dimensional circle rotations. He proved that the discrepancy function, relative to an interval, converges to a Cauchy distribution after being normalized by $\ln N$. His proof relies heavily on the continued fraction algorithm, which makes it difficult for generalizations to higher dimensions. In the 2000s, Marklof[Mar07] adopted a dynamical approach to the distribution problem, showing that such normalized discrepancy functions are often intimately related to certain orbits under the geodesic and unipotent flows on homogeneous spaces.

Following the dynamical approach, Dolgopyat and Fayad extended Kesten's results to higher-dimensional tori when $d \geq 2$. Surprisingly, they found two distinct behaviors for $\mathcal{C}$ being a ball and $\mathcal{C}$ being a box[DF12, DF14], both of which are natural generalizations of the 1-d interval. They showed that d-dimensional boxes behave similarly to 1-dimensional intervals, and the discrepancy function also converges to a Cauchy distribution after being normalized by $(\ln N)^d$. As for balls, they showed that the normalizer should be $N^{(d-1)/2d}$, and that the discrepancy function converges to a distribution function defined over the product space of infinite tori and the space of $(d+1)$-dimensional lattice.

In both cases, the first ingredient of the proof is the harmonic analysis of the discrepancy's Fourier series, which highlights the main contribution of the discrepancy and helps find the good normalizer. The second ingredient is the Dani's correspondence principle that transfers the study of small divisors which appears in the main contribution to the study of dynamics of diagonal and unipotent actions on the homogeneous space of lattices. In the case of boxes, the main contribution comes from all the Fourier modes whose norms are *less* than order N, which corresponds to the cusp excursions *up* to time $\ln N$ of the diagonal actions that code the small divisors, and this leads to a Cauchy distribution similar to the 1-dimensional case. In the case of balls, the main contribution comes only from the Fourier modes whose norms are *of* order N, which corresponds to the cusp excursions at *only* the time scale $\ln N$ of the diagonal actions that code the small divisors, and this leads to a limiting distribution that can be calculated as the level sets of the transformed discrepancy function expressed in the language of lattices.

In Chapter 2, following a similar approach as Dolgopyat and Fayad, we consider the $\mathbb{Z}^2$ action sequence $\{(n_1\alpha_1, n_2\alpha_2) \in \mathbb{T}^2\}_{1\leq n_1,n_2\leq N}$ relative to a centered disc $\mathcal{C}$ in $\mathbb{T}^2$, which corresponds to the positioning of the translated lattice $\mathbb{Z}\alpha_1 \oplus \mathbb{Z}\alpha_2 + x$ relative to the lattice $\mathbb{Z}^2$ in the Euclidean norm in $\mathbb{R}^2$, recall that x is the starting point of the action. Specifically, the discrepancy function measures how often the two lattices becomes $\mathcal{C}$-close to each other

in a growing rectangle, it is defined by:

$$D_{\mathcal{C}}(r,x,\boldsymbol{\alpha};N)=\sum_{\substack{0\le n_1\le N-1\\0\le n_2\le N-1}}\chi_{\mathcal{C}_r}(x_1+n_1\alpha_1,x_2+n_2\alpha_2)-N^2\mathrm{Vol}(\mathcal{C}_r). \tag{1.2}$$

The parameter r is the rescale of $\mathcal{C}$ to suppress possible irregular dependence of the limit distribution on $\mathcal{C}$. We will show that there are two very different behaviors for different parts in the Fourier series of $D_{\mathcal{C}}$. After choosing suitable normalizers, the part *with* 0-coordinates behaves as the ergodic sum of smooth observables, and the part *without* 0-coordinates behaves similar to the ergodic deviations of toral translations for convex bodies.

Theorem 1. *Let $\mathcal{C}$ be a symmetric (with respect to both axes), strictly convex analytic body that fits inside the unit cube of $\mathbb{R}^2$, and $D_{\mathcal{C}}$ defined as in* (1.2)*, we can decompose $D_{\mathcal{C}}$ into two parts, noted by $D_{\mathcal{C},1}$ and $D_{\mathcal{C},2}$, then we have*

(a) For $D_{\mathcal{C},1}$, let $(x,\boldsymbol{\alpha})$ be uniformly distributed in $\mathbb{T}^2\times\mathbb{T}^2$, then for every fixed r, there exists a limiting disribution for $D_{\mathcal{C},1}/N$, as $N\to\infty$.

(b) For $D_{\mathcal{C},2}$, Let $(r,x,\boldsymbol{\alpha})$ be uniformly distributed in $[a,b]\times\mathbb{T}^2\times\mathbb{T}^2$, then there exists a limiting distribution for $D_{\mathcal{C},2}/(r^{1/2}N^{1/2})$ as $N\to\infty$.

This result shows that the distribution of the discrepancy function for the $\mathbb{Z}^2$-action on $\mathbb{T}^2$ relative to a convex set can be understood similarly as a Taylor expansion. Geometrically, the main part with the normalizer N represents the resonance that naturally occurs since the two lattices $\mathbb{Z}\alpha_1\oplus\mathbb{Z}\alpha_2$ and $\mathbb{Z}^2$ are parallel. The remaining part with normalizer $N^{\frac{1}{2}}$ represents the residue error, i.e., the interaction between the smooth convex boundary of $\mathcal{C}$ and the $\mathbb{Z}^2$ action.

Although the result is stated in 2-dimension, by choosing suitable normalizers for different parts of the Fourier series, similar results can be obtained for *even* dimensions in a straightforward way. While for odd dimensions, the "middle sum" of the Fourier modes composed of $\frac{d+1}{2}$ zero-coordinates has fundamentally different behaviors, and we do not yet have a complete description for it.

In the general case of $\mathbb{Z}^2$-actions, where we consider the lattice $\mathbb{Z}\boldsymbol{\alpha}_1\oplus\mathbb{Z}\boldsymbol{\alpha}_2$ for 2-dimensional vectors $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$, the resonance part with the normalizer N does not appear, and we believe that a limit law similar to the residue part in the parallel case holds in the general $\mathbb{Z}^2$-action case with a normalizer $N^{\frac{1}{2}}$, although for the moment, we face technical difficulties caused by the extra factor $\ln N$ that arises from the measure computation for 2-dimensional cusps.

1.2 Absolute bounds for toral translations

Let $\mathcal{B}$ denote the set of all straight boxes in $\mathbb{T}^d$, then by taking the absolute upper bound of the discrepancy function (1.1) for $\mathcal{C}\in\mathcal{B}$ and all possible starting points $x\in\mathbb{T}^d$, we

arrive at the following maximal discrepancy function:

$$\Delta(\boldsymbol{\alpha}; N) = \max_{\substack{\mathcal{C}\in\mathcal{B},\\ x\in\mathbb{T}^d}} \left| \sum_{1\le n\le N} \chi_{\mathcal{C}}(x+n\boldsymbol{\alpha}) - n\mathrm{Vol}(\mathcal{C}) \right|. \tag{1.3}$$

In 1923, Khintchine[Khi23] proved the following precise bounds for $d=1$:

$$\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} < \infty \Longleftrightarrow \frac{\Delta(\boldsymbol{\alpha}; N)}{(\ln N)^d \varphi(\ln\ln N)} \text{ is bounded for almost every } \boldsymbol{\alpha}\in\mathbb{T}^d, \tag{1.4}$$

where $\varphi(n)$ is an arbitrary positive increasing function of n. The major tool in his proof is the metrical theory of irrationals which mainly relies on continued fractions. In 1964, Schmidt[Sch64] obtained an upper bound of $(\ln N)^{d+1+\epsilon}$ for $d>1$ by using the Erdős–Turán-Koksma inequality. But due to the absence of a continued fraction algorithm in higher dimensions, the case of $d>1$ was not completely solved until 1994 by Beck[Bec94]. Beck's result extended Khintchine's precise zero-one law in (1.4) to the case of straight boxes in higher dimensions. Beck observed that, similar to the 1-dimensional case, the Fourier series of the maximal discrepancy $\Delta(\boldsymbol{\alpha}; N)$ is in fact a sum of almost *pairwise independent* terms of $\boldsymbol{\alpha}$, and the study of $\Delta(\boldsymbol{\alpha}; N)$ reduced to the analysis of the small divisors of linear forms $\|n_1\alpha_1+\cdots+n_k\alpha_k\|$. By employing "the second-moment method"(See [Sch60, Sch64]), Beck showed that the contintued fraction algorithm used in the 1-d case can be effectively substituted by precise estimations of the cardinality of Fourier modes for which the divisors are restricted to small intervals. The improvement of the critical factor of $\ln N$ is achieved by the cancellations of the positive parts with the negative parts for the terms with nearly identical divisors.

In Chapter 3, going into a different direction in dimension 1, instead of irrational rotations, we can consider the $\boldsymbol{\alpha}$-related d-linear form sequence $\left\{\sum_{1\le i\le d} k_i\alpha_i\right\}_{\substack{1\le k_i\le N\\ 1\le i\le d}}$, whose discrepancy function relative to an interval $[0,x)$ is defined as

$$D(\boldsymbol{\alpha}, x; N) = \sum_{\substack{1\le k_i\le N\\ 1\le i\le d}} \chi_{[0,x)}\left(\sum_{1\le i\le d} k_i\alpha_i\right) - N^d x,$$

and the maximal discrepancy function is defined similarly:

$$\Delta(\boldsymbol{\alpha}; N) = \max_{0<x\le 1} |D(\boldsymbol{\alpha}, x; N)|.$$

By generalizing Beck's ideas to the small divisors $n\prod_{i=1}^d \|n\alpha_i\|$, we obtain an *almost optimal* upper bound for the maximal discrepancy of linear forms:

Theorem 2. *For any $\epsilon > 0$, there exist a full measure set of $\boldsymbol{\alpha} \in \mathbb{R}^d$, such that*

$$\Delta(\boldsymbol{\alpha}; N) \leq C(\boldsymbol{\alpha}, \epsilon)(\ln N)^d \cdot (\ln \ln N)^{\max\{3,d\}+\epsilon}, \tag{1.5}$$

for some constant $C(\boldsymbol{\alpha}, \epsilon) > 0$, and

$$\Delta(\boldsymbol{\alpha}; N) > (\ln N)^d \cdot \ln \ln N \tag{1.6}$$

for infinitely many $N \in \mathbb{N}$.

Compared with toral translations, the additional factor $(\ln \ln N)^{\max\{d-1,2\}}$ in the case of $d-$linear forms in our proof is needed to control the small divisors $n \prod_{i=1}^{d} \|n\alpha_i\|$. An interesting observation is that in the proof, there exists a *duality* between the case of toral translations relative to d-dimensional straight boxes and the case of d-linear forms relative to 1-dimensional intervals. The small divisors arising in the case of toral translations are d-linear forms, and the small divisors in the case of linear forms are products of coordinates of toral translations. By using this duality, we could simplify the proof for one case by applying the Erdős–Turán-Koksma inequality to obtain a weaker version of the result of its dual case, which achieves a faster estimation of the cardinality of dyadic sets for the target case, which is an important step in both proofs.

In Chapter 4, following Beck's work on straight boxes, we consider toral translations relative to right triangles with straight legs (parallel to the sides of the torus) in $\mathbb{T}^2$. By replacing $\mathcal{B}$ in (1.3) with the set of straight right triangles with hypotenuse of slope τ, we can define similarly the corresponding maximal discrepancy function $\Delta(\boldsymbol{\alpha}, \tau; N)$ for straight right triangles. Here the main difficulty comes from the hypotenuse of the right triangle. The inclination τ shifts the divisors in the Fourier series of the discrepancy function from $n_1 n_2(n_1\alpha_1 + n_2\alpha_2 - n_3)$ to $n_1(n_1\tau - n_2)(n_1\alpha + n_2\alpha_2 - n_3)$, which adds an additional possibility for the factor $|n_1\tau - n_2|$ to be small. By generalizing Beck's local criterions to the new divisors, we obtain an *almost optimal* upper bound and the same lower bound for the maximal discrepancies in the case of triangles:

Theorem 3. *Fix any $\epsilon > 0$, for almost every $\boldsymbol{\alpha} \in \mathbb{R}^2$ and almost every $\tau \in \mathbb{R}^+$, the maximal discrepancy $\Delta(\boldsymbol{\alpha}, \tau; N)$ is essentially bounded from above by $(\ln N)^2(\ln \ln N)^{2+\epsilon}$, but not by $(\ln N)^2 \ln \ln N$. Specifically, there exits $C(\boldsymbol{\alpha}, \tau, \epsilon) > 0$, such that*

$$\Delta(\boldsymbol{\alpha}, \tau; N) \leq C(\boldsymbol{\alpha}, \tau, \epsilon)(\ln N)^2 \cdot (\ln \ln N)^{2+\epsilon}, \tag{1.7}$$

and there exist infinitely many $N \in \mathbb{N}$, such that

$$\Delta(\boldsymbol{\alpha}, \tau; N) > (\ln N)^2 \cdot \ln \ln N. \tag{1.8}$$

It is easy to see that the same bounds hold for almost every polygon in $\mathbb{T}^2$, simply

by cutting the polygons into right triangles. Unfortunately, the statement does not give a zero-one law as the results of Khintchine and Beck, the additional factor $\ln\ln N$ is needed to control the small factor $n_1\tau - n_2$ that arise from the hypotenuse of the triangle, which does not occur in the case of straight rectangles. Note that when the slope τ or the translation vector $\boldsymbol{\alpha}$ is Liouville, the divisors $n_1(n_1\tau - n_2)(n_1\alpha + n_2\alpha_2 - n_3)$ could be very small, which easily breaks the logarithmic upper bound in the diophantine case. Therefore, the restriction to "almost every τ and almost every $\boldsymbol{\alpha}$" is necessary. Higher-dimensional polyhedra are expected to follow the same estimations but more involved calculations are needed.

The case of discs is significantly more difficult and seems to be a dual problem to the Gauss circle problem. If we could obtain a *uniform* upper bound for triangles of a specific class of slopes, we can approximate the disc by triangles and obtain the upper bound $N^{\frac{1}{3}+\epsilon}$, which corresponds to the classical upper bound $R^{\frac{2}{3}+\epsilon}$ of the Gauss circle problem. While the optimal bound for the error term in the Gauss circle problem is conjectured to be between $R^{\frac{1}{2}}$ and $R^{\frac{1}{2}+\epsilon}$, the optimal bound in our case, toral translations relative to balls, is conjectured to be between $N^{\frac{1}{4}}$ and $N^{\frac{1}{4}+\epsilon}$, since the result of Dolgopyat and Fayad[DF14] already shows that the *average* magnitude of the discrepancy function is $N^{\frac{1}{4}}$.

1.3 Cylindrical cascades

One interesting setting in which discrepancy functions naturally appear is the study of cylindrical cascades. Given a function $A : \mathbb{T}^d \to \mathbb{R}^r$, a cylindrical cascade above the toral translation $T_{\boldsymbol{\alpha}}$ is defined as $W_{\boldsymbol{\alpha},A}$: $\mathbb{T}^d \times \mathbb{R}^r \to \mathbb{T}^d \times \mathbb{R}^r$, given by $W_{\boldsymbol{\alpha},A}(x,y) = (x+\boldsymbol{\alpha}, y + A(x))$. When $\boldsymbol{\alpha}$ is Diophantine and A is smooth, the linear cohomological equation $A(x) - \int_{\mathbb{T}^d} A(u)du = -B(x+\boldsymbol{\alpha}) + B(x)$ has a smooth solution B, thus $W_{\boldsymbol{\alpha},A}$ is smoothly conjugated to the translation $W_{\boldsymbol{\alpha},\int_{\mathbb{T}^d} A}$. An interesting case for A non-smooth is when the components of A are zero-mean indicator functions of typical shapes in $\mathbb{T}^d$, that is, $A = (A_1, \ldots, A_r)$, where $A_i = \chi_{\mathcal{C}_i} - \mathrm{Vol}(\mathcal{C}_i)$, and each $\mathcal{C}_i$ is a typical shape such as a box or a polytope. Here the Birkhoff sums of A_i are precisely the ergodic deviations that we previous discussed.

The cylindrical cascade $W_{\boldsymbol{\alpha},A}$ can be seen as a deterministic random walk in $\mathbb{R}^r$ driven by the toral translation on the base $\mathbb{T}^d$, which makes it natural to study its recurrence and ergodicity. For $r = 1$, the *zero mean* condition $\int_{\mathbb{T}^d} A = 0$ is both necessary and sufficient for the walk $W_{\boldsymbol{\alpha},A}$ to be recurrent when α is irrational[Atk76]. For $d = 1$, we know that when A is of zero-mean with bounded variation, the walk is recurrent due to the *Denjoy-Koksma inequality.* We refer to [DF15], Section 8 for a more complete introduction for the results in 1-d.

In higher dimensions, only recurrent and transient examples are known when $d = 2$ and $\mathcal{C}_i$ are polytopes[CC09]. Chevallier and Conze[CC09] showed that for the walk of any

base space to be recurrent, a sufficient condition is that the Birkhoff sums of A_i grow more slowly than $\mathcal{O}(n^{1/r})$ along some subsequence N_n for most starting points. By using the L^2-estimation of the ergodic deviations of A, they showed that for almost every translation vector α, the walk $W_{\boldsymbol{\alpha},A}$ is recurrent for polytopes. In fact, we can see that from our absolute bounds of the discrepancy function for triangles, the growth of Birkhoff sums of A_i at time N are at most of order $(\ln N)^{2+\epsilon}$, which also leads to recurrence by the criterion of Chevallier and Conze.

The ergodicity of the walk $W_{\alpha,A}$ can be established if the Birkhoff sums of A become increasingly well distributed over $\mathbb{R}^r$ along some subsequence N_n, such that the translation $T_\alpha^{N_n}$ stays close to identity. This idea gives rise to the notion of *essential values*[Sch77], which is the major tool for our construction of ergodic walks for rectangles. In Chapter 5, as a joint work with Fatna Abdedou, we build ergodic examples of the cylindrical cascades on $\mathbb{T}^2 \times \mathbb{R}$ by using Liouville translation vectors. Specifically, we have the following:

Theorem 4 (Abdedou-W). *For almost every straight rectangle in $\mathbb{T}^2$ with a corner at the origin, let A be the centered indicator function of the rectangle, then there exists a G_δ-dense set of $\boldsymbol{\alpha}$, such that the cylindrical cascades $W_{\boldsymbol{\alpha},A}$ is ergodic on $\mathbb{T}^2 \times \mathbb{R}$.*

The statement is for a G_δ-dense set of $\boldsymbol{\alpha}$ since Liouville conditions need to be imposed on the translation vector to achieve ergodicity in our construction. Higher-dimensional counterparts in $\mathbb{T}^d \times \mathbb{R}^r$ can be obtained similarly by the subgroup property of the essential values, but again our method requires that the lengths of additional sides of the higher-dimensional boxes to be Liouville when $d > 2$ or $r > 1$. Ergodic examples using polytopes in $\mathbb{T}^d$, $d \geq 2$, are more delicate due to their inclined sides, and their constructions are still in progress. Although Liouville conditions seem to be unavoidable with our current techniques, due to the slow divergence rate of the Birkhoff sum of A, i.e., the logarithmic absolute upper bound for the maximal discrepancies in the case of polygons in Theorem 3, ergodicity of the walk is expected for typical polytopes and typical translation vectors.

Chapter 2

Multidimensional Degenerate Actions - Convex Bodies

2.1 Introduction

In a d-dimensional torus, given a translation vector $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$, we can consider the dynamical system $(\mathbb{T}^d, T_\alpha, \mu)$, where μ is the Haar measure on $\mathbb{T}^d$, and T_α is the translation from $\mathbb{T}^d \to \mathbb{T}^d$ defined by $T(x) = x + \alpha$, in the sense of modulo 1 for each coordinate. A translation T_α is called irrational if the real numbers $1, \alpha_1, \ldots, \alpha_d$ are linearly independent over $\mathbb{Q}$. Ergodic theory states that, for every irrational translation, and for almost every starting point x, the number of visits of the orbit $\{T_\alpha^n(x) : n \geq 1\}$ to a measurable set $\mathcal{C}$ before time N has a ratio converging to the measure of the set $\mathrm{Vol}(\mathcal{C})$, as $N \to \infty$. One object of interest is the discrepancy function, defined as the difference of the actual number of visits in $\mathcal{C}$ before time N and the expected visits $N\mathrm{Vol}(\mathcal{C})$. In dimension 1, Kesten [Kes60, Kes62] proved that the discrepancy function for the circle rotation relative to an interval, after being normalized by $\ln N$, converges to a Cauchy distribution.

There are different ways to extend this result to higher dimensions, one way is to study the random toral translation relative to higher dimensional counterparts of the interval, such as boxes and balls (analytic convex bodies), both of which were studied by Dolgopyat and Fayad in [DF14, DF12]. They showed that d-dimensional boxes behave similarly to 1-dimensional intervals, i.e., the discrepancy function, normalized by $(\ln N)^d$, also converges to a Cauchy distribution. They showed that for balls, the discrepancy function, normalized by $N^{(d-1)/2d}$, converges to a distribution function defined over the product space of infinite tori and the homogeneous space $\mathrm{SL}(d+1, \mathbb{R})/\mathrm{SL}(d+1, \mathbb{Z})$. Their proof consists of a combination of harmonic analysis of the discrepancy's Fourier series probability, and an important ingredient is the equidistribution of locally unstable submanifolds over the whole space of unimodular lattices.

In this chapter, we follow a similar approach as Dolgopyat and Fayad, but instead of translations, we will consider the degenerate $\mathbb{Z}^2$ action in dimension $d = 2$ (see the definition below) relative to strictly convex, symmetric, and analytic bodies. Given a convex body $\mathcal{C}$, we denote $\mathcal{C}_r$ the rescaled bodies with ratio $r > 0$ by the homothety centered at the origin, where $r < r_0$ so that $\mathcal{C}_r$ can fit inside the unit cube of $\mathbb{R}^2$. Let $\alpha = (\alpha_1, \alpha_2) \in \mathbb{T}^2$ be the action vector, we consider the following discrepancy function:

$$D_{\mathcal{C}}(r, x, \alpha; N) = \sum_{\substack{0 \le n_1 \le N-1 \\ 0 \le n_2 \le N-1}} \chi_{\mathcal{C}_r}(x_1 + n_1\alpha_1, x_2 + n_2\alpha_2) - N^2 \mathrm{Vol}(\mathcal{C}_r) \tag{2.1}$$

where $\chi_{\mathcal{C}_r}$ is the indicator function of the set $\mathcal{C}_r$.

We will show that by decomposing the discrepancy function into 2 components, each component would admit a limiting distribution after a suitable normalization, specifically:

$$D_{\mathcal{C}}(r, x, \alpha; N) = \sum_{\bar{d}=1}^{2} D_{\mathcal{C},\bar{d}}(r, x, \alpha; N) \tag{2.2}$$

where $D_{\mathcal{C},\bar{d}}$ represents the part of the Fourier series of $D_{\mathcal{C}}$ with coefficients of $\bar{d}$ non-zero coordinate(s), whose definitions will be clearer after we introduce the Fourier series of $D_{\mathcal{C}}$ in Section 3.

Our main result is the following:

Theorem 2.1. *Let $\mathcal{C}$ be a symmetric, strictly convex analytic body that fits inside the unit cube of $\mathbb{R}^2$, and $D_{\mathcal{C}}$, $D_{\mathcal{C},\bar{d}}$ defined as in (2.1) and (2.2). For $\bar{d} = 1, 2$, there exists a limiting distribution for each $D_{\mathcal{C},\bar{d}}(r, x, \alpha)$ after a suitable normalization, specifically, we have 2 distinct cases:*

(a) For $\bar{d} = 1$, assume that (x, α) are uniformly distributed in $\mathbb{T}^2 \times \mathbb{T}^2$, then for every fixed r, the random variable $D_{\mathcal{C},1}(r, x, \alpha; N)/N$ converges in distribution as $N \to \infty$.

(b) For $\bar{d} = 2$, for any $b > a > 0$, assume that (r, x, α) are uniformly distributed in $X = [a, b] \times \mathbb{T}^2 \times \mathbb{T}^2$, the random variable $D_{\mathcal{C},2}(r, x, \alpha; N)/(r^{\frac{1}{2}} N^{\frac{1}{2}})$ converges in distribution as $N \to \infty$, and the limit distribution is independent of the interval $[a, b]$.

The explicit forms of $\mathcal{D}_{\mathcal{C},\bar{d}}$ will be given in Proposition 2.1 and Proposition 2.2 of Section 2.

Remark. *In part (a) of Theorem 2.1, $D_{\mathcal{C},1}/N$ has similar behaviors to the ergodic deviation of a smooth function, in which case r does not need to be random to achieve a limiting distribution. In part (b), $D_{\mathcal{C},2}$ is similar to the case of irrational traslations relative to balls in* [DF14, DF12], *where r* needs *to be random in order to achieve asymptotic independence between variables in Sections 6.*

Remark. *Though we only treat the special case $d = 2$, it will be clear from the proof that for higher dimensions $d > 2$, $D_{\mathcal{C},\bar{d}<\frac{d+1}{2}}$ can be treated in the same way as $D_{\mathcal{C},1}$ in our case,*

and $D_{\mathcal{C},\bar{d}>\frac{d+1}{2}}$ can be treated in the same way as $D_{\mathcal{C},2}$ in our case. For odd dimensions, the sum $D_{\mathcal{C},\frac{d+1}{2}}$ exhibits distinctly different behavior and unfortunately could not be dealt with by using the same method as in this paper.

Plan of the Chapter. This chapter is organized as follows: In Section 2 we will present explicit expressions for the limiting distributions in Theorem 2.1. In Section 3 we introduce the Fourier series of the discrepancy function, and we give a proof of Theorem 2.1 (a) by showing that $D_{\mathcal{C},1}/N$ is, in fact, a coboundary. The three remaining sections are devoted to the proof of Theorem 2.1 (b). In Section 4, we show that the main contribution of the sum $D_{\mathcal{C},2}$ comes from the "averagely resonant" Fourier modes, similar to Section 3 in [DF14]. In Section 5, we use the Dani correspondence principle to relate the main contributing terms to the dynamics on the product spaces of lattices in $\mathbb{R}^2$. In Section 6, we show that the new variables, introduced in Section 5, become asymptotically independent, and therefore proving the existence of a limiting distribution in Theorem 2.1 (b).

2.2 Limiting distributions

2.2.1 Limiting Distribution for the case $\bar{d} = 1$

Proposition 2.1. *Let $\mathcal{C}$ be an analytic, symmetric (with respect to both axes), strictly convex body in $\mathbb{R}^2$, assume that (x, α) are uniformly distributed in $\mathbb{T}^2 \times \mathbb{T}^2$, then for every fixed r, there exists a function $\mathcal{D}_{\mathcal{C},1,r}(x, \alpha, \beta) : (\mathbb{T}^2)^3 \to \mathbb{R}$, such that as $N \to \infty$,*

$$D_{\mathcal{C},1}(r, x, \alpha; N)/N \Rightarrow \mathcal{D}_{\mathcal{C},1,r}(x, \alpha, \beta)$$

in distribution, where (x, α, β) is uniformly distributed on $(\mathbb{T}^2)^3$. The limiting function has the following form:

$$\mathcal{D}_{\mathcal{C},1,r}(x, \alpha, \beta) = B_{\mathcal{C}_r}(\alpha, \beta) - B_{\mathcal{C}_r}(\alpha, x),$$

where

$$B_{\mathcal{C}_r}(\alpha, x) = \sum_{k \neq 0} \frac{a_k(r)}{e^{2\pi i(k,x)} - 1} e^{2\pi i(k,x)},$$

and $a_k(r) = 0$ when $k_1 k_2 \neq 0$ and $a_k(r) = \hat{\chi}_{\mathcal{C}_r}(k)$ when $k_1 k_2 = 0$. Here $\hat{\chi}_{\mathcal{C}_r}(k) = \int_{\mathbb{R}^2} \chi_{\mathcal{C}_r}(u) e^{-2\pi i(k,u)} du$ represents the kth Fourier coefficient of $\chi_{\mathcal{C}_r}$. We will prove in Lemma 2.1 that for almost every $\alpha \in \mathbb{T}^2$, the series $B_{\mathcal{C}_r}(\alpha, x)$ converges in $L^2(\mathbb{T}^2)$ with respect to the x-variable.

2.2.2 Limiting Distribution for the case $\bar{d} = 2$

Notations. Let $M = SL(2, \mathbb{R})/SL(2, \mathbb{Z})$ denote the space of 2-dimensional unimodular lattices of $\mathbb{R}^2$. $M^2 = \prod_{2 \text{ copies}} M$. Given $L = (L_1, L_2) \in M^2$ we denote by $e_1(L_i)$ the shortest non-zero vector in L_i that lies in $\widetilde{\mathbb{R}^2} := \{v = (v_1, v_2) \in \mathbb{R}^2 : v_1 > 0$ or $[v_1 = 0$ and $v_2 >$

0]}, and $e_2(L_i)$ the shortest vector in $\widetilde{\mathbb{R}^2} \cap L_i$ among those which have the shortest nonzero projection on the orthocomplement of the line generated by $e_1(L_i)$. Clearly the vectors $e_1(L_i), e_2(L_i)$ are well defined outside a set of Haar measure 0. In fact, these vectors generate the lattice (see [Arn13]). We denote $e(L_i) = (e_1(L_i), e_2(L_i))$.

Let $\mathcal{Z}$ be the set of prime vectors $m \in \mathbb{Z}^2$ (i.e. the coordinates are coprime) with positive first nonzero coordinate, and let $\mathcal{Z}'$ be the subset of $\mathcal{Z}$ such that both coordinates are strictly positive, i.e.,

$$\mathcal{Z}' := \{p = (p_1, p_2) \in \mathcal{Z} \ : \ p_1 > 0 \text{ and } p_2 > 0\}.$$

For later usage in Section 4 and 5, we define $\mathcal{Z}^2 = \{\mathbf{m} = (m^1, m^2), m^i \in \mathcal{Z}\}$, and let

$$T_2^\infty = (\mathbb{T}^2)^2 \times \mathbb{T}^{\mathcal{Z}' \times \mathcal{Z}^2}$$

We denote elements of T_2^∞ by $(\boldsymbol{\theta}, \mathbf{b})$, where $\boldsymbol{\theta} = (\theta^1, \theta^2)$, $\theta^i \in \mathbb{T}^2$, and $\mathbf{b} = (b_{p,\mathbf{m}})_{(p,\mathbf{m}) \in \mathcal{Z}' \times \mathcal{Z}^2}$. For $\mathbf{m} = (m^1, m^2) \in \mathcal{Z}^2$ and $L = (L_1, L_2) \in M^2$, we denote by $(X_{m^i}, Z_{m^i}) = (m^i, e(L_i))$ the vector $m_1^i e_1(L_i) + m_2^i e_2(L_i)$. Given a prime vector $p = (p_1, p_2) \in \mathcal{Z}'$, we denote $X_{p,\mathbf{m}} = (p_1 X_{m^1}, p_2 X_{m^2})$ and $R_{p,\mathbf{m}} = \|X_{p,\mathbf{m}}\|$ the Euclidean norm of $X_{p,\mathbf{m}}$.

Limiting distribution. Let $\mathcal{C}$ be a strictly convex body with smooth boundary. For each vector $\xi \in \mathbb{S}^1$, we denote by $K(\xi) = K_{\mathcal{C}}(\xi)$ the standard curvature of the plane curve ∂C at the unique point $x(\xi) \in \partial C$ where the unit outer normal vector is ξ.

Denote

$$\mathcal{M}_2 = M^2 \times T_2^\infty,$$

and let μ be the Haar measure on $\mathcal{M}_2$. Define the following function on $\mathcal{M}_2$

$$\begin{aligned} \mathcal{L}_{\mathcal{C}}(L, \boldsymbol{\theta}, \mathbf{b}) = \frac{2}{\pi^3} \sum_{\check{p} \in \mathbb{Z} \setminus \{0\}}^{\infty} \sum_{p \in \mathcal{Z}'} \sum_{\mathbf{m} \in \mathcal{Z}^2} K^{-\frac{1}{2}} \left(\frac{X_{p,\mathbf{m}}}{R_{p,\mathbf{m}}} \right) \\ \times \frac{\sin\left(2\pi\left(\check{p} b_{p,\mathbf{m}} - \frac{1}{8}\right)\right)}{|\check{p}|^{\frac{7}{2}} R_{p,\mathbf{m}}^{\frac{3}{2}}} \prod_{i=1}^{2} \frac{\cos\left(2\pi \check{p} p_i \left(m^i, \theta^i\right)\right) \sin\left(\pi \check{p} p_i Z_{m^i}\right)}{p_i Z_{m^i}} \end{aligned} . \tag{2.3}$$

The above series converges almost surely, and the proof is essentially the same as Proposition 8.1 in [DF14]. The key observation is that the random variables $b_{p,\mathbf{m}}$ are independent, so $\xi_{p,\mathbf{m}}$ defined below

$$\xi_{p,\mathbf{m}} = \sum_{\check{p} \in \mathbb{Z} \setminus \{0\}} K^{-\frac{1}{2}} \left(\frac{X_{p,\mathbf{m}}}{R_{p,\mathbf{m}}} \right) \cdot \frac{\sin\left(2\pi\left(\check{p} b_{p,\mathbf{m}} - \frac{1}{8}\right)\right)}{|\check{p}|^{\frac{7}{2}} R_{p,\mathbf{m}}^{\frac{3}{2}}} \prod_{i=1}^{2} \frac{\cos\left(2\pi \check{p} p_i \left(m^i, \theta^i\right)\right) \sin\left(\pi \check{p} p_i Z_{m^i}\right)}{p_i Z_{m^i}}$$

are also independent, hence we could use Kolmogorov's three series theorem and transfer

our study to the convergence of the variances:

$$\sum_{p\in\mathcal{Z}'}\sum_{\mathbf{m}\in\mathcal{Z}^2}\mathrm{Var}(\xi_{p,\mathbf{m}}).$$

By taking $d = 2$ in the proof of Proposition 8.1 of [DF14], and replacing the original summation by the summation over $(p, \mathbf{m})$, in which case the term $R^3_{p,\mathbf{m}}$ in the denominator (of the variance) can be written as the product of two copies of $R^{\frac{3}{2}}_{p,\mathbf{m}}$, and each one can be used to control the summation over m^i, $i = 1, 2$, since the summation is in 2 dimension for each m^i, any power strictly greater than 1 suffices.

The distribution $\mathcal{D}_{\mathcal{C},2}$ of Theorem 2.1 can now be described by the level set of the function above:

Proposition 2.2. *If $\mathcal{C}$ is an analytic, symmetric, strictly convex body in $\mathbb{R}^2$, for any $b > a > 0$, assume that (r, x, α) are uniformly distributed in $X = [a, b] \times \mathbb{T}^2 \times \mathbb{T}^2$, and denote λ the normalized, Lebesgue measure on X, then there exists a distribution function $\mathcal{D}_{\mathcal{C},2}(z) : \mathbb{R} \to [0, 1]$, we have*

$$\lim_{N\to\infty}\lambda\{(r,x,\alpha)\in[a,b]\times\mathbb{T}^2\times\mathbb{T}^2|\frac{D_{\mathcal{C},2}(r,x,\alpha;N)}{r^{\frac{1}{2}}N^{\frac{1}{2}}}\le z\}=\mathcal{D}_{\mathcal{C},2}(z). \tag{2.4}$$

where $\mathcal{D}_{\mathcal{C},2}(z)$ is given by:

$$\mathcal{D}_{\mathcal{C},2}(z)=\mu\{(L,\boldsymbol{\theta},\mathbf{b})\in\mathcal{M}_2:\mathcal{L}_{\mathcal{C}}(L,\boldsymbol{\theta},\mathbf{b})\le z\},\ \forall z\in\mathbb{R}.$$

Remark. *Note that $\mathcal{C}$ only affects the term $K^{-1/2}$ in the definition of (2.3), if we scale $\mathcal{C}$ by ratio r, we have $K_{\mathcal{C}_r}(\xi) = r^{-1}K_{\mathcal{C}}(\xi)$, so immediately we have the following rescaling relation for $\mathcal{L}_{\mathcal{C}}$:*

$$\mathcal{L}_{\mathcal{C}_r}(L,\boldsymbol{\theta},\mathbf{b})=\sqrt{r}\cdot\mathcal{L}_{\mathcal{C}}(L,\boldsymbol{\theta},\mathbf{b})$$

for any $r > 0$; hence in Proposition 2.4 we have $\mathcal{D}_{\mathcal{C}_r,2}(z) = \mathcal{D}_{\mathcal{C},2}(z/\sqrt{r})$, which is consistent with the factor $\sqrt{r}$ in the normalizer when we apply Theorem 2.1(b) to both $\mathcal{C}$ and $\mathcal{C}_r$ with the intervals $[a, b]$ and $[a/r, b/r]$ respectively, and using the relation $D_{\mathcal{C}_r,2}(r_1, x, \alpha; N) = D_{\mathcal{C},2}(rr_1, x, \alpha; N)$.

2.3 Fourier series and proof of Proposition 2.1

In this section, we first introduce the Fourier series of the discrepancy function, and then we show that the first part of the discrepancy function, normalized by N, is a coboundary, thus proving the existence of the limiting distribution of $D_{\mathcal{C},1}$.

In following sections, $\epsilon > 0$ is fixed and can be arbitrarily small. The constants C may vary between inequalities but it does not depend on any variables other than the dimension

d, which is fixed to 2 in our case.

2.3.1 Fourier series for convex bodies.

We introduce the Fourier series for indicator function of the smooth strictly convex body $\mathcal{C}$ by using the asymptotic formula obtained in [Her62]. For each vector $t \in \mathbb{R}^2$, define its maximal projection on $\partial\mathcal{C}$ by $P(t) = \sup_{x\in\partial\mathcal{C}}(t,x)$, if $\mathcal{C}$ is of class $C^{\frac{9}{2}}$, then we have the following formula for the Fourier mode:

$$(2\pi i|t|)\hat{\chi}_{\mathcal{C}}(t) = \rho(\mathcal{C},t) - \bar{\rho}(\mathcal{C},-t) \tag{2.5}$$

with

$$\rho(\mathcal{C},t) = |t|^{-\frac{1}{2}}K^{-\frac{1}{2}}(t/|t|)e^{i2\pi(P(t)-\frac{1}{8})} + \mathcal{O}(|t|^{-\frac{3}{2}}).$$

By a change of variable, we have $\hat{\chi}_{\mathcal{C}_r}(k) = r^2\hat{\chi}_{\mathcal{C}}(rk)$, and by grouping the corresponding positive and negative terms in the Fourier series we get that for a symmetric body:

$$\chi_{\mathcal{C}_r}(x) - Vol(\mathcal{C}_r) = r^{\frac{1}{2}} \sum_{k\in\mathbb{Z}^2-\{0\}} c_k(r)\cos(2\pi(k,x)), \tag{2.6}$$

$$c_k(r) = d_k(r) + \mathcal{O}(|k|^{-\frac{5}{2}}),$$

$$d_k(r) = \frac{1}{\pi}\frac{g(k,r)}{|k|^{\frac{3}{2}}},$$

$$g(k,r) = K^{-\frac{1}{2}}(k/|k|)\sin(2\pi(rP(k)) - \frac{1}{8}).$$

2.3.2 Proof for the limiting distribution when $\bar{d} = 1$

We recall that $D_{\mathcal{C},1}$ is the part of the Fourier series that consists of the following modes

$$k = (k_1, 0) \text{ and } (0, k_2), \text{ where } k_1 \neq 0, \text{ and } k_2 \neq 0.$$

Since summing over n_2 (resp. n_1) from 0 to $N-1$ for the mode $k = (k_1, 0)$ (resp. $k = (0, k_2)$) results in N multiplies of the same term, and by grouping the terms for $n_1 = n_2 = n$, and thus $D_{\mathcal{C},1}$ takes the form:

$$D_{\mathcal{C},1}(r,x,\alpha;N) = N\sum_{n=0}^{N-1} A_{\mathcal{C}_r}(x+n\alpha),$$

where

$$A_{\mathcal{C}_r}(x) = \sum_{k_1\neq 0}\hat{\chi}_{\mathcal{C}_r}(k_1,0)e^{i2\pi k_1x_1} + \sum_{k_2\neq 0}\hat{\chi}_{\mathcal{C}_r}(0,k_2)e^{i2\pi k_2x_2} =: \sum_{k\in\mathbb{Z}^2\setminus\{\mathbf{0}\}} a_k(r)e^{i2\pi(k,x)},$$

here $a_k(r) = 0$ when $k_1k_2 \neq 0$ and $a_k(r) = \hat{\chi}_{\mathcal{C}_r}(k)$ when $k_1k_2 = 0$. We will show that $D_{\mathcal{C},1}/N$ is a coboundary (see (2.7)), and thus obtain its limiting distribution.

Lemma 2.1. *For almost every $\alpha \in \mathbb{T}^2$, the series defined by:*

$$B_{\mathcal{C}_r}(\alpha, x) = \sum_{k\neq 0} \frac{a_k(r)}{e^{2\pi i(k,\alpha)} - 1} e^{2\pi i(k,x)},$$

is convergent in $L^2(x)$, and we have

$$A_{\mathcal{C}_r}(x + n\alpha) = B_{\mathcal{C}_r}(\alpha, x + (n+1)\alpha) - B_{\mathcal{C}_r}(\alpha, x + n\alpha). \tag{2.7}$$

Proof. The identity (2.7) is a direct calculation. For the convergence of the series $B_{\mathcal{C}_r}(\alpha, x)$, note that $|\hat{\chi}_{\mathcal{C}_r}(k)| \leq C|k|^{-\frac{3}{2}}$, we have

$$\int_{\mathbb{T}^2} |B_{\mathcal{C}_r}(\alpha, x)|^2 dx \leq C \left(\sum_{k_1\neq 0} \frac{1}{|k_1|^3 |e^{i2\pi k_1\alpha_1} - 1|^2} + \sum_{k_2\neq 0} \frac{1}{|k_2|^3 |e^{i2\pi k_2\alpha_2} - 1|^2} \right).$$

Therefore, it suffices to show that the series

$$\sum_{k_i\neq 0} \frac{1}{|k_i|^3 \|k_i\alpha_i\|^2} \tag{2.8}$$

is convergent for almost every $\alpha_i \in \mathbb{T}$, $i = 1, 2$.

For the following proof, let the constant $C(\alpha_i, \delta)$ vary from line to line. By a standard application of Borel-Cantelli Lemma, we have for almost every $\alpha_i \in \mathbb{T}$, every $k_i > 0$ and every $\delta > 0$ we have

$$\|k_i\alpha_i\| \geq \frac{C(\alpha_i, \delta)}{|k_i|(\ln |k_i|)^{1+\delta}}, \tag{2.9}$$

which gives

$$|\ln \|k_i\alpha_i\|| \leq C(\alpha_i, \delta) \ln |k_i|, \tag{2.10}$$

where $\ln 1$ is defined as 1 for convenience.

Then (2.8) can be estimated as follows:

$$\begin{aligned} \sum_{k_i\neq 0} \frac{1}{|k_i|^3 \|k_i\alpha_i\|^2} &\underset{(2.9)}{\leq} C(\alpha_i, \delta) \sum_{k_i\neq 0} \frac{(\ln |k_i|)^{1+\delta}}{|k_i|^2 \|k_i\alpha_i\|} \\ &\leq C(\alpha_i, \delta) \sum_{k_i\neq 0} \frac{1}{|k_i|(\ln |k_i|)^{2+2\delta} \|k_i\alpha_i\|} \\ &\underset{(2.10)}{\leq} C(\alpha_i, \delta) \sum_{k_i\neq 0} \frac{1}{|k_i|(\ln |k_i|)^{1+\delta} \|k_i\alpha_i\| |\ln \|k_i\alpha_i\||^{1+\delta}}. \end{aligned} \tag{2.11}$$

Note that the integral

$$J = \int_{\mathbb{T}} \frac{1}{\|k_i\alpha_i\| |\ln(\|k_i\alpha_i\|)|^{1+\delta}} d\alpha_i$$

is convergent and its value is *independent* of $k_i \in \mathbb{N}^*$, hence

$$\int_{\mathbb{T}} \sum_{k_i \neq 0} \frac{1}{|k_i|(\ln|k_i|)^{1+\delta} \|k_i\alpha_i\| |\ln\|k_i\alpha_i\||^{1+\delta}} d\alpha_i = \sum_{k_i \neq 0} \frac{J}{|k_i|(\ln|k_i|)^{1+\delta}}$$

is convergent, thus

$$\sum_{k_i \neq 0} \frac{1}{|k_i|(\ln|k_i|)^{1+\delta} \|k_i\alpha_i\| |\ln\|k_i\alpha_i\||)|^{1+\delta}}$$

is convergent for almost every $\alpha_i \in \mathbb{T}$. The L^2 convergence of $B_{\mathcal{C}_r}(\alpha, x)$ follows from the convegence of (2.8) through (2.11). □

Proof of Proposition 2.1.

By (2.7), we have

$$\frac{D_{\mathcal{C},1}(r, x, \alpha; N)}{N} = B_{\mathcal{C}_r}(\alpha, x + N\alpha) - B_{\mathcal{C}_r}(\alpha, x).$$

Since (α, x) is uniformly distributed on $\mathbb{T}^2 \times \mathbb{T}^2$, as $N \to \infty$, the random vector $(\alpha, x, x + N\alpha)$ converges to (α, x, β), where α, x, β are independent random variables that are uniformly distributed on $\mathbb{T}^2 \times \mathbb{T}^2 \times \mathbb{T}^2$. This proves the limiting distribution for $D_{\mathcal{C},1}/N$.

2.4 Non-resonant terms

This section is devoted to highlighting the modes with main contributions in the Fourier series $D_{\mathcal{C},2}$, the final goal is to arrive at the sum (2.26) as an equivalent expression for our Fourier series in terms of limiting distributions. Throughout Section 4, we will use the formula (2.6) since we restrict ourselves to the case symmetric shapes.

For $k = (k_1, k_2)$ and $\alpha = (\alpha_1, \alpha_2)$, we use the notation $\{k_i\alpha_i\} := k_i\alpha_i + l_i$ where l_i is the unique integer such that $-1/2 < k_i\alpha_i + l_i \leq 1/2$. To evaluate $D_{\mathcal{C},2}$, we sum up term by term in the Fourier expansion (2.6) of $\chi_{\mathcal{C}_r}$ for n_1, n_2, and we will simplify by using the summation formula

$$\sum_{n=0}^{N-1} \cos(A + nB) = \frac{\cos(A + \frac{N-1}{2}B)\sin(\frac{N}{2}B)}{\sin\frac{B}{2}}.$$

The normalized term, i.e. divided by $\sqrt{rN}$, for the mode k becomes

$$f(r,x,\alpha;N,k)=c_k(r)\frac{\cos(2\pi(k,x)+\pi(N-1)(\sum_{i=1}^2\{k_i\alpha_i\}))\prod_{i=1}^2\sin(\pi N\{k_i\alpha_i\})}{N^{\frac{1}{2}}\prod_{i=1}^2\sin(\pi\{k_i\alpha_i\})}, \tag{2.12}$$

where $N^{\frac{1}{2}}$ is the remaining normalizer, since $r^{\frac{1}{2}}$ gets cancelled out.

Since the sum $D_{\mathcal{C},2}(r,x,\alpha;N)$ consists of all-non-zero coordinates modes, the normalized sum becomes the following:

$$\Delta(r,x,\alpha;N)=\sum_{k\in\mathbb{Z}^2:\prod_{i=1}^2 k_i\neq 0} f(r,x,\alpha;N,k),$$

note that we have $D_{\mathcal{C},2}(r,x,\alpha;N)=\sqrt{rN}\cdot\Delta(r,x,\alpha;N)$.

***Step* 1.** This step shows that the modes outside the circle of radius N/ϵ have a negligible combined contribution. Given a set S, for a function h defined on $\left(\mathbb{T}^2\right)^2\times S$, we denote by $\|h\|_2$ the supremum of the L^2 norms $\|h(\cdot,s)\|$ over all $s\in S$. Let

$$\Delta_1(r,x,\alpha;N)=\sum_{k\in\mathbb{Z}^2:\forall 1\le i\le 2,\ 0<|k_i|<\frac{N}{\epsilon}} f(r,x,\alpha;N,k)$$

Lemma 2.2. *We have*

$$\|\Delta-\Delta_1\|_2\le C\epsilon^{1/2} \tag{2.13}$$

Proof. Note that

$$\int_{\mathbb{T}}\left(\frac{\sin(\pi N(k_i\alpha_i))}{\sin(\pi(k_i\alpha_i))}\right)^2 d\alpha_i=\int_{\mathbb{T}}\left|\frac{e^{i\pi Nk_i\alpha_i}-e^{-i\pi Nk_i\alpha_i}}{e^{i\pi k_i\alpha_i}-e^{-i\pi k_i\alpha_i}}\right|^2 d\alpha_i$$

$$=\int_{\mathbb{T}}\left|\frac{e^{i2\pi Nk_i\alpha_i}-1}{e^{i2\pi k_i\alpha_i}-1}\right|^2 d\alpha_i=\int_{\mathbb{T}}\left|\sum_{n=0}^{N-1}e^{i2\pi nk_i\alpha_i}\right|^2 d\alpha_i=N,$$

for every $1\le i\le d$.

In the L^2 norm of $\Delta-\Delta_1$, because the integral over x kills the cross terms, only the square terms remain, and note that $|d_r(k)|=\mathcal{O}(|k|^{-\frac{3}{2}})$, we get that

$$\|\Delta-\Delta_1\|_2^2\le CN^2\frac{1}{N}\sum_{|k|\ge\frac{N}{\epsilon}}\frac{1}{|k|^3}\le CN\frac{1}{\frac{N}{\epsilon}}=C\epsilon.$$

□

***Step* 2.** We show that, within the range of $|k|<N/\epsilon$, by taking out a small measure set of α, the divisors admit a lower bound, thus the major contribution of the sum comes from the set of resonant terms in $S(N,\alpha)$ (see (2.16)).

Let

$$E_N = \bigcup_{1\le|n|\le\frac{N}{\epsilon}} \left\{\alpha\in\mathbb{T}^2 : \exists 1\le i\le 2, \quad |n|^{\frac{3}{4}}|\{n\alpha_i\}| < \frac{\epsilon^{\frac{1}{2}}}{N^{\frac{1}{4}}}\right\}. \tag{2.14}$$

and

$$E_N^{(i)} = \bigcup_{1\le|n|\le\frac{N}{\epsilon}} \left\{\alpha_i\in\mathbb{T} : \quad |n|^{\frac{3}{4}}|\{n\alpha_i\}| < \frac{\epsilon^{\frac{1}{2}}}{N^{\frac{1}{4}}}\right\}. \tag{2.15}$$

Note that

$$|E_N| \le 2|E_N^{(i)}| \le 2\sum_{n=1}^{\frac{N}{\epsilon}} \frac{\epsilon^{\frac{1}{2}}}{|n|^{\frac{3}{4}}N^{\frac{1}{4}}} = O\left(\epsilon^{\frac{1}{4}}\right).$$

Let

$$S(N,\alpha) = \left\{k\in\mathbb{Z}^2 : \forall 1\le i\le 2, \quad 0<|k_i|<\frac{N}{\epsilon}, \quad |k_i|^{\frac{3}{4}}|\{k_i\alpha_i\}| < \frac{1}{\epsilon^2 N^{\frac{1}{4}}}.\right\}, \tag{2.16}$$

$$\Delta_2(r,x,\alpha;N) = \sum_{k\in S(N,\alpha)} f(r,x,\alpha;N,k), \tag{2.17}$$

We have

Lemma 2.3. *We have*

$$\|\Delta-\Delta_2\|_{L^2(\mathbb{T}^2\times(\mathbb{T}^2-E_N))} \le C\epsilon^{1/2} \tag{2.18}$$

Proof. By (2.13) it is sufficient to show that $\|\Delta_1-\Delta_2\|^2_{L^2(\mathbb{T}^2\times(\mathbb{T}^2-E_N))} \le C\epsilon$. Note that ouside E_N, we have $N^{\frac{1}{4}}|k_i|^{\frac{3}{4}}\{k_i\alpha_i\} > \epsilon^{1/2}$, for $1\le i\le 2$, this lowerbound is in use throughout the following calculations.

We have

$$\|\Delta_1-\Delta_2\|^2_{L^2(\mathbb{T}^2\times(\mathbb{T}^2-E_N))} \le \frac{C}{N}\sum_{|k|<\frac{N}{\epsilon}} A_k$$

with

$$A_k = c_k^2\int_{\mathbb{T}^2} \frac{1}{\prod_{i=1}^2\{k_i\alpha_i\}}\chi_{\left\{\exists 1\le i\le 2,\ |k_i|^{\frac{3}{4}}|\{k_i\alpha_i\}|\ge\frac{1}{\epsilon^2 N^{\frac{1}{4}}}\right\}} d\alpha.$$

Decompose A_k into 2 parts,

$$A_k \le c_k^2\sum_{j=1}^{2} A(k,j),$$

where $A(k,j)$ denote the sum when the $j-$coordinate violates the condition in $S(N,\alpha)$:

$$\begin{aligned} A(k,j) &= \sum_{p_i\geq 1}\int_{\mathbb{T}} \frac{1}{(\{k_i\alpha_i\})^2}\chi_{\{p_i\epsilon^{\frac{1}{2}}\leq N^{\frac{1}{4}}|k_i|^{\frac{3}{4}}|\{k_i\alpha_i\}|\leq (p_i+1)\epsilon^{\frac{1}{2}}\}}d\alpha_i \\ &\times \sum_{p_j\geq 1}\int_{\mathbb{T}} \frac{1}{(\{k_j\alpha_j\})^2}\chi_{\{\frac{p_j}{\epsilon^2}\leq N^{\frac{1}{4}}|k_j|^{\frac{3}{4}}|\{k_j\alpha_j\}|\leq \frac{(p_j+1)}{\epsilon^2}\}}d\alpha_j \\ &=: \sum_{p_i\geq 1} A(k,i,p_i)\sum_{p_j\geq 1}\bar{A}(k,j,p_j), \end{aligned} \tag{2.19}$$

where the index $i\neq j$. For $p_i\geq 1$ we define

$$B(k,i,p_i) = \left\{\alpha_i\in\mathbb{T} : p_i\epsilon^{\frac{1}{2}} \leq N^{\frac{1}{4}}|k_i|^{\frac{3}{4}}|\{k_i\alpha_i\}| \leq (p_i+1)\epsilon^{\frac{1}{2}}\right\},$$

and for $p_j\geq 1$, define

$$\bar{B_L}(k,j,p_j) = \left\{\frac{p_j}{\epsilon^2} \leq N^{\frac{1}{4}}|k_j|^{\frac{3}{4}}|\{k_j\alpha_j\}| \leq \frac{(p_j+1)}{\epsilon^2}\right\}.$$

Then

$$|B(k,i,p_i)| \leq \frac{\epsilon^{\frac{1}{2}}}{N^{\frac{1}{4}}|k_i|^{\frac{3}{4}}}, \quad |\bar{B_L}(k,j,p_j)| \leq \frac{1}{\epsilon^2 N^{\frac{1}{4}}|k_j|^{\frac{3}{4}}}.$$

Thus

$$A(k,i,p_i) \leq \frac{\epsilon^{\frac{1}{2}}(N^{\frac{1}{4}}|k|^{\frac{3}{4}})^2}{(\epsilon^{\frac{1}{2}})^2 p_i^2 N^{\frac{1}{4}}|k|^{\frac{3}{4}}} \leq \frac{N^{\frac{1}{4}}|k|^{\frac{3}{4}}}{\epsilon^{\frac{1}{2}}p_i^2}, \tag{2.20}$$

similarly,

$$\bar{A}(k,j,p_j) \leq \frac{(\epsilon^2)^2(N^{\frac{1}{4}}|k_j|^{\frac{3}{4}})^2}{\epsilon^2 p_j^2 N^{\frac{1}{4}}|k_j|^{\frac{3}{4}}} \leq \epsilon^2 N^{\frac{1}{4}}|k_j|^{\frac{3}{4}},$$

By using $c_k = O\left(\frac{1}{|k|^{\frac{3}{2}}}\right)$, we obtain

$$A_k \leq C\frac{1}{|k|^3}\epsilon^{\frac{3}{2}}\prod_{i=1}^{2}\left(N^{\frac{1}{4}}|k|^{\frac{3}{4}}\right) \leq C\epsilon^{\frac{3}{2}}\frac{N^{\frac{1}{2}}}{|k|^{\frac{3}{2}}}.$$

Summing over k, we get

$$\sum_{|k|<\frac{N}{\epsilon}} A_k \leq C\epsilon^{\frac{3}{2}}N^{\frac{1}{2}}\sum_{|k|\leq\frac{N}{\epsilon}}\frac{1}{|k|^{\frac{3}{2}}} \leq C\epsilon N,$$

and the claim follows. □

***Step* 3.** In fact, with the bounded range of $\{k_i\alpha_i\}$ in Step 2, we can show that the main

contribution of the Fourier series comes from the modes of coordinates of order N. Let

$$\hat{S}(N,\alpha) = \left\{ k \in \mathbb{Z}^2 : \forall 1 \le i \le 2, \quad N\epsilon^3 < |k_i| < \frac{N}{\epsilon}, \quad |k_i|^{\frac{3}{4}} |\{k_i\alpha_i\}| < \frac{1}{\epsilon^2 N^{\frac{1}{4}}} \right\}, \tag{2.21}$$

$$\Delta_3(r,x,\alpha;N) = \sum_{k\in\hat{S}(N,\alpha)} f(r,x,\alpha;N,k). \tag{2.22}$$

We have

Lemma 2.4.

$$\|\Delta - \Delta_3\|_{L^2(\mathbb{T}^2\times(\mathbb{T}^2-E_N))} \le C\epsilon^{1/2}$$

Proof. By (2.18) it is sufficient to show that $\|\Delta_3 - \Delta_2\|^2_{L^2(\mathbb{T}^2\times(\mathbb{T}^2-E_N))} \le C\epsilon$. We have

$$\|\Delta_3 - \Delta_2\|^2_{L^2(\mathbb{T}^2\times(\mathbb{T}^2-E_N))} \le \frac{C}{N} \sum_{|k|<N\epsilon^3} \hat{A}_k$$

with

$$\hat{A}_k = c_k^2 \prod_{i=1}^{2} \int_{\mathbb{T}^2} \frac{1}{\{k_i\alpha_i\}} \chi_{\left\{ |k_i|^{\frac{3}{4}} |\{k_i\alpha_i\}| \ge \frac{\epsilon^{\frac{1}{2}}}{N^{\frac{1}{4}}} \right\}} d\alpha.$$

Repeating the argument in the Lemma 2.3 by replacing $\bar{A}(k,j,p_j)$ in (2.19) with $A(k,i,p_i)$, and using the inequality (2.20) we obtain

$$\hat{A}_k \le C \frac{1}{|k|^3} \epsilon^{-\frac{1}{2}} \prod_{i=1}^{2} \left(N^{\frac{1}{4}} |k|^{\frac{3}{4}} \right) \le C\epsilon^{-\frac{1}{2}} \frac{N^{\frac{1}{2}}}{|k|^{\frac{3}{2}}}$$

Summing over $|k| \le N\epsilon^3$, we get

$$\sum_{|k|<N\epsilon^3} \hat{A}_k \le C\epsilon^{-\frac{1}{2}} N^{\frac{1}{2}} \sum_{|k|<N\epsilon^3} \frac{1}{|k|^{\frac{3}{2}}} \le C\epsilon N,$$

and the claim follows. □

***Step* 4.** Now the error terms in the Fourier series can be safely removed. Introduce

$$\check{f}(r,x,\alpha;N,k) = \frac{d_k(r)}{c_k(r)} f(r,x,\alpha;N,k)$$

and let

$$\check{\Delta}(r,x,\alpha;N) = \sum_{k\in\hat{S}(N,\alpha)} \check{f}(r,x,\alpha;N,k). \tag{2.23}$$

Since $|c_k - d_k| = \mathcal{O}(|k|^{-\frac{5}{2}})$ and ϵ is fixed, we have

Lemma 2.5.

$$\|\check{\Delta} - \hat{\Delta}\|^2_{L^2(\mathbb{T}^2\times(\mathbb{T}^2-E_N))} \leq \sum_{\epsilon^3 N<|k|<\frac{N}{\epsilon}} \frac{C}{|k|^5}\frac{N}{N^2} \leq \mathcal{O}(N^{-1}). \tag{2.24}$$

Therefore, $\hat{\Delta}$ and $\check{\Delta}$ admit the same limiting distribution if there exists one.

***Step* 5.** Observe that when ϵ is fixed, the sum in (2.23) is limited to large k_i and small $|\{k_i\alpha_i\}|$, $i = 1, 2$. By $|\sin(\{k_i\alpha_i\}) - \{k_i\alpha_i\}| = \mathcal{O}\left(\{k_i\alpha_i\}^3\right) = \mathcal{O}\left(|k_i|^{-\frac{9}{4}}N^{-\frac{3}{4}}\right)$, and with the same calculation as in (2.24), we can replace $\check{f}$ and $\check{\Delta}$ by the following

$$g(r, x, \alpha; N, k) = d_k(r)\frac{\cos(2\pi(k, x) + \pi(N-1)(\sum_{i=1}^2\{k_i\alpha_i\}))\prod_{i=1}^2\sin(\pi N\{k_i\alpha_i\})}{\pi^2 N^{\frac{1}{2}}\prod_{i=1}^2\{k_i\alpha_i\}}.$$

Then we have the following lemma:

Lemma 2.6. *To prove Proposition 2.2, it suffices to prove that*

$$\lim_{N\to\infty}\lambda\{(r, x, \alpha) \in [a, b] \times \mathbb{T}^d \times \mathbb{T}^d \mid \Delta'(r, x, \alpha; N) \leq z\} = \mathcal{D}(z) \tag{2.25}$$

where

$$\Delta' = \sum_{k\in U(N,\alpha)} g(r, x, \alpha; N, k) \tag{2.26}$$

and $U(N, \alpha)$ is any subset of $\mathbb{Z}^2$ that contains $\hat{S}(N, \alpha)$.

2.5 Geometry of the space of lattices

2.5.1. Following [Dan85], Section 2, and [DF14], Section 4, we show by Dani correspondence that the set $\hat{S}(N, \alpha)$ corresponds to a set of short vectors in lattices in $M^2 = M \times M$, where the lattice takes the form (L_1, L_2), and $L_i \in M = SL(2, \mathbb{R})/SL(2, \mathbb{Z})$, hence the discrepancy function Δ' can be seen as a function on the homogeneous space M^2.

Let

$$g_T = \begin{pmatrix} e^{-T} & 0 \\ 0 & e^T \end{pmatrix}, \quad \Lambda_{\alpha_i} = \begin{pmatrix} 1 & 0 \\ \alpha_i & 1 \end{pmatrix}.$$

Consider the product lattice $L(N, \alpha) = L(N, \alpha_1) \times L(N, \alpha_2)$, where $L(N, \alpha_i) = g_{\ln N}\Lambda_{\alpha_i}\mathbb{Z}^2$. For each $k = (k_1, k_2) \in \mathbb{Z}^2$, we associate the vectors $\boldsymbol{k}_i = \boldsymbol{k}_i(k_i) = (k_i, l_i)$, where l_i is the unique interger such that $-\frac{1}{2} < k_i\alpha_i + l_i \leq \frac{1}{2}$. We then denote

$$(X_i, Z_i) = (k_i/N, N\{k_i\alpha_i\}) = g_{\ln N}\Lambda_{\alpha_i}\boldsymbol{k}_i \tag{2.27}$$

We have $k \in \hat{S}(N,\alpha)$ (see (2.21) if and only if :

$$\epsilon^3 < |X_i| < \frac{1}{\epsilon}, \quad |X_i|^{\frac{3}{4}}|Z_i| < \frac{1}{\epsilon^2} \tag{2.28}$$

Recall the definition of the shortest vectors $\{e_1(N,\alpha_i),\ e_2(N,\alpha_i)\}$ of $L(N,\alpha_i)$ in Section 2. The following Lemma is essentially the same as Lemma 4.1 in [DF14], since the two components of our product lattice space are independent in the following argument.

Lemma 2.7. *For every $\epsilon > 0$ there exists $K(\epsilon) > 0$ such that for α outside E_N, each $k \in \hat{S}(N,\alpha)$ corresponds to a unique pair of vectors $(m^1, m^2) \in \mathbb{Z}^2 \times \mathbb{Z}^2$ such that for $i = 1, 2$, $\|m^i\| \leq K(\epsilon)$ and*

$$g_{\ln N}\Lambda_{\alpha_i}\boldsymbol{k}_i = m_1^i e_1(N,\alpha_i) + m_2^i e_2(N,\alpha_i).$$

Conversely, for $\epsilon > 0$ fixed and N large enough, $\alpha \notin E_N$ implies that for each pair of vectors $(m^1, m^2) \in \mathbb{Z}^2 \times \mathbb{Z}^2$, where $\|m^i\| \leq K(\epsilon)$, $i = 1, 2$, there exists a unique $k = (k_1, k_2) \in \mathbb{Z}^2$ such that for $i = 1, 2$,

$$g_{\ln N}\Lambda_{\alpha_i}\boldsymbol{k}_i = (m^i, e(N,\alpha_i)) = m_1^i e_1(N,\alpha_i) + m_2^i e_2(N,\alpha_i).$$

Denote $U(N,\alpha,\epsilon)$ the set of $k = (k_1, k_2) \in \mathbb{Z}^2$ that corresponds to the set of pairs of vectors

$$V(N,\alpha,\epsilon) = \{(m^1, m^2) \in \mathbb{Z}^2 \times \mathbb{Z}^2 \mid \|m^i\| \leq K(\epsilon),\ i = 1, 2\}. \tag{2.29}$$

Remark. *Note that $m^i = 0$ iff $k_i = 0$, hence in particular, each $k \in \hat{S}(N,\alpha)$ corresponds to a unique pair of vectors $\mathbf{m} = (m^1, m^2) \in (\mathbb{Z}^2 \setminus \{\mathbf{0}\})^2$.*

Proof. From (2.28) we can deduce that $k \in \hat{S}(N,\alpha)$ implies $g_{\ln N}\Lambda_{\alpha_i}\boldsymbol{k}_i$ is shorter that $R(\epsilon) = \epsilon^{-\frac{13}{4}}$ for $i = 1, 2$. For each $L_i = L(N,\alpha_i)$, the short vectors $e_1(L_i), e_2(L_i)$ form a basis in $\mathbb{R}^2$, hence the norm $\|x\|$ is equivalent to the norm $\|\sum_{j=1,2} x_j e_j(L_i)\|$. As a consequence, for every $L = (L_1, L_2)$, there exists $K(L)$, such that if $m^i \in \mathbb{Z}^2$ satisfies $\|m^i\| \geq K(L)$, we have $\|(m^i, e(L_i))\| \geq R(\epsilon)$. Now we show that the choice of $K(L)$ can be uniform for the set of lattices $\{L \in M^2 \mid \prod_{i=1}^2 L(N,\alpha_i), \alpha \notin E_N\}$, it is enough to show that the set

$$\{\prod_{i=1}^2 L(N,\alpha_i), \alpha \notin E_N\} = \prod_{i=1}^2 \{L(N,\alpha_i), \alpha_i \notin E_N^{(i)}\} \tag{2.30}$$

is precompact, see (2.15) for the definition of $E_N^{(i)}$, or equivalently, that each component is precompact. By the definition of $E_N^{(i)}$ (see (2.15)), when $\alpha_i \notin E_N^{(i)}$, if $|X_i| < \epsilon^3$, then $|Z_i| \geq \epsilon^{-\frac{7}{4}}$. For any $l \in \mathbb{Z}$, we have $|N(k_i\alpha_i + l)| \geq |N(\{k_i\alpha_i\}| = |Z_i|$, thus for all vectors in L_i, either $|X_i|$ or $|Z_i|$ admits a lower bound $\delta = \delta(\epsilon) > 0$, take $\delta(\epsilon) = \epsilon^3$ for example. Therefore, all vectors in L_i are longer than δ. By the Mahler compactness criterion for

lattices[Rag72], the set (2.30) is precompact. This shows that in fact, the constant $K(L)$ depends only on ϵ, and therefore proves the first part of the lemma.

For the second part of the lemma, let $\epsilon > 0$ be fixed and let N be much greater than $K(\epsilon)$. If $\|m^i\| \leq K(\epsilon)$, by the equivalence between two norms, we have that $\|(m^i, e(N, \alpha_i))\|$ is bounded by some $C(\epsilon, \alpha_i)$, which is much smaller than N. For every $m^i \in \mathbb{Z}^2$, we have $(m^i, e(N, \alpha_i)) = g_{\ln N}\Lambda_{\alpha_i}\bar{\boldsymbol{k}}_i$ for some unique $\bar{\boldsymbol{k}}_i = (k_i, \tilde{l}_i) \in \mathbb{Z}^2$. This shows that $\tilde{l}_i = l_i$, where l_i satisfies $-\frac{1}{2} < k_i\alpha_i + l_i \leq \frac{1}{2}$. Indeed, if $\tilde{l}_i$ is not equal to l_i, then $|N(k_i\alpha_i + \tilde{l}_i)| \geq N/2$, contradicting the fact that $\|g_{\ln N}\Lambda_{\alpha_i}\bar{\boldsymbol{k}}_i\| = \|(m^i, e(N, \alpha_i))\|$ is much smaller than N. Therefore, we have $\tilde{l}_i = l_i$, $\bar{\boldsymbol{k}}_i = \boldsymbol{k}_i$, which proves the unicity of the vector $k = (k_1, k_2) \in \mathbb{Z}^2$.

□

2.5.2. For $n^i \in \mathbb{Z}^2$, $\alpha_i \in \mathbb{T}$, we define the coordinates of the corresponding vector as

$$(n^i, e(N, \alpha_i)) = (X_{n^i}, Z_{n^i}) \tag{2.31}$$

and define $\mathbf{n} = (n^1, n^2) \in \mathbb{Z}^2 \times \mathbb{Z}^2$, $X_{\mathbf{n}} = (X_{n^1}, X_{n^2}) \in \mathbb{R}^2$ and $R_{\mathbf{n}} = \|X_{\mathbf{n}}\|$. We introduce an equivalent function to $g(r, x, \alpha; N, k)$ with $\mathbf{n}$ as a variable:

$$h(r, x, \alpha; N, \mathbf{n}) = \frac{d_r(N, \mathbf{n}) \cos(2\pi N(X_{\mathbf{n}}, x) + \pi\frac{N-1}{N}(\sum_{i=1}^2 Z_{n^i})) \prod_{i=1}^2 \sin(\pi Z_{n^i})}{R_{\mathbf{n}}^{\frac{3}{2}} \prod_{i=1}^2 Z_{n^i}}$$

with

$$d_r(N, \mathbf{n}) = \frac{1}{\pi^3} K^{-\frac{1}{2}}(X_{\mathbf{n}}/R_{\mathbf{n}}) \sin(2\pi(rNP(X_{\mathbf{n}}) - \frac{1}{8})).$$

From Lemma 2.7, we know that for $\alpha \notin E_N$,

$$\sum_{\mathbf{n}=(n^1,n^2)\in V(N,\alpha,\epsilon)} h(r, x, \alpha; N, \mathbf{n}) = \sum_{k\in U(N,\alpha,\epsilon)} g(r, x, \alpha; N, k), \tag{2.32}$$

where $U(N, \alpha, \epsilon) \supset \hat{S}(N, \alpha)$.

Now we will reformulate the sum above as a function defined on the lattice space. If we restrict ourselves to prime vectors $\mathbf{n} \in \mathbb{Z}^2 \times \mathbb{Z}^2$, the variables $rNP(X_{\mathbf{n}})$ mod 1 will become independent random variables that are uniformly distributed on $\mathbb{T}^1$ as $N \to \infty$. In fact, any $\mathbf{n} = (n^1, n^2) \in (\mathbb{Z}^2 \setminus \{0\})^2$ could be rewritten as $\check{p}(p_1m^1, p_2m^2)$, where $\check{p}$ is the signed greatest common divisor such that the first coordinate of (p_1m^1, p_2m^2) is positive, $p = (p_1, p_2) \in \mathcal{Z}$ with strictly positive first entry, and m^1, m^2 are in $\mathcal{Z}$. Since all the vectors are multiples of the prime ones, we introduce

$$X_{p,\mathbf{m}} = (p_1X_{m^1}, p_2X_{m^2}) \tag{2.33}$$

and $R_{p,\mathbf{m}} = \|X_{p,\mathbf{m}}\|$.

Now we can express our discrepancy function using prime vectors. Define

$$q(r,x,\alpha;N,\mathbf{m},p,\check{p}) = \frac{d_r(N,\mathbf{m},p,\check{p})\cos\left(2\pi\check{p}\left(\sum_{i=1}^{2}\left(p_i\left(m_1,\gamma_i\left(\alpha,x,N\right)\right)\right)\right)+\pi\frac{N-1}{N}\check{p}\left(\sum_{i=1}^{2}\left(p_iZ_{m^i}\right)\right)\right)\prod_{i=1}^{2}\sin\left(\pi\check{p}p_iZ_{m^i}\right)}{|\check{p}|^{\frac{7}{2}}R_{p,\mathbf{m}}^{\frac{3}{2}}\prod_{i=1}^{2}\left(p_iZ_{m^i}\right)} \tag{2.34}$$

where

$$d_r(N,\mathbf{m},p,\check{p}) = \frac{1}{\pi^3}K^{-\frac{1}{2}}\left(\frac{X_{p,\mathbf{m}}}{R_{p,\mathbf{m}}}\right)\sin(2\pi(\check{p}rNP(X_{p,\mathbf{m}})-\frac{1}{8})),$$

and

$$\gamma_i(\alpha,x,N) = Nx_i(e_{11}(N,\alpha_i),e_{21}(N,\alpha_i)), \tag{2.35}$$

where e_{ij} is the j_{th} coordinate of the short vector e_i.

By using symmetry of the convex body, for $p^+ = (p_1,p_2)$ and $p^- = (p_1,-p_2)$, we have $K\left(X_{p^+,\mathbf{m}}/R_{p^+,\mathbf{m}}\right) = K\left(X_{p^-,\mathbf{m}}/R_{p^-,\mathbf{m}}\right)$ and $P(X_{p^+,\mathbf{m}}) = P(X_{p^-,\mathbf{m}})$, we can regroup the terms for p^+ and p^- for $q(r,x,\alpha;N,\mathbf{m},p,\check{p})$, hence we arrive at

$$q'(r,x,\alpha;N,\mathbf{m},p,\check{p}) = 2\cdot\frac{d_r(N,\mathbf{m},p,\check{p})}{|\check{p}|^{\frac{7}{2}}R_{p,\mathbf{m}}^{\frac{3}{2}}}\prod_{i=1}^{2}\frac{\cos\left(2\pi\check{p}p_i\left(m_1,\gamma_i\left(\alpha,x,N\right)\right)+\pi\frac{N-1}{N}\check{p}p_iZ_{m^i}\right)\sin\left(\pi\check{p}p_iZ_{m^i}\right)}{p_iZ_{m^i}}$$

Recall the definition of $\mathcal{Z}$ in **Notations** in Section 2, and that $\mathcal{Z}^2 = \mathcal{Z}\times\mathcal{Z}$. Define $\mathcal{Z}_\epsilon = \{\mathbf{m}\in\mathcal{Z}^2 : \|m^i\|\le K(\epsilon);\ i=1,2\}$. Since the $\mathcal{Z}$ consists of the primitive vectors with positive first coordinate, we need to sum over both the positive and negative $\check{p}$'s. By Lemma 2.6 and the identity (2.32) we have the following proposition:

Proposition 2.3. *Assume that (r,x,α) are uniformly distributed in $X = [a,b]\times\mathbb{T}^2\times\mathbb{T}^2$, if the sum*

$$\sum_{\check{p}\in\mathbb{Z}\setminus\{0\}}^{\infty}\sum_{p\in\mathcal{Z}'}\sum_{\mathbf{m}\in\mathcal{Z}_\epsilon} q(r,x,\alpha;N,\mathbf{m},p,\check{p}). \tag{2.36}$$

converge to the some law in distribution as $N\to\infty$ and then $\epsilon\to 0$, then the normalized discrepancy function $D_{\mathcal{C},2}$ in (2.4) of Theorem 1 converges to the same law.

2.5.3. Uniform distribution of unstable submanifold Λ_α. For each i, Λ_{α_i} is the unstable submanifold under the geodesic flow g_T and will become equidistributed over the whole manifold M as $T\to N$, naturally the same uniform distribution law of holds in the finite product space M^2, hence we have the following proposition(see [MS10], Theorem 5.3):

Proposition 2.4. *Denote by μ the Haar measure on M^2. If $\Phi : (\mathbb{R}^2\times\mathbb{R}^2)^2\times\mathbb{R}^2\to\mathbb{R}$ is*

a bounded continuous function, then

$$\begin{aligned}&\lim_{N\to\infty}\int_{\mathbb{T}^2}\Phi(e(L(N,\alpha_1)),e(L(N,\alpha_2)),\alpha)d\alpha\\&=\int_{M^2\times\mathbb{T}^2}\Phi(e(L_1),e(L_2),\alpha)d\mu(L_1\times L_2)d\alpha\end{aligned}\tag{2.37}$$

2.6 Oscillating terms

In this section we will prove that typical variables appeared in the sum (2.36) will behave like independent uniformly distributed random variables. We denote by μ_2 the distribution of $(e(L_1),e(L_2))$ when $L=(L_1,L_2)$ is distributed according to Haar measure on $M^2=\prod_{2\text{ copies}}SL(2,\mathbb{R})/SL(2,\mathbb{Z})$. We denote by $\lambda_{2,\epsilon}$ the Haar measure on $(\mathbb{T}^2)^2\times\mathbb{T}^{\mathcal{Z}'\times\mathcal{Z}_\epsilon}$.

The main result of this section is the following, from which the main theorem follows:

Proposition 2.5. *Assume that (x,α,r) are uniformly distributed on $\mathbb{T}^2\times\mathbb{T}^2\times[a,b]$, then the following random variables*

$$e(N,\alpha_1),\quad e(N,\alpha_2),\quad(\{\gamma_{11}\},\{\gamma_{21}\}),\quad(\{\gamma_{12}\},\{\gamma_{22}\}),\quad\{\{A_{p,\mathbf{m}}\}\ :\ p\in\mathcal{Z}',m\in\mathcal{Z}_\epsilon\}$$

where $A_{p,\mathbf{m}}=rNP(X_{p,\mathbf{m}})$, $(\gamma_{i1},\gamma_{i2})=\gamma_i(\alpha,x,N)\in\mathbb{R}^2$, defined as in (2.35) for $i=1,2$, and the bracket $\{\dots\}$ stands for the fractional part of the variable, converge in distribution to $\mu_2\times\lambda_{2,\epsilon}$ as $N\to\infty$.

We will prove Proposition 2.5 in Section 2.6.2. In the next section, we first prove that for different vectors $\left(p^{(1)},\mathbf{m}^{(1)}\right),\dots,\left(p^{(K)},\mathbf{m}^{(K)}\right)$ in $\mathcal{Z}'\times\mathcal{Z}^2$, $\{P(X_{p^{(i)},\mathbf{m}^{(i)}})\}_{i=1}^K$ are typically independent over $\mathbb{Q}$.

2.6.1. Exceptionally in this subsection we use the *lower* index for m_i to represent a *vector* in $\mathbb{Z}^2$, not to be confused with the coordinates in the **Notations** in section 2. For $\mathbf{m}=(m_1,m_2)\in\mathbb{Z}^2\times\mathbb{Z}^2$ with $m_i\in\mathbb{Z}^2$, and $p=(p_1,p_2)\in\mathbb{Z}^2$, $p_1\geq 1$, define the function $Q_{p,\mathbf{m}}:\ \mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}$ as $Q_{p,\mathbf{m}}(z_1,z_2)=P(\langle p_1m_1,z_1\rangle,\langle p_2m_2,z_2\rangle)$, where $z_i=(z_{i1},z_{i2})\in\mathbb{R}^2$ is a vector, and the bracket $\langle\cdot,\cdot\rangle$ means euclidean inner product in $\mathbb{R}^2$.

Proposition 2.6. *For different vectors $\left(p^{(1)},\mathbf{m}^{(1)}\right),\dots,\left(p^{(K)},\mathbf{m}^{(K)}\right)$ in $\mathcal{Z}'\times\mathcal{Z}^2$, if $l_1,\dots,l_K$ are such that*

$$\sum_{i=1}^K l_iQ_{p^{(i)},\mathbf{m}^{(i)}}\equiv 0,\tag{2.38}$$

then $l_i=0$ for $i=1,\dots,K$.

Proof. We adapt the proof for Lemma 5.2 and Proposition 5.4 in [DF14] to our case. It is not too hard to see the following sublemma (see Lemma 5.2 in [DF14]):

Sublemma 2.8. *The function P is real analytic on and not equal to a polynomial on $(\mathbb{R}\setminus\{0\})^2$. Similarly, $Q_{p,\mathbf{m}}$ is real analytic on and not equal to a polynomial on $(\mathbb{R}\setminus\{0\})^2\times(\mathbb{R}\setminus\{0\})^2$.*

Proof. Note that $P(t_1,t_2)=\sqrt{t_1^2+t_2^2}P(\frac{t_1}{\sqrt{t_1^2+t_2^2}},\frac{t_2}{\sqrt{t_1^2+t_2^2}})$. Since $P(\frac{t_1}{\sqrt{t_1^2+t_2^2}},\frac{t_2}{\sqrt{t_1^2+t_2^2}})$ is bounded, if P is a polynomial in t_1 or t_2, the degree of t_1 or t_2 is at most 1. But this is impossible if t_1 and t_2 are non-zero, because P is strictly positive and non constant. The proof for $Q_{p,\mathbf{m}}$ is similar. □

Let $S_{\mathbf{m}^{(i)}}$ denote the partiel sum of the terms with the same $\mathbf{m}^{(i)}$:

$$S_{\mathbf{m}^{(i)}}=\sum_{j:\ \mathbf{m}^{(j)}=\mathbf{m}^{(i)}} l_jQ_{p^{(j)},\mathbf{m}^{(j)}}.$$

First we prove that all these partial sums $S_{\mathbf{m}^{(i)}}$ must be zero. Let $z_1=\delta\alpha_1+\theta\beta_1$, where $\alpha_1=(\alpha_{11},\alpha_{12})$, $\beta_1=(\beta_{11},\beta_{12})\in\mathbb{R}^2$, similarly define $z_2=\delta'\alpha_2+\theta'\beta_2$. Fix β_j such that $\langle\beta_j,m_j^{(i)}\rangle\neq 0$, for $j=1,2$ and all $1\le i\le K$, and assume that $\langle m_1^{(i)},\alpha_1\rangle\neq 0$ for all $1\le i\le K$. we have $z_{1j}=\delta\alpha_{1j}+\theta\beta_{1j}$, and $z_{2j}=\delta'\alpha_{2j}+\theta'\beta_{2j}$, $j=1,2$, and

$$Q_{p^{(i)},\mathbf{m}^{(i)}}(z_1,z_2)=\left(\prod_{j=1}^{2}p_j^{(i)}|\langle m_j^{(i)},\alpha_j\rangle|\right)P\left(\pm\delta+\frac{\langle m_1^{(i)},\beta_1\rangle}{|\langle m_1^{(i)},\alpha_1\rangle|}\theta,\pm\delta'+\frac{\langle m_2^{(i)},\beta_2\rangle}{|\langle m_2^{(i)},\alpha_2\rangle|}\theta'\right),\tag{2.39}$$

where the sign $\pm$ in front of δ (δ' resp.) takes the same sign as $\langle m_1^{(i)},\alpha_1\rangle$ ($\langle m_2^{(i)},\alpha_2\rangle$ resp.). Fix all the other variables, expand the entire sum (2.38) in powers series of θ and then θ', and consider the coefficient of $\theta^2\theta'^2$, we have:

$$\sum_{i=1}^{K}l_i\left(\prod_{i=1}^{2}p_j^{(i)}\frac{\langle m_j^{(i)},\beta_j\rangle^2}{|\langle m_j^{(i)},\alpha_j\rangle|}\right)\partial_1^2\partial_2^2P(\pm\delta,\pm\delta')=0.\tag{2.40}$$

Since for $j=1,2$, all $m_j^{(i)}$'s are prime vectors in $\mathbb{Z}^2$, we can choose α_j such that for one value of $m_j^{(i)}$, $|\langle m_j^{(i)},\alpha_j\rangle|$ is arbitrarily small, while for other values of $m_j^{(k)}$'s, $|\langle m_j^{(k)},\alpha_j\rangle|$ admit a uniform lower bound. Thus the sum of $l_iQ_{p^{(i)},\mathbf{m}^{(i)}}$ with identical $m_j^{(i)}$ must be zero. Hence we have that $S_{\mathbf{m}^{(i)}}$ must be zero.

Now we assume that all $\mathbf{m}^{(i)}$ are the same, we will prove that the partial sum with the greatest $p_1^{(j)}$ is zero. For the sake of simplicity, we assume that $m_1^{(i)}=(1,0)$ and $m_2^{(i)}=(1,0)$. First we suppose that $z_{21}\neq 0$, choose j such that $p_1^{(j)}$ is the greatest among all $p_1^{(i)}$, then

$$Q_{p^{(i)},\mathbf{m}^{(i)}}(z_1,z_2)=P(p_1^{(i)}z_{11}.p_2^{(i)}z_{21})=p_1^{(j)}P\left(\frac{p_1^{(i)}}{p_1^{(j)}}z_{11},\frac{p_2^{(i)}}{p_1^{(j)}}z_{21}\right),$$

Consider the n-th partial derivative of $Q_{p^{(i)},\mathbf{m}^{(i)}}$ with respect to z_{11}, then

$$\frac{\partial^n}{\partial z_{11}} Q_{p^{(i)},\mathbf{m}^{(i)}}(z_1,z_2) = \frac{\left(p_1^{(i)}\right)^n}{\left(p_1^{(j)}\right)^{n-1}} \frac{\partial^n}{\partial z_{11}} P\left(\frac{p_1^{(i)}}{p_1^{(j)}} z_{11}, \frac{p_2^{(i)}}{p_1^{(j)}} z_{21}\right),$$

Since $\frac{\left(p_1^{(i)}\right)^n}{\left(p_1^{(j)}\right)^{n-1}} < 1$ for all $i \neq j$, we can take n sufficiently large by analyticity of P, then l_j becomes the dominant coefficient in the linear combination of n-th derivatives, we must have the linear combination of terms of identical maximal $p_1^{(j)}$ is zero. By repeating this procedure for $p_2^{(i)}$, we can deduce that the coefficient l_i in front the term of that has the greatest $p_2^{(i)}$ among those having the greatest $p_1^{(i)}$ is zero. Inductively, all coefficients l_i are zero. □

By Proposition 2.6, we can deduce the following: if we take a lattice $L \in M^2$ and let $z_i = (e_{11}(L_i), e_{21}(L_i))$, $i = 1,2$, then $P\left(X_{p,\mathbf{m}}(L)\right) = P(\langle p_1 m_1, z_1\rangle, \langle p_2 m_2, z_2\rangle) = Q_{p,\mathbf{m}}(z_1,z_2)$. By analyticity, for any different $\left(p^{(1)},\mathbf{m}^{(1)}\right),\ldots,\left(p^{(K)},\mathbf{m}^{(K)}\right)$ in $\mathcal{Z}' \times \mathcal{Z}^2$, if not all numbers l_i are zero, then we have

$$\mu\left(L : \sum_{i=1}^{K} l_i P(X_{p^{(i)},\mathbf{m}^{(i)}}(L)) = 0\right) = 0. \tag{2.41}$$

Now by Proposition 2.4 we have that

$$mes\left(\alpha \in \mathbb{T}^2 \ : \ \left|\sum_{i=1}^{K} l_i P\left(X_{p^{(i)},\mathbf{m}^{(i)}}\left(L(N,\alpha)\right)\right)\right| < \epsilon\right) \to 0 \tag{2.42}$$

as $\epsilon \to 0$, $N \to \infty$.

2.6.2. Proof of Proposition 2.5 Take integers $n_{11}, n_{21}, n_{12}, n_{22}$, $\{l_{p,\mathbf{m}}\}_{p\in\mathcal{Z}',m\in\mathcal{Z}_\epsilon}$, where $l_{p,\mathbf{m}}$ are non-zero only for finitely many $(p,\mathbf{m})$, and a function $\Phi : (\mathbb{R}^2 \times \mathbb{R}^2)^2 \to \mathbb{R}$ of compact support. It remains to show that

$$\iiint \Phi(e(L(N,\alpha_1)), e(L(N,\alpha_2)) exp\left[2\pi i\left(\sum_{j=1}^{2}(n_{j1}\gamma_{j1} + n_{j2}\gamma_{j2}) + \sum_{p\in\mathcal{Z}',m\in\mathcal{Z}_\epsilon} l_{p,\mathbf{m}} A_{p,\mathbf{m}}\right)\right]$$
$$dx d\alpha dr \to \int_{M^2} \Phi(e(L_1), e(L_2)) d\mu(L) \int_{\mathbb{T}^{2d}} e^{2\pi i \sum_j (n_{j1}\gamma_{j1} + n_{j2}\gamma_{j2})} d\gamma \int_{\mathbb{T}^{\mathcal{Z}'\times\mathcal{Z}_\epsilon}} e^{2\pi i \sum_{p,\mathbf{m}} l_{p,\mathbf{m}} A_{p,\mathbf{m}}} dA, \tag{2.43}$$

as $N \to \infty$.

Proof. This proof is very close to the proof of Proposition 5.1 in [DF14], it suffice to rewrite the original proof with the new variables and use Proposition 2.6 and (2.42). If for all j and $p,\mathbf{m}$, $n_{j1} \equiv 0$, $n_{j2} \equiv 0$ and $l_{p,\mathbf{m}} \equiv 0$, (2.43) is a special case of (2.37). Then it suffice

to prove (2.43) in the case that at least some n_j or some $l_{p,\mathbf{m}}$ are non-zero, in which case the right-hand side of (2.43) is zero, and it reduces to the following:

$$\iiint \Phi(e(L(N,\alpha_1)), e(L(N,\alpha_2)exp\left[2\pi i\left(\sum_{j=1}^{2}(n_{j1}\gamma_{j1}+n_{j2}\gamma_{j2})+\sum_{p\in\mathcal{Z}',m\in\mathcal{Z}_\epsilon} l_{p,\mathbf{m}}A_{p,\mathbf{m}}\right)\right]dxd\alpha dr \to 0 \tag{2.44}$$

If $n_{j1}\neq 0$ for at least one j, recall that $\gamma_j(\alpha,x,N)=Nx_j(e_{11}(N,\alpha_j),e_{21}(N,\alpha_j))$, then the coefficient in front of x_j in $\sum_j(n_{j1}\gamma_{j1}+n_{j2}\gamma_{j2})$ is $N(n_{j1}e_{11}(N,\alpha_j)+n_{j2}e_{21}(N,\alpha_j))$. Note that the coordinates $e_{11}(N,\alpha_j)$ and $e_{21}(N,\alpha_j)$ are typically $\mathbb{Z}$-independent outside a zero measure set of α_j. Hence (2.37) implies that

$$mes\left(\alpha_j\in\mathbb{T}:|n_{j1}e_{11}(N,\alpha_j)+n_{j2}e_{21}(N,\alpha_j)|<\frac{1}{\sqrt{N}}\right)\to 0 \tag{2.45}$$

as $N\to\infty$. This limit states that most α_j will not allow the coefficient in front of x_j to be too small, then the integral of (2.44) can be decomposed into two parts, $LHS=I_1+I_2$, where I_1 corresponds to the part of integral for α_j with $|n_{j1}e_{11}(N,\alpha_j)+n_{j2}e_{21}(N,\alpha_j)|<\frac{1}{\sqrt{N}}$ and I_2 the part for α_j with $|n_{j1}e_{11}(N,\alpha_j)+n_{j2}e_{21}(N,\alpha_j)|\geq\frac{1}{\sqrt{N}}$. Then

$$|I_1|\leq Const(\Phi)mes\left(\alpha_j\in\mathbb{T}:|n_{j1}e_{11}(N,\alpha_j)+n_{j2}e_{21}(N,\alpha_j)|<\frac{1}{\sqrt{N}}\right)$$

so it can be arbitrarily small as $N\to\infty$ by (2.45) . For I_2, since the coefficient of x_j is not too small, we use integration by parts with respect to x_j to achieve the following estimation:

$$|I_2|\leq\frac{Const(\Phi)}{\sqrt{N}}.$$

This proves the case where not all n_{j1} vanish, the case where not all n_{j2} vanish is the same.

Similarly, if there exists some $(p,\mathbf{m})$, such that $l_{p,\mathbf{m}}$ is non-zero, we can use (2.42) and integrate with respect to r to obtain (2.44), using the same decomposition and integration techniques. □

2.6.3. Proof of Theorem 2.1(b). Combining Proposition 2.3 and Proposition 2.5, by letting $N\to\infty$ and then $\epsilon\to 0$, we can subsitute the variables in (2.36) by uniformly distribtuted random variables on the infinite tori, thus we obtain Theorem 2.1(b) and Proposition 2.2.

Chapter 3

Absolute Bounds - Linear Form Sequences

3.1 Introduction

Given an irrational number $\alpha \in \mathbb{R}$, the irrational rotation T_α over $\mathbb{T} = (\mathbb{R}/\mathbb{Z}) \cong [0,1)$ is defined by $x \mapsto x + \alpha \mod 1$ for $x \in \mathbb{T}$. By Weyl's criterion (see [DT06], section 1.2.1), the sequence $\{n\alpha \mod 1\}_{1 \le n \le N}$ becomes equidistributed over $\mathbb{T}$ as $N \to \infty$, i.e., for any interval $B \subset \mathbb{T}$,

$$\frac{\sum_{n=1}^{N} \chi_B(n\alpha)}{N} \to \mathrm{Vol}(B), \quad N \to \infty.$$

To measure the rate of convergence, we introduce the discrepancy function defined as the difference between the *actual* number of hits in B before time N and the *expected* number of hits $N \cdot \mathrm{Vol}(B)$:

$$D_B(\alpha; N) = \sum_{n=1}^{N} \chi_B(n\alpha \mod 1) - N\mathrm{Vol}(B).$$

In 1920s, Khintchine [Khi23] proved that for any $\epsilon > 0$, for almost every irrational rotation T_α, the maximal discrepancy relative to all possible intervals in $[0,1)$:

$$\Delta(\alpha; N) = \max_{x \in (0,1]} |D_{[0,x)}(\alpha; N)|$$

is exactly between $\ln N \cdot \ln\ln N$ and $C(\alpha, \epsilon) \cdot \ln N (\ln\ln N)^{1+\epsilon}$. His proof used the continued fraction algorithm for irrationals. Due to the absence of the continued fraction algorithm in higher dimensions, the research about the higher-dimensional counterpart of Khintchine's theorem proved to be difficult. Schmidt [Sch64] proved that the maximal discrepancy $\Delta(\boldsymbol{\alpha}; N)$(see the definition below) in dimension k has an upper bound of $C(\boldsymbol{\alpha}, \epsilon)(\ln N)^{k+1+\epsilon}$ for almost every $\boldsymbol{\alpha} \in \mathbb{R}^k$, by using the Erdős–Turán-Koksma inequality, In 1994, by a

suprising method which consists of a combination of Fourier analysis, the "second-moment method" and combinatorics, J. Beck[Bec94] successfully got rid of the extra $\ln N$ factor and proved the following multidimensional analogue of Khintchine's theorem:

Theorem 3.1. *Let $k \geq 2$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k) \in \mathbb{R}^k$ be the translation vector, and $B(\mathbf{x}) = [0, x_1) \times \cdots \times [0, x_k) \subset [0,1)^k$, define the ergodic discrepancy:*

$$D(\boldsymbol{\alpha}, \mathbf{x}; m) = \sum_{1 \leq n \leq m} \chi_{B(\mathbf{x})}(n\boldsymbol{\alpha}) - m\,Vol(B(\mathbf{x}))$$

and the maximal discrepancy:

$$\Delta(\boldsymbol{\alpha}; N) = \max_{\substack{1 \leq m \leq N \\ \mathbf{x} \in [0,1]^d}} |D(\boldsymbol{\alpha}, \mathbf{x}; m)|.$$

Then for arbitray positive increasing function $\varphi(n)$ of n,

$$\Delta(\boldsymbol{\alpha}; N) \ll (\ln N)^k \cdot \varphi(\ln \ln N) \Longleftrightarrow \sum_{n=1}^{\infty} \frac{1}{\varphi(n)} < \infty \tag{3.1}$$

for almost every $\boldsymbol{\alpha} \in \mathbb{R}^k$, where $\ll$ denotes the Vinogradov symbol, e.g. $f(N) \ll g(N)$ means that $|f(N)| < c \cdot g(N)$ for all N with a uniform constant c.

In this chapter, we consider the sequence of linear forms $\left\{\sum_{1 \leq i \leq d} k_i \alpha_i \mod 1\right\}_{\substack{1 \leq k_i \leq N \\ 1 \leq i \leq d}}$ relative to intervals in $[0,1)$. Define the ergodic discrepancy:

$$D(\boldsymbol{\alpha}, x; N) = \sum_{\substack{1 \leq k_i \leq N \\ 1 \leq i \leq d}} \chi_{[0,x)} \left(\sum_{1 \leq i \leq d} k_i \alpha_i \mod 1 \right) - N^d x,$$

and the maximal discrepancy:

$$\Delta(\boldsymbol{\alpha}; N) = \max_{0 < x \leq 1} |D(\boldsymbol{\alpha}, x; N)|.$$

The term $N^d x$ is the expected value of the number of terms in the sequence $\left\{\sum_{1 \leq i \leq d} k_i \alpha_i\right\}_{\substack{1 \leq k_i \leq N \\ 1 \leq i \leq d}}$ whose fractional parts visit the interval $[0, x)$.

The main result of this chapter is the following:

Theorem 3.2. *Let $\varphi(n)$ be an arbitray positive increasing function of n, then for almost every $\boldsymbol{\alpha} \in \mathbb{R}^d$, we have:*

$$\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} < \infty \;\Rightarrow\; \Delta(\boldsymbol{\alpha}; N) \ll (\ln N)^d \cdot \varphi^{\max\{3,d\}}(\ln \ln N), \tag{3.2}$$

$$\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} = \infty \;\Rightarrow\; \Delta(\boldsymbol{\alpha}; N) > (\ln N)^d \cdot \varphi(\ln\ln N) \;\; i.o. \;, \tag{3.3}$$

where i.o. stands for infinitely often for N, and the constant in (3.2) *depends on* $\boldsymbol{\alpha}$.

Remark. *Our result successfully keep the main factor* $(\ln N)^d$*, but due to an absence of a second moment estimation as in Beck*[Bec94]*, we used an* L_1 *estimation, and the additional factor* $\varphi(\ln\ln N)^{\max\{2,d-1\}}$ *is needed in our proof for controlling the small divisor* $n\prod_{i=1}\|n\alpha_i\|$.

This chapter is organized as the following: in Section 2, for the convenience of following estimations, we transform the ergodic discrepancy to the Fourier series using Poisson's summation formula. In Section 3 we estimate the contribution of the "tail" of the Fourier series, i.e. the high frequency modes. Section 4 and 5 deal with the main part of the discrepancy, Section 4 is about the constant part and Section 5 deals with the exponential part, both of which will be properly defined later. Combining the section 3-5, we have an overall estimation of the discrepancy, which gives the desired result above.

3.2 Poisson's summation formula

In this section, following [Bec94], we use the Poisson formula to transform the ergodic discrepancy into a Fourier series. The main result of this section, Proposition 3.1, gives a better version of the Fourier series by taking a roof-like average, which has better convergent properties. First we adopt a heuristic way to obtain a *formal* Fourier series of the ergodic discrepancy. By the Poisson formula, without considering possible problems with convergence, we can write the ergodic sum $D(\boldsymbol{\alpha}, x; N)$ as the following formal series:

$$D(\boldsymbol{\alpha}, x; N) = \mathbf{i}^{d+1} \sum_{\mathbf{n}\in\mathbb{Z}^{d+1}\setminus\{0\}} \frac{1-e^{2\pi\mathbf{i}n_1 x}}{2\pi n_1} \prod_{i=1}^{d} \frac{1-e^{-2\pi\mathbf{i}N(n_1\alpha_i - n_{i+1})}}{2\pi(n_1\alpha_i - n_{i+1})},$$

where the term $(1-e^{2\pi\mathbf{i}n_1 x})/(2\pi n_1)$ is interpreted as $-\mathbf{i}x$ when $n_1 = 0$, and similarly $(1-e^{-2\pi\mathbf{i}N(n_1\alpha_i - n_{i+1})})/(2\pi(n_1\alpha_i - n_{i+1}))$ is interpreted as $\mathbf{i}N$.

To see this, note that the condition: $0 \le \{\sum_{1\le i\le d} k_i\alpha_i\} < x$ is equivalent to $\exists m \in \mathbb{Z}$ such that:

$$\begin{array}{c} 0 \le \sum\limits_{1\le i\le d} k_i\alpha_i - m < x; \\ 1 \le k_i \le N; \end{array},$$

Consider the lattice inside $\mathbb{R}^{d+1}$,

$$\left\{ \left(\sum_{1\le i\le d} k_i\alpha_i - m, k_1, \ldots, k_d \right) \;\middle|\; (k_1, \ldots, k_d, m) \in \mathbb{Z}^{d+1} \right\},$$

note that the fractional part of $\sum_{1\le i\le d} k_i\alpha_i$ lying in $[0,x)$ is equivalent to the lattice point inside box

$$B = [0,x) \times \prod_{\text{d copies}} (0,N].$$

So the sum becomes

$$\sum_{\substack{1\le k_i\le N\\ 1\le i\le d}} \chi_{[0,x)}\left(\sum_{1\le i\le d} k_i\alpha_i \mod 1\right) = \sum_{(k_1,\dots,k_d,m)\in\mathbb{Z}^{d+1}} \chi_B\left(\sum_{1\le i\le d} k_i\alpha_i - m, k_1,\dots,k_d\right)$$

On the other hand, we see that:

$$\sum_{1\le i\le d} k_i\alpha_i - m = \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_d & -1\\ 1 & 0 & \dots & 0 & 0\\ 0 & 1 & \dots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} k_1\\ k_2\\ \vdots\\ k_d\\ m \end{pmatrix} = \mathbf{A}\cdot\mathbf{y}$$

where

$$\mathbf{A} = \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_d & -1\\ 1 & 0 & \dots & 0 & 0\\ 0 & 1 & \dots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} k_1\\ k_2\\ \vdots\\ k_d\\ m \end{pmatrix}$$

Apply the Poisson formula to the function $f(\mathbf{y}) = \chi_B(\mathbf{A}\cdot\mathbf{y}),\ \mathbf{y} = (k_1,\dots,k_d,m) \in \mathbb{Z}^{d+1}$, we have

$$\begin{aligned}
\sum_{\mathbf{y}\in\mathbb{Z}^{d+1}} \chi_B(\mathbf{A}\cdot\mathbf{y}) &= \sum_{\mathbf{y}\in\mathbb{Z}^{d+1}} f(\mathbf{y})\\
&= \sum_{\nu\in\mathbb{Z}^{d+1}} \int_{\mathbb{R}^{d+1}} f(\mathbf{y})\cdot e^{-2\pi\mathbf{i}\nu\cdot\mathbf{y}} d\mathbf{y}\\
&= \sum_{\nu\in\mathbb{Z}^{d+1}} \int_{\mathbb{R}^{d+1}} \chi_B(\mathbf{A}\cdot\mathbf{y})\cdot e^{-2\pi\mathbf{i}\nu\cdot\mathbf{y}} d\mathbf{y}\\
&= \frac{1}{\det\mathbf{A}} \sum_{\nu\in\mathbb{Z}^{d+1}} \int_{\mathbb{R}^{d+1}} \chi_B(\mathbf{z}) e^{-2\pi\mathbf{i}\nu\cdot(\mathbf{A}^{-1}\cdot\mathbf{z})} d\mathbf{z}
\end{aligned}$$

Integrate with respect to each coordinate of $\mathbf{z}$ and we get

$$\begin{aligned} D(\boldsymbol{\alpha}, x; N) &= \sum_{\mathbf{y}\in\mathbb{Z}^{d+1}} f(\mathbf{y}) - N^d x \\ &= \mathbf{i}^{d+1} \sum_{\mathbf{n}\in\mathbb{Z}^{d+1}\setminus\{0\}} \frac{1-e^{2\pi\mathbf{i}n_1 x}}{2\pi n_1} \prod_{i=1}^{d} \frac{1-e^{-2\pi\mathbf{i}N(n_1\alpha_i - n_{i+1})}}{2\pi(n_1\alpha_i - n_{i+1})}. \end{aligned}$$

In fact, in order to avoid technical problems with the convergence, we will not study $D(\boldsymbol{\alpha}, x; N)$ directly, instead, we will follow Beck[Bec94] and use a special weighted average of $D(\boldsymbol{\alpha}, x; N)$ over a $\frac{1}{N^d}$ neighborhood. To this end, we oscillate the target interval $[0, x)$ with an amplitude of $\frac{1}{N^d}$, and also the range for summation $\{1, \ldots, N\}$ with amplitude of 2. Specifically, let $\mathbf{u} = (u_2, \ldots, u_{d+1})$, $\mathbf{u} + N = (u_2 + N, \ldots, u_{d+1} + N)$, and define :

$$D(\boldsymbol{\alpha}; a, b; \mathbf{u}, \mathbf{u}+N) = \sum_{\substack{u_{i+1}<k_i<N+u_{i+1} \\ 1\le i\le d}} \chi_{[a,b)}\left(\sum_{1\le i\le d} k_i\alpha_i \mod 1\right) - N^d(b-a) \tag{3.4}$$

and define the $\frac{1}{N^d}$ average:

$$\begin{aligned} \bar{D}(\boldsymbol{\alpha}, x; N) = \frac{N^d}{2}\left(\frac{1}{2}\right)^d \int_{-\frac{2}{N^d}}^{\frac{2}{N^d}} \int_{-2}^{2} \cdots \int_{-2}^{2} \left(1 - \frac{N^d}{2}|u_1|\right) \prod_{i=1}^{d} \left(1 - \frac{|u_{i+1}|}{2}\right) \\ \times D(\boldsymbol{\alpha}; u_1, x+u_1; \mathbf{u}, \mathbf{u}+N) du_1 du_2 \ldots du_{d+1}. \end{aligned}$$

Using the Féjer kernel identity

$$\frac{N^d}{2}\int_{-\frac{2}{N^d}}^{\frac{2}{N^d}} \left(1 - \frac{N^d|y|}{2}\right) e^{2\pi\mathbf{i}ky} dy = \left(\frac{\sin(2\pi\frac{k}{N^d})}{2\pi\frac{k}{N^d}}\right)^2,$$

$$\frac{1}{2}\int_{-2}^{2} \left(1 - \frac{|y|}{2}\right) e^{2\pi\mathbf{i}ky} dy = \left(\frac{\sin 2\pi k}{2\pi k}\right)^2.$$

We arrive at:

$$\begin{aligned} \bar{D}(\boldsymbol{\alpha}, x; N) = \mathbf{i}^{d+1} \sum_{\mathbf{n}\in\mathbb{Z}^{d+1}} \frac{1-e^{2\pi\mathbf{i}n_1 x}}{2\pi n_1} \left(\frac{\sin 2\pi\left(\frac{n_1}{N^d}\right)}{2\pi\left(\frac{n_1}{N^d}\right)}\right)^2 \\ \cdot \prod_{i=1}^{d} \left(\frac{1-e^{-2\pi\mathbf{i}N(n_1\alpha_i - n_{i+1})}}{2\pi(n_1\alpha_i - n_{i+1})}\right) \left(\frac{\sin 2\pi(n_1\alpha_i - n_{i+1})}{2\pi(n_1\alpha_i - n_{i+1})}\right)^2. \end{aligned} \tag{3.5}$$

We claim that if $\bar{D}(\boldsymbol{\alpha}, x; N)$ is bounded from above by $(\ln N)^d \varphi^{d+1}(\ln\ln N)$, then so is $D(\boldsymbol{\alpha}, x; N)$.

Proposition 3.1. *For almost every $\boldsymbol{\alpha} \in [0,1)^d$, we have*

$$|\bar{D}(\boldsymbol{\alpha}, x; N) - D(\boldsymbol{\alpha}, x; N)| \ll (\ln N)^{1+\epsilon},$$

Proof. Note that we have the following identity

$$\mathrm{Leb}\left\{\boldsymbol{\alpha} \in [0,1)^d \;\middle|\; \left\|\sum_{1\le i\le d} k_i\alpha_i\right\| \le \frac{1}{\left(\sum\limits_{1\le i\le d} k_i^2\right)^{d/2}\left(\ln\left(\sum\limits_{1\le i\le d} k_i^2\right)\right)^{1+\epsilon}}\right\}$$
$$=\frac{2}{\left(\sum\limits_{1\le i\le d} k_i^2\right)^{d/2}\left(\ln\left(\sum\limits_{1\le i\le d} k_i^2\right)\right)^{1+\epsilon}}.$$

Since the right hand side of the identity above is summable over $\mathbf{k} \in \mathbb{Z}^d$ such that $\mathbf{k} > 1$, by a standard application of Borel-Cantelli Lemma, for almost every $\boldsymbol{\alpha} \in [0,1)^d$, there exists a constant $C(\boldsymbol{\alpha}) > 0$, and every $\mathbf{k} = (k_1, \ldots, k_d) \in \mathbb{Z}^d$ such that for $|\mathbf{k}| > 1$, we have:

$$\left\|\sum_{1\le i\le d} k_i\alpha_i\right\| > \frac{C(\boldsymbol{\alpha})}{\left(\sum\limits_{1\le i\le d} k_i^2\right)^{d/2}\left(\ln\left(\sum\limits_{1\le i\le d} k_i^2\right)\right)^{1+\epsilon}}.$$

Therefore, $\|\sum_{1\le i\le d} k_i\alpha_i\| \gg 1/(N^d(\ln N)^{1+\epsilon})$ for $|k| > 1$ such that $-2 \le k_1, \ldots, k_d \le N+2$. Since the numbers of elements of the sequence $\{\sum_{1\le i\le d} k_i\alpha_i\}$ which lie inside $[0, x)$ and $[-2/N^d, x+2/N^d)$ differ only by the number of elements which lie inside the intervals $[-2/N^d, 0)$ and $[x, x+2/N^d)$, and the "gap" between two different elements of the sequence is at least $1/(N^d(\ln N)^{1+\epsilon})$, by using the Dirichlet principle, we have

$$D(\boldsymbol{\alpha}; u_1, x+u_1; \mathbf{u}, \mathbf{u}+N) = D(\boldsymbol{\alpha}, x; N) + O\left((\ln N)^{1+\epsilon}\right),$$

for almost every $\boldsymbol{\alpha} \in [0,1)^d$, and every $u_1 \in [-2/N^d, 2/N^d]$, $u_2, u_3 \in [-2, 2]$. The constant above only depend on $\boldsymbol{\alpha}$ and ϵ, therefore the same bound hold after integration over u_1 and $\mathbf{u}$. □

By Proposition 3.1, we can now shift our attention to the asymptotic behavior of $\bar{D}(\boldsymbol{\alpha}, x; N)$, which has better convergence property as shown in Section 3.

3.3 Estimating the "tail" of the discrepancy function

This section is devoted to estimating the "tail" of $\bar{D}(\boldsymbol{\alpha}, x; N)$, note that $\bar{D}(\boldsymbol{\alpha}, x; N)$ is a sum of the products (where $\mathbf{n} \in \mathbb{Z}^{d+1} \backslash \mathbf{0}$)

$$\begin{aligned} f(\mathbf{n}, x, \boldsymbol{\alpha}) =& \mathbf{i}^{d+1} \frac{1 - e^{2\pi \mathbf{i} n_1 x}}{2\pi n_1} \left(\frac{\sin 2\pi \left(\frac{n_1}{N^d}\right)}{2\pi \left(\frac{n_1}{N^d}\right)} \right)^2 \\ & \cdot \prod_{i=1}^{d} \left(\frac{1 - e^{-2\pi \mathbf{i} N(n_1 \alpha_i - n_{i+1})}}{2\pi(n_1 \alpha_i - n_{i+1})} \right) \left(\frac{\sin 2\pi(n_1 \alpha_i - n_{i+1})}{2\pi(n_1 \alpha_i - n_{i+1})} \right)^2 \end{aligned} \tag{3.6}$$

let:

$$\bar{D}_4(\boldsymbol{\alpha}, x; N) = \sum_{\mathbf{n} \in U_4(\boldsymbol{\alpha}; N)} f(\mathbf{n}, x, \boldsymbol{\alpha}),$$

where

$$U_4(\boldsymbol{\alpha}; N) = \left\{ \mathbf{n} \in \mathbb{Z}^{d+1} \left| \begin{array}{l} 1 < |n_1| < N^d/4, \\ |n_1 \alpha_i - n_{i+1}| = \|n_1 \alpha_i\|, \ 1 \le i \le d \\ |n_1| \prod_{i=1}^{d} \|n_1 \alpha_i\| > (\ln N)^s \end{array} \right. \right\} \tag{3.7}$$

$\|\cdot\|$ denotes the distance to the closest integer, and s is a large enough but fixed integer to be determined later(see Lemma 3.7).

The main result of this section is the following:

Proposition 3.2. *Let $\varphi(n)$ be a positive increasing function such that $\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} < \infty$, then for almost every $\boldsymbol{\alpha} \in [0,1]^d$, we have*

$$|\bar{D}_4(\boldsymbol{\alpha}, x; N) - \bar{D}(\boldsymbol{\alpha}, x, N)| \ll (\ln N)^d \varphi^{\max\{3,d\}}(\ln \ln N)$$

To prove Proposition 3.2, we need to control the components of the difference $\bar{D}_4 - \bar{D}$ step by step.

3.3.1 Estimation for the sum when $|n_1|$ is large.

First we will show that the sum of the terms when $|n_1|$ is large does not contribute much. Define

$$\bar{D}_1(\boldsymbol{\alpha}, x; N) = \sum_{\mathbf{n} \in U_1(N)} f(\mathbf{n}, x, \boldsymbol{\alpha}), \tag{3.8}$$

where

$$U_1(N) = \{\mathbf{n} \in \mathbb{Z}^{d+1} \backslash \{\mathbf{0}\} : |n_1| < N^d (\ln N)^d\}$$

We show the following:

Proposition 3.3. *For almost every $\boldsymbol{\alpha} \in \mathbb{R}^d$ we have*

$$\left|\bar{D}_1(\boldsymbol{\alpha}, x; N) - \bar{D}(\boldsymbol{\alpha}, x; N)\right| = \mathcal{O}(1),$$

where $\mathcal{O}(1)$ represents an absolute bound which may depend on $\boldsymbol{\alpha}$, but does not depend on x or N.

Proof. We will bound individual terms in $f(\mathbf{n}, \mathbf{x}, \boldsymbol{\alpha})$ with different methods according to their magnitudes.

When $|n_1\alpha_i - n_{i+1}| < 1/2$ or $|n_1/N^d| < 1$, i.e., *small*, we use the bound

$$\left|\frac{\sin(2\pi(n_1\alpha_i - n_{i+1}))}{2\pi(n_1\alpha_i - n_{i+1})}\right| \leq 1.$$

and

$$\left|\frac{\sin(2\pi(\frac{n_1}{N^d})}{2\pi(\frac{n_1}{N^d})}\right| \leq 1.$$

When $|n_1\alpha_i - n_{i+1}| \geq 1/2$ or $|n_1/N^d| > 1$, i.e., *large*, we bound the numerator by 1, that is,

$$\left|\frac{\sin(2\pi(n_1\alpha_i - n_{i+1}))}{2\pi(n_1\alpha_i - n_{i+1})}\right| \leq \frac{1}{2\pi|n_1\alpha_i - n_{i+1}|}.$$

and

$$\left|\frac{\sin(2\pi(\frac{n_1}{N^d}))}{2\pi(\frac{n_1}{N^d})}\right| \leq \frac{1}{2\pi|\frac{n_1}{N^d}|}.$$

By using the above inequalities, we can bound the partial sums where the terms have zero or one *large* factor of $|n_1\alpha_i - n_{i+1}|$. For partial sums where the terms have two or more *large* factors of $|n_1\alpha_i - n_{i+1}|$, they are even smaller and thus can be bounded by the partial sums with only one *large* factor of $|n_1\alpha_i - n_{i+1}|$, the same way as $\bar{D}_{1,j} \ll \bar{D}_{1,0}$ as shown with details in the inequality (3.10).

we have the following upper bound, (see the explanation after the definition of $U_{1,j}$):

$$\left|\bar{D}_1(\boldsymbol{\alpha}, x; N) - \bar{D}(\boldsymbol{\alpha}, x; N)\right| \ll \bar{D}_{1,0} + \sum_{j=1}^{d} \bar{D}_{1,j} \tag{3.9}$$

where

$$\bar{D}_{1,0} = \sum_{\mathbf{n} \in U_{1,0}(\boldsymbol{\alpha}; N)} \frac{N^{2d}}{|n_1|^3} \prod_{i=1}^{d} \frac{1}{|n_1\alpha_i - n_{i+1}|}$$

where

$$U_{1,0}(\boldsymbol{\alpha}; N) = \left\{\mathbf{n} \in \mathbb{Z}^{d+1} : |n_1| > N^d(\ln N)^d, \quad |n_1\alpha_i - n_{i+1}| < \frac{1}{2}, \quad 1 \leq i \leq d\right\},$$

and

$$\bar{D}_{1,j} = \sum_{\mathbf{n}\in U_{1,j}(\alpha;N)} \frac{N^{2d}}{|n_1|^3} \prod_{\substack{1\le i\le d\\ i\neq j}} \frac{1}{|n_1\alpha_i - n_{i+1}|} \cdot \frac{1}{|n_1\alpha_j - n_{j+1}|^{d+1}},$$

where

$$U_{1,j}(\boldsymbol{\alpha};N) = \left\{\mathbf{n}\in\mathbb{Z}^{d+1} : |n_1| > N^d(\ln N)^d, \quad |n_1\alpha_i - n_{i+1}| < \frac{1}{2},\ i\neq j, \quad |n_1\alpha_j - n_{j+1}| > \frac{1}{2}\right\}.$$

Define $r_{\alpha_i} = \lceil |n_1\alpha_i - n_{i+1}| \rceil$, $1\le i\le d$, we have that for $1\le j\le d$,

$$\bar{D}_{1,j} \ll \sum_{r_{\alpha_j}=1}^{\infty} \sum_{\substack{\mathbf{n}\in\mathbb{Z}^{d+1}:\\ |n_1|>N^d(\ln N)^d\\ |n_1\alpha_i - n_{i+1}|<1/2\\ i\neq j}} \frac{N^{2d}}{|n_1|^3} \prod_{\substack{1\le i\le d\\ i\neq j}} \frac{1}{|n_1\alpha_i - n_{i+1}|} \frac{1}{r_{\alpha_j}^3} \ll \bar{D}_{1,0}, \tag{3.10}$$

Therefore it suffices to show that $\bar{D}_{1,0} = \mathcal{O}(1)$, we will first prove a lemma:

Lemma 3.1. *Let $\varphi(n)$ be a positive increasing function such that $\sum_{n=1}^{\infty}\frac{1}{\varphi(n)} < \infty$, then for almost every $\boldsymbol{\alpha}\in\mathbb{R}^d$, the sum*

$$\sum_{n=2}^{\infty} \frac{1}{n\varphi(\ln n)\prod_{i=1}^d (\|n\alpha_i\|\varphi(|\ln\|n\alpha_i\||))}$$

converges.

Proof. Note that the integral

$$J(n) = \int_{[0,1]^d} \frac{1}{\prod_{i=1}^d (\|n\alpha_i\|\varphi(|\ln\|n\alpha_i\||))} d\boldsymbol{\alpha}$$

is finite and independent of n, the series $\sum\limits_{n=2}^{\infty}\frac{1}{n\varphi(\ln n)}$ also converges, so the integral

$$\int_{[0,1]^d} \left(\sum_{n=2}^{\infty} \frac{1}{n\varphi(\ln n)\prod_{i=1}^d (\|n\alpha_i\|\varphi(|\ln\|n\alpha_i\||))}\right) d\boldsymbol{\alpha} = \sum_{n=2}^{\infty}\frac{J(n)}{n\varphi(\ln n)} \tag{3.11}$$

is finite. Therefore we have that the series

$$\sum_{n=2}^{\infty} \frac{1}{n\varphi(\ln n)\prod_{i=1}^d (\|n\alpha_i\|\varphi(|\ln\|n\alpha_i\||))}$$

converges for almost every $\boldsymbol{\alpha}\in[0,1]^d$ and ever $\epsilon>0$, the periodicity of $\|\cdot\|$ gives the result for almost every $\boldsymbol{\alpha}\in\mathbb{R}^d$. □

Continuing with the proof for Proposition 3.3, since $\sum_{n=1}^{\infty} 1/n^2 = \mathcal{O}(1)$, by Borel-Cantelli Lemma, we have for almost every $\boldsymbol{\alpha} \in \mathbb{R}^d$,

$$\|n\alpha_i\| \geq \frac{1}{n^2}, \quad 1 \leq i \leq d,$$

for all but finitely many $n \in \mathbb{N}$, or equivalently,

$$|\ln \|n\alpha_i\|| \leq 2 \ln n, \quad 1 \leq i \leq d, \tag{3.12}$$

for all but finitely many $n \in \mathbb{N}^+$.

Take $\varphi(n) = n^{1+\epsilon}$ where $\epsilon > 0$ is small, then $\bar{D}_{1,0}$ can be estimated as follows:

$$\begin{aligned}
\bar{D}_{1,0} &= \sum_{n_1 > N^d (\ln N)^d} \frac{1}{(\frac{n_1}{N^d})^2} \frac{1}{n_1 \prod_{i=1}^d \|n_1\alpha_i\|} \\
&\ll \sum_{n \geq N^d (\ln N)^d} \frac{1}{(\ln n)^{(d+1)(1+\epsilon)}} \frac{1}{n \prod_{i=1}^d \|n\alpha_i\|} \\
&\ll \sum_{n \geq N^d (\ln N)^d} \frac{1}{|n| \varphi(\ln n) \prod_{i=1}^d (\|n\alpha_i\| \varphi(|\ln \|n\alpha_i\||))} \\
&\leq \sum_{n=2}^{\infty} \frac{1}{|n| \varphi(\ln n) \prod_{i=1}^d (\|n\alpha_i\| \varphi(|\ln \|n\alpha_i\||))} \\
&= \mathcal{O}(1),
\end{aligned}$$

finishing the proof. □

3.3.2 Estimation for the sum when $|n_1\alpha_i - n_{i+1}|$ is larger than $1/2$ for one of $1 \leq i \leq d$.

Let

$$\bar{D}_2(\boldsymbol{\alpha}, x; N) = \sum_{\mathbf{n} \in U_2(\boldsymbol{\alpha}; N)} f(\mathbf{n}, x, \boldsymbol{\alpha})$$

where $\mathbf{n} = (n_1, \ldots, n_{d+1})$ satisfies

$$U_2(\boldsymbol{\alpha}; N) = \left\{ \mathbf{n} \in \mathbb{Z}^{d+1} \backslash \{\mathbf{0}\} \;\middle|\; \begin{array}{l} |n_1| < N^d (\ln N)^d, \\ |n_1\alpha_i - n_{i+1}| = \|n_1\alpha_i\|, \; 1 \leq i \leq d \end{array} \right\} \tag{3.13}$$

We show that $\bar{D}_1$ can be replaced by $\bar{D}_2$:

Proposition 3.4. *Let $\varphi(n)$ be a positive increasing function such that $\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} < \infty$, then for almost every $\boldsymbol{\alpha} \in [0,1]^d$, we have*

$$\left| \bar{D}_2(\boldsymbol{\alpha}, x; N) - \bar{D}_1(\boldsymbol{\alpha}, x; N) \right| \ll (\ln N)^d \varphi^d (\ln \ln N)$$

Proof. Since we can bound those $\mathbf{n} \in U_1(\boldsymbol{\alpha}; N)$ for which $|n_1\alpha_i - n_{i+1}| \geq 1/2$ holds for two or more indices i by those $\mathbf{n} \in U_1(\boldsymbol{\alpha}; N)$ for which $|n_1\alpha_i - n_{i+1}| \geq 1/2$ holds for just one index, as in (3.10), we have the following inequality:

$$\begin{aligned}\left|\bar{D}_2(\boldsymbol{\alpha}, x; N) - \bar{D}_1(\boldsymbol{\alpha}, x; N)\right| &\ll \sum_{j=1}^{d} \sum_{\mathbf{n} \in U_{2,j}} \frac{1}{|n_1|} \prod_{\substack{1 \leq i \leq d \\ i \neq j}} \frac{1}{|n_1\alpha_i - n_{i+1}|} \cdot \frac{1}{|n_1\alpha_j - n_{j+1}|^3} \\ &\ll \sum_{j=1}^{d} \sum_{n=1}^{N^d(\ln N)^d} \frac{1}{n \prod_{i \neq j} \|n\alpha_i\|},\end{aligned} \tag{3.14}$$

where

$$U_{2,j} = \left\{ |n_1| \leq N^d(\ln N)^d, \quad |n_1\alpha_i - n_{i+1}| < \frac{1}{2},\ i \neq j, \quad |n_1\alpha_j - n_{j+1}| \geq \frac{1}{2} \right\}.$$

The proposition follows from the following lemma:

Lemma 3.2. *Let $\varphi(n)$ be an arbitrary positive increasing function of n with $\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} < \infty$, then for almost every $\boldsymbol{\alpha} \in \mathbb{R}^d$ and every $1 \leq j \leq d$, we have*

$$\sum_{n=1}^{N^d(\ln N)^d} \frac{1}{n \prod_{i \neq j} \|n\alpha_i\|} \ll (\ln N)^d \varphi^d(\ln \ln N),$$

Proof. It suffices to prove for every $1 \leq j \leq d$ the bound above holds, we employ the same technique as in Lemma 3.1. Denote $d\boldsymbol{\alpha}_j = d\alpha_1 \cdots \hat{d\alpha_j} \cdots d\alpha_d$, where $d\alpha_j$ is omitted. Note that the integral

$$J(n) = \int_{[0,1]^{d-1}} \frac{1}{\prod_{i \neq j} \left(\|n\alpha_i\| |\ln \|n\alpha_i\|| \varphi(|\ln \ln \|n\alpha_i\||)\right)} d\alpha_1 \cdots \hat{d\alpha_j} \cdots d\alpha_d$$

is finite and independent of n, the series $\sum\limits_{n=2}^{\infty} \frac{1}{n(\ln n)\varphi(\ln \ln n)}$ also converges, so the integral

$$\begin{aligned}&\int_{[0,1]^{d-1}} \left(\sum_{n=2}^{\infty} \frac{1}{n(\ln n)\varphi(\ln \ln n) \prod_{i \neq j} \left(\|n\alpha_i\| |\ln \|n\alpha_i\|| \varphi(|\ln \ln \|n\alpha_i\||)\right)} \right) d\boldsymbol{\alpha}_j \\ &= \sum_{n=2}^{\infty} \frac{J(n)}{n(\ln n)\varphi(\ln \ln n)}\end{aligned} \tag{3.15}$$

is finite. Therefore we have that the series

$$\sum_{n=2}^{\infty} \frac{1}{n(\ln n)\varphi(\ln \ln n) \prod_{i \neq j} \left(\|n\alpha_i\| |\ln \|n\alpha_i\|| \varphi(|\ln \ln \|n\alpha_i\||)\right)}$$

converges for almost every $\boldsymbol{\alpha} \in [0,1]^d$ and ever $\epsilon > 0$, the periodicity of $\|\cdot\|$ gives the

convergence for almost every $\boldsymbol{\alpha} \in \mathbb{R}^d$. Using the inequality (3.12), we have

$$\begin{aligned}
&\sum_{n=1}^{N^d(\ln N)^d} \frac{1}{n\prod_{i\neq j}\|n\alpha_i\|} \\
&\ll \left(\sum_{n=2}^{\infty} \frac{(\ln N)^d\varphi(\ln\ln N)^d}{n(\ln n)\varphi(\ln\ln n)\prod_{i\neq j}\left(\|n\alpha_i\|\,|\ln\|n\alpha_i\||\varphi(|\ln\ln\|n\alpha_i\||)\right)}\right) \\
&\ll (\ln N)^d\varphi(\ln\ln N)^d.
\end{aligned}$$

□

Since the sum in (3.14) is a finite sum over $1 \le j \le d$, the proof for the proposition is completed. □

3.3.3 Estimation for the sum when $|n_1|\prod_{i=1}^d \|n_1\alpha_i\|$ is small.

Define:

$$\bar{D}_3(\boldsymbol{\alpha}, x; N) = \sum_{\mathbf{n}\in U_3(\boldsymbol{\alpha}:N)} f(\mathbf{n}, x, \boldsymbol{\alpha}) \tag{3.16}$$

where

$$U_3(\boldsymbol{\alpha}; N) = \left\{ \mathbf{n} \in \mathbb{Z}^{d+1}\backslash\{\mathbf{0}\} \;\middle|\; \begin{array}{l} 1 \le |n_1| \le N^d(\ln N)^d, \\ |n_1|\displaystyle\prod_{i=1}^{d}\|n_1\alpha_i| > (\ln N)^s, \\ |n_1\alpha_i - n_{i+1}| = \|n_1\alpha_i\|,\ 1 \le i \le d \end{array} \right\}. \tag{3.17}$$

Note that in (3.17), $\{n_{i+1}\}_{1\le i\le d}$ are respectively the closest integers to $\{n_1\alpha_i\}_{1\le i\le d}$, therefore in the sequel we only need to discuss n_1 instead of $\mathbf{n} \in \mathbb{Z}^{d+1}$. The main result of this step is the following:

Proposition 3.5. *Let $\varphi(n)$ be an arbitrary positive increasing function of n with $\sum_{n=1}^{\infty}\frac{1}{\varphi(n)} < \infty$, then for almost every $\boldsymbol{\alpha}$, we have*

$$\left|\bar{D}_3(\boldsymbol{\alpha}, x; N) - \bar{D}_2(\boldsymbol{\alpha}, x; N)\right| \ll (\ln N)^d\varphi^3(\ln\ln N).$$

First we need some preparation lemmas about the lower bound of the small divisors.

Lemma 3.3. *Let $\varphi(n)$ be an arbitrary positive increasing function of n with $\sum_{n=1}^{\infty}\frac{1}{\varphi(n)} < \infty$, then for almost every $\boldsymbol{\alpha} \in [0,1]^d$, we have for every $|n| > 1$,*

$$|n|\prod_{i=1}^{d}\|n\alpha_i\| \gg \frac{1}{(\ln|n|)^d\varphi(\ln\ln|n|)}$$

and

$$\|n\alpha_i\| \gg \frac{1}{|n|\varphi(\ln|n|)}, \quad 1 \le i \le d,$$

with a constant that depends on $\boldsymbol{\alpha} \in [0,1]^d$.

Proof. For $\epsilon > 0$ fixed, define

$$E_n := \{\boldsymbol{\alpha} \in [0,1]^d : |n| \prod_{i=1}^{d} \|n\alpha_i\| < 1/(\ln|n|)^d \varphi(\ln\ln|n|)\},$$

by direct calculation, we have:

$$Leb(E_n) \le \frac{2}{n(\ln n)\varphi(\ln\ln|n|)}$$

therefore $\sum\limits_{n\in\mathbb{Z}\setminus\{0\}}^{\infty} Leb(E_n) < \infty$, by Borel-Cantelli Lemma, we have that for for almost every $\boldsymbol{\alpha}$,

$$|n| \prod_{i=1}^{d} \|n\alpha_i\| \gg 1/(\ln|n|)^d \varphi(\ln\ln|n|)$$

with a constant that depends on $\boldsymbol{\alpha} \in [0,1]^d$. The proof for $|n|\|n\alpha\|$ is similar. □

Proposition 3.5 follows easily from the following lemma:

Lemma 3.4. *Let* $\varphi(n)$ *be an arbitrary positive increasing function of* n *with* $\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} < \infty$*, then for almost every* $\boldsymbol{\alpha}$*, and* $s \in \mathbb{N}$ *fixed, we have the following estimation:*

$$\sum_{\substack{1 \le n \le N^d(\ln N)^d \\ n\prod_{i=1}^{d}\|n\alpha_i\| < (\ln N)^s}} \frac{1}{n\prod_{i=1}^{d}\|n\alpha_i\|} \ll (\ln N)^d \varphi^3(\ln\ln N) \tag{3.18}$$

where the inequality holds for a constant that depends on s, $\boldsymbol{\alpha}$.

Proof. Divide n in discussion into the following sets:

$$S_{\boldsymbol{\alpha}}(p,v) = \left\{ e^{e^{p-1}} \le n < e^{e^p} \,\middle|\, \frac{2^{v-1}}{(\ln n)^d \varphi(\ln\ln|n|)} \le n\prod_{i=1}^{d}\|n\alpha_i\| < \frac{2^v}{(\ln n)^d\varphi(\ln\ln|n|)} \right\}$$

Where $0 \le p \le \ln\ln N + 1$, $v \ll \ln\ln N$. By Lemma 3.3, there exists a constant $K = K(\boldsymbol{\alpha},\varphi) > 0$, such that $v \ge -K$. We prove a sublemma for the upper bound of the number of elements inside each $S_{\boldsymbol{\alpha}}(p,v)$:

Sublemma 3.5. *Let* $\varphi(n)$ *be an arbitrary positive increasing function of* n *with* $\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} < \infty$*, then for almost every* $\boldsymbol{\alpha}$*, and every* $p > 1, v \in \mathbb{Z}$*,*

$$\#S_{\boldsymbol{\alpha}}(p,v) \ll 2^v \varphi(|v|), \tag{3.19}$$

with a constant that depends only on $\boldsymbol{\alpha}$, φ, *where* $|v|$ *should be understood as* $\max\{|v|, 1\}$ *for* $\phi(|v|)$ *to be well defined.*

Proof. Define the corresponding set

$$C(n,v) = \left\{ \boldsymbol{\alpha} \in [0,1]^d \;\middle|\; \frac{2^{v-1}}{(\ln n)^d \varphi(\ln\ln|n|)} \leq n \prod_{i=1}^{d} \|n\alpha_i\| < \frac{2^v}{(\ln n)^d \varphi(\ln\ln|n|)} \right\},$$

we have :

$$Leb(C(n,v)) \ll \frac{2^v}{n(\ln n)\varphi(\ln\ln n)},$$

and

Since

$$\#S_{\boldsymbol{\alpha}}(p,v) = \sum_{e^{e^{p-1}} \leq n < e^{e^p}} \chi_{C(n,v)}(\boldsymbol{\alpha})$$

We have:

$$\mathbb{E}(\#S_{\boldsymbol{\alpha}}(p,v)) = \sum_{e^{e^{p-1}} \leq n < e^{e^p}} Leb(C(n,v)) \ll \frac{2^v}{\varphi(p)}$$

and

$$Leb\{\#S_{\boldsymbol{\alpha}}(p,v) \geq 2^v \varphi(|v|)\} \leq \frac{\mathbb{E}(\#S_{\boldsymbol{\alpha}}(p,v))}{2^v \varphi(|v|)} \ll \frac{1}{\varphi(p)\varphi(|v|)}$$

So

$$\sum_{\substack{p>1 \\ v \in \mathbb{Z}}} Leb\{\#S_{\boldsymbol{\alpha}}(p,v) \geq 2^v \varphi(|v|)\} < \infty.$$

By Borel-Cantelli Lemma, we have that for almost every $\boldsymbol{\alpha}$,

$$\#S_{\boldsymbol{\alpha}}(p,v) \ll 2^v \varphi(v).$$

□

Continuing with the proof for Lemma 3.4, we have the following estimation:

$$\begin{aligned}
&\sum_{\substack{1 \le n \le N^d (\ln N)^d \\ n\prod_{i=1}^d \|n\alpha_i\| < (\ln N)^s}} \frac{1}{n\prod_{i=1}^d \|n\alpha_i\|} \\
&\ll \sum_{p=1}^{\ln\ln N+1} \sum_{-K \le v \ll \ln\ln N} \frac{(e^p)^d \varphi(p)}{2^v} \cdot 2^v \varphi(v) \\
&\ll \sum_{p=1}^{\ln\ln N+1} (e^p)^d \varphi(p) \sum_{-K \le v \ll \ln\ln N} \varphi(v) \\
&\ll (\ln N)^d \varphi(\ln\ln N) \cdot (\varphi(\ln\ln N)\ln\ln N) \\
&\ll (\ln N)^d \varphi^3(\ln\ln N).
\end{aligned} \tag{3.20}$$

Therefore Lemma 3.4 is proven. □

Proof for Proposition 3.5:

Note that

$$\left|\bar{D}_3(\boldsymbol{\alpha}, x; N) - \bar{D}_2(\boldsymbol{\alpha}, x; N)\right| \ll \sum_{\substack{1 \le n \le N^d (\ln N)^d \\ n \prod_{i=1}^d \|n\alpha_i\| < (\ln N)^s}} \frac{1}{n \prod_{i=1}^d \|n\alpha_i\|},$$

Proposition 3.5 follows from Lemma 3.4.

3.3.4 Control the sum when n_1 is between $N^d/4$ and $N^d(\ln N)^d$.

The goal of this step is to prove the following:

Proposition 3.6. *For almost every $\boldsymbol{\alpha}$, we have*

$$\left|\bar{D}_4(\boldsymbol{\alpha}, x; N) - \bar{D}_3(\boldsymbol{\alpha}, x; N)\right| \ll (\ln N)^d \ln\ln N.$$

With the proposition above, we could limit the range of $(2\pi n_1/N^d)$ in $f(\mathbf{n}, x, \boldsymbol{\alpha})$ to $(-\pi/2, \pi/2)$, which makes f better-behaved for later estimations.

The range for n_1 in the difference $\bar{D}_4 - \bar{D}_3$ can be decomposed as follows:

$$T_{\boldsymbol{\alpha}}(\mathbf{l}; N) = \left\{ n_1 \in \mathbb{Z} \;\middle|\; \begin{array}{l} 2^{l_1} \le |n_1| < 2^{l_1+1}, \\ 2^{-l_{i+1}} \le \|n_1\alpha_i\| < 2^{-l_{i+1}+1},\ 1 \le i \le d-1 \\ 2^{\sum_{i=1}^d l_{i+1} - l_1} \le \|n_1\alpha_d\| < 2^{\sum_{i=1}^d l_{i+1} - l_1 + 1} \end{array} \right\},$$

where l_1 defines the range for $|n_1|$, l_{i+1} defines the range for $\|n_1\alpha_i\|$, $1 \le i \le d-1$, and l_{d+1} defines the range for $|n_1| \prod_{i=1}^d \|n_1\alpha_i\|$. By Lemma 3.3, the range for $\mathbf{l} = (l_1, \ldots, l_{d+1}) \in L_1(N)$ is defined by:

$$L_1(N) := \left\{ \mathbf{l} \in \mathbb{Z}^{d+1} \;\middle|\; \begin{array}{l} 2\ln_2 N - 2 \le l_1 \le 2\ln_2 N + 2\ln_2 \ln N, \\ 2 \le l_{i+1} \ll \ln N,\ 1 \le i \le d-1, \\ \ln\ln N \ll l_{d+1} \le l_1 \ll \ln N, \end{array} \right\}. \tag{3.21}$$

We will first prove a uniform upper bound for $\#T_{\boldsymbol{\alpha}}(\mathbf{l}; N)$:

Lemma 3.6. *If $s > 3$, then for almost every $\boldsymbol{\alpha} \in \mathbb{R}^d$ and all the $\mathbf{l}$ in $L_1(N)$, we have*

$$\#T_{\boldsymbol{\alpha}}(\mathbf{l}; N) \ll 2^{l_{d+1}}$$

where the constant is uniform for $\mathbf{l}$ in $L_1(N)$.

Proof. The number of elements inside $T_{\boldsymbol{\alpha}}(\mathbf{l}; N)$ corresponds to the number of $n_1 \in [2^{l_1}, 2^{l_1+1})$, such that vector $n_1\boldsymbol{\alpha} \mod 1 = (n_1\alpha_1 \mod 1, \ldots, n_1\alpha_d \mod 1)$ visit one of the 2^d target

boxes:

$$\prod_{i=1}^{d-1} \pm[2^{-l_{i+1}}, 2^{-l_{i+1}+1}) \times \pm[2^{\sum_{i=1}^{d} l_{i+1}-l_1}, 2^{\sum_{i=1}^{d} l_{i+1}-l_1+1}),$$

where the negative signs deal with the case when $\frac{1}{2} \leq \{n_1\alpha_i\} < 1$. Therefore the expected value for $\#T_{\boldsymbol{\alpha}}(\mathbf{l}; N)$ is $\mathcal{O}(2^{l_{d+1}})$, from Theorem 3.1 we know that the error term for the cardinality is $\mathcal{O}((\ln 2^{l_1})^{2+\epsilon})$ for any target box, we have that

$$\#T_{\boldsymbol{\alpha}}(\mathbf{l}; N) = 2^{l_{d+1}} + \mathcal{O}(l_1^{2+\epsilon}).$$

Note that $l_1^{2+\epsilon} \ll (\ln N)^{2+\epsilon} < (\ln N)^s < 2^{l_{d+1}}$, the claim follows. □

Now we can estimate the contribution of the terms for which $N^d/4 \leq n_1 \leq N^d(\ln N)^d$.

Proof of Proposition 3.6

Proof. Inside each $T_{\boldsymbol{\alpha}}(\mathbf{l}; N)$, the divisors $n_1 \prod_{i=1}^{d} \|n\alpha_i\|$ is between $2^{l_{d+1}}$ and $2^{l_{d+1}+d+1}$, therefore

$$\begin{aligned}
&\left|\bar{D}_4 - \bar{D}_3\right| \\
&\leq \sum_{\mathbf{l}\in L_1(N)} \sum_{T_{\boldsymbol{\alpha}}(\mathbf{l};N)} \frac{1}{|n_1| \prod_{i=1}^{d} \|n_1\alpha_i\|} \\
&\leq \sum_{\mathbf{l}\in L_1(N)} \sum_{T_{\boldsymbol{\alpha}}(\mathbf{l};N)} \frac{1}{2^{l_{d+1}}} \\
&\ll \sum_{\mathbf{l}\in L_1(N)} \frac{1}{2^{l_{d+1}}} \#T_{\boldsymbol{\alpha}}(\mathbf{l}; N) \\
&\ll \sum_{\mathbf{l}\in L_1(N)} 1 \\
&\ll (\ln N)^d \ln\ln N.
\end{aligned}$$

The last inequality follows from the range of l_i's in $L_1(N)$ (see (3.21)), which gives the number of possible l_1's is $\mathcal{O}(\ln\ln N)$, and the number of possible l_i's are $\mathcal{O}(\ln N)$ for $2 \leq i \leq d+1$. □

Proof of Proposition 3.2

Proof. Combining Propositions 3.2-3.5, we have

$$\begin{aligned}
|\bar{D}_4 - \bar{D}| &\ll \mathcal{O}(1) + (\ln N)^d \varphi^d (\ln\ln N) + (\ln N)^d \varphi^3 (\ln\ln N) + (\ln N)^d \ln\ln N \\
&\ll (\ln N)^d \varphi^{\max\{d,3\}} (\ln\ln N).
\end{aligned}$$

□

3.4 Cancellation of the main terms.

To estimate the contribution of the remaining terms, we first decompose the product $f(\mathbf{n}, x, \boldsymbol{\alpha})$ into 2 distinct parts,

$$\bar{D}_4 = \frac{\mathbf{i}^{d+1}}{(2\pi)^{d+1}} \Bigg(\sum_{\mathbf{n} \in U_4(\boldsymbol{\alpha}; N)} \frac{1}{n_1 \prod_{i=1}^{d} (n_1 \alpha_i - n_{i+1})} \cdot g(\boldsymbol{n}, \boldsymbol{\alpha}; N) + \sum_{\mathbf{s}} \pm \sum_{\mathbf{n} \in U_4(\boldsymbol{\alpha}; N)} \frac{e^{2\pi \mathbf{i} \mathcal{L}_{\mathbf{s}}(\mathbf{n})}}{n_1 \prod_{i=1}^{d} (n_1 \alpha_i - n_{i+1})} \cdot g(\boldsymbol{n}, \boldsymbol{\alpha}; N) \Bigg). \tag{3.22}$$

Now $g(\mathbf{n}, \boldsymbol{\alpha}; N)$ is the product below (observe that $|g(\mathbf{n}, \boldsymbol{\alpha}; N)| \leq 1$):

$$\left(\frac{\sin 2\pi (\frac{n_1}{N^d})}{2\pi (\frac{n_1}{N^d})} \right)^2 \prod_{i=1}^{d} \left(\frac{\sin 2\pi (n_1 \alpha_i - n_{i+1})}{2\pi (n_1 \alpha_i - n_{i+1})} \right)^2 \tag{3.23}$$

and finally, $\mathcal{L}_{\mathbf{s}} = \mathcal{L}_{\mathbf{s}, x, N, \boldsymbol{\alpha}}$ is one of the $2^{d+1} - 1$ linear forms of $d+1$ variables:

$$\mathcal{L}_{\mathbf{s}}(\mathbf{n}) = \mathcal{L}_{\mathbf{s}}(n_1, \ldots, n_{d+1}) = \delta_1 n_1 x - \sum_{i=1}^{d} \delta_i N (n_1 \alpha_i - n_{i+1}) \tag{3.24}$$

where $\mathbf{s} = (\delta_1, \ldots, \delta_{d+1}) \in \{0, 1\}^{d+1}$ and $\mathbf{s} \neq \mathbf{0}$.

Note that the sign $\pm$ in the second part of (3.22) is in fact $\pm = (-1)^{\sum_{i=1}^{d+1} \delta_i}$, and so it is independent of $\mathbf{n} \in \mathbb{Z}^{d+1}$.

We begin with the constant part of $\bar{D}_4$:

Let

$$\bar{D}_5(\boldsymbol{\alpha}, x; N) = \sum_{\mathbf{n} \in U_4(\boldsymbol{\alpha}; N)} \frac{g(\boldsymbol{n}, \boldsymbol{\alpha}; N)}{n_1 \prod_{i=1}^{d} (n_1 \alpha_i - n_{i+1})}. \tag{3.25}$$

The goal of this section is to prove the following:

Proposition 3.7. *For almost every $\boldsymbol{\alpha} \in [0, 1]^d$, we have*

$$\left| \bar{D}_5(\boldsymbol{\alpha}, x; N) \right| \ll \ln N$$

The idea is to decompose the Fourier modes into sets where all the factors of the divisor are controlled in a small range, and by estimating of the cardinality of these sets, we can cancel the positive terms with with the negative terms and achieve a cancellation of the extra power of $\ln N$.

Let $\delta_N = 1/(\ln N)^d$, $\boldsymbol{\epsilon} = \{\epsilon_1, \ldots, \epsilon_{d+1}\} \in \{\pm 1\}^{d+1}$, and $\mathbf{l} = \{l_1, \ldots, l_{d+1}\} \in \mathbb{Z}^{d+1}$, and

define the sets:

$$S_{\boldsymbol{\alpha}}(\mathbf{l},\boldsymbol{\epsilon};N)=\left\{\begin{array}{l}\mathbf{n}\in\mathbb{Z}^{d+1}:(1+\delta_N)^{l_1}\leq\epsilon_1 n_1<(1+\delta_N)^{l_1+1},\\(1+\delta_N)^{-l_{i+1}}\leq\epsilon_{i+1}(n_1\alpha_i-n_{i+1})<(1+\delta_N)^{-l_{i+1}+1},\ 1\leq i\leq d-1\\(1+\delta_N)^{\sum_{i=1}^d l_{i+1}-l_1}\leq\epsilon_{d+1}(n_1\alpha_d-n_{d+1})\leq(1+\delta_N)^{\sum_{i=1}^d l_{i+1}-l_1+1},\\n_{i+1}\text{'s are respectively the closest integer to } n_1\alpha_i,\ 1\leq i\leq d\end{array}\right\},\tag{3.26}$$

where l_i's are positive integers, and the range of $\mathbf{l}=(l_1,\ldots,l_{d+1})$ is as the following:

$$L_2(N)=\left\{\mathbf{l}\in\mathbb{N}^{d+1}\ \middle|\ \begin{array}{l}(\ln N)^s\leq(1+\delta_N)^{l_{d+1}}\leq(1+\delta_N)^{l_1}\leq N^d/4,\\(1+\delta_N)^{l_1-l_{i+1}}\geq(\ln N)^s,\ 1\leq i\leq d-1.\end{array}\right\}$$

which gives

$$\ln\ln N/\delta_N\ll l_{d+1}\leq l_1\ll\ln N/\delta_N,\quad l_1-l_{i+1}\gg\ln\ln N/\delta_N,\ 1\leq i\leq d-1.\tag{3.27}$$

By integration over $\boldsymbol{\alpha}\in[0,1)$, the expected value for $\#S_{\boldsymbol{\alpha}}(\mathbf{l},\boldsymbol{\epsilon};N)$ is:

$$E(\mathbf{l},\boldsymbol{\epsilon};N)=\int_{\boldsymbol{\alpha}\in[0,1]^d}\#S_{\boldsymbol{\alpha}}(\mathbf{l},\boldsymbol{\epsilon};N)\ d\boldsymbol{\alpha}=\mathcal{O}(\delta_N^{d+1}(1+\delta_N)^{l_{d+1}}).$$

A more precise description for the number of elements in $\#S_{\boldsymbol{\alpha}}(\mathbf{l},\boldsymbol{\epsilon};N)$ is the following:

Lemma 3.7. *If $s>(d+2)d+3$, then for almost every $\boldsymbol{\alpha}\in\mathbb{R}^d$, and every $\mathbf{l}\in L_2(N)$, we have*

$$|\#S_{\boldsymbol{\alpha}}(\mathbf{l},\boldsymbol{\epsilon};N)-E(\mathbf{l},\boldsymbol{\epsilon};N)|\ll\delta_N E(\mathbf{l},\boldsymbol{\epsilon};N)$$

where the constant is uniform for $\mathbf{l}\in L_2(N)$ and N.

Proof. With the same reasoning as in Lemma 3.6, the number of elements inside $S_{\boldsymbol{\alpha}}(\mathbf{l},\boldsymbol{\epsilon};N)$ corresponds to the number of

$$n_1\in[(1+\delta_N)^{l_1},(1+\delta_N)^{l_1+1}),$$

such that vector $n_1\boldsymbol{\alpha}$ lies inside the target box

$$\prod_{i=1}^{d-1}[(1+\delta_N)^{-l_{i+1}},(1+\delta_N)^{-l_{i+1}+1})\times[(1+\delta_N)^{\sum_{i=1}^d l_{i+1}-l_1},(1+\delta_N)^{\sum_{i=1}^d l_{i+1}-l_1+1}),$$

therefore, the expected value for $\#S_{\boldsymbol{\alpha}}(\mathbf{l},\boldsymbol{\epsilon};N)$ is of order $\delta_N^{d+1}(1+\delta_N)^{l_{d+1}}$, from Theorem 3.1 we know that the error term is $\mathcal{O}(\ln(1+\delta_N)^{l_1})^{2+\epsilon})$ for any box, i.e.

$$\#S_{\boldsymbol{\alpha}}(\mathbf{l},\boldsymbol{\epsilon};N)=E(\mathbf{l},\boldsymbol{\epsilon};N)+\mathcal{O}\left((\ln(1+\delta_N)^{l_1})^{2+\epsilon}\right).$$

Note that $(\ln(1+\delta_N)^{l_1})^{2+\epsilon} \ll (\ln N)^{2+\epsilon} < \delta_N^{d+2}(\ln N)^s \leq \delta_N^{d+2}(1+\delta_N)^{l_{d+1}} \ll \delta_N E(\mathbf{l}, \boldsymbol{\epsilon}; N)$, the claim follows. □

For the sake of simplicity, we abbreviate $S_{\boldsymbol{\alpha}}(\mathbf{l}, \boldsymbol{\epsilon}; N)$ and note it by $S(\mathbf{l}, \boldsymbol{\epsilon}; N)$ in later discussions. Using the Lemma above, we can estimate the sum $\bar{D}_5$ by cancelling out the main terms.

Proof of Proposition 3.7

Proof. Let $\boldsymbol{\epsilon}^+$ and $\boldsymbol{\epsilon}^-$ be two vectors in $\{-1, +1\}^{d+1}$ such that one and only one coordinate is different, and the sign of the divisor is $+$ for $S(\mathbf{l}, \boldsymbol{\epsilon}^+; N)$, and $-$ for $S(\mathbf{l}, \boldsymbol{\epsilon}^-; N)$.

$$\bar{D}_5 = \sum_{\substack{\mathbf{l} \in L_2(N), \\ \text{pairs of } \boldsymbol{\epsilon}^\pm}} \left(\sum_{\mathbf{n} \in S(\mathbf{l}, \boldsymbol{\epsilon}^+; N)} \frac{g(\boldsymbol{n}, \boldsymbol{\alpha}; N)}{n_1 \prod_{i=1}^d (n_1\alpha_i - n_{i+1})} + \sum_{\mathbf{n} \in S(\mathbf{l}, \boldsymbol{\epsilon}^-; N)} \frac{g(\boldsymbol{n}, \boldsymbol{\alpha}; N)}{n_1 \prod_{i=1}^d (n_1\alpha_i - n_{i+1})} \right).$$

Inside each $S(\mathbf{l}, \boldsymbol{\epsilon}; N)$, denote by $g(\boldsymbol{n}, \boldsymbol{\alpha}; N)_{max/min}$ the maximal/minimal value of $g(\boldsymbol{n}, \boldsymbol{\alpha}; N)$ for $\mathbf{n} \in S(\mathbf{l}, \boldsymbol{\epsilon}; N)$, we have:

$$|g(\boldsymbol{n}, \boldsymbol{\alpha}; N)_{max} - g(\boldsymbol{n}, \boldsymbol{\alpha}; N)_{min}| \ll \delta_N g(\boldsymbol{n}, \boldsymbol{\alpha}; N)_{min}$$

and

$$0 < g(\boldsymbol{n}, \boldsymbol{\alpha}; N)_{max} \leq 1.$$

For each pair, the sum would cancel out as follows:

$$\begin{aligned}
&\sum_{\mathbf{n} \in S(\mathbf{l}, \boldsymbol{\epsilon}^+; N)} \frac{g(\boldsymbol{n}, \boldsymbol{\alpha}; N)}{n_1 \prod_{i=1}^d (n_1\alpha_i - n_{i+1})} + \sum_{\mathbf{n} \in S(\mathbf{l}, \boldsymbol{\epsilon}^-; N)} \frac{g(\boldsymbol{n}, \boldsymbol{\alpha}; N)}{n_1 \prod_{i=1}^d (n_1\alpha_i - n_{i+1})} \\
&\leq \frac{g(\boldsymbol{n}, \boldsymbol{\alpha}; N)_{max}}{(1+\delta_N)^{l_{d+1}}}(1+\delta_N)E(\mathbf{l}, \boldsymbol{\epsilon}; N) - \frac{g(\boldsymbol{n}, \boldsymbol{\alpha}; N)_{min}}{(1+\delta_N)^{l_{d+1}+d+1}}(1-\delta_N)E(\mathbf{l}, \boldsymbol{\epsilon}; N) \\
&\ll \frac{\left(g(\boldsymbol{n}, \boldsymbol{\alpha}; N)_{max}(1+\delta_N)^{d+1} - g(\boldsymbol{n}, \boldsymbol{\alpha}; N)_{min}\right)}{(1+\delta_N)^{l_{d+1}+d+1}} E(\mathbf{l}, \boldsymbol{\epsilon}; N) + \frac{g(\boldsymbol{n}, \boldsymbol{\alpha}; N)_{max}}{(1+\delta_N)^{l_{d+1}}} \delta_N E(\mathbf{l}, \boldsymbol{\epsilon}; N) \\
&\ll \delta_N \frac{E(\mathbf{l}, \boldsymbol{\epsilon}; N)}{(1+\delta_N)^{l_{d+1}}} \\
&\ll \delta_N^{d+2}
\end{aligned}$$

The other direction is the same:

$$\sum_{\mathbf{n} \in S(\mathbf{l}, \boldsymbol{\epsilon}^+; N)} \frac{g(\boldsymbol{n}, \boldsymbol{\alpha}; N)}{n_1 \prod_{i=1}^d (n_1\alpha_i - n_{i+1})} + \sum_{\mathbf{n} \in S(\mathbf{l}, \boldsymbol{\epsilon}^-; N)} \frac{g(\boldsymbol{n}, \boldsymbol{\alpha}; N)}{n_1 \prod_{i=1}^d (n_1\alpha_i - n_{i+1})} \gg -\delta_N^{d+2}.$$

Summing over all the possible pairs of $\boldsymbol{\epsilon}^\pm$ and $l_1, \ldots, l_{d+1}$, note that from (3.27),

$0 \le l_{i+1} \le l_1,\ 1 \le i \le d$, we have:

$$\left|\bar{D}_5\right| \ll \sum_{\mathbf{l}\in L_2(N)} \delta_N^{d+2} \ll \sum_{\frac{\ln\ln N}{\delta_N} \le l_1 \le \frac{\ln N}{\delta_N}} \delta_N^{d+2} l_1^d \ll \delta_N^{d+2} \cdot \frac{(\ln N)^{d+1}}{\delta_N^{d+1}} = \ln N.$$

□

3.5 Estimation of small exponentials.

Finally we study the contribution of the linear forms $\mathcal{L}_{\mathbf{s}}$ in (3.22).

Let

$$\begin{aligned} \bar{D}_6 = \bar{D}_6^{(\mathbf{s})} &= \sum_{\mathbf{n}\in U_4(\boldsymbol{\alpha};N)} \frac{e^{2\pi \mathbf{i}\mathcal{L}_{\mathbf{s}}(\mathbf{n})}}{n_1 \prod_{i=1}^d (n_1\alpha_i - n_{i+1})} \cdot g(\boldsymbol{n}, \boldsymbol{\alpha}; N) \\ &= \sum_{\substack{\mathbf{l}\in L_2(N),\\ \text{pairs of } \boldsymbol{\epsilon}^{\pm}}} \left(\sum_{\mathbf{n}\in S(\mathbf{l},\boldsymbol{\epsilon}^+;N)} \bar{f}(\mathbf{n}, x, \boldsymbol{\alpha}) + \sum_{\mathbf{n}\in S(\mathbf{l},\boldsymbol{\epsilon}^-;N)} \bar{f}(\mathbf{n}, x, \boldsymbol{\alpha}), \right) \end{aligned} \tag{3.28}$$

where

$$\bar{f}(\mathbf{n}, x, \boldsymbol{\alpha}) = \frac{e^{2\pi \mathbf{i}\mathcal{L}_{\mathbf{s}}(\mathbf{n})}}{n_1 \prod_{i=1}^d (n_1\alpha_i - n_{i+1})} \cdot g(\boldsymbol{n}, \boldsymbol{\alpha}; N) \tag{3.29}$$

and $\mathcal{L}$ is one of the $2^{d+1}-1$ linear forms $\mathcal{L}_{\mathbf{s},x,N,\boldsymbol{\alpha}}$ defined in (3.24). Following the same line of reasoning as in J. Beck[Bec94], we prove Lemma 3.8 as a version of Key Lemma as in [Bec94] that can be adopted in our case. We shall emphasize the key ingredient in the proof is that arithmetic progressions will contribute like a single term inside the progression. Using Lemma 3.8, we can estimate the number of $\boldsymbol{\epsilon}^{\pm}$-big vectors defined as follows, which can be used to cancel the extra $\ln N$ in the crude estimation by bounding every numerator by 1.

Definition 3.1. *We say* $\mathbf{l} = (l_1, \ldots, l_{d+1})$ *is "*$\boldsymbol{\epsilon}^{\pm}$*-big" vector if*

$$\frac{|S(\mathbf{l}, \boldsymbol{\epsilon}^+; N)| + |S(\mathbf{l}, \boldsymbol{\epsilon}^-; N)|}{\ln N} \le \left| \sum_{\mathbf{n}\in S(\mathbf{l},\boldsymbol{\epsilon}^+;N)} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{n})} - \sum_{\mathbf{n}\in S(\mathbf{l},\boldsymbol{\epsilon}^-;N)} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{n})} \right|, \tag{3.30}$$

where $|S| = \#S$ *denotes the number of elements inside the set* S*, and as in the proof of Proposition 3.7,* $\boldsymbol{\epsilon}^+$ *and* $\boldsymbol{\epsilon}^-$ *are a pair of vectors in* $\{-1,+1\}^{d+1}$ *such that one and only one coordinate is different, and the sign of the divisor is* $+$ *for* $S(\mathbf{l}, \boldsymbol{\epsilon}^+; N)$*, and* $-$ *for* $S(N, \mathbf{l}, \boldsymbol{\epsilon}^-)$*.*

For convenience denote

$$S(\mathbf{l}, \boldsymbol{\epsilon}^{\pm}; N) = S(\mathbf{l}, \boldsymbol{\epsilon}^+; N) \cup S(\mathbf{l}, \boldsymbol{\epsilon}^-; N).$$

Definition 3.2. *Two integral vectors* $\mathbf{l} = (l_1, \ldots, l_{d+1})$ *and* $\mathbf{h} = (h_1, \ldots, h_{d+1})$ *satisfying* (3.27) *are called "**neighbors**" if*

$$(1+\delta_N)^{h_1-l_1} = (\ln N)^9, \tag{3.31}$$

$$(1+\delta_N)^{h_{i+1}-l_{i+1}} = (\ln N)^{-9},\ 1 \le i \le d-1, \tag{3.32}$$

$$(1+\delta_N)^{h_{d+1}-l_{d+1}} = (\ln N)^{9(d+1)}. \tag{3.33}$$

The notation $\mathbf{l}\to\mathbf{h}$ means that the ordered pair $\langle \mathbf{l}, \mathbf{h}\rangle$ of vectors satisfies (3.31) to (3.33).

Note that by slightly modifying the value of $\delta_N \approx (\ln N)^{-d}$, we can make sure that the above definitions are met for integer vectors $\mathbf{l}$ and $\mathbf{h}$.

Definition 3.3. *A sequence* $H = \langle \mathbf{h}^{(1)}, \mathbf{h}^{(2)}, \mathbf{h}^{(3)}, \ldots\rangle$ *of vectors satisfying* (3.27) *is called a "special line" if* $\mathbf{h}^{(1)}\to\mathbf{h}^{(2)}\to\mathbf{h}^{(3)}\to\ldots$, *that is, any two consecutive vectors in* H *are neighbors.*

Lemma 3.8. *For almost every* $\boldsymbol{\alpha}$ *every special line contains at most one* $\boldsymbol{\epsilon}^{\pm}$*-big vector.*

Proof. Let $H = \langle \mathbf{h}^{(1)}, \mathbf{h}^{(2)}, \mathbf{h}^{(3)}, \ldots\rangle$ be a special line with two $\boldsymbol{\epsilon}^{\pm}$-big vectors $\mathbf{h}^{(p)}$ and $\mathbf{h}^{(q)}$, $1 \le p < q$. If

$$\|\mathcal{L}(\mathbf{n})\| \le (\ln N)^{-2} \text{ for every } \mathbf{n} \in S(\mathbf{h}^{(p)}, \boldsymbol{\epsilon}^{\pm}; N), \tag{3.34}$$

then

$$|1 - e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{n})}| \ll (\ln N)^{-2} \text{ for every } \mathbf{n} \in S(\mathbf{h}^{(p)}, \boldsymbol{\epsilon}^{\pm}; N). \tag{3.35}$$

By repeating the argument of the cancellation of the main term with the above equation, we obtain (Err means error for the number of elements in the set S)

$$\begin{aligned}
&\left| \sum_{\mathbf{n}\in S(\mathbf{h}^{(p)},\boldsymbol{\epsilon}^{+};N)} e^{2\pi\mathbf{i}\mathcal{L}(\mathbf{n})} - \sum_{\mathbf{n}\in S(\mathbf{h}^{(p)},\boldsymbol{\epsilon}^{-};N)} e^{2\pi\mathbf{i}\mathcal{L}(\mathbf{n})} \right| \\
&\ll (1+(\ln N)^{-2})(E(\mathbf{h}^{(p)}, \epsilon; N)) + |\mathrm{Err}|) - (1-(\ln N)^{-2})(E(\mathbf{h}^{(p)}, \epsilon; N)) - |\mathrm{Err}|) \\
&\ll |\mathrm{Err}| + (\ln N)^{-2} E(\mathbf{h}^{(p)}, \epsilon; N)) \\
&\ll \delta_N (E(\mathbf{h}^{(p)}, \epsilon; N)) \\
&\ll \frac{|S(\mathbf{h}^{(p)}, \boldsymbol{\epsilon}^{\pm}; N)|}{(\ln N)^2}.
\end{aligned}$$

But this contradicts the assumption that $\mathbf{h}^{(p)}$ is $\boldsymbol{\epsilon}^{\pm}$-big, see (3.30). So there is an $\mathbf{n}^* \in S(\mathbf{h}^{(p)}, \boldsymbol{\epsilon}^{\pm}; N)$ such that

$$\|\mathcal{L}(\mathbf{n})\| > (\ln N)^{-2}. \tag{3.36}$$

For every $\mathbf{m} \in S(\mathbf{h}^{(q)}, \boldsymbol{\epsilon}^{\pm}; N)$ (another $\boldsymbol{\epsilon}^{\pm}$-big vector), consider the "arithmetic progres-

sion" with difference $\mathbf{n}^*$:

$$\mathbf{m} + r \cdot \mathbf{n}^* = (m_1 + r \cdot \mathbf{n}^*{}_1, \ldots, m_{d+1} + r \cdot \mathbf{n}^*{}_{d+1}), \quad r = 0, \pm 1, \pm 2, \ldots$$

We will estimate how many consecutive members $\mathbf{m} + r \cdot \mathbf{n}^*$ are contained in $S(N, \mathbf{h}^q, \boldsymbol{\epsilon}^{\pm})$. Since $n^* \in S(\mathbf{h}^p, \boldsymbol{\epsilon}^{\pm}; N)$, the definition for $S(\mathbf{h}^q, \boldsymbol{\epsilon}^{\pm}; N)$ (see (3.26)) gives the following:

$$(1+\delta_N)^{h_1^{(p)}} \leq \epsilon_1 n_1^* < (1+\delta_N)^{h_1^{(p)}+1}, \tag{3.37}$$

$$(1+\delta_N)^{-h_{i+1}^{(p)}} \leq \epsilon_{i+1}(n_1^* \alpha_i - n_{i+1}^*) < (1+\delta_N)^{-h_{i+1}^{(p)}+1}, \tag{3.38}$$

$$(1+\delta_N)^{\sum_{i=1}^d h_{i+1}^{(p)} - h_1^{(p)}} \leq \epsilon_{d+1}(n_1^* \alpha_d - n_{d+1}^*) < (1+\delta_N)^{\sum_{i=1}^d h_{i+1}^{(p)} - h_1^{(p)}+1}. \tag{3.39}$$

Definition 3.4. *An* $\mathbf{m} \in S(\mathbf{h}^q, \boldsymbol{\epsilon}^{\pm}; N)$ *is called an inner point if*

$$(1+\delta_N)^{h_1^{(q)}}(1 + \frac{\delta_N}{(\ln N)^2}) \leq \epsilon_1 m_1 < (1 - \frac{\delta_N}{(\ln N)^2})(1+\delta_N)^{h_1^{(q)}+1}, \tag{3.40}$$

$$(1+\delta_N)^{-h_{i+1}^{(q)}}(1 + \frac{\delta_N}{(\ln N)^2}) \leq \epsilon_{i+1}(m_1 \alpha_i - m_{i+1}) < (1 - \frac{\delta_N}{(\ln N)^2})(1+\delta_N)^{-h_{i+1}^{(q)}+1}, \tag{3.41}$$

$$(1+\delta_N)^{\sum_{i=1}^d h_{i+1}^{(p)} - h_1^{(q)}}(1 + \frac{\delta_N}{(\ln N)^2}) \leq \epsilon_{d+1}(m_1 \alpha_{d+1} - m_{d+1}) < (1 - \frac{\delta_N}{(\ln N)^2})(1+\delta_N)^{\sum_{i=1}^d h_{i+1}^{(p)} - h_1^{(q)}+1}. \tag{3.42}$$

The rest of the points in $S(\mathbf{h}^q, \boldsymbol{\epsilon}^{\pm}; N)$ *are called border points.*

For every inner point $\mathbf{m} \in S(\mathbf{h}^q, \boldsymbol{\epsilon}^{\pm}; N)$, and for every $|r| \leq (\ln N)^4$, it follows from (3.31), (3.37) and (3.40) that,

$$\begin{aligned} &(1+\delta_N)^{h_1^{(q)}} < (1+\delta_N)^{h_1^{(q)}}(1 + \frac{\delta_N}{(\ln N)^2}) - (\ln N)^4 (1+\delta_N)^{h_1^{(p)}+1} \\ &\leq \epsilon_1(m_1 + r \cdot n_1^*) \\ &< (1 - \frac{\delta_N}{(\ln N)^2})(1+\delta_N)^{h_1^{(q)}+1} + (\ln N)^4 (1+\delta_N)^{h_1^{(p)}+1} < (1+\delta_N)^{h_1^{(q)}+1} \end{aligned} \tag{3.43}$$

Similarly, from (3.32), (3.38) and (3.41), and from (3.33), (3.39) and (3.42) we obtain the following, for $1 \leq i \leq d-1$,

$$(1+\delta_N)^{-h_{i+1}^{(q)}} < \epsilon_{i+1}\left((m_1 + r n_1^*)\alpha_{i+1} - (m_{i+1} + r \cdot n_{i+1}^*)\right) < (1+\delta_N)^{-h_{i+1}^{(q)}+1}, \tag{3.44}$$

and

$$(1+\delta_N)^{\sum_{i=1}^d h_{i+1}^{(q)} - h_1^{(q)}} < \epsilon_{d+1}\left((m_1 + r n_1^*)\alpha_{d+1} - (m_{d+1} + r \cdot n_{d+1}^*)\right) < (1+\delta_N)^{\sum_{i=1}^d h_{i+1}^{(q)} - h_1^{(q)}+1} \tag{3.45}$$

In view of (3.43)-(3.45), for any inner point $\mathbf{m} \in S(\mathbf{h}^q, \boldsymbol{\epsilon}^{\pm}; N)$, at least $(\ln N)^4$ consecutive

members in the progression $\mathbf{m} + r \cdot \mathbf{n}^*$ are contained in $S(\mathbf{h}^q, \boldsymbol{\epsilon}^{\pm}; N)$. Therefore, we can decommpose $S(\mathbf{h}^q, \boldsymbol{\epsilon}^{\pm}; N)$ into three parts:

$$S(\mathbf{h}^q, \boldsymbol{\epsilon}^{\pm}; N) = AP^+ \cup AP^- \cup BP \tag{3.46}$$

where $AP^{\pm}$ denotes the family of arithmetic progreesions $\{\mathbf{m} + r \cdot \mathbf{n}^* : 0 \leq r \leq l-1\}$ in $S(\mathbf{h}^q, \epsilon^+; N)$ and $S(\mathbf{h}^q, \epsilon^-; N)$ respectively, where $l = l(\mathbf{m})$ is the length of the progression starting from $\mathbf{m}$, and $l \geq (\ln N)^4$. BP denotes a set of border points of $S(\mathbf{h}^q, \boldsymbol{\epsilon}^{\pm}; N)$ that are not included in any arithmetic progressions. Using $\|\mathcal{L}(\mathbf{n})\| > (\ln N)^{-2}$ (see (3.36)), the linearity of $\mathcal{L}$, we obtain

$$\begin{aligned}
|\sum_{\mathbf{AP}^+} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{n})}| &\leq \sum_{\text{arithmetic progressions}} |\sum_{r=0}^{l-1} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{m}+r\mathbf{n}^*)}| \\
&= \sum_{\text{arithmetic progressions}} |\sum_{r=0}^{l-1} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{m})+r\mathcal{L}(\mathbf{n}^*)}| \\
&\ll \sum_{\text{arithmetic progressions}} \frac{1}{\|\mathcal{L}(\mathbf{n})\|} < \sum_{\text{arithmetic progressions}} (\ln N)^2 \\
&\leq \sum_{\text{arithmetic progressions}} \frac{\text{length}}{(\ln N)^2} \leq \frac{|S(\mathbf{h}^q, \epsilon^+; N)|}{(\ln N)^2}
\end{aligned} \tag{3.47}$$

since each length $\geq (\ln N)^4$. Similarly,

$$|\sum_{\mathbf{AP}^-} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{n})}| \ll \frac{|S(\mathbf{h}^q, \epsilon^-; N)|}{(\ln N)^2} \tag{3.48}$$

Finally, for border points, at least one of the inequalities in definition 3.4 is violated, thus the range for at least one components is shrunk with ratio $\frac{\delta_N}{(\ln N)^2}$. Using the same reasoning as in Lemma 3.6 and 3.7, and the cardinality of the set BP can be controlled by the total number of border points of $S(\mathbf{h}^q, \boldsymbol{\epsilon}^{\pm}; N)$, we have for almost every $\boldsymbol{\alpha}$,

$$|BP| \ll \frac{|S(\mathbf{h}^q, \boldsymbol{\epsilon}^{\pm}; N)|}{(\ln N)^2} \tag{3.49}$$

Combining (3.46)-(3.49), for almost every $\boldsymbol{\alpha}$, we obtain

$$\left| \sum_{\mathbf{n} \in S(\mathbf{h}^{(p)}, \boldsymbol{\epsilon}^+; N)} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{n})} - \sum_{\mathbf{n} \in S(\mathbf{h}^{(p)}, \boldsymbol{\epsilon}^-; N)} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{n})} \right| \ll \frac{|S(\mathbf{h}^{(p)}, \boldsymbol{\epsilon}^{\pm}; N)|}{(\ln N)^2}$$

which contradicts the assumption that $h^{(q)}$ is $\boldsymbol{\epsilon}^{\pm}$-big (see (3.30)), therefore for almost very $\boldsymbol{\alpha}$, every special line contains at most one $\boldsymbol{\epsilon}^{\pm}$-big vector, which proves the lemma. □

Corollary 3.3. *The number of $\boldsymbol{\epsilon}^{\pm}$-big vectors is* $\ll (\ln\ln N)(\ln N)^d \delta_N^{-(d+1)}$.

Proof. First we estimate the number of maximal special lines. Let $H = \langle \mathbf{h}^{(1)}, \mathbf{h}^{(2)}, \mathbf{h}^{(3)}, \ldots \rangle$ be a "special line", here $\mathbf{h}^{(1)}$ is the first element of H, that is, if $\mathbf{h}^{(0)} \to \mathbf{h}^{(1)}$ holds for some $\mathbf{h}^{(0)}$, then at least one of the following inequalities is violated:

$$(1+\delta_N)^{h_1^{(0)}} \geq (\ln N)^s,$$

$$(1+\delta_N)^{h_1^{(0)} - h_{i+1}^{(0)}} \geq (\ln N)^s, 1 \leq i \leq d-1,$$

$$(1+\delta_N)^{h_{d+1}^{(0)}} \geq (\ln N)^s.$$

Thus, by (3.31) and (3.33), one of the inequalities below holds:

$$(\ln N)^s \leq (1+\delta_N)^{h_1^{(1)}} \leq (\ln N)^{s+9},$$

$$(\ln N)^s \leq (1+\delta_N)^{h_1^{(1)} - h_{i+1}^{(1)}} \leq (\ln N)^{s+18},\ 1 \leq i \leq d-1,$$

$$(\ln N)^s \leq (1+\delta_N)^{h_{d+1}^{(1)}} \leq (\ln N)^{s+9(d+1)},$$

So at least one coordinate of $h_1^{(1)}$, $h_{d+1}^{(1)}$, or $h_1^{(1)} - h_{i+1}^{(1)}$ of the first element $\mathbf{h}^{(1)}$ of H is restricted to a short interval of length const$\cdot \ln\ln N \cdot \delta_N^{-1}$, the rest, by the condition (3.27), are restricted to an interval of length const$\cdot \ln N \cdot \delta_N^{-1}$. Since the starting vector determines the whole special line, the number of special line is

$$\ll (\ln\ln N) \cdot (\ln N)^d \cdot \delta_N^{-(d+1)}.$$

By Lemma 3.8, the total number of $\boldsymbol{\epsilon}^{\pm}$-big vectors is also

$$\ll (\ln\ln N) \cdot (\ln N)^d \cdot \delta_N^{-(d+1)}.$$

□

With the help of the Lemma 3.8, we can estimate the contribution of the exponential terms, we have the following claim:

Proposition 3.8. *For almost every $\boldsymbol{\alpha}$, we have*

$$\left|\bar{D}_6\right| \ll (\ln N)^d (\ln\ln N)$$

Proof.

$$\bar{D}_6 = \sum_{\text{small}} + \sum_{\text{big}}$$

where

$$\sum_{\text{small}} = \sum_{\boldsymbol{\epsilon}} \sum_{\mathbf{l} \text{ not } \boldsymbol{\epsilon}^{\pm}\text{-big}} \sum_{\mathbf{n} \in S(N,\mathbf{l},\boldsymbol{\epsilon})} \bar{f}(\mathbf{n}, x, \boldsymbol{\alpha}) \tag{3.50}$$

$$\sum_{\text{big}} = \sum_{\boldsymbol{\epsilon}} \sum_{\mathbf{l} \text{ is } \boldsymbol{\epsilon}^{\pm}\text{-big}} \sum_{\mathbf{n} \in S(N,\mathbf{l},\boldsymbol{\epsilon})} \bar{f}(\mathbf{n}, x, \boldsymbol{\alpha}) \tag{3.51}$$

By Lemma 3.7, range for $\mathbf{l}$:(3.27), and (3.30),

$$\begin{aligned}
\left|\sum_{\text{small}}\right| &\ll \sum_{\boldsymbol{\epsilon}} \sum_{\mathbf{l} \text{ is not } \boldsymbol{\epsilon}^{\pm}\text{-big}} (1+\delta_N)^{-l_{d+1}} \cdot \frac{E(\mathbf{l}, \boldsymbol{\epsilon}; N)}{\ln N} \\
&\ll \sum_{\mathbf{l} \in L_2(N)} \frac{\delta_N^{d+1}}{\ln N} \\
&\ll \frac{\delta_N^{d+1}}{\ln N} \cdot \frac{(\ln N)^{d+1}}{\delta_N^{d+1}} \\
&\ll (\ln N)^d
\end{aligned}$$

By Corollary 3.3,

$$\begin{aligned}
|\sum_{\text{big}}| &\ll \sum_{\mathbf{l} \text{ is } \boldsymbol{\epsilon}^{\pm}\text{-big}} (1+\delta_N)^{-l_{d+1}} \cdot E(\mathbf{l}, \boldsymbol{\epsilon}; N) \\
&\ll [(\ln\ln N)(\ln N)^d \delta_N^{-(d+1)}] \cdot [(1+\delta_N)^{-l_{d+1}} \cdot (1+\delta_N)^{l_{d+1}} \delta_N^{d+1}] \\
&\ll (\ln N)^d \ln\ln N
\end{aligned}$$

□

Combining Proposition 3.1, 3.2, 3.7, 3.8, we finally proved the convergent part (3.2) of Theorem 3.2.

Proof for the divergent part (3.3) **of Theorem 3.2**: The idea for the divergent part is straightforward, we search for a "large" Fourier coefficient that dominates. For technical simplicity, we consider a roof-like average of the discrepancy function $D(\boldsymbol{\alpha}; a, b; \mathbf{u}, \mathbf{u}+N)$ in (3.4) over the interval $[-x, x]$ instead $[0, x]$, from which we get a symmetric version of $\bar{D}(\boldsymbol{\alpha}, x; N)$:

$$\bar{D}(\boldsymbol{\alpha}, -x, x; 0, N) = \sum_{n \in \mathbb{Z}^3 \setminus \{0\}} f_{\mathbf{n},N}(x, \boldsymbol{\alpha}) \tag{3.52}$$

where

$$f_{\mathbf{n},N}(x, \boldsymbol{\alpha}) = \frac{\sin(2\pi n_1 x)}{\pi n_1} g(\mathbf{n}, \boldsymbol{\alpha}; N) \cdot \prod_{i=1}^{d} \frac{1 - e^{-2\pi \mathbf{i} N(n_1\alpha_i - n_{i+1})}}{\pi(n_1\alpha_i - n_{i+1})}. \tag{3.53}$$

By a result of Gallagher[Gal62] about metric simultaneous diophantine approximation, we have the following lemma:

Lemma 3.9. *Let $\varphi(n)$ be an arbitrary positive increasing function of n with $\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} =$*

∞, then for almost every $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$, there exist infinitely many $n \in \mathbb{N}$, such that:

$$n \prod_{i=1}^{d} \|n\alpha_i\| < \frac{1}{(\ln n)^d \varphi(\ln \ln n)}. \tag{3.54}$$

Proof. Let n_1^* be one of the infinite integers that satisfies (3.54), and let n_{i+1}^* be the closest integer to $n_1^* \alpha_i$ for $1 \leq i \leq d$, then by Lemma 3.3 and (3.54), we have for n_1^* large enough:

$$\frac{1}{n_1^*(\ln n_1^*)^{d+\epsilon}} \leq \prod_{i=1}^{d} \|n_1^* \alpha_i\| \leq \frac{1}{n_1^*(\ln n_1^*)^d \varphi(\ln \ln n_1^*)} \tag{3.55}$$

Now we are looking for an integer N such that:

$$\prod_{i=1}^{d} \left| 1 - e^{2\pi \mathbf{i} N(n_1^* \alpha_i - n_{i+1}^*)} \right| \geq \frac{1}{2}, \tag{3.56}$$

that is, $N(n_1^* \alpha_i - n_{i+1}^*)$ is *not* close to 0.

By the same reasoning as in Lemma 3.3 and the upper bound for $\prod_{i=1}^{d} \|n_1^* \alpha_i\|$ in (3.55) we have:

$$\frac{1}{n_1^*(\ln n_1^*)^{1+\epsilon}} \leq \|n_1^* \alpha_i\| \leq \frac{1}{(\ln n_1^*)^{\frac{1}{2}}}, \quad 1 \leq i \leq d, \tag{3.57}$$

thus it is possible to choose

$$n_1^* \ln n_1^* \leq N \leq n_1^* (\ln n_1^*)^2, \tag{3.58}$$

such that (3.56) is satisfied by pigeonhole principle.

With the choice of $\mathbf{n}^* = (n_1^*, \ldots, n_{d+1}^*)$ and N^* as above such that the inequalities (3.55) and (3.56) are satisfied, we are able to estimate the lower bound of the discrepancy function by its projection of to $e^{2\pi \mathbf{i} n_1^* x}$.

$$\int_{[0,1]} \bar{D}(\boldsymbol{\alpha}, -x, x; 0, N^*) \sin(2\pi n_1^* x) dx = S_1 + S_2, \tag{3.59}$$

where S_1 is the integral over the Fourier mode $\mathbf{n}^*$, and S_2 the rest:

$$\begin{aligned} S_1 &= g(\mathbf{n}^*, N^*, \alpha) \int_{[0,1]} f_{\mathbf{n}^*, N^*}(x) \sin(2\pi n_1^* x) dx \\ &= g(\mathbf{n}^*, N^*, \alpha) \frac{\prod_{i=1}^{d} (1 - e^{2\pi \mathbf{i} N(n_1^* \alpha_i - n_{i+1}^*)})}{n_1^* \prod_{i=1}^{d} (n_1^* \alpha_i - n_{i+1}^*)} \end{aligned} \tag{3.60}$$

$$\begin{aligned} S_2 &= \sum_{(n_2,\ldots,n_{d+1})\neq(n_2^*,\ldots,n_{d+1}^*)} g(\mathbf{n},N^*,\boldsymbol{\alpha}) \int_{[0,1]} f_{\mathbf{n},N^*}(x)\sin(2\pi n_1^* x)dx \\ &= \sum_{(n_2,\ldots,n_{d+1})\neq(n_2^*,\ldots,n_{d+1}^*)} g(\mathbf{n},N^*,\boldsymbol{\alpha}) \frac{\prod_{i=1}^d (1-e^{2\pi \mathbf{i} N(n_1^*\alpha_i - n_{i+1})})}{n_1^* \prod_{i=1}^d (n_1^*\alpha_i - n_{i+1})} \end{aligned} \tag{3.61}$$

Let $r_{\alpha_i} = \lceil |n_1^*\alpha_i - n_{i+1}| \rceil$, $1 \le i \le d$, then

$$\begin{aligned} |S_2| &\ll \sum_{i=1}^d \sum_{r_{\alpha_i}=1}^{\infty} \frac{1}{n_1^* \|n_1^*\alpha_i\| \prod_{j\neq i} r_{\alpha_j}^3} + \sum_{i=1}^d \sum_{r_{\alpha_i}=1}^{\infty} \frac{1}{n_1^* \prod_{i=1}^d r_{\alpha_i}^3} \\ &\ll \sum_{i=1}^d \frac{1}{n_1^* \|n_1^*\alpha_i\|} \ll (\ln n_1^*)^{1+\epsilon} \ll (\ln N^*)^{1+\epsilon}, \end{aligned} \tag{3.62}$$

and on the other hand, note that since $\|n_1^*\alpha_i\|$ is small, $g(n^*, N^*, \boldsymbol{\alpha})$ admits an absolute lower bound, then by (3.55) and (3.56),

$$|S_1| \gg (\ln n_1^*)^d \varphi(\ln\ln n_1^*) \gg (\ln N^*)^d \varphi(\ln\ln N^*). \tag{3.63}$$

Combining the inequalities (3.63) and (3.62), we have the divergent part of Theorem 3.2 □

Chapter 4

Absolute Bounds - Irrational translations - Triangles

4.1 Introduction

In this chapter, we consider the ergodic discrepancies of irrational toral translations relative to *triangles* inside $\mathbb{T}^2$. The main ideas of the proof remain very similar to the case of linear forms, with modifications to treat the new divisors. nevertheless we include the full proof for the sake of completeness.

Let $\boldsymbol{\alpha} = (\alpha_1, \alpha_2) \in [0,1]^2$ be the translation vector, where $1, \alpha_1, \alpha_2$ are rationally independent, let $\boldsymbol{a} = (a_1, a_2) \in \mathbb{R}^2$ be the starting point, and denote the right triangle of sides $\mathbf{x} = (x_1, x_2) \in (0,1]^2$ by:

$$\triangle_{\mathbf{x}} = \{(y_1, y_2) \mid 0 \leq y_1/x_1 + y_2/x_2 < 1,\ 0 \leq y_i \leq x_i, i = 1,2\}. \tag{4.1}$$

Define the ergodic discrepancy of the toral translation $\boldsymbol{\alpha}$ starting from the point $\boldsymbol{a}$ relative to the triangle $\triangle_{\mathbf{x}}$:

$$D_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; m) = \sum_{n=1}^{m} \chi_{\triangle_{\mathbf{x}}}(\boldsymbol{a} + n\boldsymbol{\alpha} \mod 1) - m\mathrm{Vol}(\triangle_{\mathbf{x}})$$

Since the sequence $\{n\boldsymbol{\alpha} \mod 1\}$ is equistributed over $\mathbb{T}^2$, the "expected number" of points in the sequence that visit the triangle $\triangle_{\mathbf{x}}$ before time m is $m\mathrm{Vol}(\triangle_{\mathbf{x}})$, and we would like to prove a similar bound for the error term $D_{\triangle_{\mathbf{x}}}$.

Notations. Through out this chapter, $\|\cdot\|$ denotes the distance to the closest integer. For a given triangle $\triangle_{\mathbf{x}}$ defined as above (4.1), let $\tau = x_1/x_2$ denote the slope of the hypotenuse of the triangle. Without loss of generality, we assume that $\tau \in [0,1]$. For $\mathbf{n} = (n_1, n_2, n_3) \in \mathbb{Z}^3$, and a vector $\boldsymbol{\alpha} = (\alpha_1, \alpha_2) \in \mathbb{R}^2$, let $\mathbf{n}\boldsymbol{\alpha} = n_1\alpha_1 + n_2\alpha_2$ denote the inner product of their first two coordinates. From Section 5, the $\mathbf{n}$ would denote

$(n_1, n_2) \in \mathbb{Z}^2$ instead of (n_1, n_2, n_3), and n_3 would denote the closest integer to $n_1\alpha_1+n_2\alpha_2$, since in Section 4 the other possibilities for n_3 have already been estimated.

The main result of this chapter is the following:

Theorem 4.1. *Let $\boldsymbol{\alpha} = (\alpha_1, \alpha_2) \in \mathbb{R}^2$ be the translation vector, $\boldsymbol{a} = (a_1, a_2) \in [0,1]^2$ be the starting point, and for $\mathbf{x} = (x_1, x_2) \in (0,1]^2$ define the triangle $\triangle_{\mathbf{x}} = \{(y_1, y_2) \mid 0 \leq y_1/x_1 + y_2/x_2 < 1,\ 0 \leq y_i < x_i, i = 1, 2\}$, define the ergodic discrepancy:*

$$D_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; m) = \sum_{n=1}^{m} \chi_{\triangle_{\mathbf{x}}}(\boldsymbol{a} + n\boldsymbol{\alpha} \mod 1) - m\, Vol(\triangle_{\mathbf{x}})$$

and the maximal discrepancy for triangles of a fixed slope τ:

$$\Delta(\boldsymbol{\alpha}, \tau; N) = \max_{\substack{1 \leq m \leq N; \\ \boldsymbol{a} \in [0,1]^2; \\ x_1/x_2 = \tau;}} |D_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; m)|.$$

Given an arbitray positive increasing function $\varphi(n)$ of n, then for almost every $\boldsymbol{\alpha} \in \mathbb{R}^2$ and almost every $\tau \in \mathbb{R}^+$, we have:

$$\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} < \infty \ \Rightarrow\ \Delta(\boldsymbol{\alpha}, \tau; N) \ll (\ln N)^2 \cdot \varphi^2(\ln\ln N), \tag{4.2}$$

$$\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} = \infty \ \Rightarrow\ \Delta(\boldsymbol{\alpha}, \tau; N) > (\ln N)^2 \cdot \varphi(\ln\ln N) \ \textit{i.o.}\ , \tag{4.3}$$

where i.o. stands for infinitely often, and the constant in (4.2) depends on $\boldsymbol{\alpha}$ and τ.

Remark. *Unfortunately the statement does not give a perfect equivalence, the additional factor $\varphi(\ln\ln N)$ is needed for controlling the extra small divisors that arise from the hypotenuse of the triangle, which does not occur in the case of straight rectangles.*

Remark. *In fact, after fixing a toral translation, the bound of the discrepancy is uniform for any starting point $\boldsymbol{a}$, therefore it is equivalent to translate the right triangle $\triangle_{\mathbf{x}}$ inside $\mathbb{T}^2$, and since every general shaped triangle could be decompsed into a finite sum or difference of right triangles, it is immediate that the theorem above hold true for almost every triangle inside $\mathbb{T}^2$, hence we obtain a corollary:*

Corollary 4.2. *The same bounds in Theorem 4.1 hold for almost every triangle $\in \mathbb{T}^2$ (or more generally almost every polygon), with a constant that depends on the slopes of the sides of the triangle (or the polygon).*

This chapter is organized as the following: in Section 2, for the convenience of later estimations, we transform the ergodic discrepancy to its Fourier series by using Poisson's

summation formula. In Section 3 we estimate the contribution of the "tail" of the Fourier series, i.e. the high frequency modes or the extremely small divisors. Section 4 and 5 deal with the main part of the discrepancy, Section 4 is about the constant part and Section 5 deals with the exponential part, both of which will be properly defined later. Combining the Section 3-5, we have an overall estimation of the discrepancy, which proves our main theorem.

4.2 Poisson's summation formula

In this section, following [Bec94], we use the Poisson formula to transform the ergodic discrepancy into a Fourier series. The main result of this section is Proposition 4.1. It gives the main subject of our analysis: a roof-like average of the original Fourier series of the discrepancy, which has better convergence properties. Again, we adopt a heuristic way to obtain a *formal* Fourier series of the ergodic discrepancy. By the Poisson formula, without considering possible problems with convergence, we can write the ergodic sum $D(\boldsymbol{\alpha}, x; N)$ as the following formal series:

$$D_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; N) = \frac{\mathbf{i}^3}{(2\pi)^3} \sum_{\mathbf{n}\in\mathbb{Z}^3\setminus\{0\}} \prod_{j=1}^{2} e^{-2\pi\mathbf{i}n_j a_j} \left(\frac{1 - e^{2\pi\mathbf{i}n_2 x_2}}{n_1 n_2} \frac{e^{-2\pi\mathbf{i}(n_1\alpha_1 + n_2\alpha_2 - n_3)N} - 1}{n_1\alpha_1 + n_2\alpha_2 - n_3} \right)$$
$$+ \frac{\mathbf{i}^3}{(2\pi)^3} \sum_{\mathbf{n}\in\mathbb{Z}^3\setminus\{0\}} \prod_{j=1}^{2} e^{-2\pi\mathbf{i}n_j a_j} \left(\frac{e^{2\pi\mathbf{i}n_1 x_1} - e^{2\pi\mathbf{i}n_2 x_2}}{n_1(n_1 \frac{x_1}{x_2} - n_2)} \frac{e^{-2\pi\mathbf{i}(n_1\alpha_1 + n_2\alpha_2 - n_3)N} - 1}{n_1\alpha_1 + n_2\alpha_2 - n_3} \right).$$

To see this, the ergodic sum could be calculated as follows, note that the condition: $n\boldsymbol{\alpha}$ mod $1 \in \triangle_{\mathbf{x}}$ is equivalent to $\exists \mathbf{m} = (m_1, m_2, m_3) \in \mathbb{Z}^3$ such that:

$$\begin{gathered} 0 \le m_1\alpha_i - m_{i+1} < x_i;\ i = 1, 2; \\ \textstyle\sum_{i=1}^{2} \frac{m_1\alpha_i - m_{i+1}}{x_i} < 1; \\ 1 \le m_1 \le N; \end{gathered}$$

Consider the translated lattice in $\mathbb{R}^3$,

$$L(\boldsymbol{a}, \boldsymbol{\alpha}) = \{(a_1 + m_1\alpha_1 - m_2, a_2 + m_1\alpha_2 - m_3, m_1) \mid \mathbf{m} = (m_1, m_2, m_3) \in \mathbb{Z}^3\},$$

note that the fractional part of the vector $n\boldsymbol{\alpha}$ lying in $\triangle_{\mathbf{x}}$ is equivalent to the lattice point inside the following set:

$$B(\mathbf{x}, N) = \triangle_{\mathbf{x}} \times (0, N]$$

Let $\chi(\mathbf{y}) = \chi_{\mathbf{x},N}(\mathbf{y})$ be the characteristic function of the box $B(\mathbf{x}, N)$, so the ergodic

sum becomes

$$\sum_{n=1}^{N} \chi_{\triangle_{\mathbf{x}}}(\boldsymbol{a}+n\boldsymbol{\alpha} \mod 1) = \sum_{\mathbf{m}\in\mathbb{Z}^3} \chi(a_1+m_1\alpha_1-m_2, a_2+m_1\alpha_2-m_3, m_1)$$

Writing in matrix form:

$$(a_1+m_1\alpha_1-m_2, a_2+m_1\alpha_2-m_3, m_1) = \begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} + \begin{pmatrix} \alpha_1 & -1 & 0 \\ \alpha_2 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} = \boldsymbol{a} + \mathbf{A}\cdot\mathbf{m}$$

where

$$\mathbf{A} = \begin{pmatrix} \alpha_1 & -1 & 0 \\ \alpha_2 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{m} = \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix},$$

Apply the Poisson formula to the function $f_{\boldsymbol{a}}(\mathbf{y}) = \chi_{\mathbf{x},N}(\boldsymbol{a}+\mathbf{A}\cdot\mathbf{y})$, we have

$$\begin{aligned}
\sum_{\mathbf{m}\in\mathbb{Z}^3} \chi(\boldsymbol{a}+\mathbf{A}\cdot\mathbf{m}) &= \sum_{\mathbf{m}\in\mathbb{Z}^3} f_{\boldsymbol{a}}(\mathbf{m}) \\
&= \sum_{\boldsymbol{\nu}\in\mathbb{Z}^3} \int_{\mathbb{R}^3} f_{\boldsymbol{a}}(\mathbf{y}) \cdot e^{-2\pi\mathbf{i}\boldsymbol{\nu}\cdot\mathbf{y}} d\mathbf{y} \\
&= \sum_{\boldsymbol{\nu}\in\mathbb{Z}^3} \int_{\mathbb{R}^3} \chi(\boldsymbol{a}+\mathbf{A}\cdot\mathbf{y}) \cdot e^{-2\pi\mathbf{i}\boldsymbol{\nu}\cdot\mathbf{y}} d\mathbf{y} \\
&= \frac{1}{\det \mathbf{A}} \sum_{\boldsymbol{\nu}\in\mathbb{Z}^3} \int_{\mathbb{R}^3} \chi(\mathbf{z}) e^{-2\pi\mathbf{i}\boldsymbol{\nu}\cdot(\mathbf{A}^{-1}\cdot(\mathbf{z}-\boldsymbol{a}))} d\mathbf{z}
\end{aligned}$$

Integrate over the triangle $\triangle_{\mathbf{x}}$ for the first 2 coordinates of $\mathbf{z}$ and integrate over $(0, N]$ for the last coordinate z_3, we get:

$$\begin{aligned}
D_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; N) &= \sum_{\mathbf{m}\in\mathbb{Z}^3} f(\mathbf{m}) - N\mathrm{Vol}(\triangle_{\mathbf{x}}) \\
&= \frac{\mathbf{i}^3}{(2\pi)^3} \sum_{\mathbf{n}\in Z^3\setminus\{0\}} \prod_{j=1}^{2} e^{-2\pi\mathbf{i}n_j a_j} \left(-\frac{e^{2\pi\mathbf{i}n_1x_1}-1}{n_1n_2} \frac{e^{-2\pi\mathbf{i}(n_1\alpha_1+n_2\alpha_2-n_3)N}-1}{n_1\alpha_1+n_2\alpha_2-n_3} \right) \\
&+ \frac{\mathbf{i}^3}{(2\pi)^3} \sum_{\mathbf{n}\in Z^3\setminus\{0\}} \prod_{j=1}^{2} e^{-2\pi\mathbf{i}n_j a_j} \left(\frac{e^{2\pi\mathbf{i}n_1x_1}-e^{2\pi\mathbf{i}n_2x_2}}{(n_1-\frac{x_2}{x_1}n_2)n_2} \frac{e^{-2\pi\mathbf{i}(n_1\alpha_1+n_2\alpha_2-n_3)N}-1}{n_1\alpha_1+n_2\alpha_2-n_3} \right).
\end{aligned}$$

In order to avoid technical problems with the convergence, we will not study $D_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; N)$ directly, instead, we will follow Beck[Bec94] and use a special weighted average of $D_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; m)$ over a $\frac{1}{N^2}$ neighborhood. To this end, we oscillate starting point $\boldsymbol{a}$ with an amplitude of $\frac{1}{N^2}$, and also the range for summation $\{1, \ldots, N\}$ with amplitude of 2, and then by using

the Féjer kernal, we obtain our main object of interest. Specifically, let $\mathbf{u} = (u_1, u_2)$ be the oscillation of the starting point, u_3 be the oscillation of the summation range, then for the oscillated discrepancy:

$$D_{\triangle_{\mathbf{x}}}(\boldsymbol{a}+\mathbf{u}, \boldsymbol{\alpha}; u_3, N+u_3) = \sum_{u_3 < n \leq N+u_3} \chi_{\triangle_{\mathbf{x}}}(\boldsymbol{a}+\mathbf{u}+n\boldsymbol{\alpha} \mod 1) - N\mathrm{Vol}(\triangle_{\mathbf{x}}),$$

we define the $\frac{1}{N^2}$ average:

$$\bar{D}_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; N) = \left(\frac{N^2}{2}\right)^2 \cdot \frac{1}{2} \int_{-\frac{2}{N^2}}^{\frac{2}{N^2}} \int_{-\frac{2}{N^2}}^{\frac{2}{N^2}} \int_{-2}^{2} \prod_{j=1}^{2} \left(1 - \frac{N^2}{2}|u_j|\right) \cdot \left(1 - \frac{|u_3|}{2}\right)$$
$$\cdot D_{\triangle_{\mathbf{x}}}(\boldsymbol{a}+\mathbf{u}, \boldsymbol{\alpha}; u_3, N+u_3) du_1 du_2 du_3.$$

Using the Féjer kernel identity

$$\frac{N^2}{2} \int_{-\frac{2}{N^2}}^{\frac{2}{N^2}} \left(1 - \frac{N^2|y|}{2}\right) e^{2\pi \mathbf{i} k y} dy = \left(\frac{\sin(2\pi \frac{k}{N^2})}{2\pi \frac{k}{N^2}}\right)^2,$$

$$\int_{-2}^{2} \left(1 - \frac{|y|}{2}\right) e^{2\pi \mathbf{i} k y} dy = \left(\frac{\sin 2\pi k}{2\pi k}\right)^2.$$

We arrive at:

$$\begin{aligned}
&\bar{D}_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; N) \\
&= \frac{\mathbf{i}^3}{(2\pi)^3} \sum_{\mathbf{n} \in \mathbb{Z}^3 \setminus \{0\}} \left(\frac{1 - e^{2\pi \mathbf{i} n_2 x_2}}{n_1 n_2} \left(\frac{e^{-2\pi \mathbf{i}(n_1\alpha_1 + n_2\alpha_2 - n_3)N} - 1}{n_1\alpha_1 + n_2\alpha_2 - n_3}\right)\right) \\
&\cdot \prod_{j=1}^{2} \left(e^{-2\pi \mathbf{i} n_j a_j} \left(\frac{\sin(2\pi \frac{n_j}{N^2})}{2\pi \frac{n_j}{N^2}}\right)^2\right) \cdot \left(\frac{\sin 2\pi(n_1\alpha_1 + n_2\alpha_2 - n_3)}{2\pi(n_1\alpha_1 + n_2\alpha_2 - n_3)}\right)^2 \\
&+ \frac{\mathbf{i}^3}{(2\pi)^3} \sum_{\mathbf{n} \in \mathbb{Z}^3 \setminus \{0\}} \left(\frac{e^{2\pi \mathbf{i} n_1 x_1} - e^{2\pi \mathbf{i} n_2 x_2}}{n_1(n_1 \frac{x_1}{x_2} - n_2)} \frac{e^{-2\pi \mathbf{i}(n_1\alpha_1 + n_2\alpha_2 - n_3)N} - 1}{n_1\alpha_1 + n_2\alpha_2 - n_3}\right) \\
&\cdot \prod_{j=1}^{2} \left(e^{-2\pi \mathbf{i} n_j a_j} \left(\frac{\sin(2\pi \frac{n_j}{N^2})}{2\pi \frac{n_j}{N^2}}\right)^2\right) \cdot \left(\frac{\sin 2\pi(n_1\alpha_1 + n_2\alpha_2 - n_3)}{2\pi(n_1\alpha_1 + n_2\alpha_2 - n_3)}\right)^2 \\
&=: \sum_{\mathbf{n} \in \mathbb{Z}^3 \setminus \{0\}} f_1(\mathbf{n}, \boldsymbol{a}, \mathbf{x}, \boldsymbol{\alpha}, N) + \sum_{\mathbf{n} \in \mathbb{Z}^3 \setminus \{0\}} f_2(\mathbf{n}, \boldsymbol{a}, \mathbf{x}, \boldsymbol{\alpha}, N) \\
&=: \bar{D}_{\triangle_{\mathbf{x}},1}(\boldsymbol{a}, \boldsymbol{\alpha}; N) + \bar{D}_{\triangle_{\mathbf{x}},2}(\boldsymbol{a}, \boldsymbol{\alpha}; N).
\end{aligned} \tag{4.4}$$

We claim that the difference between $\bar{D}_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; N)$ and $D_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; N)$ is bounded by $C(\boldsymbol{\alpha}, \tau, \epsilon)(\ln N)^{1+\epsilon}$, with a constant that depends on $\boldsymbol{\alpha}$ and τ.

Proposition 4.1. *For almost every $\boldsymbol{\alpha} \in [0,1)^2$ and every $\epsilon > 0$, for a fixed slope $\tau = x_1/x_2$,*

there exists a constant $C(\boldsymbol{\alpha}, \tau, \epsilon)$, such that

$$|\bar{D}_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; N) - D_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; N)| < C(\boldsymbol{\alpha}, \tau, \epsilon)(\ln N)^{1+\epsilon}.$$

Proof. We will prove that for $(u_1, u_2, u_3) \in [-\frac{2}{N^2}, \frac{2}{N^2}] \times [-2, 2]$,

$$|D_{\triangle_{\mathbf{x}}}(\boldsymbol{a} + \mathbf{u}, \boldsymbol{\alpha}; u_3, N + u_3) - D_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; N)| < C(\boldsymbol{\alpha}, \tau, \epsilon)(\ln N)^{1+\epsilon}.$$

The important thing is that for given a slope τ, the above constant does not depend on u_i or x_i.

Note that translating the starting point is equivalent to translating the triangle, the difference above counts in fact the number of points of the translation which lie in the difference of two translated triangles. The difference between two triangles can be decomposed into three parts, the horizontal strip denoted by $H_{u;N}$, the vertical strip denoted by $V_{u;N}$, and the tilted strip denoted by $T_{u;N}$, it is sufficient to prove that each strip does not contain more points than the desired bound.

For $H_{u;N}$ and $V_{u;N}$, the number of points of the sequence $\{n\boldsymbol{\alpha} \mod 1 \mid 1 \leq n \leq N\}$ inside the strips is limited by the number of n's such that $\|n\alpha_1\| \leq 2/N^2$ or $\|n\alpha_2\| \leq 2/N^2$. Since the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ is convergent, by a standard application of Borel-Cantelli Lemma, the number of n's that solve the inequality $\|n\alpha_i\| < 2/n^2$ is finite for almost every $\alpha_i \in [0, 1)$, therefore there exists a constant $C(\boldsymbol{\alpha})$ such that:

$$\#\{1 \leq n \leq N \mid \|n\alpha_i\| < 2/N^2,\ i = 1 \text{ or } 2\} < C(\boldsymbol{\alpha}),$$

thus the number of points inside $H_{u;N}$ and $V_{u;N}$ has an upper bound $C(\boldsymbol{\alpha})$.

As for the tilted strip $T_{u;N}$, consider the vectors formed by any two different points inside $T_{u;N}$, then the vectors form a subset of the sequence $\{n\boldsymbol{\alpha} \mod 1\}_{-N \leq n \leq N}$, and they fall inside the the strip $T_{\tau;N}$ for any $|\mathbf{u}| \leq 2/N^2$, where $T_{\tau;N}$ is a parallelogram centered at the origin, with the slope τ, and of volume $\mathcal{O}(1/N^2)$. Since for any n, $\int_{\mathbb{T}^2} \chi_{T_{\tau;N}}(n\boldsymbol{\alpha} \mod 1) d\boldsymbol{\alpha} = \mathrm{Vol}(T_{\tau;N}) = \mathcal{O}(1/N^2)$, we have

$$\begin{aligned}
&\mathrm{mes}\left\{\boldsymbol{\alpha} \in \mathbb{T}^2 \;\middle|\; \sum_{n=-N}^{N} \chi_{T_{\tau;N}}(n\boldsymbol{\alpha} \mod 1) \geq (\ln N)^{1+\epsilon}\right\} \\
&\leq \frac{1}{(\ln N)^{1+\epsilon}} \sum_{n=-N}^{N} \mathrm{Vol}(T_{\tau;N}) \\
&= O\left(\frac{1}{(\ln N)^{1+\epsilon} N}\right),
\end{aligned}$$

which summable for $N \geq 1$, therefore by Borel-Cantelli Lemma, it happens only finitely many times that the desired bound $\#\{-N \leq n \leq N \mid (n\boldsymbol{\alpha} \mod 1) \in T_{\tau;N}\} < (\ln N)^{1+\epsilon}$ is

violated for almost every $\boldsymbol{\alpha} \in [0,1]^2$, which means that there exists a constant $C(\boldsymbol{\alpha}, \tau, \epsilon)$, such that:

$$\#\{-N \leq n \leq N \mid n\boldsymbol{\alpha} \mod 1 \in T_{\tau,N}\} \leq C(\boldsymbol{\alpha}, \tau, \epsilon)(\ln N)^{1+\epsilon}.$$

Since the vectors are no fewer than the points inside $T_{u;N}$ for any $\mathbf{u}$, we have the desired bound for all strips $T_{u;N}$ with a given slope τ. □

By Proposition 4.1, we can now shift our attention to the asymptotic behavior of $\bar{D}_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; N)$, which has better convergence property as shown in Section 4.

4.3 Local lemmas

In this section we introduce four lemmas that describes some "almost always" properties of $\boldsymbol{\alpha} \in \mathbb{R}^k$, these are modified versions of the lemmas in Section 4 of Beck[Bec94], and can be proved in the exact same way by modifying the denominators in the original proof of Beck.

The first lemma estimates a sum of the "large terms" in $\bar{D}_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; N)$ (4.4).

Lemma 4.1. *Let $\varphi(n)$ be an arbitrary positive increasing function of n with $\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} < \infty$. Then for almost every $\boldsymbol{\alpha} \in \mathbb{R}^2$, with $\mathbf{n}\boldsymbol{\alpha} = n_1\alpha_1 + n_2\alpha_2$, and every $\tau \in [0,1]$ irrational,*

$$\sum_{\mathbf{n} \in U(\boldsymbol{\alpha}, \tau; N)} (\max\{1, |n_1|\}|n_1\tau - n_2| \cdot \|\mathbf{n}\boldsymbol{\alpha}\|)^{-1} \ll (\ln N)^2 \cdot \varphi(\ln \ln N) \tag{4.5}$$

holds for all N, where

$$U(\boldsymbol{\alpha}, \tau; N) = \left\{ \mathbf{n} \in \mathbb{Z}^2 \backslash \{\mathbf{0}\} \;\middle|\; \begin{array}{l} |n_j| \leq N^2 \cdot (\ln N)^2, 1 \leq j \leq 2; \\ \max\{1, |n_j|\}|n_1\tau - n_2| \cdot \|\mathbf{n}\boldsymbol{\alpha}\| \leq (\ln N)^{40} \\ |n_1\tau - n_2| \geq 1/2 \end{array} \right\} \tag{4.6}$$

Proof. This is a modified version of Lemma 4.1 of Beck[Bec94], the key difference here is that n_2 is replaced by $n_1\tau - n_2$, but since the additional condition $|n_1\tau - n_2| \geq 1/2$ helps avoid the difficult situation when $\|n_1\tau\|$ is small, Beck's proof works here in the exact same way after replacing n_2 by $n_1\tau - n_2$. □

The following lemma is used to handle the "small terms" in $\bar{D}_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; N)$.

Let $\mathbf{V} = (v_1, v_2) \in \mathbb{Z}^2$ and $\mathbf{W} = (w_1, w_2) \in \mathbb{Z}^2$ satisfy $v_j < w_j$, $1 \leq j \leq 2$, that is ,

$\mathbf{V} \leq \mathbf{W}$. Let $Q(\mathbf{V}, \mathbf{W})$ denote the lattice box

$$Q(\tau; \mathbf{V}, \mathbf{W}) = \left\{ \mathbf{n} \in \mathbb{Z}^2 \;\middle|\; \begin{array}{l} v_1 \leq n_1 < w_1, \\ v_2 \leq n_1\tau - n_2 < w_2 \\ |n_1\tau - n_2| \geq 1/2 \end{array} \right\}. \tag{4.7}$$

For every real vector $\boldsymbol{\alpha} \in \mathbb{R}^2$, constant $C \geq 1$, integral vectors $\mathbf{V} \in \mathbb{Z}^2$ and $\mathbf{W} \in \mathbb{Z}^2$ with $\mathbf{V} < \mathbf{W}$, let $Z(\boldsymbol{\alpha}, \tau; C; \mathbf{V}, \mathbf{W})$ denote the number of integral solutions $\mathbf{n} \in Q(\tau; \mathbf{V}, \mathbf{W})$ of inequality below for the fractional part:

$$0 < \{\mathbf{n}\boldsymbol{\alpha}\} < \min\left\{\frac{1}{2}, \frac{C}{\max\{1, |n_1|\} \cdot |n_1\tau - n_2|}\right\}. \tag{4.8}$$

The "expected value" of $Z(\boldsymbol{\alpha}, \tau; C; \mathbf{V}, \mathbf{W})$, $\boldsymbol{\alpha} \in \mathbb{R}^2$, is

$$E(\tau; C; \mathbf{V}, \mathbf{W}) = \sum_{\mathbf{n} \in Q(\tau; \mathbf{V}, \mathbf{W})} \min\left\{\frac{1}{2}, \frac{C}{\max\{1, |n_j|\} \cdot |n_1\tau - n_2|}\right\}. \tag{4.9}$$

Lemma 4.2. *For almost every $\boldsymbol{\alpha} \in \mathbb{R}^2$, and every $\tau \in [0, 1]$ irrational,*

$$Z(\boldsymbol{\alpha}, \tau; C; \mathbf{V}, \mathbf{W}) = E(\tau, C; \mathbf{V}, \mathbf{W}) + \mathcal{O}(C^{\frac{3}{4}+\epsilon} \cdot (\ln N)^{1+\epsilon})$$

holds for all $C = q^4$, $q = 1, 2, 3, \ldots$, and for all lattice boxes $Q(\tau; \mathbf{V}, \mathbf{W})$ with $-N \leq \mathbf{V} < \mathbf{W} \leq N$, $N = 2, 3, 4 \ldots$.

It is evident that Lemma 4.2 remains true if we replace (4.8) by its complement:

$$1 - \min\left\{\frac{1}{2}, \frac{C}{\max\{1, |n_j|\} \cdot |n_1\tau - n_2|}\right\} < \{\mathbf{n}\boldsymbol{\alpha}\} < 1.$$

Proof. This is a modified version of Lemma 4.2 of Beck[Bec94], but similar to Lemma 3.1, the condition $|n_1\tau - n_2| \geq 1/2$ helps avoid the difficult situation when $\|n_1\tau\|$ is small, and Beck's proof works here in the exact same way after replacing n_2 by $n_1\tau - n_2$. □

The next lemma is special case of Khintchine's local criterion for linear forms.

Lemma 4.3. *Let $\varphi(n)$ be an arbitrary positive increasing function of n with $\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} = \infty$. Then for almost every $\boldsymbol{\alpha} \in \mathbb{R}^k$, there are infinitely many strictly positive integeral vectors $\mathbf{n} = (n_1, \ldots, n_k) \in \mathbb{N}^k$ such that*

$$n_1 \ldots n_k \cdot (\ln(n_1 \ldots n_k))^k \cdot \varphi(\ln\ln(n_1^2 \ldots n_k^2)) \cdot \|\mathbf{n}\boldsymbol{\alpha}\| < 1. \tag{4.10}$$

The last lemma helps control the divisors when $|n_1\tau - n_2| = \|n_1\tau\|$ is small and $|\mathbf{n}|$ is large:

Lemma 4.4. *Suppose that τ is irrational, then for almost every $\boldsymbol{\alpha} \in \mathbb{R}^2$, and almost every $\tau \in [0,1]$, the series*

$$\sum_{\mathbf{n}\in\mathbb{Z}^2\setminus\{\mathbf{0}\}} \left(\|\mathbf{n}\boldsymbol{\alpha}\| \cdot (\max\{1, |\ln|n_1|, \ln|n_2|\})^4 \cdot \max\{1, |n_1|\} \cdot |n_1\tau - n_2|\right)^{-1}$$

converges.

Proof. We denote the sum above by S, S can be divided into 2 parts: $S = S_1 + S_2$, where

$$S_1 = \sum_{\substack{n_1 \neq 0 \\ n_2 \,:\, |n_1\tau - n_2| = \|n_1\tau\|}} \left(\|\mathbf{n}\boldsymbol{\alpha}\| \cdot (\max\{1, |\ln|n_1|, \ln|n_2|\})^4 \cdot \max\{1, |n_1|\} \cdot |n_1\tau - n_2|\right)^{-1},$$

and

$$S_2 = \sum_{\mathbf{n}:|n_1\tau - n_2|\geq 1/2} \left(\|\mathbf{n}\boldsymbol{\alpha}\| \cdot (\max\{1, |\ln|n_1|, \ln|n_2|\})^4 \cdot \max\{1, |n_1|\} \cdot |n_1\tau - n_2|\right)^{-1}.$$

Let $\delta > 0$, and consider the integral with $\mathbf{n} \in \mathbb{Z}^2\setminus\{0\}$.

$$J_1(n) = \int_0^1 \int_0^1 \int_0^1 \left(\|\mathbf{n}\boldsymbol{\alpha}\| \, |\ln\|\mathbf{n}\boldsymbol{\alpha}\||^{1+\delta} \, \|n_1\tau\| \, |\ln\|n_1\tau\||^{1+\delta}\right)^{-1} d\alpha_1 d\alpha_2 d\tau.$$

The integral has an finite value independent of $\mathbf{n}$. Hence the sum

$$\sum_{\substack{n_1 \neq 0 \\ n_2 \,:\, |n_1\tau - n_2| = \|n_1\tau\|}} \frac{J_1(\mathbf{n})}{\max\{1, |n_1| \cdot (\ln|n_1|)^{1+\delta}\}}$$

is convergent, and so

$$\sum_{\substack{n_1 \neq 0 \\ n_2 \,:\, |n_1\tau - n_2| = \|n_1\tau\|}} \left(\|\mathbf{n}\boldsymbol{\alpha}\| \, |\ln\|\mathbf{n}\boldsymbol{\alpha}\||^{1+\delta} \, \|n_1\tau\| \, |\ln\|n_1\tau\||^{1+\delta} \max\{1, |n_1| \cdot (\ln|n_1|)^{1+\delta}\right)^{-1} \tag{4.11}$$

is convergent for almost every $\boldsymbol{\alpha} \in \mathbb{R}^2$, and almost every $\tau \in [0,1]$. Since $\sum_{n=1}^{\infty} n^{-2} = \mathcal{O}(1)$, the inequality

$$\|\mathbf{n}\boldsymbol{\alpha}\| < \left(\prod_{j=1}^{2} \max\{1, |n_j|\}\right)^{-2}$$

has only a finite number of integral solutions $\mathbf{n} \in \mathbb{Z}^2$ for almost every $\boldsymbol{\alpha} \in \mathbb{R}^2$. Hence

$$|\ln\|\mathbf{n}\boldsymbol{\alpha}\|| < 2\sum_{j=1}^{2} \max\{0, \ln|n_j|\} \ll \max\{0, \ln|n_1|\}, \tag{4.12}$$

the last part of the inequality comes from $|n_1\tau - n_2| \leq 1/2$. By the same reasoning we have:

$$|\ln \|n_1\tau\|| < 2\max\{0, \ln|n_1|\} \tag{4.13}$$

By (4.11), (4.12), and (4.13), we have the series:

$$S_1 \ll \sum_{\substack{n_1 \neq 0 \\ n_2 \,:\, |n_1\tau - n_2| = \|n_1\tau\|}} \left(\|\mathbf{n}\boldsymbol{\alpha}\| \, |\ln \|\mathbf{n}\boldsymbol{\alpha}\||^{1+\delta} \, \|n_1\tau\| \, |\ln \|n_1\tau\||^{1+\delta} \max\{1, |n_1| \cdot (\ln|n_1|)^{1+\delta}\right)^{-1},$$

which is convergent by taking $\delta = 1/3$.

For S_2, we employ the same method, the integral

$$J_2(n) = \int_0^1 \int_0^1 \left(\|\mathbf{n}\boldsymbol{\alpha}\| \, |\ln \|\mathbf{n}\boldsymbol{\alpha}\||^{1+\delta} \right)^{-1} d\alpha_1 d\alpha_2.$$

The integral has an finite value independent of $\mathbf{n}$. Hence the sum

$$\sum_{\mathbf{n} \,:\, |n_1\tau - n_2| \geq 1/2} \frac{J_2(\mathbf{n})}{n_1 \max\{1, (\ln|n_1|)^{1+\delta}\} |n_1\tau - n_2| \max\{1, (\ln|n_1\tau - n_2|)^{1+\delta}\}}$$

is convergent, and so

$$\sum_{\mathbf{n}:|n_1\tau - n_2| \geq 1/2} \frac{1}{\|\mathbf{n}\boldsymbol{\alpha}\| \, |\ln \|\mathbf{n}\boldsymbol{\alpha}\||^{1+\delta}} \cdot \frac{1}{n_1 \max\{1, (\ln|n_1|)^{1+\delta}\}} \cdot \frac{1}{|n_1\tau - n_2| \max\{1, (\ln|n_1\tau - n_2|)^{1+\delta}\}}$$

is convergent for almost every $\boldsymbol{\alpha}$ and every τ irrational. By (4.12) and $|n_1\tau - n_2| \leq |n_1| + |n_2| \leq 2\max\{|n_1|, |n_2|\}$, the series S_2 is convergent by taking $\delta = 1/3$. □

4.4 Estimating the "tail" of the discrepancy function

Recall that $\bar{D}_{\triangle_{\mathbf{x}}}(\boldsymbol{a}, \boldsymbol{\alpha}; N)$ is the sum of $\bar{D}_{\triangle_{\mathbf{x}},1}(\boldsymbol{a}, \boldsymbol{\alpha}; N)$ and $\bar{D}_{\triangle_{\mathbf{x}},2}(\boldsymbol{a}, \boldsymbol{\alpha}; N)$(see (4.4)), where $\bar{D}_{\triangle_{\mathbf{x}},1}(\boldsymbol{a}, \boldsymbol{\alpha}; N)$ is already dealt with in Beck[Bec94], see Lemma 4.1 below, this section is devoted to estimating the "tail" of the second part of sum $\bar{D}_{\triangle_{\mathbf{x}},2}(\boldsymbol{a}, \boldsymbol{\alpha}; N)$.

Lemma 4.5. *Let $\varphi(n)$ be an arbitrary positive increasing function of n with $\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} < \infty$, then for almost every $\boldsymbol{\alpha} \in \mathbb{T}^2$,*

$$\max_{\mathbf{x}, \boldsymbol{a} \in \mathbb{T}^2} |\bar{D}_{\triangle_{\mathbf{x}},1}(\boldsymbol{a}, \boldsymbol{\alpha}; N)| \ll (\ln N)^2 \varphi(\ln\ln N). \tag{4.14}$$

Proof. Note that $\bar{D}_{\triangle_{\mathbf{x}},1}(\boldsymbol{a}, \boldsymbol{\alpha}; N)$ in (4.4) is in fact part of the sum (3.12) in Beck[Bec94],

whose estimation is given by (6.1) for $\sum_9$ and (8.14) for $\sum_{10}$ in Beck[Bec94]. Our sum $\bar{D}_{\triangle_{\mathbf{x}},1}(\boldsymbol{a},\boldsymbol{\alpha};N)$ has an additional factor of $\prod_{j=1}^{2} e^{-2\pi \mathbf{i} n_j a_j}$, which can be estimated in the same fashion, or it could be understood as part of the discrepancy function for the translated boxes $\prod_{i=1}^{2}[a_i, a_i + x_i]$, which is a finite sum or difference of the original boxes of the form $\prod_{i=1}^{2}[0, x_i]$, therefore satisfies the same estimation as the original boxes. □

Note that $\bar{D}_{\triangle_{\mathbf{x}},2}(\boldsymbol{a},\boldsymbol{\alpha};N)$ is the sum of the products (where $\mathbf{n} \in \mathbb{Z}^3 \backslash \mathbf{0}$)

$$\begin{aligned} f_2(\mathbf{n},\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N) =& \frac{\mathbf{i}^3}{(2\pi)^3} \sum_{\mathbf{n}\in\mathbb{Z}^3\backslash\{0\}} \left(\frac{e^{2\pi\mathbf{i}n_1x_1} - e^{2\pi\mathbf{i}n_2x_2}}{n_1(n_1\frac{x_1}{x_2} - n_2)} \frac{e^{-2\pi\mathbf{i}(n_1\alpha_1+n_2\alpha_2-n_3)N} - 1}{n_1\alpha_1 + n_2\alpha_2 - n_3} \right) \\ & \cdot \prod_{j=1}^{2} \left(e^{-2\pi\mathbf{i}n_j a_j} \left(\frac{\sin(2\pi\frac{n_j}{N^2})}{2\pi\frac{n_j}{N^2}} \right)^2 \right) \cdot \left(\frac{\sin 2\pi(n_1\alpha_1 + n_2\alpha_2 - n_3)}{2\pi(n_1\alpha_1 + n_2\alpha_2 - n_3)} \right)^2 \end{aligned} \tag{4.15}$$

let:

$$\bar{D}_5(\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N) = \sum_{\mathbf{n}\in U_5(\boldsymbol{\alpha},\tau)} f_2(\mathbf{n},\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N),$$

and

$$U_5(\mathbf{n},\tau;N) = \left\{ \mathbf{n} \in \mathbb{Z}^3\backslash\{\mathbf{0}\} \left| \begin{array}{l} \max\{n_1, |n_1\tau - n_2|\} \le N^2/4, \\ \min\{n_1, |n_1\tau - n_2|\} \ge (\ln N)^{40}, \\ |\mathbf{n}\boldsymbol{\alpha} - n_3| = \|\mathbf{n}\boldsymbol{\alpha}\|, \ |n_1\tau - n_2| \ge 1/2, \\ |n_1|n_1\tau - n_2|\|\mathbf{n}\boldsymbol{\alpha}\| > (\ln N)^{40}. \end{array} \right. \right\} \tag{4.16}$$

The main result of this section is the following:

Proposition 4.2. *Let $\varphi(n)$ be an arbitrary positive increasing function of n with $\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} < \infty$, then for almost every $\boldsymbol{\alpha} \in [0,1]^2$, and almost every $\tau \in [0,1]$, we have*

$$|\bar{D}_5(\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N) - \bar{D}_{\triangle_{\mathbf{x}},2}(\boldsymbol{a},\boldsymbol{\alpha};N)| \ll (\ln N)^2 \varphi^2(\ln\ln N)$$

To prove Proposition 4.2, we need to control different components of $\bar{D}_5 - \bar{D}_{\triangle_{\mathbf{x}},2}$ step by step.

4.4.1 Estimation for the sum when $|n_1|$ or $|n_2|$ is large

Define

$$\bar{D}_1(\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N) = \sum_{\mathbf{n}\in U_1(\boldsymbol{\alpha};N)} f_2(\mathbf{n},\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N), \tag{4.17}$$

where

$$U_1(\boldsymbol{\alpha};N) = \left\{ \mathbf{n} \in \mathbb{Z}^3\backslash\{\mathbf{0}\} \mid \max\{|n_1|, |n_2|\} \le N^2(\ln N)^2 \right\}.$$

We define another set that is parallel to $U_{\boldsymbol{\alpha}}$ and abuse the notation $\mathbf{n}$ for (n_1, n_2) if n_3 is the closest integer to $n_1\alpha_1 + n_2\alpha_2$:

$$U_1'(\boldsymbol{\alpha}; N) = \left\{\mathbf{n} \in \mathbb{Z}^2 \backslash \{\mathbf{0}\} \mid \max\{|n_1|, |n_2|\} \leq N^2 (\ln N)^2\right\}.$$

We show the following:

Proposition 4.3. *For almost every $\boldsymbol{\alpha} \in \mathbb{R}^2$, and almost every $\tau \in [0,1]$, we have*

$$\left|\bar{D}_1(\mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N) - \bar{D}_{\triangle_{\mathbf{x}},2}(\boldsymbol{a}, \boldsymbol{\alpha}; N)\right| = \mathcal{O}(1),$$

where $\mathcal{O}(1)$ represents an absolute bound which depends on $\boldsymbol{\alpha}$, τ, but does not depend on $\boldsymbol{a}$ or N.

Proof. First we decompose the difference into 2 parts:

$$\left|\bar{D}_1(\mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N) - \bar{D}_{\triangle_{\mathbf{x}},2}(\boldsymbol{a}, \boldsymbol{\alpha}; N)\right| \ll \bar{D}_{1,1} + \bar{D}_{1,2}, \tag{4.18}$$

where

$$\bar{D}_{1,1} = \sum_{\mathbf{n} \notin U_1'(\boldsymbol{\alpha})} \left(\|\mathbf{n}\boldsymbol{\alpha}\| |n_1| |n_1\tau - n_2| \prod_{j=1} (\max\{1, \frac{n_j}{N^2}\})^2 \right)^{-1}$$

$$\bar{D}_{1,2} = \sum_{r=1}^{\infty} \sum_{\mathbf{n} \notin U_1'(\boldsymbol{\alpha})} \left(|n_1| |n_1\tau - n_2| \prod_{j=1} (\max\{1, \frac{n_j}{N^2}\})^2 \right)^{-1} \cdot \frac{1}{r^3},$$

where $r = \lceil |n_1\alpha_1 + n_2\alpha_2 - n_3| \rceil$.

To estimate $\bar{D}_{1,1}$, we employ Lemma 4.4. For every $\boldsymbol{\alpha} \in \mathbb{R}^2$ and every $\tau \in [0,1]$ that satisfy Lemma 4.4,

$$\begin{aligned}
\bar{D}_{1,1} &\ll \sum_{\mathbf{n} \notin U_1'(\boldsymbol{\alpha})} \left(\|\mathbf{n}\boldsymbol{\alpha}\| \cdot (\max\{1, |\ln|n_1|, \ln|n_2|\})^4 \cdot \max\{1, |n_1|\} \cdot |n_1\tau - n_2|\right)^{-1} \\
&\leq \sum_{\mathbf{n} \in \mathbb{Z}^2 \backslash \{\mathbf{0}\}} \left(\|\mathbf{n}\boldsymbol{\alpha}\| \cdot (\max\{1, |\ln|n_1|, \ln|n_2|\})^4 \cdot \max\{1, |n_1|\} \cdot |n_1\tau - n_2|\right)^{-1} \\
&= \mathcal{O}(1).
\end{aligned}$$

$\bar{D}_{1,2}$ can be easily controlled by $\bar{D}_{1,1}$:

$$\begin{aligned}
\bar{D}_{1,2} &\ll \left(\sum_{r=1}^{\infty} \frac{1}{r^3}\right) \left(\sum_{\mathbf{n} \notin U_1'(\boldsymbol{\alpha})} \left(|n_1| |n_1\tau - n_2| \prod_{j=1} (\max\{1, \frac{n_j}{N^2}\})^2 \right)^{-1} \right) \\
&\ll \bar{D}_{1,1} \\
&= \mathcal{O}(1).
\end{aligned}$$

□

From the proof it is easy to see that the series $\bar{D}_{\triangle_{\mathbf{x}},2}(\boldsymbol{a},\boldsymbol{\alpha};N)$ is absolutely convergent for almost every $\alpha \in \mathbb{R}^2$ and $\tau \in [0,1]$.

4.4.2 Estimation for the sum when $|\mathbf{n}\boldsymbol{\alpha} - n_3|$ is larger than $1/3$.

Let

$$\bar{D}_2(\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N) = \sum_{\mathbf{n}\in U_2(\boldsymbol{\alpha};N)} f_2(\mathbf{n},\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N)$$

where $\mathbf{n} = (n_1, n_2, n_3)$ satisfies

$$U_2(\boldsymbol{\alpha};N) = \left\{\mathbf{n} \in \mathbb{Z}^3 \mid \max\{|n_1|,|n_2|\} \le N^2(\ln N)^2, \quad |\mathbf{n}\boldsymbol{\alpha} - n_3| < \frac{1}{3}\right\} \tag{4.19}$$

We show that the difference between $\bar{D}_1$ and $\bar{D}_2$ can be well controlled, hence we only need to focus on the sum over (n_1, n_2) instead of (n_1, n_2, n_3):

Proposition 4.4. *Let $\varphi(n)$ be an arbitrary positive increasing function of n with $\sum_{n=1}^{\infty}\frac{1}{\varphi(n)} < \infty$, then for almost every $\boldsymbol{\alpha} \in \mathbb{R}^2$ and almost every $\tau \in [0,1]$, we have*

$$\left|\bar{D}_2(\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N) - \bar{D}_1(\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N)\right| \ll (\ln N)^2\varphi(\ln\ln N),$$

Proof. Similar to the previous proposition, the difference above can be decomposed into 2 parts:

$$\begin{aligned}
&\left|\bar{D}_1(\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N) - \bar{D}_2(\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N)\right| \\
&\ll \sum_{r=1}^{\infty}\sum_{\mathbf{n}\in U_1'(\boldsymbol{\alpha})}\left(|n_1||n_1\tau - n_2|\prod_{j=1}(\max\{1,\frac{n_j}{N^2}\})^2\right)^{-1}\cdot\frac{1}{r^3} \qquad (4.20)\\
&\ll \bar{D}_{2,1} + \bar{D}_{2,2},
\end{aligned}$$

where

$$\bar{D}_{2,1} = \sum_{1\le|n_1|\le N^2(\ln N)^2}\frac{1}{|n_1|\|n_1\tau\|},$$

and

$$\bar{D}_{2,2} = \sum_{\mathbf{n}\in U_1'(\boldsymbol{\alpha}):|n_1\tau-n_2|\ge 1/2}\frac{1}{|n_1||n_1\tau - n_2|}.$$

$\bar{D}_{2,2}$ can be estimated easily:

$$\bar{D}_{2,2} \ll \prod_{j=1}^{2}\left(\sum_{1\le|n_j|\le N^2(\ln N)^2}\frac{1}{|n_j|}\right) \ll (\ln N)^2.$$

For $\bar{D}_{2,1}$, we apply the continued fraction algorithm as in the proof for (9.1) in Beck[Bec94],

see [Lan95] for the basics of the theory of continued fractions. Let a_j be the j-th partial quotient of τ and q_j be the dnominator of the j-th convergent, then by Dirichlet's "box principle" and the estimation for the j-th convergent:

$$\frac{1}{2q_{j+1}} \leq \|q_j\tau\| \leq \frac{1}{q_{j+1}},$$

we have:

$$\sum_{n_1=q_j}^{q_{j+1}} \frac{1}{n_1\|n_1\tau\|} \ll \frac{q_{j+1}}{q_j} \sum_{k=1}^{q_{j+1}-q_j-1} \frac{1}{k} \ll a_{j+1} \cdot \ln(q_{j+1}).$$

Let s be the integer such that $q_s \leq N^{2+\epsilon} < q_{s+1}$, then for almost every τ, we have

$$s \ll \ln N.$$

combined with $\ln q_j \ll j$ for almost every τ, we have

$$\sum_{n_1=1}^{N^2(\ln N)^2} \frac{1}{n_1\|n_1\tau\|} \ll \left(\sum_{j=1}^{s} a_{j+1}\right) \ln N.$$

By a theorem of Khintchine about the the sum of partial quotient (see [Kok13], p.46), if $\psi(n)$ is any increasing function, then

$$a_1(\tau) + \cdots + a_s(\tau) \ll s \cdot \psi(s) \Longleftrightarrow \sum_{n=1}^{\infty} \frac{1}{n \cdot \psi(n)} < \infty$$

for almost every τ. Let $\psi(n) = \varphi(\ln n)$, then

$$\sum_{n_1=1}^{N^2(\ln N)^2} \frac{1}{n_1\|n_1\tau\|} \ll \left(\sum_{j=1}^{s} a_{j+1}\right) \ln N \ll (\ln N)^2 \varphi(\ln\ln N).$$

□

4.4.3 Estimation for the sum when $|n_1\tau - n_2| = \|n_1\tau\|$.

An important difference between the case of triangles and the case of boxes is that the divisor changes from $n_1n_2(n_1\alpha_1 + n_2\alpha_2 - n_3)$ to $n_1(n_1\tau - n_2)(n_1\alpha_1 + n_2\alpha_2 - n_3)$, which adds a possibility for $|n_1\tau - n_2|$ to be small. In this section, we deal with the difference by some metrical properties of $n_1\|n_1\alpha_1 + n_2\alpha_2\|$, where n_2 is the closest integer to $n_1\tau$. Unfortunately, this would lead to a loss of $\varphi(\ln\ln N)$ in the estimation.

Let

$$\bar{D}_3(\mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N) = \sum_{\mathbf{n} \in U_3(\alpha,\tau)} f_2(\mathbf{n}, \mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N)$$

where $\mathbf{n} = (n_1, n_2, n_3)$ satisfies

$$U_3(\boldsymbol{\alpha}, \tau; N) = \left\{ \mathbf{n} \in \mathbb{Z}^3 \;\middle|\; \begin{array}{l} \max\{|n_1|, |n_2|\} \leq N^2(\ln N)^2, \\ |\mathbf{n}\boldsymbol{\alpha} - n_3| < \dfrac{1}{3}, \\ |n_1\tau - n_2| = \|n_1\tau\| \end{array} \right\} \tag{4.21}$$

We show that the difference between $\bar{D}_3$ and $\bar{D}_2$ can be controlled by the following proposition, hence we only need to focus on the sum over (n_1, n_2) when $|n_1\tau - n_2| \geq 1/2$, which could be treated in the same way as Beck did.

Proposition 4.5. *Let $\varphi(n)$ be an arbitrary positive increasing function of n with $\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} < \infty$, then for almost every $\boldsymbol{\alpha} \in \mathbb{R}^2$ and almost every $\tau \in [0, 1]$, we have*

$$\left|\bar{D}_3(\mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N) - \bar{D}_2(\mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N)\right| \ll (\ln N)^2 \varphi^2(\ln\ln N),$$

Proof. Note that $\tau = x_1/x_2$, and $x_2 \in (0, 1]$, we use the inequality:

$$\left|\frac{e^{2\pi \mathbf{i} n_1 x_1} - e^{n_2 x_2}}{n_1\tau - n_2}\right| = \left|x_2 \frac{e^{2\pi \mathbf{i} n_1 x_1} - e^{2\pi \mathbf{i} n_2 x_2}}{n_1 x_1 - n_2 x_2}\right| \leq 1,$$

therefore $|f_2(\mathbf{n}, \mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N)| \ll \frac{1}{|n_1|\|\mathbf{n}\boldsymbol{\alpha}\|}$.

Let n_2^* denote the closest integer to $n_1\tau$, then

$$|\bar{D}_3(\mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N) - \bar{D}_2(\mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N)| \ll \sum_{1 \leq |n_1| \leq N^2(\ln N)^2} \frac{1}{|n_1|\|n_1\alpha_1 + n_2^*\alpha_2\|}$$

we employ the same method as in Lemma 4.4, the integral

$$J_2(\mathbf{n}) = \int_0^1 \int_0^1 \left(\|\mathbf{n}\boldsymbol{\alpha}\| \, |\ln \|\mathbf{n}\boldsymbol{\alpha}\|| \, \varphi(\ln\ln \|\mathbf{n}\boldsymbol{\alpha}\|)\right)^{-1} d\alpha_1 d\alpha_2.$$

has an finite value independent of $\mathbf{n}$. Hence the sum

$$\sum_{|n_1| = 1}^{\infty} \frac{J_2(n_1, n_2^*)}{|n_1| \ln |n_1| \varphi(\ln\ln |n_1|)}$$

is convergent, and so for almost every $\boldsymbol{\alpha}$, the series

$$S_3 = \sum_{|n_1|=1}^{\infty} \left(|n_1| \ln |n_1| \varphi(\ln\ln |n_1|) \|\mathbf{n}\boldsymbol{\alpha}\| \, |\ln \|\mathbf{n}\boldsymbol{\alpha}\|| \, \varphi(\ln\ln \|\mathbf{n}\boldsymbol{\alpha}\|)\right)^{-1}$$

is convergent, here $\mathbf{n}\boldsymbol{\alpha}{=}n_1\alpha_1 + n_2^*\alpha_2$. Combined with the inequality (4.12), we have for

almost every $\boldsymbol{\alpha} \in \mathbb{R}^2$

$$\begin{aligned}
&\sum_{1\le|n_1|\le N^2(\ln N)^2} \frac{1}{|n_1|\|n_1\alpha_1+n_2^*\alpha_2\|} \\
&\ll \sum_{|n_1|=1}^{N^2(\ln N)^2} \frac{(\ln N)^2\varphi^2(\ln\ln N)}{|n_1|\ln|n_1|\varphi(\ln\ln|n_1|)\|\mathbf{n}\boldsymbol{\alpha}\|\,|\ln\|\mathbf{n}\boldsymbol{\alpha}\||\,\varphi(\ln\ln\|\mathbf{n}\boldsymbol{\alpha}\|)} \\
&\ll (\ln N)^2\varphi^2(\ln\ln N)\cdot S_3 \\
&\ll (\ln N)^2\varphi^2(\ln\ln N)
\end{aligned}$$

□

4.4.4 Estimation for the sum when $|n_1||n_1\tau - n_2|\|\mathbf{n}\boldsymbol{\alpha}\|$ is small.

From now on, we can restrcit the range of our discussion to $\mathbf{n} = (n_1, n_2)$, Define:

$$\bar{D}_4(\mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N) = \sum_{\mathbf{n}\in U_4(\mathbf{n},\tau;N)} f_2(\mathbf{n}, \mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N) \tag{4.22}$$

where

$$U_4(\mathbf{n}, \tau; N) = \left\{ \mathbf{n} \in \mathbb{Z}^3\backslash\{\mathbf{0}\} \left| \begin{array}{l} \max\{n_1, n_2\} \le N^2(\ln N)^2, \\ |\mathbf{n}\boldsymbol{\alpha} - n_3| = \|\mathbf{n}\boldsymbol{\alpha}\|,\ |n_1\tau - n_2| \ge 1/2, \\ |n_1|n_1\tau - n_2|\|\mathbf{n}\boldsymbol{\alpha}\| < (\ln N)^{40} \end{array} \right. \right\} \tag{4.23}$$

From Lemma 4.1, we have the following estimation:

Proposition 4.6. *Let $\varphi(n)$ be an arbitrary positive increasing function of n with $\sum_{n=1}^{\infty}\frac{1}{\varphi(n)} < \infty$, then for almost every $\boldsymbol{\alpha} \in \mathbb{R}^2$ and every $\tau \in [0, 1]$ irrational, we have*

$$\left|\bar{D}_4(\mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N) - \bar{D}_3(\mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N)\right| \ll (\ln N)^2\varphi(\ln\ln N).$$

where the constant may depend on $\boldsymbol{\alpha}$ and τ.

4.4.5 Estimation for the sum when n_1 or n_2 is between $N^2/4$ and $N^2(\ln N)^2$.

Define:

$$\bar{D}_5(\mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N) = \sum_{\mathbf{n}\in U_5(\mathbf{n},\tau;N)} f_2(\mathbf{n}, \mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N)$$

where we recall that

$$U_5(\mathbf{n},\tau;N)=\left\{\mathbf{n}\in\mathbb{Z}^3\backslash\{\mathbf{0}\}\left|\begin{array}{l}\max\{n_1,|n_1\tau-n_2|\}\le N^2/4,\\ \min\{n_1,|n_1\tau-n_2|\}\ge(\ln N)^{40},\\ |\mathbf{n}\boldsymbol{\alpha}-n_3|=\|\mathbf{n}\boldsymbol{\alpha}\|,\ |n_1\tau-n_2|\ge 1/2,\\ |n_1|n_1\tau-n_2|\|\mathbf{n}\boldsymbol{\alpha}\|>(\ln N)^{40}.\end{array}\right.\right\}$$

The goal of this step is to prove the following:

Proposition 4.7. *For almost every $\boldsymbol{\alpha}\in\mathbb{R}^2$, we have*

$$\left|\bar{D}_5(\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N)-\bar{D}_4(\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N)\right|\ll(\ln N)^2\ln\ln N$$

By limiting the range of $(2\pi n_1/N^2)$ in $f_2(\mathbf{n},\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N)$ to $(-\pi/2,\pi/2)$, f_2 becomes better-behaved for later estimations.

Proof. For every $\mathbf{v}\in\mathbb{Z}^2$, with $1\le 2^{v_j}\le N^2\cdot(\ln N)^k$, $1\le j\le 2$, and every $\epsilon=(\epsilon_1,\epsilon_2)\in\{-1,+1\}^2$, let

$$T(\tau;\mathbf{v},\epsilon)=\{\mathbf{n}\in\mathbb{Z}^2:2^{v_1-1}\le\epsilon_1 n_1<2^{v_1},\ 2^{v_2-1}\le\epsilon_1(n_1\tau-n_2)<2^{v_2}\}$$

with the convention that if $v_1=0$ then the requirement above means $n_1=0$, and since $|n_1\tau-n_2|\ge 1/2$ we have $v_2\ge 0$. Note that $0\le v_j\ll\ln N$, $1\le j\le 2$. Let

$$V_1(N)=\{\mathbf{v}\in\mathbb{Z}^2:\min_{1\le j\le 2}v_j\le c^*\ln\ln N\text{ with }c^*=\frac{20k}{\ln 2}\},$$

and

$$V_2(N)=\{\mathbf{v}\in\mathbb{Z}^2:\exists j\in\{1,2\}\text{ such that }\frac{N^2}{4}\le 2^{v_j}\le N^2\cdot(\ln N)^k\}.$$

Let $C=2^q$, q an integer with $2^q>(\ln N)^{40}=2^{c^*\ln\ln N}$, and write

$$T(\tau;\mathbf{v},\epsilon,q)=\left\{\mathbf{n}\in T(\tau;\mathbf{v},\epsilon):2^{q-1}<\max\{1,|n_1|\}|n_1\tau-n_2|\cdot\|\mathbf{n}\boldsymbol{\alpha}\|\le 2^q\right\}$$

We use Lemma 3.2 to have the estimation for the cardinality for $T(\tau;\mathbf{v},\epsilon,q)$: for almost every $\boldsymbol{\alpha}\in\mathbb{R}^2$, since $2^q>(\ln N)^{40}$, we have

$$\begin{aligned}|T(\tau;\mathbf{v},\epsilon,q)|\ll&\sum_{\mathbf{n}\in T(\tau;\mathbf{v},\epsilon)}\min\left\{1,\frac{2^q}{\max\{1,|n_1|\}|n_1\tau-n_2|}\right\}\\&+\mathcal{O}((2^q)^{\frac{3}{4}+\epsilon}\cdot(\ln N)^{\frac{1}{2}+\epsilon})\ll 2^q.\end{aligned}$$

Note that $|V_j|=\mathcal{O}(\ln N\cdot\ln\ln N)$, $j=1,2$; and there are $\mathcal{O}(\ln N)$ choices for q since $2^q\le(N^2(\ln N)^2)^2$. Combined with the fact that each set of $T(\tau;\mathbf{v},\epsilon,q)$ would contribute

$\mathcal{O}(1)$, the difference can be estimated by the following:

$$\begin{aligned}
&|\bar{D}_5 - \bar{D}_4| \\
&\ll \sum_{\epsilon\in\{-1,+1\}^2} \left(\sum_{\mathbf{v}\in V_1(N)} + \sum_{\mathbf{v}\in V_2(N)} \right) \\
&\sum_{q:2^q\le(N^2\cdot(\ln N)^2)^2} \sum_{\mathbf{n}\in T(\tau;\mathbf{v},\epsilon,q)} \frac{1}{\|\mathbf{n}\boldsymbol{\alpha}\|\cdot\max\{1,|n_1|\}|n_1\tau - n_2|} \\
&\ll \ln N\cdot\ln\ln N\cdot \sum_{q:2^q\le(N^2\cdot(\ln N)^2)^2} \frac{2^q}{2^{q-1}} \\
&\ll (\ln N)^2\cdot\ln\ln N.
\end{aligned}$$

□

In next section we will tackle the main difficulty for the sum of $f_2(\mathbf{n},\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N)$ in U_5.

4.5 Estimation of the exponential sums.

The form of the product $f_2(\mathbf{n},\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N)$ saves us the work for the constant part which is dealt with in section 6 of Beck[Bec94], and it remains to verify that everything in Beck's proof for the exponential part also works for the shifted divisor $n_1(n_1\tau - n_2)(\mathbf{n}\boldsymbol{\alpha} - n_3)$. Since in Section 4 we already controlled the terms where $|n_1\tau - n_2| = \|n_1\tau\|$ which could be very small, the remaining sum would behave similarly as the exponential sums in the case of boxes.

To estimate the contribution of the remaining sum, first we highlight the essentials in the sum $\bar{D}_5$, which can be represented as follows:

$$\bar{D}_5(\mathbf{x},\boldsymbol{a},\boldsymbol{\alpha};N) = \frac{\mathbf{i}^3}{(2\pi)^3}\left(\sum_{\mathbf{s}} \pm \sum_{\mathbf{n}\in U_5(\mathbf{n},\tau;N)} \frac{e^{2\pi\mathbf{i}\mathcal{L}_{\mathbf{s}}(\mathbf{n})}}{n_1(n_1\tau - n_2)(\mathbf{n}\boldsymbol{\alpha} - n_3)}\cdot g(\mathbf{n},N,\boldsymbol{\alpha}) \right) \tag{4.24}$$

where $g(\mathbf{n},N,\boldsymbol{\alpha})$ is the product below (observe that $|g(\mathbf{n},N,\boldsymbol{\alpha})| \le 1$):

$$\left(\frac{\sin 2\pi(\mathbf{n}\boldsymbol{\alpha} - n_3)}{2\pi(\mathbf{n}\boldsymbol{\alpha} - n_3)}\right)^2 \cdot \prod_{j=1}^{2}\left(\frac{\sin 2\pi(\frac{n_j}{N^2})}{2\pi(\frac{n_j}{N^2})}\right)^2 \tag{4.25}$$

and $\mathcal{L}_{\mathbf{s}} = \mathcal{L}_{\mathbf{s},\boldsymbol{a},x,N,\boldsymbol{\alpha}}$ is one of the 4 linear forms of 3 variables:

$$\mathcal{L}_{\mathbf{s}}(\mathbf{n}) = \mathcal{L}_{\mathbf{s}}(n_1,n_2,n_3) = n_1(\delta_1 x_1 - a_1) + n_2(\delta_2 x_2 - a_2) - \delta_3 N(\mathbf{n}\boldsymbol{\alpha} - n_3) \tag{4.26}$$

where $\mathbf{s} = (\delta_1,\delta_2,\delta_3) \in \{0,1\}^3$, and $\delta_1 + \delta_2 = 1$.

Note from (4.15) that the sign $\pm$ in (4.24) is in fact $\pm = (-1)^{\delta_2+\delta_3+1}$, and so it is

independent of $\mathbf{n} \in \mathbb{Z}^3$.

The main idea is to divide the divisors into small ranges and cancel out the positives with the negatives, with the aim of getting rid of the extra $\ln N$ in the usual estmations by the Erdős–Turán-Koksma inequality.

Let $\delta_N = (\ln N)^{-2}$, and for every $\epsilon = (\epsilon_1, \epsilon_2, \epsilon_3) \in \{-1, +1\}^3$ and $\mathbf{l} = (l_1, l_2, l_3) \in \mathbb{N}^3$ with

$$\begin{aligned} &(\ln N)^{40} \leq (1+\delta_N)^{l_j} \leq \frac{N^2}{4}, \ 1 \leq j \leq 2, \text{ and} \\ &(\ln N)^{40} \leq (1+\delta_N)^{l_3} \leq \left(\frac{N^2}{4}\right)^2 \end{aligned} \tag{4.27}$$

and write

$$U(\mathbf{l}, \epsilon) = \left\{ \mathbf{n} \in \mathbb{Z}^2 \left| \begin{array}{l} (1+\delta_N)^{l_1} \leq \epsilon_1 n_1 < (1+\delta_N)^{l_1+1}, \\ (1+\delta_N)^{l_2} \leq \epsilon_2 (n_1\tau - n_2) < (1+\delta_N)^{l_2+1}, \\ (1+\delta_N)^{l_2} \leq \epsilon_3 n_1 (n_1\tau - n_2)(\mathbf{n}\boldsymbol{\alpha} - n_3) < (1+\delta_N)^{l_3+1} \end{array} \right. \right\} \tag{4.28}$$

where $\mathbf{n} = (n_1, n_2)$, and n_3 is the nearest integer to $\mathbf{n}\boldsymbol{\alpha} = n_1\alpha_1 + n_2\alpha_2$. Note that by (4.27) the range for $\mathbf{l}$ is of order $(\ln N/\delta_N)^3$.

Let ϵ^+ and ϵ^- be two vectors in $\{-1, +1\}^3$ such that the corresponding first 2 coordinates are equal: $\epsilon_j^+ = \epsilon_j^-$, $1 \leq j \leq k$, but the last coordinates are different: $\epsilon_3^+ = 1$, $\epsilon_3 = -1$. Since in Section 4 we have done all the estimations by bounding the absolute values of the terms, we can assume without loss of generality that the remaining terms in $\bar{D}_5$ form a perfect union of boxes $U(\mathbf{l}, \epsilon^+) \cup U(\mathbf{l}, \epsilon^-)$ by attaching additional border points. Therefore we have

$$\begin{aligned} &\bar{D}_5(\mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N) \\ &= \sum_s \sum_{(\mathbf{l}, \epsilon^\pm)} \left(\sum_{\mathbf{n} \in U(\mathbf{l}, \epsilon^+)} f_2(\mathbf{n}, \mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N) + \sum_{\mathbf{n} \in U(\mathbf{l}, \epsilon^-)} f_2(\mathbf{n}, \mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N) \right) \end{aligned} \tag{4.29}$$

By Lemma 3.2, for both $\epsilon^\pm$ we have precisely the same estimation:

Lemma 4.6. *For both ϵ^+ and ϵ^-, we have*

$$|U(\mathbf{l}, \epsilon^{+or-})| = E(\mathbf{l}) + Err, \tag{4.30}$$

where $E(\mathbf{l})$ is the main term, and Err is the Error term and specifically:

$$\begin{aligned} &E(\mathbf{l}) \ll \delta_N^3 (1+\delta_N)^{l_3} \\ &Err \ll \delta_N^4 (1+\delta_N)^{l_3} \end{aligned} \tag{4.31}$$

Proof. In Lemma 3.2, by taking $C_1 = (1+\delta_N)^{l_3+1}$, $C_2 = (1+\delta_N)^{l_3}$, and V, W, with

$$v_j = (1+\delta_N)^{l_j},\ w_j = (1+\delta_N)^{l_j+1},\ 1 \le j \le 2,$$

we have that the main term

$$\begin{aligned} E(\mathbf{l}) &= \sum_{\mathbf{n}\in U(\mathbf{l},\epsilon)} \left(\min\{\frac{1}{2}, \frac{C_1}{|n_1||n_1\tau - n_2|}\} - \min\{\frac{1}{2}, \frac{C_2}{|n_1||n_1\tau - n_2|}\} \right) \\ &\ll (C_1 - C_2)\delta_N^2 \ll (1+\delta_N)^{l_3}\delta_N^3 \end{aligned}$$

and the error term

$$Err \ll C_1^{\frac{3}{4}+\epsilon} \cdot (\ln N)^{1+\epsilon} \ll \delta_N^{2d}(1+\delta_N)^{l_3},$$

since $C_1 > (\ln N)^{40}$. □

We adopt the same definition as Beck in [Bec94] for the ϵ-big vectors, ϵ-neighbors and ϵ-lines, although here the "neighbors" and "lines" do not depend on ϵ.

Definition 4.1. *We say that* $\mathbf{l} = (l_1, l_2, l_3) \in \mathbb{N}^3$ *satisfying* (4.27) *is an* ϵ***-big*** *vector if*

$$\frac{|U(\mathbf{l},\epsilon^+)| + |U(\mathbf{l},\epsilon^-)|}{\ln N} \le \left| \sum_{\mathbf{n}\in U(\mathbf{l},\epsilon^+)} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{n})} - \sum_{\mathbf{n}\in U(\mathbf{l},\epsilon^-)} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{n})} \right|, \tag{4.32}$$

where $|U| = \#U$ *denotes the cardinality of the set* U. *Let* $U(\mathbf{l},\epsilon^\pm) = U(\mathbf{l},\epsilon^+) \cup U(\mathbf{l},\epsilon^-)$

Definition 4.2. *Two integral vecotrs* $\mathbf{l} = (l_1, l_2, l_3)$ *and* $\mathbf{h} = (h_1, h_2, h_3)$ *satisfying* (4.27) *are called* ***neighbors*** *if*

$$(1+\delta_N)^{h_j - l_j} = (\ln N)^9,\ 1 \le j \le 2,\ \textit{and} \tag{4.33}$$

$$(1+\delta_N)^{h_3 - l_3} = (\ln N)^{27}, \tag{4.34}$$

The notation $\mathbf{l}\to\mathbf{h}$ means that the ordered pair $\langle \mathbf{l}, \mathbf{h}\rangle$ of vectors satisfies (4.33) and (4.34).

Note that by slightly modifying the value of $\delta_N \approx (\ln N)^{-2}$, we can make sure that the above definitions are met for integer vectors $\mathbf{l}$ and $\mathbf{h}$.

Definition 4.3. *A sequence* $H = \langle \mathbf{h}^{(1)}, \mathbf{h}^{(2)}, \mathbf{h}^{(3)}, \ldots\rangle$ *of vectors satisfying* (4.27) *is called a "special line" if* $\mathbf{h}^{(1)}\to\mathbf{h}^{(2)}\to\mathbf{h}^{(3)}\to\ldots$, *that is, any two consecutive vectors in* H *are neighbors.*

Lemma 4.7. *For almost every* $\boldsymbol{\alpha}$, *almost every* $\tau > 0$, *every special line contains at most one* $\boldsymbol{\epsilon}^\pm$*-big vector.*

Proof. Let $H = \langle \mathbf{h}^{(1)}, \mathbf{h}^{(2)}, \mathbf{h}^{(3)}, \ldots \rangle$ be a special line with two $\boldsymbol{\epsilon}^{\pm}$-big vectors $\mathbf{h}^{(p)}$ and $\mathbf{h}^{(q)}$, $1 \le p < q$. If

$$\|\mathcal{L}(\mathbf{n})\| \le (\ln N)^{-2} \text{ for every } \mathbf{n} \in U(\mathbf{h}^{(p)}, \epsilon^{\pm}), \tag{4.35}$$

then

$$|1 - e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{n})}| \ll (\ln N)^{-2} \text{ for every } \mathbf{n} \in U(\mathbf{h}^{(p)}, \epsilon^{\pm}). \tag{4.36}$$

By repeating the argument of the cancellation of the main term with the above equation, we obtain (Err means error for the number of elements in S)

$$\begin{aligned}
&\left| \sum_{\mathbf{n} \in U(\mathbf{h}^{(p)}, \epsilon^{+})} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{n})} - \sum_{\mathbf{n} \in U(\mathbf{h}^{(p)}, \epsilon^{-})} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{n})} \right| \\
&\ll (1 + (\ln N)^{-2})(E(\mathbf{h}^{(p)})) + |\mathrm{Err}|) - (1 - (\ln N)^{-2})(E(\mathbf{h}^{(p)})) - |\mathrm{Err}|) \\
&\ll |\mathrm{Err}| + (\ln N)^{-2} E(\mathbf{h}^{(p)})) \\
&\ll \delta_N (E(\mathbf{h}^{(p)})) \\
&\ll \frac{|U(\mathbf{h}^{(p)}, \epsilon^{\pm})|}{(\ln N)^2}.
\end{aligned}$$

But this contradicts the assumption that $\mathbf{h}^{(p)}$ is $\boldsymbol{\epsilon}^{\pm}$-big, see (4.32). So there is an $\mathbf{n}^* \in U(\mathbf{h}^{(p)}, \boldsymbol{\epsilon}^{\pm})$ such that

$$\|\mathcal{L}(\mathbf{n})\| > (\ln N)^{-2}. \tag{4.37}$$

For every $\mathbf{m} \in U(\mathbf{h}^{(q)}, \epsilon^{\pm})$ (another $\boldsymbol{\epsilon}^{\pm}$-big vector), consider the "arithmetic progression" with difference $\mathbf{n}^*$:

$$\mathbf{m} + r \cdot \mathbf{n}^* = (m_1 + r \cdot \mathbf{n}^*{}_1, m_2 + r \cdot \mathbf{n}^*{}_2, m_3 + r \cdot \mathbf{n}^*{}_3), \quad r = 0, \pm 1, \pm 2, \ldots$$

We will estimate how many consecutive members $\mathbf{m} + r \cdot \mathbf{n}^*$ are contained in $U(\mathbf{h}^q, \epsilon^{\pm})$. Since $n^* \in U(\mathbf{h}^p, \epsilon^{\pm})$, the definition for $U(\mathbf{h}^q, \epsilon^{\pm})$ (see (4.28)) gives the following:

$$(1 + \delta_N)^{h_1^{(p)}} \le \epsilon_1 n_1^* < (1 + \delta_N)^{h_1^{(p)}+1}, \tag{4.38}$$

$$(1 + \delta_N)^{h_2^{(p)}} \le \epsilon_2 (n_1^* \tau - n_2^*) < (1 + \delta_N)^{h_2^{(p)}+1}, \tag{4.39}$$

$$(1 + \delta_N)^{h_3^{(p)}} \le |n_1^*| |n_1^* \tau - n_2^*| \|\mathbf{n}^* \boldsymbol{\alpha}\| < (1 + \delta_N)^{h_3^{(p)}+1}. \tag{4.40}$$

Definition 4.4. *An* $\mathbf{m} \in U(\mathbf{h}^q, \epsilon^{\pm})$ *is called an inner point if*

$$(1 + \delta_N)^{h_1^{(q)}} \left(1 + \frac{\delta_N}{(\ln N)^2}\right) \le \epsilon_1 m_1 < \left(1 - \frac{\delta_N}{(\ln N)^2}\right) (1 + \delta_N)^{h_1^{(q)}+1}, \tag{4.41}$$

$$(1 + \delta_N)^{h_2^{(q)}} \left(1 + \frac{\delta_N}{(\ln N)^2}\right) \le \epsilon_2 (m_1 \tau - m_2) < \left(1 - \frac{\delta_N}{(\ln N)^2}\right) (1 + \delta_N)^{h_2^{(q)}+1}, \tag{4.42}$$

$$(1+\delta_N)^{h_3^{(q)}}\left(1+\frac{\delta_N}{(\ln N)^2}\right) \le |m_1||m_1\tau - m_2|\cdot\|\mathbf{m}\boldsymbol{\alpha}\| < \left(1-\frac{\delta_N}{(\ln N)^2}\right)(1+\delta_N)^{h_3^{(q)}+1}. \tag{4.43}$$

The rest of the points in $U(\mathbf{h}^q,\epsilon^{\pm})$ are called border points.

For every inner point $\mathbf{m}\in U(\mathbf{h}^q,\epsilon^{\pm})$, and for every $|r|\le(\ln N)^4$, it follows from (4.33), (4.38) and (4.41) that,

$$\begin{aligned}(1+\delta_N)^{h_1^{(q)}} &< (1+\delta_N)^{h_1^{(q)}}\left(1+\frac{\delta_N}{(\ln N)^2}\right) - (\ln N)^4(1+\delta_N)^{h_1^{(p)}+1}\\ &\le \epsilon_1(m_1+r\cdot n_1^*)\\ &< \left(1-\frac{\delta_N}{(\ln N)^2}\right)(1+\delta_N)^{h_1^{(q)}+1} + (\ln N)^4(1+\delta_N)^{h_1^{(p)}+1}\\ &< (1+\delta_N)^{h_1^{(q)}+1}.\end{aligned} \tag{4.44}$$

Similarly, from (4.33), (4.39) and (4.42), we obtain the following,

$$(1+\delta_N)^{-h_2^{(q)}} < \epsilon_2((m_1+r\cdot n_1^*)\tau-(m_2+rn_2^*)) < (1+\delta_N)^{-h_2^{(q)}+1}, \tag{4.45}$$

and from (4.34), (4.40) and (4.43) we have:

$$\begin{aligned}&\frac{(1+\delta_N)^{h_3^{(q)}}}{|m_1+r\cdot n_1^*||(m_1+r\cdot n_1^*)\tau-(m_2+r\cdot n_2^*)|}\\ &< \|\mathbf{m}\boldsymbol{\alpha}\| - |r|\cdot\|\mathbf{n}^*\boldsymbol{\alpha}\|\\ &\le \|(\mathbf{m}+r\cdot\mathbf{n}^*)\boldsymbol{\alpha}\|\\ &\le \|\mathbf{m}\boldsymbol{\alpha}\| + |r|\cdot\|\mathbf{n}^*\boldsymbol{\alpha}\|\\ &< \frac{(1+\delta_N)^{h_3^{(q)}+1}}{|m_1+r\cdot n_1^*||(m_1+r\cdot n_1^*)\tau-(m_2+r\cdot n_2^*)|},\end{aligned} \tag{4.46}$$

where we use the In view of (4.44)-(4.46), for any inner point $\mathbf{m}\in U(\mathbf{h}^q,\epsilon^{\pm})$, at least $(\ln N)^4$ consecutive members in the progression $\mathbf{m}+r\cdot\mathbf{n}^*$ are contained in $U(\mathbf{h}^q,\epsilon^{\pm})$. Therefore, we can decommpose $U(\mathbf{h}^q,\epsilon^{\pm})$ into three parts:

$$U(\mathbf{h}^q,\epsilon^{\pm}) = AP^+\cup AP^-\cup BP \tag{4.47}$$

where $AP^{\pm}$ denotes the family of arithmetic progreesions $\{\mathbf{m}+r\cdot\mathbf{n}^* : 0\le r\le l-1\}$ in $U(\mathbf{h}^q,\epsilon^+)$ and $U(\mathbf{h}^q,\epsilon^-)$ respectively, where $l=l(\mathbf{m})$ is the length of the progression starting from $\mathbf{m}$, and $l\ge(\ln N)^4$. BP denotes a set of border points of $U(\mathbf{h}^q,\epsilon^{\pm})$ that are not included in any arithmetic progressions. Using $\|\mathcal{L}(\mathbf{n})\| > (\ln N)^{-2}$ (see (4.37)), the

linearity of $\mathcal{L}$, we obtain

$$\begin{aligned}
|\sum_{\mathbf{AP}^+} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{n})}| &\leq \sum_{\text{arithmetic progressions}} |\sum_{r=0}^{l-1} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{m}+r\mathbf{n}^*)}| \\
&= \sum_{\text{arithmetic progressions}} |\sum_{r=0}^{l-1} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{m})+r\mathcal{L}(\mathbf{n}^*)}| \\
&\ll \sum_{\text{arithmetic progressions}} \frac{1}{\|\mathcal{L}(\mathbf{n})\|} < \sum_{\text{arithmetic progressions}} (\ln N)^2 \\
&\leq \sum_{\text{arithmetic progressions}} \frac{\text{length}}{(\ln N)^2} \leq \frac{|U(\mathbf{h}^q, \epsilon^+)|}{(\ln N)^2}
\end{aligned} \tag{4.48}$$

since each length $\geq (\ln N)^4$. Similarly,

$$|\sum_{\mathbf{AP}^-} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{n})}| \ll \frac{|U(\mathbf{h}^q, \epsilon^-)|}{(\ln N)^2} \tag{4.49}$$

Finally, for border points, at least one of the inequalities in definition 4.4 is violated, thus the range for at least one components is shrunk by the ratio $\frac{\delta_N}{(\ln N)^2}$. Using the same reasoning as in Lemma 4.6, and the fact that the cardinality of the set BP can be controlled by the total number of border points of $U(\mathbf{h}^q, \epsilon^{\pm})$, we have for almost every $\boldsymbol{\alpha}, \tau > 0$,

$$|BP| \ll \frac{|U(\mathbf{h}^q, \epsilon^{\pm})|}{(\ln N)^2} \tag{4.50}$$

Combining (4.47)-(4.50), for almost every α, β, we obtain

$$\left| \sum_{\mathbf{n} \in U(\mathbf{h}^{(p)}, \boldsymbol{\epsilon}^+)} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{n})} - \sum_{\mathbf{n} \in U(\mathbf{h}^{(p)}, \boldsymbol{\epsilon}^-)} e^{2\pi \mathbf{i}\mathcal{L}(\mathbf{n})} \right| \ll \frac{|U(\mathbf{h}^{(p)}, \boldsymbol{\epsilon}^{\pm})|}{(\ln N)^2}$$

which contradicts the assumption that $h^{(q)}$ is $\boldsymbol{\epsilon}^{\pm}$-big (see (4.32)), therefore for almost very $\boldsymbol{\alpha}, \tau > 0$, every special line contains at most one $\boldsymbol{\epsilon}^{\pm}$-big vector, which proves the lemma. □

Corollary 4.3. *The number of $\boldsymbol{\epsilon}^{\pm}$-big vectors is $\ll \delta_N^{-3} \cdot (\ln N)^2 (\ln \ln N)$.*

Proof. First we estimate the number of maximal special lines. Let $H = \langle \mathbf{h}^{(1)}, \mathbf{h}^{(2)}, \mathbf{h}^{(3)}, \ldots \rangle$ be a "special line", here $\mathbf{h}^{(1)}$ is the first element of H, that is, if $\mathbf{h}^{(0)} \to \mathbf{h}^{(1)}$ holds for some $\mathbf{h}^{(0)}$, then at least one of the following inequalities is violated(see (4.27):

$$(1+\delta_N)^{\epsilon_j h_j^{(0)}} \geq (\ln N)^{40}, \ 1 \leq j \leq 2$$

$$(1+\delta_N)^{h_3^{(0)}} \geq (\ln N)^{40}.$$

Thus, by (4.33) and (4.34), one of the inequalities below holds:

$$(\ln N)^{40} \leq (1+\delta_N)^{\epsilon_j h_1^{(1)}} \leq (\ln N)^{40},\ 1 \leq j \leq 2$$

$$(\ln N)^{40} \leq (1+\delta_N)^{h_3^{(1)}} \leq (\ln N)^{40+27}.$$

So at least one coordinate $h_j^{(1)}$ of the first element $\mathbf{h}^{(1)}$ of H is restricted to an interval of length const$\cdot \ln\ln N \cdot \delta_N^{-1}$, the rest, by the condition (4.27), are restricted to an interval of length const$\cdot \ln N \cdot \delta_N^{-1}$. Since the starting vector determines the whole special line, the number of special line is

$$\ll (\ln\ln N) \cdot (\ln N)^2 \cdot \delta_N^{-3}.$$

By Lemma 4.7, the total number of $\boldsymbol{\epsilon}^{\pm}$-big vectors is also

$$\ll (\ln\ln N) \cdot (\ln N)^2 \cdot \delta_N^{-3}.$$

□

With the help of Lemma 4.7, we can estimate the contribution of the exponential terms, we have the following:

Proposition 4.8. *For almost every $\boldsymbol{\alpha}$ we have*

$$\left|\bar{D}_5\right| \ll (\ln N)^2 (\ln\ln N)$$

Proof.

$$\bar{D}_5 = \sum_{\text{small}} + \sum_{\text{big}}$$

where

$$\sum_{\text{small}} = \sum_{\epsilon} \sum_{\mathbf{l} \text{ not } \epsilon\text{-big}} \sum_{\mathbf{n} \in U(\mathbf{l},\epsilon)} f_2(\mathbf{n}, \mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N) \tag{4.51}$$

$$\sum_{\text{small}} = \sum_{\epsilon} \sum_{\mathbf{l} \text{ is } \epsilon\text{-big}} \sum_{\mathbf{n} \in U(\mathbf{l},\epsilon)} f_2(\mathbf{n}, \mathbf{x}, \boldsymbol{a}, \boldsymbol{\alpha}; N) \tag{4.52}$$

By Lemma 4.6, range for $\mathbf{l}$: (4.27), and (4.32),

$$\begin{aligned}
\left|\sum_{\text{small}}\right| &\ll \sum_{\epsilon} \sum_{\mathbf{l} \text{ is not } \epsilon\text{-big}} (1+\delta_N)^{-l_3} \cdot \frac{E(\mathbf{l}) + Err}{\ln N} \\
&\ll \sum_{\mathbf{l}:(4.27)} \frac{\delta_N^3}{\ln N} \\
&\ll \frac{\delta_N^3}{\ln N} \cdot \frac{(\ln N)^3}{\delta_N^3} \\
&\ll (\ln N)^2
\end{aligned}$$

By Corollary 4.3,

$$
\begin{aligned}
|\sum_{\text{big}}| &\ll \sum_{\mathbf{l} \text{ is } \boldsymbol{\epsilon}^{\pm}\text{-big}} (1+\delta_N)^{-l_3} \cdot (E(\mathbf{l}) + Err) \\
&\ll [(\ln\ln N)(\ln N)^2 \delta_N^{-3}] \cdot [(1+\delta_N)^{-l_3} \cdot (1+\delta_N)^{l_3} \delta_N^3] \\
&\ll (\ln N)^2 \ln\ln N
\end{aligned}
$$

□

Combining Proposition 4.1, 4.2, 4.8 and Lemma 4.1, 4.5, we finally arrive at the convergent part of our main Theorem 4.1. The divergent part is easy:

Proof of the divergent part of Theorem 4.1

One way is to employ the same strategy as in Beck[Bec94], and try to find a "large" Fourier coefficient using Lemma 3.3. But an easier way is to use Beck's result directly, since Beck proved that for $\sum_{n=1}^{\infty} \frac{1}{\varphi(n)} = \infty$, the maximal discrepancy function for boxes will be greater than $(\ln N)^2 \varphi(\ln\ln N)$ infinitely often, by cutting the box diagonally we know that at least the discrepancy function for one of the triangles is greater than $1/2(\ln N)^2 \varphi(\ln\ln N)$, since the maximal descrepancy is defined over all starting points $\boldsymbol{a} \in [0,1]^2$, we can translate the triangle with the large discrepancy to the origin, hence the maximal discrepancy defined as in the theorem 1.2 is also greater than $1/2(\ln N)^2 \varphi(\ln\ln N)$ infinitely often. Note that the coefficient $1/2$ does not change the convergence of the series of $\sum_{n=1}^{\infty} \frac{1}{\varphi(n)}$. The proof for the main theorem 1.2 is completed. □

Chapter 5

Ergodic Examples of Cylindrical Cascades

5.1 Introduction

Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ denote the torus, parameterized by $[0, 1)$. Given an irrational vector $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$, let $T_\alpha : \mathbb{T}^d \to \mathbb{T}^d$ denote the translation by α on $\mathbb{T}^d$, given by: $x \mapsto x + \alpha \pmod 1$. Given an observable $A(\cdot)$ on $\mathbb{R}^r$, the cylindrical cascade above T_α relative to $A(\cdot)$ will be denoted $W_{\alpha,A} : \mathbb{T}^d \times \mathbb{R}^r \to \mathbb{T}^d \times \mathbb{R}^r$. It is given by $W_{\alpha,A}(x, y) = (x + \alpha, y + A(x))$. The dynamics of $W_{\alpha,A}$ is closely related to the Birkhoff sum of A, denoted by $A_N(\cdot)$:

$$A_N(x) = \sum_{n=0}^{N-1} A(T_\alpha^n x).$$

We can see cylindrical cascade $W_{\alpha,A}$ as a random walk on the fiber $\mathbb{R}^r$ driven by the translation T_α on the base $\mathbb{T}^d$, which makes it natural to study its recurrence and ergodicity. For the walk $W_{\alpha,A}$ to be *recurrent*, A has to have zero mean, since otherwise A_N is unbounded by the ergodic theorem. In fact, zero mean is also a sufficient condition if $r = 1$ (see [Atk76]). For $r > 1$, a sufficient condition for recurrence, together with recurrent and transient examples, is given by Chevallier and Conze[CC09].

The ergodicity of the walk $W_{\alpha,A}$ can be established if the sums A_{N_n} are increasingly well distributed over $\mathbb{R}^r$ along some subsequence N_n, while the translation $T_\alpha^{N_n}$ stays close to identity. This idea gives rise to the notion of essential values, introduced by Klaus Schmidt in [Sch77]: $a \in \mathbb{R}^r$ is called an **essential value** of A if for each $B \in \mathbb{T}^d$ of positive measure, for each $\epsilon > 0$ there exists $N \in \mathbb{Z}$ such that

$$\mu(B \cap T_\alpha^{-N} B \cap [\mid A_N(\cdot) - a \mid < \epsilon]) > 0.$$

Denote $E(A)$ the set of all essential values of A. The following lemma characterizes

$E(A)$ and gives a criterion for ergodicity,

Lemma 5.1 ([Aar97, Sch77]).

1. *$E(A)$ is a closed subgroup of $\mathbb{R}^r$.*
2. *$E(A) = \mathbb{R}^r$ iff $W_{\alpha,A}$ is ergodic.*
3. *If A is integer valued and $E(A) = \mathbb{Z}^r$ then $W_{\alpha,A}$ is ergodic in $\mathbb{T}^d \times \mathbb{Z}^r$.*

In this chapter, as a joint work with Fatna Abdedou, we construct ergodic examples of cylindrical cascades $W_{\alpha,A}$ over $\mathbb{T}^2$ for sufficiently Liouville translation vectors α with A being the centered indicator functions of *straight* rectangles with the origin as a corner. Specifically, we prove the following:

Theorem 5.1 (Abdedou-W.). *For almost every straight rectangle in $\mathbb{T}^2$ with a corner at the origin, let A be the centered indicator function of the rectangle, then there exists a G_δ-dense set of $\boldsymbol{\alpha}$, such that the cylindrical cascades $W_{\boldsymbol{\alpha},A}$ is ergodic on $\mathbb{T}^2 \times \mathbb{R}$.*

A more precise statement will be given in Section 5.3, together with a version extended for the case of $\mathbb{T}^2 \times \mathbb{R}^2$, where we achieve ergodicity at the price of the genericity of the side lengths of the second rectangle.

5.2 Arithmetic Notations

Let x be a real number,

1. Denote $[x]$ the integer part of x.
2. Denote $\{x\}$ the signed distance of the x to the closest integer, i.e., $\{x\} = x - n$ where n is the only integer such that $x - n \in [-1/2, 1/2)$.
3. Denote $\|x\| = |\{x\}|$ the distance of x to the closest integer.

5.3 Main results and Lemmata

A slightly more precise statement of Theorem 5.1 is the following:

Theorem 5.2 (Abdedou-W.). *For almost every $L = (l, l') \in (0, 1]^2$, there exists a G_δ-dense set of $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2 - \mathbb{Q}^2$, such that the cylinder cascade $W_{\alpha,A}$ constructed over the zero-mean indicator function $A(\cdot) = \chi_{B_L}(\cdot) - Vol(B_L)$ is ergodic.*

For the case of $\mathbb{T}^2 \times \mathbb{R}^2$, given two pairs of lengths (l_1, l_1'), (l_2, l_2'), define the corresponding rectangles $B_1 = [0, l_1] \times [0, l_1']$ and $B_2 = [0, l_2] \times [0, l_2']$ and the zero-means indicator functions

$$\begin{cases} A^{(1)}(\cdot) = \chi_{B_1}(\cdot) - \mathrm{vol}(B_1) \\ A^{(2)}(\cdot) = \chi_{B_2}(\cdot) - \mathrm{vol}(B_2) \end{cases} \tag{5.1}$$

Then we have the following result for the $\mathbb{R}^2$ case.

Theorem 5.3 (Abdedou-W.). *For almost every $l_1, l_1' \in (0,1]$, there exist l_2, l_2', both belongs to a G_δ-dense set of (0,1], and a G_δ-dense set of $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2 - \mathbb{Q}^2$, such that the cylindrical cascade $W_{\alpha,A}$ constructed over $A(\cdot) = (A^1(\cdot), A^2(\cdot))$(defined in (5.1)) is ergodic.*

It is not hard to see that by restricting the discussion in Theorem 5.3 to the first coordinate of $A(\cdot)$, we obtain Theorem 5.2 as a corollary. In the next section, we will prove Theorem 5.3 in detail and keep Theorem 5.2 as a corollary.

Our construction is based on the approximation of the rectangle by small rational rectangles that tile the torus and the approximation of the translation vector by their vertices. We start with the following observation:

Lemma 5.2. *Let v denote the rational vector $(p/q, p'/q')$ on $\mathbb{T}^2$, where the coordinates are in reduced forms, i.e., $(p,q) = 1$ and $(p',q') = 1$. If in addition, the two denominators are relatively prime, that is, $(q,q') = 1$, then we have an identity in $\mathbb{T}^2$ between the orbit of translations by v and the vertices of the rectangles of side lengths $1/q$ and $1/q'$:*

$$\{mv \mid 0 \le m \le qq' - 1\} = \left\{ \left(\frac{i}{q}, \frac{j}{q'}\right) \;\middle|\; 0 \le i \le q-1,\ 0 \le j \le q'-1 \right\}. \tag{5.2}$$

Proof. Because the cardinalities of the two sets are both qq', it is enough to show that for $0 \le m, n \le qq' - 1$,

$$mv = nv \text{ in } \mathbb{T}^2 \Rightarrow m = n. \tag{5.3}$$

Note that $mv = nv$ in $\mathbb{T}^2$ is equivalent to

$$(m-n)p/q \in \mathbb{Z}, \text{ and } (m-n)p'/q' \in \mathbb{Z}. \tag{5.4}$$

Since p and q are relatively prime, q divides $m-n$, similarly q' divides $m-n$, but q and q' are also relatively prime, so we have that qq' divides $m-n$, which is only possible when $m = n$, as $|m-n| \le qq' - 1$. □

Divide the torus $\mathbb{T}^2$ by disjoint rectangles $R^{i,j}$ of the form $[i/q, (i+1)/q) \times [j/q', (j+1)/q')$, where $0 \le i \le q-1$, $0 \le j \le q'-1$. If the translation vector α is close to the rational vector $(p/q, p'/q')$, the above lemma implies that starting from the origin, within a time range of qq', the orbit of the translation visit a close neighborhood of each and every vertice of $R^{i,j}$ exactly once. By approximating the rectangle $B_L = [0,l] \times [0,l']$ by a union of $R^{i,j}$, we can compute the Birkhoff sums for $A = \chi_{B_L} - \mathrm{Vol}(B_L)$ as follows.

Let

$$l = (b+\delta)/q,\ l' = (b'+\delta')/q', \tag{5.5}$$

where $\delta = \{ql\}$, $\delta' = \{q'l'\}$, and let

$$\alpha_1 = (p+\eta)/q,\ \alpha_2 = (p'+\eta')/q', \tag{5.6}$$

where $\eta = \{q\alpha_1\}$, $\eta' = \{q'\alpha_2\}$.

Lemma 5.3. *Denote $N = qq'$, if there exists a $K > 0$ such that*

$$\eta < \delta/(Kq'), \quad \eta' < \delta'/(Kq).$$

then there exists a set $\mathcal{F} \subset \mathbb{T}^2$ of measure $(1-4q\delta)(1-4q'\delta')$, such that for every $x \in \mathcal{F}$, every k between 0 and $K-1$, the Birkhoff sum $A_{kN}(x)$ has the following form:

$$A_{kN}(x) = kA_N(x) = -k(b\delta' + b'\delta + \delta\delta'). \tag{5.7}$$

Proof. Define $\mathcal{F}$ to be the set of points that are *not close* to the boundaries of any $R^{i,j}$:

$$\mathcal{F} = \bigcup_{\substack{0\le i\le q-1\\ 0\le j\le q'-1}} \left\{ (x_1,x_2)\in\mathbb{T}^2 \,\middle|\, \begin{array}{l} i/q+2\delta \le x_1 \le (i+1)/q - 2\delta, \\ j/q'+2\delta' \le x_2 \le (j+1)/q' - 2\delta' \end{array} \right\}. \tag{5.8}$$

$\mathcal{F}$ is the disjoint union of qq' rectangles, each is of measure $(1/q-4\delta)(1/q-4\delta')$, so the measure of $\mathcal{F}$ is $(1-4q\delta)(1-4q'\delta')$.

Note that for $0 \le m \le KN$, so $m\alpha$ is $\delta-$close to $m(p/q, p'/q')$:

$$\begin{aligned} |m\alpha_1 - mp/q| \le & m\eta/q \le \delta, \\ |m\alpha_2 - mp'/q'| \le & m\eta'/q' \le \delta'. \end{aligned} \tag{5.9}$$

Since the points in $\mathcal{F}$ are 2δ-away from the boundaries of $R^{i,j}$, the orbit $\{x+m\alpha \mid 0 \le m \le kN-1\}$ stays δ-away from the boundaries of $R^{i,j}$. By Lemma 5.2, for $0 \le k \le K-1$, the orbit $\{x+m\alpha|\ kN \le m \le (k+1)N-1\}$ visits each $R^{i,j}$ exactly once and stays d-away from the boundaries, thus the orbit visits the rectangle B_L bb' times within a time range of N. Therefore, for $0 \le k \le K-1$, we have

$$\begin{aligned} A_{kN}(x) &= kA_N(x) \\ &= k(bb' - qq'll') \\ &= k(bb' - (b+\delta)(b'+\delta')) \\ &= -k(b\delta' + b'\delta + \delta\delta'). \end{aligned} \tag{5.10}$$

□

If we can show that $A_N(x)$ can be increasingly small along a subsequence, then by choosing the translation vector α to be Liouville enough, we can choose K to be large,

and thus the sequence $\{A_{kN}\}_{0\le k\le K-1}$ becomes increasingly well-distributed in $\mathbb{R}$, while keeping the translation T_α^{kN} close to identity. The following lemma states that this is indeed the case for generic $L=(l,l')\in\mathbb{T}^2$.

Lemma 5.4. *For almost every $l,l'\in(0,1]$, there exist two sequences $\{q_n\}_{n\in\mathbb{N}}$ and $\{q'_n\}_{n\in\mathbb{N}}$ such that the following properties hold,*

1. *$q_n, q'_n\to\infty$, as $n\to\infty$,*
2. *q_n and q'_n are relatively prime, i.e. $(q_n,q'_n)=1$, $\forall n\in\mathbb{N}$,*
3. $$\frac{l}{2(\ln q_n)^2}\le\mid b'_n(q_nl-b_n)+b_n(q'_nl'-b'_n)\mid\le\frac{2}{(\ln\ln\ln q_n)^{1/2}},\tag{5.11}$$
   where b_n and b'_n are the closest integers to q_nl and q'_nl'.
4. *q'_n and q_n are approximately of the same magnitude:*
   $$\frac{q_n}{(\ln q_n)^2}\le q'_n\le q_n(\ln\ln\ln q_n)^{\frac12}\tag{5.12}$$

Lemma 5.4 states that the Birkhoff sums for A can indeed be small along a subsequence, while the coprime condition between q_n and q'_n is satisfied. For examples in higher dimensions, we need similar bounds for the second rectangle, but in order to approximate the essential values along the two axes, we need the magnitude of the Birkhoff sums for the second rectangle to alternate between large and small relative to the Birkhoff sums of the first rectangle. This can be achieved by restricting the length of the second rectangle in a G_δ-dense set, as stated in the next lemma.

Lemma 5.5. *For almost every $l_1,l'_1\in(0,1]$, there exist $l_2,l'_2\in(0,1]$ (each belongs to a G_δ-dense set of (0,1]), such that for the two sequences $\{q_n\}_{n\in\mathbb{N}}$, $\{q'_n\}_{n\in\mathbb{N}}$ defined as in Lemma 5.4,*

$$\frac{1}{4(\ln q_{2n})^3}\le\left|q'_{2n}(q_{2n}l_2-r_{2n})+q_{2n}(q'_{2n}l'_2-r'_{2n})\right|\le\frac{3}{2(\ln q_{2n})^3},\tag{5.13}$$

$$\frac{1}{4(\ln\ln\ln q_{2n+1})^{1/4}}\le\left|q'_{2n+1}(q_{2n+1}l_2-r_{2n+1})+q_{2n+1}(q'_{2n+1}l'_2-r'_{2n+1})\right|\le\frac{3}{2(\ln\ln\ln q_{2n+1})^{1/4}},\tag{5.14}$$

where r_n and r'_n are respectively the closest integer to q_nl_2 and $q'_nl'_2$.

Lemma 5.4 is the key to our construction of ergodic examples on $\mathbb{T}^2\times\mathbb{R}$, where the main difficulty is the coprime condition between q_n and q'_n. If we do not require the coprime condition, Lemma 5.4 can be easily obtained by using the classical version of Khintchine's divergence theorem for Diophantine approximation and Dirichlet principle. We dedicate the last section of the chapter for its proof, together with the proof of Lemma 5.5, which is relatively simple.

5.4 Proof of the Theorems using the Lemmata

We first prove the inequality of essential value criterion for the special rectangles $R_n^{i,j}$, which can then naturally generalized to positive measure sets.

Lemma 5.6. *For the irrationals l_1, l_1', l_2, l_2', and the sequences $\{q_n\}$, $\{q_n'\}$ in Lemma 5.5, there exists a G_δ-dense set of $\alpha = (\alpha_1, \alpha_2) \in [0,1)^2$ such that for every box $R_n^{i,j} = [i/q_n, (i+1)/q_n] \times [j/q_n', (j+1)/q_n']$, $0 \le i \le q_n - 1$, $0 \le j \le q_n' - 1$, for all $a \in \mathbb{R}$, and $\epsilon > 0$, there exists $N \in \mathbb{N}^*$, such that*

$$\mu\left(R_n^{i,j} \cap T_\alpha^{-N} R_n^{i,j} \cap [\mid A_N(\cdot) - (a,0) \mid < \epsilon]\right) > \frac{1}{2q_n q_n'}, \tag{5.15}$$

and the same inequality holds if we change $(a,0)$ to $(0,a)$.

Proof. By choosing a subsequence of n, we can assume that $q_{n+1} \ge 2n(\ln q_n)^2 q_n^3$, also note that q_n' and q_n are of the same magnitude(see (5.12)), so it is possible to choose $\alpha = (\alpha_1, \alpha_2)$ Liouville enough such that there exist two sequences $\{p_n\}$ and $\{p_n'\}$ such that $(p_n, q_n) = 1$, $(p_n', q_n') = 1$, and

$$\left|\alpha_1 - \frac{p_n}{q_n}\right| \le \frac{1}{n(\ln q_n)^2 q_n^2 q_n'}, \qquad \left|\alpha_2 - \frac{p_n'}{q_n'}\right| \le \frac{1}{n(\ln q_n)^2 q_n q_n'^2}. \tag{5.16}$$

For any number γ in $\mathbb{T}$ and $\epsilon > 0$, we can choose q_n large enough, such that there exists a certain p_n relatively prime to q_n with the property that $|p_n/q_n - \gamma| < \epsilon$. Since α_1 is sufficiently close to p_n/q_n when q_n is large, the choice of α_1 naturally form a G_δ-dense set of $\mathbb{T}$, similarly, the choice of α_2 also form a G_δ-dense set of $\mathbb{T}$.

For $M_n = \sqrt{n}(\ln q_n)^2 q_n q_n'$, and for every $0 \le m \le M_n$, we have

$$\left|m\alpha_1 - m\frac{p_n}{q_n}\right| = m\left|\alpha_1 - \frac{p_n}{q_n}\right| \le \frac{1}{\sqrt{n} q_n}, \tag{5.17}$$

and

$$\left|m\alpha_2 - m\frac{p_n'}{q_n'}\right| = m\left|\alpha_2 - \frac{p_n'}{q_n'}\right| \le \frac{1}{\sqrt{n} q_n'}. \tag{5.18}$$

For any $i, j \in \mathbb{N}$, denote by $\mathcal{F}_{i,j}(q_n, q_n')$ the set of points (x_1, x_2) satisfying the following

$$\frac{2}{\sqrt{n} q_n} + \frac{i}{q_n} \le x_1 \le \frac{i+1}{q_n} - \frac{2}{\sqrt{n} q_n}, \tag{5.19}$$

and

$$\frac{2}{\sqrt{n} q_n'} + \frac{j}{q_n'} \le x_2 \le \frac{j+1}{q_n'} - \frac{2}{\sqrt{n} q_n'}. \tag{5.20}$$

Then the Haar measure of the set $\mathcal{F}_{i,j}$ has the following lower bound,

$$\mu\{\mathcal{F}_{i,j}(q_n, q'_n)\} \geq \left(1 - \frac{2}{\sqrt{n}}\right)\left(1 - \frac{2}{\sqrt{n}}\right)\frac{1}{q_n q'_n} \tag{5.21}$$

$$\geq \frac{3}{4}\frac{1}{q'_n q_n}, \tag{5.22}$$

for $n \geq 4$.

Note that the condition of $x \in \mathcal{F}_{i,j}(q_n, q'_n)$ means that x is far away from the boundary of the box $R_n^{i,j}$, while the condition for α (see (5.16)) means that $\{m\alpha\}_{1\leq m\leq M_n}$ is close to the boundaries of the rectangles $[i'/q_n, (i'+1)/q_n] \times [j'/q'_n, (j'+1)/q'_n]$, this means that the sequence of points $\{x + m\alpha\}_{1\leq m\leq M_n}$ are not close to the boundaries of the rectangle B_1 and B_2, i.e. the following,

$$\|x_1 + m\alpha_1 - l_i\| \geq \max\left\{\left|l_1 - \frac{b_n}{q_n}\right|, \left|l_2 - \frac{r_n}{q_n}\right|\right\}$$

and

$$\|x_2 + m\alpha_2 - l'_i\| \geq \max\left\{\left|l'_1 - \frac{b'_n}{q'_n}\right|, \left|l'_2 - \frac{r'_n}{q'_n}\right|\right\}$$

where $0 \leq m \leq M_n$.

Therefore, for $x \in \mathcal{F}_{i,j}(q_n, q'_n)$, $N_n = q_n q'_n$, by the coprime conditions of (p_n, q_n) and (p'_n, q'_n), $\{x+m\alpha\}_{1\leq m\leq N_n}$ will visit the center part of all the small rectangles $[i/q_n, (i+1)/q_n]\times [j/q'_n, (j+1)/q'_n]$ exactly once, then the number of visits inside the rectangle B_1 is equal to $b_n b'_n$, then we have

$$\mid A_{N_n}^{(1)}(x) \mid = \mid b_n b'_n - q_n q'_n l_1 l'_1 \mid,$$

Let

$$l_1 = \frac{b_n + \delta_k}{q_n} \quad \text{and} \quad l'_1 = \frac{b'_n + \delta'_k}{q'_n}.$$

Thus, from (5.11)

$$\begin{aligned}
\mid A_{N_n}^{(1)}(x) \mid &= \mid b_n b'_n - q_n q'_n l_1 l'_1 \mid \\
&= \mid b_n b'_n - (b_n + \delta_n)(b'_n + \delta'_n) \mid \\
&\leq 2 \mid b_n \delta'_n + b_n \delta'_n \mid \\
&\leq 2\text{RHS of (5.11)} \\
&= \frac{4}{(\ln\ln\ln q_n)^{1/2}},
\end{aligned}$$

and similarly,

$$A_{N_n}^{(1)}(x) \geq \frac{l_1}{4(\ln q_n)^2}.$$

Therefore, for $x \in \mathcal{F}_{i,j}(q_n, q'_n)$, we have

$$\frac{l_1}{4(\ln q_n)^2} \leq \left|A^{(1)}_{N_n}(x)\right| \leq \frac{4}{(\ln\ln\ln q_n)^{1/2}}. \tag{5.23}$$

Similarly, from (5.13) and (5.14) we have

$$\left|A^{(2)}_{N_{2n}}(x)\right| \leq \frac{3}{(\ln q_n)^3}, \tag{5.24}$$

$$\frac{1}{4(\ln\ln\ln q_{2n+1})^{1/4}} \leq \left|A^{(2)}_{N_{2k+1}}(x)\right| \leq \frac{3}{(\ln\ln\ln q_{2n+1})^{1/4}}, \tag{5.25}$$

For $a \in \mathbb{R}$ and $\epsilon > 0$, by choosing n large enough, there exists $K_{2n} = K_{2n}(a)$ such that $|K_{2n}| \leq \sqrt{2n}(\ln q_{2n})^2$ and the following inequalities hold:

$$\left|K_{2n}A^{(1)}_{N_{2n}}(x) - a\right| \leq \frac{4}{(\ln\ln\ln q_{2n})^{1/2}} \leq \frac{1}{\sqrt{2}}\epsilon, \tag{5.26}$$

$$\left|K_{2n}A^{(2)}_{N_{2n}}(x)\right| \leq \frac{3\sqrt{2n}(\ln q_{2n})^2}{(\ln q_{2n})^3} \leq \frac{3\sqrt{2n}}{\ln q_{2n}} \leq \frac{3\sqrt{2n}}{2n} \leq \frac{1}{\sqrt{2}}\epsilon, \tag{5.27}$$

where we use the fact that $\ln q_n \geq n$, which follows from the choice of $q_{n+1} \gg q_n$. Therefore, by Lemma 5.3, for $N = K_{2n}N_{2n} = K_{2n}q_{2n}q'_{2n}$, $x \in \mathcal{F}_{i,j}(q_{2n}, q'_{2n}) \subset R^{i,j}_{2n}$, we have

$$A_N(x) = K_{2n}(A^{(1)}_{N_{2n}}(x), A^{(2)}_{N_{2n}}(x)),$$

which implies that

$$|A_N(x) - (a, 0)| \leq \epsilon. \tag{5.28}$$

So we have

$$\begin{aligned} &\mathcal{F}_{i,j}(q_{2n}, q'_{2n}) \cap T_\alpha^{-N}(\mathcal{F}_{i,j}(q_{2n}, q'_{2n}) \\ &\subset T_\alpha^{-N}(R^{i,j}_{2n}) \cap R^{i,j}_{2n} \cap \left\{x \in [0,1)^2, |A_N(x) - (a,0)|\} \leq \epsilon\right\} \end{aligned} \tag{5.29}$$

Note that by inequality (5.17) and (5.18)

$$\begin{aligned} \mu\left(\mathcal{F}_{i,j}\left(q_{2n}, q'_{2n}\right) \cap T_\alpha^{-N}\left(\mathcal{F}_{i,j}(q_{2n}, q'_{2n})\right)\right) &\geq \left(1 - \frac{1}{\sqrt{2n}}\right)\left(1 - \frac{1}{\sqrt{2n}}\right)\mu\left(\mathcal{F}_{i,j}(q_{2n}, q'_{2n})\right), \\ &\geq \frac{1}{2q_{2n}q'_{2n}} > 0. \end{aligned}$$

Similarly, we can obtain the positive measure when we substitute $2n$ with $2n+1$, and $(a, 0)$ with $(0, a)$. This finishes the proof of Lemma 5.6. □

Corollary 5.4. *For almost every $l_1, l'_1 \in (0, 1]$, a G_δ-dense set of $l_2, l'_2 \in (0, 1]$, and a G_δ-dense set of $\alpha = (\alpha_1, \alpha_2) \in [0, 1)^2$, we have $E(A) = \mathbb{R}^2$.*

Proof. By Theorem 5.1, it suffices to prove that a similar conclusion in Lemma 5.6 holds

after we substitute $R_n^{i,j}$ by any set B of positive measure. In fact, for any positive measure set B, there exists n large enough, such that

$$\mu\left(R_n^{i,j} \cap B\right) > \frac{9}{10}\frac{1}{q_n' q_n}, \tag{5.30}$$

for some $0 \le i \le q_n - 1$ and $0 \le j \le q_n' - 1$, thus we can expect a slightly weaker lower bound for a general positive measure set B.

For given $a \in \mathbb{R}$, $\epsilon > 0$, choose n large enough such that inequality(5.30) holds, denote $N = N(a, \epsilon) = K_{2n} N_{2n} = K_{2n} q_{2n}' q_{2n}$ as in Lemma 5.6. By inequality(5.28), we have

$$\mathcal{F}_{i,j}(q_{2n}, q_{2n}') \subset \{x \mid |A_N(x) - (a, 0)| < \epsilon\}. \tag{5.31}$$

Now with standard measure computations, we are ready to show the following desired bound,

$$\mu\left(\mathcal{F}_{i,j}(q_{2n}, q_{2n}') \cap T_\alpha^{-N}\left(B \cap R_{2n}^{i,j}\right) \cap \left(B \cap R_{2n}^{i,j}\right)\right) > 0,$$

Let $\mathcal{X} = T_\alpha^{-N}\left(B \cap R_{2n}^{i,j}\right) \cap \left(B \cap R_{2n}^{i,j}\right)$, then we have the following bound for $\mu(\mathcal{X})$:

$$\begin{aligned}
\mu(\mathcal{X}) &= \mu\left(T_\alpha^{-N}\left(B \cap R_{2n}^{i,j}\right) \cap \left(B \cap R_{2n}^{i,j}\right)\right) \\
&= \mu\left(T_\alpha^{-N}\left(B \cap R_{2n}^{i,j}\right)\right) + \mu\left(B \cap R_{2n}^{i,j}\right) - \mu\left(T_\alpha^{-N}\left(B \cap R_{2n}^{i,j}\right) \cup \left(B \cap R_{2n}^{i,j}\right)\right) \\
&\ge 2\mu\left(B \cap R_{2n}^{i,j}\right) - \mu\left(T_\alpha^{-N}\left(R^{i,j}\right) \cup R_{2n}^{i,j}\right) \\
&\ge \frac{18}{10}\mu\left(R_{2n}^{i,j}\right) - \frac{5}{4}\mu\left(R_{2n}^{i,j}\right) \\
&= \frac{1}{2}\mu\left(R_{2n}^{i,j}\right),
\end{aligned}$$

where the inequality $\mu\left(T_\alpha^{-N}(R^{i,j}) \cup R_{2n}^{i,j}\right) \le \frac{5}{4}\mu\left(R_{2n}^{i,j}\right)$ comes from the fact that T_α^{-N} is close to identity, as $q_{2n} q_{2n}'$ divides N(see (5.16)).

Denote $\mathcal{Y} = \mathcal{F}_{i,j}(q_{2n}, q_{2n}')$. Note that $\mathcal{X}$ and $\mathcal{Y}$ are both subsets of $R_{2n}^{i,j}$, by using the lower bound (5.21) of $\mu(\mathcal{Y})$, we have

$$\begin{aligned}
\mu(\mathcal{X} \cap \mathcal{Y}) &= \mu(\mathcal{X}) + \mu(\mathcal{Y}) - \mu(\mathcal{X} \cup \mathcal{Y}) \\
&\ge \frac{1}{2}\mu\left(R_{2n}^{i,j}\right) + \frac{3}{4}\mu\left(R_{2n}^{i,j}\right) - \mu\left(R_{2n}^{i,j}\right) \\
&\ge \frac{1}{4}\mu\left(R_{2n}^{i,j}\right).
\end{aligned}$$

Finally we obtain the following bound,

$$\mu\left(T_\alpha^{-N}\left(B \cap R_{2n}^{i,j}\right) \cap \left(B \cap R_{2n}^{i,j}\right) \cap \mathcal{F}_{i,j}(q_{2n}, q_{2n}')\right) > \frac{1}{4}\mu\left(R_{2n}^{i,j}\right) > 0.$$

This proves that for every $a \in \mathbb{R}$, $(a, 0) \in E(A)$, similarly we can prove that $(0, a)$ belongs to

$E(A)$, by the subgroup property of essential value in Theorem 5.1, we have $E(A) = \mathbb{R}^2$. $\square$

Corollary 5.5. *By restricting the above discussions to the first coordinate of $A^{(1)}$, it is easy to see that in the case of $\mathbb{T}^2 \times \mathbb{R}$, for almost every $l_1, l_1' \in (0,1]$, there exists a G_δ-dense set of $\alpha = (\alpha_1, \alpha_2) \in [0,1)^2$, we have $E(A^{(1)}) = \mathbb{R}$.*

5.5 Proofs of Lemmata 5.4 and 5.5

We can see that the *coprime* condition is essential for Lemma 5.2, and to achieve it, we need a stronger version of Khintchine's divergence theorem, one that allows us to impose prime conditions for the denominators.

Lemma 5.7 ([DS41]). *If there exists a function $\psi : \mathbb{N} \to \mathbb{R}_+$ such that*

(i) $\sum \psi(q) = \infty$,

(ii) there exists a strictly positive constant c, such that the following inequality

$$\sum_{q=1}^{n} \frac{\psi(q)\varphi(q)}{q} > c \sum_{q=1}^{n} \psi(q),$$

holds for infinitely many n, where $\varphi(q)$ is the Euler's totient function, then for almost every $l \in \mathbb{R}$, there exist infinitely many pairs of integers (b, q) such that

$$\|ql\| = |ql - b| < \psi(q).$$

With the help of Lemma 5.7, we could restrict the approximation of $l \in \mathbb{R}$ by rationals with *prime* denominators, as shown in the next lemma.

Corollary 5.6. *For almost every $l \in \mathbb{R}$, there exist infinitely many pairs of integers (b, q) such that q is* prime *and*

$$\|ql\| = |ql - p| < \frac{1}{q \ln\ln\ln q} \tag{5.32}$$

Proof of Corollary 5.6. Define

$$\psi(q) = \begin{cases} 1/(q \ln\ln\ln q) & \text{if } q \text{ is prime and } \ln\ln\ln q \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

It suffices to check that condition (i) and (ii) of Lemma 5.7 are satisfied for $\psi(q)$.

Note that if q is prime, we have $\varphi(q) = q - 1$ (φ is the Euler's totient function), then $\varphi(q)/q \geq 1/2$, and condition (ii) of Lemma 5.7 is satisfied.

For condition (i), by the following inequality of the sum of the reciprocals of primes

$$\sum_{\substack{q\le n\\ q \text{ prime}}} \frac{1}{q} \ge \ln\ln(n+1) - \ln(\pi^2/6), \tag{5.33}$$

we have

$$\begin{aligned}\sum_{\substack{q\le n\\ q \text{ prime}}} \psi(q) &\ge \frac{1}{\ln\ln\ln n} \sum_{\substack{q\le n\\ q \text{ prime}}} \frac{1}{q}\\ &\underset{(5.33)}{\ge} \frac{\ln\ln(n+1) - \ln(\pi^2/6)}{\ln\ln\ln n} \to \infty,\end{aligned}$$

as $n \to \infty$, thus condition (i) is also satisfied. □

Now we proceed with the proof of Lemma 5.4.

Proof of Property 1. By Corollary 5.6, we have that for almost every $l \in (0,1]$, there exist an increasing sequence of prime numbers $\{q_n\}$ such that $q_n \to \infty$, as $n \to \infty$, and

$$\|q_n l\| \le \frac{1}{q_n \ln\ln\ln q_n}. \tag{5.34}$$

By choosing a subsequence of q_n, we can assume that $q_n \ge n^2$ and $\ln\ln\ln q_n \ge 1$. By Dirichlet's principle, for every $l' \in [0,1]$, for all $n \in \mathbb{N}$, there exists $q'_n \in \mathbb{N}$ such that

$$q'_n \le q_n (\ln\ln\ln q_n)^{1/2}, \tag{5.35}$$

and

$$\|q'_n l'\| \le \frac{1}{q_n (\ln\ln\ln q_n)^{1/2}}. \tag{5.36}$$

Since for almost every $l' \in (0,1]$, there exists a constant $C(l')$ such that $\|nl'\| \ge C(l')/n^2$, the inequality above implies that $q'_n \to \infty$ as $n \to \infty$ for almost every $l' \in (0,1]$. This proves Property 1 of Lemma 5.7. □

Proof of the right hand side of (5.11). Combining (5.34) and (5.36), we have

$$q'_n \|q_n l\| + q_n \|q'_n l'\| \le \frac{2}{(\ln\ln\ln q_n)^{1/2}}.$$

Since b_n, b'_n are respectively the closest integers to $q_n l$ and $q'_n l'$ for $l, l' \in (0,1]$, we have $b_n \le q_n$ and $b'_n \le q'_n$, thus we obtain the right hand side of (5.11) from the inequality above. □

Proof of Property 2. We prove that for almost every $l' \in (0,1]$, we have $(q_n, q'_n) = 1$ when

n is large enough, then Property 2 follows by choosing a subsequence of n. Define

$$G_{n,k} = \left\{ l' \in [0,1] \;\middle|\; \|kl'\| \leq \frac{1}{q_n(\ln\ln\ln q_n)^{1/2}} \right\},$$

and

$$G_n = \bigcup_{k \in S_n} G_{n,k},$$

where

$$S_n = \left\{ k \in \mathbb{N} \;\middle|\; (k, q_n) \neq 1,\ k \leq q_n(\ln\ln\ln n)^{1/2} \right\}$$

denotes the set of k's that violates the coprime condition. Note that when l' does not belong to G_n, the corresponding q_n' that solves (5.35) and (5.36) is coprime to q_n, so it suffices to show that almost every l' belongs to at most finitely many G_n.

Since q_n is prime, if $(k, q_n) \neq 1$, then $(k, q_n) = q_n$ and $q_n \mid k$. So we have

$$\#S_n \leq (\ln\ln\ln q_n)^{1/2}.$$

Thus

$$\begin{aligned}
\mu(G_n) &\leq \sum_{k \in S_n} \mu(G_{n,k}) \\
&\leq \#S_n \cdot \frac{1}{q_n(\ln\ln\ln q_n)^{1/2}} \\
&\leq \frac{(\ln\ln\ln q_n)^{1/2}}{q_n(\ln\ln\ln q_n)^{1/2}} \\
&\leq \frac{1}{n^2},
\end{aligned}$$

and

$$\sum_{n=1}^{\infty} \mu(G_n) \leq \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

By Borel-Cantelli Lemma, almost every $l' \in (0,1]$ belongs to at most finitely many G_n, which proves the coprime condition. □

To prove the left hand side of (5.11), it suffices to show that the two signed distances, $q_n l - b_n$ and $q_n' l - b_n'$, can be of the same sign infinitely many times, and that the lower bound hold for one of the distances. First we introduce a Lemma about the asymptotic estimation about the number of solution of the approximatiion inequality.

Lemma 5.8 ([Lan95, Chapter II, Theorem 7]). *Let $\phi : \mathbb{N} \to [0,1]$ be a decreasing function,*

such that $\sum_{k=1}^{\infty}\phi(k)$ *diverges. For each positive integer* N *and irrational number* l*, let*

$$\Phi(N)=\sum_{k=1}^{N}\phi(k),$$

and let $\lambda(l,N)$ *denotes the number of solutions in integers* b *and* q *of the inequalities*

$$0<ql-b<\phi(k)\quad \textit{and}\quad 1\le q<N.$$

Then for almost every $l\in\mathbb{R}$*, we have*

$$\lambda(l,N)=\Phi(N)+o(\Phi(N)).$$

Therefore, for almost every $l\in\mathbb{R}$*, the number of solutions* (b,q) *for the above inequality is infinite.*

Proof of the left hand side of (5.11). By substituting by a subsequence, we can assume that $q_n\ge 2q_{n-1}$, and q_nl-b_n is positive for all n, i.e.

$$0<q_nl-p_n\le\frac{1}{q_n\ln\ln\ln q_n}.$$

We first show that for almost every $l'\in(0,1]$, there exist infinitely many pairs (b'_n,q'_n) such that $q'_nl'-b'_n$ are positive.

Define $Q_0=q_0=0$, $Q_n=q_n(\ln\ln\ln q_n)^{1/2}$ for $n\ge 1$.

Define $\phi:\mathbb{N}^*\to(0,1]$ by

$$\phi(k)=\frac{1}{Q_n},\quad \text{for}\quad Q_{n-1}<k\le Q_n$$

then $\phi(k)$ is decreasing. Note that

$$\sum_{k=1}^{q_n}\phi(k)=\sum_{i=1}^{n}\left(1-\frac{q_{i-1}(\ln\ln\ln q_{i-1})^{1/2}}{q_i(\ln\ln\ln q_i)^{1/2}}\right)\ge\frac{n}{2},$$

So $\sum\phi(k)$ diverges. By Lemma 5.8, for almost every $l'\in(0,1]$, there exist infinitely many pairs of solutions (b'_n,q'_n), such that

$$0<q'_nl'-b'_n\le\frac{1}{q_n(\ln\ln\ln q_n)^{1/2}},\quad \text{and}\quad 1\le q'_n<q_n(\ln\ln\ln q_n)^{1/2}.$$

This shows that there are infinitely many n's such that q_nl-b_n and $q'_nl'-b'_n$ have the same sign.

It remains to show the lower bound for one of the distances, here we show it for $b_n|q'_nl'-b'_n|$.

By Borel-Cantelli Lemma, for almost every $l' \in (0,1]$, we have

$$\|kl'\| \geq \frac{1}{k(\ln k)^{3/2}},$$

for all k large enough. By the inequality $q'_n \leq q_n(\ln\ln\ln q_n)^{1/2}$, we have

$$\|q'_n l'\| \geq \frac{1}{q_n(\ln q_n)^2} \tag{5.37}$$

for all n large enough, which gives the desired lower bound

$$b_n\|q'_n l'\| \geq \frac{b_n}{q_n}\frac{1}{(\ln q_n)^2} \geq \frac{l}{2(\ln q_n)^2}.$$

□

Proof of Property 4 of Lemma 5.4. From the upper bound (5.35) for q'_n and the lower and upper bounds (5.37) (5.36), we have that q'_n and q_n are approximately of the same magnitude:

$$\frac{q_n}{(\ln q_n)^2} \leq q'_n \leq q_n(\ln\ln\ln q_n)^{\frac{1}{2}}$$

□

Proof of Lemma 5.5. By choosing a subsequence of $\{q_n\}$ if needed, we can assume that

$$q_{n+1} \geq q_n^2(\ln q_n)^5. \tag{5.38}$$

With (5.12) and (5.38), we can choose l_2 such that

$$\frac{1}{2(\ln q_{2n})^3} \leq q'_{2n}\|q_{2n}l_2\| \leq \frac{1}{(\ln q_{2n})^3}, \tag{5.39}$$

$$\frac{1}{2(\ln\ln\ln q_{2n+1})^{1/4}} \leq q'_{2n+1}\|q_{2n+1}l_2\| \leq \frac{1}{(\ln\ln\ln q_{2n+1})^{1/4}}, \tag{5.40}$$

and l'_2 such that

$$q_{2n}\|q'_{2n}l'_2\| \leq \frac{1}{4(\ln q_{2n})^3}, \tag{5.41}$$

$$q_{2n+1}\|q'_{2n+1}l'_2\| \leq \frac{1}{4(\ln\ln\ln q_{2n+1})^{1/4}}. \tag{5.42}$$

The above choice of l_2 (and l'_2) forms naturally a G_δ-dense set of $(0,1]$, as we can approximate any given number well enough by a rational with a denominator that is large enough.

From inequalities (5.39) and (5.41), we have

$$\begin{aligned}\frac{1}{4(\ln q_{2n})^3} &\leq \frac{1}{2(\ln q_{2n})^3} - \frac{1}{4(\ln q_{2n})^3} \\ &\leq \left|q'_{2n}(q_{2n}l_2 - r_{2n}) + q_{2n}(q'_{2n}l'_2 - r'_{2n})\right| \\ &\leq \frac{1}{(\ln q_{2n})^3} + \frac{1}{4(\ln q_{2n})^3} \\ &\leq \frac{3}{2(\ln q_{2n})^3}.\end{aligned}$$

Similarly, from inequalities (5.40) and (5.42), we have

$$\frac{1}{4(\ln\ln\ln q_{2n+1})^{1/4}} \leq \left|q'_{2n+1}(q_{2n+1}l_2 - r_{2n+1}) + q_{2n+1}(q'_{2n+1}l'_2 - r'_{2n+1})\right| \leq \frac{3}{2(\ln\ln\ln q_{2n+1})^{1/4}}.$$

□

Bibliography

[Aar97] J. Aaronson, *An introduction to infinite ergodic theory*, American Mathematical Soc., 1997.

[Arn13] V. I. Arnol'd, *Mathematical methods of classical mechanics*, Vol. 60, Springer Science & Business Media, 2013.

[Atk76] G. Atkinson, *Recurrence of co-cycles and random walks*, Journal of the London Mathematical Society **2** (1976), no. 3, 486–488.

[Bec94] J. Beck, *Probabilistic diophantine approximation, i. kronecker sequences*, Annals of Mathematics (1994), 449–502.

[CC09] N. Chevallier and J.-P. Conze, *Examples of recurrent or transient stationary walks in rd over a rotation of t2*, Contemp. Math **485** (2009), 71–84.

[Dan85] S. G. Dani, *Divergent trajectories of flows on homogeneous spaces and diophantine approximation.* (1985).

[DF12] D. Dolgopyat and B. Fayad, *Deviations of ergodic sums for toral translations ii. boxes*, arXiv preprint arXiv:1211.4323 (2012).

[DF14] ———, *Deviations of ergodic sums for toral translations i. convex bodies*, Geometric and Functional Analysis **24** (2014), no. 1, 85–115.

[DF15] ———, *Limit theorems for toral translations*, Hyperbolic dynamics, fluctuations and large deviations **89** (2015), 227–277.

[DS41] R. J Duffin and A. C Schaeffer, *Khintchine's problem in metric diophantine approximation*, Duke Mathematical Journal **8** (1941), no. 2, 243–255.

[DT06] M. Drmota and R. F Tichy, *Sequences, discrepancies and applications*, Springer, 2006.

[Gal62] P. Gallagher, *Metric simultaneous diophantine approximation*, Journal of the London Mathematical Society **1** (1962), no. 1, 387–390.

[Her62] C. Herz, *Fourier transforms related to convex sets*, Annals of Mathematics (1962), 81–92.

[Kes60] H Kesten, *Uniform distribution mod 1*, Annals of Mathematics (1960), 445–471.

[Kes62] ———, *Uniform distribution mod 1 (ii)*, Acta Arithmetica **7** (1962), no. 4, 355–380 (eng).

[Khi23] A Khintchine, *Ein satz über kettenbrüche, mit arithmetischen anwendungen*, Mathematische Zeitschrift **18** (1923), no. 1, 289–306.

[Kok13] J. F. Koksma, *Diophantische approximationen*, Vol. 4, Springer-Verlag, 2013.

[Lan95] S. Lang, *Introduction to diophantine approximations*, Springer Science & Business Media, 1995.

[Mar07] J. Marklof, *Distribution modulo one and ratner's theorem*, Equidistribution in number theory, an introduction, 2007, pp. 217–244.

[MS10] J. Marklof and A. Strömbergsson, *The distribution of free path lengths in the periodic lorentz gas and related lattice point problems*, Annals of Mathematics (2010), 1949–2033.

[Rag72] M. S. Raghunathan, *Discrete subgroups of lie groups*, Vol. 3, Springer, 1972.

[Sch60] W. Schmidt, *A metrical theorem in diophantine approximation*, Canadian Journal of Mathematics **12** (1960), 619–631.

[Sch64] W. M Schmidt, *Metrical theorems on fractional parts of sequences*, Transactions of the American Mathematical Society **110** (1964), no. 3, 493–518.

[Sch77] K. Schmidt, *Cocycles on ergodic transformation groups*, Vol. 1, SG Wasani, 1977.