Université de Lorraine Ecole doctorale IAEM

Résumé de travaux

présenté pour obtenir

l'**Habilitation à Diriger des Recherches**

en

Mathématiques, CNU 25

par

Manfred Madritsch

Méthodes analytiques, probabilistes et dynamiques dans l'étude des systèmes de numération.

Date de soutenance : 13 décembre 2018

Jury :

Valérie Berthé	Directrice de recherche, Université Paris-Diderot	Rapporteur
Yann Bugeaud	Professeur, Université de Strasbourg	Examinateur
Cécile Dartyge	Maître de Conférences HDR, Université de Lorraine	Examinatrice
Michael Drmota	Professeur, Université technique de Vienne	Examinateur
Christian Mauduit	Professeur, Aix-Marseille Université	Rapporteur
Joël Rivat	Professeur, Aix-Marseille Université	Examinateur
András Sárközy	Professeur, Université Loránd Eötvös	Rapporteur
Gérald Tenenbaum	Professeur, Université de Lorraine	Examinateur

RÉSUMÉ

Dans la vie quotidienne, nous avons souvent besoin de représenter des nombres à l'aide de systèmes de numération. Dans ce mémoire de HDR nous envisageons de tels systèmes sous plusieurs aspects différents. Nous commençons par étudier les bases possibles pour ces systèmes et la distribution de la longueur des partitions en entiers dans l'écriture desquels certains chiffres n'apparaissent pas. Puis nous focalisons notre attention sur la distribution des chiffres d'une écriture et, en particulier, sur les fonctions qui opèrent seulement sur les chiffres de l'écriture. Pour les valeurs d'une telle fonction nous démontrons un théorème limite central. Une méthode combinatoire nous permet de prouver une conjecture de Stolarsky sur le ratio des sommes des chiffres. En utilisant des résultats d'équirépartitions ainsi que la méthode du cercle, nous présentons deux applications : l'une concerne les ensembles intersectifs et l'autre les formes de degrés différents sur un corps de nombres. Après cette digression nous nous consacrons aux nombres normaux et à leurs constructions. Nous donnons une construction en utilisant des nombres premiers, puis une construction pour une mesure arbitraire et enfin une construction d'un nombre normal par rapport à toute base. Bien presque tout nombre soit normal, les nombres non normaux jouent un rôle important en topologie. Nous concluons notre travail par le calcul de la dimension d'Hausdorff de certains ensembles des nombres non normaux.

ABSTRACT

In everyday live we need to represent numbers with a numeration system. The present habilitation deals with server aspects of these systems. Starting with the determination of the possible bases we look at the length of partitions of an integer in integers with missing digits. Then we focus on the distribution of the digits of in the representation and, in particular, on function operating only on the digits of the expansion. We show a central limit theorem for the function values. Using a combinatorial argument we are able to settle a conjecture of Stolarsky on the ratio of sum of digits functions. Similar methods in equidistribution theory and the circle method allows us to present two application on intersective sets and forms of different degrees over number fields. After this short excursion we return to normal numbers and their construction. We present a construction involving primes, one for arbitrary measure as well as a construction of a number, which is normal to any base. Even though almost every number is normal, non-normal numbers play a curcial rôle in topology. The calculation of the Hausdorff-dimension of certain sets of non-normal numbers closes the cercle of our research.

Table des matières

Chapitre 1

Introduction

Un système de numération est un ensemble de règles pour représenter les nombres entiers. Cette représentation permet de comptabiliser des récoltes, faire du commerce ou établir un calendrier. Le concept de système de numération est donc aussi ancien que l'écriture. Pour l'aspect historique, nous nous référons souvent au livre de Ifrah [**111**].

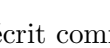
Le système le plus ancien est le système unaire où un symbole unique représente l'unité. Si "|" est le symbole utilisé, alors 4 s'écrit "||||". Avec ce système, on peut très facilement représenter et faire des calculs pour des nombres petits comme 4, 7 ou même 20. Mais pour des nombres plus grands il est nécessaire de combiner un certain nombre de ces symboles représentants l'unité par un autre symbole. Un exemple de système de numération avec plusieurs symboles est la numération romaine où

I	V	X	L	C	D	M
l'unité	5	10	50	100	500	1000

Un tel système est appelé additif, parce que les nombres sont représentés par l'addition des valeurs des différents symboles. Dans la numération romaine le nombre 2016 s'écrit "MMXVI" qui signifie $1000 + 1000 + 10 + 5 + 1 = 2016$. Un inconvénient du système romain est que la valeur dépend de la position : quatre s'écrit IV et six s'écrit VI. On utilise les même symboles mais la valeur finale est différente.

Les Égyptiens de l'Antiquité utilisaient déjà un système décimal sans zéro (voir aussi [**248**]). Pour chaque puissance de 10 ils employaient un symbole différent :

					
1	10	100	1000	10000	100000

Des groupes de symboles sont parfois séparés par un espace pour améliorer la lisibilité. Par exemple, le nombre 2016 s'écrit comme 

Une écriture similaire est développée par les Grecs de l'Antiquité. Ils utilisaient les lettres courantes de leur alphabet. Mais l'alphabet grec a seulement 25 lettres ; trois lettres archaïques ont donc été ajoutées : le digamma ϝ, le koppa ϟ, et le sampi ϡ. Nous avons ainsi

α	1	alpha	ι	10	iota	ρ	100	rô
β	2	bêta	κ	20	kappa	σ	200	sigma
γ	3	gamma	λ	30	lambda	τ	300	tau
δ	4	delta	μ	40	mu	υ	400	upsilon
ϵ	5	epsilon	ν	50	nu	φ	500	phi
ϝ	6	digamma	ξ	60	ksi	χ	600	khi
ζ	7	dzêta	o	70	omicron	ψ	700	psi
η	8	êta	π	80	pi	ω	800	oméga
θ	9	thêta	ϟ	90	koppa	ϡ	900	sampi

Dans un système où les lettres sont aussi des chiffres, un énoncé comme « Soit x la solution de l'équation ... » est malheureusement impossible à résoudre, parce que x a déjà une valeur définie.

Un autre défaut est qu'on on peut seulement représenter des nombres jusqu'à 999. L'introduction de la myriade M' représentant mille (10^3) et une notation plus complexe permettent de représenter des nombres plus grands. Aristarque de Samos, Diophante d'Alexandrie et Apollonios de Perga ont notamment proposé d'autres méthodes.

À cette époque les Grecs aimaient montrer la supériorité de leur savoir. Un tel exemple est le problème des bœufs d'Hélios attribué à Archimède dont le but est déterminer la taille du troupeau des bœufs d'Hélios (*cf.* [**134**, **222**]). En 1769, Gotthold Ephraim Lessing était responsable de la bibliothèque August Herzog à Wolfenbüttel (Allemagne). Quelques années plus tard y il a retrouvé une lettre d'Archimède à Ératosthène qui contient ce problème sous forme de poème. Lessing et d'autres doutent que ce poème soit authentique. En revanche, le problème lui-même est si difficile qu'on l'attribue effectivement à Archimède. Le problème comporte deux parties :

(1) sept équations pour les taureaux blancs, noirs, pies et jaunes et les vaches blanches, noirs, pies et jaunes ;

(2) il faut que la somme des nombres des taureaux blancs et noirs soit un carré parfait et que la somme des effectifs des taureaux pies et des taureaux jaunes soit un nombre triangulaire.

La première partie donne un nombre total de $50389082 \cdot k$ têtes de bétail, où k est un entier positif. La résolution de la deuxième revient à résoudre l'équation (de Pell)

$$y^2 - 410286423278424x^2 = 1.$$

On a constaté à travers ces problèmes qu'Archimède était non seulement capable de poser un tel problème mais aussi qu'il pouvait calculer avec des nombres aussi élevés que 410286423278424. D'une part, il n'est absolument pas clair qu'une telle équation ait une solution en nombres entiers. D'autre part, la taille des coefficients a rendu le problème insoluble pendant une longue période. En particulier une méthode de résolution a été trouvée par August Amthor en 1880 (voir [**134**]). En 1965, Hugh Williams, Gus German et Bob Zarke [**253**] utilisent deux ordinateurs (IBM 7040 et IBM 1620) pour calculer la représentation en base 10 en 7 heures et 49 minutes.

Pour obtenir la plus petite solution de cette équation il faut calculer des réduites du développement en fraction continue de $\sqrt{4729494}$. Mais la plus petite solution ne résout pas le problème des bœufs d'Hélios, parce que le nombre correspondant pour les têtes n'est pas entier. On peut déterminer toutes les solutions en itérant et après un nombre incroyable de 2329 itérations on obtient une solution entière à 206545 chiffres.

1. Notation positionnelle

Après les systèmes additifs nous évoluons vers les systèmes positionnels. Nous suivons ici l'exposé très profond du Chapitre 4.1 de Knuth [**124**]. Une notation positionnelle en base q est définie par la règle

$$(\ldots a_3a_2a_1a_0.a_{-1}a_{-2}\ldots)_q = \cdots + a_3q^3 + a_2q^2 + a_1q^1 + a_0 + a_{-1}q^{-1} + a_{-2}q^{-2} + \cdots;$$

par exemple $(520.3)_6 = 5\cdot 6^2 + 2\cdot 6^1 + 0 + 3\cdot 6^{-1} = 192\frac{1}{2}$. Le grand avantage d'un tel système de numération par rapport aux systèmes précédents est l'utilisation d'un nombre fini de symboles

pour la représentation ainsi que la possibilité de représenter n'importe quel élément, grand ou petit. Même si, avec son système, Archimède pouvait représenter des nombres très grands, il y avait une limite pour leur taille maximale.

Supposons que $0 \leq a_i < q$. Cela donne le système binaire ($q = 2$), le système ternaire ($q = 3$), le système quaternaire ($q = 4$), etc. En général nous pouvons choisir les a_i dans n'importe quel ensemble et pour q prendre n'importe quel nombre réel strictement plus grand que 1. Nous étudierons ces notations positionnelles générales plus loin.

La notation positionnelle apparaît pour la première fois vers le III$^{\text{e}}$ millénaire av. J.-C. en Mésopotamie, donc avant les anciens Grecs. Les nombres utilisés quotidiennement sont en notation additive comme précédemment. Mais les mathématiciens utilisent un système sexagésimal (en base 60) pour leurs calculs. Ce système était très avancé pour 1750 av. J.-C. du fait qu'il s'agit d'un système à virgule flottante. Il est fascinant que, dans ce système, les nombres 2, 120, 7200 et $\frac{1}{30}$ s'écrivent de la même manière.

Il y a 2000 ans que les Mayas et les Aztèques utilisent le chiffre zéro dans leur système vigésimal (base vingt). Notre chiffre 0 est venu de l'Inde où le zéro (ou plutôt le néant), comme le nirvana, a une valeur positive. Cela est peut-être aussi la raison pour laquelle les anciens Grecs ne l'ont pas utilisé. La date du premier usage du chiffre 0 n'est pas claire ce serait environ 600 ap. J.-C. La science hindoue était très avancée et, dans les premiers manuscrits, l'écriture des nombres est renversée. Mais on a rapidement utilisé l'écriture moderne. En 750 ap. J.-C. environ les Perses commencent à traduire les principes hindous d'arithmétique. Même si le sens de leur écriture est de droite à gauche, ils utilisent l'écriture de gauche à droite pour les nombres en base dix.

La notation en base 10 est, au début seulement, utilisée pour écrire des nombres entiers. Pour leurs calculs les astronomes arabes ont besoin des nombres rationnels dans leurs tableaux et leur calendriers. En choisissant le système sexagésimal (en base 60) ils continuent le travail de Ptolémée (le célèbre astronome grec). Ce système est encore présent dans notre mesure d'angle plan (degré, minute, seconde, ...) ou la chronométrie (heure, minute, seconde, ...). Les premiers mathématiciens européens utilisent aussi le système sexagésimal pour leurs calculs. Par exemple, Fibonacci a donné la valeur

$$1^\circ\, 22'\, 7''\, 42'''\, 33^{IV}\, 4^{V}\, 40^{VI}$$

pour approcher une racine du polynôme $x^3 + 2x^2 + 10x - 20$. Les fractions décimales apparaissent sporadiquement au XVI$^{\text{è}}$ siècle. Après la découverte de la fonction logarithme, les fractions décimales sont utilisées partout pendant le XVII$^{\text{è}}$ siècle.

Les nombres négatifs et le signe ont toujours joué un rôle important. Pour les ordinateurs, il y a plusieurs façons d'écrire un entier négatif en base 2 : avec le signe moins, le complément à un ou le complément à deux.

- La méthode la plus simple est d'ajouter le symbole « - » à l'écriture. Ainsi -3 est l'opposé de 3, tel que $(-3) + 3 = 3 + (-3) = 0$.
- Une autre méthode est le complément à un qui inverse tout les chiffres : $\overline{1101} = 0010$. Cette représentation possède un inconvénient : on ne peut pas sommer deux représentations de manière binaire. Par exemple, 4 est représenté par 0100 et 3 par 0011. Donc -3 est $\overline{0011} = 1100$. Leur somme est $0100 + 1100 = 0000$ qui n'est pas la somme de $4 + (-3) = 1$.
- La troisième méthode est le complément à deux (exposant N). On associe à l'écriture en base 2 de longueur N les nombres de -2^{N-1} à $2^{N-1} - 1$. Cette représentation a

l'avantage que l'on peut ajouter deux valeurs selon la méthode standard et le résultat est correct. Par exemple, soit $N = 4$. Alors la représentation de 4 est 0100 et celle de -3 est 1101. Si nous sommons, $4+(-3) = 0100+1101$, nous obtenons bien $0001 = 1$.

Une idée différente consiste à choisir -2 comme base :

$$\begin{aligned}(\ldots a_3a_2a_1a_0.a_{-1}a_{-2}\ldots)_{-2}&\\ &= \cdots + a_3(-2)^3 + a_2(-2)^2 + a_1(-2)^1 + a_0 + a_{-1}(-2)^{-1} + a_{-2}(-2)^{-2} + \cdots\\ &= \cdots - 8a_3 + 4a_2 - 2a_1 + a_0 - \tfrac{1}{2}a_{-1} + \tfrac{1}{4}a_{-2} - \cdots .\end{aligned}$$

En 1885, Vittorio Grünwald [**99**] a considéré les systèmes de bases négatives. Il a donné des algorithmes pour calculer les racines, pour tester la divisibilité et pour changer de base. Mais il a publié ses résultats dans un journal peu connu et ils sont tombés dans l'oubli. La publication suivante sur les systèmes aux base négatives est due à Kempner [**121**] en 1936. Ensuite, il y a eu de nombreux travaux sur cette question.

Les bases négatives fournissent une représentation pour les entiers (positifs et négatifs) pour les entiers mais on peut aussi considérer les entiers de Gauss $\mathbb{Z}[i]$ ou d'Eisenstein $\mathbb{Z}[j]$. Knuth [**126**] a montré que la base $2i$ donne un système de numération "quarter-imaginary" (quart-imaginaire). Par exemple,

$$(11210.31)_{2i} = 1\cdot 16 + 1\cdot(-8i) + 2\cdot(-4) + 1\cdot(2i) + 3\cdot(-\tfrac{1}{2}i) + 1(-\tfrac{1}{4}) = 7\tfrac{3}{4} - 7\tfrac{1}{2}i.$$

Le nom "quart" vient du fait que

$$\begin{aligned}(a_{2n}\ldots a_1a_0.a_{-1}\ldots a_{-2k})_{2i}&\\ &= (a_{2n}\ldots a_2a_0.a_{-2}\ldots a_{-2k})_{-4} + 2i(a_{2n-1}\ldots a_3a_1.a_{-1}\ldots a_{-2k+1})_{-4}.\end{aligned}$$

Kátai et Szabó [**120**] montrent que les seuls systèmes possibles dans $\mathbb{Z}[i]$ ont des bases de la forme $-m \pm i$ avec $m \geq 1$ et l'ensemble des chiffres $\{0, 1, \ldots, m^2\}$. Tous les systèmes dans les corps quadratiques sont caractérisés par Gilbert [**92**] et Kátai et Kovács [**118**, **119**] indépendamment. On peut en déduire que, pour l'anneau d'Eisenstein $\mathbb{Z}[j] = \mathbb{Z}\left[\frac{1+\sqrt{-3}}{2}\right]$, les bases possibles sont de la forme $-m \pm j$ avec $m \geq 1$.

Pour des systèmes de numération dans un anneau d'entiers algébriques ou dans un ordre (au sens de la théorie des anneaux), il faut élargir notre vision. Soit

$$P = p_dX^d + \cdots + p_1X + p_0 \in \mathbb{Z}[X]$$

un polynôme à coefficients entiers et soit $\mathcal{R} = \mathbb{Z}[X]/(P)$ l'anneau quotient. Nous cherchons des critères tels que tout polynôme $g \in \mathcal{R}$ ait une représentation de la forme

$$g \equiv \sum_{k=0}^{n} a_kX^k \bmod P$$

avec $a_k \in \{0, 1, \ldots, |p_0| - 1\}$. Il est clair que les polynômes $P = X + q$ donnent les systèmes q-adiques et les polynômes $P = X^2 + 2mX + (m^2 + 1)$ donnent les bases dans $\mathbb{Z}[i]$. De même Kovacs et Pethő [**129**] ont trouvé un algorithme qui détermine si un entier algébrique est une base. L'algorithme comporte deux étapes :

(1) la première détermine si la base est bien contractante, c'est-à-dire, qu'il existe une version de la division euclidienne qui à chaque étape de diminuer la « taille » jusqu'à ce que le reste apparaisse dans un ensemble fini ;

(2) la deuxième étape montre que tous les éléments de cet ensemble fini possèdent une représentation.

Ces systèmes ont aussi une jolie connexion avec la topologie des fractales et des pavages. L'ensemble $[0,1]$ contient tous les représentations q-adiques avec des puissances négatives. Si nous considérons l'ensemble correspondant pour la base $(-1+i)$:

$$\mathcal{F} := \left\{ \sum_{k\geq 1} a_k(-1+i)^{-k} \colon a_k \in \{0,1\} \right\}$$

nous obtenons le dragon de Knuth (Figure 1). En particulier cette fractale permet un pavage du plan complexe et les intersections à la frontière coïncident avec les éléments ayant plusieurs représentations. D'autres propriétés sont encodées dans ce domaine fondamental et nous renvoyons le lecteur intéressé à l'ouvrage de Berthé et Rigo [**33**].

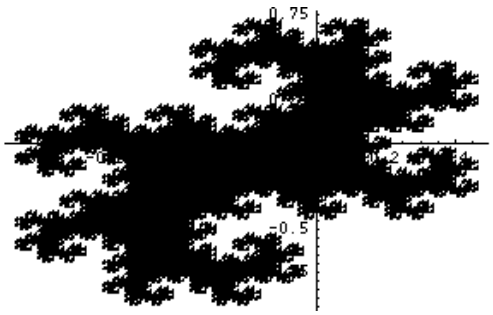


FIGURE 1. Le dragon de Knuth $\mathcal{F}$

Considérons à présent la connexion entre des systèmes de numération et l'informatique, ou plus précisément la cryptographie. Pour calculer les multiples d'un point sur une courbe elliptique ou la puissance pour le chiffrement RSA, on utilise toujours un algorithme de « square and multiply » (élever au carré et multiplier). Par exemple, si l'on veut calculer x^{27} on note la représentation binaire de $(27)_2 = 11011$. On a alors

$$x^{27} = \Big(\big(x^2 \cdot x \big)^4 \cdot x \Big)^2 \cdot x.$$

Plus généralement, si

$$n = \sum_{k=0}^{\ell} a_k 2^k \quad \text{avec } a_k \in \{0,1\}$$

est l'écriture binaire de n, alors

$$x^n = \left(\cdots \Big((x^{a_\ell})^2 \cdot x^{a_{\ell-1}} \Big)^2 \cdots x^{a_1} \right)^2 \cdot x^{a_0}.$$

Autrement dit, nous avons l'algorithme récursif suivant :

$$\text{puissance}(x,n) = \begin{cases} x & \text{si } n = 1, \\ \text{puissance}\left(x^2, n/2\right) & \text{si } n \text{ est pair}, \\ x \times \text{puissance}\left(x^2, (n-1)/2\right) & \text{si } n > 2 \text{ est impair}. \end{cases}$$

Dans chaque itération, l'algorithme élève au carré et si la représentation binaire de la puissance a le chiffre 1 à la position correspondante il multiplie par x, sinon il ne fait rien. Comme chaque multiplication consomme du temps et de l'énergie, on peut lancer une attaque par un canal auxiliaire et en déduire la clé secrète. Le but de la cryptanalyse est de cacher cette consommation. Une méthode consiste à ajouter le chiffre -1, écrit $\overline{1}$. Avec ce chiffre, l'écriture binaire signée de 27 devient $(27)_2 = 100\overline{1}0\overline{1}$ et nous avons diminué le nombre de chiffres non nuls (ce nombre est appelé le poids de Hamming). Pour calculer l'inverse de x nous utilisons l'algorithme d'Euclide : le calcul de la puissance devient

$$x^{27} = \left(x^4 \cdot x^{-1}\right)^4 \cdot x^{-1}.$$

De plus si on suppose que $a_k a_{k+1} = 0$ pour $k \geq 0$, alors l'écriture est unique. À cause de cette restriction, on désigne une telle écriture par le terme de forme non adjacente (« non adjacent form » ou NAF). En effet, cette écriture possède deux avantages importants pour l'informatique :

(1) son poids de Hamming est minimal parmi les écritures équivalentes possibles ;

(2) on peut obtenir cette écriture en parcourant l'écriture en base 2 avec un transducteur fini.

Une autre généralisation est l'utilisation d'une suite différente pour la base. Par exemple, soient $F_0 = 1$ et $F_1 = 2$ et définissons $F_k = F_{k-1} + F_{k-2}$ pour $k \geq 2$ récursivement. Le théorème de Zeckendorf [**254**] dit que tout entier positif n possède une représentation de la forme

$$n = \sum_{k \geq 0} a_k F_k \quad \text{avec } a_k \in \{0, 1\}.$$

Comme précédement, cette écriture devient unique si l'on suppose que $a_k \cdot a_{k+1} = 0$ pour $k \geq 0$. Fraenkel [**82**] a généralisé ce système de numération aux suites strictement croissantes $1 = G_0 < G_1 < G_2 < \cdots$. Sous certaines conditions on peut montrer que l'écriture est unique. Ici presque toute la littérature se concentre sur les suites récurrentes.

Le défaut de ces systèmes de numération est qu'il ne fournit des représentations que pour les entiers positifs. Supposons que nous ayons la suite des nombres Fibonacci ($F_0 = 0$ et $F_1 = 1$). Si l'on continue la récurrence pour des valeurs négatives de k ($F_{k-2} = F_k - F_{k-1}$), on obtient la suite complète

$$\ldots, -8, 5, -3, 2, -1, 1, 0, 1, 1, 2, 3, 5, 8, \ldots$$

En particulier, Knuth [**125**] a montré que tout entier n relatif possède une représentation avec des nombres de Fibonacci d'indices strictement négatifs. Ces systèmes sont considérés en toute généralité par Anne Bertrand-Mathis [**37**]. Son résultat utilise la connexion entre les systèmes dont les bases sont les éléments de suites récurrentes et les systèmes de numération β.

Comme pour la suite des puissances d'un entier (*cf.* les écritures q-adiques), on peut choisir une suite des puissances d'un réel. Soit $\beta > 1$, alors tout entier positif n possède une représentation de la forme

$$n = \sum_{k \geq -\ell} a_k \beta^{-k} \quad \text{avec } a_k \in \{0, 1, \ldots, \lceil \beta \rceil - 1\}.$$

Si β est la racine dominante du polynôme caractéristique d'une équation de récurrence linéaire $(G_0)_{k \geq 0}$, les représentations dans les systèmes de β et de G sont liées. Ces dernières années

la recherche s'est plutôt intéressée aux systèmes $-\beta$, parce que ces derniers fournissent des représentations pour des entiers négatifs également.

2. Les bases des systèmes de numération

Après avoir exploré les différents systèmes de numération, revenons sur les systèmes q-adiques. Soit $q \in \mathbb{Z}$ et $\mathcal{N}_q = \mathcal{N} = \{0, 1, \ldots, |q| - 1\}$. Nous disons que $(q, \mathcal{N})$ est un système de numération dans $\mathbb{Z}$ si chaque entier $n \in \mathbb{Z}$ possède une représentation unique de la forme

$$n = \sum_{k=0}^{\ell} a_k q^k \quad \text{avec } a_k \in \mathcal{N} \text{ et } a_\ell \neq 0.$$

Nous désignons la base par q et l'ensemble de chiffres par $\mathcal{N}$.

Une première question porte sur la détermination des bases possibles. Il est clair que $q \geq 0$ ne donne pas de représentation pour des entiers négatifs. Le choix $q = -1$ induit des ambiguïtés. Nous considérons donc $q \leq -2$. Nous démontrons que $(q, \mathcal{N})$ est un système de numération en $\mathbb{Z}$ en deux étapes.

— **Existence :** Pour $0 \leq n < |q|$ on a la représentation $n = a_0$. Supposons maintenant qu'on a une représentation pour $|m| < |n|$. Alors il existe ℓ unique tel que $q^\ell \leq n < q^{\ell+1}$. Par la unicité de la division euclidienne, il existe a et b avec $n = aq^\ell + b$ avec $1 \leq a < q$ et $0 \leq b < q^\ell \leq n$. Alors b admet une représentation $b = \sum_{k=0}^{\ell-1} a_k q^k$ et donc

$$n = aq^\ell + \sum_{k=0}^{\ell-1} a_k q^k$$

est une représentation de n.

— **Unicité :** Supposons que n admettent deux représentations différentes :

$$n = \sum_{k=0}^{\ell} a_k q^k = \sum_{k=0}^{m} b_k q^k \text{avec } a_k, b_k \in \{0, 1, \ldots, |q| - 1\}.$$

Soit j le plus petit entier tel que $a_k \neq b_k$. On obtient

$$n \equiv \sum_{k=0}^{j} a_k q^k \equiv \sum_{k=0}^{j} b_k q^k \bmod q^{j+1}.$$

Alors $a_j = b_j + kq$ avec $k \neq 0$, ce qui contredit $a_j, b_j \in \{0, 1, \ldots, |q| - 1\}$.

Cette idée fonctionne si la division euclidienne donne un quotient et un reste uniques. Dans l'anneau des entiers de Gauss $\mathbb{Z}[i]$, il faut donc faire autrement. Comme ci-dessus nous disons que le couple $(\theta, \mathcal{N})$ avec $\mathcal{N} = \{0, 1, \ldots, |\theta|^2 - 1\}$ est un système de numération canonique dans $\mathbb{Z}[i]$ si tout $\gamma \in \mathbb{Z}[i]$ admet une représentation de la forme

$$\gamma = \sum_{k=0}^{\ell} a_k \theta^k \quad \text{avec } a_k \in \mathcal{N}. \tag{2.1}$$

Knuth [**124**] a montré que $-1 + i$ et $-1 - i$ sont des bases avec l'ensemble des chiffres $\{0, 1\}$. Si nous considérons les nombres complexes dont la représentation n'inclut que des puissances

négatives de la base

$$\mathcal{F} := \left\{ \sum_{k\geq 1} a_k(-1+i)^{-k} \colon a_k \in \{0,1\} \right\},$$

alors nous obtenons le dragon de Knuth. On peut montrer que toute base canonique dans $\mathbb{Z}[i]$ est de la forme $-m \pm i$ avec $m \geq 1$ un entier.

THÉORÈME 1.1 ([**120**, Theorem 1]). *Le couple $(\theta, \mathcal{N})$ est un système de numération dans $\mathbb{Z}[i]$ si et seulement si*

$$\mathrm{Re}(\theta) < 0 \quad \textit{et} \quad \mathrm{Im}(\theta) = \pm 1.$$

De même, on peut montrer que toute base de l'anneau des entiers d'Eisenstein $\mathbb{Z}[j]$ avec $j = \exp\left(\frac{2\pi i}{3}\right) = \frac{-1+\sqrt{-3}}{2}$ est de la forme $-m \pm j$. Avant de poursuivre, nous aimerions donner une définition assez générale d'un système de numération canonique.

DÉFINITION 1.1 (Système de numération canonique, *cf.* [**179**, **211**]). Soit $P = p_d x^d + p_{d-1}x^{d-1} + \cdots + p_0 \in \mathbb{Z}[x]$ un polynôme (non nécessairement unitaire) et $\mathcal{Z} := \mathbb{Z}[x]/(P)$ l'anneau quotient. Si tout élément $\gamma \in \mathcal{Z}$ a une représentation comme

$$\gamma = \sum_{j=0}^{\ell} a_j X^j \text{ avec } a_j \in \mathcal{N} := \{0, 1, \ldots, |p_0| - 1\} \tag{2.2}$$

(où $X \in \mathcal{Z}$ est l'image de $x \in \mathbb{Z}[x]$ par l'épimorphisme canonique) alors la paire $(P, \mathcal{N})$ est appelée un *système canonique de numération* (CNS)) et P est un *polynôme CNS*.

Pour des polynômes irréductibles P, il est clair que $(P, \mathcal{N})$ est un CNS si et seulement si $(q, \mathcal{N})$ est un système de numération dans $\mathbb{Z}[q]$, pour tout $q \in \mathbb{C}$ avec $P(q) = 0$. Nous avons vu que le couple $(q, \{0, \ldots, |q| - 1\})$ avec $q \leq -2$ un entier est un système de numération dans $\mathbb{Z}$. De plus, $-m \pm i$ et $-m \pm j$ sont les bases dans l'anneau des entiers de Gauss $\mathbb{Z}[i]$ ou d'Eisenstein $\mathbb{Z}[j]$.

Kátai et Szabó [**120**] montrent que les systèmes possibles dans $\mathbb{Z}[i]$ ont des bases $-m \pm i$ avec $m \geq 1$ et l'ensemble des chiffres $\{0, 1, \ldots, m^2\}$. Tous les systèmes dans les corps quadratiques ont été caractérisés par Gilbert [**92**] et par Kátai et Kovács [**118**, **119**] indépendamment. Le cas des anneaux des entiers pour des corps de plus haut degré est plus difficile. Même si l'on connaît des algorithmes qui déterminent si un couple donné $(\beta, \{0, 1, \ldots, \mathrm{N}(\beta)\})$ est un système de numération, cela fournit un résultat pour le polynôme donné et pas pour une classe de polynômes. Cependant, par exemple, Kovács et Pethő [**129**] obtiennent le résultat suivant.

THÉORÈME 1.2 ([**129**, Theorem 3]). *Soit α un entier algébrique sur $\mathbb{Q}$. Soit $\beta \in \mathbb{Z}[\alpha]$, $\mathcal{N} \subset \mathbb{Z}$ et posons $A = \max_{a\in\mathcal{N}} |a|$. Alors $\{\beta, \mathcal{N}\}$ est un système de numération dans $\mathbb{Z}[\alpha]$ si et seulement si*

(1) $\left|\beta^{(j)}\right| > 1$ pour $j = 1, 2, \ldots, n$.

(2) $\mathcal{N}$ est un ensemble de restes modulo $|\mathrm{N}(\beta)|$ contenant 0,

(3) $\alpha \in \mathbb{Z}[\beta]$,

(4) chaque $\gamma \in \mathbb{Z}[\alpha]$ avec

$$\left|\gamma^{(j)}\right| \leq \frac{A}{\left|\beta^{(j)}\right| - 1}, \quad (j = 1, \ldots, n)$$

a une représentation dans $(\beta, \mathcal{N})$,

où $\beta^{(j)}$ avec $j = 1, 2, \ldots, n$ sont les conjugués de β dans le corps $\mathbb{Q}(\alpha)$.

Comme l'anneau des entiers de Gauss et l'anneau d'Eisenstein sont aussi les anneaux des entiers des corps cyclotomiques d'ordre 4 et 3, respectivement, nous nous demandons si pour tout anneau des entiers d'un corps cyclotomique $\mathbb{Z}[\zeta_k]$ les bases sont de la forme $-m+\zeta_k$ avec $m \geq 1$. Il est clair que le théorème de Kovács et Pethő ne sera pas le bon outil pour ce cas spécial.

Kovács [**128**] a montré que pour tout corps de nombres K et un ordre $\mathcal{O} = \mathbb{Z}[\alpha]$ dans K il existe β tel que $(\beta, \mathcal{N})$ est un système de numération canonique dans $\mathcal{O}$. De plus, supposons que $1 \leq p_{d-1} \leq \cdots \leq p_1 \leq p_0$, $p_0 \geq 2$ et que P soit irréductible. Alors Kovács a également montré que $(P, \mathcal{N})$ est un système de numération canonique dans $\mathcal{O}$. Pethő [**179**] a omis la condition sur le polynôme d'être irréductible en supposant qu'il n'a pas de racine d'unité. En vérifiant cette condition des coefficients croissants, Volker Ziegler et moi-même [**154**] pouvons démontrer le théorème suivant.

THÉORÈME 1.3. *Soient $k > 2$ et m des entiers positifs. Soient ζ_k une racine primitive de l'unité et $\mathcal{D} = \{0, 1, \ldots, \mathrm{N}(\zeta_k)\}$ où* N *est la norme algébrique. Si $m \geq \varphi(k) + 1$, alors $(-m + \zeta_k, \mathcal{D})$ est un système de numération.*

La borne $\varphi(k) + 1$ est optimale pour la méthode utilisée. En particulier, nous montrons que le polynôme minimal de la base possède des coefficients croissants. Pour $m = \varphi(k)$, il existe des valeurs de k pour lesquelles le polynôme minimal a des coefficients croissants et des valeurs de k pour lesquelles le polynôme minimal n'a pas de coefficients croissants. Par exemple, les polynômes minimaux de $-6 + \zeta_9$ et $-4 + \zeta_{10}$ sont

$$x^6 + 36x^5 + 540x^4 + 4321x^3 + 19458x^2 + 46764x + 46873$$

et

$$x^4 + 15x^3 + 85x^2 + 215x + 205$$

respectivement.

Soit β un entier algébrique. Hormis notre résultat on ne connaît pas de système de numération pour des corps de nombres de degré plus élevé.

Pour les constructions des nombres normaux en plusieurs bases et le théorème de Cobham (que nous allons considérer ci-dessous) il faut décider si deux bases sont multiplicativement indépendantes. On dit que deux nombres algébriques α et β sont multiplicativement dépendants si l'équation $\alpha^m = \beta^n$ a une solution différente de la solution triviale $(m, n) = (0, 0)$. Pour les bases $-m+\zeta_k$ dans $\mathbb{Z}[\zeta_k]$ Volker Ziegler et moi-même [**154**] montrons le théorème suivant.

THÉORÈME 1.4. *Soit $k \geq 3$ un entier positif. Alors les entiers algébriques $-m + \zeta_k$ et $-n + \zeta_k$ sont multiplicativement indépendants si $m > n > C(k)$, où $C(k)$ est une constante effectivement calculable qui ne dépend que de k.*

De plus si k est une puissance de $2, 3, 5, 6, 7, 11, 13,$ $17, 19$ ou 23, alors $-m+\zeta_k$ et $-n+\zeta_k$ sont multiplicativement indépendants pour $m > n > 0$ si $k \neq 6$ et $m > n > 1$ sinon.

Ce théorème repose sur l'équation de Nagell-Ljunggren

$$\frac{x^k - 1}{x - 1} = y^q \quad x, y > 1, q \geq 2, k > 2.$$

Il est conjecturé que

$$\frac{3^5-1}{3-1}=11^2,\quad \frac{7^4-1}{7-1}=20^2,\quad \frac{18^3-1}{18-1}=7^3,$$

sont les seules solutions de l'équation. Cela nous incite à conjecturer que $-m+\zeta_k$ et $-n+\zeta_k$ sont multiplicativement indépendants pour tout $m>n>1$ (sauf dans certains cas pathologiques). En utilisant une méthode différente nous allons montrons ce résultat ci-dessous.

Un autre avantage de ce dernier théorème est que nous avons montré l'indépendance si on fixe k et qu'on laisse varier m et n. Dans un deuxième article, Volker Ziegler et moi [**148**] montrons l'indépendance si on fixe la différence entre m et n en laissant varier k. En particulier, si m et a sont deux entiers positifs fixés, nous considérons l'indépendance des bases $-m+\zeta_k$ et $-(m+a)+\zeta_k$ pour $k\geq 2$.

THÉORÈME 1.5. *Étant donné un entier $a>0$, il existe un nombre fini de couples $(m,k)\in\mathbb{Z}^+\times\mathbb{Z}_{\geq 3}$ telles que $-m+\zeta_k$ et $-(m+a)+\zeta_k$ soient multiplicativement dépendants. De plus les couples exceptionnels (m,k) sont effectivement calculables.*

Pour de petites valeurs de a nous avons déterminé ces paires exceptionnels et nous avons montré le théorème suivant.

THÉORÈME 1.6. *Soit $1\leq a\leq 10^6$ et supposons que $m\neq 0,-a$ et que $(m,k)\neq(1,6)$, $(-1,3)$, $(-a+1,6)$, $(-a-1,3)$. Alors $-m+\zeta_k$ et $-(m+a)+\zeta_k$ sont multiplicativement indépendant ou $m=-1$, $a=2$ et $k=4$.*

Après ces premiers résultats sur les bases dans l'anneau des entiers d'un corps cyclotomique, nous avons encore deux problèmes. D'une part la constante $C(k)$ dans le théorème est effectivement calculable mais gigantesque, et, d'autre part, nous avons seulement considéré les bases de la forme $-m+\zeta_k$. Soit β une base d'un système de numération dans $\mathbb{Z}[\zeta_k]$. Alors l'anneau des entiers $\mathbb{Z}[\zeta_k]$ est engendré par β. Györy [**100**] a montré que, dans tout corps de nombres, il existe seulement un nombre fini d'éléments primitifs non équivalents qui engendrent son anneau des entiers. Deux éléments primitifs α et β sont équivalents si $\alpha-\beta\in\mathbb{Z}$. Le théorème suivant caractérise les bases possibles.

THÉORÈME 1.7 ([**129**, Theorem 5]). *Soient $\alpha_1,\dots,\alpha_t\in\mathcal{O}$, $n_1,\dots,n_t\in\mathbb{Z}$ et $N_1,\dots,N_t$ des sous-ensembles de $\mathbb{Z}$ (tous effectivement calculables). Alors $\{\alpha,\mathcal{N}(\alpha)\}$ est un système de numération dans $\mathcal{O}$ si et seulement si $\alpha=\alpha_i-h$ pour deux entiers i,h tels que $1\leq i\leq t$ et soit $h\geq n_i$ soit $h\in N_i$.*

Nos premiers résultats ont montré que $n_i\leq\phi(k)$ pour $\beta_i=\zeta_k$. Dans le contexte des corps cyclotomiques, Bremner [**46**] et Robertson [**199**] conjecturent que les seuls éléments primitifs (à équivalence près) sont ζ_k, $\eta_k:=(1+\zeta_k)^{-1}$ et $\theta_k:=(1-\zeta_k)^{-1}$. Cette conjecture est vérifié pour $k=$ ... et avec les méthodes de Györy [**100**] on peut la résoudre pour n'importe quel k si on a suffisamment de temps.

Dans un troisième article, Paul Surer, Volker Ziegler et moi [Chapter 2] avons attaqué le calcul des n_i pour η_k et θ_k. Nous avons obtenu le théorème suivant.

THÉORÈME 1.8. *Soient $k\in\mathbb{N}$ (avec $k\geq 3$) et $a\in\mathbb{Q}$. Si $a\geq\varphi(k)+\frac{1}{2}$, alors $-a+\eta_k$ et $-a+\theta_k$ sont des bases d'un système de numération. Si $a\geq\varphi(k)-\frac{1}{2}$, alors $-a-\eta_k$ et $-a-\theta_k$ sont des bases d'un système de numération.*

En fait, on peut prendre $a \in \mathbb{Q}$ (cela a été démontré), mais pour cela il faudrait généraliser la définition des systèmes de numération alors que les autres résultats ne nécessitent pas une définition si générale.

De même, nous avons considéré l'indépendance pour deux bases. En réalité nous avions commencé avec ce résultat car nous avions déjà une estimation pour n_i ($n_i \leq k^2$) avec l'ancienne méthode..

THÉORÈME 1.9. *Soit ζ_k une racine du k-ième polynôme cyclotomique $\Phi_k(x)$ avec $k \notin \{1,2,3,4,6\}$ et $a, a' \in \mathbb{Q}$ avec $a' < a$ tels que*

$$|a + \zeta_k|^p = \left|a' + \zeta_k\right|^{p'}$$

pour certains entiers $p, p' > 0$. Alors une des conditions suivantes est satisfaite :

(i) $a' < a < -\delta_2$ *et* $aa' < 1$;

(ii) $-\delta_2 < a' < a < 0$ *et* $a + a' > -\delta_2$;

(iii) $0 < a' < a < -\delta_1$ *et* $a + a' < -\delta_1$;

(iv) $-\delta_1 < a' < a$ *et* $aa' < 1$,

où

$$\delta_1 := \begin{cases} 2\cos\left(\frac{(k-1)\pi}{k}\right) & \text{si } k \equiv 1 \pmod 2 \\ 2\cos\left(\frac{(k-2)\pi}{k}\right) & \text{si } k \equiv 0 \pmod 4 \\ 2\cos\left(\frac{(k-4)\pi}{k}\right) & \text{si } k \equiv 2 \pmod 4 \end{cases} < 0,$$

$$\delta_2 := 2\cos\left(\tfrac{2\pi}{k}\right) > 0.$$

Dans le Chapitre 2, nous présentons ces nouvelles idées et démontrons les deux derniers théorèmes. De plus, nous présentons une variante du théorème de Cobham comme application. Pour bien comprendre cette application il faut se plonger dans le domaine des ensembles reconnaissables. Un sous-ensemble $S \subset \mathbb{N}$ est dit q-reconnaissable si l'ensemble des mots sur l'alphabet $\{0, 1, \ldots, q-1\}$ qui représentent les éléments de S en base q est une partie reconnaissable par un automate du monoïde libre $\{0, 1, \ldots, q-1\}^*$. En 1969, Cobham [**59**] a démontré le théorème fondamental suivant.

THÉORÈME 1.10. *Soient p et q deux entiers supérieurs à 1 multiplicativement indépendants et soit S un sous-ensemble de $\mathbb{N}$. L'ensemble S est à la fois p-reconnaissable et q-reconnaissable si et seulement si S est une réunion d'un ensemble fini et d'un nombre fini de progressions arithmétiques.*

D'autres travaux ont simplifié la démonstration initiale, considéré le cas des dimensions multiples et généralisé ce résultat dans des systèmes de numération différents. Pour une vue globale des travaux effectués, nous renvoyons le lecteur à l'article de Durand et Rigo [**72**].

Dans ce mémoire, nous nous concentrons sur les généralisations à des systèmes de numération dans les corps cyclotomiques. Soient $k \geq 2$ un entier et ζ_k une k-ième racine d'unité. Alors pour deux bases $\alpha = -p + \zeta_k$ et $\beta = -q + \zeta_k$ multiplicativement indépendantes, nous faisons la conjecture suivante : une partie S est α-reconnaissable et β-reconnaissable si et seulement si c'est une partie périodique de $\mathbb{Z}[\zeta_k]$ à un ensemble fini près. Cette conjecture est une généralisation de celle de Hansel et Safer pour les entiers de Gauss ($k = 4$). Allouche *et al.* [**11**] ont formulé cette conjecture dans le cas spécial $\alpha = -1 + i$ et $\beta = -q + i$.

Une direction de la démonstration de l'énoncé est très simple : une partie périodique de $\mathbb{Z}[\zeta_k]$ est α-reconnaissable dans toute base $\alpha = -m + \zeta_k$. Les difficultés apparaissent si nous démontrons l'autre direction. Pour deux bases $\alpha = -p + i$ et $\beta = -q + i$ dans les entiers de Gauss, il est facile de démontrer qu'elles sont multiplicativement indépendantes. Dans le cas général, nous utilisons le Théorème 1.9.

Toutes les démonstrations du théorème de Cobham utilisent une étape intermédiaire en montrant que toute partie finie et p- et q-reconnaissable de $\mathbb{N}$ est une partie syndétique de $\mathbb{N}$. Une partie $S \subset \mathbb{N}$ est dite syndétique s'il existe $r \in \mathbb{N}$, tel que $S \cap [n, n + r] \neq \emptyset$ pour tout $n \in \mathbb{N}$. L'équivalent dans les entiers d'un corps cyclotomique (sous-ensemble de $\mathbb{C}$) est l'existence d'un réel $r > 0$ tel que pour tout $x \in \mathbb{C}$ l'intersection $S \cap B(x, r) \neq \emptyset$ où $B(x, r)$ est la boule de centre x et de rayon r dans $\mathbb{C}$.

Comme Hansel et Safer [**101**], nous ne pouvons montrer cette première étape que sous réserve de la validité de la conjecture des quatre exponentielles.

CONJECTURE 1.11. *Soient $\{\lambda_1, \lambda_2\}$ et $\{x_1, x_2\}$ deux paires de nombres complexes formées chacune de deux nombres rationnellement indépendants). Alors, l'un des quatre nombres*

$$e^{\lambda_1 x_1}, \quad e^{\lambda_1 x_2}, \quad e^{\lambda_2 x_1} \quad ou \quad e^{\lambda_2 x_2}$$

est transcendant.

La raison de cette restriction est la suivante. Dans une partie de la démonstration, il faut passer de l'écriture en base α à l'écriture en base β. Pour ce changement, il faut que l'ensemble des rapports α^m/β^n soit dense dans $\mathbb{C}$. Une manière d'établir cette densité est d'utiliser la conjecture des quatre exponentielles.

Alors la conjecture implique la densité. L'inverse n'est pas clair mais semble vrai. Nous conjecturons que, si α^m/β^n est dense, alors la conjecture de quatre exponentielles est vraie.

Par contre l'équivalent pour six exponentielles (le théorème des six exponentielles) est démontré indépendamment par Serge Lang [**132**] et Kanakanahalli Ramachandra [**192**] dans les années 1960.

THÉORÈME 1.12. *Soient x_1, x_2, x_3 trois complexes $\mathbb{Q}$-linéairement indépendants et λ_1, λ_2 deux complexes également $\mathbb{Q}$-linéairement indépendants. Alors, l'un au moins des six nombres*

$$e^{\lambda_1 x_1}, \quad e^{\lambda_1 x_2}, \quad e^{\lambda_1 x_3}, \quad e^{\lambda_2 x_1}, \quad e^{\lambda_2 x_2} \quad ou \quad e^{\lambda_2 x_3}$$

est transcendant.

Cela donne une version sans restriction avec trois bases. Mais nous n'avons pas réussi à énoncer proprement ce que veut dire « avec trois bases ».

Dans le Chaptire 2, nous allons montrer la généralisation suivante du résultat de Hansel et Safer [**101**].

THÉORÈME 1.13. *Supposons vraie la conjecture des quatre exponentielles. Soit $k \geq 2$, soient α et β deux bases dans $\mathbb{Z}[\zeta_k]$ multiplicativement indépendantes et soit S un sous-ensemble de $\mathbb{Z}[\zeta_k]$. Si l'ensemble S est à la fois α-reconnaissable et β-reconnaissable, alors S est syndétique.*

3. Longueur des partitions d'un entier en entiers ellipséphiques

Après la recherche des bases possibles nous focalisons notre attention sur les chiffres d'une représentation. Le premier résultat concerne les partitions d'un entier en entiers ayant certaines restrictions sur les chiffres. En général, une partition d'un entier est une décomposition

de cet entier en une somme d'entiers strictement positifs. Le nombre de partitions de l'entier n est classiquement noté $p(n)$. Nous posons $p(n) = 0$ si $n < 0$ et $p(0) = 1$, puisque 0 possède exactement une partition : la somme vide. La suite $(p(n))_{n\geq 0}$ est déterminée par une fonction récursive. Hardy et Ramanujan [**103**] ont donné un développement asymptotique pour $p(n)$. Leur considération est aussi à l'origine de la méthode du cercle ultérieurement développée en détail par Hardy et Littlewood (voir [**245**] ou [**169**]). Vinogradov a par la suite étendu la technique en remplaçant la série génératrice par une somme trigonométrique (*cf.* [**246**]). Rademacher [**190**] a amélioré le résultat de Hardy et Ramanujan [**103**] en obtenant une formule explicite. En particulier, il a donné une série convergente pour $p(n)$.

Nous présentons des méthodes modernes pour démontrer une formule asymptotique pour $p(n)$. Soit P la série génératrice de $p(n)$. Alors

$$P(z) = \sum_{n=1}^{\infty} p(n) z^n = \prod_{n=1}^{\infty} \frac{1}{1-z^n}. \tag{3.1}$$

Posons $z = e^{-t}$ et supposons que $t > 0$. La fonction

$$L(t) := \log P\left(e^{-t}\right) = \sum_{n\geq 1} \frac{e^{-nt}}{n\left(1-e^{-nt}\right)}$$

est une somme harmonique. Nous pouvons donc appliquer la transformation de Mellin. Nous renvoyons le lecteur au survol de Flajolet, Gourdon et Dumas [**79**]. La transformée de $e^{-t}/\left(1-e^{-t}\right)$ est $\Gamma(s)\zeta(s)$ et la série de Dirichlet associée est $\sum n^{-1}n^{-s} = \zeta(s+1)$. Au total on obtient

$$L^*(s) = \zeta(s)\zeta(s+1)\Gamma(s).$$

On en déduit que L^* est méromorphe avec un pôle simple en $s = 1$, un pôle double en $s = 0$ et des pôles simple en $s = -m$ pour $m \in \mathbb{N}^*$. Pour l'analyse, il suffit de regarder le développement des singularités dans $\operatorname{Re} s > -2$. Cela nous donne

$$L^*(s) \asymp \frac{\pi^2}{6}\frac{1}{s-1} - \frac{1}{2}\frac{1}{s^2} - \frac{\log\sqrt{2\pi}}{s} - \frac{1}{24}\frac{1}{s+1}.$$

Avec la transformation inverse on obtient (*cf.* [**79**, Figure 4])

$$L(t) = \frac{\pi^2}{6t} + \frac{1}{2}\log t - \log\sqrt{2\pi} - \frac{1}{24}t + \mathcal{O}(t^2), \quad t \to 0^+.$$

En remplaçant dans P nous avons

$$\log P(z) = \frac{\pi^2}{6}\frac{1}{1-z} + \frac{1}{2}\log(1-z) - \frac{\pi^2}{12} - \log\sqrt{2\pi} + \mathcal{O}(1-z).$$

Pour une analyse complète nous avons besoin des estimations de P dans le domaine non central. Considérons la représentation en produit de (3.1). La moitié des facteurs est infinie pour $z = -1$, un tiers l'est pour $z = \exp(\pm 2\pi i/3)$, et ainsi de suite. Il est possible d'élargir la méthode de la transformation de Mellin au cas $z = e^{-t-i\phi}$ pour $t \to 0$ et $\phi = 2\pi\frac{p}{q}$. Dans ce cas, il faut utiliser avec les sommes

$$L_\phi(t) = \sum_{m\geq 1}\frac{1}{m}\frac{e^{-m(t+i\phi)}}{1-e^{-m(t+i\phi)}} = \sum_{m\geq 1}\sum_{k\geq 1}\frac{1}{m}e^{-mk(t+i\phi)}.$$

Le résultat final est le suivant : si z tend vers $e^{2\pi i \frac{p}{q}}$, alors $P(z)$ se comporte comme

$$\exp\left(\frac{\pi^2}{6q^2(1-|z|)}\right),$$

qui est une puissance $1/t^2$ du comportement asymptotique de P pour $z \to 1^-$. L'analyse se poursuit sur les petits arcs. Enfin, nous considérons un recouvrement du cercle par des arcs dont l'argument du centre est $2\pi j/N$, $j = 1, \ldots, N-1$, avec N suffisamment grand. Cette estimation est une amélioration des idées de Hardy et Ramanujan [**103**] par Rademacher [**190**] mentionnée ci-dessus. Pour plus de détails, le lecteur intéressé peut consulter l'ouvrage de Flajolet et Sedgewick [**81**].

Les deux estimations – au centre et en dehors du centre – nous donnent pour le nombre p_n des partitions d'un entier n en entiers positifs

$$p_n \equiv [z^n] \prod_{k=1}^{\infty} \frac{1}{1-z^k} \sim \frac{1}{4n\sqrt{3}} \exp\left(\pi\sqrt{2n/3}\right).$$

Le principe de cette démonstration est généralisé par Meinardus [**161**], qui ramène les étapes de la démonstration ci-dessus aux trois conditions à vérifier.

THÉORÈME 1.14 ([**161**, Satz 1]). *On suppose les trois conditions suivantes remplies.*

(1) Le produit

$$f(\tau) = \prod_{k=1}^{\infty} (1-e^{-k\tau})^{-\lambda_k}$$

est convergent pour $\mathrm{Re}(\tau) > 0$ *et on écrit*

$$f(\tau) = 1 + \sum_{k=1}^{\infty} r(k) e^{-k\tau}$$

son développement où les puissances λ_k *sont des réels positifs.*

(2) La série de Dirichlet $D(s) = \sum_{k\geq 1} \lambda_k^{-s}$ *est convergente dans le demi-plan* $\mathrm{Re}\, s > \alpha > 0$ *et prolongeable analytiquement jusqu'à la droite* $\mathrm{Re}\, s = -c_0$ *avec* $0 < c_0 < 1$. *Pour* $\mathrm{Re}(s) \geq -c_0$, *la fonction* $D(s)$ *est holomorphe sauf pour* $s = \alpha$ *où elle a un pôle simple avec résidu* A. *De plus, il existe une constante* c_1 *telle que*

$$D(\sigma + it) \ll |t|^{c_1}$$

pour $|t| \to \infty$.

(3) Soit $\tau = y + 2\pi i x$ *la décomposition de* τ *en partie réelle et imaginaire,* $y > 0$ *et*

$$g(\tau) = \sum_{k\geq 1} a_k exp\,(-k\tau).$$

Pour $|\arg \tau| > \frac{\pi}{4}$ *et* $|x| \leq \frac{1}{2}$ *on a*

$$\mathrm{Re}\, g(\tau) - g(y) \geq -c_2 y^{-\varepsilon}$$

pour y *suffisamment petit,* ε *arbitraire positive et* $c_2 > 0$ *appropriée.*

Alors

$$r(n) = C \cdot n^{\kappa} \cdot \exp\left(n^{\frac{\alpha}{\alpha+1}} \left(1+\frac{1}{\alpha}\right) (A\Gamma(\alpha+1)\zeta(\alpha+1))^{\frac{1}{\alpha+1}}\right) \cdot \left(1+\mathcal{O}\left(n^{-\kappa_1}\right)\right).$$

Ces trois conditions sont adaptées et transformées dans les applications. Par exemple, Hwang [**109**] a donné des conditions plus faibles que celles-ci de Meinardus.

L'application sur laquelle nous nous focalisons concerne les partitions d'un entier en entiers ayant des chiffres manquant parfois appelés entiers ellipséphiques. En particulier, soient $b \geq 2$ un entier et $\mathcal{D} \subset \{0, 1, \ldots, b-1\}$ tel que $0 \in \mathcal{D}$ et $2 \leq |\mathcal{D}| \leq b-1$. Alors l'ensemble des entiers ellipséphiques $\mathcal{MD}(b, \mathcal{D}) \subset \mathbb{N}$ contient tout entier positif n tel que

$$n = \sum_{k=0}^{\ell} a_k b^k \quad \text{avec } a_k \in \mathcal{D}.$$

Dans le Chapitre 3 nous considérons la distribution des longueurs des partitions d'un entier en entiers ellipséphiques. Ces entiers sont en dehors du domaine d'application des résultats de Hwang [**109**] parce que leur série génératrice de Dirichlet a des pôles équidistants sur la droite $\operatorname{Re} z = \alpha$. En effet, soit $D(s)$ la série génératrice de Dirichlet définie par

$$D(s) = \sum_{m \in \mathcal{MD}(b,D)} m^{-s}.$$

En notant que

$$\mathcal{MD}(b, D) = \{bn_0 + a_0 \,|\, n_0 \in \mathcal{MD}(b, D) \cup \{0\},\, a_0 \in D\} \setminus \{0\}$$

nous obtenons

$$\begin{aligned} R(s) &= \left(1 - |D| b^{-s}\right) D(s) \\ &= \sum_{n \in \mathcal{MD}(b,D)} \sum_{a \in D} \left(\frac{1}{(bn+a)^s} - \frac{1}{(bn)^s} \right) + \sum_{a \in D \setminus \{0\}} \frac{1}{a^s}, \end{aligned}$$

où $R(s)$ est une série de Dirichlet holomorphe sur le demi-plan. Les pôles de $D(s)$ sont absorbés par les racines de $1 - |D|\, b^{-s}$ qui sont de la forme $s = \frac{\log|D|}{\log b} + \frac{2k\pi}{\log b} i$ avec $k \in \mathbb{Z}$.

Conjointement avec Stephan Wagner nous étendons la méthode aux partitions en entiers de Λ la série génératrice de Dirichlet a des pôles équidistants sur une droite verticale. Plus précisément, soit $\Lambda = (\Lambda_1, \Lambda_2, \ldots)$ une suite croissante d'entiers avec $\Lambda_k \to \infty$. Alors une partition de n en entiers de Λ est une somme de la forme

$$\sum_{j=1}^{s} \Lambda_{i_j} = n$$

avec $i_1 < i_2 < \ldots < i_s$. Pour la longueur s des partitions nous montrons le théorème central limite suivant.

THÉORÈME 1.15. *Supposons que la suite Λ satisfait les conditions suivantes.*

(M1) La série de Dirichlet $D(s) = \sum_{k \geq 1} \Lambda_k^{-s}$ converge dans le demi-plan $\operatorname{Re} s > \alpha > 0$ et on peut la prolonger analytiquement pour $\operatorname{Re} s \geq \alpha - \varepsilon$ avec $\varepsilon > 0$. Sur la droite $\operatorname{Re} s = \alpha$, $D(s)$ a des pôles équidistants (dont la distance est notée ω) et simples en $s = \alpha + 2\pi i \omega$ pour $k \in \mathbb{Z}$; A_k est le résidu de $D(s)$ en $s = \alpha + 2\pi i k \omega$. De plus nous supposons qu'il n'existe pas d'autre pôles pour $\operatorname{Re} s \geq \alpha - \varepsilon$.

(M2) Il existe une suite $T_j \to \infty$ et une constante positive c_1 telles que

$$D(s) \ll |T_j|^{c_1}$$

uniformément pour tout $s \in \{z \in \mathbb{C}\colon \alpha - \varepsilon \leq \mathrm{Re} z \leq \alpha\}$ *avec* $|\mathrm{Im}\, s| = T_j$. *De plus nous supposons que* D *satisfait*

$$D(\alpha - \varepsilon + it) \ll |t|^{c_1}.$$

(M3) Soit $g(\tau) = \sum_{k\geq 1} exp\,(-\lambda_k \tau)$ *où* $\tau = r + iy$ *avec* $r > 0$ *et* $-\pi \leq y \leq \pi$. *Il existe une constante positive* c_2 *telle que*

$$g(r) - \mathrm{Re} g(\tau) \geq c_2 \left(\log \frac{1}{r}\right)^{2+4/\alpha}$$

uniformément pour $\pi/2 \leq |y| \leq \pi$ *quand* $r \to 0^+$.

Soit ϖ *le nombre des termes d'une partition aléatoire. Alors* ϖ *suit la loi normale d'espérance* $\mathbb{E}(\varpi) \sim \mu_n$ *et de variance* $\mathbb{V}(\varpi) \sim \sigma_n^2$ *:*

$$\mathbb{P}\left(\frac{\varpi - \mu_n}{\sigma_n} < x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2} \mathrm{d}t + o(1),$$

uniformément pour tout x *quand* $n \to \infty$ *;* μ_n *et* σ_n *sont définis par :*

$$\mu_n = \sum_{k\geq 1} \frac{1}{e^{\eta \Lambda_k} + 1},$$

$$\sigma_n^2 = \sum_{k\geq 1} \frac{e^{\eta\Lambda_k}}{(e^{\eta\Lambda_k}+1)^2} - \frac{\left(\sum_{k\geq 1} \frac{\Lambda_k e^{\eta\Lambda_k}}{(e^{\eta\Lambda_k}+1)^2}\right)^2}{\sum_{k\geq 1} \frac{\Lambda_k^2 e^{\eta\Lambda_k}}{(e^{\eta\Lambda_k}+1)^2}},$$

et η *est définie implicitement par*

$$n = \sum_{k\geq 1} \frac{\Lambda_k}{e^{\eta\Lambda_k}+1};$$

μ_n *et* σ_n *satisfont les formules asymptotiques suivantes :*

$$\mu_n \sim n^{\alpha/(1+\alpha)} \Psi_\mu \left(\frac{\omega \log n}{\alpha+1}\right),$$

$$\sigma_n^2 \sim n^{\alpha/(1+\alpha)} \Psi_\sigma \left(\frac{\omega \log n}{\alpha+1}\right),$$

pour certaine fonctions Ψ_μ *et* Ψ_σ *de période* 1.

L'outil central est la méthode du col conjointement avec la transformation de Mellin appliquée à la série génératrice de Dirichlet.

Ce travail est la base de la thése de Ralaivaosaona [**191**]. Il a étendu nos conditions $M1$ à $M3$ en considérant le cas des partitions en nombres premiers. La difficulté est le pôle essentiel de la série génératrice.

4. Distribution des chiffres et fonctions q-additives

Pour analyser la distribution des chiffres on utilise très souvent des fonctions qui opèrent seulement sur les chiffres d'une représentation. Regardons les systèmes dans $\mathbb{N}$. Soit $(q, \mathcal{N})$

avec $q \geq 2$ et $\mathcal{N} = \{0, 1, \ldots, q-1\}$ une base. Alors une fonction $f\colon \mathbb{N} \to \mathbb{C}$ est appelée q-additive si

$$f(n) = \sum_{k=0}^{\ell} f(a_k q^k) \quad \text{pour } n = \sum_{k=0}^{\ell} a_k q^k \text{ avec } a_k \in \mathcal{N}.$$

L'exemple le plus simple est la somme des chiffres en base q, notée s_q, qui est définie par

$$s_q(n) = \sum_{k=0}^{\ell} a_k \quad \text{pour } n = \sum_{k=0}^{\ell} a_k q^k \text{ et } a_k \in \mathcal{N}.$$

On peut très simplement généraliser ce concept aux autres systèmes de numération mentionnés au début de l'introduction.

Dans cette partie nous nous concentrons sur des anneaux quotients sur l'anneau des polynômes sur $\mathbb{Z}[X]$. Soit $p \in \mathbb{Z}[X]$. Pethő [**180**] a considéré des systèmes de numération dans $\mathcal{R} := \mathbb{Z}[X]/(p(X))$. Il a obtenu des conditions suffisantes pour que $(p, \{0, 1, \ldots, |p(0)| - 1\})$ soit un système de numération dans $\mathcal{R}$: tout $n \in \mathcal{R}$ a une représentation unique de la forme

$$n = \sum_{k=0}^{\ell} a_k X^k \quad \text{avec } a_k \in \{0, 1, \ldots, |p(0)| - 1\}.$$

En remplaçant X par un entier rationnel ou algébrique on retrouve respectivement les systèmes de numération dans $\mathbb{Z}$ et l'anneau des entiers d'un corps de nombres.

Nous commençons par la distribution en classes de résidus de la somme des chiffres sur des progressions arithmétiques. Gelfond [**91**] a montré que l'ensemble

$$S_{h,m}(N) := \{n \leq N\colon s_q(n) \equiv h \mod m\}$$

est équiréparti dans les progressions arithmétiques.

THÉORÈME 1.16. *Soient q, h, a, p et m des entiers telle que $q > 1$, $0 \leq a < p$, $0 \leq h < m$, et $(p, q-1) = 1$. Alors le nombre $T_a(x)$ des entiers $n \leq x$ satisfaisant les conditions*

$$n = \sum_{k=0}^{\ell} a_k q^k \equiv h \bmod m; \quad s_q(n) = \sum_{k=0}^{\ell} a_k \equiv a \bmod p$$

est donné par la formule

$$T_a(x) = \frac{x}{mp} + \mathcal{O}(x)^{\lambda},$$

où $\lambda < 1$ ne dépend pas de x, m, h et a.

L'origine de la condition $(p, q-1) = 1$ est la preuve par neuf qui se généralise à la congruence

$$n = \sum_{k=0}^{\ell} a_k q^k \equiv \sum_{k=0}^{\ell} a_k = s_q(n) \bmod q - 1.$$

Un résultat similaire pour des sommes d'ensembles est établi par Mauduit et Sárközy [**157**] qui ont montré pour deux ensembles $\mathcal{A}, \mathcal{B} \subset \{1, \ldots, N\}$ que

$$\left| \#\{(a, b) \in \mathcal{A} \times \mathcal{B}\colon s_q(a+b) \equiv h \bmod m\} - \frac{|\mathcal{A}|\,|\mathcal{B}|}{m} \right| \ll N^{\theta} \left(|\mathcal{A}|\,|\mathcal{B}|\right)^{\frac{1}{2}}$$

où $\theta < 1$ et la constante implicite sont absolus. Thuswaldner [**239**] a transféré ces résultats aux systèmes de numération dans l'anneau des entiers d'un corps de nombres.

Pour des systèmes dans $\mathcal{R}$, je peux montrer les résultats suivants.

THÉORÈME 1.17. *Soit $(p,\mathcal{N})$ un système de numération. Pour tout idéal $\mathfrak{s}$ de $\mathcal{R}$ on note $\mathcal{V}_p(\mathcal{R}(T))$, le nombre d'éléments de $\mathcal{U}_{h,m}(\mathcal{R}(T))$ tels que*

$$z \equiv a \mod \mathfrak{s}.$$

Si $(p(1),m)=1$, alors

$$\mathcal{V}_p(\mathcal{R}(T)) = \frac{|\mathcal{U}_{h,m}(\mathcal{R}(T))|}{m\mathrm{N}(\mathfrak{s})} + \mathcal{O}\left(|\mathcal{U}_{h,m}(\mathcal{R}(T))|^{\lambda}\right), \quad (\lambda<1),$$

où λ ne dépend pas de T, h, a et $\mathfrak{s}$.

De même, je généralise le résultat de Mauduit et Sárközy [**157**].

THÉORÈME 1.18. *Soit $(p,\mathcal{N})$ un système de numération. Si $(p(1),m)=1$, alors pour deux sous-ensembles $\mathcal{A},\mathcal{B}\subset\mathcal{R}(T)$ on a*

$$\left||\{(x,y)\in\mathcal{A}\times\mathcal{B} : x+y\in\mathcal{U}_{h,m}(\mathcal{R}(T))\}| - \frac{|\mathcal{A}|\,|\mathcal{B}|}{m}\right| \ll |\mathcal{R}(\mathcal{T})|^{\mu}\left(|\mathcal{A}|\,|\mathcal{B}|\right)^{\frac{1}{2}},$$

où la constante implicite est absolue et $\mu<1$.

Comme les chiffres sont des entiers toute représentation est commune à tous les conjugués. Donc la longueur de représentation croît par rapport aux valeurs absolues des conjugués de la base, et non par rapport à la norme de la base. Cette interaction entre les deux « mesures » pose des difficultés dans la démonstration.

Après la distribution en classes de résidus, nous considérons la moyenne arithmétique de la somme des chiffres. Pour les systèmes q-adiques Delange [**64**] a montré le résultat suivant.

THÉORÈME 1.19. *Il existe une fonction Φ continue sur $\mathbb{R}$ et périodique de période 1, telle que, pour tout entier $x\geq 1$,*

$$\frac{1}{x}\sum_{n\leq x} s_q(n) = \frac{q-1}{2\log q}\log x + \Phi\left(\frac{\log m}{\log q}\right).$$

Le point essentiel est que la formule est exacte – le terme d'erreur est caché dans la fonction Φ dont Delange a calculé les coefficients de Fourier.

Ce résultat a été généralisé dans plusieurs articles dans plusieurs directions différentes. Grabner, Kirschenhofer, Prodinger et Tichy [**96**] estiment le moment d'ordre d de la somme des chiffres. Dans le cadre des systèmes de numération dans l'anneau des entiers d'un corps de nombres, Thuswaldner [**238**] a généralisé le résultat de Delange [**64**] et Gittenberger et Thuswaldner [**93**] ont estimé le moment d'ordre d.

Avec Pethő [**152**], nous pouvons estimer le moment d'ordre d pour des fonctions additives dans des systèmes de numération d'anneau quotient.

THÉORÈME 1.20. *Soient $(p,\mathcal{N})$ un système de numeration et $M=|p(0)|$. De plus soit f une fonction additive dans $(p,\mathcal{N})$ et μ_f la moyenne arithmétique des valeurs de f, i.e.,*

$$\mu_f := \frac{1}{|\mathcal{N}|}\sum_{a\in\mathcal{N}} f(a).$$

Si on pose

$$N = M^{\ell}\prod_{i=1}^{r}\prod_{k=1}^{s_i+t_i}(x_{ik}(x))^{m_i},$$

alors on obtient

$$\frac{1}{N}\sum_{z\in\mathcal{M}(p,\ell,x)}(f(z))^d = c(p)\mu_f^d \log_M^d(N) + \sum_{j=0}^{d-1}\log_M^j(N)\Phi_j(\log_M N) + \mathcal{O}\left(N^{-\frac{1}{n}}\log_M^d N\right),$$

où $c(p)$ est une constante qui dépend seulement de p et $\Phi_0,\dots,\Phi_{d-1}$ sont des fonctions continues périodiques de période 1.

Le terme d'erreur vient des estimations du domaine fondamental. En particulier, la structure factorielle de ce domaine perturbe le calcul des coefficients de Fourier de la fonction indicatrice. A part ce terme d'erreur, ce résultat se ramène aux résultats mentionnés ci-dessus.

Après les moments nous nous intéressons à la démonstration d'un théorème central limite pour des fonctions q-additives. Cette démonstration repose sur un théorème montré à l'origine par Bassily et Kátai [**19**] pour des fonctions q-additives dans les systèmes d'entiers positifs. Plusieurs généralisations ont été effectuées par Gittenberger et Thuswaldner [**94**] (cas des entiers de Gauss) et moi-même [**143**] (cas des corps de nombres). Avec Pethő [**151**], nous investissions la distribution des chiffres pour les systèmes de numération dans l'anneau quotient $\mathcal{R} := \mathbb{Z}[X]/(p(X))$. Dans le Chapitre 4 nous montrons le résultat suivant.

THÉORÈME 1.21. *Soient $(p,\mathcal{N})$ un système de numération dans $\mathcal{R}$ et f une fonction additive telle que $f(aX^h) = \mathcal{O}(1)$ pour $h\to\infty$ et $a\in\mathcal{N}$. De plus, soit*

$$m_h := \frac{1}{|\mathcal{N}|}\sum_{a\in\mathcal{N}} f(aX^h), \quad \sigma_h^2 := \frac{1}{|\mathcal{N}|}\sum_{a\in\mathcal{N}} f^2(aX^h) - m_h^2,$$

et

$$M(x) := \sum_{h=0}^{L} m_h, \quad D^2(x) := \sum_{h=0}^{L}\sigma_h^2,$$

où $L = \lfloor \log_{p(0)} x\rfloor$. Supposons qu'il existe $\varepsilon>0$ tel que $D(x)/(\log x)^\varepsilon\to\infty$ pour $x\to\infty$ et soit $P\in\overline{\mathcal{K}}[Y]$ un polynôme de degré d. Alors, pour $T\to\infty$, on a

$$\frac{1}{\#\mathcal{R}(T)}\#\left\{z\in\mathcal{R}(T) : \frac{f(\lfloor P(z)\rfloor) - M(T^d)}{D(T^d)} < y\right\} \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y}\exp(-x^2)\mathrm{d}x.$$

L'idée centrale de la démonstration est d'utiliser le théorème de Fréchet et Shohat (*cf.* [**74**, Lemma 1.43]) : soit F_n une suite de fonctions de répartition ; si pour tout n les moments de tout ordre de F_n existent et déterminent la fonction de répartition limite F uniquement, alors F_n converge vers F. Cette idée apparaît déjà dans la démonstration de Bassily et Kátai [**19**] et on appelle donc parfois cette méthode la méthode de Bassily et Kátai.

Pour terminer cette section, nous considérons une conjecture de Stolarsky. Dans un article [**236**], Stolarsky a analysé les ratios $s_q(n^h)/s_q(n)$ pour des entiers $h,q\geq 2$. Il a montré que

$$\limsup_{n\to\infty}\frac{s_q(n^h)}{s_q(n)} = +\infty.$$

Dans le même article, il a conjecturé que

$$\liminf_{n\to\infty}\frac{s_q(n^h)}{s_q(n)} = 0$$

et que

$$\lim_{T\to\infty}\sum_{n=1}^{T}\frac{s_q(n^h)}{s_q(n)}=h'$$

avec $1<h'\leq h$. La première conjecture est démontrée par Hare, Laishram et Stoll [**104**]. Avec Thomas Stoll, nous avons démontré la deuxième conjecture.

THÉORÈME 1.22. *Soient $q_1,q_2\geq 2$ des entiers et $P_1(X),P_2(X)\in\mathbb{C}[X]$ des polynômes de degrés $r_1,r_2\geq 1$, respectivement, avec $P_1(\mathbb{N}),P_2(\mathbb{N})\subset\mathbb{N}$. Alors*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\leq N}\frac{s_{q_1}(P_1(n))}{s_{q_2}(P_2(n))}=\frac{q_1-1}{q_2-1}\cdot\left(\frac{\log q_1}{\log q_2}\right)^{-1}\cdot\frac{r_1}{r_2}.$$

De plus,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\leq N}\frac{s_{q_1}(P_1(p_n))}{s_{q_2}(P_2(p_n))}=\frac{q_1-1}{q_2-1}\cdot\left(\frac{\log q_1}{\log q_2}\right)^{-1}\cdot\frac{r_1}{r_2}$$

et

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\leq N}\frac{s_{q_1}(P_1(p_n))}{s_{q_2}(P_2(n))}=\frac{q_1-1}{q_2-1}\cdot\left(\frac{\log q_1}{\log q_2}\right)^{-1}\cdot\frac{r_1}{r_2}$$

où $(p_n)_{n\geq 1}$ est la suite des nombres premiers.

Pour la démonstration, nous regroupons les éléments dont la somme de chiffres est loin de la moyenne :

$$A=\left\{1\leq n\leq N\big|\,\big|s_q(P(n))-\mu_q r\log_q N\big|>\sigma_q(r\log_q N)^{\frac{3}{4}}\right\}.$$

En utilisant le théorème de Bassily et Kátai [**19**], nous montrons que l'ensemble A est négligeable. Le reste donne les estimations demandées. L'argumentation est très élémentaire. Les variantes du théorème de Bassily et Kátai nous permettent de généraliser ce résultat aux autres systèmes de numération.

5. Équirépartition et nombres normaux

Maintenant nous nous focalisons sur le concept de l'équirépartition dont la connexion avec les systèmes de numération est établie par des nombres normaux considérés ci-dessous. Soit $\mathbb{T}:=\mathbb{R}/\mathbb{Z}$ le tore. On dit qu'une suite $(u_n)_{n\geq 1}$ d'éléments de $\mathbb{T}$ est équirépartie modulo 1 si pour tout $\alpha\in[0,1)$ la fréquence d'apparition des éléments de la suite dans l'intervalle $[0,\alpha]$ tend vers α, c'est-à-dire

$$\frac{1}{N}\#\left\{n\leq N\colon u_n\in[0,\alpha)\right\}\xrightarrow[N\to\infty]{}\alpha.$$

Weyl [**252**] a démontré le critère suivant qui connecte l'équirépartition avec les séries de Fourier sur $[0,1]$.

LEMME 1.23 (Critère de Weyl). *Une suite $(u_n)_{n\geq 1}$ est équirépartie modulo 1 si et seulement si pour tout $h\in\mathbb{Z}\setminus\{0\}$ on a que*

$$\frac{1}{N}\sum_{n\leq N}e\left(hu_n\right)\xrightarrow[N\to\infty]{}0$$

où $e(x)=\exp(2\pi ix)$.

Dans le même article Weyl [**252**] applique son critère pour montrer que la suite $(f(n))_{n\geq 1}$ avec $f \in \mathbb{R}[X]$ est équirépartie modulo 1 si et seulement si $f(X) - f(0)$ a au moins un coefficient irrationnel. Cette méthode est étendue aux suites plus compliquées. Conjointement avec Vitaly Bergelson, Grigori Kolesnik, Son Younghwan et Robert Tichy [**30**], nous avons appliqué cette méthode aux suites pseudo-polynômiales.

THÉORÈME 1.24. *Soit $\xi(x) = \sum_{j=1}^{r} \alpha_j x^{\theta_j}$, où $0 < \theta_1 < \theta_2 < \cdots < \theta_r$, α_j sont des réels non nuls et supposons que si tous les $\theta_j \in \mathbb{Z}^+$, alors au moins un α_j est irrationnel. Alors la suite $(\xi(p))_{p\in\mathcal{P}}$ est équirépartie modulo* 1.

D'après le critère de Weyl, il suffit de montrer pour tout $h \in \mathbb{Z}^*$

$$\sum_{p\leq N} e\left(h\xi(p)\right) \xrightarrow[N\to\infty]{} 0,$$

où la somme est sur les nombres premiers $p \leq N$. Notons que, si ξ est un polynôme ($\theta_j \in \mathbb{Z}$ pour $1 \leq j \leq r$), alors cela se réduit à un résultat de Rhin [**195**]. Nous supposons qu'il existe $1 \leq j \leq r$ tel que $\theta_j \notin \mathbb{Z}$.

En appliquant la sommation par parties, la somme d'exponentielles se réduit à

$$\sum_{p\leq N} e\left(h\xi(p)\right) = \sum_{n\leq N} \Lambda(n) e\left(h\xi(n)\right) + \mathcal{O}(\sqrt{N}),$$

où Λ est la fonction de von Mangoldt

$$\Lambda(n) = \begin{cases} \ln p & \text{si } n = p^k \text{ pour un nombre premier } p \text{ et un entier } k \geq 1, \\ 0 & \text{sinon.} \end{cases}$$

L'identité de Vaughan [**244**] (utilisée dans la version de Heath-Brown [**107**]) nous donne que

$$\left|\sum_n \Lambda(n) e(h\xi(n))\right| \ll K \log N + F_0 + L(\log N)^8,$$

où

$$K = \max_M \sum_{m=1}^{\infty} d_3(m) \left|\sum_{Z<n\leq M} e(h\xi(mn))\right|,$$

$$L = \sup \sum_{m=1}^{\infty} d_4(m) \left|\sum_{U<n<V} b(n) e(h\xi(mn))\right|,$$

où d_3 et d_4 sont respectivement la troisième et quatrième convolution de la fonction 1, le supremum est sur toutes les fonctions b telles que $|b(n)| \leq d_3(n)$ et U, V, Z sont des paramètres satisfaisant quelques conditions techniques. Enfin nous appliquons la méthode de Weyl-van der Corput (voir Graham et Kolesnik [**97**]) pour montrer la distribution uniforme.

Avec cette clé en main nous pouvons montrer un théorème ergodique de von Neumann. L'ingrédient principal est le théorème de Bochner et Herglotz.

THÉORÈME 1.25 (Bochner et Herglotz). *Soient $U_1, \ldots, U_k$ des opérateurs commutatifs unitaires sur un espace de Hilbert $\mathcal{H}$ et $f \in \mathcal{H}$. Alors il existe une mesure ν_f sur $\mathbb{T}^k$ telle que*

$$\langle U_1^{n_1} U_2^{n_2} \cdots U_k^{n_k} f, f\rangle = \int_{\mathbb{T}^k} \exp\left(2\pi i(n_1\gamma_1 + \cdots + n_k\gamma_k)\right) \mathrm{d}\nu_f(\gamma_1, \ldots, \gamma_k),$$

pour tout $(n_1, n_2, \dots, n_k) \in \mathbb{Z}^k$.

En associant ce théorème avec le Théorème 1.24, nous obtenons la variante du théorème de von Neumann suivante.

THÉORÈME 1.26. *Soient* $c_1, \dots, c_k \in \mathbb{R}_+^* \setminus \mathbb{N}$ *des réels positifs distincts non entiers. Soient* $U_1, \dots, U_k$ *des opérateurs commutatifs unitaires sur un espace de Hilbert* $\mathcal{H}$. *Alors*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} U_1^{\lfloor p_n^{c_1} \rfloor} \cdots U_k^{\lfloor p_n^{c_k} \rfloor} f = f^*,$$

où p_n *est le* n*-ième nombre premier et* f^* *est la projection de* f *sur l'espace* $\mathcal{H}_{inv} := \{f \in \mathcal{H} : U_i f = f \text{ pour tout } i\}$.

Notre article [**30**] suppose que les c_i sont distincts. Dans l'article avec Robert Tichy [**153**] nous avons supprimé cette hypothèse.

En utilisant la correspondance de Furstenberg, nous déduisons du théorème de von Neumann une généralisation des résultats de Furstenberg [**88**] et Sárközy [**206**–**208**].

THÉORÈME 1.27. *Soient* $c_1, \cdots, c_k$ *des entiers positifs. Soit* $E \subset \mathbb{Z}^k$ *de densité de Banach positive, i.e.*

$$d^*(E) := \limsup_{|I|\to\infty} \frac{|E \cap I|}{|I|} > 0,$$

où la limite est sur tous les parallélépipèdes $I \subset \mathbb{Z}^k$ *dont la longueur des arêtes tendent vers l'infinie.*

Alors il existe un nombre premier p *tel que* $([p^{c_1}], \cdots, [p^{c_k}]) \in E - E$. *De plus,*

$$\liminf_{N\to\infty} \frac{|\{p \le N : ([p^{c_1}], \cdots, [p^{c_k}]) \in E - E\}|}{\pi(N)} \ge d^*(E)^2,$$

où $\pi(N)$ *est la fonction comptant des nombre premiers inférieurs ou égaux à* N.

La méthode de Weyl-van der Corput est une des deux méthodes appliquées pour l'estimation des somme d'exponentielles de la forme :

$$\sum_{n\le N} e\left(\alpha p(n)\right),$$

où p désigne un polynôme à coefficients réel. L'autre méthode est celle de Vinogradov (*cf.* [**247**]). L'avantage de la méthode de Weyl-van der Corput se manifeste pour des polynômes aux degrés petits. A l'inverse, la méthode de Vinogradov donne une bonne estimation dans le cas de degrés suffisamment grands. Entre autres Browning et Heath-Brown [**48**] l'ont utilisée pour estimer le nombre de points sur un système de formes algébriques aux degrés différents.

Soit K un corps de nombres de degré n sur $\mathbb{Q}$ et soit $\mathcal{O}_K$ son anneau des entiers. Soient $D \in \mathbb{N}^*$ et $t_d \in \mathbb{N}$ pour $1 \le d \le D$ avec $t_D \ge 1$. Supposons que pour tout $d \le D$ nous avons des polynômes

$$F_{d,1}(x_1, \dots, x_s), \dots, F_{d,t_d}(x_1, \dots, x_s) \in \mathcal{O}_K[x_1, \dots, x_s]$$

de degré d en s inconnues, telles que leur nombre total est

$$T = t_1 + \cdots + t_D.$$

Posons

$$\Delta := \{d \in \mathbb{N} \colon t_d \ge 1\} \subset \{1, 2, \dots, D\}.$$

Pour tout $d \in \Delta$ nous définissons la matrice

$$J_d(\mathbf{x}) := \begin{pmatrix} \nabla F_{d,1}(\mathbf{x}) \\ \vdots \\ \nabla F_{d,t_d}(\mathbf{x}) \end{pmatrix}$$

et posons

$$B_d := \dim\{\mathbf{x} \in \mathbb{A}^n \colon \operatorname{rank}(J_d(\mathbf{x})) < r_d\}.$$

Fixons un idéal $\mathfrak{n}$ de $\mathcal{O}_K$. Supposons que $\omega_1, \dots, \omega_n$ est une $\mathbb{Z}$-base de $\mathfrak{n}$. Cette base est aussi une $\mathbb{R}$-base de l'espace $V := K \otimes_{\mathbb{Q}} \mathbb{R}$. Soit $\mathcal{B} \subset [-1;1]^n$ une boîte alignée à la base $\omega_1, \dots, \omega_n$, *i.e.* l'ensemble des $\mathbf{x} = (x_1, \dots, x_s) \in V^s$ tels que tout x_i est de la forme $x_{i,1}\omega_1 + \cdots + x_{i,n}\omega_n$ avec des coordonnées $(x_{i,j})_{i,j} \in \mathbb{R}^{ns}$ dans une boîte alignée aux axes de coordonnées. Pour tout P suffisamment grand nous posons

$$N(P) := \#\left\{\mathbf{x} \in \mathfrak{n}^s \cap P\mathcal{B} \colon F_{d,i}(\mathbf{x}) = 0 \quad \forall i, d\right\}.$$

Pour $1 \leq j \leq D$ nous définissons

$$\mathcal{D}_j := t_1 + 2t_2 + \cdots + jt_j$$

et $\mathcal{D}_0 := 0$. Enfin nous posons

$$s_d := \sum_{k=d}^{D} \frac{2^{k-1}(k-1)t_k}{n - B_k}.$$

Conjointement avec Christopher Frei [**84**] nous avons généralisé le résultat de Browning et Heath-Brown [**48**] aux formes sur un corps de nombres.

Théorème 1.28. *Supposons que*

$$\mathcal{D}_d\left(\frac{2^{d-1}}{n - B_d} + s_{d+1}\right) + s_{d+1} + \sum_{j=d+1}^{D} s_j t_j < 1$$

pour $d \in \Delta \cup \{0\}$. Alors il existe un $\delta > 0$ tel que

$$N(P) = \sigma_\infty \left(\prod_p \sigma_p\right) P^{n-\mathcal{D}_D} + \mathcal{O}\left(P^{n-\Delta_D-\delta}\right),$$

où σ_∞ et σ_p ont les densités locales.

Ce travail est une continuation de travail de Birch [**40**] et Schmidt [**220**] qui ont considéré le cas $t_1 = t_2 = \cdots = t_{D-1} = 0$ où la formule asymptotique donnée se ramène aux résultats précédents. Browning et Heath-Brown [**48**] ont le même résultat pour les rationnels (*i.e.* $K = \mathbb{Q}$). Ce théorème est aussi la motivation pour notre généralisation.

En passant de $\mathbb{Q}$ vers un corps de nombres il y a plusieurs obstacles. La méthode utilisée est la méthode du cercle. Écrivons $e(x) = \exp(2\pi i x)$ pour $x \in \mathbb{R}$ et $\Phi(x) = e(\operatorname{Tr}(x))$ où $\operatorname{Tr}\colon V \to \mathbb{R}$ est la généralisation de la trace d'un élément sur $\mathbb{Q}$. Alors la méthode du cercle porte sur l'identité suivante :

$$N(P) = \int_{\alpha \in R^T} S(\alpha) \mathrm{d}\alpha,$$

où

$$S(\alpha) := \sum_{\mathbf{x} \in \mathfrak{n}^s \cap P\mathfrak{B}} \Phi\left(\sum_{d=1}^{D} \sum_{i=1}^{t_D} \alpha_{d,i} G_{d,i}(\mathbf{x})\right).$$

L'idée centrale est que la fonction $S(\alpha)$ est d'autant plus grande que α est proche d'un « rationnel » avec un petit dénominateur. En pratique, nous divisons R^T en deux parties appelées les arcs majeurs et les arcs mineurs. Soit $\varpi \in (0, 1/3)$ une constante fixée. Pour tout $\gamma \in K$ nous définissons l'idéal de dénominateur (par rapport à $\mathfrak{n}$) par $\mathfrak{a}_\gamma := \{\beta \in \mathcal{O}_K : \beta\gamma \in \mathfrak{n}\}$. De plus pour $\gamma = (\gamma_{d,i})_{d,i} \in (R \cap K)^T$ posons $\mathfrak{a}_\gamma := \bigcap_{d,i} \mathfrak{a}_{\gamma_{d,i}}$ l'idéal de dénominateur commun. L'arc majeur associé à γ est

$$\mathfrak{M}_\gamma := \left\{\alpha \in R^T \,:\, |\alpha_{d,i} - \gamma_{d,i}| \leq P^{-d+\varpi} \text{ pour tout } 1 \leq d \leq D \text{ et } 1 \leq i \leq t_d\right\},$$

où il faut interpréter la distance $|\alpha_{d,i} - \gamma_{d,i}|$ modulo $\mathfrak{n}$. Alors les arcs majeurs sont définis par

$$\mathfrak{M} := \bigcup_{\substack{\boldsymbol{\gamma} \in (R\cap K)^T \\ \mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}} \leq P^{\varpi}}} \mathfrak{M}_{\boldsymbol{\gamma}}$$

et les arcs mineurs par

$$\mathfrak{m} := R^T \smallsetminus \mathfrak{M}.$$

La contribution des arc mineurs disparaît dans le terme d'erreur. Les difficultés de cette estimation viennent de la multiplication dans un corps de nombres qui est une multiplication matricielle. Alors l'approximation des coefficients se ramène à la solution d'un système d'équations linéaires. On peut déjà trouver cette partie dans le travail de Skinner [**230**].

En revanche, la contribution des arc majeurs donne la formule asymptotique :

$$\mathfrak{S}\mathfrak{J}P^{n(s-\mathcal{D})} + O(P^{n(s-\mathcal{D})-\delta}),$$

où

$$\mathfrak{S} = \prod_{\mathfrak{p}} \sum_{j=0}^{\infty} \frac{1}{\mathfrak{N}\,\mathfrak{p}^{js}} \sum_{\substack{\boldsymbol{\gamma} \in (R\cap K)^T \\ \mathfrak{a}_{\boldsymbol{\gamma}} = \mathfrak{p}^j}} \sum_{\mathbf{x} \in (\mathfrak{n}/\mathfrak{p}^j\mathfrak{n})^s} \Phi\left(\sum_{d=1}^{D}\sum_{i=1}^{t_d} \gamma_{d,i} G_{d,i}(\mathbf{x})\right)$$

est la série singulière et

$$\mathfrak{J} = \int_{\boldsymbol{\gamma} \in V^T} \int_{\mathbf{x} \in \mathcal{B}} \Phi\left(\sum_{d=1}^{D}\sum_{i=1}^{t_d} \gamma_{d,i} F_{d,i}(\mathbf{x})\right) \mathrm{d}\mathbf{x}\, \mathrm{d}\boldsymbol{\gamma}$$

est l'intégrale singulière, respectivement.

La série singulière porte sur la distribution des solutions locales. L'argument de Browning et Heath-Brown [**48**] s'applique parce qu'il est géométrique. En revanche, dans le traitement de l'intégrale singulière de Skinner [**230**], nous avons trouvé une erreur. En particulier, la difficulté ici vient de la norme des éléments. Dans $\mathbb{Q}$, on utilise que $|n| \geq 1$ pour $n \in \mathbb{Z}^*$. Cette inégalité reste valable pour la norme d'un élément mais pas pour la valeur absolue de ses conjugués. On peut alors diviser par une valeur absolue arbitrairement petite qui rend l'intégrale singulière arbitrairement grande.

Avant d'expliquer notre idée j'aimerais évoquer les chemins qui sont utilisés dans des problèmes similaires. Pour éviter l'obstacle, Skinner [**230**] a transféré le problème d'un système de r formes sur un corps de nombres vers un système de $n \cdot r$ formes sur $\mathbb{Q}$, où n est le degré du corps K. Une autre source de méthodes dans les corps de nombres sont les généralisations du résultat remarquable de Heath-Brown [**106**] sur les formes cubiques en dix variables. Il peut montrer que ces formes ont une racine non triviale. Ce résultat repose sur une version discrète de la méthode du cercle. Cette version est généralisée sur les corps des nombres par Skinner [**231**] (pour 11 variables) et très récemment par Browning et Vishe [**47**]. Cependant,

ces derniers introduisent une version pondérée de l'intégrale singulière qui élude les difficultés sur un chemin et n'est donc pas applicable à notre problème.

Si les valeurs absolues des conjuguées sont proches de la racine n-ième de la norme, alors l'estimation est la même que dans le cas de Browning et Heath-Brown [**48**]. Notre idée est de séparer le cas où les valeurs absolues des conjuguées sont très différentes – très grandes et très petites. En utilisant l'existence d'une racine de l'unité μ telle que les valeurs absolues des conjuguées sont aussi très grandes et très petites nous pouvons effectuer une rotation de l'élément u tel que les valeurs absolues des conjuguées de μu sont autour de la racine n-ième de la norme. Le reste de la démonstration consiste à estimer les deux intégrales singulières.

Nous revenons aux systèmes de numération et à une autre application de la méthode de Weyl-van der Corput. Comme pour la construction du dragon de Knuth, nous allons considérer les nombres dont la représentation n'a que des puissances négatives de la base. Pour le cas $q \geq 2$ entier, cela donne l'intervalle $[0, 1]$. En particulier, tout $x \in [0, 1]$ admet une représentation de la forme

$$x = \sum_{k \geq 1} a_k q^{-k}.$$

En notant que $\sum_{k \geq 1}(q-1)q^{-k} = 1$ pour tout $q \geq 2$, on peut supposer l'unicité de cette représentation en ajoutant la condition que seulement un nombre fini d'éléments sont égaux à $q - 1$.

Nous nous intéressons à la distribution des chiffres et des blocs de chiffres dans un nombre aléatoire. En particulier, on dit qu'un nombre $x \in [0, 1]$ est normal en base q si pour tout $k \geq 1$ et tout bloc de chiffres $b_1 \dots b_k \in \{0, \dots, q-1\}^k$, la fréquence d'apparition de ce bloc tend vers q^{-k}, c'est-à-dire

$$\lim_{N \to \infty} \frac{1}{N} \# \{n \leq N \colon a_n = b_1, \dots, a_{n+k-1} = b_k\} = \frac{1}{q^k}.$$

Borel [**43**] a montré que presque tout $x \in [0, 1]$ est normal dans toute base $q \geq 2$. Il a conjecturé que "presque tous" les nombres algébriques sont normaux. La connexion entre les nombres normaux et l'équirépartition est établie par le lemme suivant, montré par Borel [**43**].

LEMME 1.29. *Un réel $x \in [0, 1]$ est normal en base q si et seulement si la suite $(q^n x)_{n \geq 1}$ est équirépartie modulo 1.*

Une construction très simple d'un nombre normal est celle de Champernowne [**58**]. Il a montré que le nombre dont la représentation est la concaténation des nombres naturels, *i.e.*

$$0.1\,2\,3\,4\,5\,6\,7\,8\,9\,10\,11\,12\,13\,14\,15\,16\,17\,18\,19\,20\,\dots,$$

est normal en base 10. Mahler a montré que ce nombre est transcendant.

On peut voir la construction de Champernowne comme une concaténation des valeurs de $f(n)$ avec $f(x) = x$ en base 10. Cette idée a été généralisée pour les polynômes sur $\mathbb{Z}[X]$, sur $\mathbb{Q}[X]$ et $\mathbb{R}[X]$ par Davenport and Erdős [**63**], Schiffer [**214**] et Nakai and Shiokawa [**167**], respectivement. Conjointement avec Jörg Thuswaldner et Robert Tichy [**146**], nous avons considéré le cas des fonctions entières transcendantes (non polynomiales).

De même, Champernowne a conjecturé que le nombre dont la représentation est la concaténation des nombres premiers, *i.e.*

$$0.2\,3\,5\,7\,11\,13\,17\,19\,23\,29\,31\,37\,41\,43\,53\,59\,61\,67\,71\,\dots,$$

est normale en base 10. Copeland et Erdős [**60**] ont démontré cette conjecture. Nakai et Shiokawa [**168**] ont considéré la construction avec la suite $f(p)$ où f est un polynôme tel que $f(\mathbb{N}) \subset \mathbb{N}$ et p parcourt les nombres premiers.

Dans le Chapitre 8 nous présentons une construction pour des fonctions pseudo-polynomiales. Soit $f(x) = \alpha_d x^{\beta_d} + \cdots + \alpha_1 x^{\beta_1}$ avec $\alpha_i \in \mathbb{R}$ et $\beta_d > \cdots > \beta_1 > 1$. On dit que f est un pseudo-polynôme s'il existe $1 \leq i \leq d$ tel que $\beta_i \notin \mathbb{Z}$. Nakai et Shiokawa [**166**] ont montré que la concaténation des valeurs $f(n)$ où f est un pseudo polynôme donne un nombre normal. Dans le Chapitre 8, nous généralisons ce résultat à la suite des nombres premiers.

THÉORÈME 1.30. *Soient $q \geq 2$ un entier et $f(x) = \sum_{j=1}^{d} \alpha_j x^{\beta_j}$, où $0 < \beta_1 < \beta_2 < \cdots < \beta_d$. Supposons qu'il existe $1 \leq j \leq d$ tel que $\beta_j \notin \mathbb{Z}$. Alors le nombre*

$$\tau_f = 0.\lfloor f(2) \rfloor \lfloor f(3) \rfloor \lfloor f(5) \rfloor \lfloor f(7) \rfloor \lfloor f(11) \rfloor \lfloor f(13) \rfloor \cdots$$

est normal en base q.

Un bloc de chiffres $\mathbf{b} = b_1 b_2 \dots b_k$ peut apparaître dans τ_f dans un $\lfloor f(p_n) \rfloor$ ou entre deux $\lfloor f(p_n) \rfloor$, $\lfloor f(p_{n+1}) \rfloor$ consécutifs (p_n est le n-ième nombre premier). La structure polynomiale de f entraîne que la longueur de l'écriture de $\lfloor f(p_n) \rfloor$ est croissante, donc que la contribution des apparitions entre deux écritures $\lfloor f(p_n) \rfloor$ et $\lfloor f(p_{n+1}) \rfloor$ est négligeable. Soit $\mathcal{N}(x, \mathbf{b})$ le nombre d'apparitions du bloc $\mathbf{b}$ dans l'écriture de $\lfloor x \rfloor$ en base q.

Pour tout bloc de chiffres $\mathbf{b}$ nous définissons la fonction indicatrice des réels dans $[0, 1]$ dont les k premiers chiffres de l'écriture en base q sont $b_1, \dots, b_k$

$$\mathcal{I}_{\mathbf{b}}(x) = \begin{cases} 1 & \text{si } \sum_{i=1}^{k} \frac{b_i}{q^i} \leq x < \sum_{i=1}^{k} \frac{b_i}{q^i} + \frac{1}{q^k}, \\ 0 & \text{sinon.} \end{cases}$$

Maintenant nous rassemblons les $\lfloor f(p) \rfloor$ en groupes dont l'écriture en base q est de même taille. Comme f est de forme polynomiale il existe j_0 et X_j tels que

$$\lfloor f(p) \rfloor = a_{j-1} q^{j-1} + a_{j-2} q^{j-2} + \cdots + a_1 q + a_0 \quad \text{pour tout } X_j \leq p < X_{j+1} \text{ et } j \geq j_0.$$

On a donc pour le nombre total d'apparitions

$$\sum_{n \leq N} \mathcal{N}(f(p), \mathbf{b}) = \sum_{j=j_0}^{J} \sum_{\ell=0}^{j-k} \sum_{X_j \leq p < X_{j+1}} \mathcal{I}_{\mathbf{b}} \left(\frac{f(p)}{q^\ell} \right).$$

L'approximation des fonctions indicatrices par Vinogradov (voir Chapitre 1 de [**247**]) réduit le problème à une estimation d'une somme d'exponentielles de la forme

$$\sum_{X_j \leq p < X_{j+1}} e \left(\frac{\nu f(p)}{q^\ell} \right).$$

Ces sommes sont similaires à celles qui apparaissent dans l'équirépartition ci-dessus. La différence, le dénominateur q^j, est aussi source de grandes difficultés. Pour estimer les sommes d'exponentielles, il faut distinguer plusieurs cas selon la taille de j et selon le ratio entre la plus grande puissance entière et plus grande puissance non entière. Chaque combinaison demande un traitement individuel.

6. Systèmes dynamiques symboliques

Nous avons déjà vu que la construction d'un nombre normal repose sur la concaténation de mots. Dans cette section nous prenons le point de vue dynamique symbolique sur les systèmes de numération. Nous utilisons la notation de Lind et Marcus [**138**].

Dans les trois derniers chapitres, nous considérons les systèmes dynamiques symboliques. Nous commençons par les mots et les langages. Soit $\mathcal{A}$ un ensemble fini ou dénombrable appelé alphabet. Les éléments de $\mathcal{A}$ sont appelés les lettres. Soit $\mathcal{A}^k$ l'ensemble des mots de longueur k et $\mathcal{A}^* = \bigcup_{k=0}^{\infty} \mathcal{A}^k$, où $\mathcal{A}^0 = \{\varepsilon\}$ contient seulement le mot vide, l'ensemble des mots finis.

Un décalage (unilatère) sur un alphabet $\mathcal{A}$ (un ensemble fini ou dénombrable) se définit comme l'espace de suites

$$\Sigma = \Sigma(\mathcal{A}) := \mathcal{A}^{\mathbb{N}}$$

muni de la transformation $\sigma\colon \Sigma \to \Sigma$ définie par $\sigma(\alpha) = (\alpha_{n+1})_{n\in\mathbb{N}}$. Avec la topologie discrète (qui est équivalente à la distance d définie par $d(x,y) = 0$ si $x = y$ et $d(x,y) = 1$ sinon) sur l'alphabet $\mathcal{A}$, on munit l'espace Σ d'une topologie produit équivalente à celle donnée par la distance suivante :

$$\forall x, y \in \Sigma, d(x,y) = \begin{cases} 2^{-\min\{|n|, x_n \neq y_n\}} & \text{si } x \neq y, \\ 0 & \text{sinon.} \end{cases}$$

Pour bien expliquer la connexion avec les systèmes de numération, il faut faire appel à la dynamique topologique. Soient M un espace topologique et $\varphi\colon M \to M$ une application continue. Alors on appelle le couple (M, φ) un système dynamique topologique. Une partition topologique (de M) est une famille $\mathcal{P} = \{P_0, \ldots, P_{q-1}\}$ des sous-ensembles disjoints et ouverts, telle que

$$M = \overline{P_0} \cup \overline{P_1} \cup \cdots \cup \overline{P_{q-1}}.$$

Nous fixons un système dynamique topologique (M, φ) et une partition topologique $\mathcal{P} = \{P_0, \ldots, P_{q-1}\}$ de M pour le reste de cette section. On dit qu'un mot $\omega = a_1 a_2 \ldots a_n$ est admissible pour $\mathcal{P}, \varphi$ si $\bigcap_{j=1}^{n} \varphi^{-j}(P_{a_j}) \neq \emptyset$. Soit $\mathcal{L} \subset \mathcal{A}^*$ l'ensemble de tous les mots admissibles pour $\mathcal{P}, \varphi$. L'ensemble $\mathcal{L}$ est le langage de l'espace de décalage unique X (qui contient toutes les représentations possibles).

Pour $\omega = a_1 a_2 a_3 \ldots \in X$ et $n \geq 1$ nous posons

$$D_n(\omega) = \bigcap_{k=1}^{n} \varphi^{-k}(P_{a_k}) \subset M.$$

Les adhérences $\overline{D_n(\omega)}$ sont compactes et décroissantes avec $n : \overline{D_1(\omega)} \supseteq \overline{D_2(\omega)} \supseteq \overline{D_3(\omega)} \supseteq \cdots$ D'après le théorème de Baire, on a $\bigcap_{n=1}^{\infty} \overline{D_n(x)} \neq \emptyset$. On dit que la partition topologique $\mathcal{P}$ de M donne une représentation symbolique de (M, φ) si pour tout $x \in X$ l'intersection $\bigcap_{n=1}^{\infty} \overline{D_n(\omega)}$ consiste en un seul point. Alors il existe une application $\pi\colon X \to M$ qui envoie tout mot infini $\omega = a_1 a_2 a_3 \ldots$ vers le point unique $\pi(\omega)$ dans l'intersection $\bigcap_{n=1}^{\infty} \overline{D_n(\omega)}$. On appelle ω la représentation symbolique de $\pi(\omega)$.

Par exemple l'ensemble $M = [0,1]$ et l'application φ définie par $x \mapsto 10x - \lfloor 10x \rfloor$ forment un système dynamique topologique. La partition topologique

$$\mathcal{P} = \left\{ \left(0, \frac{1}{10}\right), \ldots, \left(\frac{9}{10}, 1\right) \right\}$$

nous donne le système décimal comme système symbolique associé. On peut obtenir tous les systèmes q-adique de la même manière. Comme il n'y a pas de restrictions, la représentation symbolique présente peu d'intérêt.

Soient $\beta = \frac{1+\sqrt{5}}{2}$ le nombre d'or et φ définie par $x \mapsto \beta x - \lfloor \beta x \rfloor$. La partition

$$\mathcal{P} = \left\{ \left(0, \frac{1}{\beta}\right), \left(\frac{1}{\beta}, 1\right) \right\}$$

engendre le système de base le nombre d'or. Comme $\frac{1}{\beta} + \frac{1}{\beta^2} = 1$, on obtient que

$$\varphi^{-1}\left(\left(0, \frac{1}{\beta}\right)\right) = (0, 1)\,.$$

Le mot 11 ne peut donc pas apparaître dans les représentations. Autrement dit le décalage associé X est défini par

$$X = \left\{ \omega = a_1 a_2 a_3 \ldots \in \{0,1\}^{\mathbb{N}} \colon a_k \cdot a_{k+1} = 0 \ \forall k \geq 1 \right\}.$$

Maintenant nous nous concentrons sur des systèmes dynamiques symboliques dont le langage satisfait la propriété de spécification. On dit qu'une langage $\mathcal{L}$ satisfait la propriété de spécification s'il existe une constante c, qui ne dépend que du langage $\mathcal{L}$, telle que pour tout $a, b \in \mathcal{L}$, il existe un mot $u \in \mathcal{L}$ de longueur $|u| \leq c$ tel que aub est aussi un élément de $\mathcal{L}$. Il est clair que cette condition nous permet de concaténer les mots, ce qui est l'opération essentielle pour toutes les constructions présentées ci-dessus.

Pour finir nous aimerions transférer le concept de la normalité aux systèmes dynamiques symboliques. Soit μ une mesure probabiliste σ-invariante (c'est-à-dire $\mu(\sigma(A)) = \mu(A)$ pour tout $A \subset X$ mesurable). Pour tout $k \geq 1$, $\mathbf{b} = b_1 b_2 \ldots b_k \in \mathcal{A}^k$, $\omega = a_1 a_2 a_3 \ldots \in X$ et $n \geq 1$ on pose

$$\mathrm{P}(\omega, \mathbf{b}, n) = \frac{\#\left\{0 \leq i < n \colon a_{i+1} = b_1, \ldots, a_{i+k} = b_k\right\}}{n}$$

la fréquence d'apparitions du bloc $\mathbf{b}$ parmi les n premiers chiffres de ω. On dit que $\omega = a_1 a_2 a_3 \ldots \in X$ est normal par rapport à μ ou ω est générique pour la mesure μ si pour tout $k \geq 1$ et tout mot $\mathbf{b} = b_1 b_2 \ldots b_k \in \mathcal{A}^k$ on a

$$\lim_{n \to \infty} \mathrm{P}(\omega, \mathbf{b}, n) = \mu(\mathbf{b}).$$

Dans le Chapitre 9 nous construisons pour toute mesure probabiliste σ-invariante μ donnée un mot infini ω qui est normal par rapport à μ. L'idée centrale est d'utiliser la construction de Champernowne et de répéter les mots suivant leurs propriétés pour la distribution μ. Pour démontrer que le mot est générique pour la mesure μ, on compte le nombre d'apparitions d'un bloc dans les n premiers chiffres. Cette fois-ci, il faut distinguer trois cas : le bloc apparaît dans un mot, entre deux mots égaux ou entre deux mots différents. Ce dernier cas est négligeable (comme auparavant) et pour les deux autres cas nous trouvons des bornes supérieures et inférieures.

Nous appliquons cette construction aussi aux fractions continues avec un nombre infini de chiffres. Dans ce cas, nous accroissons la base d'expansion à chaque étape pour obtenir finalement un nombre normal. D'autres applications et exemples portent sur l'écriture en base q, la β-numération, le développement en série de Lüroth.

Les constructions mentionnées jusqu'à maintenant nous donnent des nombres normaux dans une seule base (ou par rapport à un seul couple ϕ,$\mathcal{P}$). Comme les systèmes β sont définis sur le même ensemble $[0, 1]$ et que Borel a démontré que presque tout réel est normal par

rapport à toute base, il serait intéressant de construire des nombres qui sont normaux dans plusieurs bases. S'ils sont normaux par rapport à toute base, il répondent à la question d'une construction d'un nombre absolument normal. Il existe essentiellement les constructions de Schmidt [**217**] et Turing [**241**] de nombres absolument normaux. Celle de Turing donne des intervalles où l'écriture en base q est fixe jusqu'à un certain rang. L'intersection des intervalles pour $q \leq q_i$ donne des nombres dont l'écriture est fixée pour les bases premières q_i. Sous certaines conditions, cela donne un nombre absolument normal pour $q_i \to \infty$.

Verónika Becher, Pablo Ariel Heiber et Théodore Slaman [**22**] ont construit un algorithme utilisant l'idée de Turing qui donne un nombre normal en temps polynomial. Cette construction utilise en fait qu'il suffit de montrer qu'un nombre est simplement normal en toute base, *i.e.* que tout chiffre (et ne pas tout bloc) apparaît avec la bonne fréquence. Conjointement avec Scheerer et Tichy nous avons généralisé ce résultat aux nombres absolument Pisot normaux. L'idée centrale de la construction de Turing [**241**] est une cascade d'intervalles (plus précisément des cylindres) telle que les blocs de plus en plus longs sont bien distribués. En particulier, à l'étape i nous avons une cascade

$$I_2^{(i)} \subset I_3^{(i)} \subset \cdots \subset I_{b_i}^{(i)}$$

d'intervalles telle que dans chaque intervalle I_j^i les blocs en base j de longueur $\leq k_i$ sont bien distribués. La difficulté est maintenant de trouver un intervalle $I_2^{(i+1)} \supset I_{b_i}^{(i)}$ qui n'est ni trop grand ni trop petit.

Dans notre version sur les bases Pisot nous avons deux obstacles principaux. Le premier est que Becher, Heiber et Slaman utilisent le fait qu'il suffit de démontrer qu'un nombre est simplement normal en base q, q^2, q^3 *etc.* pour démontrer qu'il est normal en base q. Alors pour leur construction il suffit de montrer que le nombre est simplement normal en tout base pour obtenir qu'il est en effet absolument normal. Cette méthode ne fonctionne plus pour des bases Pisot. Pour obtenir notre résultat, il nous faut donc montrer que le nombre construit est normal par rapport à tout base β_i. Le deuxième obstacle n'est pas vraiment un obstacle. C'est plutôt que dans la littérature le développement β est considéré du point vue dynamique, ce qui donne des résultats qualitatifs et non quantitatifs. Dans notre travail, il faut quantifier beaucoup de résultats pour démontrer que le nombre est construit en temps polynômial.

Maintenant nous quittons les nombres normaux et passons aux nombres non normaux. Comme indiqué plus haut, Borel [**43**] a montré que presque tout nombre est normal. D'un autre côté, les nombres normaux forment un ensemble maigre (ou de première catégorie). Un ensemble est dit maigre s'il est réunion dénombrable d'ensembles nulle part denses. Un ensemble comaigre (ou résiduel) est le complémentaire d'un ensemble maigre. Les nombres non normaux (un ensemble comaigre) sont donc plus intéressants au sens topologique. Dans le dernier chapitre, nous considérons deux classes des nombres non normaux.

Pour $k, n \geq 1$ des entiers et $\omega = a_1 a_2 a_3 \ldots \in X$, on pose

$$\mathrm{P}_k(\omega, n) = (\mathrm{P}(\omega, \mathbf{b}, n))_{\mathbf{b} \in \mathcal{A}^k}$$

le vecteur des fréquences d'apparitions des blocs de longueur k parmi les n premiers chiffres de ω. Soit $A_k(\omega)$ l'ensemble de tous les points d'adhérence de la suite $(\mathrm{P}_k(\omega, n))_n$. De plus, soit S_k l'union de tous les points d'adhérence possibles, *i.e.*

$$S_k := \bigcup_{\omega \in X} A_k(\omega).$$

On dit qu'un nombre est essentiellement non normal si, pour tout chiffre $i \in \mathcal{A}$, la limite

$$\lim_{n\to\infty} \mathrm{P}(\omega, i, n)$$

n'existe pas.

THÉORÈME 1.31. *Soit $\mathcal{P} = \{P_0, \ldots, P_{N-1}\}$ une partition topologique de (M, φ). Supposons que*

- *$\bigcap_{n=0}^{\infty} \overline{D_n(\omega)}$ consiste en un seul point ;*
- *$X_{\mathcal{P},\phi}$ satisfait la propriété de spécification ;*
- *pour tout $0 \leq i < N$, il existe $\mathbf{q}_{i,1} = (q_{1,1}, ...q_{1,N-1}), \mathbf{q}_{i,2} = (q_{2,1}, ...q_{2,N-1}) \in S_1$ tels que $|q_{1,i} - q_{2,i}| > 0$.*

Alors l'ensemble des nombres essentiellement non normaux est un ensemble comaigre.

Il est clair que, pour tout $\omega \in X$, on a que $A_k(\omega) \subset S_k$. On dit qu'un nombre $\omega \in X$ est extrêmement non normal si $A_k(\omega) = S_k$.

THÉORÈME 1.32. *Soit $\mathcal{P} = \{P_0, \ldots, P_{N-1}\}$ une partition topologique de (M, φ). Si le langage $\mathcal{L}$ associé satisfait la propriété de spécification, alors l'ensemble des nombres extrêmement non normaux est comaigre.*

La démonstration dans les deux cas est constructive. En effet nous construisons un sous-ensemble de nombres essentiellement (resp. extrêmement) non normaux qui est une intersection d'ensembles ouverts et denses, si bien que ce sous-ensemble est comaigre. Ci-dessous, j'expose seulement la construction de l'ensemble des mots extrêmement non normaux. En effet, l'autre construction est très similaire.

Chaque vecteur $\mathrm{P}_k(\omega, n)$ engendre une distribution sur les blocs de longueur k et donc une mesure sur le langage. On peut définir une distance entre deux mesures comme une distance entre les vecteurs qui les engendrent. Cela donne une topologie compacte dont il existe une suite dense $(\mathbf{q}_i)_i$. Dans la construction, nous concaténons des mots qui sont proches de $\mathbf{q}_i$. Plus précisément, soit $\phi_1(x) = 2^x$ et $\phi_m(x) = \phi_1(\phi_{m-1}(x))$ pour tout $m \geq 2$. On dit qu'une suite $(\mathbf{x}_n)_n$ dans $\mathbb{R}^{N^k}$ satisfait la propriété P si pour tout $\mathbf{q} \in \mathbb{Q}^n \cap S_k$, $m \in \mathbb{N}$, $i \in \mathbb{N}$ et $\varepsilon > 0$, il existe $j \in \mathbb{N}$ tel que

(1) $j \geq i$,
(2) $j/2^j < \varepsilon$,
(3) si $j < n < \phi_m(j)$, alors $\|\mathbf{x}_n - \mathbf{q}\| < \varepsilon$.

Soit E l'ensemble des mots infinis dont la suite des fréquences d'apparitions satisfait la propriété P, *i.e.*

$$E = \{x \,:\, (P_k(x; n))_{n=1}^{\infty} \text{ satisfait la propriété } P\}.$$

Dans la démonstration, nous montrons que E est comaigre et que les éléments de E sont extrêmement non normaux.

Chapitre 2

On multiplicative independent bases for canonical number systems in cyclotomic number fields

This chapter is joint work with Paul Surer and Volker Ziegler and appeared in *Number theory – Diophantine problems, uniform distribution and applications*, 313 – 332, Springer, Cham, 2017.

1. Introduction

Let $q \geq 2$ be a positive integer. Then every $n \in \mathbb{N}$ admits a unique q-adic representation of the form

$$n = \sum_{j=0}^{\ell} a_j q^j \qquad \text{with } a_j \in \mathfrak{N} := \{0, \dots, q-1\} \text{ for } 0 \leq j \leq \ell.$$

We call the pair $(q, \mathfrak{N})$ a number system with base q and set of digits $\mathfrak{N}$. A set of positive integers $E \subset \mathbb{N}$ is called q-recognizable, if the language (that is the set of the occurring digit strings) of q-adic representations of the elements of E is recognizable by an automaton (*i.e.* the language is regular). As an example we consider the automaton in Figure 1, which determines the congruence class modulo 3 of integers in binary representation. In particular, starting with the least significant bit of the base-2 expansion of n we follow the path whose labels correspond to the digit string. Then the number in the final state yields the congruence class modulo 3 of n. It is easy to construct a similar automaton that recognises any given

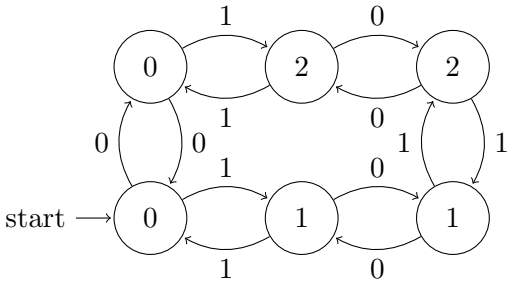


FIGURE 1. The automaton recognizing the congruence classes modulo 3 in base 2.

arithmetic progression with respect to any given base. Therefore ultimately periodic sets, *i.e.* unions of a finite set and a finite union of arithmetic progressions, are the easiest example of sets that are q-recognisable for every $q \geq 2$.

The converse direction, however, is more involved. Suppose that a set $E \subset \mathbb{N}$ is q-recognisable as well as q'-recognisable with $q \neq q'$. If q and q' are multiplicatively dependent, *i.e.* the equation $q^p = q'^{p'}$ has non-trivial solutions over the non-negative integers (solutions

of the form $(p, p') \neq (0, 0)$), then the q-adic representation is a recoding of the q'-adic one and vice versa. Thus, any set that is q-recognisable is trivially q'-recognisable, too. Therefore, we suppose that q and q' are multiplicatively independent. In this case Cobham's theorem [**59**] states that arithmetic progressions and unions thereof are the only sets which are q-recognizable as well as q'-recognizable.

THEOREM 2.1 ([**59**, Cobham's Theorem]). *Let $q, q' \geq 2$ be two multiplicatively independent integers. A set $E \subset \mathbb{N}$ is both, q-recognizable and q'-recognizable, if and only if E is a finite union of arithmetic progressions.*

Cobham's theorem has been generalised in several directions as U-numerations [**36**], fractals [**1**], automata [**42**], and substitutions [**71**]. A more detailed overview can be found in the survey of Durand and Rigo [**72**].

The aim of the present article is to show a variant of Cobham's theorem in the spirit of Hansel and Safer [**101**] for number systems in the ring of integers of cyclotomic fields. This will be realised in three parts : first we have to find bases that induce a number system, secondly we will study their multiplicative independence, and finally we have to make minor adaptions in order to transfer the proof of Cobham's theorem to our generalised notion of number systems. Especially the first two parts are done in a very general way and yield interesting results on their own that we want to outline quickly.

In the first part we are concerned with appropriate generalisations of number systems. Knuth [**126**] suggested to use the base $2i$ and the set of digits $\{0, 1, 2, 3\}$. Since $(2i)^2 = -4$ this corresponds to an expansion of the real and imaginary part in base -4 (see page 205f in volume 2 of Knuth's famous monograph "The art of computer programming" [**127**]). This idea was independently extended to $-1 \pm i$ by Khmelnik [**122**] and Penney [**178**]. In particular, they show that every Gaussian integer z can be represented uniquely as

$$z = \sum_{j=0}^{\ell} a_j(-1+i)^j \qquad \text{with } a_j \in \{0, 1\} \text{ for } 0 \leq j \leq \ell.$$

Analogously to above, we call $(-1+i, \{0, 1\})$ a number system in the Gaussian integers $\mathbb{Z}[i]$. This concept can be generalised in a straightforward way. Indeed, some years later Kátai and Szabó [**120**] proved that $(q, \mathcal{N})$ is a number system for $\mathbb{Z}[i]$ if and only if $q = -m \pm i$ with $m \geq 1$, and $\mathcal{N} = \{0, \ldots, m^2\}$. In subsequent researches, Gilbert [**92**] and Kátai and Kovács [**118**, **119**] independently characterized all number systems for the ring of integers in arbitrary quadratic fields.

As already mentioned we are interested in number systems in $\mathbb{Z}[\zeta]$ where ζ is a primitive kth root of unity. Since the Gaussian integers and the Eisenstein integers are the ring of integers in the fourth and sixth cyclotomic field, respectively, this can be seen as a generalisation of the above mentioned results for quadratic number fields.

Every base of a number system in the ring of integers of a number field is clearly also a power base of the ring of integers. Thus, the most obvious bases have the shape $a \pm \zeta$ with $a \in \mathbb{Z}$. Less obvious are bases of the shape $a \pm \eta$ and $a \pm \theta$ where $\eta = (1+\zeta)^{-1}$ and $\theta = (1-\zeta)^{-1}$. This is motivated by a conjecture of Bremner [**46**] and Robertson [**199**] stating that these are the only possible power integral bases (see also [**89**, **193**, **196**–**198**] for details and partial results). Hence, we concentrate on these three types of potential bases and have to find out for which integers a we obtain number systems.

We consider the problem from the point of view of *canonical number system*. This generalisation of the concept of number systems has been presented in 1991 by Pethő [**179**] : For a given monic polynomial $P \in \mathbb{Z}[x]$ let $\mathcal{N} := \{0, \dots, |P(0)| - 1\}$. We call $(P, \mathcal{N})$ a canonical number system (CNS for short) if for each $Q \in \mathbb{Z}[x]$ there exists a polynomial $A \in \mathcal{N}[x]$ such that $Q \equiv A \pmod{P}$. For irreducible polynomials P we clearly have that $(P, \mathcal{N})$ is a CNS if and only if $(q, \mathcal{N})$ is a number system in $\mathbb{Z}[q]$ for each $q \in \mathbb{C}$ with $P(q) = 0$.

Due to their complex structure the characterization of CNS polynomials, (*i.e.* polynomials that induce a CNS) is a challenging task. While the quadratic case is completely solved by the above cited results, there exist only partial results and algorithms for polynomials of higher degree (*e.g.*, [**4**, **6**–**8**, **49**–**55**, **128**, **129**, **213**]). Furthermore, canonical number systems are closely related with so-called *shift radix systems* (see [**3**]). From this interaction one obtains further characterisation results for CNS. For an overview concerning shift radix systems we refer the interested reader to the survey of Kirschenhofer and Thuswaldner [**123**].

In view of our initial problem we have to check for which $a \in \mathbb{Z}$ the minimal polynomials of $a \pm \zeta$, $a \pm \eta$ and $a \pm \theta$ are CNS-polynomials. The problem is completely solved for the quadratic cases $k = 3, 4, 6$ in [**92**, **120**]. Concerning cyclotomic fields of higher degree the first and the last author showed in [**154**] that $\Phi_k(a+x)$ induces a CNS for $a \leq -\varphi(k)-1$, where Φ_k is the kth cyclotomic polynomial and φ denotes Euler's totient function. We show analogous results for the minimal polynomials of η and θ. In fact, this will follow from our main result of the first part which is a much stronger statement. For a given polynomial $R \in \mathbb{Z}[x]$ we will give a bound for a such that $R(a + x)$ induces a CNS. As a further generalisation we will not only consider the case $a \in \mathbb{Z}$ but also the case $a \in \mathbb{Q}$. This is justified by a recently published research [**211**] that shows that the notion of canonical number system can be easily generalised to non-monic polynomials.

In the second part we analyse under which conditions pairs of bases are multiplicatively independent. The concept of multiplicative independence appears in other contexts, too. Senge and Straus [**223**] showed that the set of integers whose sum of digits is bounded with respect to different bases is finite if and only if these bases are multiplicatively independent. Steward [**235**] obtained an effective version of this result and Schlickewei [**216**] extended this to $t > 2$ multiplicative independent bases. Finally, Pethő and Tichy [**181**] solved the CNS analog.

The present article is organised in the following way. In Section 2 we give all necessary definitions and state the main results. Then, in Section 3, we show the results on canonical number systems. The statements concerning the multiplicative independence of the bases are proved in Section 4. Finally, Section 5 contains the proof of our variant of Cobham's theorem for number systems in the ring of integers of cyclotomic fields.

2. Definitions and Statement of results

In the introduction we already defined the notion of canonical number system and noted that it can be generalised to non-monic polynomials. Therefore we want to give a unique definition which we will use in the sequel.

Définition 2.1 (Canonical number system, *cf.* [**179**, **211**]). Let $P = p_d x^d + p_{d-1} x^{d-1} + \cdots + p_0 \in \mathbb{Z}[x]$ be a (not necessarily monic) polynomial and $\mathcal{Z} := \mathbb{Z}[x]/(P)$ the factor ring. If

each element $\gamma \in \mathcal{Z}$ can be represented as

$$\gamma = \sum_{j=0}^{\ell} a_j X^j \text{ with } a_j \in \mathcal{N} := \{0, 1, \ldots, |p_0| - 1\} \tag{2.1}$$

(where $X \in \mathcal{Z}$ denotes the image of $x \in \mathbb{Z}[x]$ under the canonical epimorphism) then the pair $(P, \mathcal{N})$ is called a *canonical number system* (CNS for short) and P is a *CNS polynomial*.

Kovacs [**128**] remarked that for each (monic, irreducible) polynomial $R \in \mathbb{Z}[x]$ there exists a (sufficiently large) integer a such that the polynomial $R(x - a)$ induces a CNS. In our first main theorem we restate this result in a more general way and give an explicit bound for a.

THEOREM 2.2. *Let $R \in \mathbb{Z}[x]$ and $p, q \in \mathbb{N}$ with $(p, q) = 1$. If*

$$p/q \geq \deg(R) + \max\{\mathrm{Re}(\lambda) : R(\lambda) = 0\}$$

then $P := q^{\deg(R)} R\,(-p/q + x)$ is a CNS-Polynomial.

For a better understanding of the background we want to state several details concerning CNS. Consider a polynomial P and an element $\gamma \in \mathcal{Z}$. We obtain the representation (2.1) by using the backward division algorithm, *i.e.* by successive application of the maps

$$\delta_{\mathcal{N}} : \mathcal{Z} \longrightarrow \mathcal{N}, \gamma \longmapsto a \text{ the unique element of } \mathcal{N} \text{ with } \gamma \equiv a \,(\mathrm{mod}\, X),$$

$$T_P : \mathcal{Z} \longrightarrow \mathcal{Z}, \gamma \longmapsto \frac{\gamma - \delta_{\mathcal{N}}(\gamma)}{X}.$$

In particular, if $T_P^{\ell}(\gamma) = 0$ for some $\ell \in \mathbb{N}$ then we have

$$\gamma = \sum_{j=0}^{\ell} \delta_{\mathcal{N}}(T_P^j(\gamma)) X^j.$$

Thus, the pair $(P, \mathcal{N})$ is a CNS if and only if for each $\gamma \in \mathcal{Z}$ there exists an $\ell \in \mathbb{N}$ such that $T_P^{\ell}(\gamma) = 0$.

It is well known that for P an expanding polynomial (that is, all roots are outside the complex unit circle) the backward division algorithm is eventually periodic for each $\gamma \in \mathcal{Z}$. On the other hand, if there is at least one root inside the unit circle then we can find elements of $\mathcal{Z}$ with infinite T-orbit. Thus, CNS polynomials are necessarily expanding polynomials. For polynomials whose roots have absolute value greater or equal to 1 and equality holds at least once it is up to now not clear whether all T_P-orbits are finite or not. An overview of existing results in this context can be found in the survey [**123**] (from the point of view of SRS).

The ring $\mathcal{Z}$ is a finitely generated free $\mathbb{Z}$-module if and only if P is monic. This fact seems to cause difficulties in the non-monic case. However, the next lemma states that in the context of CNS polynomials the monic and the non-monic case can be treated quite analogously.

LEMMA 2.3 (*cf.* [**211**, Theorem 4.9]). *Let $P = p_d x^d + p_{d-1} x^{d-1} + \cdots + p_0 \in \mathbb{Z}[x]$ and $\mathcal{N} := \{0, \ldots, |p_0| - 1\}$. For each $m \in \{0, \ldots, d-1\}$ let $\omega_m := \sum_{j=0}^{m} p_{d-j} X^j$. Then the following statements are equivalent :*

(1) The pair $(P, \mathcal{N})$ is a canonical number system.

(2) For each $\gamma \in \mathcal{Z}_1$ there exists an ℓ such that $T_P^{\ell}(\gamma) = 0$ where $\mathcal{Z}_1$ is the free $\mathbb{Z}$-module generated by $\{1, X, \ldots, X^{d-1}\}$.

(3) For each $\gamma \in \mathcal{Z}_2$ there exists an ℓ such that $T_P^\ell(\gamma) = 0$ where $\mathcal{Z}_2$ is the free $\mathbb{Z}$-module generated by $\{\omega_0, \ldots, \omega_{d-1}\}$ (this basis is sometimes referred to as Brunotte basis).

Observe that $\mathcal{Z}_2 \subseteq \mathcal{Z}_1 \subseteq \mathcal{Z}$ where equality holds if and only if P is monic. Furthermore, $\mathcal{Z}_1$ as well as $\mathcal{Z}_2$ are closed with respect to the application of T_P.

Our most important tool for proving Theorem 2.2 will be the following monotonicity condition which appears in several articles on CNS in slightly different formulations and contexts (*cf.* [**54**]). The present form is, in fact, a non-monic version of [**54**, Corollary 7] which follows immediately from [**211**, Theorem 5.3] and [**5**, Theorem 3.5].

PROPOSITION 2.4. *Let $P = p_d x^d + p_{d-1}x^{d-1} + \cdots + p_0 \in \mathbb{Z}[x]$. If $p_0 \geq 2$ and $0 < p_d \leq p_{d-1} \leq \cdots \leq p_1 < p_0$ then P is a CNS polynomial.*

Note that the bound provided in Theorem 2.2 ensures that P satisfies the condition of this proposition. But observe that Theorem 2.2 is only a sufficiency result, that is, the monotonicity condition is fulfilled even for smaller rationals p/q. However, without additional assumptions (for examples on the roots of R) it seems to be difficult to improve this bound (see also [**154**, Remark 1]). Furthermore we want to note that Proposition 2.4 itself provides sufficient conditions on polynomials in order to induce a CNS, that is a polynomial P can be a CNS-polynomial even when P does not satisfy the monotonicity condition. We want to summarise these considerations in the following result (which is actually a corollary to Theorem 2.2).

PROPOSITION 2.5. *Let $P \in \mathbb{Z}[x]$ such that for each root λ we have $\mathrm{Re}(\lambda) \leq -\deg(P)$ ($\mathrm{Re}(\lambda) < -\deg(P)$ if P does not have at least 2 distinct roots). Then P satisfies the conditions of Proposition 2.4, especially, P induces a CNS.*

Finally we should mention that for a given polynomial $R \in \mathbb{Z}[x]$ it is absolutely not clear that there exists a bound $K \in \mathbb{R}$ such that $q^{\deg(R)}R\left(-p/q + x\right)$ induces a CNS for $p/q > K$ while for $p/q < K$ it does not.

Let $k \in \mathbb{N}$, denote by ζ a primitive kth root of unity, and by $\Phi_k \in \mathbb{Z}[x]$ its minimal polynomial (the kth cyclotomic polynomial). Then Theorem 2.2 immediately yields that $\Phi_k(-a+x)$ is a CNS-polynomial for an integer $a \geq \varphi(k) + 1$ (*cf.* [**154**, Theorem 1.1]).

In view of Bremner's conjecture we are also interested in the two other power integral bases induced by

$$\eta = \eta(\zeta) := (1+\zeta)^{-1} \quad \text{and} \quad \theta = \theta(\zeta) := (1-\zeta)^{-1}.$$

Since $\mathrm{Re}(\eta) = \mathrm{Re}(\theta) = 1/2$ (*cf.* Lemma 2.13) we obtain as corollary to Theorem 2.2.

COROLLARY 2.6. *Let $k \in \mathbb{N}$ (with $k \geq 3$) and $a \in \mathbb{Q}$. If $a \geq \varphi(k) + 1/2$ then the minimal polynomials of $-a + \eta$ and $-a + \theta$ are CNS polynomials. If $a \geq \varphi(k) - 1/2$ then the minimal polynomials of $-a - \eta$ and $-a - \theta$ are CNS polynomials.*

REMARK 1. Observe that for odd k the kth cyclotomic field equals the $(2k)$th one. More precisely, we have $\Phi_{2k}(x) = \Phi_k(-x)$. On the other hand, if $k \equiv 0 \pmod 4$, then for each primitive kth root of unity ζ, $-\zeta$ is also a primitive kth root of unity. Finally observe that $\theta(\zeta) = \eta(-\zeta)$ for each primitive kth root of unity ζ. Thus, the results up to now already give us a quite good overview of bases for number systems in the ring of integers of the kth cyclotomic field, namely, $-a \pm \zeta$, $-a + \eta(\zeta)$ and $-a + \theta(\zeta)$ for integers $a \geq \varphi(k) + 1$ as well as $-a - \eta(\zeta)$ and $-a - \theta(\zeta)$ for integers $a \geq \varphi(k)$.

Now let us turn to conditions for the multiplicative independence of two bases.

DÉFINITION 2.2. Let α and β be two algebraic numbers. Then we call α and β multiplicatively independent if $(p, p') = (0, 0)$ is the only pair of integers that solves the equation $\alpha^p = \beta^{p'}$.

´ Note that if α and β are algebraic integers or $|\alpha|, |\beta| > 1$ or $|\alpha|, |\beta| < 1$, then it is sufficient to restrict to pairs of non-negative integers (p, p') that solve $\alpha^p = \beta^{p'}$.

In [**148**, **154**] the authors ask whether $a + \zeta$ and $a' + \zeta$ are multiplicatively independent numbers for negative integers $a \neq a'$ and ζ a primitive kth root of unity. In fact, the two papers already contain partial results concerning this problem. The following considerations complete this question. Actually, we are able to show much more with our idea.

THEOREM 2.7. *Let ζ be a root of the kth cyclotomic polynomial $\Phi_k(x)$ with $k \notin \{1, 2, 3, 4, 6\}$ and $a, a' \in \mathbb{Q}$ rational numbers with $a' < a$ such that*

$$|a + \zeta|^p = |a' + \zeta|^{p'} \tag{2.2}$$

holds for a pair of non-negative integers $(p, p') \neq (0, 0)$. Then one of the following conditions is necessarily satisfied.

(i) $a' < a < -\delta_2$ *and* $aa' < 1$;

(ii) $-\delta_2 < a' < a < 0$ *and* $a + a' > -\delta_2$;

(iii) $0 < a' < a < -\delta_1$ *and* $a + a' < -\delta_1$;

(iv) $-\delta_1 < a' < a$ *and* $aa' < 1$,

where

$$\delta_1 := \begin{cases} 2\cos(^{(k-1)\pi}/_k) & \text{if } k \equiv 1 \pmod 2 \\ 2\cos(^{(k-2)\pi}/_k) & \text{if } k \equiv 0 \pmod 4 \\ 2\cos(^{(k-4)\pi}/_k) & \text{if } k \equiv 2 \pmod 4 \end{cases} < 0,$$

$$\delta_2 := 2\cos(^{2\pi}/_k) > 0.$$

This Theorem allows several conclusions. In particular we want to note that :

— The multiplicative independence of $|a+\zeta|$ and $|a'+\zeta|$ implies that of $a+\zeta$ and $a'+\zeta$.

— Multiplicative independence follows as soon as we have that $|a + \zeta|$ and $|a' + \zeta|$ are both larger than 1.

— For a, a' integers with $a, a' \neq 0$ the numbers $a+\zeta$ and $a'+\zeta$ are always multiplicatively independent (which answers the initial question – the excluded cases $2, 3, 4, 6$ are treated in [**148**, **154**]).

— For $k \neq 5$ the numbers $|a + \zeta|$ and $|a' + \zeta|$ are multiplicatively independent as soon as a and a' are smaller than $-\delta_2$ since the first case can only occur for $k = 5$.

— For $k \neq 10$ the numbers $|a + \zeta|$ and $|a' + \zeta|$ are multiplicatively independent as soon as a and a' are larger than $-\delta_1$ since the last case can only occur for $k = 10$.

Finally, we want to remark that it would be interesting whether there exist a pair of rational $(a, a') \neq (0, 0)$ (that necessarily satisfy one of the conditions ((i))–((iv))) such that $a + \zeta$ are $a' + \zeta$ multiplicatively dependent.

Now we are interested in a similar result for $\eta(\zeta)$ and $\theta(\zeta)$. The situation is easier due to the fact that all Galois conjugates have real part $a + 1/2$ and $a' + 1/2$, respectively (see Lemma 2.13 in Section 4).

THEOREM 2.8. *Let k be a positive integer with $k \notin \{1,2,3,4,6\}$, ζ a root of Φ_k, $\theta = (1-\zeta)^{-1}$, and $a, a' \in \mathbb{Q}$ rational numbers. Then $|a+\theta|$ and $|a'+\theta|$ are multiplicatively independent provided that*

(i) $a + a' + 1 \neq 0$ and

(ii) if $k = 5$, then we assume that $(2a+1)^2 > 3 + \frac{2\sqrt{5}}{5}$ and $(2a'+1)^2 > 3 + \frac{2\sqrt{5}}{5}$.

Due to the close relation of $\eta(\zeta)$ and $\theta(\zeta)$ (see Remark 1) an analogue result for $\eta(\zeta)$ follows immediately.

COROLLARY 2.9. *Let k be a positive integer with $k \notin \{1,2,3,4,6\}$, ζ a root of Φ_k, and $a, a' \in \mathbb{Q}$ rational numbers. Then $|a+\eta|$ and $|a'+\eta|$ are multiplicatively independent provided that*

(i) $a + a' + 1 \neq 0$ and

(ii) if $k = 10$, then we assume that $(2a+1)^2 > 3 + \frac{2\sqrt{5}}{5}$ and $(2a'+1)^2 > 3 + \frac{2\sqrt{5}}{5}$.

Observe that in both cases the assumption that a, a' are non-zero integers ensures multiplicative independence. As before, the multiplicative independence of $|a+\theta|$ and $|a'+\theta|$ ($|a+\eta|$ and $|a'+\eta|$, respectively) immediately implies the multiplicative independence of $a+\theta$ and $a'+\theta$ ($a+\eta$ and $a'+\eta$, respectively).

Observe that in both cases integers a, a' with $a, a' \neq 0$ ensure multiplicative independence. As before, the multiplicative independence of $|a+\theta|$ and $|a'+\theta|$ ($|a+\eta|$ and $|a'+\eta|$, respectively) immediately implies the multiplicative independence of $a+\theta$ and $a'+\theta$ ($a+\eta$ and $a'+\eta$, respectively).

REMARK 2. Let ζ be a primitive kth root of unity and remind that our initial purpose was to find multiplicative independent bases of number systems in $\mathbb{Z}[\zeta]$ (see Remark 1). From Theorem 2.7 we see that numbers of the shape $a+\zeta$ and $a'+\zeta$ with integers a, a' such that $a, a' \neq 0$ are multiplicative independent. Now observe that $a+\zeta$ and $-a-\zeta$ are clearly multiplicatively dependent and multiplicative dependence is a transitive property. Therefore, all bases of the form $a \pm \zeta$ are pairwise multiplicatively independent since $a = 0$ does not yield a base of a number system.

With the same argumentation we see that all bases of the shape $a \pm \theta(\zeta)$ are pairwise multiplicative independent (where $a \neq 0$ is an integer). The same holds for all bases of the shape $a \pm \eta(\zeta)$.

For completeness it would be an interesting whether numbers $a+\zeta$, $a'+\eta(\zeta)$ and $a''+\theta(\zeta)$ are multiplicatively independent. We also do not know anything about the multiplicative independence of $a+\zeta$ and $a'+\zeta'$ for two different primitive kth roots of unity ζ and ζ' (and, in an analogous way, for the corresponding values θ and η).

With these infinite families of multiplicative independent bases we want to turn to the last part and show a result analogue to Cobham's theorem for number systems in the ring of integers $\mathbb{Z}[\zeta]$ where ζ is a primitive kth root of unity.

Consider two multiplicatively independent bases α and β. A main ingredient in the proof of Cobham's theorem is that the set of all numbers of the form $\alpha^m \beta^{-n}$ with $m, n \in \mathbb{N}$ lie dense in $\mathbb{C}$. For real α and β it is easy to show that the corresponding result holds true. However, in the complex case we do not have such a result and we are not even close to one. We circumvent this issue by using the four exponentials conjecture to obtain the desired density result.

CONJECTURE 2.1 (Four exponentials conjecture). *Let x_1, x_2 and y_1, y_2 be two pairs of complex numbers such that each pair is linearly independent over $\mathbb{Q}$. Then at least one of the four numbers*

$$e^{x_1y_1}, \quad e^{x_1y_2}, \quad e^{x_2y_1}, \quad e^{x_2y_2}$$

is transcendental.

For a detailed account to the four exponentials conjectjure we refer the reader to the book of Waldschmidt [**250**, Chapter 1.3 resp. Chapter 11].

THEOREM 2.10. *Let $\alpha, \beta \in \mathbb{C} \setminus \{0\}$ algebraic numbers such that $|\alpha|$ and $|\beta|$ are multiplicatively independent. If the four exponentials conjecture is true then the set*

$$P_{\alpha,\beta} := \left\{ \frac{\alpha^m}{\beta^n} \colon m, n \in \mathbb{N} \right\}$$

is dense in $\mathbb{C}$.

Theorem 2.10 was proved by Hansel and Safer [**101**] in the special case that α and β are of the form $\alpha = -a + i$ and $\beta = -a' + i$, with a an a' positive rational integers.

Since transferring our density conjecture one obtains something very similar to the four exponentials conjecture it seems that they are very close or even equivalent. It is tempting to use the six exponentials theorem to obtain a variant of Cobham's theorem for more than two bases, but currently we have no idea how to do that.

As α is a basis of a number system in $\mathbb{Z}[\zeta]$ (with digit set $\mathcal{N}$), each element $z \in \mathbb{Z}[\zeta]$ has a unique representation of the form

$$z = \sum_{j=0}^{\ell} a_j \alpha^j \quad \text{with } a_j \in \mathcal{N} \text{ for } 0 \le j \le \ell \text{ and } a_\ell \neq 0.$$

We denote by $\rho_\alpha(z)$ the corresponding digit string over the alphabet $\mathcal{N}$, that is

$$\rho_\alpha(z) := a_0 a_1 \dots a_\ell \in \mathcal{N}^*.$$

For a set $S \subset \mathbb{Z}[\zeta]$ we define the language $\rho_\alpha(S)$ by

$$\rho_\alpha(S) := \{\rho_\alpha(x) \colon x \in S\}$$

We call S α-recognizable if the language $\rho_\alpha(S)$ is recognizable by a finite automaton. Moreover let α and β be two multiplicatively independent bases of canonical number systems in $\mathbb{Z}[\zeta]$. Then we call a set (α, β)-recognizable if it is α-recognizable and β-recognizable.

Our final result is a weaker form of Cobham's theorem. In particular, we do not show, that the set is ultimately periodic but syndetic. For better understanding this terminology let S be a subset of the positive integers $\mathbb{N}$. Then we call S syndetic (or with bounded gaps) if there exists $r \in \mathbb{N}$ such that $S \cap [n, n + r] \neq \emptyset$ for each $n \in \mathbb{N}$. The analog for a lattice Λ in $\mathbb{C}$ is that for a given set $S \subset \Lambda$ there exists $r \in \mathbb{R}$ such that $S \cap B(n, r) \neq \emptyset$ for each $n \in \Lambda$, where $B(n, r)$ is the closed disc with center n and radius r.

THEOREM 2.11. *Let ζ be a primitive kth root of unity and consider two multiplicatively independent bases α and β for digit systems in $\mathbb{Z}[\zeta]$ such that $P_{\alpha,\beta}$ is dense in $\mathbb{C}$. If S is an infinite (α, β)-recognizable subset of $\mathbb{Z}[\zeta]$, then S is syndetic.*

3. Proof of Theorem 2.2

The following lemma estimates the coefficients of a polynomial if we shift the center from 0 to a.

LEMMA 2.12. *Let $R(x) \in \mathbb{Z}[x]$ and $a = \frac{p}{q} \in \mathbb{Q}$. Then*

$$R(x-a) = \sum_{n=0}^{\deg(R)} \frac{R^{(n)}(-a)x^n}{n!} \in \mathbb{Q}[x]$$

(where $R^{(n)}$ denotes the nth derivation of $R = R^{(0)}$). If R is the minimal polynomial of an algebraic $z \in \mathbb{C}$ then $P(x) := q^{\deg(R)}R(x-a)$ is the minimal polynomial of $z+a$.

DÉMONSTRATION. The first assertion is clear by Taylor's theorem. The second part follows from the observation that P is an integer polynomial, $P(\hat{z}+a) = q^{\deg(R)}R(\hat{z}) = 0$ for each Galois conjugate $\hat{z}$ of z, and the degree of P coincides with the degree of the minimal polynomial of $z-a$. □

PROOF OF THEOREM 2.2. Let $a := -p/q$. By assumption we have

$$a \le -\deg(R) - \max\{\mathrm{Re}(\lambda) : R(\lambda) = 0\}.$$

We will show that

$$\frac{R^{(n+1)}(-a)n!}{R^{(n)}(-a)(n+1)!} = \frac{R^{(n+1)}(-a)}{R^{(n)}(-a)(n+1)} < 1$$

holds for all $n \in \{0, \dots, \deg(k)-1\}$. Then, by Lemma 2.12 and Proposition 2.4, $P(x) = q^{\deg(k)}R(a+x)$ induces a CNS.

We first show the case $n = 0$. Denote by $\xi_1, \dots, \xi_t$ the t (not necessarily distinct) real roots and by $u_1 \pm iv_1, \dots, u_s \pm iv_s$ the s (not necessarily distinct) pairs of complex conjugate roots of R (that is, $2s+t = \deg(R)$). Thus, $R(x) = \prod_{j=1}^{s}(x^2 - 2u_jx + u_j^2 + v_j^2) \cdot \prod_{j=1}^{t}(x - \xi_j)$. From this we easily obtain the logarithmic derivative

$$\begin{aligned}\frac{R'(-a)}{R(-a)} &= \sum_{j=1}^{s} \frac{-2a-2u_j}{a^2+2u_ja+u_j^2+v_j^2} + \sum_{j=1}^{t}\frac{1}{-a-\xi_j} \\ &\le \sum_{j=1}^{s}\frac{2(-a-u_j)}{a^2+2u_ja+u_j^2} + \sum_{j=1}^{t}\frac{1}{-a-\xi_j} = \sum_{j=1}^{s}\frac{2}{-a-u_j} + \sum_{j=1}^{t}\frac{1}{-a-\xi_j}.\end{aligned} \tag{3.1}$$

Now observe that by the choice of a we have $-a-\xi_j \ge \deg(R)$ for all $j \in \{1, \dots, t\}$ and $-a-u_j \ge \deg(R)$ for all $j \in \{1, \dots, s\}$. Therefore we obtain

$$\sum_{j=1}^{s}\frac{2}{-a-u_j} + \sum_{j=1}^{t}\frac{1}{-a-\xi_j} \le \sum_{j=1}^{s}\frac{2}{\deg(R)} + \sum_{j=1}^{t}\frac{1}{\deg(R)} = \frac{2s+t}{\deg(R)} = 1. \tag{3.2}$$

Note that the inequality in (3.1) is sharp when $s \neq 1$ (hence, R has non-real roots) while the inequality in (3.2) is sharp when at least two real parts are different. Thus, $\frac{R'(-a)}{R(-a)} < 1$ provided that R has at least 2 different roots.

For $n \ge 1$ the situation is less critical. Again we denote by $\xi_1, \dots, \xi_t$ the real roots of $R^{(n)}$ and by $u_1 \pm iv_1, \dots, u_s \pm iv_s$ the pairs of complex conjugate roots (again, $2s+t = \deg(R)-n$).

As before we consider the logarithmic derivative and obtain analogously

$$\frac{R^{(n+1)}(-a)}{R^{(n)}(-a)} \leq \sum_{j=1}^{s} \frac{2}{-a-u_j} + \sum_{j=1}^{t} \frac{1}{-a-\xi_j}.$$

By the Gauss-Lucas theorem, the roots of $R^{(n)}(x)$ are contained in the convex hull of the roots of $R^{(n-1)}(x)$. Hence, $\max\{\mathrm{Re}(\lambda) : R(\lambda)=0\} \geq \max\{\mathrm{Re}(\lambda) : R^{(n)}(\lambda)=0\}$. Similarly as above this immediately yields

$$\frac{R^{(n+1)}(-a)}{R^{(n)}(-a)(n+1)} < \frac{R^{(n+1)}(-a)}{R^{(n)}(-a)} \leq \frac{2s+t}{\deg(R)} = \frac{\deg(R)-n}{\deg(R)} < 1.$$

□

REMARK 3. Observe that if R does not have at least 2 different roots (that is, $R(x) = (x-\xi)^n$ for some $n \in \mathbb{N}$) then Theorem 2.2 holds if we require

$$p/q > \deg(R) + \max\{\mathrm{Re}(\lambda) : R(\lambda)=0\} (= \deg(R) + \xi).$$

4. Multiplicative independent bases

In this section we want to collect the proofs of the results concerning multiplicative independence.

PROOF OF THEOREM 2.7. It is obvious that $|a+\zeta|$ and $|a'+\zeta|$ are multiplicatively independent if and only if $|a+\zeta|^2$ and $|a'+\zeta|^2$ are multiplicatively independent. Therefore we may concentrate on $|a+\zeta|^2$ and $|a'+\zeta|^2$, respectively. In particular, we prove the theorem by showing that if a pair a, a' of rational numbers does not satisfy one of the conditions ((i))–((iv)), then

$$|a+\zeta|^{2p} = |a'+\zeta|^{2p'} \tag{4.1}$$

cannot hold for a pair $(p,p') \neq (0,0)$ of non-negative integers.

At first we claim that $aa' = 0$ contradicts (4.1). Indeed, suppose that one of the both rationals, say a, were 0. Then $|a+\zeta|^2 = 1$ and $|a'+\zeta|^2 = a'^2 + 1 + 2a\mathrm{Re}(\zeta)$. Since $\mathrm{Re}(\zeta) \notin \mathbb{Q}$ for the considered values of k, (4.1) cannot hold.

Before we continue we need some further considerations. Denote by ζ_1 and ζ_2, respectively, two roots of $\Phi_k(x)$ such that

$$\mathrm{Re}(\zeta_1) \leq \mathrm{Re}\left(\hat{\zeta}\right) \leq \mathrm{Re}(\zeta_2)$$

holds for each root $\hat{\zeta}$ of $\Phi_k(x)$. Since the degree of $\Phi_k(x)$ is at least 4 we clearly have $\delta_1/2 = \mathrm{Re}(\zeta_1) < 0 < \mathrm{Re}(\zeta_2) = \delta_2/2$. Define for each $i \in \{1,2\}$

$$\begin{aligned} \alpha_i &:= |a+\zeta_i|^2 = a^2 + 2a\mathrm{Re}(\zeta_i) + 1, \\ \alpha_i' &:= \left|a'+\zeta_i\right|^2 = a'^2 + 2a'\mathrm{Re}(\zeta_i) + 1. \end{aligned} \tag{4.2}$$

Observe that $|a+\zeta|^2$, α_1 and α_2 as well as $|a'+\zeta|^2$, α_1' and α_2' are algebraically conjugate, hence, from (4.1) we conclude that for $i \in \{1,2\}$

$$\alpha_i^p = \alpha_i'^{p'}. \tag{4.3}$$

holds. Furthermore, this clearly implies

$$(\alpha_1/\alpha_2)^p = (\alpha_1'/\alpha_2')^{p'}. \tag{4.4}$$

With this we show that our initial assumption (4.1) cannot hold if a and a' have different signs. Indeed, suppose that $a' < 0 < a$. From (4.2) we easily see that $\alpha_1 < \alpha_2$, hence, $\alpha_1/\alpha_2 < 1$. On the other hand we have $\alpha_1' > \alpha_2'$. Thus, $\alpha_1'/\alpha_2' > 1$. This is a contradiction to (4.4) and, hence, (4.1).

Up to now we have shown that (4.1) implies either $0 < a' < a$ or $a' < a < 0$. We first concentrate on the latter case. We obviously have $0 < \alpha_2 < \alpha_1$ as well as $0 < \alpha_2' < \alpha_1'$. Furthermore, we clearly have $1 < \alpha_1 < \alpha_1'$. Thus, (4.3) implies

$$p > p'. \tag{4.5}$$

Now we distinguish three cases.

Case 1 : $-\delta_2 < a' < a < 0$**:** In this case we have $\alpha_2, \alpha_2' < 1$. Suppose that $a + a' < -2\mathrm{Re}(\zeta_2)$. Since $(a - a') > 0$ this implies $(a - a')(a + a') + 2(a - a')\mathrm{Re}(\zeta_2) < 0$ and, hence

$$\alpha_2 = a^2 + 1 + 2a\mathrm{Re}(\zeta_2) < a'^2 + 1 + 2a'\mathrm{Re}(\zeta_2) = \alpha_2' < 1$$

which contradicts (4.3) since we have $p > p'$. Thus, we necessarily have $a + a' > -2\mathrm{Re}(\zeta_2) = -\delta_2$ (since a and a' are rational numbers, we always have $a + a' \neq -2\mathrm{Re}(\zeta_2)$ for the possible values of n).

Case 2 : $a' < -\delta_2 < a < 0$**:** We see that this yields $\alpha_2 < 1$ and $\alpha_2' > 1$ which contradicts (4.3).

Case 3 : $a' < a < -\delta_2$**:** We obviously have $\alpha_2 < \alpha_2'$. We calculate

$$\begin{aligned}
\frac{\alpha_1}{\alpha_2} - \frac{\alpha_1'}{\alpha_2'} &= \frac{\alpha_1\alpha_2' - \alpha_1'\alpha_2}{\alpha_2\alpha_2'} \\
&= \frac{(a^2+1)2a'\mathrm{Re}(\zeta_2) + (a'^2+1)2a\mathrm{Re}(\zeta_1) - (a'^2+1)2a\mathrm{Re}(\zeta_2) - (a^2+1)2a'\mathrm{Re}(\zeta_1)}{\alpha_2\alpha_2'} \\
&= \frac{(a^2a' + a' - a'^2a - a)2\mathrm{Re}(\zeta_2) + (a'^2a + a - a^2a' - a')2\mathrm{Re}(\zeta_1)}{\alpha_2\alpha_2'} \\
&= \frac{(aa'-1)(a-a')2\mathrm{Re}(\zeta_2) + (aa'-1)(a'-a)2\mathrm{Re}(\zeta_1)}{\alpha_2\alpha_2'} \\
&= \frac{(aa'-1)(a-a')(2\mathrm{Re}(\zeta_2) - 2\mathrm{Re}(\zeta_1))}{\alpha_2\alpha_2'} = \frac{1}{\alpha_2\alpha_2'}(aa'-1)(a-a')(\delta_2 - \delta_1).
\end{aligned}$$

For $aa' > 1$ the latter expression is strictly positive, hence, $1 < \frac{\alpha_1'}{\alpha_2'} < \frac{\alpha_1}{\alpha_2}$. This violates (4.4) since $p > p'$.

By exploiting symmetries the case $0 < a' < a$ runs analogously. □

Now we consider two bases of the form $a + \theta$ and $a' + \theta$. We start with the following basic lemma which will be useful in the sequel.

LEMMA 2.13. *Let $z \in \mathbb{C} \setminus \{1\}$ with $|z| = 1$ and $a \in \mathbb{R}$. Then*

$$\mathrm{Re}\left(a + (1-z)^{-1}\right) = a + \frac{1}{2}, \tag{4.6}$$

$$\mathrm{Im}\left(a + (1-z)^{-1}\right) = \frac{\mathrm{Im}(\zeta)}{2(1-\mathrm{Re}(z))}, \tag{4.7}$$

$$\left|a + (1-z)^{-1}\right|^2 = a^2 + a + \frac{1}{2(1-\mathrm{Re}(z))}. \tag{4.8}$$

DÉMONSTRATION. Easy exercise. □

PROOF OF THEOREM 2.8. The idea of the proof is essentially the same as that of Theorem 2.7. We concentrate on $|a+\theta|^2$ and $|a'+\theta|^2$ and we show the assertion indirectly, thus, we suppose that $|a+\theta|^2$ and $|a'+\theta|^2$ were not multiplicatively independent. Then there exists a pair of non-negative integers $(p, p') \neq (0, 0)$ such that

$$|a+\theta|^{2p} = \left|a'+\theta\right|^{2p'}. \tag{4.9}$$

At first we claim that under our conditions we always have $p \neq p'$. Indeed, $p = p'$ implies that $|a+\theta|^2 = |a'+\theta|^2$. By observing Lemma 2.13 this is possible if and only if we either have the trivial case $a = a'$ or the excluded one $a + a' + 1 = 0$.

Thus, we can suppose that $p \neq p'$ and, without loss of generality, we may assume that $0 < p < p'$. By the assumption on k there exist roots of $\Phi_k(x)$ with different real parts. We let ζ_1 be a root with the minimal real part and ζ_2 a root with the maximal real part, hence $\mathrm{Re}(\zeta_1) = {}^{\delta_1}/_2 < 0 < {}^{\delta_2}/_2 = \mathrm{Re}(\zeta_2)$ where δ_1 and δ_2 are defined in the statement of Theorem 2.7. We define for each $j \in \{1, 2\}$

$$\alpha_j := \left|a + (1-\zeta_j)^{-1}\right|^2 = a^2 + a + \frac{1}{2(1-\mathrm{Re}(\zeta_j))} > 0,$$

$$\alpha'_j := \left|a' + (1-\zeta_j)^{-1}\right|^2 = a'^2 + a' + \frac{1}{2(1-\mathrm{Re}(\zeta_j))} > 0,$$

(where we used the results of Lemma 2.13) and note that, $|a+\theta|^2$, α_1 and α_2 (and $|a'+\theta|^2$, α'_1 and α'_2, respectively) are algebraic conjugates. Therefore, from (4.9) follows that $\alpha_j^p = \alpha_j'^{p'}$ holds for each $j \in \{1, 2\}$. From the assumption $p < p'$ we deduce that either $\alpha_j < \alpha'_j < 1$ or $\alpha_j > \alpha'_j > 1$ holds for both, $j = 1$ as well as $j = 2$.

We claim that the first case cannot occur within the terms on the theorem. At first we observe that $\alpha'_2 < 1$ only if

$$1 > \mathrm{Im}(a' + (1-\zeta_2)^{-1})^2 = \frac{\mathrm{Im}(\zeta_2)^2}{4(1-\mathrm{Re}(\zeta_2))^2} = \frac{1+\mathrm{Re}(\zeta_2)}{4(1-\mathrm{Re}(\zeta_2))}.$$

This immediately yields the condition $\mathrm{Re}(\zeta_2) = \frac{\delta_2}{2} < {}^3/_5$ which is satisfied in the case $k = 5$ only. Now, if $k = 5$, then $\alpha_2 < 1$ as well as $\alpha'_2 < 1$ must hold. By using (4.8) and the well-known identity $\mathrm{Re}(\zeta_2) = \cos({}^{2\pi}/_5) = {}^{(\sqrt{5}-1)}/_4$ one readily verifies that this implies that $(2a+1)^2 < 3 + \frac{2\sqrt{5}}{5}$ and $(2a'+1)^2 < 3 + \frac{2\sqrt{5}}{5}$.

Thus we may concentrate on the case that $\alpha_j > \alpha'_j > 1$ holds for $j = 1$ as well as $j = 2$, which immediately implies

$$0 > \alpha'_j - \alpha_j = a'^2 - a^2 + a' - a \tag{4.10}$$

for $j \in \{1,2\}$. Now observe that from (4.9) we also obtain that

$$(\alpha_2\alpha_1^{-1})^p = (\alpha_2'\alpha_1'^{-1})^{p'}. \tag{4.11}$$

must hold. We compute

$$\alpha_2 - \alpha_1 = \frac{1}{2(1-\mathrm{Re}(\zeta_2))} - \frac{1}{2(1-\mathrm{Re}(\zeta_1))} > 0$$

which shows that $\alpha_2\alpha_1^{-1} > 1$, and we estimate

$$\begin{aligned}\alpha_2\alpha_1' - \alpha_1\alpha_2' =& \left(a'^2 + a' - a^2 - a\right)\frac{1}{2(1-\mathrm{Re}(\zeta_2))} + \left(a^2 + a - a'^2 - a'\right)\frac{1}{2(1-\mathrm{Re}(\zeta_1))}\\ =& \frac{1}{2}\left(a'^2 + a' - a^2 - a\right)\left(\frac{1}{1-\mathrm{Re}(\zeta_2)} - \frac{1}{1-\mathrm{Re}(\zeta_1)}\right) < 0,\end{aligned}$$

where we used (4.10) for the last inequality. Thus, $1 < \alpha_2\alpha_1^{-1} < \alpha_2'\alpha_1'^{-1}$. This contradicts (4.11) since we assumed $p < p'$. □

Remark 4. We have seen that when $a + a' + 1 = 0$, then we possibly have $|a+\theta|^p = |a'+\theta|^p$ for a positive integer p. The question remains whether this is possible for distinct $a, a' \in \mathbb{Q}$? The methods from above do not seem to work here.

5. Complex Bases and density properties

This short section is devoted to the proof of Theorem 2.10. Throughout the section we suppose that $\alpha = ae^{i\psi}$ and $\beta = be^{i\omega}$ are bases of number systems in $\mathbb{Z}[\zeta]$ (with ζ a primitive kth root of unity) such that $a = |\alpha|$ and $b = |\beta|$ are multiplicatively independent. With these notations at hand the proof of Theorem 2.10 is an immediate consequence of the following two lemmas.

Lemma 2.14 ([**101**, Lemme 1]). *The set*

$$P_{\alpha,\beta} = \left\{ \frac{\alpha^m}{\beta^n} \colon m, n \in \mathbb{N} \right\}$$

is dense in $\mathbb{C}$*, if*

$$\frac{\log b}{\log a}, \quad \frac{\psi}{2\pi}\frac{\log b}{\log a} - \frac{\omega}{2\pi}, \quad 1$$

are linearly independent over $\mathbb{Q}$*.*

Lemma 2.15 ([**101**, Lemme 2]). *If the four exponentials conjecture, Conjecture 2.1, holds, then*

$$\frac{\log b}{\log a}, \quad \frac{\psi}{2\pi}\frac{\log b}{\log a} - \frac{\omega}{2\pi}, \quad 1$$

are linearly independent over $\mathbb{Q}$*.*

Recall that a set $S \subset \mathbb{Z}[\zeta]$ is α-recognizable if the set of representations $\rho_\alpha(S) = \{\rho_\alpha(x) \colon x \in S\}$ is recognizable in $\mathcal{N}^*$. Using the Nerode equivalence (*cf.* Sakarovitch [**202**]) this means that a set S is α-recognizable if and only if the equivalence relation $\mathbb{Z}[\zeta]_S^\alpha$ on $\mathcal{N}^*$ defined by

$$u\mathbb{Z}[\zeta]_S^\alpha v :\Leftrightarrow (\forall w \in \mathcal{N}^* \colon uw \in \rho_\alpha(S) \Leftrightarrow vw \in \rho_\alpha(S))$$

is of finite index (*cf.* Proposition 9.3.3 of [**73**]).

Since ρ_α is a bijection from $\mathbb{Z}[\zeta] \setminus \{0\}$ into $(\mathcal{N} \setminus \{0\})\mathcal{N}^*$, we can pull this definition back to $\mathbb{Z}[\zeta]$. In particular, a set $S \subset \mathbb{Z}[\zeta]$ is α-recognizable if the equivalence relation $\mathbb{Z}[\zeta]_S^\alpha$ on $\mathbb{Z}[\zeta]$ defined by

$$x\mathbb{Z}[\zeta]_S^\alpha y \Leftrightarrow (\forall w \in \mathcal{N}^* \colon \rho_\alpha(x)w \in \rho_\alpha(S) \Leftrightarrow \rho_\alpha(y)w \in \rho_\alpha(S))$$

is of finite index.

For an $s \in \mathbb{N}$ denote by $\mathbb{Z}[\zeta]_s$ the subset of elements of $\mathbb{Z}[\zeta]$ whose expansion with respect to the base α has at most length s, *i.e.*

$$\mathbb{Z}[\zeta]_s := \left\{ \sum_{j=0}^{s-1} a_j \alpha^j \colon a_j \in \mathcal{N} \text{ for } 0 \le j < s \right\}.$$

PROPOSITION 2.16. *Let S' be an infinite equivalence class of the relation $\mathbb{Z}[\zeta]_S^\alpha$ and $s \in \mathbb{N}^*$. There is a finite set of positive integers F_s such that for each $x \in \mathbb{Z}[\zeta]$, there exists $\lambda \in F_s$ and $z \in \mathbb{Z}[\zeta]_{\lambda s}$ such that*

$$x\alpha^{\lambda s} + z \in S'.$$

DÉMONSTRATION. This is Proposition 8 of Hansel and Safer [**101**] replacing the estimate for the length of expansion by the corresponding one for canonical number systems due to Kovács and Pethő [**130**]. □

PROPOSITION 2.17 ([**101**, Proposition 9]). *Let S' be an infinite equivalence class of the relation $\mathbb{Z}[\zeta]_S^\alpha$. There exist $s \in \mathbb{N}$ and $z \in \mathbb{Z}[\zeta]_s$ such that $S'\alpha^s + z \subset S'$.*

Now we have all the tools needed for the

PROOF OF THEOREM 2.11. Let S' be one of the infinite equivalence classes of $\mathbb{Z}[\zeta]_S^\alpha$ contained in S. It suffices to show the theorem for S'.

An application of Proposition 2.17 yields the existence of $s \in \mathbb{N}$ and $z \in \mathbb{Z}[\zeta]_s$ such that

$$S'\alpha^s + z \subset S'. \tag{5.1}$$

Let F_s be the finite set whose existence is guaranteed by an application of Proposition 2.16 and set $\overline{\lambda} = \sup F_s$. Furthermore let $x \in \mathbb{Z}[\zeta]$. Then there exists $\lambda \in F_s$ and $z_0 \in \mathbb{Z}[\zeta]_{\lambda s}$ such that

$$x\alpha^{\lambda s} + z_0 \in S'. \tag{5.2}$$

Now we recursively show for each $n \in \mathbb{N}$ the existence of $z_n \in \mathbb{Z}[\zeta]_{(\lambda+n)s}$ such that

$$x\alpha^{(\lambda+n)s} + z_n \in S'. \tag{5.3}$$

For $n = 0$ the existence follows from (5.2). Now suppose that we already found $z_0, z_1, \ldots, z_{n-1}$ satisfying (5.3). Then it follows from (5.1) that

$$\left(x\alpha^{(\lambda+(n-1)s)} + z_{n-1}\right)\alpha^s + z \in S'.$$

Since $z_{n-1} \in \mathbb{Z}[\zeta]_{(\lambda+(n-1))s}$ and $z \in \mathbb{Z}[\zeta]_s$, we have that $z_{n-1}\alpha^s + z \in \mathbb{Z}[\zeta]_{(\lambda+n)s}$. Therefore (5.3) is satisfied for $z_n = z_{n-1}\alpha^s + z$.

We set $n = \overline{\lambda} - \lambda$. Then we get that for each $x \in \mathbb{Z}[\zeta]$ there exists a $z_x \in \mathbb{Z}[\zeta]_{\overline{\lambda}s}$ such that

$$x\alpha^{\overline{\lambda}s} + z_x \in S'.$$

Let $y \in \mathbb{Z}[\zeta]$ and let q and r be such that

$$y = q\alpha^{\overline{\lambda}s} + r \quad \text{with} \quad N(r) < N\left(\alpha^{\overline{\lambda}s}\right).$$

Therefore $q\alpha^{\overline{\lambda}s} + z_q \in S'$ and we have that

$$\left|y - \left(q\alpha^{\overline{\lambda}s} + z_q\right)\right| = |r - z_q| < 4\,|\alpha|^{\overline{\lambda}s+1}.$$

Thus for each $y \in \mathbb{Z}[\zeta]$ we have that $B\left(y, 4\,|\alpha|^{\overline{\lambda}s+1}\right) \cap S' \neq \emptyset$ and therefore S' is syndetic. $\square$

Chapitre 3

A central limit theorem for integer partitions

This chapter is joint work with Stephan Wagner and appeared in *Monatshefte für Mathematik*, **161** (2010) 85 – 114.

1. Introduction

For a nondecreasing sequence $\Lambda = (\Lambda_1, \Lambda_2, \ldots)$ of positive integers with $\Lambda_k \to \infty$, a *restricted Λ-partition* of n is a subsequence of Λ that sums to n, *i.e.*,

$$\sum_{j=1}^{s} \Lambda_{i_j} = n$$

with $i_1 < i_2 < \ldots < i_s$. On the other hand, if repetitions are allowed (i.e. $i_1 \leq i_2 \leq \ldots \leq i_s$), one speaks of *unrestricted Λ-partitions*. The number of restricted/unrestricted Λ-partitions is denoted by $q_\Lambda(n)$ and $p_\Lambda(n)$, respectively. There is a wealth of literature on the enumeration of restricted or unrestricted Λ-partitions, see for instance [**16**] and the references therein. Ingham [**112**] provides a Tauberian theorem that results in an asymptotic formula for the number of Λ-partitions under certain technical conditions. A different approach was used by Meinardus [**161**], who applied methods from complex analysis and was able to remove a monotonicity condition necessary in Ingham's approach. Essentially, if the Dirichlet generating function of Λ has only a simple pole at $\alpha > 0$ and can be analytically continued into a half-plane $\mathrm{Re} s \geq -\alpha_0$ with $\alpha_0 > 0$, and some other (fairly mild) conditions are satisfied, one obtains an asymptotic formula of the type

$$p_\Lambda(n) \sim A \cdot n^\kappa \exp\left(B \cdot n^{\alpha/(\alpha+1)}\right)$$

for unrestricted Λ-partitions. Under more general conditions, Roth and Szekeres [**201**] were able to prove slightly weaker theorems for both restricted and unrestricted partitions. The following holds for restricted partitions, the theorem for unrestricted partitions is similar :

THEOREM 3.1 ([**201**, Theorem 2]). *Assume that the following conditions hold :*

— $\alpha^{-1} = \lim\limits_{k\to\infty} \dfrac{\log \Lambda_k}{\log k}$ *exists,*

— $J_k = \inf\left\{(\log k)^{-1} \sum\limits_{\nu=1}^{k} \|\Lambda_\nu \beta\|^2\right\} \to \infty$ *as* $k \to \infty$, *where the infimum is taken over all* β *with* $\frac{1}{2}\Lambda_k^{-1} < \beta \leq \frac{1}{2}$, *and* $\|x\|$ *denotes the distance of* x *from the nearest integer.*

Then we have an asymptotic formula for $q_\Lambda(n)$, namely

$$q_\Lambda(n) = (2\pi A)^{-1/2} \exp\left(\sum_{k=1}^{\infty}\left(\frac{\eta\Lambda_k}{e^{\eta\Lambda_k}+1} + \log(1+e^{-\eta\Lambda_k})\right)\right) \cdot \left(1 + O(n^{-1/(\alpha+1)+\delta})\right),$$

where $\eta = \eta(n)$ is determined from

$$n = \sum_{k=1}^{\infty} \frac{\Lambda_k}{e^{\eta\Lambda_k}+1}$$

and

$$A = A(n) = \sum_{k=1}^{\infty} \frac{\Lambda_k^2 e^{\eta\Lambda_k}}{(e^{\eta\Lambda_k}+1)^2}.$$

The *length* (number of summands) of a partition is one of the most natural parameters to study. For unrestricted partitions, it was shown by Erdős and Lehner [**75**] that the number of summands asymptotically follows an extreme-value distribution in the special case $\Lambda = \mathfrak{N}$, a result that was further extended in many directions, see for instance [**78**]. Similar results are known for the distribution of distinct elements in unrestricted partitions, see [**95**, **221**]. On the other hand, the limit distribution is Gaussian if restricted partitions are considered, as was shown by Hwang [**109**], who extended a previous result of Erdős and Lehner. His conditions are essentially taken from the aforementioned paper of Meinardus [**161**]. Specifically, Hwang's central limit theorem reads as follows :

THEOREM 3.2 ([**109**, Theorem 1]). *Suppose that the sequence Λ satisfies the following conditions :*

- *(M1) The Dirichlet series $D(s) = \sum_{k\geq 1} \Lambda_k^{-s}$ converges in the half-plane $\mathrm{Re}\, s > \alpha > 0$, and can be analytically continued into the half-plane $\mathrm{Re}\, s \geq -\alpha_0$ for some $\alpha_0 > 0$. In $\mathrm{Re}\, s \geq -\alpha_0$, $D(s)$ is analytic except for a simple pole at $s = \alpha$ with residue A.*
- *(M2) There exists an absolute constant c_1 such that $D(s) \ll |t|^{c_1}$ uniformly for $\mathrm{Re}\, s \geq -\alpha_0$ as $|t| \to \infty$.*
- *(M3) Define $g(\tau) = \sum_{k\geq 1} e^{-\Lambda_k \tau}$, where $\tau = r + iy$ with $r > 0$ and $-\pi \leq y \leq \pi$. There exists a positive constant c_2 such that $g(r) - \mathrm{Re}\, g(\tau) \geq c_2(\log(1/r))^{2+4/\alpha^2}$ uniformly for $\pi/2 \leq |y| \leq \pi$ as $r \to 0^+$.*

Let ϖ_n be the random variable counting the number of summands in a random restricted Λ-partition of n. Set $\kappa = A\Gamma(\alpha)(1-2^{-\alpha})\zeta(\alpha+1)$, where Γ and ζ are the Gamma function and the Riemann zeta function, respectively. Furthermore, set

$$\mu_n = (\kappa\alpha)^{1/(\alpha+1)} \frac{(1-2^{1-\alpha})\zeta(\alpha)}{\alpha(1-2^{-\alpha})\zeta(\alpha+1)} n^{\alpha/(\alpha+1)},$$

$$\sigma_n^2 = (\kappa\alpha)^{1/(\alpha+1)} \left(\frac{(1-2^{2-\alpha})\zeta(\alpha-1)}{\alpha(1-2^{-\alpha})\zeta(\alpha+1)} - \frac{(1-2^{1-\alpha})^2\zeta(\alpha)^2}{(\alpha+1)(1-2^{-\alpha})^2\zeta(\alpha+1)^2}\right) n^{\alpha/(\alpha+1)}.$$

Then ϖ_n is asymptotically normally distributed with mean $\mathbb{E}(\varpi_n) \sim \mu_n$ and variance $\mathbb{V}(\varpi_n) \sim \sigma_n^2$:

$$\mathbb{P}\left(\frac{\varpi_n - \mu_n}{\sigma_n} < x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} \mathrm{d}t + o(1),$$

uniformly for all x as $n \to \infty$.

It is generally required in all the mentioned results that Meinardus' condition on the Dirichlet generating function (analyticity on a half-plane except for a simple pole at α) is satisfied. The examples given by Hwang include, for instance, powers ($\Lambda_j = j^\ell$) or arithmetic progressions ($\Lambda_j = a+bj$, where a and b are coprime). However, there are fairly natural integer sequences which do not satisfy the conditions (M1)-(M3) of Theorem 3.2. Specifically, we consider integers satisfying certain conditions on their digits in base b; various authors, most notably Gel'fond [**91**], studied arithmetical properties of integers defined by such conditions. A typical example is the set of numbers with *missing digits*, which were treated, among others, by Erdős, Mauduit and Sárközy [**76**, **77**] : if, for instance, one considers only those integers which do not contain the digit 2 in their base-3 expansion, one obtains the sequence

$$\Lambda = (1, 3, 4, 9, 10, 12, 13, 27, \ldots),$$

which is Sloane's A005836 [**232**]. It is not difficult to see (and will be shown later) that the corresponding Dirichlet generating function does not only have a pole at $\alpha = \frac{\log 2}{\log 3}$, but also further poles along the line $\mathrm{Re} s = \alpha$. This and similar examples form our motivation for replacing Hwang's conditions (M1) and (M2) by slightly weaker assumptions that allow for further poles with nonnegative real part. It will turn out that the limit distribution is still Gaussian under these assumptions and thus specifically for numbers with missing digits.

In order to get a flavor of the new phenomena that occur, let us apply the result of Roth and Szekeres (Theorem 3.1) to the sequence of integers with missing digits. Let $b > 2$ be an integer, and let $D \subseteq \{0, 1, \ldots, b-1\}$ be a set of digits ($1 < |D| < b$). Without loss of generality, we will always assume that the digits in D do not have a common divisor (otherwise, simply divide everything by the greatest common divisor). Let $\mathcal{MD}(b, D)$ be the set of positive integers with the property that all digits in the b-ary representation come from the set D, *i.e.*,

$$\mathcal{MD}(b, D) := \left\{ \sum_{i=0}^{k-1} a_i b^i \mid k \in \mathfrak{N},\, a_i \in D \right\} \setminus \{0\}.$$

Now we would like to determine (asymptotically) the number of restricted $\mathcal{MD}(b, D)$-partitions. To this end, we need some information on the Dirichlet generating function of such a set, which is provided by the following lemma.

LEMMA 3.3. *Let $D(s)$ be defined by*

$$D(s) = \sum_{m \in \mathcal{MD}(b,D)} m^{-s}.$$

Then we have

$$D(s) = \left(1 - |D| b^{-s}\right)^{-1} R(s),$$

where R is analytic within the right half-plane

$$\left\{ s \in \mathbb{C} \,\middle|\, \mathrm{Re}\, s > \frac{\log |D|}{\log b} - 1 \right\}$$

and satisfies $R(s) \ll |s|$ uniformly on

$$\left\{ s \in \mathbb{C} \,\middle|\, \mathrm{Re}\, s \geq \frac{\log |D|}{\log b} - 1 + \varepsilon \right\}.$$

REMARK 5. It should be mentioned that $D(s)$ is an instance of what is called an *automatic Dirichlet series* in [**12**].

DÉMONSTRATION. Note that

$$\mathcal{MD}(b,D) = \{bn_0 + a_0 \,|\, n_0 \in \mathcal{MD}(b,D) \cup \{0\},\, a_0 \in D\} \setminus \{0\}.$$

Therefore, we have

$$\begin{aligned} R(s) &= \left(1 - |D|b^{-s}\right) D(s) \\ &= \sum_{n \in \mathcal{MD}(b,D)} \sum_{a \in D} \left(\frac{1}{(bn+a)^s} - \frac{1}{(bn)^s} \right) + \sum_{a \in D \setminus \{0\}} \frac{1}{a^s}. \end{aligned}$$

So $R(s)$ is a Dirichlet series again, which means that it is analytic within a right half-plane. Hence, in order to prove the theorem, it is sufficient to show that the estimate $R(s) \ll |s|$ holds (uniformly) for $\sigma = \mathrm{Re} s \geq \frac{\log |D|}{\log b} - 1 + \varepsilon$. To this end, note that

$$\sum_{a \in D} \left(\frac{1}{(bn)^s} - \frac{1}{(bn+a)^s} \right) \ll |s| n^{-\sigma-1},$$

and this estimate holds uniformly as σ is restricted to a compact set. Furthermore, $\left|\mathcal{MD}(b,D) \cap [b^{l-1}, b^l)\right| \leq |D|^l$, so that we obtain

$$\begin{aligned} \sum_{n \in \mathcal{MD}(b,D)} \sum_{a \in D} \left(\frac{1}{(bn)^s} - \frac{1}{(bn+a)^s} \right) &\ll |s| \sum_{n \in \mathcal{MD}(b,D)} n^{-\sigma-1} \\ &\leq |s| \sum_{l=0}^{\infty} \frac{|D|^l}{b^{(l-1)(\sigma+1)}} \\ &= |s| b^{\sigma+1} \frac{1}{1 - |D| b^{-(\sigma+1)}} < \infty \end{aligned}$$

as long as $|D|b^{-(\sigma+1)} < 1$ or $\sigma > \frac{\log |D|}{\log b} - 1$, and the estimate is uniform for

$$\sigma = \mathrm{Re} s \geq \frac{\log |D|}{\log b} - 1 + \varepsilon.$$

□

Thus, in order to obtain asymptotic estimates for sums of the type

$$\sum_{m \in \mathcal{MD}(b,D)} f(m\eta),$$

as given in Theorem 3.1, we can use the Mellin inversion formula together with Lemma 3.3. For a survey on Mellin transforms we refer the reader to [**79**, **80**, **237**]. We denote the Mellin transform of a function $h(x)$ by

$$h^*(s) = \mathcal{M}\left[h(x); s\right] = \int_0^\infty h(x) x^{s-1} \mathrm{d}s.$$

First of all, we want to know the asymptotics of $\eta = \eta(n)$ in Theorem 3.1. Note that the Mellin transform of $f_1(x) = \frac{x}{e^x+1}$ is given by

$$f_1^*(s) = (1 - 2^{-s}) \Gamma(s+1) \zeta(s+1),$$

which is analytic for $\operatorname{Re} s > -1$. By the properties of the Mellin transform, we have

$$\mathcal{M}\left[\sum_{m\in\mathcal{MD}(b,D)} \frac{mx}{e^{mx}+1}; s\right] = (1-2^{-s})\Gamma(s+1)\zeta(s+1)D(s)$$

Thus, shifting the path of integration in the Mellin inversion formula yields

$$\sum_{m\in\mathcal{MD}(b,D)} \frac{m}{e^{\eta m}+1} = \eta^{-\alpha-1}U_1\left(\frac{\log\eta}{\log b}\right) + O(\eta^{-\alpha}),$$

where $\alpha = \frac{\log|D|}{\log b}$ and U_1 is a 1-periodic function given by its Fourier series (for details, see Lemma 3.6 in Section 2)

$$U_1(t) = \sum_{k\in\mathbb{Z}} \frac{1}{\log b}\left(1-2^{-\alpha+\frac{2k\pi i}{\log b}}\right)\Gamma\left(1+\alpha-\frac{2k\pi i}{\log b}\right)\zeta\left(1+\alpha-\frac{2k\pi i}{\log b}\right)R\left(\alpha-\frac{2k\pi i}{\log b}\right)\exp(2k\pi i t).$$

It follows that

$$\eta = n^{-1/(\alpha+1)}V\left(\frac{\log n}{\log|D|+\log b}\right) + O(n^{-2/(\alpha+1)}),$$

where V is also 1-periodic. Similarly, the Mellin transform of $f_2(x) = \log(1+e^{-x})$ is given by

$$f_2^*(s) = (1-2^{-s})\Gamma(s)\zeta(s+1),$$

and we obtain

$$\sum_{m\in\mathcal{MD}(b,D)} \left(\frac{m\eta}{e^{m\eta}+1} + \log(1+e^{-\eta m})\right) = \eta^{-\alpha}U_2\left(\frac{\log\eta}{\log b}\right)C + O(\eta^{1-\alpha}),$$

where U_2 is a 1-periodic function and C a constant. Finally, the Mellin transform of $f_3(x) = \frac{x^2e^x}{(e^x+1)^2}$ is given by

$$f_3^*(s) = (1-2^{-s})\Gamma(s+2)\zeta(s+1).$$

Summing up, we obtain the following asymptotic formula :

THEOREM 3.4. *The number $q_{\mathcal{MD}}(n)$ of partitions into distinct elements of $\mathcal{MD}(b,D)$ is asymptotically*

$$\begin{aligned} q_{\mathcal{MD}}(n) = {} & n^{-(\alpha+2)/(2\alpha+2)}\exp\left(n^{\alpha/(\alpha+1)}W_1\left(\frac{\log\eta}{\log|D|+\log b}\right)\right)W_2\left(\frac{\log\eta}{\log|D|+\log b}\right) \\ & \cdot\left(1+O(n^{-\min(\alpha,1-\alpha)/(\alpha+1)+\delta})\right) \end{aligned}$$

for some 1-periodic functions W_1, W_2.

Given this asymptotic result for the number of partitions, it is natural to consider distributions as well. However, Hwang's Theorem (Theorem 3.2) is not directly applicable since there is not only a single pole on the abscissa of convergence of the relevant Dirichlet series, but rather a countable set of poles, as can be seen from Lemma 3.3. Therefore, we aim to extend Hwang's result in order to make it applicable to partitions into integers with missing digits and similar sequences of integers (see the examples in Section 5).

2. Preliminaries and statement of the main result

In the following, we typically consider the case that the sequence Λ is strictly increasing, i.e.

$$\Lambda_1 < \Lambda_2 < \Lambda_3 < \dots,$$

and so it will be convenient to write $\mathcal{S}$ for the set $\{\Lambda_1, \Lambda_2, \Lambda_3, \dots\}$ (e.g. the set of integers with certain missing digits). However, all our theorems and proofs also apply in the case that $\mathcal{S}$ is a *multiset*, i.e. elements are allowed to occur with a certain (finite) multiplicity. In order to study the number of summands in partitions into distinct elements of $\mathcal{S}$, we define the bivariate generating function

$$Q(u,z) = \prod_{m\in\mathcal{S}} (1+uz^m)\,. \tag{2.1}$$

It is clear that the power of u indicates the number of summands. For convenience, we mostly work with the logarithm of Q and thus define the function

$$f(u,\tau) = \log Q(u,e^{-\tau}) = \sum_{m\in\mathcal{S}} \log\left(1+ue^{-m\tau}\right). \tag{2.2}$$

We write $D(s)$ for the Dirichlet generating function of $\mathcal{S}$, *i.e.*,

$$D(s) = \sum_{m\in\mathcal{S}} m^{-s},$$

and we will use the notation

$$\langle \alpha,\beta\rangle := \{z\in\mathbb{C} : \alpha < \mathrm{Re} z < \beta\}$$

for strips in the complex plane. Throughout this paper, we will assume that $\mathcal{S}$ satisfies the following conditions, which are slight modifications of Hwang's conditions (M1)-(M3) :

(M1') $D(s)$ converges in the half-plane $\mathrm{Re}s > \alpha > 0$ and can be analytically continued to $\mathrm{Re}s \geq \alpha-\varepsilon$ with $\varepsilon > 0$. On the line $\mathrm{Re}s = \alpha$, $D(s)$ has equidistant (the distance is denoted by ω) simple poles at $s = \alpha + 2\pi i k\omega$ with $k\in\mathbb{Z}$; A_k is the residue of $D(s)$ at $s = \alpha+2\pi ik\omega$. Furthermore we assume that there are no further poles with $\mathrm{Re}s \geq \alpha-\varepsilon$.

(M2') There exists a sequence $T_j\to\infty$ and a positive constant c_1 such that

$$D(s) \ll |T_j|^{c_1}$$

uniformly for all $s\in\langle\alpha-\varepsilon,\alpha\rangle$ with $|\mathrm{Im}s| = T_j$. Furthermore we assume that D satisfies

$$D(\alpha-\varepsilon+it) \ll |t|^{c_1}\,.$$

(M3') Let $g(\tau) = \sum_{m\in\mathcal{S}} e^{-m\tau}$, where $\tau = r+iy$ with $r>0$ and $-\pi\leq y\leq\pi$. There exists a positive constant c_2 such that

$$g(r) - \mathrm{Re}g(\tau) \geq c_2\left(\log\frac{1}{r}\right)^{2+4/\alpha}$$

uniformly for $\pi/2\leq|y|\leq\pi$ as $r\to 0^+$.

We will show that these conditions are sufficient for a central limit theorem as follows :

THEOREM 3.5 (Main Theorem). *Suppose that (M1')–(M3') hold. As in Theorem 3.2, let ϖ_n be the number of summands of a random partition. Then ϖ_n is asymptotically normally distributed with mean $\mathbb{E}(\varpi_n) \sim \mu_n$ and variance $\mathbb{V}(\varpi_n) \sim \sigma_n^2$:*

$$\mathbb{P}\left(\frac{\varpi_n - \mu_n}{\sigma_n} < x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} \mathrm{d}t + o(1),$$

uniformly for all x as $n \to \infty$. μ_n and σ_n are defined as follows :

$$\mu_n = \sum_{m \in \mathcal{S}} \frac{1}{e^{\eta m} + 1}, \tag{2.3}$$

$$\sigma_n^2 = \sum_{m \in \mathcal{S}} \frac{e^{\eta m}}{(e^{\eta m}+1)^2} - \frac{\left(\sum_{m \in \mathcal{S}} \frac{m e^{\eta m}}{(e^{\eta m}+1)^2}\right)^2}{\sum_{m \in \mathcal{S}} \frac{m^2 e^{\eta m}}{(e^{\eta m}+1)^2}}, \tag{2.4}$$

and η is implicitly given by

$$n = \sum_{m \in \mathcal{S}} \frac{m}{e^{\eta m} + 1}.$$

μ_n and σ_n satisfy the following asymptotic formulas (recall that ω is the distance between two poles as defined in (M1')) :

$$\mu_n \sim n^{\alpha/(1+\alpha)} \Psi_\mu \left(\frac{\omega \log n}{\alpha + 1}\right),$$

$$\sigma_n^2 \sim n^{\alpha/(1+\alpha)} \Psi_\sigma \left(\frac{\omega \log n}{\alpha + 1}\right),$$

for certain 1-periodic functions Ψ_μ and Ψ_σ. Finally, we have the following exponential bounds for the tails :

$$\mathbb{P}\left(\frac{\varpi_n - \mu_n}{\sigma_n} \geq x\right) \leq \begin{cases} e^{-x^2/2}\left(1 + O((\log n)^{-3})\right) & \text{if } 0 \leq x \leq n^{\alpha/(6\alpha+6)}/\log n, \\ e^{-n^{\alpha/(6\alpha+6)} x/(2\log n)}\left(1 + O((\log n)^{-3})\right) & \text{if } x \geq n^{\alpha/(6\alpha+6)}/\log n, \end{cases}$$

and the same inequalities for $\mathbb{P}\left(\frac{\varpi_n - \mu_n}{\sigma_n} \leq -x\right)$.

The proof makes use of the saddle point method that is applied to the generating function $Q(u, z)$ (*cf.* [**109**, **201**]). Note that the definition of η is analogous to that in Theorem 3.1—as we will see from the proof, this is precisely the choice for the saddle point. Harmonic sums over all elements of $\mathcal{S}$ (as in the definitions of η, μ_n, σ_n^2) will occur repeatedly, and so we will make frequent use of the following important lemma :

LEMMA 3.6 ([**79**, Theorem 4]). *Let $f(x)$ be continuous in $(0, \infty)$ with Mellin transform $f^*(s)$ having a nonempty fundamental strip $\langle \alpha, \beta \rangle$.*

Assume that $f^*(s)$ admits a meromorphic continuation to the strip $\langle \gamma, \beta \rangle$ for some $\gamma < \alpha$ with at most a countable set of poles P there, and is analytic on $\mathrm{Re}s = \gamma$. Assume also that there exists a real number $\eta \in (\alpha, \beta)$ and a sequence of horizontal segments $|\mathrm{Im}s| = T_j$ with $T_j \to +\infty$ such that

$$f^*(s) = \mathcal{O}\left(|s|^{-r}\right) \quad \text{with } r > 1$$

holds on these segments uniformly for $\gamma \leq \mathrm{Re}s \leq \eta$. If $f^*(s)$ admits the singular expansion

$$f^*(s) \asymp \sum_{\xi \in P} \frac{A_\xi}{s - \xi}$$

for $s \in \langle \gamma, \alpha \rangle$, then an asymptotic expansion of $f(x)$ at 0 is

$$f(x) = \sum_{\xi \in P} A_\xi x^{-\xi} + \mathcal{O}\left(x^{-\gamma}\right).$$

In order to apply the Mellin calculus to the function f defined in (2.2), we need the Mellin transform of $\log(1+ue^{-x})$, which we denote by $Y(u,s)$. The following lemma collects some of its important properties.

LEMMA 3.7 ([**109**, Lemma 1]). *For each fixed u lying in the cut-plane $\mathbb{C} \setminus (-\infty, -1]$, the function $Y(u,s)$ can be meromorphically continued into the whole s-plane with simple poles at $s = 0, -1, -2, \ldots$. Moreover, $Y(u,s)$ satisfies the estimate*

$$|Y(u, \sigma + it)| \ll e^{-(\pi/2-\varepsilon)|t|} \text{ for any } \varepsilon > 0 \text{ as } |t| \to +\infty,$$

uniformly as σ and u are restricted to compact sets.

By partial integration of $Y(u,s)$ we get

$$Y(u,s) = \int_0^\infty \log(1+ue^{-x}) x^{s-1} \mathrm{d}x = \frac{1}{s} \int_0^\infty \frac{1}{u^{-1}e^x + 1} x^s \mathrm{d}x$$

and analogously

$$Y(u,s) = \frac{1}{s(s+1)} \int_0^\infty \frac{u^{-1}e^x}{(u^{-1}e^x+1)^2} x^{s+1} \mathrm{d}x.$$

Note that the integrals can also be interpreted as Mellin transforms (of $\frac{x}{u^{-1}e^x+1}$ and $\frac{x^2 u^{-1} e^x}{(u^{-1}e^x+1)^2}$, respectively), which will be needed later.

3. Proof of the main theorem

In order to prove Theorem 7.1, we need an asymptotic formula for

$$Q_n(u) = [z^n] Q(u,z).$$

Using Cauchy's residue theorem and the substitution $z = e^{-(r+it)}$, this can be written as

$$Q_n(u) = \frac{1}{2\pi i} \oint_{|z|=e^{-r}} z^{-n-1} Q(u,z) \mathrm{d}z = \frac{e^{nr}}{2\pi} \int_{-\pi}^{\pi} \exp(int + f(u, r+it)) \mathrm{d}t \tag{3.1}$$

for any $r > 0$. Let $\delta > 0$ be any fixed number in the unit interval. Throughout the proof, we assume that $\delta \le u \le \delta^{-1}$. Thus, "uniformly in u" means "uniformly as $\delta \le u \le \delta^{-1}$". Now we apply the saddle-point method : in the following, $r = r(n,u)$ is chosen in such a way that

$$\left.\frac{\partial(int + f(u, r+it))}{\partial t}\right|_{t=0} = in - i \sum_{m \in \mathcal{S}} \frac{m}{u^{-1}e^{rm}+1} = 0.$$

or equivalently

$$n = \sum_{m \in \mathcal{S}} \frac{m}{u^{-1}e^{rm}+1}.$$

Note that the right hand side is strictly decreasing and thus bijective as a function of r. Therefore there is a unique r that satisfies this equation. Furthermore, r is strictly decreasing as a function of n (and tends to 0 as $n \to \infty$) and strictly increasing as a function of u.

We can make use of Lemma 3.6 to find an asymptotic formula for the sum in the definition of r. Recall that the Mellin transform of $\frac{x}{u^{-1}e^x+1}$ is given by

$$\mathcal{M}\left[\frac{x}{u^{-1}e^x+1};\, s\right] = sY(u,s),$$

and thus

$$\mathcal{M}\left[\sum_{m\in\mathcal{S}} \frac{mx}{u^{-1}e^{xm}+1};\, s\right] = sY(u,s)D(s).$$

So Lemma 3.6 yields

$$\begin{aligned} n &= \sum_{m\in\mathcal{S}} \frac{m}{u^{-1}e^{rm}+1} = \frac{1}{r}\sum_{m\in\mathcal{S}} \frac{rm}{u^{-1}e^{rm}+1} \\ &= \frac{1}{r}\sum_{j\in\mathbb{Z}} A_j(\alpha+2\pi ij\omega)Y(u,\alpha+2\pi ij\omega)r^{-(\alpha+2\pi ij\omega)} + \mathcal{O}(r^{-(\alpha+1)+\varepsilon}) \\ &= r^{-(\alpha+1)}\sum_{j\in\mathbb{Z}} A_j(\alpha+2\pi ij\omega)Y(u,\alpha+2\pi ij\omega)\exp(-2\pi ij\omega\log r) + \mathcal{O}(r^{-(\alpha+1)+\varepsilon}) \\ &= r^{-(\alpha+1)}\Phi_1\left(u,\omega\log r\right) + \mathcal{O}(r^{-(\alpha+1)+\varepsilon}) \end{aligned}$$

for a 1-periodic function Φ_1 that is given by the Fourier series

$$\Phi_1(u,v) = \sum_{j\in\mathbb{Z}} A_j(\alpha+2\pi ij\omega)Y(u,\alpha+2\pi ij\omega)\exp(-2\pi ijv).$$

The properties of Y summarized in Lemma 3.7 guarantee that this series is absolutely and uniformly convergent and infinitely differentiable. Also note that

$$H(u,r) = \sum_{m\in\mathcal{S}} \frac{m}{u^{-1}e^{rm}+1}$$

is a positive and monotonic function of r. Therefore, $\Phi_1(u,v)$ can never be 0 : otherwise, there are sequences $r_{1,k}$ and $r_{2,k}$ both tending to 0 such that

$$H(u,r_{1,k}) \ll r_{1,k}^{-(\alpha+1)+\varepsilon} \quad \text{and} \quad H(u,r_{2,k}) \gg r_{2,k}^{-(\alpha+1)}$$

(the latter simply follows from the fact that Φ_1 is not identically 0), contradicting the monotonicity. Thus, $\Phi_1(u,v)$ must be bounded above and below by strictly positive constants (uniformly in u), which means that $r = \Theta(n^{-1/(\alpha+1)})$. More precisely, one has

$$r \sim n^{-\frac{1}{\alpha+1}}\Psi_1\left(u, \frac{\omega\log n}{\alpha+1}\right)$$

for a 1-periodic function Ψ_1, which will be used later.

For our application of the saddle point method, we need a uniform estimate as t in the integral representation (3.1) is away from 0. This is the main objective of the following two lemmas :

Lemma 3.8. *For every integer $\ell \geq 0$ we have*

$$h(X) = \sum_{\substack{m\leq X\\ m\in\mathcal{S}}} m^\ell \gg X^{\alpha+\ell}.$$

DÉMONSTRATION. For a nonnegative integer k we set

$$G_k(X) := \sum_{\substack{m \le X \\ m \in \mathcal{S}}} m^\ell \left(1 - \frac{m}{X}\right)^k.$$

Sums of this type can be written as integrals by means of the Mellin transform (see [**80**, Theorem 2.1]) :

$$G_k(X) = \frac{k!}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{D(s-\ell)X^s}{s(s+1)\dots(s+k)}\,\mathrm{d}s$$

for any $c > \alpha + \ell$. We choose k large enough ($k > c_1$, where the constant c_1 is taken as in (M2')) so as to make the resulting integral converge and shift the line of integration to the left (collecting residues at $s = \alpha + \ell + 2\pi i j\omega$ for every $j \in \mathbb{Z}$) to obtain

$$\begin{aligned} G_k(X) &= k! \sum_{j\in\mathbb{Z}} \frac{A_j X^{\alpha+\ell+2\pi i j\omega}}{(\alpha+\ell+2\pi i j\omega)(\alpha+\ell+1+2\pi i j\omega)\dots(\alpha+\ell+k+2\pi i j\omega)} \\ &\quad + \frac{k!}{2\pi i} \int_{\alpha+\ell-\varepsilon-i\infty}^{\alpha+\ell-\varepsilon+i\infty} \frac{D(s-\ell)X^s}{s(s+1)\dots(s+k)}\,\mathrm{d}s \\ &= X^{\alpha+\ell}\Xi_k\left(\omega \log X\right) + \mathcal{O}\left(X^{\alpha+\ell-\varepsilon}\right), \end{aligned}$$

where Ξ_k is a function of period 1, given by its Fourier series. Here, we made use of the fact that the integral can be estimated as follows :

$$\begin{aligned} \left|\int_{\alpha+\ell-\varepsilon-i\infty}^{\alpha+\ell-\varepsilon+i\infty} \frac{D(s-\ell)X^s}{s(s+1)\dots(s+k)}\,\mathrm{d}s\right| &\le X^{\alpha+\ell-\varepsilon} \int_{\alpha+\ell-\varepsilon-i\infty}^{\alpha+\ell-\varepsilon+i\infty} \left|\frac{D(s-\ell)}{s(s+1)\dots(s+k)}\right| \mathrm{d}s \\ &\ll X^{\alpha+\ell-\varepsilon} \int_{\alpha+\ell-\varepsilon-i\infty}^{\alpha+\ell-\varepsilon+i\infty} |s|^{c_1-k-1}\,\mathrm{d}s \\ &\ll X^{\alpha+\ell-\varepsilon}. \end{aligned}$$

Since $G_k(X)$ is nonnegative for all X, $\Xi_k(\omega \log X)$ must be nonnegative for sufficiently large X (and thus for all X, since it is periodic). Now assume that $\Xi_k(\omega \log X)$ is 0 for some X. Equivalently,

$$G_k(X_n) X_n^{-(\alpha+\ell)} \to 0$$

for $X_n = Xe^{n/\omega}$. Note that

$$G_k(X) \le h(X) \le G_k(\beta X)\left(1 - \frac{1}{\beta}\right)^{-k}. \tag{3.2}$$

for every k and $\beta > 1$. Therefore, we also have

$$G_k(X_n/\beta) X_n^{-(\alpha+\ell)} \to 0$$

for any $\beta > 1$, implying $\Xi_k(\omega(\log X - \log\beta)) = 0$. But then, Ξ_k is identically 0, an obvious contradiction. Hence, Ξ_k is bounded above and below by strictly positive constants. Now, the left hand side inequality in (3.2) shows that $h(X) \gg X^{\alpha+\ell}$, as claimed. □

The following simple corollary will be needed later :

COROLLARY 3.9. *There is a constant $C > 1$ such that the cardinality of $\mathcal{S} \cap (X, CX]$ satisfies*

$$|\mathcal{S} \cap (X, CX]| \gg X^\alpha.$$

Démonstration. Set $\ell = 0$ in the lemma; the proof shows that $h(X) \ll X^\alpha$ holds as well as $h(X) \gg X^\alpha$. Hence,

$$|\mathcal{S} \cap (X, CX]| = h(CX) - h(X) \gg X^\alpha$$

for sufficiently large C. □

Lemma 3.10. *For any constant c_3 with $0 < c_3 < \alpha/2$, there is a constant c_4 such that*

$$\frac{\left|Q(u, e^{-(r+iy)})\right|}{Q(u, e^{-r})} \le \exp\left(-\frac{c_4 u}{(1+u)^2}\left(\log\frac{1}{r}\right)^2\right),$$

for $r^{1+c_3} \le |y| \le \pi$ as $r \to 0^+$.

Démonstration. We start by rewriting the quotient under consideration.

$$\left(\frac{\left|Q(u, e^{-(r+iy)})\right|}{Q(u, e^{-r})}\right)^2 = \prod_{m\in\mathcal{S}}\left(1 - \frac{2ue^{-mr}(1-\cos my)}{(1+ue^{-mr})^2}\right)$$
$$\le \exp\left(-\frac{2u}{(1+u)^2}\sum_{m\in\mathcal{S}} e^{-mr}(1-\cos my)\right).$$

Using the definition of g in (M3') we set

$$G(r,y) := g(r) - \mathrm{Re}g(r+iy) = \sum_{m\in\mathcal{S}}\left(e^{-mr} - \mathrm{Re}e^{-m(r+iy)}\right) = \sum_{m\in\mathcal{S}} e^{-mr}(1-\cos my).$$

Now (M3') yields

$$G(r,y) \ge c_2\left(\log\frac{1}{r}\right)^{2+4/\alpha} \text{ for } \frac{\pi}{2} \le |y| \le \pi,$$

and so it suffices to show

$$G(r,y) \ge c_5\left(\log\frac{1}{r}\right)^2$$

for some constant $c_5 > 0$ uniformly for $r^{1+c_3} \le |y| \le \frac{\pi}{2}$ as $r \to 0^+$.

We split this interval into three parts according to the size of $|y|$.

— $r \le |y| \le \left(\log\frac{1}{r}\right)^{-\frac{2}{\alpha}}$: Note that for $|t| \le \pi$

$$1 - \cos t \ge \frac{2}{\pi^2}t^2.$$

Thus we can apply Lemma 3.8 to find

$$G(r,y) \ge \sum_{\substack{1\le m\le |y|^{-1}\\ m\in\mathcal{S}}} e^{-mr}(1-\cos my) \ge \sum_{\substack{1\le m\le |y|^{-1}\\ m\in\mathcal{S}}} e^{-mr}\frac{2}{\pi^2}m^2y^2$$
$$\ge \frac{2}{\pi^2}y^2 \sum_{\substack{1\le m\le |y|^{-1}\\ m\in\mathcal{S}}} e^{-\frac{r}{|y|}}m^2 \ge \frac{2}{\pi^2}e^{-1}y^2\cdot c_6|y|^{-(\alpha+2)}$$
$$\ge c_5\,|y|^{-\alpha} \ge c_5\left(\log\frac{1}{r}\right)^2$$

if c_5 is chosen sufficiently small.

— $r^{1+c_3} \le |y| \le r$: In the same manner as before, we get

$$G(r,y) \ge \sum_{\substack{1\le m\le r^{-1}\\ m\in\mathcal{S}}} e^{-mr}(1-\cos my) \ge \sum_{\substack{1\le m\le r^{-1}\\ m\in\mathcal{S}}} \frac{2}{\pi^2}e^{-1}y^2m^2$$
$$\ge \frac{2}{\pi^2}e^{-1}y^2\cdot c_6 r^{-(\alpha+2)} \ge c_7 r^{2c_3-\alpha} \ge c_5\left(\log\frac{1}{r}\right)^2.$$

— $\left(\log\frac{1}{r}\right)^{-\frac{2}{\alpha}} \le |y| \le \frac{\pi}{2}$: It is clear that there exists an integer ℓ such that

$$\frac{\pi}{2} \le 2^\ell |y| \le \pi.$$

From the inequality $1-\cos\theta \ge \frac{1}{4}(1-\cos 2\theta)$ we get by iteration

$$1-\cos\theta \ge 4^{-\ell}(1-\cos 2^\ell\theta),$$

which allows us to apply (M3') again :

$$G(r,y) \ge 4^{-\ell}\sum_{m\in\mathcal{S}} e^{-mr}(1-\cos 2^\ell my) \ge 4^{-\ell}c_2\left(\log\frac{1}{r}\right)^{2+4/\alpha}$$
$$\ge \frac{c_2}{\pi^2}y^2\left(\log\frac{1}{r}\right)^{2+4/\alpha} \ge \frac{c_2}{\pi^2}\left(\log\frac{1}{r}\right)^2,$$

so that we obtain the desired estimate in this case as well.

□

Now we return to the integral representation (3.1) : choose c_3 such that $\frac{\alpha}{3} < c_3 < \frac{\alpha}{2}$, e.g. $c_3 = \frac{3\alpha}{7}$, and split the interval into the part $|t| \le r^{1+c_3}$ and the remaining two intervals. For the latter, Lemma 3.10 shows that

$$\int_{r^{1+c_3}}^{\pi} \exp(int + f(u,r+it))\mathrm{d}t \ll \exp\left(f(u,r) - c_8\left(\log\frac{1}{r}\right)^2\right) \ll \exp(f(u,r) - c_9\log^2 n)$$

for certain positive constants c_8 and c_9 (uniformly in u), and the same estimate holds for $-\pi \le t \le -r^{1+c_3}$. For the central integral, we have to expand $f(u,r+it)$ around $t=0$: by our choice of r, the first derivative with respect to t is $-in$, and the second derivative is given by

$$-B^2 = -B^2(u,r) = \left.\frac{\partial^2 f(u,r+it)}{\partial^2 t}\right|_{t=0} = -\sum_{m\in\mathcal{S}} \frac{m^2u^{-1}e^{rm}}{(u^{-1}e^{rm}+1)^2}.$$

Now we can apply the Mellin transform technique again : the transform of $\frac{x^2u^{-1}e^x}{(u^{-1}e^x+1)^2}$ is given by

$$\mathcal{M}\left[\frac{x^2u^{-1}e^x}{(u^{-1}e^x+1)^2};\, s\right] = s(s+1)Y(u,s),$$

and thus

$$\mathcal{M}\left[\sum_{m\in\mathcal{S}} \frac{x^2m^2u^{-1}e^{xm}}{(u^{-1}e^{xm}+1)^2};\, s\right] = s(s+1)Y(u,s)D(s).$$

Thus, applying Lemma 3.6 yields

$$B^2 = \sum_{m\in\mathcal{S}} \frac{m^2 u^{-1} e^{rm}}{(u^{-1}e^{rm}+1)^2} = r^{-(\alpha+2)}\Phi_2\left(u, \omega \log r\right) + \mathcal{O}(r^{-(\alpha+2)+\varepsilon}),$$

where Φ_2 is a periodic function ; again, a simple argument shows that Φ_2 is bounded below by a positive constant (uniformly for $\delta \leq u \leq \delta^{-1}$, as it is the case for all our estimates), implying that B^2 is of order $r^{-(\alpha+2)}$: just note that

$$\begin{aligned} B^2 = \sum_{m\in\mathcal{S}} \frac{m^2 u^{-1} e^{rm}}{(u^{-1}e^{rm}+1)^2} &\geq \sum_{\substack{m\leq r^{-1}\\ m\in\mathcal{S}}} \frac{m^2 u^{-1} e^{rm}}{(u^{-1}e^{rm}+1)^2} \\ &\geq \sum_{\substack{m\leq r^{-1}\\ m\in\mathcal{S}}} \frac{\delta m^2}{(\delta^{-1}e+1)^2} \gg \sum_{\substack{m\leq r^{-1}\\ m\in\mathcal{S}}} m^2 \gg r^{-(\alpha+2)} \end{aligned}$$

by Lemma 3.8. Finally, we estimate the third derivative as follows : it is given by

$$\frac{\partial^3 f(u, r+it)}{\partial^3 t} = -i \sum_{m\in\mathcal{S}} \frac{m^3 u^{-1} e^{mr(1+it)}\left(1 - u^{-1}e^{mr(1+it)}\right)}{(u^{-1}e^{mr(1+it)}+1)^3}.$$

Now let $m_0 = r^{-(1+c_{10})}$ for some constant $c_{10} > 0$, and write $v = u^{-1}$ for short. Then we split up the sum into two parts according to whether $m \leq m_0$ or not. For the latter we get

$$\begin{aligned} \left| \sum_{\substack{m>m_0\\ m\in\mathcal{S}}} \frac{m^3 v e^{mr(1+it)}\left(1 - v e^{mr(1+it)}\right)}{(v e^{mr(1+it)}+1)^3} \right| &\leq \sum_{\substack{m>m_0\\ m\in\mathcal{S}}} \frac{m^3 v e^{mr}\left(1+v e^{mr}\right)}{\left|v e^{mr(1+it)}+1\right|^3} \\ &\leq \sum_{\substack{m>m_0\\ m\in\mathcal{S}}} \frac{m^3 v e^{mr}\left(1+v e^{mr}\right)}{(v e^{mr}-1)^3} \ll \sum_{\substack{m>m_0\\ m\in\mathcal{S}}} \frac{m^3}{e^{mr}} \ll \frac{r^{-4-3c_{10}}}{e^{r^{-c_{10}}}}. \end{aligned}$$

For the remaining sum we note that

$$\left|1 + v e^{mr(1+it)}\right| \geq (1 + v e^{mr}) \cos\left(\frac{mrt}{2}\right).$$

Therefore we get

$$\begin{aligned} &\left| \sum_{\substack{m\leq m_0\\ m\in\mathcal{S}}} \frac{m^3 v e^{mr(1+it)}\left(1 - v e^{mr(1+it)}\right)}{(v e^{mr(1+it)}+1)^3} \right| \leq \sum_{\substack{m\leq m_0\\ m\in\mathcal{S}}} \frac{m^3 v e^{mr}\left(1+v e^{mr}\right)}{\left|v e^{mr(1+it)}+1\right|^3} \\ &\leq \sum_{\substack{m\leq m_0\\ m\in\mathcal{S}}} \frac{m^3 v e^{mr}\left(1+v e^{mr}\right)}{(v e^{mr}+1)^3}\left(1 + \mathcal{O}\left((rmt)^2\right)\right) \leq \sum_{\substack{m\leq m_0\\ m\in\mathcal{S}}} \frac{m^3}{v e^{mr}}\left(1 + \mathcal{O}\left((rmt)^2\right)\right) \\ &\leq \sum_{m\in\mathcal{S}} \frac{m^3}{v e^{mr}} + \mathcal{O}\left(\sum_{m\in\mathcal{S}} \frac{m^5 r^2 t^2}{e^{mr}}\right) \ll r^{-3-\alpha} + r^{-3-\alpha} t^2, \end{aligned}$$

where the last estimate is obtained by means of Lemma 3.6 again. So finally,

$$\frac{\partial^3 f(u, r+it)}{\partial^3 t} \ll r^{-3-\alpha}$$

for $|t| \leq r^{1+c_3}$, and so we have the expansion

$$f(u, r+it) = f(u,r) - int - \frac{B^2}{2}t^2 + \mathcal{O}\left(r^{-3-\alpha}t^3\right).$$

Hence, the corresponding integral can be estimated as follows :

$$\begin{aligned}
&\frac{e^{nr}}{2\pi}\int_{-r^{1+c_3}}^{r^{1+c_3}} \exp(int + f(u, r+it))\mathrm{d}t \\
&\qquad = \frac{e^{nr+f(u,r)}}{2\pi}\int_{-r^{1+c_3}}^{r^{1+c_3}} \exp\left(-\frac{B^2}{2}t^2 + \mathcal{O}\left(r^{3c_3-\alpha}\right)\right)\mathrm{d}t \\
&\qquad = \frac{e^{nr+f(u,r)}}{2\pi}\left(\int_{-r^{1+c_3}}^{r^{1+c_3}} \exp\left(-\frac{B^2}{2}t^2\right)\mathrm{d}t\right)\left(1 + \mathcal{O}\left(r^{2\alpha/7}\right)\right) \\
&\qquad = \frac{e^{nr+f(u,r)}}{2\pi}\left(\int_{-r^{1+c_3}}^{r^{1+c_3}} \exp\left(-\frac{B^2}{2}t^2\right)\mathrm{d}t\right)\left(1 + \mathcal{O}\left(n^{-2\alpha/(7\alpha+7)}\right)\right)
\end{aligned}$$

by our choice of c_3. Also note that

$$\begin{aligned}
\int_{-r^{1+c_3}}^{r^{1+c_3}} \exp\left(-\frac{B^2}{2}t^2\right)\mathrm{d}t &= \int_{-\infty}^{\infty} \exp\left(-\frac{B^2}{2}t^2\right)\mathrm{d}t - 2\int_{r^{1+c_3}}^{\infty} \exp\left(-\frac{B^2}{2}t^2\right)\mathrm{d}t \\
&= \frac{\sqrt{2\pi}}{B} + \mathcal{O}\left(\int_{r^{1+c_3}}^{\infty} \exp\left(-\frac{B^2 r^{1+c_3}}{2}t\right)\mathrm{d}t\right) \\
&= \frac{\sqrt{2\pi}}{B} + \mathcal{O}\left(r^{-1-c_3}B^{-2}\exp\left(-\frac{B^2 r^{2(1+c_3)}}{2}\right)\right).
\end{aligned}$$

We know that $B^2 \gg r^{-(\alpha+2)}$, which implies

$$B^2 r^{2(1+c_3)} \gg r^{2c_3-\alpha} = r^{-\alpha/7} \gg n^{\alpha/(7\alpha+7)}.$$

Hence, the error term tends to 0 faster than any power of n.

Putting everything together, we find that

$$Q_n(u) = \frac{1}{\sqrt{2\pi B^2}} e^{nr+f(u,r)}\left(1 + \mathcal{O}\left(n^{-2\alpha/(7\alpha+7)}\right)\right)$$

uniformly in u. Now we study the moment generating function of the random variable ϖ_n (the number of parts in a random partition), which can be expressed in terms of $Q_n(u)$: let $M_n(t) = \mathbb{E}(e^{(\varpi_n-\mu_n)t/\sigma_n})$, where t is real and μ_n and σ_n are chosen as in (2.3) and (2.4). Then we get

$$\begin{aligned}
M_n(t) &= \exp\left(-\frac{\mu_n t}{\sigma_n}\right)\frac{Q_n(e^{t/\sigma_n})}{Q_n(1)} \\
&= \sqrt{\frac{B^2(1, r(n,1))}{B^2(e^{t/\sigma_n}, r(n, e^{t/\sigma_n}))}}\exp\Big(-\frac{\mu_n t}{\sigma_n} + nr(n, e^{t/\sigma_n}) + f(e^{t/\sigma_n}, r(n, e^{t/\sigma_n})) \\
&\qquad - nr(n,1) - f(1, r(n,1)) + \mathcal{O}\left(n^{-2\alpha/(7\alpha+7)}\right)\Big).
\end{aligned} \tag{3.3}$$

Now, we want to determine the expansion around $t = 0$; for this purpose, we need the partial derivatives of $r(n, u)$ with respect to u : recall that $r = r(n, u)$ was defined by

$$n = -f_\tau(u, r) = \sum_{m\in\mathcal{S}} \frac{m}{u^{-1}e^{rm} + 1}.$$

Hence, the partial derivatives can be determined by means of implicit differentiation :

$$r_u = r_u(n, u) = -\frac{f_{u\tau}(u, r)}{f_{\tau\tau}(u, r)} = \frac{\sum_{m\in\mathcal{S}} \frac{me^{mr}}{(e^{mr}+u)^2}}{\sum_{m\in\mathcal{S}} \frac{um^2e^{mr}}{(e^{mr}+u)^2}}$$

and similarly

$$r_{uu} = r_{uu}(n, u) = \frac{-f_{\tau\tau\tau}(u, r)f_{u\tau}(u, r)^2 + 2f_{u\tau\tau}(u, r)f_{u\tau}(u, r)f_{\tau\tau}(u, r) - f_{uu\tau}(u, r)f_{\tau\tau}(u, r)^2}{f_{\tau\tau}(u, r)^3},$$

$$\begin{aligned} r_{uuu} = r_{uuu}(n, u) = f_{\tau\tau}(u, r)^{-5}\Big(& - f_{uuu\tau}(u, r)f_{\tau\tau}(u, r)^4 \\ & + \left(3f_{uu\tau\tau}(u, r)f_{u\tau}(u, r) + 3f_{uu\tau}(u, r)f_{u\tau\tau}(u, r)\right) f_{\tau\tau}(u, r)^3 \\ & + \big(-3f_{u\tau\tau\tau}(u, r)f_{u\tau}(u, r)^2 - 6f_{u\tau\tau}(u, r)^2 f_{u\tau}(u, r) \\ & \qquad -3f_{uu\tau}(u, r)f_{\tau\tau\tau}(u, r)f_{u\tau}(u, r)\big) f_{\tau\tau}(u, r)^2 \\ & + \left(f_{\tau\tau\tau\tau}(u, r)f_{u\tau}(u, r)^3 + 9f_{u\tau\tau}(u, r)f_{\tau\tau\tau}(u, r)f_{u\tau}(u, r)^2\right) f_{\tau\tau}(u, r) \\ & - 3f_{u\tau}(u, r)^3 f_{\tau\tau\tau}(u, r)^2\Big), \end{aligned}$$

where the derivatives of f are given as follows :

$$f_{u\tau}(u, r) = -\sum_{m\in\mathcal{S}} \frac{me^{mr}}{(e^{mr} + u)^2} \ll r^{-(1+\alpha)} \ll n,$$

$$f_{\tau\tau}(u, r) = \sum_{m\in\mathcal{S}} \frac{um^2e^{mr}}{(e^{mr} + u)^2} \ll r^{-(2+\alpha)} \ll n^{1+1/(1+\alpha)},$$

$$f_{uu\tau}(u, r) = \sum_{m\in\mathcal{S}} \frac{2me^{mr}}{(e^{mr} + u)^3} \ll r^{-(1+\alpha)} \ll n,$$

$$f_{u\tau\tau}(u, r) = \sum_{m\in\mathcal{S}} \frac{m^2e^{mr}(e^{mr} - u)}{(e^{mr} + u)^3} \ll r^{-(2+\alpha)} \ll n^{1+1/(1+\alpha)},$$

$$f_{\tau\tau\tau}(u, r) = -\sum_{m\in\mathcal{S}} \frac{um^3e^{mr}(e^{mr} - u)}{(e^{mr} + u)^3} \ll r^{-(3+\alpha)} \ll n^{1+2/(1+\alpha)},$$

$$f_{uuu\tau}(u, r) = -\sum_{m\in\mathcal{S}} \frac{6me^{mr}}{(e^{mr} + u)^4} \ll r^{-(1+\alpha)} \ll n,$$

$$f_{uu\tau\tau}(u, r) = -\sum_{m\in\mathcal{S}} \frac{2m^2e^{mr}(2e^{mr} - u)}{(e^{mr} + u)^4} \ll r^{-(2+\alpha)} \ll n^{1+1/(1+\alpha)},$$

$$f_{u\tau\tau\tau}(u, r) = -\sum_{m\in\mathcal{S}} \frac{m^3e^{mr}(e^{2mr} - 4ue^{mr} + u^2)}{(e^{mr} + u)^4} \ll r^{-(3+\alpha)} \ll n^{1+2/(1+\alpha)},$$

$$f_{\tau\tau\tau\tau}(u,r) = \sum_{m\in\mathcal{S}} \frac{um^4e^{mr}(e^{2mr}-4ue^{mr}+u^2)}{(e^{mr}+u)^4} \ll r^{-(4+\alpha)} \ll n^{1+3/(1+\alpha)}.$$

The asymptotic estimates are all obtained by means of the usual Mellin transform method. Furthermore, note that

$$f_{\tau\tau}(u,r) = B^2(u,r) \gg r^{-(2+\alpha)} \gg n^{1+1/(1+\alpha)},$$

from which it follows that $r_u, r_{uu}, r_{uuu} \ll n^{-1/(1+\alpha)}$, all uniformly in u. Thus, we have the following expansions :

$$r(n,e^{t/\sigma_n}) - r(n,1) = r_u(n,1)\cdot\frac{t}{\sigma_n} + \frac{r_u(n,1)+r_{uu}(n,1)}{2}\cdot\left(\frac{t}{\sigma_n}\right)^2 + \mathcal{O}\left(n^{-1/(1+\alpha)}\frac{t^3}{\sigma_n^3}\right)$$

and

$$\begin{aligned} f(e^{t/\sigma_n}, r(n,e^{t/\sigma_n})) - f(1,r(n,1)) &= \Big(f_\tau(1,r(n,1))r_u(n,1) + f_u(1,r(n,1))\Big)\cdot\frac{t}{\sigma_n} \\ &+ \frac{1}{2}\Big(f_\tau(1,r(n,1))(r_u(n,1)+r_{uu}(n,1)) + f_{\tau\tau}(1,r(n,1))r_u(n,1)^2 + 2f_{u\tau}(1,r(n,1))r_u(n,1) \\ &+ f_u(1,r(n,1)) + f_{uu}(1,r(n,1))\Big)\cdot\left(\frac{t}{\sigma_n}\right)^2 + \mathcal{O}\left(n^{\alpha/(1+\alpha)}\frac{t^3}{\sigma_n^3}\right). \end{aligned}$$

Altogether, this means that the exponent in (3.3) can be written as

$$\begin{aligned} &\Big(nr_u(n,1) + f_\tau(1,\eta)r_u(n,1) + f_u(1,\eta) - \mu_n\Big)\cdot\frac{t}{\sigma_n} \\ &+ \frac{1}{2}\Big(n(r_u(n,1)+r_{uu}(n,1)) + f_\tau(1,\eta)(r_u(n,1)+r_{uu}(n,1)) + f_{\tau\tau}(1,\eta)r_u(n,1)^2 \\ &\quad + 2f_{u\tau}(1,\eta)r_u(n,1) + f_u(1,\eta) + f_{uu}(1,\eta)\Big)\cdot\left(\frac{t}{\sigma_n}\right)^2 + \mathcal{O}\left(n^{\alpha/(1+\alpha)}\frac{t^3}{\sigma_n^3} + n^{-2\alpha/(7\alpha+7)}\right), \end{aligned}$$

where we use η as an abbreviation for $r(n,1)$. Now, we make use of the fact that $n = -f_\tau(1,\eta)$ and that $r_u(n,1) = -\frac{f_{u\tau}(1,\eta)}{f_{\tau\tau}(1,\eta)}$ to simplify this expression :

$$\begin{aligned} \Big(f_u(1,\eta) - \mu_n\Big)\cdot\frac{t}{\sigma_n} + \frac{1}{2}\left(f_u(1,\eta) + f_{uu}(1,\eta) - \frac{f_{u\tau}(1,\eta)^2}{f_{\tau\tau}(1,\eta)}\right)\cdot\left(\frac{t}{\sigma_n}\right)^2 \\ + \mathcal{O}\left(n^{\alpha/(1+\alpha)}\frac{t^3}{\sigma_n^3} + n^{-2\alpha/(7\alpha+7)}\right). \end{aligned}$$

It is not difficult to show in a similar way that

$$\frac{B^2(1,r(n,1))}{B^2(e^{t/\sigma_n}, r(n,e^{t/\sigma_n}))} = 1 + \mathcal{O}\left(\frac{t}{\sigma_n}\right),$$

and so we obtain the following asymptotic formula for the moment generating function from (3.3) :

$$\begin{aligned} M_n(t) = \exp\Bigg(\Big(f_u(1,\eta) - \mu_n\Big)\cdot\frac{t}{\sigma_n} + \frac{1}{2}\left(f_u(1,\eta) + f_{uu}(1,\eta) - \frac{f_{u\tau}(1,\eta)^2}{f_{\tau\tau}(1,\eta)}\right)\cdot\left(\frac{t}{\sigma_n}\right)^2 \\ + \mathcal{O}\left(\frac{t}{\sigma_n} + n^{\alpha/(1+\alpha)}\frac{t^3}{\sigma_n^3} + n^{-2\alpha/(7\alpha+7)}\right)\Bigg). \end{aligned}$$

Now, note that μ_n and σ_n were chosen in such a way that

$$\mu_n = f_u(1,\eta) = \sum_{m\in\mathcal{S}} \frac{1}{e^{\eta m}+1},$$

$$\sigma_n^2 = f_u(1,\eta) + f_{uu}(1,\eta) - \frac{f_{u\tau}(1,\eta)^2}{f_{\tau\tau}(1,\eta)} = \sum_{m\in\mathcal{S}} \frac{e^{\eta m}}{(e^{\eta m}+1)^2} - \frac{\left(\sum_{m\in\mathcal{S}} \frac{me^{\eta m}}{(e^{\eta m}+1)^2}\right)^2}{\sum_{m\in\mathcal{S}} \frac{m^2 e^{\eta m}}{(e^{\eta m}+1)^2}}.$$

We only have to prove that the error term is small, and so we need a lower estimate for σ_n : first of all, our usual Mellin transform technique shows that

$$\mu_n = \eta^{-\alpha}\Phi_\mu(\omega\log\eta) + \mathcal{O}\left(\eta^{-\alpha+\varepsilon}\right) \sim n^{\alpha/(1+\alpha)}\Psi_\mu\left(\frac{\omega\log n}{\alpha+1}\right)$$

and

$$\sigma_n^2 = \eta^{-\alpha}\Phi_\sigma(\omega\log\eta) + \mathcal{O}\left(\eta^{-\alpha+\varepsilon}\right) \sim n^{\alpha/(1+\alpha)}\Psi_\sigma\left(\frac{\omega\log n}{\alpha+1}\right)$$

for certain 1-periodic functions Φ_μ, Φ_σ and Ψ_μ, Ψ_σ (note that since $\log\eta \sim -\frac{\log n}{\alpha+1}$, the periods differ by a factor of $\alpha+1$).

$\mu_n = \sum_{m\in\mathcal{S}} \frac{1}{e^{\eta m}+1}$ is obviously a positive monotonic function of η, showing immediately that Φ_μ and Ψ_μ must be bounded above and below by positive constants (in the same way as it was shown that $r \gg n^{-1/(1+\alpha)}$). However, this approach does not apply that easily to σ_n (in particular, it is less obvious that Φ_σ cannot be identically 0), so we proceed a little differently : using the abbreviation $q(x) = \frac{e^x}{(e^x+1)^2}$, we can write the numerator of σ_n^2 in (2.4) as

$$\begin{aligned}
&\left(\sum_{m\in\mathcal{S}} \frac{e^{\eta m}}{(e^{\eta m}+1)^2}\right)\left(\sum_{m\in\mathcal{S}} \frac{m^2 e^{\eta m}}{(e^{\eta m}+1)^2}\right) - \left(\sum_{m\in\mathcal{S}} \frac{me^{\eta m}}{(e^{\eta m}+1)^2}\right)^2 \\
&\qquad = \sum_{m_1\in\mathcal{S}}\sum_{m_2\in\mathcal{S}} (m_2^2 - m_1 m_2) q(\eta m_1) q(\eta m_2) \\
&\qquad = \sum_{m_1\in\mathcal{S}}\sum_{\substack{m_2>m_1 \\ m_2\in\mathcal{S}}} (m_2-m_1)^2 q(\eta m_1) q(\eta m_2) \\
&\qquad = \frac{1}{2}\sum_{m_1\in\mathcal{S}}\sum_{m_2\in\mathcal{S}} (m_2-m_1)^2 q(\eta m_1) q(\eta m_2).
\end{aligned}$$

This can be estimated as follows :

$$\begin{aligned}
\frac{1}{2}\sum_{m_1\in\mathcal{S}}\sum_{m_2\in\mathcal{S}} (m_2-m_1)^2 q(\eta m_1) q(\eta m_2) &\geq \frac{1}{2}\sum_{\substack{m_1\leq\eta^{-1}/2 \\ m_1\in\mathcal{S}}}\sum_{\substack{\eta^{-1}\leq m_2\leq C\eta^{-1} \\ m_2\in\mathcal{S}}} (m_2-m_1)^2 q(\eta m_1) q(\eta m_2) \\
&\geq \frac{1}{2}\sum_{\substack{m_1\leq\eta^{-1}/2 \\ m_1\in\mathcal{S}}}\sum_{\substack{\eta^{-1}\leq m_2\leq C\eta^{-1} \\ m_2\in\mathcal{S}}} \frac{1}{4\eta^2} q(\eta m_1) q(\eta m_2) \\
&\gg \eta^{-2}\left(\sum_{\substack{m_1\leq\eta^{-1}/2 \\ m_1\in\mathcal{S}}} 1\right)\left(\sum_{\substack{\eta^{-1}\leq m_2\leq C\eta^{-1} \\ m_2\in\mathcal{S}}} 1\right)
\end{aligned}$$

$$\gg \eta^{-2-2\alpha}$$

by Lemma 3.8 and the corollary thereafter. The denominator has already been shown earlier to be of order $\eta^{-2-\alpha}$. Hence, $\sigma_n^2 \gg \eta^{-\alpha} \gg n^{\alpha/(1+\alpha)}$. Putting everything together, we arrive at

$$\begin{aligned} M_n(t) &= \exp\left(\frac{t^2}{2} + \mathcal{O}\left(n^{-\alpha/(2+2\alpha)}(t+t^3) + n^{-2\alpha/(7\alpha+7)}\right)\right) \\ &= \exp\left(\frac{t^2}{2} + \mathcal{O}\left(n^{-2\alpha/(7\alpha+7)}\right)\right) \end{aligned}$$

for bounded t. Now, Curtiss's Theorem [**61**] shows that the distribution of ϖ_n is indeed asymptotically normal. For the remaining parts of the theorem, we can again follow the lines of Hwang [**109**] : note that if $t = o(n^{\alpha/(6\alpha+6)})$, the above equation, together with Markov's inequality, yields

$$\begin{aligned} \mathbb{P}\left(\frac{\varpi_n - \mu_n}{\sigma_n} \geq x\right) &\leq e^{-tx} M_n(t) \\ &= e^{-tx+t^2/2}\left(1 + \mathcal{O}\left(n^{-\alpha/(2+2\alpha)}(t+t^3) + n^{-2\alpha/(7\alpha+7)}\right)\right). \end{aligned}$$

We set $T = n^{\alpha/(6\alpha+6)}/\log n$ and $t = x$ for $x \leq T$ (minimizing $-tx + t^2/2$) to obtain

$$\mathbb{P}\left(\frac{\varpi_n - \mu_n}{\sigma_n} \geq x\right) \leq e^{-x^2/2}\left(1 + \mathcal{O}\left((\log n)^{-3}\right)\right)$$

and for $x \geq T$, by setting $t = T$,

$$\mathbb{P}\left(\frac{\varpi_n - \mu_n}{\sigma_n} \geq x\right) \leq e^{-Tx/2}\left(1 + \mathcal{O}\left((\log n)^{-3}\right)\right).$$

The probability $\mathbb{P}\left(\frac{\varpi_n - \mu_n}{\sigma_n} \leq -x\right)$ can be estimated in an analogous way. Finally, we can also apply Hwang's method that was used in [**109**] to show that the mean and variance of ϖ_n are indeed asymptotic to μ_n and σ_n^2 respectively.

Remark 6. If the only pole of the Dirichlet series $D(s)$ is at $s = \alpha$ (so that the periodic functions Φ_μ and Φ_σ are actually constant), we obtain the asymptotic expressions for μ_n and σ_n^2 given in Theorem 3.2 : in this case, Lemma 3.6 yields

$$n = \sum_{m \in \mathcal{S}} \frac{m}{e^{\eta m} + 1} \sim A_0(1 - 2^{-\alpha})\Gamma(\alpha+1)\zeta(\alpha+1)\eta^{-(\alpha+1)} = \kappa\alpha\eta^{-(\alpha+1)},$$

where κ is taken as in Theorem 3.2, and

$$\begin{aligned} \mu_n = \sum_{m \in \mathcal{S}} \frac{1}{e^{\eta m} + 1} &\sim A_0(1 - 2^{1-\alpha})\Gamma(\alpha)\zeta(\alpha)\eta^{-\alpha} \\ &\sim A_0(1 - 2^{1-\alpha})\Gamma(\alpha)\zeta(\alpha)(\kappa\alpha)^{-\alpha/(\alpha+1)} n^{\alpha/(\alpha+1)} \\ &= (\kappa\alpha)^{1/(\alpha+1)} \frac{(1 - 2^{1-\alpha})\zeta(\alpha)}{\alpha(1 - 2^{-\alpha})\zeta(\alpha+1)} n^{\alpha/(\alpha+1)}, \end{aligned}$$

and an asymptotic formula for σ_n^2 follows in a similar manner.

REMARK 7. It might be possible to relax the conditions of our theorem, in particular (M1'), even further, so that the poles do not necessarily have to be evenly spaced. However, we are not aware of any natural example for which this actually occurs. Finally, the core of the proof, which is the application of the saddle point method, is in principle also applicable if the Dirichlet series has more complicated singularities (e.g. if $\mathcal{S}$ is the set of all primes), as long as one is able to obtain sufficiently strong upper and lower estimates for the harmonic sums involved.

4. Examples

As mentioned in the introduction, our initial motivating example was the set of integers with certain missing digits. However, there are also other natural examples for which Theorem 7.1 is applicable. At the end of this section, we also exhibit an example where our theorem fails because one of the conditions does not hold.

4.1. Missing digits. Recall that $\mathcal{MD}(b, D)$ denotes the set of positive integers with the property that all digits in the b-ary representation come from the set D. Lemma 3.3 shows immediately that condition (M1') is satisfied, since we have simple poles at $\alpha + 2\pi i k\omega$ with $\alpha = \frac{\log|D|}{\log b}$ and $\omega = (\log b)^{-1}$. (M2') is also obvious in view of Lemma 3.3. Condition (M3') can also be proved by elementary means : for $r > 0$ and $\frac{\pi}{2} \le |y| \le \pi$, choose k in such a way that $b^k \le r^{-1} < b^{k+1}$. Then we have

$$\begin{aligned}
g(r) - \mathrm{Re}g(r+iy) &= \sum_{m\in\mathcal{S}} e^{-mr}(1-\cos(my)) \ge \sum_{\substack{m<b^k\\ m\in\mathcal{S}}} e^{-mr}(1-\cos(my)) \\
&\ge e^{-1}\sum_{\substack{m<b^k\\ m\in\mathcal{S}}} (1-\cos(my)) = e^{-1}\mathrm{Re}\sum_{\substack{m<b^k\\ m\in\mathcal{S}}} (1-\exp(imy)) \\
&= e^{-1}\mathrm{Re}\left(|D|^k - \prod_{j=0}^{k-1}\sum_{d\in D}\exp(idb^j y)\right) \ge e^{-1}\left(|D|^k - \left|\prod_{j=0}^{k-1}\sum_{d\in D}\exp(idb^j y)\right|\right) \\
&\ge e^{-1}\left(|D|^k - \left|\sum_{d\in D}\exp(idy)\right||D|^{k-1}\right) \ge e^{-1}\left(1 - \frac{1}{|D|}\sup_{\frac{\pi}{2}\le|y|\le\pi}\left|\sum_{d\in D}\exp(idy)\right|\right)|D|^k \\
&\gg r^{-\log|D|/\log b},
\end{aligned}$$

which proves (M3'). Note that

$$\left|\sum_{d\in D}\exp(idy)\right| < |D|$$

for $\frac{\pi}{2} \le |y| \le \pi$: Since it was assumed that the digits in D do not have a common divisor > 1, dy cannot be a multiple of 2π for all d.

Figure 1 illustrates the periodic fluctuations in an example : here, the set of integers which do not contain the digit 2 in their ternary representation is considered. The plot shows the normalized mean of the length (i.e. $n^{-\alpha/(1+\alpha)}\mathbb{E}(\varpi_n)$, where $\alpha = \frac{\log 2}{\log 3}$ in this case) on a logarithmic scale ; the main term of the asymptotical formula is shown for comparison.

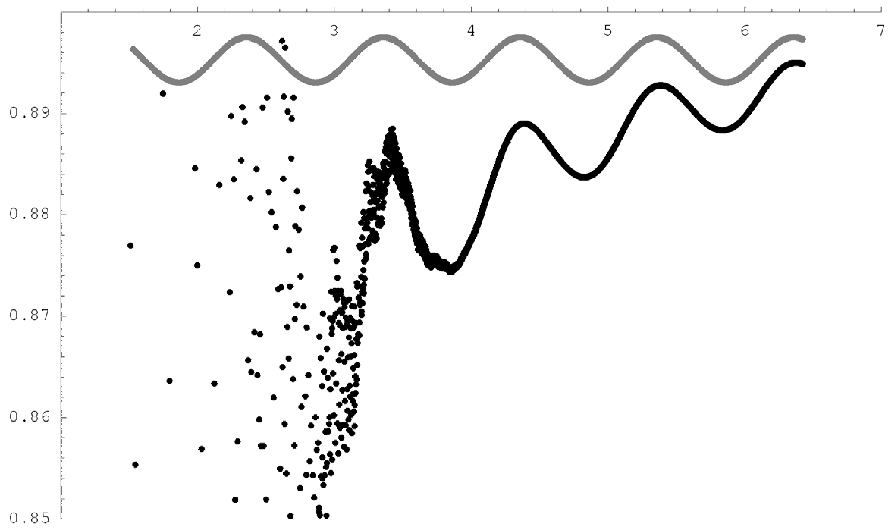


FIGURE 1. Periodic fluctuations of the mean in an example : $n^{-\alpha/(1+\alpha)}\mathbb{E}(\varpi_n)$ (black dots) and the periodic function Ψ_μ (gray line)

4.2. Missing blocks. The preceding example can easily be extended to integers with missing blocks in their digital expansion, such as the so-called "Fibbinary numbers" (see [**232**], sequence A003714, or [**13**]), *i.e.*, integers whose binary representation does not contain the

block 11 : $\{1,2,4,5,8,9,10,16,\ldots\}$. Write $\mathcal{F}$ for this set :

$$\mathcal{F} := \left\{ \sum_{i=0}^{k} a_i 2^i \middle| k \in \mathfrak{N}, a_i \in \{0,1\}, a_i \cdot a_{i+1} = 0 \right\}.$$

It is not difficult to show that the associated Dirichlet series satisfies our hypotheses (M1')-(M3') : noting that

$$\mathcal{F} = 2\mathcal{F} \cup (4\mathcal{F}+1) \cup \{1\},$$

we get

$$\begin{aligned}\left(1-2^{-s}-4^{-s}\right) D(s) &= 1 + \sum_{m\in\mathcal{F}} (4m+1)^{-s} - (4m)^{-s} \\ &\ll 1 + |s| \sum_{m\in\mathcal{F}} m^{-(\mathrm{Re}s+1)},\end{aligned}$$

which converges for $\mathrm{Re}s > 0$. Therefore, we have

$$D(s) = \left(1-2^{-s}-4^{-s}\right)^{-1} R(s)$$

for a Dirichlet series $R(s)$ that satisfies $R(s) \ll |s|$ uniformly for $\mathrm{Re}s \geq \varepsilon$. Hence $\alpha = \frac{\log((\sqrt{5}+1)/2)}{\log 2}$, and $\omega = \frac{1}{\log 2}$. Conditions (M1') and (M2') are satisfied as before. Furthermore, in order to prove that (M3') is also fulfilled, we note that

$$\mathcal{MD}(4,\{0,1\}) \subseteq \mathcal{F},$$

and so the considerations of the previous example show that

$$g(r) - \mathrm{Re}g(r+iy) = \sum_{m\in\mathcal{F}} e^{-mr}(1-\cos my) \geq \sum_{m\in\mathcal{MD}(4,\{0,1\})} e^{-mr}(1-\cos my) \gg r^{-1/2}.$$

4.3. Numbers with even/odd length. Let us now consider numbers whose b-ary representation has odd length : we obtain the sequence

$$\{1,4,5,6,7,16,17,\ldots\}$$

in the binary case, which is Sloane's A053738 [**232**]. Of course, this example can be generalized in many directions as well. Write $\mathcal{L}$ for the set of all such numbers, i.e.

$$\mathcal{L} := \left\{ \sum_{i=0}^{2k} a_i b^i \middle| k \in \mathfrak{N}, a_i \in \{0,1,\ldots,b-1\}, a_{2k} \neq 0 \right\}.$$

Thus we get for the Dirichlet generating function

$$D(s) = \sum_{k\geq 0} \sum_{m=b^{2k}}^{b^{2k+1}-1} m^{-s}.$$

Noting that

$$\mathcal{L} = \bigcup_{i=0}^{b^2-1} (b^2\mathcal{L}+i) \cup \{1,\ldots,b-1\},$$

we find

$$\begin{aligned} D(s) &= \sum_{i=0}^{b^2-1} \sum_{m\in\mathcal{L}} \big(b^2m+i\big)^{-s} + \sum_{m=1}^{b-1} m^{-s} \\ &= b^{2-2s}D(s) + \sum_{m=1}^{b-1} m^{-s} + \sum_{i=0}^{b^2-1} \sum_{m\in\mathcal{L}} \Big(\big(b^2m+i\big)^{-s} - \big(b^2m\big)^{-s}\Big) \\ &= b^{2-2s}D(s) + R(s). \end{aligned}$$

By the same method as above we see that $R(s)$ converges for $\mathrm{Re}s > 0$, and we get

$$D(s) = (1-b^{2-2s})^{-1}R(s),$$

which has poles at $s = 1 + \frac{\pi i}{\log b}$. As before, (M1') and (M2') hold with $\alpha = 1$ and $\omega = \frac{1}{2\log b}$. In order to prove that (M3') holds as well, we can use an elementary argument : choose K such that $b^{2K+1} \leq r^{-1} < b^{2K+3}$. Then we have

$$\begin{aligned} g(r) - \mathrm{Re}g(r+iy) &= \sum_{k\geq 0} \sum_{m=b^{2k}}^{b^{2k+1}-1} e^{-mr}(1-\cos my) \geq \sum_{m=b^{2K}}^{b^{2K+1}-1} e^{-mr}(1-\cos my) \\ &\geq e^{-1} \sum_{m=b^{2K}}^{b^{2K+1}-1} (1-\cos my) \\ &= e^{-1}\left((b-1)b^{2K} - \frac{\sin((b^{2K+1}-1/2)y) - \sin((b^{2K}-1/2)y)}{2\sin(y/2)}\right) \\ &= e^{-1}(b-1)b^{2K} + \mathcal{O}(1) \gg r^{-1}. \end{aligned}$$

4.4. Numbers with restricted sum of digits. Numbers whose b-ary sum of digits has to satisfy a certain congruence have been studied by Gel'fond [**91**] and Mauduit and Sárközy [**158**]. Their additive properties have been discussed in a paper by Thuswaldner and Tichy [**240**] and subsequent papers. This is actually an example for which there is only a single pole : it is not difficult to show that the Dirichlet series associated with the set of all integers whose b-ary sum of digits is $\equiv h \mod k$ is essentially $\frac{1}{k}\zeta(s)$. Hence, $\alpha = 1$, and the periodic functions Ψ_μ and Ψ_σ are actually constants. Let us illustrate this in the binary case : let $\mathcal{C}_0$ and $\mathcal{C}_1$ be the sets of positive integers for which the binary sum of digits is even resp. odd, and let $D_0(s)$ and $D_1(s)$ be the associated Dirichlet series. Then,

$$\mathcal{C}_0 = 2\mathcal{C}_0 \cup (2\mathcal{C}_1 + 1)$$

and thus

$$\begin{aligned} D_0(s) &= \sum_{m\in\mathcal{C}_0} (2m)^{-s} + \sum_{m\in\mathcal{C}_1} (2m+1)^{-s} = 2^{-s}(D_0(s)+D_1(s)) + \sum_{m\in\mathcal{C}_1} \big((2m+1)^{-s} - (2m)^{-s}\big) \\ &= 2^{-s}\zeta(s) + \sum_{m\in\mathcal{C}_1} (2m)^{-s} \sum_{k\geq 1} \binom{-s}{k} (2m)^{-k} = 2^{-s}\zeta(s) + \sum_{k\geq 1} \binom{-s}{k} 2^{-(s+k)} D_1(s+k). \end{aligned}$$

Remark 8. In all the aforementioned examples, one can also work with the squares, cubes, etc. of the numbers in $\mathcal{S}$, since the associated Dirichlet series is essentially the same. There are many other examples of fairly natural sets of integers whose associated Dirichlet

series has equidistant poles on a line of the form $\mathrm{Re} s = \alpha$; for instance, consider *palindromes* : the binary palindromes are

$$1, 3, 5, 7, 9, 15, 17, 21, \ldots$$

(Sloane's A006995 [**232**]). Since all of them are odd, the length of a partition of n will always have the same parity as n, but the central limit theorem still holds. Another example are numbers whose digital representation is the juxtaposition of two identical strings : in base 2, these are

$$3, 10, 15, 36, 45, 54, 63, 136, \ldots$$

(Sloane's A020330 [**232**]). In all these cases, (M1') and (M2') are fairly easy to check, proving (M3') is the difficult part.

Let us finally consider a trivial example for which (M1') and (M2') are satisfied, but (M3') is not :

4.5. An example that does not satisfy all conditions. Finally, we want to discuss an example of a sequence that does not satisfy all of our conditions. Let us consider the sequence $1, 4, 4, 16, 16, 16, 16, 64, \ldots$, i.e. 4^k appears with multiplicity 2^k. The corresponding Dirichlet series is extremely simple and obviously satisfies (M1') and (M2') (with $\alpha = \frac{1}{2}$ and $\omega = \frac{1}{2 \log 2}$) :

$$D(s) = \sum_{k \geq 0} \frac{2^k}{4^{ks}} = \frac{1}{1 - 2^{1-2s}}.$$

(M3') is violated, however, and there are even infinitely many integers which cannot be partitioned in this case (all those which are $\equiv 2, 3 \mod 4$, for instance), so that Theorem 7.1 cannot hold any longer.

Acknowledgment

We would like to thank Peter Grabner for his kind help with the proof of Lemma 3.8.

This paper was written while M. Madritsch was a visitor at the Centre of Experimental Mathematics at the University of Stellenbosch. He thanks the centre for its hospitality. He is also supported by the Austrian Science Foundation FWF, project S9603, that is part of the Austrian National Research Network "Analytic Combinatorics and Probabilistic Number Theory".

Chapitre 4

Asymptotic normality of additive functions on polynomial sequences in canonical number systems

This chapter is joint work with Attila Pethő and appeared in the *Journal of Number Theory*, **131** (2011) 1553 – 1574.

1. Introduction

In this paper we investigate the asymptotic behaviour of q-additive functions. But before we start we need an idea of additive functions and the number systems they are living in. Note that a function f is said to be *q-additive* if it acts only on the q-adic digits, *i.e.*, $f(0) = 0$ and

$$f(n) = \sum_{h=0}^{\ell} f(a_h(n)q^h) \quad \text{for} \quad n = \sum_{h=0}^{\ell} a_h(n)q^h,$$

where $a_h(n) \in \mathcal{N} := \{0, \dots, q-1\}$ are the *digits* of the q-adic expansion of n.

One of the first results dealing with the asymptotic behavior of such a q-additive function is the following, which is due to Bassily and Kátai [**19**].

THEOREM. *Let f be a q-additive function such that $f(aq^h) = \mathcal{O}(1)$ as $h \to \infty$ and $a \in \mathcal{N}$. Furthermore let*

$$m_{h,q} := \frac{1}{q} \sum_{a \in \mathcal{N}} f(aq^h), \quad \sigma_{h,q}^2 := \frac{1}{q} \sum_{a \in \mathcal{N}} f^2(aq^h) - m_{h,q}^2,$$

and

$$M_q(x) := \sum_{h=0}^{N} m_{h,q}, \quad D_q^2(x) = \sum_{h=0}^{N} \sigma_{h,q}^2$$

with $N = [\log_q x]$. Assume that $D_q(x)/(\log x)^{1/3} \to \infty$ as $x \to \infty$ and let P be a polynomial with integer coefficients, degree d and positive leading term. Then, as $x \to \infty$,

$$\frac{1}{x} \# \left\{ n < x \,\middle|\, \frac{f(P(n)) - M_q(x^d)}{D_q(x^d)} < y \right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} \exp(-x^2) \mathrm{d}x.$$

A first step towards a generalization of this concept is based on number systems living in an order in an algebraic number field.

DEFINITION 4.1. *Let $\mathcal{R}$ be an integral domain, $b \in \mathcal{R}$, and $\mathcal{N} = \{n_1, \dots, n_m\} \subset \mathbb{Z}$. Then we call the pair $(b, \mathcal{N})$ a number system in $\mathcal{R}$ if every $g \in \mathcal{R}$ admits a unique and finite representation of the form*

$$g = \sum_{h=0}^{\ell} a_h(g) b^h \quad \textit{with} \quad a_h(g) \in \mathcal{N}$$

and $a_h(g) \neq 0$ if $h \neq 0$. We call b the base and $\mathcal{N}$ the set of digits.

If $\mathcal{N} = \mathcal{N}_0 = \{0, 1, \dots, m\}$ for $m \geq 1$ then we call the pair $(b, \mathcal{N})$ a canonical number system.

When extending the number system to the complex plane one has to face effects such as amenability, *i.e.*, there may exist two or more different expansions of one number. In fact, one can construct a graph (the connection graph) which characterizes all the amenable expansions. This has been done by Müller *et al.* [**164**] (with a direct approach) and by Scheicher and Thuswaldner [**212**] (consideration of the odometer).

A different view on digits in number systems is done by normal numbers. These are numbers in which expansion every possible block occurs asymptotically equally often. Constructions of such numbers have been considered by Dumont *et al.* [**70**] and the first author in [**141**, **142**].

In this paper we mainly concentrate on additive functions. Thus we define additive functions in these number systems as follows.

Definition 4.2. *Let $(b, \mathcal{N})$ be a number system in the integral domain $\mathcal{R}$. A function f is called b-additive if $f(0) = 0$ and*

$$f(g) = \sum_{h \geq 0} f(a_h(g) b^h) \quad \text{for} \quad g = \sum_{h=0}^{\ell} a_h(g) b^h.$$

The simplest version of an additive function is the sum-of-digits function s_b defined by

$$s_b(g) := \sum_{h \geq 0} a_h(g).$$

The result by Bassily and Kátai was first generalized to number systems in the Gaussian integers by Gittenberger and Thuswaldner [**94**] who gained the following

Theorem. *Let $b \in \mathbb{Z}[i]$ and $(b, \mathcal{N})$ be a canonical number system in $\mathbb{Z}[i]$. Let f be a b-additive function such that $f(ab^h) = \mathcal{O}(1)$ as $h \to \infty$ and $a \in \mathcal{N}$. Furthermore let*

$$m_{h,b} := \frac{1}{\mathrm{N}(b)} \sum_{a \in \mathcal{N}} f(ab^h), \quad \sigma_{h,b}^2 := \frac{1}{\mathrm{N}(b)} \sum_{a \in \mathcal{N}} f^2(ab^h) - m_{h,b}^2,$$

and

$$M_b(x) := \sum_{h=0}^{L} m_{h,b}, \quad D_b^2(x) = \sum_{h=0}^{L} \sigma_{h,b}^2$$

with N *the norm of an element over $\mathbb{Q}$ and $L = [\log_{\mathrm{N}(b)} x]$.*

Assume that $D_b(x)/(\log x)^{1/3} \to \infty$ as $x \to \infty$ and let P be a polynomial of degree d with coefficients in $\mathbb{Z}[i]$. Then, as $N \to \infty$,

$$\frac{1}{\#\{z \in \mathbb{Z}[i] | \mathrm{N}(z) < N\}} \# \left\{ \mathrm{N}(z) < N \,\middle|\, \frac{f(P(z)) - M_b(N^d)}{D_b(N^d)} < y \right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} \exp(-x^2) \mathrm{d}x,$$

where z runs over the Gaussian integers.

This build the base for further considerations of b-additive functions in algebraic number fields in general. Therefore let $\mathcal{K} = \mathbb{Q}(\beta)$ be an algebraic number field and denote by $\mathcal{O}_\mathcal{K}$ its ring of integers (aka its maximal order). Furthermore let $\beta \in \mathcal{O}_\mathcal{K}$ then we set $\mathcal{R} = \mathbb{Z}[\beta]$ to be an order in $\mathcal{K}$. We now want to analyze additive functions for number systems in $\mathcal{R}$.

We need some more parameters in order to successfully generalize the theorem from above. Thus let $\mathcal{K}^{(\ell)}$ $(1 \leq \ell \leq r_1)$ be the real conjugates of $\mathcal{K}$, while $\mathcal{K}^{(m)}$ and $\mathcal{K}^{(m+r_2)}$ $(r_1 < m \leq r_1 + r_2)$ are the pairs of complex conjugates of $\mathcal{K}$, where $r_1 + 2r_2 = n$.

For $\gamma \in \mathcal{K}$ we denote by $\gamma^{(i)}$ $(1 \leq i \leq n)$ the conjugates of γ. In order to extend the term of conjugation to the completion $\overline{\mathcal{K}}$ of $\mathcal{K}$ we define for $\gamma_j \in \mathcal{K}$ and $x_j \in \mathbb{R}$ $(1 \leq j \leq n)$ $\lambda = \sum_{1 \leq j \leq n} x_j \gamma_j$ and $\lambda^{(i)} := \sum_{1 \leq j \leq n} x_j \gamma_j^{(i)}$.

Next we have to guarantee that we choose the increasing set for our asymptotic distribution. In the integer case we had the logarithm of the value, since the length of expansion grows with the logarithm. Since $\mathcal{R}$ is of dimension n we need a way to enlarge the area under consideration such that the expansion grows also in a smooth way. This is motivated by the following

LEMMA 4.3 ([**130**, Theorem]). *Let $\ell(\gamma)$ be the length of the expansion of γ to the base b. Then*

$$\left| \ell(\gamma) - \max_{1 \leq i \leq n} \frac{\log \left|\gamma^{(i)}\right|}{\log \left|b^{(i)}\right|} \right| \leq C.$$

Therefore we define $\mathcal{R}(T_1, \ldots, T_r)$ to be the set

$$\mathcal{R}(T_1, \ldots, T_r) := \left\{ \lambda \in \mathcal{R} : \left|\lambda^{(i)}\right| \leq T_i, 1 \leq i \leq r \right\}. \tag{1.1}$$

Now we use Lemma 4.3 to bound the area $\mathcal{R}(T_1, \ldots, T_n)$ such that we reach all elements of a certain length. Thus for a fixed T we set T_i for $1 \leq i \leq n$ such that

$$\log T_i = \log T \frac{\log \left|b^{(i)}\right|^n}{\log |\mathrm{N}(b)|}. \tag{1.2}$$

Furthermore we will write for short $\mathcal{R}(\mathbf{T}) := \mathcal{R}(T_1, \ldots, T_r)$ with T_i as in (1.2).

Finally one can extend the definition of a number system also for negative powers of b. Then for $\gamma \in \overline{\mathcal{K}}$ such that

$$\gamma = \sum_{h=-\infty}^{\ell} a_h b^h \quad \text{with} \quad a_h \in \mathcal{N}$$

we call

$$\lfloor \gamma \rfloor := \sum_{h=0}^{\ell} a_h b^h \quad \text{and} \quad \{\gamma\} := \sum_{h \geq 1} a_h b^{-h}$$

the *integer part* and *fractional part* of γ, respectively.

With all these tools we now can state the generalization of the theorem of Bassily and Kátai to arbitrary number fields.

THEOREM 4.4 ([**143**]). *Let $(b, \mathcal{N})$ be a number system in $\mathcal{R}$ and f be a b-additive function such that $f(ab^h) = \mathcal{O}(1)$ as $h \to \infty$ and $a \in \mathcal{N}$. Furthermore let*

$$m_{h,b} := \frac{1}{\mathrm{N}(b)} \sum_{a \in \mathcal{N}} f(ab^h), \quad \sigma_{h,b}^2 := \frac{1}{\mathrm{N}(b)} \sum_{a \in \mathcal{N}} f^2(ab^h) - m_{h,b}^2,$$

and

$$M_b(x) := \sum_{h=0}^{L} m_{h,q}, \quad D_b^2(x) = \sum_{h=0}^{L} \sigma_{h,q}^2$$

with $L = [\log_{\mathrm{N}(b)} x]$.

Assume that there exists an $\varepsilon > 0$ *such that* $D_b(x)/(\log x)^\varepsilon \to \infty$ *as* $x \to \infty$ *and let* $P \in \overline{\mathcal{K}}[X]$ *be a polynomial of degree* d. *Then, as* $T \to \infty$ *let* T_i *be as in* (1.2),

$$\frac{1}{\#\mathcal{R}(\mathbf{T})}\#\left\{z \in \mathcal{R}(\mathbf{T}) \middle| \frac{f(\lfloor P(z)\rfloor) - M_b(T^d)}{D_b(T^d)} < y\right\} \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} \exp(-x^2)\mathrm{d}x.$$

2. Definitions and result

The objective of this paper are generalizations of number systems to quotient rings of the ring of polynomials over the integers. Our aim is to extend Theorem 4.4 to such rings. To formulate our results we have to introduce the relevant notions. In particular we use the following definition in order to describe number systems in quotient rings of the ring of polynomials over the integers.

Definition 4.5. *Let* $p \in \mathbb{Z}[X]$ *be monic of degree* n *and let* $\mathcal{N}$ *be a subset of* $\mathbb{Z}$. *The pair* $(p, \mathcal{N})$ *is called a number system if for every* $g \in \mathbb{Z}[X] \setminus \{0\}$ *there exist unique* $\ell \in \mathfrak{N}$ *and* $a_h \in \mathcal{N}, h = 0, \ldots, \ell; a_\ell \neq 0$ *such that*

$$g \equiv \sum_{h=0}^{\ell} a_h(g) X^h \pmod{p}. \tag{2.1}$$

In this case a_h *are called the digits and* $\ell = \ell(a)$ *the length of the representation.*

This concept was introduced in [**179**] and was studied among others in [**3**, **6**, **129**, **130**]. It was proved in [**6**], that $\mathcal{N}$ must be a complete residue system modulo $p(0)$ including 0 and the zeroes of p are lying outside or on the unit circle. However, following the argument of the proof of Theorem 6.1 of [**179**], which dealt with the case p square free, one can prove that non of the zeroes of p are lying on the unit circle.

If p is irreducible then we may replace X by one of the roots β of p. Then we are in the case of $\mathbb{Z}[X]/(p) \cong \mathbb{Z}[\beta]$ being an integral domain in an algebraic number field (*cf.* Section 1). Then we may also denote the number system by the pair$(\beta, \mathcal{N})$ instead of $(p, \mathcal{N})$. For example, let $q \geq 2$ be a positive integer, then $(p, \mathcal{N})$ with $p = X - q$ gives a number system in $\mathbb{Z}$, which corresponds to the number systems $(q, \mathcal{N})$. Furthermore for n a positive integer and $p = X^2 + 2n\,X + (n^2 + 1)$ we get number systems in $\mathbb{Z}[i]$.

Now we want to come back to these more general number systems and consider additive functions within them.

Definition 4.6. *Let* $(p, \mathcal{N})$ *be a number system. A function* f *is called additive if* $f(0) = 0$ *and*

$$f(g) \equiv \sum_{h=0}^{\ell} f(a_h(g) X^h) \pmod{p} \quad \textit{for} \quad g \equiv \sum_{h=0}^{\ell} a_h(g) X^h \pmod{p}.$$

Since we have defined the analogues of number systems and additive functions to the definitions for number fields above, we now need to extend the length estimation of Lemma 4.3 in order to successfully state the result. But before we start we need a little linear algebra. We fix a number system $(p, \mathcal{N})$ and factor p by

$$p := \prod_{i=1}^{t} p_i^{m_i}$$

with $p_i \in \mathbb{Z}[X]$ irreducible and $\deg p_i = n_i$. Furthermore we denote by β_{ik} the roots of p_i for $i = 1, \ldots, t$ and $k = 1, \ldots, n_i$.

Then we define by

$$\mathcal{R} := \mathbb{Z}[X]/(p) = \bigoplus_{i=1}^{t} \mathcal{R}_i \quad \text{with} \quad \mathcal{R}_i = \mathbb{Z}[X]/\left(p_i^{m_i}\right)$$

for $i = 1, \ldots, t$ the $\mathbb{Z}$-module under consideration and in the same manner by

$$\mathcal{K} := \mathbb{Q}[X]/(p) = \bigoplus_{i=1}^{t} \mathcal{K}_i \quad \text{with} \quad \mathcal{K}_i = \mathbb{Q}[X]/\left(p_i^{m_i}\right)$$

for $i = 1, \ldots, t$ the corresponding vector space. Finally we denote by $\overline{\mathcal{K}}$ the completion of $\mathcal{K}$ according to the usual Euclidean distance.

Obviously $\mathcal{R}$ is a free $\mathbb{Z}$-module of rank n. Let $\lambda : \mathcal{R} \to \mathcal{R}$ be a linear mapping and $\{z_1, \ldots, z_n\}$ be any basis of $\mathcal{R}$. Then

$$\lambda(z_j) = \sum_{i=1}^{n} a_{ij} z_i \quad (j = 1, \ldots, n)$$

with $a_{ij} \in \mathbb{Z}$. The matrix $M(\lambda) = (a_{ij})$ is called the matrix of λ with respect to the basis $\{z_1, \ldots, z_n\}$. For an element $r \in \mathcal{R}$ we define by $\lambda_r : \mathcal{R} \to \mathcal{R}$ the mapping of multiplication by r; that is $\lambda_r(z) = rz$ for every $z \in \mathcal{R}$. Then we define the norm $\mathrm{N}(r)$ and the trace $\mathrm{Tr}(r)$ of an element $r \in \mathcal{R}$ as the determinant and the trace of $M(\lambda_r)$, respectively, *i.e.*,

$$\mathrm{N}(r) := \det(M(\lambda_r)), \quad \mathrm{Tr}(r) := \mathrm{Tr}(M(\lambda_r)).$$

Note that these are unique despite of the used basis $\{z_1, \ldots, z_n\}$. We can canonically extend these notions to $\mathcal{K}$ and $\overline{\mathcal{K}}$ by everywhere replacing $\mathbb{Z}$ by $\mathbb{Q}$ and $\mathbb{R}$, respectively.

In the following we will need parameters which help us bounding the length of the expansion of an element $g \in \mathcal{R}$. Therefore let $g \in \mathbb{Z}[X]$ be a polynomial, then we put

$$B_{ijk}(g) := \left.\frac{\mathrm{d}^{j-1} g}{\mathrm{d}X^{j-1}}\right|_{X=\beta_{ik}} \qquad (i = 1, \ldots, t; j = 1, \ldots, m_i; k = 1, \ldots, n_i).$$

In connection with these values we define the "house" function H as

$$H(g) := \max_{i=1}^{t} \max_{j=1}^{m_i} \max_{k=1}^{n_i} |B_{ijk}(g)| .$$

We want to investigate the elements with bounded maximum length of expansion. To this end we need a proposition which estimates the length of expansion in connection with properties of the number itself. The proof of this proposition will be presented in the following section.

PROPOSITION 4.7. *Assume that $(p, \mathcal{N})$ is a number system. Let $N = \max\{|a| : a \in \mathcal{N}\}$ and we set*

$$M(g) := \max\left\{ \frac{\log |B_{ijk}(g)|}{\log |\beta_{ik}|} : i = 1, \ldots, t; j = 1, \ldots, m_i; k = 1, \ldots, n_i \right\}.$$

If $g \in \mathbb{Z}[X]$ is of degree at most $n-1$, then for any $\varepsilon > 0$ there exists $L = L(\varepsilon)$ such that if $\ell(g) > L$ then

$$|\ell(g) - M(g)| \leq C. \tag{2.2}$$

This provides us with an estimation for the length of the expansion and motivates us to look at subsets of $\mathcal{R}$ where the absolute values B_{ijk} are bounded. For a vector $\mathbf{T} := (T_1, \dots, T_n) = (T_{111}, \dots, T_{11n_1}, T_{121}, \dots, T_{1,m_i,n_1}, T_{211}, \dots, T_{t,m_t,n_t})$ we denote by

$$\mathcal{R}(\mathbf{T}) := \{g \in \mathcal{R} : |B_{ijk}(g)| \le T_{ijk}\} \tag{2.3}$$

$$\mathcal{R}_i(\mathbf{T}) := \{g \in \mathcal{R}_i : |B_{ijk}(g)| \le T_{ijk}\}. \tag{2.4}$$

We want to let the length of expansion to smoothly increase. Therefore we fix a T and set T_{ijk} for $i = 1, \dots, t$, $j = 1, \dots, m_i$, $k = 1, \dots, n_i$ such that

$$\log T_{ijk} = \log T \frac{\log |\beta_{ik}|^n}{\log \prod_{i=1}^{t} \prod_{k=1}^{n_i} |\beta_{ik}|^{m_i}}. \tag{2.5}$$

Remark that T_{ijk} is independent from j, which will be important in Lemma 4.12. In view of Proposition 4.7 we get that the expansions of the elements in $\mathcal{R}(\mathbf{T})$ almost have the same maximum length. If not stated otherwise we denote by $\mathbf{T}$ the vector $(T_{111}, \dots, T_{t,m_t,n_t})$ where the T_{ijk} are as in (2.5).

Since X is an invertible element in $\mathcal{K}$ we may extend the definition of a number system for negative powers of X. Then for $\gamma \in \overline{\mathcal{K}}$ such that

$$\gamma = \sum_{h=-\infty}^{\ell} a_h X^h \quad \text{with} \quad a_h \in \mathcal{N}$$

we call

$$\lfloor \gamma \rfloor := \sum_{h=0}^{\ell} a_h X^h \quad \text{and} \quad \{\gamma\} := \sum_{h=-\infty}^{-1} a_h X^h$$

the *integer part* and *fractional part* of γ, respectively.

Now we have collected all the tools to state our main result.

THEOREM 4.8. *Let $(p, \mathcal{N})$ be a number system and f be an additive function such that $f(aX^h) = \mathcal{O}(1)$ as $h \to \infty$ and $a \in \mathcal{N}$. Furthermore let*

$$m_h := \frac{1}{|\mathcal{N}|} \sum_{a \in \mathcal{N}} f(aX^h), \quad \sigma_h^2 := \frac{1}{|\mathcal{N}|} \sum_{a \in \mathcal{N}} f^2(aX^h) - m_h^2,$$

and

$$M(x) := \sum_{h=0}^{L} m_h, \quad D^2(x) := \sum_{h=0}^{L} \sigma_h^2,$$

where $L = \lfloor \log_{p(0)} x \rfloor$. Assume that there exists an $\varepsilon > 0$ such that $D(x)/(\log x)^\varepsilon \to \infty$ as $x \to \infty$ and let $P \in \overline{\mathcal{K}}[Y]$ be a polynomial of degree d. Then, as $T \to \infty$ let T_{ijk} be as in (2.5),

$$\frac{1}{\#\mathcal{R}(T)} \# \left\{ z \in \mathcal{R}(T) : \frac{f(\lfloor P(z) \rfloor) - M(T^d)}{D(T^d)} < y \right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} \exp(-x^2) \mathrm{d}x.$$

Our theorem shows that the distribution properties of patterns in the sequence of digits depend neither on the polynomial p nor on the quotient ring $\mathcal{R}$. They are intimate properties of the "backward division algorithm" defined in [**179**].

We will show the main theorem in several steps.

(1) In the following section we will show properties of number systems which we need on the one hand to estimate the length of expansion and on the other hand to provide us with an Urysohn function, that helps us counting the occurrences of a fixed pattern of digits in the expansion.

(2) Equipped with these tools we will estimate the exponential sums occurring in the proof in Section 4. Therefore we need to split the module $\mathcal{R}$ up into its components and consider each of them separately. We also show that we may neglect the nilpotent elements.

(3) Now we take a closer look at the Urysohn function, which will count the occurences of our pattern in the expansion, and estimate the number of hits of the border of this function in Section 5. In particular, we count the number of hits of the area, where the function value lies between 0 and 1 as this area corresponds to the error term.

(4) In Section 6 we will show that any chosen patterns of digit and position occurs uniformly in the expansions. This will be our central tool in the proof of Theorem 4.8.

(5) Finally we draw all the thinks together. The main idea here is to use the growth rate of the deviation together with the Fréchet-Shohat Theorem to cut of the head and the tail of the expansion. Then an application of the central Proposition 4.16 and juxtaposition of the moments will prove the result.

3. Number system properties

In this section we want to show two properties we need in the sequel. The first deals with the above mentioned estimation of the length of an expansion (Proposition 4.7). We will need this result in order to justify our choice of $\mathbf{T}$ as in (2.5). Secondly we construct the Urysohn function for indicating the elements starting with a certain digit. The main idea is to embed the elements of $\mathcal{R}$ in $\mathbb{R}^n$ and to use the properties of matrix number systems in this field.

We start with the

Proof of Proposition 4.7. In the proof we combine ideas from [**6**] and [**130**].

We may assume $g \neq 0$. As $\{p, \mathcal{N}\}$ is a number system there are $\ell = \ell(g)$ and $a_h \in \mathcal{N}$ for $h = 0, \dots, \ell$; $a_\ell \neq 0$ such that

$$g \equiv \sum_{h=0}^{\ell} a_h X^h \pmod{p},$$

i.e.,

$$g = \sum_{h=0}^{\ell} a_h X^h + r\,p$$

with a polynomial $r \in \mathbb{Z}[X]$. For $j \geq 1$ this implies

$$\frac{\mathrm{d}^{j-1} g}{\mathrm{d}X^{j-1}} = \sum_{h=j-1}^{\ell} \frac{h!}{(h-j+1)!} a_h X^{h-j+1} + \sum_{s=0}^{j-1} \binom{j-1}{s} \frac{\mathrm{d}^s r}{\mathrm{d}X^s} \frac{\mathrm{d}^{j-1-s} p}{\mathrm{d}X^{j-1-s}}. \tag{3.1}$$

Consider a zero β_{ik} of p, which has multiplicity m_i. As we noticed in the Introduction, the argument of the proof of Theorem 6.1. of [**179**] allows to prove that $|\beta_{ik}| > 1$ for all $i = 1, \dots, t; k = 1, \dots, n_i$. Inserting β_{ik} into (3.1) we obtain

$$B_{ijk}(g) = \left.\frac{\mathrm{d}^{j-1}g}{\mathrm{d}X^{j-1}}\right|_{X=\beta_{ik}} = \sum_{h=j-1}^{\ell} \frac{h!}{(h-j+1)!} a_h \beta_{ik}^{h-j+1}$$

for $i = 1, \dots, t$ and $j = 1, \dots, m_i$. This implies by taking absolute value

$$\begin{aligned} |B_{ijk}(g)| &\le N \sum_{h=j-1}^{\ell} \frac{h!}{(h-j+1)!} |\beta_{ik}|^{h-j+1} \\ &\le N \frac{\ell!}{(\ell-j+1)!} |\beta_{ik}^{\ell-j+1}| \sum_{h=j-1}^{\ell} \frac{h(h-1)\dots(h-j+1)}{\ell(\ell-1)\dots(\ell-j+1)} |\beta_{ik}|^{h-\ell} \\ &\le N \frac{\ell^{j-1}|\beta_{ik}|^{\ell}}{|\beta_{ik}|-1}, \end{aligned}$$

which verifies the lower bound for ℓ, because $|\beta_{ik}| > 1$.

Now we turn to prove the upper bound. Denote by $V = V_p$ the following mapping : for $g \in \mathbb{Z}[X]$ of degree at most $n-1$ choose an $a \in \mathcal{N}$ such that $g(0) \equiv a \pmod{p(0)}$. Such an a exists by Theorem 6.1 of [**179**]. Putting $q = \frac{g(0)-a}{p(0)}$, let $V(g) = \frac{g-q\cdot p-a}{X}$. Obviously $V(g) \in \mathbb{Z}[X]$ and has degree at most $n-1$, thus V can be iterated. Moreover we have

$$g \equiv \sum_{h=0}^{u} a_h X^h + X^{u+1} V^{u+1}(g) \pmod{p} \tag{3.2}$$

with $a_h \in \mathcal{N}$ for $h = 0, \dots, u$.

Choose u the largest integer satisfying $|B_{ijk}(g)| \ge \frac{u^j |\beta_{ik}|^{u-j+1}}{|\beta_{ik}|-1}$ for all $i = 1, \dots, t$, $j = 1, \dots, m_i$ and $k = 1, \dots, n_i$. Then $u \le (1+\varepsilon/2)M(A)$. Proceeding like in the previous case we get

$$\begin{aligned} B_{ijk}(g) = \left.\frac{\mathrm{d}^{j-1}g}{\mathrm{d}X^{j-1}}\right|_{X=\beta_{ik}} &= \sum_{h=j}^{u} \frac{h!}{(h-j+1)!} a_h \beta_{ik}^{h-j} \\ &+ \sum_{s=0}^{j} \binom{j-1}{s} \frac{(u+1)!}{(u+1-s)!} \beta_{ik}^{u+1-s} \left.\frac{\mathrm{d}^{j-s-1}V^{u+1}(g)}{\mathrm{d}X^{j-s-1}}\right|_{X=\beta_{ik}}. \end{aligned}$$

By its definition $V^{u+1}(g)$ has integer coefficients. Dividing the last equation by $\frac{(u+1)!}{(u+1-j)!}\beta_{ik}^{u+1-j}$ and consider the obtained equations for $i = 1, \dots, t$ and $k = 1, \dots, n_i$ and for fixed i and k for $j = 1, \dots, m_i$ successively, then using the choice of u and that $|\beta_{ik}| > 1$ we conclude that

$$\left.\frac{\mathrm{d}^{j-s-1}V^{u+1}(g)}{\mathrm{d}X^{j-s-1}}\right|_{X=\beta_{ik}} < c,$$

where c is a constant depending only on N as well as the size and the multiplicities of the zeroes of p. These can be considered as n inequalities for the n unknown coefficients of $V^{u+1}(g)$. Furthermore the determinant of the coefficient matrix is not zero (c.f. [**6**]). Thus the solutions are bounded. As they are integers there are only finitely many possibilities for $V^{u+1}(g)$. As $\{p, \mathcal{N}\}$ is a number system, $V^{u+1}(g)$ has a representation, which length is bounded by a

constant, say c_1, which depends only on N as well as the size and the multiplicities of the zeroes of p. Thus $\ell(g) \leq u + c_1 \leq (1+\varepsilon)M(g)$ and the proposition is proved. □

Now we turn our attention back to the counting of the numbers and in particular to the construction of the Urysohn function. In order to properly count the elements we need the fundamental domain, which is defined as the set of all numbers whose integer part is zero. Since this is not so easy to define in this context we want to consider its embedding in $\mathbb{R}^n$. The main idea is to use the corresponding matrix of the polynomial p and to use properties of matrix number systems. This idea essentially goes back to Gröchenig and Haas [**98**]. The following definitions are standard in that area and we mainly follow Gittenberger and Thuswaldner [**94**] and Madritsch [**143**].

We note that if $(p, \mathcal{N})$ is a number system then X is a integral power base of $\overline{\mathcal{K}}$, *i.e.*, $\{1, X, \ldots, X^{n-1}\}$ is an $\mathbb{R}$-basis for $\overline{\mathcal{K}}$. Thus we define the embedding ϕ by

$$\phi : \begin{cases} \overline{\mathcal{K}} & \to \quad \mathbb{R}^n, \\ a_1 + a_2 X + \cdots + a_n X^{n-1} & \mapsto \quad (a_1, \ldots, a_n). \end{cases}$$

Now let $p = b_{n-1}X^{n-1} + \cdots + b_1 X + b_0$. Then we define the corresponding matrix B by

$$B = \begin{pmatrix} 0 & 0 & \cdots & \cdots & \cdots & -b_0 \\ 1 & 0 & \cdots & \cdots & 0 & \vdots \\ 0 & 1 & \ddots & & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & 1 & 0 & \vdots \\ 0 & 0 & \cdots & 0 & 1 & -b_{n-1} \end{pmatrix}. \tag{3.3}$$

One easily checks that $\phi(X \cdot p) = B \cdot \phi(p)$. Since B is invertible we can extend the definition of ϕ by setting for an integer h

$$\phi(X^h \cdot p) := B^h \phi(p). \tag{3.4}$$

By this we define the (embedded) fundamental domain by

$$\mathcal{F} := \left\{ z \in \mathbb{R}^n \middle| z = \sum_{h \geq 1} B^{-h} a_h, a_h \in \phi(\mathcal{N}) \right\}.$$

Following Gröchenig and Haas [**98**] we get that

$$\lambda((\mathcal{F} + g_1) \cap (\mathcal{F} + g_2)) = 0$$

for every $g_1, g_2 \in \mathbb{Z}^n$ with $g_1 \neq g_2$, where λ denotes the n dimensional Lebesgue measure. Thus $(B, \phi(\mathcal{N}))$ is a matrix number system and a so called *just touching covering system*. Therefore we are allowed to apply the results of the paper by Müller et al. [**164**].

We now follow the lines of Madritsch [**143**] where the ideas of Gittenberger and Thuswaldner [**94**] were combined with the results of Kátai and Környei [**117**] and Müller et al. [**164**].

Our main interest is the fundamental domain consisting of all numbers whose first digit equals $a \in \mathcal{N}$, *i.e.*,

$$\mathcal{F}_a = B^{-1}(\mathcal{F} + \phi(a)).$$

Imitating the proof of Lemma 3.1 of [**94**] we get the following.

LEMMA 4.9. *For all $a \in \mathcal{N}$ and all $v \in \mathfrak{N}$ there exist a $1 \leq \mu < |\det B|$ and an axe-parallel tube $P_{v,a}$ with the following properties :*

- *$\partial \mathcal{F}_a \subset P_{v,a}$ for all $v \in \mathfrak{N}$,*
- *the Lebesgue measure of $P_{v,a}$ is a $\mathcal{O}\left(\frac{\mu^v}{|\det B|^v}\right)$,*
- *$P_{v,a}$ consists of $\mathcal{O}(\mu^v)$ axe-parallel rectangles, each of which has Lebesgue measure $\mathcal{O}(|\det B|^v)$,*

where λ denotes the Lebesgue measure.

As in the proof of Lemma 3.1 of [**94**] we can construct for each pair (v, a) an axe-parallel polygon $\Pi_{v,a}$ and the corresponding tube

$$P_{v,a} := \left\{ z \in \mathbb{R}^n \middle| \|z - \Pi_{v,a}\|_\infty \leq 2c_p \, |\det B|^{-v} \right\},$$

where c_p is an arbitrary constant. Furthermore we denote by $I_{v,a}$ the set of all points inside $\Pi_{v,a}$. Now we define our Urysohn function u_a by

$$u_a(x_1, \ldots, x_n) = \frac{1}{\kappa^n} \int_{-\frac{\kappa}{2}}^{\frac{\kappa}{2}} \cdots \int_{-\frac{\kappa}{2}}^{\frac{\kappa}{2}} \psi_a(x_1 + y_1, \ldots, x_n + y_n) \, \mathrm{d}y_1 \cdots \mathrm{d}y_n, \tag{3.5}$$

where

$$\kappa := 2c_u \, |\det B|^{-v} \tag{3.6}$$

with c_u a constant and

$$\psi_a(x_1, \ldots, x_n) = \begin{cases} 1 & \text{if } (x_1, \ldots, x_n) \in I_{v,a} \\ \frac{1}{2} & \text{if } (x_1, \ldots, x_n) \in \Pi_{v,a} \\ 0 & \text{otherwise.} \end{cases}$$

Thus u_a is the desired Urysohn function which equals 1 for $z \in I_{v,a} \setminus P_{v,a}$, 0 for $z \in \mathbb{R}^n \setminus (I_{v,a} \cup P_{v,a})$, and linear interpolation in between.

We now do a Fourier transform of u_a and estimate the coefficients in the same way as in Lemma 3.2 of [**94**].

LEMMA 4.10. *Let $u_a(x_1, \ldots, x_n) = \sum_{(m_1,\ldots,m_n) \in \mathbb{Z}^n} c_{m_1,\ldots,m_n} e(m_1 x_1 + \cdots + m_n x_n)$ be the Fourier series of u_a. Then the Fourier coefficients $c_{m_1,\ldots,m_n}$ can be estimated by*

$$c_{0,\ldots,0} = \frac{1}{|\det B|}, \quad c_{m_1,\ldots,m_n} \ll \mu^v \prod_{i=1}^{n} \frac{1}{r(m_i)}$$

with

$$r(m_i) = \begin{cases} \kappa m_i & m_i \neq 0, \\ 1 & m_i = 0. \end{cases}$$

4. Estimation of the Weyl Sum

Before we continue with the estimation of the number of points inside the fundamental domain and those hitting the boarder, we want to estimate the exponential sums, which will occur in the following sections. In particular we want to prove the following.

PROPOSITION 4.11. *Let $T \geq 0$ and T_{ijk} as in (2.5). Let L be the maximum length of the b-adic expansion of $z \in \mathcal{R}(\mathbf{T})$ and let C_1 and C_2 be sufficiently large constants. Furthermore let $l_1, \dots, l_h$ be positions and $\mathbf{h}_1, \dots, \mathbf{h}_h$ be corresponding n-dimensional vectors. If*

$$C_1 \log L \leq l_1 < l_2 < \cdots < l_h \leq dL - C_2 \log L \tag{4.1}$$

and

$$\|\mathbf{h}_r\|_\infty \leq (\log T)^{\sigma_1} \tag{4.2}$$

for $1 \leq r \leq h$, then we have

$$\sum_{z \in \mathcal{R}(\mathbf{T})} e\left(\sum_{r=1}^{h} \left\langle \mathbf{h}_r, B^{-l_r-1}\phi(P(z)) \right\rangle\right) \ll T^n (\log T)^{-t\sigma_0}$$

where σ_0 depends on σ_1, C_1 and C_2.

Our main idea consists in several steps. First we will split the ring $\mathcal{R}$ up into the $\mathcal{R}_i$ and consider each of them separately. Then we distinguish two cases according to whether $m_i > 1$ or not. The latter reduces to an estimation of the sum in an algebraic number field. Whereas for the case of $m_i > 1$ we have to deal with nilpotent elements. Therefore we divide $\mathcal{R}_i$ into the radical and the nilpotent elements. Thus we define $\widetilde{\mathcal{R}_i}$ as

$$\widetilde{\mathcal{R}_i} := \mathbb{Z}[X]/(p_i) \quad \text{and} \quad \widetilde{\mathcal{R}_i}(\mathbf{T}) := \left\{ g \in \widetilde{\mathcal{R}_i} : |B_{ijk}(g)| \leq T_{ijk} \right\} \tag{4.3}$$

and the set $\mathcal{N}_i$ to be the nilpotent elments, *i.e.*,

$$\mathcal{N}_i := \{g \in \mathcal{R}_i : g \equiv 0 \bmod p_i\} \quad \text{and} \quad \mathcal{N}_i(\mathbf{T}) := \{g \in \mathcal{R}_i(\mathbf{T}) : g \equiv 0 \bmod p_i\}. \tag{4.4}$$

But before we start with the proof we need to show that the estimation is good compared with the trivial one. Thus we will show the following.

LEMMA 4.12. *Let T_{ijk} for $i = 1, \dots, t$, $j = 1, \dots, m_i$, $k = 1, \dots, n_i$ be positive reals. Then*

$$\begin{aligned}
\#\mathcal{R}(\mathbf{T}) &= \prod_{i=1}^{t} \#\mathcal{R}_i(\mathbf{T}), \\
\#\mathcal{R}_i(\mathbf{T}) &= c_i \left(\prod_{k=1}^{n_i} T_{i1k} \right)^{m_i} + \mathcal{O}\left(T_0^{m_i n_i - 1} \right), \\
\#\mathcal{N}_i(\mathbf{T}) &= c_i \left(\prod_{k=1}^{n_i} T_{i1k} \right)^{m_i - 1} + \mathcal{O}\left(T_0^{(m_i-1)n_i - 1} \right),
\end{aligned}$$

where

$$T_0 = \max_{i=1}^{t} \max(1, (T_{i11} \cdots T_{i,m_i,n_i})^{\frac{1}{m_i n_i}})$$

and the constants c_i will be defined in Lemma 4.13.

DÉMONSTRATION. The first assertion follows immediately from the definition of $\mathcal{R}(\mathbf{T})$. Since the $\mathcal{R}_i$ are independent we fix an i and focus on $\mathcal{R}_i(\mathbf{T})$. Obviously we have that

$$\mathcal{R}_i = \mathbb{Z}[X]/(p_i^{m_i}) \cong (\mathbb{Z}[X]/(p_i))^{m_i}. \tag{4.5}$$

Thus we concentrate on $\mathbb{Z}[X]/(p_i\mathbb{Z}[X])$ which is an order in a number field $\mathcal{K}_i$ of degree n_i over $\mathbb{Q}$. For $\gamma \in \mathcal{K}_i$ let $\gamma^{(\ell)}$ $(1 \leq \ell \leq r_1)$ be the real conjugates and $\gamma^{(m)}$ and $\gamma^{(m+r_2)}$

$(r_1+1 \le m \le r_1+r_2)$ be the pairs of complex conjugates of γ. Note that $r_1+2r_2=n_i$. We will apply the following lemma.

LEMMA 4.13 ([**143**, Lemma 3.3]). *Let T_k $(1 \le k \le r_1+r_2)$ be positive integers and set $T_{r_1+r_2+k}=T_{r_1+k}$ for $1 \le k \le r_2$. Then*

$$\#\left\{a \in \mathbb{Z}[X]/(p_i) : \left|a^{(k)}\right| \le T_k\right\} = c_i T_1 \cdots T_{n_i} + \mathcal{O}\left(T_0^{n_i-1}\right),$$

where $T_0 = \max\left(1, (T_1 \cdots T_{n_i})^{1/n_i}\right)$ and c_i is a constant depending on $\mathbb{Z}[X]/(p_i)$.

Furthermore since (4.5) holds, we get that there exists a $\mathbb{Z}$ linear mapping M_i such that

$$M_i \cdot (T_{i11}, \ldots, T_{i,m_i,n_i}) = (\widetilde{T}_{i11}, \ldots, \widetilde{T}_{i,m_i,n_i})$$

and

$$\#\mathcal{R}_i(\mathbf{T}) = \prod_{j=1}^{m_i} \#\left\{a \in \mathbb{Z}[X]/(p_i) : \left|a^{(k)}\right| \le \widetilde{T}_{ijk} \quad 1 \le k \le n_i\right\}.$$

As the value of T_{ijk} is independent from j, an application of Lemma 4.13 yields

$$\#\mathcal{R}_i(\mathbf{T}) = \widetilde{c}_i \left(\prod_{k=1}^{n_i} T_{i1k}\right)^{m_i} + \mathcal{O}\left(T_{i0}^{m_i n_i - 1}\right),$$

where $\widetilde{c}_i$ depends on c_i and M_i and

$$T_{i0} = \max\left(1, (T_{i11} \cdots T_{i,m_i,n_i})^{\frac{1}{m_i n_i}}\right).$$

For the estimate involving $\mathcal{N}_i(\mathbf{T})$ we note that

$$\mathcal{N}_i = \{g \in \mathcal{R}_i : g \equiv 0 \bmod p_i\} \cong (\mathbb{Z}[X]/(p_i))^{m_i-1}$$

and the result follows in the same way as for $\mathcal{R}_i(\mathbf{T})$. □

With help of all these tools we can show Proposition 4.11.

PROOF OF PROPOSITION 4.11. The first step consists in splitting the sum over $\mathcal{R}(\mathbf{T})$ up into those over $\mathcal{R}_i(\mathbf{T})$. Therefore let $\pi_i : \mathcal{R} \to \mathcal{R}_i$ be the canonical projections. Then $\pi := (\pi_1, \ldots, \pi_t)$ is an isomorphism by the Chinese Remainder Theorem. Furthermore let ϕ_i be the embedding defined by

$$\phi_i : \begin{cases} \overline{\mathcal{K}_i} & \to & \mathbb{R}^{n_i m_i}, \\ a_1 + a_2 X + \cdots + a_{m_i n_i} X^{n_i m_i - 1} & \mapsto & (a_1, \ldots, a_{m_i n_i}) \end{cases}$$

for $i = 1, \ldots, t$. Finally we define the matrix M to be such that

$$M \cdot \phi(z) := (\phi_1 \circ \pi_1(z), \ldots, \phi_t \circ \pi_t(z)).$$

Furthermore we note that for $P_i := \pi_i \circ P$ and $l \in \mathbb{Z}$ we have

$$M \cdot \phi(P(z)X^l) = (\phi_1(P_1(z_1)X^l), \ldots, \phi_t(P(z_t)X^l)).$$

Thus

$$\sum_{z \in \mathcal{R}(\mathbf{T})} e\left(\sum_{r=1}^{h} \left\langle \mathbf{h}_r, \phi(P(z)X^{-l_r-1})\right\rangle\right) = \prod_{i=1}^{t} \sum_{z_i \in \mathcal{R}_i(\mathbf{T})} e\left(\sum_{r=1}^{h} \left\langle \mathbf{h}_{ri}, \phi_i(P_i(z_i)X^{-l_r-1})\right\rangle\right)$$

where $(\mathbf{h}_{r1}, \ldots, \mathbf{h}_{rt}) := \mathbf{h}_r M^{-1}$.

Now we will consider each sum over $\mathcal{R}_i(\mathbf{T})$ separately. Therefore we fix until the end of the proof an $1 \leq i \leq t$ and distinguish two cases according to whether $m_i = 1$ or not.

— **Case 1,** $m_i = 1$: In this case we set $\beta = \beta_{i1}$ and observe that $K = \mathcal{K}_i = \mathbb{Q}(\beta_{i1})$ and $\mathcal{R}_i \cong \mathbb{Z}[\beta]$. Furthermore let $\mathcal{O}_K$ be the maximum order aka the ring of integers of K, then clearly $\mathbb{Z}[\beta_{i1}] \subset \mathcal{O}_K$. We denote by $\beta_{ik} = \beta^{(k)}$ the conjugates of β. Now we will proceed as in the proof of Proposition 6.1 of Madritsch [**143**].

Therefore we need some parameters of the field K and for short we set $n = n_i$ during this case. Then we order the conjugates by denoting with $\beta^{(k)}$ for $1 \leq k \leq r_1$ the real conjugates, whereas $\beta^{(k)}$ and $\beta^{(k+r_2)}$ denote the pairs of complex conjugates, where $n = r_1 + 2r_2$. Let Tr be the trace of an element of K over $\mathbb{Q}$, then we define

$$\tau(z) := \big(\mathrm{Tr}(z), \mathrm{Tr}(\beta z), \dots, \mathrm{Tr}(\beta^{n-1} z)\big) = \Xi \phi_i(z), \tag{4.6}$$

where $\Xi = VV^T$ and V is the Vandermonde matrix

$$V = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \beta & \beta^{(2)} & \cdots & (\beta^{(n)})^{n-1} \\ \vdots & \vdots & & \vdots \\ \beta^{n-1} & (\beta^{(2)})^{n-1} & \cdots & (\beta^{(n)})^{n-1} \end{pmatrix}.$$

We set $(\widetilde{h}_{r1}, \dots, \widetilde{h}_{rn}) := \mathbf{h}_{ri} \Xi^{-1}$ and note that

$$\left\langle \mathbf{h}_{ri}, \phi_i(P_i(z_i) X^{-l_r-1}) \right\rangle = \mathbf{h}_{ri}^T \Xi^{-1} \tau(P_i(z_i)\beta^{-l_r-1}) = \mathrm{Tr}\left(\sum_{k=1}^{n} \widetilde{h}_{rk} \beta^{k-l_r-2} P_i(z_i) \right).$$

Thus we may rewrite the sum under consideration as follows

$$\sum_{z \in \mathcal{R}_i(\mathbf{T})} e\left(\sum_{r=1}^{h} \left\langle \mathbf{h}_{ri}, \phi_i(P_i(z_i) X^{-l_r-1}) \right\rangle \right) = \sum_{z \in \mathcal{R}_i(\mathbf{T})} e\left(\mathrm{Tr}\left(\sum_{r=1}^{h} \sum_{k=1}^{n} \widetilde{h}_{rk} \beta^{k-l_r-2} P_i(z_i) \right) \right).$$

Now we need an approximation lemma which essentially goes back to Siegel [**227**]. Therefore let δ be the different of K over $\mathbb{Q}$ and Δ be the absolute value of the discriminant of K. Then we have the following.

LEMMA 4.14. *Let $N_1, \dots, N_{r_1+r_2}$ be real numbers and let $N = \sqrt[n]{N_1 \cdots N_{r_1} (N_{r_1+1} \cdots N_{r_1+r_2})^2}$ be their geometric mean. Suppose that $N > \Delta^{\frac{1}{n}}$, then, corresponding to any $\xi \in K$, there exist $q \in \mathcal{O}_K$ and $a \in \delta^{-1}$ such that*

$$\begin{aligned} &\left| q^{(k)} \xi^{(k)} - a^{(k)} \right| < N_k^{-1}, \quad 0 < \left| q^{(k)} \right| \leq N_k, \quad 1 \leq k \leq r_1 + r_2, \\ &\max\left(N_k \left| q^{(k)} \xi^{(k)} - a^{(k)} \right|, \left| q^{(k)} \right| \right) \geq \Delta^{-\frac{1}{2}}, \quad 1 \leq k \leq r_1 + r_2, \end{aligned}$$

and

$$\mathrm{N}((q, a\delta)) \leq \Delta^{\frac{1}{2}}.$$

For $1 \leq r \leq h$ we set ξ_r to be the leading coefficient of $\sum_{k=1}^{n} \widetilde{h}_{rk} P_i(z)$. Then we apply Lemma 4.14 with $N_k = T_{i,1,k}^d (\log T)^{-\sigma_2}$ for $1 \leq k \leq r_1 + r_2$ in order to get

that there exist $a \in \delta^{-1}$ and $q \in \mathcal{O}_K$ such that

$$\left|\sum_{r=1}^{h} \frac{\xi_r^{(k)}}{(\beta^{(k)})^{l_r+1}}\, q^{(k)} - a^{(k)}\right| < \frac{(\log T)^{\sigma_2}}{T_{i,1,k}^d} \quad \text{and} \quad 0 < \left|q^{(k)}\right| < \frac{T_{i,1,k}^d}{(\log T)^{\sigma_2}} \quad \text{for} \quad 1 \le k \le n.$$

LEMMA 4.15 ([**143**, Proposition 3.2]). *Suppose that*

$$Q(X) = \alpha_d X^d + \cdots + \alpha_1 X \tag{4.7}$$

is a polynomial of degree d with coefficients in K. If for the leading coefficient α_d there exist $a \in \delta^{-1}$ and $q \in \mathcal{O}_K$ as in Lemma 4.14 with $N_k = T_{i1k}^d (\log T)^{-\sigma_2}$ and

$$(\log T)^{\sigma_2} \le \left|q^{(k)}\right| \le T_{i1k}^d (\log T)^{-\sigma_2} \quad 1 \le k \le r_1 + r_2,$$

then

$$\sum_{x \in \mathcal{R}_i(\mathbf{T})} e(\mathrm{Tr}(Q(x))) \ll T^{n_i} (\log T)^{-\sigma_0}$$

with $\sigma_2 \ge 2^{d-1}\left(\sigma_0 + r 2^{2d}\right)$.

Now we distinguish several cases according to the quality of approximation by Lemma 4.14, which is represented by the size of $H(q)$:

— **Case 1.1,** $\lceil q \rceil \ge (\log T)^{\sigma_2}$: We apply Lemma 4.15 and get

$$\sum_{z_i \in \mathcal{R}_i(\mathbf{T})} e\left(\sum_{r=1}^{h}\sum_{k=1}^{n} \frac{\widetilde{h}_{rk} P_i(z_i))}{\beta^{l_r+1}}\right) \ll T^n (\log T)^{-\sigma_0}.$$

— **Case 1.2,** $2 \le \lceil q \rceil < (\log T)^{\sigma_2}$: In the last two cases we need Minkowski's lattice theory (*cf.* [**108**]). Let λ_1 be the first successive minimum of the $\mathbb{Z}$-lattice δ^{-1}. Then we get

$$\left|\sum_{r=1}^{h} \frac{\xi_r^{(k)}}{(\beta^{(k)})^{l_r+1}}\right| \ge \left|\frac{a^{(k)}}{q^{(k)}}\right| - \frac{1}{\left|q^{(k)}\right|^2} \ge \lambda_1 \left(\frac{1}{\left|q^{(k)}\right|} - \frac{1}{\left|q^{(k)}\right|^2}\right) \ge \lambda_1 \frac{1}{2\left|q^{(k)}\right|} \gg (\log T)^{-\sigma_2},$$

which implies

$$(\log T)^{\sigma_2} \ll \left|\sum_{r=1}^{h} \frac{\xi_r^{(k)}}{(\beta^{(k)})^{l_r+1}}\right| \le \frac{\sum_{r=1}^{h} \left|\xi_r^{(k)}\right|}{\left|\beta^{(k)}\right|^{l_1+1}}.$$

Since $\left|\beta^{(k)}\right| > 1$ and (4.2) we have

$$\left|\beta^{(k)}\right|^{l_1+1} \ll \left|\xi^{(k)}\right| (\log T)^{\sigma_2} \ll n (\log T)^{\sigma_2+\sigma_1},$$

which yields

$$l_1 \ll \frac{(\sigma_2 + \sigma_1)}{\log\left|\beta^{(k)}\right|} \log\log T$$

contradicting the lower bound of l_1 for sufficiently large C_1 in (4.1).

— **Case 1.3,** $0 < \lceil q \rceil < 2$: In this case we will again use Minkowski's lattice theory (*cf.* [**108**]). Let λ_1 be the first successive minimum of the $\mathbb{Z}$-lattice δ^{-1}, then we have to consider two subcases

— **Case 1.3.1,** $\left|\sum_{r=1}^{h} \frac{\xi_r}{\beta^{l_r+1}} q\right| \geq \frac{\lambda_1}{2}$: Let $1 \leq k \leq n$ be such that

$$\frac{\lambda_1}{2} \leq \left|\sum_{r=1}^{h} \frac{\xi_r^{(k)}}{(\beta^{(k)})^{l_r+1}} q^{(k)}\right| \leq \frac{\sum_{r=1}^{h} \left|\xi_r^{(k)}\right|}{\left|\beta^{(k)}\right|^{l_1+1}} \left|q^{(k)}\right|,$$

then

$$l_1 + 1 \ll \log\log T$$

again contradicts the lower bound of l_1 for sufficiently large C_1 in (4.1).

— **Case 1.3.2,** $\left|\sum_{r=1}^{h} \frac{\xi_r}{\beta^{l_r+1}} q\right| < \frac{\lambda_1}{2}$: By Minkowski's theorem (*cf.* [**108**]) we get that $a = 0$. Thus for $1 \leq k \leq n$

$$\left|\sum_{r=1}^{h} \frac{\xi_r^{(k)}}{(\beta^{(k)})^{l_r+1}} q^{(k)}\right| = \left|\frac{1}{(\beta^{(k)})^{l_r+1}} \sum_{r=1}^{h} \xi_r^{(k)} (\beta^{(k)})^{l_h - l_r} q^{(k)}\right| \leq \frac{(\log T)^{\sigma_2}}{T_{i1k}^d}$$

which implies (taking the norm of the left side)

$$l_h + 1 \geq nd \log_{|\mathrm{N}(\beta)|} T - c(\log\log_{|\mathrm{N}(\beta)|} T)$$

contradicting the upper bound for sufficiently large C_2.

— **Case 2,** $m_i > 1$: Now we have to go one step further and to take a closer look at $\mathcal{R}_i$. In particular we divide every element $z_i \in \mathcal{R}_i$ into its radical and its nilpotent part. We fix an element $z \in \mathcal{R}$ and set $z_i := \pi_i(z)$.

On the one hand, since $\mathcal{R}_i = \widetilde{\mathcal{R}_i} \oplus \mathcal{N}_i$ we have for $z_i \in \mathcal{R}_i$ the unique representation

$$z_i = z_{i1} + z_{i2} \tag{4.8}$$

with $z_{i1} \in \widetilde{\mathcal{R}_i}$ and $z_{i2} \in \mathcal{N}_i$. This motivates the definition of the linear map π_{ij} such that $\pi_{ij}(z) := z_{ij}$ for $i = 1, \ldots, t$ and $j = 1, 2$.

On the other hand, since $\mathcal{R}_i \cong (\mathbb{Z}[X]/(p_i))^{m_i}$ we have for every $z_i \in \mathcal{R}_i$ the unique representation

$$z_i = \sum_{j=1}^{m_i} a_{ij} p_i^{j-1} = \sum_{j=1}^{m_i} \sum_{k=1}^{n_i} a_{ijk} X^{k-1} p_i^{j-1}$$

with $a_{ij} \in \mathbb{Z}[X]$ and $a_{ijk} \in \mathbb{Z}$, respectively. We clearly have

$$\pi_{i1}(z) = \sum_{k=1}^{n_i} a_{i1k} X^{k-1} \quad \text{and} \quad \pi_{i2}(z) = \sum_{j=2}^{m_i} \sum_{k=1}^{n_i} a_{ijk} X^{k-1} p_i^{j-1}.$$

Thus we define for $z_i \in \mathcal{R}_i$ the embeddings ψ_{i1} and ψ_{i2} by

$$\psi_{i1}(\pi_{i1}(z_i)) = (a_{i11}, \ldots, a_{i1n_i}) \quad \text{and} \quad \psi_{i2}(\pi_{i2}(z_i)) = (a_{i21}, \ldots, a_{i,m_i,n_i}).$$

Then there exists an invertible matrix $\widetilde{M_i}$ such that

$$\widetilde{M_i}(\phi_i \circ \pi_i(z)) = (\psi_{i1} \circ \pi_{i1}(z), \psi_{i2} \circ \pi_{i2}(z)).$$

Now we can divide the sum up as follows.

$$\begin{aligned}
&\sum_{z_i\in\mathcal{R}_i(\mathbf{T})} e\left(\sum_{r=1}^{h}\left\langle \mathbf{h}_{ri}, \phi_i(P_i(z_i)X^{-l_r-1})\right\rangle\right)\\
&= \sum_{z_{i1}\in\widetilde{\mathcal{R}_i}(\mathbf{T})}\sum_{z_{i2}\in\mathcal{N}_i(\mathbf{T})} e\left(\sum_{r=1}^{h}\left\langle \mathbf{h}_{ri}, \phi_i(P_i(z_{i1}+z_{i2})X^{-l_r-1})\right\rangle\right)\\
&= \sum_{z_{i1}\in\widetilde{\mathcal{R}_i}(\mathbf{T})} e\left(\sum_{r=1}^{h}\left\langle \mathbf{h}_{ri1}, \psi_{i1}(P_{i1}(z_{i1})X^{-l_r-1})\right\rangle\right)\sum_{z_{i2}\in\mathcal{N}_i(\mathbf{T})} e\left(\sum_{r=1}^{h}\left\langle \mathbf{h}_{ri2}, \psi_{i2}(P_{i2}(z_{i2})X^{-l_r-1})\right\rangle\right),
\end{aligned}$$

where we have set $P_{ij} = \pi_{ij}\circ P$ for $j = 1, 2$.

Since for the first sum we have that $m_i = 1$ we may follow **Case 1** above and use Lemma 4.13 for trivially estimating the second one to prove the proposition for this case.

□

5. Treatment of the border

In Section 3 above we have constructed the Urysohn function we need in order to properly count the number of elements within the fundamental domain. In this construction we also used an axe-parallel tube in order to cover the border of the fundamental domain. The number of hits of this tube gives rise to the error term which we will consider in this section.

We fix a positive integer v, which will be chosen later, and a real vector $\mathbf{T}$. Furthermore for $l \geq 0$ we define F_l to be the number of hits of the border of the Urysohn function which is

$$F_l := \#\left\{ z\in\mathcal{R}(\mathbf{T}) \middle| B^{-l-1}\phi\left(P(z)\right) \in \bigcup_{a\in\mathcal{N}} P_{v,a} \mod B^{-1}\mathbb{Z}^n \right\}. \tag{5.1}$$

As indicated above we are interested in an estimation of F_l.

Proposition 4.16. *Let $\mu < |\det B|$ be as in Section 3 and C_1 and C_2 be sufficiently large positive reals. Suppose that l is a positive integer such that*

$$C_1 \log\log T \leq l \leq d\log_{|\det B|} T - C_2 \log\log T. \tag{5.2}$$

Then for any positive σ_3 we have

$$F_l \ll \mu^v T^n \left(|\det B|^{-v} + (\log T)^{-t\sigma_3}\right).$$

In order to estimate F_l we need the Erdős-Turán-Koksma Inequality.

Lemma 4.17 ([**66**, Theorem 1.21]). *Let $x_1, \ldots, x_S$ be points in the n-dimensional real vector space $\mathbb{R}^n$ and H an arbitrary positive integer. Then the discrepancy $D_S(x_1, \ldots, x_S)$ fulfills the inequality*

$$D_S(x_1,\ldots,x_S) \ll \frac{2}{H+1} + \sum_{0<\|\mathbf{h}\|_\infty\leq H} \frac{1}{r(\mathbf{h})}\left|\frac{1}{S}\sum_{s=1}^{S} e(\langle \mathbf{h}, x_s\rangle)\right|,$$

where $\mathbf{h}\in\mathbb{Z}^n$ and $r(\mathbf{h}) = \prod_{i=1}^{n}\max(1, |h_i|)$.

Proof of Proposition 4.16. We want to proceed in three steps. First we subdivide the tube $P_{v,a}$ into rectangles in order to apply the Erdős-Turán-Koksma Inequality in the second step. Finally we put them together to gain the desired result.

Recall that the tube $P_{v,a}$ defined in Lemma 4.9 consists of a family of rectangles. Let R_a be one of them, then we want to estimate

$$F_l(R_a) := \#\left\{ z \in \mathcal{R}(\mathbf{T}) \middle| B^{-l-1}\phi\left(P(z)\right) \in \bigcup_{a\in\mathcal{N}} R_a \mod B^{-1}\mathbb{Z}^n \right\}.$$

Using the definition of the discrepancy we get that

$$F_l(R_a) \ll T^n \left(\lambda(R_a) + D_S\left(\left\{ B^{-l-1}\phi\left(P(z)\right) \right\}_{z\in\mathcal{R}(\mathbf{T})} \right) \right), \tag{5.3}$$

where λ is the n-dimensional Lebesgue measure and S is the number of elements in $\mathcal{R}(\mathbf{T})$. By Lemma 4.12 we have that

$$S = \prod_{i=1}^{t} c_i \left(\prod_{k=1}^{n_i} T_{i1k} \right)^{m_i} + \mathcal{O}\left(T_0^{n-1}\right). \tag{5.4}$$

Now we apply Lemma 4.17 to get

(5.5)

$$D_S\left(\left\{ B^{-l-1}\phi\left(P(z)\right) \right\}_{z\in\mathcal{R}(\mathbf{T})} \right) \ll \frac{2}{H+1} + \sum_{0<\|\mathbf{h}\|_\infty \le H} \frac{1}{r(\mathbf{h})} \left| \frac{1}{S} \sum_{z\in\mathcal{R}(\mathbf{T})} e\left(\langle \mathbf{h}, B^{-l-1}\phi\left(P(z)\right)\rangle \right) \right|.$$

The next step consists in an application of Proposition 4.11 which yields

$$\left| \sum_{z\in\mathcal{R}(\mathbf{T})} e\left(\langle \mathbf{h}, B^{-l-1}\phi\left(P(z)\right)\rangle \right) \right| \ll T^n (\log T)^{-t\sigma_0} \tag{5.6}$$

Putting (5.4), (5.5), and (5.6) togetter in (5.3) gives

$$\begin{aligned} F_l(R_a) &\ll T^n\lambda(R_a) + \frac{T^n}{(\log T)^{\sigma_1}} + T^n(\log T)^{-t\sigma_0} \sum_{0<\|\mathbf{h}\|_\infty \le H} \frac{1}{r(\mathbf{h})} \\ &\ll T^n\lambda(R_a) + \frac{T^n}{(\log T)^{\sigma_1}} + T^n(\log T)^{-t\sigma_0}(\log\log T)^n. \end{aligned}$$

Setting $\sigma_1 := t\sigma_0/2$ and summing over all rectangles R_a yields

$$F_l \ll \mu^v T^n \left(|\mathrm{N}(b)|^{-v} + (\log T)^{-t\sigma_0/2} \right).$$

Finally we set $\sigma_3 = t\sigma_0/2$ which proves the proposition. □

6. The main proposition

The main idea is to understand the additive function as putting weights on the digits. Thus if we can show that the digits are uniformly distributed the same is true for the values of the additive functions. Therefore we look at patterns in the expansion of $P(z)$. In particular, we count the number of occurrences of certain digits at certain positions in the expansions.

PROPOSITION 4.18. *Let f be an additive function. Let $T \geq 0$ and T_{ijk} as in (2.5). Let L be the maximum length of the b-adic expansion of $z \in \mathcal{R}(\mathbf{T})$ and let C_1 and C_2 be sufficiently large. Then for*

$$C_1 \log L \leq l_1 < l_2 < \cdots < l_h \leq dL - C_2 \log L \tag{6.1}$$

we have,

$$\begin{aligned}\Theta &:= \#\left\{z \in \mathcal{R}(\mathbf{T}) | a_{l_r}(f(z)) = b_r, r = 1, \ldots, h\right\} \\ &= \frac{c_1 \cdots c_t}{|\det B|^h} T^n + \mathcal{O}\left(T^n (\log T)^{-t\sigma_0}\right).\end{aligned}$$

uniformly for $T \to \infty$, where $(l_r, b_r) \in \mathfrak{N} \times \mathcal{N}$ are given pairs of position and digit and σ_0 is an arbitrary positive constant.

DÉMONSTRATION. We recall our Urysohn function u_a (defined in (3.5)) and set for $\nu \in \mathbb{R}^n$

$$t(\nu) = u_{b_1}(B^{-l_1-1}\nu) \cdots u_{b_h}(B^{-l_h-1}\nu),$$

where B is the matrix defined in (3.3).

Now we want to apply the Fourier transformation, which we developed in Lemma 4.10. Therefore we set

$$\mathcal{M} := \left\{M = (h_1, \ldots, h_h) | h_r \in \mathbb{Z}^n, \text{ for } r = 1, \ldots, h\right\}.$$

An application of Lemma 4.10 yields

$$t(\nu) = \sum_{M \in \mathcal{M}} T_M e\left(\sum_{r=1}^{h} \mathbf{h}_r B^{-l_r-1} \nu\right), \tag{6.2}$$

where $T_M = \prod_{r=1}^{h} c_{m_{r1},\ldots,m_{rn}}$. Combining this with the definition of F_l in (5.1) we get

$$\left|\Theta - \sum_{z \in \mathcal{R}(\mathbf{T})} t(\phi(P(z)))\right| \leq F_{l_1} + \cdots + F_{l_h}. \tag{6.3}$$

Plugging (6.2) into (6.3) together with an application of Lemma 4.10 for the coefficients yields

$$\Theta = \frac{c_1 \cdots c_t}{|\det(B)|^h} T^n + \sum_{0 \neq M \in \mathcal{M}} T_M e\left(\sum_{r=1}^{h} \langle \mathbf{h}_r, B^{-l_r-1} \phi(P(z))\rangle\right) + \mathcal{O}\left(\sum_{r=1}^{h} F_{l_r}\right).$$

Now an application of Proposition 4.11 to treat the exponential sums, of Proposition 4.16 for the border F_l with $v \ll \log\log T$ and the observation that

$$\sum_{M \in \mathcal{M}} |T_M| \ll \kappa^{-2h} \ll |\det B|^{2hv} \ll (\log T)^{t\sigma_0/2},$$

where we used the definition of κ in (3.6), proves the proposition. □

7. Proof of Theorem 4.8

For this proof we mainly follow the proof of the Theorem of Bassily and Kátai [**19**]. In the same manner we cut of the head and tail of the expansion and show the theorem for a truncated version of the additive function. In particular we set $C := \max(C_1, C_2)$, $A := [C \log L]$ and $B := L - A$, where L, C_1 and C_2 are defined in the statement of Proposition 4.18. Furthermore we define the truncated function f' to be

$$f'(P(z)) = \sum_{j=A}^{B} f(a_j(P(z))b^j).$$

By the definition of A and $f(ab^j) \ll 1$ with $a \in \mathcal{N}$ we get that $f'(P(z)) = f(P(z)) + \mathcal{O}(\log L)$. In the same manner we define the truncated mean and standard deviation

$$M'(T) := \sum_{j=A}^{B} m_j \quad \text{and} \quad D'^2(T) := \sum_{j=A}^{B} \sigma_j^2.$$

At this point we need that the deviation D tends sufficiently fast do infinity. In particular, we could refine the statement, if we shrink the part, which is cut of. Since $M(T) - M'(T) = \mathcal{O}(\log L)$ and $D^2(T) - D'^2(T) = \mathcal{O}(\log L)$ we get that it suffices to show that

$$\frac{1}{\#\mathcal{R}(\mathbf{T})} \# \left\{ z \in \mathcal{R}(\mathbf{T}) \middle| \frac{f'(P(z)) - M'(T^d)}{D'(T^d)} < y \right\} \longrightarrow \Phi(y).$$

By the Fréchet-Shohat Theorem (*cf.* [**74**, Lemma 1.43]) this holds true if and only if the moments

$$\xi_k(T) := \frac{1}{\#\mathcal{R}(\mathbf{T})} \sum_{z \in \mathcal{R}(\mathbf{T})} \left(\frac{f'(P(z)) - M'(T^d)}{D'(T^d)} \right)^k$$

converge to the moments of the normal law for $T \to \infty$. We will show the last statement by comparing the moments ξ_k with

$$\eta_k(T) := \frac{1}{\#\mathcal{R}(\mathbf{T}^d)} \sum_{z \in N(T^d)} \left(\frac{f'(z) - M'(T^d)}{D'(T^d)} \right)^k,$$

where $\mathbf{T}^d = (T_1^d, \dots, T_n^d) = (T_{111}^d, \dots, T_{t,n_t,m_t}^d)$.

An application of Proposition 4.18 gives that

$$\xi_k(T) - \eta_k(T) \to 0 \quad \text{for} \quad T \to \infty.$$

Furthermore we get by Proposition 4.7 that these sums consist of independently identically distributed random variables (with possible $2C$ exceptions). By the central limit theorem we get that their distribution converges to the normal law. Thus the $\eta_k(T)$ converge to the moments of the normal law. This yields

$$\lim_{T \to \infty} \xi_k(T) = \lim_{T \to \infty} \eta_k(T) = \int x^k \mathrm{d}\Phi.$$

We apply the Fréchet-Shohat Theorem again to prove the theorem.

Acknowledgment

This paper was written while M. Madritsch was a visitor at the Faculty of Informatics of the University of Debrecen. He thanks the centre for its hospitality. During his stay he was supported by the project HU 04/2010 founded by the ÖAD. The second author was supported by the Hungarian National Foundation for Scientific Research Grant No.T67580 and by the TÁMOP 4.2.1/B-09/1/KONV-2010-0007 project. The second project is implemented through the New Hungary Development Plan co-financed by the European Social Fund, and the European Regional Development Fund.

Chapitre 5

On a second conjecture of Stolarsky : the sum of digits of polynomial values

This chapter is joint work with Thomas Stoll and appeared in the *Archiv der Mathematik*, **102** (2014), no. 1, 49 – 57.

1. Introduction and statement of results

Let $q \geq 2$ be an integer. Then every positive integer n has a unique q-adic representation of the form

$$n = \sum_{k=0}^{\ell} n_k q^k \quad \text{with} \quad n_\ell \neq 0.$$

We call a function f a q-additive function if it acts only on the digits of this expansion, *i.e.*,

$$f\left(\sum_{k=0}^{\ell} n_k q^k\right) = \sum_{k=0}^{\ell} f(n_k q^k).$$

Moreover, if this action is independent of the position of the digit, *i.e.*, $f(aq^k) = f(aq^j)$ for $k, j \geq 0$ and $a \in \{0, 1, \ldots, q-1\}$, then we call f strictly q-additive. The most famous example of a strictly q-additive function is the sum-of-digits function s_q defined by

$$s_q\left(\sum_{k=0}^{\ell} n_k q^k\right) = \sum_{k=0}^{\ell} n_k.$$

In 1978, K. B. Stolarsky [**236**] studied the distribution properties of the sequence of fractions

$$(s_2(n^r)/s_2(n))_{n\geq 1}\,,$$

where $r \geq 2$ denotes a fixed integer. At the end of his paper, he posed two conjectures. His first conjecture was to give a proof that for all fixed $r \geq 2$ one has

$$\liminf_{n\to\infty} \frac{s_2(n^r)}{s_2(n)} = 0.$$

Hare, Laishram and Stoll [**104**] recently settled this conjecture and proved, more generally, that for any polynomial $P(X) \in \mathbb{Z}[X]$ with $P(\mathbb{N}) \subset \mathbb{N}$ of degree $r \geq 2$,

$$\liminf_{n\to\infty} \frac{s_q(P(n))}{s_q(n)} = 0.$$

Stolarsky also showed that the sequence $s_2(n^r)/s_2(n)$ is unbounded as $n \to \infty$ (this is also true for $\frac{s_q(P(n))}{s_q(n)}$, see [**104**]). In very recent work, the authors [**140**] show that the ratio

$$s_q(P_1(n))/s_q(P_2(n))$$

for polynomials $P_1(X)$ and $P_2(X)$ of distinct degrees lies indeed dense in $\mathbb{R}^+$.

In [**236**] Stolarsky posed a second question, namely, whether

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\le N}\frac{s_2(n^r)}{s_2(n)} \tag{1.1}$$

exists and — if it exists — to determine its value. He conjectured that the limit (1.1) exists and that it is included in the interval $]1,h]$. The purpose of this paper is to prove this conjecture. More precisely, as in Hare, Laishram and Stoll [**104**], we show a general version. We also present a variant to polynomial values of prime numbers. Let p_n denote the n–th prime, *i.e.*, $p_1=2$, $p_2=3$, $p_3=5$ etc.

Our main result is the following :

THEOREM 5.1. *Let $q_1,q_2\ge 2$ be integers and $P_1(X),P_2(X)\in\mathbb{C}[X]$ be polynomials of degrees $r_1,r_2\ge 1$, respectively, with $P_1(\mathbb{N}),P_2(\mathbb{N})\subset\mathbb{N}$. Then*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\le N}\frac{s_{q_1}(P_1(n))}{s_{q_2}(P_2(n))}=\frac{q_1-1}{q_2-1}\cdot\left(\frac{\log q_1}{\log q_2}\right)^{-1}\cdot\frac{r_1}{r_2}. \tag{1.2}$$

Moreover,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\le N}\frac{s_{q_1}(P_1(p_n))}{s_{q_2}(P_2(p_n))}=\frac{q_1-1}{q_2-1}\cdot\left(\frac{\log q_1}{\log q_2}\right)^{-1}\cdot\frac{r_1}{r_2}. \tag{1.3}$$

REMARK 9. In exactly the same way one can show that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\le N}\frac{s_{q_1}(P_1(p_n))}{s_{q_2}(P_2(n))}=\frac{q_1-1}{q_2-1}\cdot\left(\frac{\log q_1}{\log q_2}\right)^{-1}\cdot\frac{r_1}{r_2}.$$

All these results easily extend to strictly q_1- resp. q_2-additive functions provided that the variance and the image set of the functions satisfy some suitable conditions (see [**19**]). These conditions are automatically verified by the sum-of-digits function.

Furthermore we remark that by replacing the theorem of Bassily and Kátai (Theorem 5.3) by analogous results due to Drmota and Steiner [**68**], Madritsch [**143**], Madritsch and Pethő [**151**], Drmota and Gutenbrunner [**65**], and Madritsch and Thuswaldner [**145**], respectively, one can prove analogous results for additive functions of polynomials in numeration systems that are defined via linear recurrent sequences (such as the Zeckendorf expansion), for additive functions of polynomials in numeration systems in algebraic number fields, for additive functions of polynomials in numeration systems in the quotient ring of polynomials over $\mathbb{Z}$, for additive functions of polynomials in numeration systems in the ring of polynomials over a finite field and for additive functions of polynomials in numeration systems in function fields, respectively.

2. Preliminaries

For the proof, we need some notation. We denote by

$$\mu_q=\frac{q-1}{2}\quad\text{and}\quad\sigma_q^2=\frac{q^2-1}{12},$$

the mean and the variance of the values of the sum-of-digits function (see [**19**] or [**67**]). We will use the letter p to refer to a prime number, and use $\pi(N)$ for the number of primes up to N. We write $f\ll_\omega g$ or $f=O_\omega(g)$ if there exists a constant C depending at most on ω such

that $f(x) \leq Cg(x)$ for sufficiently large x. If there is no such ω then the implied constant is meant to be absolute. We write $\log_q x$ for the logarithm to base q. Finally, for $A \subset \mathbb{N}$ we denote by $d(A)$ the asymptotic density of A, *i.e.*,

$$d(A) = \lim_{N\to\infty} \frac{A \cap [1, N]}{N}.$$

The idea of the proof of Theorem 5.1 is to use Cesàro means (see [**102**]). However, we cannot apply these means directly since the summands in (1.2) and (1.3) could be arbitrarily large. We will therefore divide the sequence into two parts. The first part corresponds to terms where the ratio stays close to the mean value whereas the second part is made up by terms that are far away from the mean (this will be made precise in a moment). For the first part, we use Cesàro means and the following lemma, which helps us replacing the unbounded sequence by a bounded one.

LEMMA 5.2. *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of reals and $A \subset \mathbb{N}$ a set with asymptotic density one. If*

$$\lim_{\substack{n\to\infty \\ n\in A}} x_n = x < \infty,$$

then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{\substack{n\leq N \\ n\in A}} x_n = x.$$

DÉMONSTRATION. We define the sequence $(y_n)_{n\in\mathbb{N}}$ by

$$y_n = \begin{cases} x_n & \text{if } n \in A, \\ x & \text{if } n \notin A. \end{cases}$$

Since A has asymptotic density one, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{\substack{n\leq N \\ n\in A}} x_n = \lim_{N\to\infty} \frac{1}{N} \sum_{n\leq N} y_n - \lim_{N\to\infty} \frac{1}{N} \sum_{\substack{n\leq N \\ n\notin A}} x = x.$$

□

For the second part, we make use of a minor refinement of a result of Bassily and Kátai [**19**]. The difference to the original result is that we only suppose $P(X) \in \mathbb{C}[X]$ instead of $P(X) \in \mathbb{Z}[X]$.

THEOREM 5.3. *Let $q \geq 2$ be an integer and $P(X) \in \mathbb{C}[X]$ be a polynomial of degree $r \geq 1$ with $P(\mathbb{N}) \subset \mathbb{N}$. Then*

$$\frac{1}{N} \# \left\{ 1 \leq n \leq N \colon \frac{s_q(P(n)) - \mu_q \log_q(N^r)}{\sigma_q (\log_q N^r)^{\frac{1}{2}}} < t \right\} \xrightarrow[N\to\infty]{} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x$$

and

$$\frac{1}{\pi(N)} \# \left\{ 1 \leq p \leq N \colon \frac{s_q(P(p)) - \mu_q \log_q(N^r)}{\sigma_q (\log_q N^r)^{\frac{1}{2}}} < t \right\} \xrightarrow[N\to\infty]{} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x.$$

DÉMONSTRATION. To prove this refinement we proceed in two steps, starting with an application of the following

LEMMA 5.4 ([**182**, Satz 6.9]). *Let $P(X) \in \mathbb{C}[X]$ be a polynomial of degree r. If $P(n) \in \mathbb{Z}$ for all $n \in \mathbb{Z}$, then there exists $c_0, \dots, c_r \in \mathbb{N}$ such that*

$$P(x) = \sum_{i=0}^{r} c_i \binom{x}{i}.$$

Thus, in order to have $P(\mathbb{N}) \subset \mathbb{N}$, we get that $P(X) \in \mathbb{Q}[X]$.

Secondly, we take a closer look at the proof the Lemma 5 in [**19**]. We use the same notation throughout the rest of the proof. Let $c \in \mathbb{N}$ be the smallest positive integer such that $cP(X) = \widetilde{P}(X) \in \mathbb{Z}[X]$. By replacing P everywhere with $\widetilde{P}/c$ we reach at Equation 4.10 of [**19**], which then reads

$$\sum_{n\le x} t\left(\frac{\widetilde{P}(n)}{c}\right) = \sum_{M} T_M \sum_{n\le x} e\left(\frac{A_M}{H_M}\frac{\widetilde{P}(n)}{c}\right) = \sum_{M} T_M \sum_{n\le x} e\left(\frac{A'_M}{H_M c'}\widetilde{P}(n)\right)$$

and

$$\sum_{p\le x} t\left(\frac{\widetilde{P}(p)}{c}\right) = \sum_{M} T_M \sum_{p\le x} e\left(\frac{A_M}{H_M}\frac{\widetilde{P}(p)}{c}\right) = \sum_{M} T_M \sum_{p\le x} e\left(\frac{A'_M}{H_M c'}\widetilde{P}(p)\right)$$

with

$$A'_M = \frac{A_M}{(A_M, c)}, \quad c' = \frac{c}{(A_M, c)} \quad \text{and} \quad (A'_M, H_M c') = 1.$$

Let $q = p_1^{e_1} \cdots p_s^{e_s}$ be the prime decomposition of q. If $q \nmid m_h$, then $p_t^{e_t} \nmid m_h$ for some $1 \le t \le s$. Thus

$$H_M c' \left(m_h + q^{\ell_h - \ell_{h-1}} m_{h-1} + \cdots + m_1 q^{\ell_h - \ell_1}\right) = A'_M q^{\ell_h + 1}.$$

Since $p_t^{e_t} \nmid m_h$ and $(A'_M, H_M c') = 1$ we get that $p_t^{\ell_t e_t} \mid H_M c'$. Thus $H_M c' \ge q^{\eta \ell_h}$ with $\eta = e_t \log_q p_t$. Similarly we get that $H_M \ge q^{\eta \ell_s}$ holds if $q \nmid m_s$ and $m_{s+1} = \cdots = m_h = 0$. Therefore we may apply Lemma 1 and 2 of [**19**] and the Theorem follows in the same way. □

3. Proof of Theorem 5.1

In the sequel, assume that $N \ge 2$ is a fixed real number. For a given integer $q \ge 2$ and a given polynomial $P(X) \in \mathbb{Z}[X]$ (with $P(\mathbb{N}) \subset \mathbb{N}$) of degree $r \ge 1$ we define $A_{q,P} = A_{q,P}(N)$ to be the set of integers n with $1 \le n \le N$ such that $s_q(P(n))$ is close to its mean value, *i.e.*,

$$A_{q,P} := \left\{1 \le n \le N\colon \ \left|s_q(P(n)) - \mu_q r \log_q N\right| \le \sigma_q (r \log_q N)^{\frac{3}{4}}\right\}.$$

(We remark that in fact any exponent larger than $\frac{1}{2}$ in place of $\frac{3}{4}$ would have done the job.) In a similar way, we define $B_{q,P} = B_{q,P}(N)$ to be the set of integers n with $1 \le n \le N$ such that $s_q(P(p_n))$ is close to its mean value (note that by the prime number theorem we have $p_N \sim N \log N$, as $N \to \infty$), *i.e.*,

$$B_{q,P} := \left\{1 \le n \le N\colon \ \left|s_q(P(p_n)) - \mu_q r \log_q (N \log N)\right| \le \sigma_q (r \log_q (N \log N))^{\frac{3}{4}}\right\}.$$

In order to be able to apply the properties of the Cesàro mean we need that both the numerator and the denominator of the ratios in (1.2) and (1.3) are near the mean. We first show that for $N \to \infty$ we have $\#B_{q,P} \sim N$ and $\#A_{q,P} \sim N$. We then use asymptotic densities to show that there are only few elements in $[1, N]\backslash(A_{q_1,P_1} \cap A_{q_2,P_2})$ resp. $[1, N]\backslash(B_{q_1,P_1} \cap B_{q_2,P_2})$. We will then be able to restrict our attention to $A_{q_1,P_1} \cap A_{q_2,P_2}$ resp. $B_{q_1,P_1} \cap B_{q_2,P_2}$ in the end.

We start with an application of Theorem 5.3. As $N \to \infty$, we get that

$$\#\left([1,N] \setminus A_{q,P}\right) = \#\left\{1 \le n \le N\colon \left|\frac{s_q(P(n)) - \mu_q r \log_q N}{\sigma_q (r\log_q N)^{\frac{1}{2}}}\right| > (r\log_q N)^{\frac{1}{4}}\right\} \ll N \int_{(r\log_q N)^{\frac{1}{4}}}^{\infty} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x.$$

Thus the number of elements that lie not in $A_{q,P}$ can be estimated by the tail of the normal distribution. We have

$$\int_t^{\infty} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x \le \int_t^{\infty} \frac{x}{t} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x = \frac{\exp\left(-\frac{t^2}{2}\right)}{t},$$

where we have used that $0 < t \le x$. Therefore,

$$\begin{aligned} \#\left([1,N] \setminus A_{q,P}\right) &\ll N \exp\left(-\frac{(r\log_q N)^{\frac{1}{2}}}{2}\right) (r\log_q N)^{-\frac{1}{4}} \\ &\ll \frac{N}{(r\log_q N)^{\frac{5}{4}}} \ll_{q,P} \frac{N}{(\log N)^{\frac{5}{4}}}. \end{aligned} \tag{3.1}$$

The same calculation also shows that

$$\begin{aligned} &\#\left([1,N] \setminus B_{q,P}\right) \\ &\quad = \#\left\{1 \le n \le N\colon \left|\frac{s_q(P(p_n)) - \mu_q r \log_q(N\log N)}{\sigma_q (r\log_q(N\log N))^{\frac{1}{2}}}\right| > (r\log_q(N\log N))^{\frac{1}{4}}\right\} \\ &\quad \ll_{q,P} \frac{N}{(\log N)^{\frac{5}{4}}}. \end{aligned} \tag{3.2}$$

Recall the setting of Theorem 5.1. The prime number theorem (in the form $p_N \sim N\log N$) and a comparison of the lengths of the expansions give

$$\max\left(\frac{s_{q_1}(P_1(n))}{s_{q_2}(P_2(n))}, \frac{s_{q_1}(P_1(p_n))}{s_{q_2}(P_2(p_n))}\right) \ll_{q_1,P_1} \log N, \qquad N \to \infty,$$

uniformly for all n with $1 \le n \le N$. Hence, we get from (3.1) that

$$\begin{aligned} \frac{1}{N} \sum_{\substack{n\le N \\ n\notin A_{q_1,P_1}\cap A_{q_2,P_2}}} \frac{s_{q_1}(P_1(n))}{s_{q_2}(P_2(n))} &\ll_{q_1,P_1} \frac{\log N}{N} \left(\sum_{\substack{n\le N \\ n\notin A_{q_1,P_1}}} 1 + \sum_{\substack{n\le N \\ n\notin A_{q_2,P_2}}} 1\right) \\ &\ll_{q_1,q_2,P_1} (\log N)(\log N)^{-\frac{5}{4}} = o(1), \end{aligned} \tag{3.3}$$

and similarly from (3.2) that

$$\frac{1}{N} \sum_{\substack{n\le N \\ n\notin B_{q_1,P_1}\cap B_{q_2,P_2}}} \frac{s_{q_1}(P_1(p_n))}{s_{q_2}(P_2(p_n))} = o(1). \tag{3.4}$$

Now we turn to the elements which are in $A_{q_1,P_1} \cap A_{q_2,P_2}$. By the definitions of these sets we get for all $n \in A_{q_1,P_1} \cap A_{q_2,P_2}$ (note that the denominator is positive),

$$\begin{aligned}\frac{s_{q_1}(P_1(n))}{s_{q_2}(P_2(n))} &\leq \frac{\mu_{q_1} r_1 \log_{q_1} N + \sigma_{q_1}(r_1 \log_{q_1} N)^{\frac{3}{4}}}{\mu_{q_2} r_2 \log_{q_2} N - \sigma_{q_2}(r_2 \log_{q_2} N)^{\frac{3}{4}}} \\ &= \frac{\frac{\mu_{q_1} r_1}{\log q_1} + \sigma_{q_1} r_1^{\frac{3}{4}} (\log_{q_1} N)^{-\frac{1}{4}} (\log q_1)^{-\frac{3}{4}}}{\frac{\mu_{q_2} r_2}{\log q_2} - \sigma_{q_2} r_2^{\frac{3}{4}} (\log_{q_2} N)^{-\frac{1}{4}} (\log q_2)^{-\frac{3}{4}}},\end{aligned}$$

and

$$\frac{s_{q_1}(P_1(n))}{s_{q_2}(P_2(n))} \geq \frac{\frac{\mu_{q_1} r_1}{\log q_1} - \sigma_{q_1} r_1^{\frac{3}{4}} (\log_{q_1} N)^{-\frac{1}{4}} (\log q_1)^{-\frac{3}{4}}}{\frac{\mu_{q_2} r_2}{\log q_2} + \sigma_{q_2} r_2^{\frac{3}{4}} (\log_{q_2} N)^{-\frac{1}{4}} (\log q_2)^{-\frac{3}{4}}}.$$

This can be rephrased as follows. Let $(N_k)_{k\geq 0}$ be any sequence of reals with $\lim_{k\to\infty} N_k = \infty$ and let $(n_k)_{k\geq 0}$ be any sequence of integers with $n_k \in A_{q_1,P_1}(N_k) \cap A_{q_2,P_2}(N_k)$. Then

$$\lim_{k\to\infty} \frac{s_{q_1}(P_1(n_k))}{s_{q_2}(P_2(n_k))} = \frac{\mu_{q_1} \log q_2}{\mu_{q_2} \log q_1} \cdot \frac{r_1}{r_2}.$$

A similar calculation shows that

$$\begin{aligned}\frac{s_{q_1}(P_1(p_n))}{s_{q_2}(P_2(p_n))} &\leq \frac{\mu_{q_1} r_1 \log_{q_1}(N \log N) + \sigma_{q_1} \left(r_1 \log_{q_1}(N \log N)\right)^{\frac{3}{4}}}{\mu_{q_2} r_2 \log_{q_2}(N \log N) - \sigma_{q_2} \left(r_2 \log_{q_2}(N \log N)\right)^{\frac{3}{4}}} \\ &= \frac{\frac{\mu_{q_1} r_1}{\log q_1}\left(1 + \frac{\log\log N}{\log N}\right) + \sigma_{q_1} r_1^{\frac{3}{4}} (\log q_1)^{-\frac{3}{4}} (\log N)^{-\frac{1}{4}} \left(1 + \frac{\log\log N}{\log N}\right)^{\frac{3}{4}}}{\frac{\mu_{q_2} r_2}{\log q_2}\left(1 + \frac{\log\log N}{\log N}\right) - \sigma_{q_2} r_2^{\frac{3}{4}} (\log q_2)^{-\frac{3}{4}} (\log N)^{-\frac{1}{4}} \left(1 + \frac{\log\log N}{\log N}\right)^{\frac{3}{4}}}.\end{aligned}$$

In a similar way we get the lower bound with the signs reversed. Again, we obtain as limit

$$\lim_{k\to\infty} \frac{s_{q_1}(P_1(p_{n_k}))}{s_{q_2}(P_2(p_{n_k}))} = \frac{\mu_{q_1} \log q_2}{\mu_{q_2} \log q_1} \cdot \frac{r_1}{r_2}.$$

Now, since by (3.1) and (3.2) we have that

$$\#A_{q,P}(N)/N \sim \#B_{q,P}(N)/N \sim 1$$

as $N \to \infty$, the sets $\mathcal{A}_{q,P} = \bigcup_{N\geq 1} A_{q,P}(N)$ and $\mathcal{B}_{q,P} = \bigcup_{N\geq 1} B_{q,P}(N)$ satisfy $d(\mathcal{A}_{q,P}) = d(\mathcal{B}_{q,P}) = 1$, and therefore

$$d(\mathcal{A}_{q_1,P_1} \cap \mathcal{A}_{q_2,P_2}) = d(\mathcal{B}_{q_1,P_1} \cap \mathcal{B}_{q_2,P_2}) = 1.$$

By Lemma 5.2, the limit for $n \to \infty$ is not altered when we only look at those n that lie in these subsets of asymptotic density one. Thus we get for the Cesàro mean that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{\substack{n\leq N \\ n\in\mathcal{A}_{q_1,P_1}\cap\mathcal{A}_{q_2,P_2}}} \frac{s_{q_1}(P_1(n))}{s_{q_2}(P_2(n))} = \frac{\mu_{q_1} \log q_2}{\mu_{q_2} \log q_1} \cdot \frac{r_1}{r_2}. \tag{3.5}$$

A combination of (3.3) and (3.5) yields

$$\begin{aligned}
&\lim_{N\to\infty}\frac{1}{N}\sum_{n\le N}\frac{s_{q_1}(P_1(n))}{s_{q_2}(P_2(n))}\\
&\quad=\lim_{N\to\infty}\frac{1}{N}\sum_{\substack{n\le N\\ n\in\mathcal{A}_{q_1,P_1}\cap\mathcal{A}_{q_2,P_2}}}\frac{s_{q_1}(P_1(n))}{s_{q_2}(P_2(n))}+\lim_{N\to\infty}\frac{1}{N}\sum_{\substack{n\le N\\ n\notin\mathcal{A}_{q_1,P_1}\cap\mathcal{A}_{q_2,P_2}}}\frac{s_{q_1}(P_1(n))}{s_{q_2}(P_2(n))}\\
&\quad=\frac{\mu_{q_1}\log q_2}{\mu_{q_2}\log q_1}\cdot\frac{r_1}{r_2},
\end{aligned}$$

which proves (1.2), and similarly we get (1.3). This completes the proof of Theorem 5.1. □

4. Concluding remarks

The above results do not hold in general for arbitrary q-additive functions. For example, let $f\colon\mathbb{N}\to\mathbb{N}$ be such that $f(n)=1$ for $n\ge 0$. This function is q-additive (for any $q\ge 2$) as it puts 1 on the least significant digit of n and 0 on all other digits. Then, for each $r\ge 1$, we have that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\frac{f(n^r)}{f(n)}=1.$$

On the other hand, if $f\colon\mathbb{N}\to\mathbb{N}$ is q-additive with $f(aq^k)=2^k$ for $a\in\{0,\dots,q-1\}$ and $k\ge 0$ then for each $r\ge 2$,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\frac{f(n^r)}{f(n)}=\infty.$$

In order to get other non-trivial values for the limit, the values of f must depend on the position k as well as on the digit a. We conclude our discussion with the following

Conjecture 5.5. *Let $q,r\ge 2$ be integers. Then for each real $\ell\in[1,\infty)$ there exists a q-additive function $f\colon\mathbb{N}\to\mathbb{N}$ such that*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\frac{f(n^r)}{f(n)}=\ell.$$

Acknowledgment

The authors want to thank Peter J. Grabner (Graz University of Technology) and Kevin G. Hare (University of Waterloo) for fruitful discussions on this problem.

Chapitre 6

Uniform Distribution of Prime Powers and sets of Recurrence and van der Corput sets in $\mathbb{Z}^k$

This chapter is joint work with Vitaly Bergelson, Grigori Kolesnik, Younghwan Son and Robert Tichy and appeared in the *Israel Journal of Mathematics*, **201** (2014), no. 2, 729 – 760.

1. Introduction

A. Sárkőzy established in [**206**], [**207**] and [**208**] the following surprising results :

THEOREM 6.1. *Let $E \subset \mathbb{N}$ be a set of positive upper density :*

$$\overline{d}(E) := \limsup_{N\to\infty} \frac{|E \cap \{1, 2, \cdots, N\}|}{N} > 0.$$

(1) Let $k \in \mathbb{N} = \{1, 2, 3, \ldots\}$. Then one can find arbitrary large $n \in \mathbb{N}$ such that for some $x, y \in E$, $x - y = n^k$.

(2) Denote by $\mathcal{P}$ be the set of prime numbers $\{2, 3, 5, 7, 11, \cdots\}$. One can find arbitrary large $p \in \mathcal{P}$ such that for some $x, y \in E$, $x - y = p - 1$. Also one can find arbitrarily large $q \in \mathcal{P}$ such that $x - y = q + 1$.

REMARK 10.

(1) In [**206**] the case of the equation $x - y = n^2$ is considered and a quantitative refinement of statement (i) is proved by an application of the Hardy-Littlewood method. Let $A(N) = |E \cap \{1, \ldots, N\}|$ and assume that the difference set of E does not contain a square of an integer. It is proved in [**206**] that

$$\frac{A(N)}{N} = O\left(\frac{(\log\log N)^{\frac{2}{3}}}{(\log N)^{\frac{1}{3}}}\right) = o(1),$$

which implies assertion (i) of Theorem 1.1. In [**207**] a lower bound for $A(N)$ is established and in [**208**] similar results are given for n^k, $k \in \mathbb{N}$, as well as a quantitative version of assertion (ii) of Theorem 1.1.

The best bound on square differences is by Pintz, Steiger and Szemerédi [**186**]. In particular, they combined the Hardy-Littlewood method with a combinatorial construction in order to show that

$$\frac{A(N)}{N} = O\left((\log N)^{-c_n}\right),$$

where $c_n \to \infty$.

(2) It is not hard to see that only shifts by 1 or -1 can "work" for the part (ii) of Theorem 6.1. (Just consider $E = 4\mathbb{N}$).

Theorem 6.1 can also be obtained with the help of the ergodic method introduced by H. Furstenberg in [**88**]. While the ergodic method does not provide sharp finitistic bounds, it allows us to see Sárkőzy's results as statements about recurrence in measure preserving systems and leads to a variety of strong extensions of Theorem 6.1.

To illustrate how the ergodic method works, let us consider, for example, the following polynomial refinement, due to Furstenberg, of the classical Poincaré recurrence theorem.

THEOREM 6.2 ([**87**], Theorem 3.16). *Let $(X, \mathcal{B}, \mu)$ be a probability space and let T be an invertible measure preserving transformation.* [1] *Let $A \in \mathcal{B}$ with $\mu(A) > 0$. For any $g(t) \in \mathbb{Z}[t]$ with $g(0) = 0$, there are arbitrarily large $n \in \mathbb{N}$ such that $\mu(A \cap T^{-g(n)}A) > 0$.*

Theorem 6.2 implies the following combinatorial result which generalizes Theorem 6.1 (i).

THEOREM 6.3 ([**87**]. Proposition 3.19). *Let $E \subset \mathbb{N}$ have positive upper Banach density :*

$$d^*(E) := \limsup_{N-M\to\infty} \frac{|E \cap \{M, M+1, \cdots, N-1\}|}{N-M} > 0.$$

For any $g(t) \in \mathbb{Z}[t]$ with $g(0) = 0$, there are arbitrarily large n such that

$$d^*(E \cap (E - g(n))) > 0.$$

To derive Theorem 6.3 from Theorem 6.2 one can utilize Furstenberg's correspondence principle (see [**25**]), which for the case in question says that for any $E \subset \mathbb{N}$ with $d^*(E) > 0$ there exist an invertible measure preserving system $(X, \mathcal{B}, \mu, T)$ and $A \in \mathcal{B}$ with $\mu(A) = d^*(E)$ such that for any $n \in \mathbb{Z}$ one has

$$d^*(E \cap E - n) \geq \mu(A \cap T^{-n}A).$$

One can also show that Theorem 6.3 implies Theorem 6.2. To see this one can utilize Theorem 6.1 from [**23**] (see also [**34**].)

DEFINITION 6.4. *A set $R \subset \mathbb{N}$ is called a set of recurrence if for any invertible measure preserving system $(X, \mathcal{B}, \mu, T)$ and any $A \in \mathcal{B}$ with $\mu(A) > 0$, there exists $n \in R$ such that $\mu(A \cap T^{-n}A) > 0$.*

Applications of ergodic theory to combinatorics and number theory bring to life various natural refinements of Definition 6.4. Here is a sample of some notions of recurrence relevant to this paper. [2]

— *Nice recurrence* (See [**24**]). A set $R \subset \mathbb{N}$ is called a set of nice recurrence if for any measure preserving system $(X, \mathcal{B}, \mu, T)$, any $A \in \mathcal{B}$ with $\mu(A) > 0$, and $\epsilon > 0$, there exist infinitely many $n \in R$ such that $\mu(A \cap T^{-n}A) \geq \mu(A)^2 - \epsilon$.

— *vdC sets* (See [**115**]). A set $H \subset \mathbb{N}$ is called a van der Corput set, or a vdC set if the uniform distribution mod 1 of the sequence $(x_{n+h} - x_n)_{n\in\mathbb{N}}$ for any $h \in H$ implies the uniform distribution mod 1 of the sequence $(x_n)_{n\in\mathbb{N}}$. Equivalently (see [**31**]), $H \subset \mathbb{N}$ is a vdC set if for any sequence of complex numbers $(u_n)_{n\in\mathbb{N}}$ of modulus 1, such that for any $h \in H$ $\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} u_{n+h}\overline{u_n} = 0$, one has $\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} u_n = 0$.

1. We will refer to the quadruple $(X, \mathcal{B}, \mu, T)$ as a measure preserving system and will tacitly assume that T is invertible and $\mu(X) = 1$.

2. For convenience of the discussion we define these notions for $\mathbb{N}$. We will introduce later the more general notions in $\mathbb{Z}^k$.

Clearly any set of nice recurrence is a set of recurrence. It is somewhat less obvious that any vdC set is a set of recurrence. (See [**115**] for the proof.) One can also show that not every set of recurrence is a set of nice recurrence (see [**160**]) and that not every set of recurrence is a vdC set (see [**44**].)

It turns out that the sets mentioned above, namely the sets $\mathcal{P}-1$, $\mathcal{P}+1$ as well as the sets of the form $\{g(n) : n \in \mathbb{Z}\}$, where $g(t) \in \mathbb{Z}[t]$ and $g(0) = 0$, are sets of nice recurrence and also *vdC* sets. (See, for example, [**27**] and [**31**].)

As a matter of fact the following simultaneous extension of Theorem 6.1 and Theorem 6.2 holds true. (See Proposition 1.22 and Corollary 2.13 in [**31**]. See also Theorem 6.31 below.)

THEOREM 6.5. *For any $g(t) \in \mathbb{Z}[t]$ with $g(0) = 0$, the sets $\{g(p-1) : p \in \mathcal{P}\}$ and $\{g(p+1) : p \in \mathcal{P}\}$ are sets of nice recurrence and also are vdC sets.*

One of the goals of this paper is to obtain a number of n-dimensional refinements and generalizations of Theorem 6.5.

Our proofs of the results on sets of (nice) recurrence and (various enhanced versions of) van der Corput sets rely on the following general result about uniform distribution, which is of independent interest.

THEOREM 6.9. Let $\xi(x) = \sum_{j=1}^{m} \alpha_j x^{\theta_j}$, where $0 < \theta_1 < \theta_2 < \cdots < \theta_m$, α_j are non-zero reals and assume that if all $\theta_j \in \mathbb{Z}^+$, then at least one α_j is irrational. Then the sequence $(\xi(p))_{p\in\mathcal{P}}$ is u.d. mod 1. [3]

One of the applications of Theorem 6.9 is the following von Neumann-type theorem along primes.

THEOREM 6.22. Let $c_1, \ldots, c_k$ be positive real numbers such that $c_i \notin \mathbb{N}$ for $i = 1, 2, \ldots, k$. Let $U_1, \ldots, U_k$ be commuting unitary operators on a Hilbert space $\mathcal{H}$. Then,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} U_1^{[p_n^{c_1}]} \cdots U_k^{[p_n^{c_k}]} f = f^*,$$

where p_n denotes n-th prime and f^* is the projection of f on $\mathcal{H}_{inv}(:= \{f \in \mathcal{H} : U_i f = f \text{ for all } i\})$.

Theorem 6.22, in turn, has the following corollaries.

COROLLARY 6.25. Let $c_1, c_2, \ldots, c_k$ be positive, non-integers. Let $T_1, T_2, \ldots, T_k$ be commuting, invertible measure preserving transformations on a probability space $(X, \mathcal{B}, \mu)$. Then, for any $A \in \mathcal{B}$ with $\mu(A) > 0$, one has

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T_1^{-[p_n^{c_1}]} \cdots T_k^{-[p_n^{c_k}]} A) \geq \mu^2(A),$$

where p_n denotes the n-th prime.

COROLLARY 6.26. Let $c_1, \cdots, c_k$ be positive non-integers. If $E \subset \mathbb{Z}^k$ with $d^*(E) > 0$, then there exists a prime p such that $([p^{c_1}], \cdots, [p^{c_k}]) \in E - E$. Moreover,

$$\liminf_{N\to\infty} \frac{|\{p \leq N : ([p^{c_1}], \cdots, [p^{c_k}]) \in E - E\}|}{\pi(N)} \geq d^*(E)^2,$$

3. We are tacitly assuming that the set $\mathcal{P} = (p_n)_{n\in\mathbb{N}}$ is naturally ordered, so that $(f(p))_{p\in\mathcal{P}}$ is just another way of writing $(f(p_n))_{n\in\mathbb{N}}$.

where $\pi(N)$ is the number of primes less than or equal to N.

Before formulating additional results to be proved in this paper we have to introduce some pertinent definitions. (A detailed discussion of various additional notions of sets of recurrence in $\mathbb{Z}^k$ is provided in Section 4.)

DEFINITION 6.6. *A set $D \subset \mathbb{Z}^k$ is a set of nice recurrence if given any ergodic measure preserving $\mathbb{Z}^k$-action $T = (T^{\mathbf{m}})_{(\mathbf{m}\in\mathbb{Z}^k)}$ on a probability space $(X, \mathcal{B}, \mu)$, any set $A \in \mathcal{B}$ with $\mu(A) > 0$ and any $\epsilon > 0$, we have*

$$\mu(A \cap T^{-\mathbf{d}}A) \geq \mu^2(A) - \epsilon$$

for infinitely many $\mathbf{d} \in D$.

DEFINITION 6.7 (cf. [**31**], Definition 1.2.1). *A subset D of $\mathbb{Z}^k\backslash\{0\}$ is a van der Corput set (vdC set) if for any family $(u_{\mathbf{n}})_{\mathbf{n}\in\mathbb{Z}^k}$ of complex numbers of modulus 1 such that*

$$\forall d \in D,\ \lim_{N_1,\cdots,N_k\to\infty} \frac{1}{N_1\cdots N_k} \sum_{\mathbf{n}\in\prod_{i=1}^k[0,N_i)} u_{\mathbf{n}+\mathbf{d}}\overline{u_{\mathbf{n}}} = 0$$

we have

$$\lim_{N_1,\cdots,N_k\to\infty} \frac{1}{N_1\cdots N_k} \sum_{\mathbf{n}\in\prod_{i=1}^k[0,N_i)} u_{\mathbf{n}} = 0.$$

The following results are obtained in Sections 4 and 5.

THEOREM 6.8 (cf. Theorem 6.31). *If α_i are positive integers and β_i are positive and non-integers, then*

$$D_1 = \{\Big((p-1)^{\alpha_1}, \cdots, (p-1)^{\alpha_k}, [(p-1)^{\beta_1}], \cdots, [(p-1)^{\beta_l}]\Big) : p \in \mathcal{P}\},$$

and

$$D_2 = \{\Big((p+1)^{\alpha_1}, \cdots, (p+1)^{\alpha_k}, [(p+1)^{\beta_1}], \cdots, [(p+1)^{\beta_l}]\Big) : p \in \mathcal{P}\}$$

are vdC sets and also sets of nice recurrence in $\mathbb{Z}^{k+l}$.

COROLLARY 6.36. Let D_1 and D_2 be as in Theorem 6.31. If $E \subset \mathbb{Z}^{k+l}$ with $d^*(E) > 0$, then for any $\epsilon > 0$

$$\{\mathbf{d} \in D_i : d^*(E \cap E - \mathbf{d}) \geq d^*(E)^2 - \epsilon\}$$

has positive lower relative density [4] in D_i for $i = 1, 2$. Furthermore,

$$\liminf_{N\to\infty} \frac{\big|\{p \leq N : \big((p-1)^{\alpha_1}, \cdots, (p-1)^{\alpha_k}, [(p-1)^{\beta_1}], \cdots, [(p-1)^{\beta_l}]\big) \in E - E\}\big|}{\pi(N)} > 0,$$

and

$$\liminf_{N\to\infty} \frac{\big|\{p \leq N : \big((p+1)^{\alpha_1}, \cdots, (p+1)^{\alpha_k}, [(p+1)^{\beta_1}], \cdots, [(p+1)^{\beta_l}]\big) \in E - E\}\big|}{\pi(N)} > 0.$$

4. For sets $A \subset B \subset \mathbb{Z}^m$, the lower relative density of A with respect to B is defined as

$$\liminf_{n\to\infty} \frac{|A \cap [-n, n]^m|}{|B \cap [-n, n]^m|}.$$

2. equidistribution

The goal of this section is to prove the following simultaneous extension of the results in [**195**] and [**234**] and to derive from it some useful corollaries.

THEOREM 6.9. *Let $\xi(x) = \sum_{j=1}^{m} \alpha_j x^{\theta_j}$, where $0 < \theta_1 < \theta_2 < \cdots < \theta_m$, α_j are non-zero reals and assume that if all $\theta_j \in \mathbb{Z}^+$, then at least one α_j is irrational. Then the sequence $(\xi(p))_{p\in\mathcal{P}}$ is u.d. mod 1.*

The following notation will be used throughout this paper.

(1) $e(x) = \exp(2\pi i x)$.

(2) $X \ll Y$ (or $X = O(Y)$) means $X \le CY$ for some positive constant C.

(3) $X \asymp Y$ for $C_1 X \le Y \le C_2 X$ for some positive constants C_1, C_2.

(4) $f(x) \lll A$ means that for any $\epsilon > 0$, there is a positive constant $C(\epsilon)$ such that
$$|f(x)| \le C(\epsilon) A x^{\epsilon}.$$

(5) $\sum_{p\le N}$ denotes the sum over primes.

(6) The von Mangoldt function is defined as
$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^k \text{ for some prime } p \text{ and integer } k \ge 1 \\ 0 & \text{otherwise} \end{cases}$$

(7) For $s \in \mathbb{N}$, the s-fold divisor function is defined as
$$d_s(n) = \sum_{n_1\cdots n_s = n} 1,$$
where the sum is extended over all products with s factors.

Before giving the proof of Theorem 6.9 we formulate some necessary auxiliary results. We start with the classical Weyl - van der Corput inequality.

LEMMA 6.10 (cf. [**97**, Lemma 2.7]). *Let k be a positive integer and $K = 2^k$. Assume that $X, X_1 \in \mathbb{N}$ and $X < X_1 < 2X$. For any positive $H_1, \cdots, H_k \ll X_1 - X$ and*
$$S = \left| \sum_{X\le x\le X_1} e(f(x)) \right|$$
we have
$$\left(\frac{S}{X_1 - X}\right)^K \le 8^{K-1} \left\{ \frac{1}{H_1^{K/2}} + \frac{1}{H_2^{K/4}} + \cdots + \frac{1}{H_k} + \frac{1}{H_1 \cdots H_k (X_1 - X)} \sum_{h_1=1}^{H_1} \cdots \sum_{h_k=1}^{H_k} \left| \sum_{x\in I(\underline{h})} e(f_1(x)) \right| \right\},$$
where $f_1(x) := f(\underline{h}, x) = h_1 \cdots h_k \int_0^1 \cdots \int_0^1 \frac{\partial^k}{\partial x^k} f(x + \underline{h}\cdot\underline{t})\, d\underline{t}$, $\underline{h} = (h_1, \cdots, h_k)$, $\underline{t} = (t_1, \cdots, t_k)$ and $I(\underline{h}) = (X, X_1 - h_1 - \cdots - h_k]$.

The next lemma provides a useful estimate for polynomial-like functions.

LEMMA 6.11 ([**97**, Theorem 2.9]). *Let $q \geq 0$ be an integer and $X \in \mathbb{N}$. Suppose that $f(x)$ has $(q+2)$ continuous derivatives on an interval $I \subset (X, 2X]$. Assume also that there is some constant G such that $|f^{(r)}(x)| \asymp GX^{-r}$ for $r = 1, \ldots, q+2$. Then*

$$S := \left| \sum_{x \in I} e(f(x)) \right| \ll G^{\frac{1}{4Q-2}} X^{1-\frac{q+2}{4Q-2}} + \frac{X}{G},$$

where $Q = 2^q$ and the implied constant in $\ll$ depends only on q and on the implied constants in $\asymp$.

We will also need the following estimate involving the von Mangoldt function, the proof of which is based on an identity of Vaughan's type.

LEMMA 6.12. *Assume $F(x)$ to be any function defined on the real line, supported on $[N/2, N]$ and bounded by F_0. Let further U, V, Z be any parameters satisfying $3 \leq U < V < Z < N$, $Z \geq 4U^2$, $N \geq 64Z^2U$, $V^3 \geq 32N$ and $Z - \frac{1}{2} \in \mathbb{N}$. Then*

$$\left| \sum_n \Lambda(n) F(n) \right| \ll K \log N + F_0 + L(\log N)^8,$$

where the summation over n is restricted to the interval $[N/2, N]$, and K and L are defined by

$$K = \max_M \sum_{m=1}^{\infty} d_3(m) \left| \sum_{Z < n \leq M} F(mn) \right|,$$

$$L = \sup \sum_{m=1}^{\infty} d_4(m) \left| \sum_{U < n < V} b(n) F(mn) \right|,$$

where the supremum is taken over all arithmetic functions $b(n)$ satisfying $|b(n)| \leq d_3(n)$.

DÉMONSTRATION. The inequality in question can be easily derived from Lemma 2 and Lemma 3 of Heath-Brown [**107**]. The reader should be warned that in [**107**] F, U, V and Z are denoted by f, u, v and z and our parameters N and M correspond to x and N, respectively. From Lemma 2 in [**107**] (which is of combinatorial nature) we immediately obtain the representation :

$$\sum_n \Lambda(n) F(n) = \Sigma_1 + \Sigma_1' - \Sigma_2 - \Sigma_2' - \Sigma_3 - \Sigma_3',$$

where the quantities on the right hand side satisfy the following estimates (see Lemma 3 in [**107**], Equations (7) and (8)) :

$$\Sigma_1, \Sigma_1', \Sigma_2, \Sigma_2' \ll K \log N,$$

$$\Sigma_3, \Sigma_3' \ll F_0 + L(\log N)^8.$$

Combining these estimates, the triangle inequality immediately yields our Lemma 6.12. □

Given a sequence $(x_n)_{n=1}^{\infty}$, its *discrepancy* is defined by

$$D_N(x_n) = \sup_I \left| \frac{\#\{n \leq N : x_n \in I \,(\mathrm{mod} 1)\}}{N} - (b-a) \right|,$$

where the sup is taken over all intervals $I = [a, b) \subset [0, 1)$.

Note that the sequence (x_n) is u.d. (mod1) if and only if $\lim_{N\to\infty} D_N(x_n) = 0$.

The proof of Theorem 6.9 will be achieved by showing that $\lim_{N\to\infty} D_N(f(p_n)) = 0$. In doing so we will be using the following version of Erdős-Turán Inequality.

LEMMA 6.13 (cf. [**66**, Theorem 1.21], [**131**, Theorem 2.5]). *For any real sequence $(x_n)_{n=1}^{\infty}$ and any positive integer N and $H \leq N$, one has :*

$$D_N(x_n) \ll \frac{1}{H} + \sum_{h=1}^{H} \frac{1}{h} \left| \frac{1}{N} \sum_{n=1}^{N} e(hx_n) \right|.$$

The following lemma will serve as the central tool in the proof of Theorem 6.9.

LEMMA 6.14. *Let $X, k, q \in \mathbb{N}$ with $k, q \geq 0$ and set $K = 2^k$ and $Q = 2^q$. Let $P(x)$ be a polynomial of degree k with real coefficients. Let $f(x)$ be a real $(q+k+2)$ times continuously differentiable function on $[X/2, X]$ such that $\left|f^{(r)}(x)\right| \asymp FX^{-r}$ $(r = 1, \ldots, q+k+2)$. Then, if $F = o(X^{q+2})$ for F and X large enough, we have*

$$S := \left| \sum_{X/2<x\leq X} e(f(x) + P(x)) \right| \ll X^{1-\frac{1}{K}} + X \left(\frac{\log^k X}{F} \right)^{\frac{1}{K}} + X \left(\frac{F}{X^{q+2}} \right)^{\frac{1}{(4KQ-2K)}}.$$

DÉMONSTRATION. Using Lemma 6.10 with $H_i = \frac{X}{2K}$, we obtain

$$\left| \frac{S}{X} \right|^K \ll \frac{1}{H_k} + \frac{1}{H_1 \cdots H_k X} \sum_{h_1=1}^{H_1} \cdots \sum_{h_k=1}^{H_k} \left| \sum_{x\in I(\underline{h})} e(f_1(x)) \right|,$$

where $I(\underline{h}) = (X/2, X - h_1 - \cdots - h_k]$ and

$$f_1(x) := f_1(\underline{h}, x) = h_1 \cdots h_k \left[\int_0^1 \cdots \int_0^1 \frac{\partial^k}{\partial x^k} f(x + \underline{h} \cdot \underline{t}) d\,\underline{t} + a_k k! \right],$$

where a_k is the leading coefficient of $P(x)$. The function $f_1(x)$ satisfies the conditions of Lemma 6.11 with $G = h_1 \cdots h_k F/X^k$. Thus its application yields

$$\begin{aligned} \left| \frac{S}{X} \right|^K &\ll \frac{1}{X} + \frac{1}{H_1 \cdots H_k X} \sum_{h_1=1}^{H_1} \cdots \sum_{h_k=1}^{H_k} \left(F^{\frac{1}{4Q-2}} X^{1-\frac{q+2}{4Q-2}} + \frac{X^{k+1}}{F h_1 \cdots h_k} \right) \\ &\ll \frac{1}{X} + \left(\frac{F}{X^{q+2}} \right)^{\frac{1}{4Q-2}} + \frac{\log^k X}{F}. \end{aligned}$$

This proves the Lemma. □

REMARK 11. Using a better choice of parameters $\underline{H} = (H_1, \cdots, H_q)$, we can easily improve the estimate in Lemma 6.14 but since our aim is to prove uniform distribution, the obtained estimate will be sufficient.

PROPOSITION 6.15. *Let $P(x)$ be a polynomial of degree k and $f(x) = \sum_{j=1}^{r} d_j x^{\theta_j}$ with $r \geq 1$, $d_r \neq 0$, d_j real, $0 < \theta_1 < \theta_2 < \cdots < \theta_r$ and $\theta_j \notin \mathbb{Z}^+$. Assume that $l < \theta_r < l+1$ for some l. Let $1 \leq |m| \leq N^{1/10}$. Then*

$$\left| \sum_{p\leq N} e(mf(p) + mP(p)) \right| \lll N^{1-\frac{1}{3K}} + \frac{N}{(mN^{\theta_r})^{1/K}} + N^{1-\frac{1}{64KL^5-4K}} + N^{1-1/10},$$

where $K = 2^k$ *and* $L = 2^l$.

DÉMONSTRATION. We split the sum S into $\leq \log N$ subsums of the form $\left|\sum_{X\leq p\leq 2X} e(mf(p)+mP(p))\right|$ with $2X \leq N$ and evaluate a typical one of them. We can obviously assume that $X \geq N^{9/10}$. By using partial summation formula we obtain

$$\begin{aligned} S &:= \left|\sum_{X\leq p\leq 2X} e(mf(p)+mP(p))\right| \\ &= \left|\sum_n \frac{\Lambda(n)}{\log n} e(mf(n)+mP(n))\right| + O(\sqrt{X}) \\ &\ll \frac{1}{\log X}\left|\sum_{n\in I} \Lambda(n) e(mf(n)+mP(n))\right| + O(\sqrt{X}), \end{aligned}$$

where I is a subinterval of $(X, 2X]$. Denote the last sum by S_1 and use Lemma 6.12 with $U = \frac{1}{4}X^{1/5}$, $V = 4X^{1/3}$ and Z the unique number in $\frac{1}{2} + \mathbb{N}$, which is closest to $\frac{1}{4}X^{2/5}$. We obtain

$$\begin{aligned} S_1 \ll 1 + \log X &\left|\sum_{x<\frac{2X}{Z}} d_3(x) \sum_{y>Z, \frac{X}{x}<y<\frac{2X}{x}} e(mf(xy)+mP(xy))\right| \\ + \log^8 X &\left|\sum_x d_4(x) \sum_{U<y<V, \frac{X}{x}<y\leq\frac{2X}{x}} b(y) e(mf(xy)+mP(xy))\right|. \end{aligned}$$

Denote the first sum by S_2 and the second sum by S_3. To evaluate S_2, we use Lemma 6.14 to estimate, for a fixed x, the sum over y. Here (denoting $Y = \frac{X}{x}$) we have

$$\left|\frac{\partial^j f(xy)}{\partial y^j}\right| \asymp X^{\theta_r} Y^{-j}$$

for any j. Furthermore for $j \geq 5(l+1)$ we have

$$\left|m\frac{\partial^j f(xy)}{\partial y^j}\right| \ll mX^{\theta_r - \frac{2}{5}j} \ll X^{\frac{1}{10}+l+1-\frac{2}{5}j} \leq X^{-1/2},$$

where we have used that $y > Z \gg X^{2/5}$. Thus an application of Lemma 6.14 yields the following estimate :

$$\begin{aligned} S_2 &\lll \sum_{x\leq 2X/Z} X/x \left[(\frac{x}{X})^{\frac{1}{K}} + (\frac{1}{mX^{\theta_r}})^{\frac{1}{K}} + X^{-\frac{1}{2}\frac{1}{4K\cdot 8L^5-2K}}\right] \\ &\lll X\left(X^{-\frac{2}{5K}} + \frac{1}{(mX^{\theta_r})^{\frac{1}{K}}} + X^{-\frac{1}{64KL^5-4K}}\right). \end{aligned}$$

Now we need to estimate S_3 :

$$S_3 \lll \sum_{\frac{X}{V}<x\le\frac{2X}{U}} \left| \sum_{\substack{U<y<V \\ \frac{X}{x}<y\le\frac{2X}{x}}} b(y)e(mf(xy)+mP(xy)) \right|.$$

We split the interval $(\frac{X}{V}, \frac{2X}{U}]$ into $\le \log X$ subintervals of the form $I = (X_1, 2X_1]$ and take one of them. Denote the corresponding sum by S_4 and use Cauchy's inequality :

$$\begin{aligned} |S_4|^2 &\le X_1 \sum_{x\in I} \left| \sum_y b(y)e(mf(xy)+mP(xy)) \right|^2 \\ &\ll X_1^2 \frac{X}{X_1} + X_1 \left| \sum_{x\in I} \sum_{A<y_1<y_2\le B} b(y_1)\overline{b(y_2)}e(m(f(xy_1)-f(xy_2)+P(xy_1)-P(xy_2))) \right|, \end{aligned}$$

where $A = \max\{U, \frac{X}{x}\}$ and $B = \min\{U, \frac{2X}{x}\}$. Changing the order of summation, we get

$$|S_4|^2 \lll XX_1 + X_1 \sum_{y_1,y_2} \left| \sum_x e(m(f(xy_1)-f(xy_2)+P(xy_1)-P(xy_2)) \right|.$$

Now we fix y_1 and $y_2 \ne y_1$. The function $g(x) := m(f(xy_1)-f(xy_2))$ satisfies the conditions of Lemma 6.14 :

$$|g^{(j)}(x)| \asymp m\frac{|y_1-y_2|}{y_1} X^{\theta_r} X_1^{-j} \ll mX^{\theta_r}\left(\frac{X}{V}\right)^{-j} \ll X^{\theta_r+\frac{1}{10}-\frac{2}{3}j} \ll X^{-\frac{1}{2}}$$

if $j \le 2l+3$. Using Lemma 6.14 with $q = 2l+3$ we obtain

$$\begin{aligned} |S_4|^2 &\lll XX_1 + X_1 \sum_{y_1,y_2} \left(X_1^{1-\frac{1}{K}} + X_1 \left(\frac{y_1 \log^k X}{m|y_1-y_2|X^{\theta_r}} \right)^{1/K} \right) + X_1 X^{-\frac{1}{2}\frac{1}{4K\cdot 2L^2-2K}} \\ &\lll XX_1 + X^{2-\frac{2}{3k}} + X^2 \left(\frac{\log^k X}{mX^{\theta_r}} \right)^{1/K} + X^{2-\frac{1}{16KL^2-4K}}. \end{aligned}$$

Summing over all the subintervals completes the proof. □

PROPOSITION 6.16. *Let $P(x)$ and $f(x)$ be as in Proposition 6.15. Then the discrepancy of the sequence $(f(p)+P(p))$ satisfies*

$$D_N \lll N^{-\frac{1}{10}} + N^{-\frac{1}{3K}} + N^{-\frac{\theta_r}{K}} + N^{-\frac{1}{64KL^5}}.$$

Consequently, the sequence $(f(p)+P(p))$ is u.d. mod1.

DÉMONSTRATION. We use Lemma 6.13 with $H = N^{1/10}$ and obtain

$$D_N \ll \frac{1}{H} + \sum_{h=1}^{H} \frac{1}{h} \left| \frac{1}{N} \sum_{p\le N} e(hf(p)+hP(p)) \right|.$$

Applying Proposition 6.15 we obtain the claimed result :

$$D_N \lll N^{-\frac{1}{10}} + N^{-\frac{1}{3K}} + N^{-\frac{\theta_r}{K}} + N^{-\frac{1}{64KL^5}}.$$

□

PROOF OF THEOREM 6.9. We will consider two cases.

Assume first that at least one $\theta_j \notin \mathbb{Z}^+$. Then the function $\xi(x)$ can be rewritten as $f(x)+P(x)$ as in Proposition 6.15, namely, $P(x)$ is a polynomial and $f(x) = \sum_{j=1}^{r} d_j x^{\theta_j}$ with $r \geq 1$, $d_r \neq 0$, d_j real, $0 < \theta_1 < \cdots < \theta_r$ and $\theta_j \notin \mathbb{Z}^+$, so $(\xi(p))$ is u.d. (mod1).

Now we assume that all $\theta_j \in \mathbb{Z}^+$, i.e. $\xi(x)$ is a polynomial and at least one coefficient α_j is irrational. Then $(\xi(p))$ is u.d. mod1 due to the result of Rhin. (See [**195**].) □

We list now some corollaries of Theorem 6.9

COROLLARY 6.17. *Let $\xi(x) = \sum_{j=1}^{m} \alpha_j x^{\theta_j}$ be as in Theorem 6.9. Then for any $h \in \mathbb{Z}$ $(\xi(p-h))_{p\in\mathcal{P}}$ is u.d.* mod1.

DÉMONSTRATION. Note that for $k < \theta < k+1$, where k is a non-negative integer, there are $a_1, a_2, \cdots, a_k$ and $g(x)$ such that

$$(x-h)^\theta = x^\theta + a_1 x^{\theta-1} + \cdots + a_k x^{\theta-k} + g(x) \quad \text{and} \quad \lim_{x\to\infty} g(x) = 0.$$

Then $\sum_{j=1}^{m} \alpha_j (p-h)^{\theta_j}$ can be written as the sum of $\widetilde{\xi}(p) + G(p)$, where $\widetilde{\xi}(x)$ is the function as in Theorem 6.9 and $\lim_{x\to\infty} G(x) = 0$. So the result follows. □

The following result follows from Corollary 6.17 via the classical Weyl criterion (see Theorem 6.2 in Chapter 1 of [**131**].)

COROLLARY 6.18. *Let $0 < \theta_1 < \theta_2 < \cdots < \theta_m$ and let $\gamma_1, \gamma_2, \ldots, \gamma_m$ be non-zero real numbers such that $\gamma_i \notin \mathbb{Q}$ if $\theta_i \notin \mathbb{N}$. Let h be an integer. Then*

$$((\gamma_1(p-h)^{\theta_1}, \gamma_2(p-h)^{\theta_2}, \cdots, \gamma_m(p-h)^{\theta_m}))_{p\in\mathcal{P}}$$

is u.d. mod1 *in $\mathbb{T}^m$.*

COROLLARY 6.19. *Let $\theta_1, \cdots, \theta_m$ and $\gamma_1, \cdots, \gamma_m$ be as in Corollary 6.18. Let q and t be positive integes such that $(t,q) = 1$ and let h be an integer. If $\theta_i \notin \mathbb{Q}$ for all i, then*

$$\{(\gamma_1(p-h)^{\theta_1}, \gamma_2(p-h)^{\theta_2}, \cdots, \gamma_m(p-h)^{\theta_m})\}$$

is u.d. mod 1 *in $\mathbb{T}^m$, where p describes the increasing sequence of prime numbers belonging to the congruence class $t + q\mathbb{N}$.*

The proof of Corollary 6.19 hinges on the following classical identity (see p.34 in [**162**]).

LEMMA 6.20. *For any $q \in \mathbb{N}$ and $b \in \mathbb{N}$ with $1 \leq b \leq q$, one has*

$$\frac{1}{q}\sum_{j=1}^{q} e\left(\frac{(n-b)j}{q}\right) = \begin{cases} 1 & n \equiv b \ (\mathrm{mod} q) \\ 0 & \textit{otherwise} \end{cases}$$

PROOF OF COROLLARY 6.19. Let $A_N = \{p \leq N : p \equiv t \mod q\}$. We need to show that for $(a_1, a_2, \cdots, a_m) \neq (0, 0, \cdots, 0)$,

$$\lim_{N\to\infty} \frac{1}{|A_N|} \sum_{p\in A_N} e\left(\sum_{i=1}^{m} a_i r_i (p-h)^{\theta_i}\right) = 0.$$

The result follows from

$$(i)\quad \sum_{\substack{p\le N\\ p\equiv t \mod q}} e\left(\sum_{i=1}^{m} a_i\gamma_i(p-h)^{\theta_i}\right) = \sum_{p\le N} e\left(\sum_{i=1}^{m} a_i\gamma_i(p-h)^{\theta_i}\right)\frac{1}{q}\sum_{j=1}^{q} e\left(\frac{(p-t)j}{q}\right)$$

$$= \frac{1}{q}\sum_{j=1}^{q}\sum_{p\le N} e\left(\sum_{i=1}^{m} a_i\gamma_i(p-h)^{\theta_i} + \frac{j}{q}(p-h) + \frac{j}{q}(t-h)\right)$$

and (ii)

$$\lim_{N\to\infty}\frac{|A_N|}{\pi(N)} = \lim_{N\to\infty}\frac{|\{p\le N : p\equiv t \mod q\}|}{\pi(N)} = \frac{1}{\phi(q)},$$

where ϕ is Euler's totient function. □

We will utilize Corollary 6.17 in the proof of the following proposition, which will be used in the next sections.

PROPOSITION 6.21. *Let $g(x) = \sum_{j=1}^{m}\alpha_j[x^{\theta_j}]$, where $\theta_1, \theta_2, \ldots, \theta_m$ are distinct positive real numbers and $\alpha_1, \alpha_2, \ldots, \alpha_m$ are non-zero reals. Let h be an integer.*

(1) If $\theta_j \notin \mathbb{Z}$ for all j and $\alpha_j \notin \mathbb{Z}$ for all j, then

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} e(g(p_n - h)) = 0. \tag{2.1}$$

(2) If one of α_j is irrational, then $(g(p-h))_{p\in\mathcal{P}}$ is u.d. mod1.

DÉMONSTRATION. Our argument is similar to that used in the proof of Lemma 5.12 in [**29**]. We will prove the case $h = 0$ with the help of Theorem 6.9. The case of non-zero h can be done similarly by invoking Corollary 6.17 instead of Theorem 6.9 and is omitted.

(i) Without loss of generality, we can assume that there exists l such that $\alpha_1, \ldots, \alpha_l \notin \mathbb{Q}$ and $\alpha_{l+1}, \ldots, \alpha_m \in \mathbb{Q}$. Furthermore we also assume that $\alpha_{l+1}, \ldots, \alpha_m$ have a common denominator q, thus denote $\alpha_j = \frac{c_j}{q}$ for $l+1 \le j \le m$.

We have

$$e(g(p_n)) = e([p_n^{\theta_1}]\alpha_1 + \cdots + [p_n^{\theta_m}]\alpha_m) = \prod_{j=1}^{l} f_j(p_n^{\theta_j}\alpha_j, p_n^{\theta_j}) \prod_{j=l+1}^{m} g_j([p_n^{\theta_j}]),$$

where $f_j(x,y) = e(x - \{y\}\alpha_j)$ $(1 \le j \le l)$, and $g_j(z) = e(c_j\frac{z}{q})$ $(l+1 \le j \le m)$.

Note that $f_j(x,y)$ are Riemann-integrable on $\mathbb{T}^2$ and $g_j(z)$ are continuous functions on $\mathbb{Z}_q = \mathbb{Z}/q\mathbb{Z}$, hence the function $\prod_{j=1}^{l} f_j \prod_{j=l+1}^{m} g_j$ is Riemann-integrable on $\mathbb{T}^{2l}\times\mathbb{Z}_q^{m-l}$.

It follows from Theorem 6.9 and the classical Weyl criterion that, for any $u \in \mathbb{N}$,

$$(p_n^{\theta_1}\alpha_1, p_n^{\theta_1}, \cdots, p_n^{\theta_l}\alpha_l, p_n^{\theta_l}, \frac{p_n^{\theta_{l+1}}}{u}, \cdots, \frac{p_n^{\theta_m}}{u})$$

is u.d. in $\mathbb{T}^{2l}\times\mathbb{T}^{m-l}$. Since $[x] \equiv a \pmod q$ is equivalent to $\frac{a}{q} \le \{\frac{x}{q}\} < \frac{a+1}{q}$, we have

$$(p_n^{\theta_1}\alpha_1, p_n^{\theta_1}, \cdots, p_n^{\theta_l}\alpha_l, p_n^{\theta_l}, [p_n^{\theta_{l+1}}], \cdots, [p_n^{\theta_m}])$$

is u.d. in $\mathbb{T}^{2l}\times\mathbb{Z}_q^{m-l}$. Hence, (2.1) follows.

(ii) By rearranging θ_i, we can write

$$g(x) = \sum_{i=1}^{s} a_i x^{\gamma_i} + \sum_{j=1}^{t} b_j [x^{\delta_j}],$$

where $\gamma_i \in \mathbb{N}$ and $\delta_j \in \mathbb{R}^+ \backslash \mathbb{N}$.

Then, for any non-zero integer r, we need to show

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e(rg(p_n)) = 0.$$

Without loss of generality, we assume that $b_1, \dots b_l$ are irrational and $b_{l+1}, \dots, b_t$ are rational. We also assume that $b_j = \frac{c_j}{q}$ $(j = l+1, \dots, t)$. Let $P(x) = \sum_{i=1}^{s} a_i x^{\gamma_i}$. Now consider the following two cases.

Case I. Suppose that some a_i is irrational. Note that

$$\begin{aligned} e(rg(p_n)) &= e\left(rP(p_n)\right) \prod_{j=1}^{l} e\left(rb_j (p_n)^{\delta_j} - rb_j \{(p_n)^{\delta_j}\}\right) \prod_{j=l+1}^{t} e\left(rb_j [(p_n)^{\delta_j}]\right) \\ &= f_0(P(p_n)) \prod_{j=1}^{l} f_j(b_j (p_n)^{\delta_j}, (p_n)^{\delta_j}) \prod_{j=l+1}^{t} g_j([(p_n)^{\delta_j}]), \end{aligned}$$

where $f_0(x) = e(rx)$, $f_j(x,y) = e(r(x - b_j\{y\}))$ $(1 \le j \le l)$, and $g_j(x) = e(rc_j \frac{x}{q})$ $(l+1 \le j \le t)$. Using the above argument with Theorem 6.9

$$\left(P(p_n), b_1 (p_n)^{\delta_1}, (p_n)^{\delta_1}, \cdots, b_l (p_n)^{\delta_l}, (p_n)^{\delta_l}, [(p_n)^{\delta_{l+1}}], \cdots, [(p_n)^{\delta_t}]\right)$$

is uniformly distributed on $\mathbb{T}^{2l+1} \times \mathbb{Z}_q^{t-l}$. Hence, $(g(p))_{p\in\mathcal{P}}$ is uniformly distributed mod1.

Case II. Suppose that all a_i are rational. Note that b_1 is irrational. Using the same method in Case I, the result follows from that

$$\left(P(p_n) + b_1 (p_n)^{\delta_1}, (p_n)^{\delta_1}, b_2 (p_n)^{\delta_2}, (p_n)^{\delta_2}, \cdots, b_l (p_n)^{\delta_l}, (p_n)^{\delta_l}, [(p_n)^{\delta_{l+1}}], \cdots, [(p_n)^{\delta_t}]\right)$$

is uniformly distributed on $\mathbb{T}^{2l} \times \mathbb{Z}_q^{t-l}$. □

3. Recurrence along non-integer prime powers

In this section we will prove the following ergodic theorem along the prime powers and derive some corollaries pertaining to sets of recurrence and sets of differences of positive upper Banach density in $\mathbb{Z}^k$.

THEOREM 6.22. *Let $c_1, \dots, c_k$ be positive real numbers such that $c_i \notin \mathbb{N}$ for $i = 1, 2, \dots, k$. Let $U_1, \dots, U_k$ be commuting unitary operators on a Hilbert space $\mathcal{H}$. Then,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} U_1^{[p_n^{c_1}]} \cdots U_k^{[p_n^{c_k}]} f = f^*,$$

where p_n denotes the n-th prime and f^* is the projection of f on $\mathcal{H}_{inv}$(:= $\{f \in \mathcal{H} : U_i f = f$ for all $i\}$).

For the proof of this theorem, we will need the following Hilbert space splitting theorem :

THEOREM 6.23 (cf. [**26**]). *Let $U_1, U_2, \dots, U_k$ be commuting unitary operators on a Hilbert space $\mathcal{H}$. Then we can split $\mathcal{H}$ in the following ways.*

(1) $\mathcal{H} = \mathcal{H}_{inv} \oplus \mathcal{H}_{erg}$, where

$$\mathcal{H}_{inv} = \{f \in \mathcal{H} : U_i f = f \text{ for all } i\},$$

and

$$\mathcal{H}_{erg} = \{f \in \mathcal{H} : \lim_{N_1,\cdots,N_k\to\infty} \left|\left| \frac{1}{N_1 \cdots N_k} \sum_{n_1=0}^{N_1-1} \cdots \sum_{n_k=0}^{N_k-1} U_1^{n_1} \cdots U_k^{n_k} f \right|\right| = 0\}.$$

(2) $\mathcal{H} = \mathcal{H}_{rat} \oplus \mathcal{H}_{tot}$, where

$$\mathcal{H}_{rat} = \overline{\{f \in \mathcal{H} : \text{there exists non-zero } k\text{-tuple } (m_1, m_2, \dots, m_k) \in \mathbb{Z}^k,\ U_i^{m_i} f = f \text{ for all } i\}},$$

and

$$\mathcal{H}_{tot} = \{f \in \mathcal{H} \ : \ \text{for any non-zero } (m_1, m_2, \dots, m_k)$$
$$\lim_{N_1,\cdots,N_k\to\infty} \left|\left| \frac{1}{N_1 \cdots N_k} \sum_{n_1=0}^{N_1-1} \cdots \sum_{n_k=0}^{N_k-1} U_1^{m_1 n_1} \cdots U_k^{m_k n_k} f \right|\right| = 0\}.$$

We will also need the following version of the classical Bochner-Herglotz theorem.

THEOREM 6.24. *Let $U_1, \cdots, U_k$ be commuting unitary operators on a Hilbert space $\mathcal{H}$ and $f \in \mathcal{H}$. Then there is a measure ν_f on $\mathbb{T}^k$ such that*

$$< U_1^{n_1} U_2^{n_2} \cdots U_k^{n_k} f, f > = \int_{\mathbb{T}^k} e^{2\pi i (n_1\gamma_1 + \cdots + n_k\gamma_k)}\, d\nu_f(\gamma_1, \cdots, \gamma_k),$$

for any $(n_1, n_2, \cdots, n_k) \in \mathbb{Z}^k$.

PROOF OF THEOREM 6.22. Without loss of generality we can assume that $c_1, \cdots, c_k$ are distinct. Consider Hilbert space splitting $\mathcal{H} = \mathcal{H}_{inv} \oplus \mathcal{H}_{erg}$. For $f \in \mathcal{H}_{inv}$, $U_1^{[p_n^{c_1}]} \cdots U_k^{[p_n^{c_k}]} f = f$. So let us assume that $f \in \mathcal{H}_{erg}$ and show that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} U_1^{[p_n^{c_1}]} \cdots U_k^{[p_n^{c_k}]} f = 0. \tag{3.1}$$

This will follow from Proposition 6.21 and Theorem 6.24. We have

$$\begin{aligned}
\left|\left| \frac{1}{N} \sum_{n=1}^{N} U_1^{[p_n^{c_1}]} \cdots U_k^{[p_n^{c_k}]} f \right|\right|_2^2 &= \frac{1}{N^2} \sum_{m,n=1}^{N} \langle U_1^{[p_m^{c_1}]} \cdots U_k^{[p_m^{c_k}]} f, U_1^{[p_n^{c_1}]} \cdots U_k^{[p_n^{c_k}]} f \rangle \\
&= \frac{1}{N^2} \sum_{m,n=1}^{N} \langle U_1^{[p_m^{c_1}]-[p_n^{c_1}]} \cdots U_k^{[p_m^{c_k}]-[p_n^{c_k}]} f, f \rangle \\
&= \frac{1}{N^2} \sum_{m,n=1}^{N} \int e(([p_m^{c_1}] - [p_n^{c_1}], \cdots, [p_m^{c_k}] - [p_n^{c_k}]) \cdot \gamma)\, d\nu_f(\gamma) \\
&= \int \left| \frac{1}{N} \sum_{n=1}^{N} e(([p_n^{c_1}], \cdots, [p_n^{c_k}]) \cdot \gamma) \right|^2 d\nu_f(\gamma)
\end{aligned}$$

Since $f \in H_{erg}$, we have $\nu_f(\{(0,\dots,0)\}) = 0$, so that, for our f, (3.1) follows. □

COROLLARY 6.25. *Let $c_1, c_2, \dots, c_k$ be positive, non-integers. Let $T_1, T_2, \dots, T_k$ be commuting, invertible measure preserving transformations on a probability space $(X, \mathcal{B}, \mu)$. Then, for any $A \in \mathcal{B}$ with $\mu(A) > 0$, one has*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T_1^{-[p_n^{c_1}]} \cdots T_k^{-[p_n^{c_k}]} A) \geq \mu^2(A).$$

DÉMONSTRATION. Let $f = 1_A$. A measure preserving transformation T_i can be considered as a unitary operator $T_i f = f \circ T_i$. Denote by P the projection on $\mathcal{H}_{inv}$ for $T_1, \dots, T_k$. Then we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T_1^{-[p_n^{c_1}]} \cdots T_k^{-[p_n^{c_k}]} A) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int f\, T_1^{[p_n^{c_1}]} \cdots T_k^{[p_n^{c_k}]} f \, d\mu$$

$$= \int fPf \, d\mu = \langle f, Pf\rangle = \langle f, P^2 f\rangle = \langle Pf, Pf\rangle \geq \left(\int Pf \, d\mu\right)^2 = \mu^2(A).$$

□

Recall that the upper Banach density of a set $E \subset \mathbb{Z}^k$ is defined to be

$$d^*(E) = \sup_{\{\Pi_n\}_{n\in\mathbb{N}}} \limsup_{n\to\infty} \frac{|E \cap \Pi_n|}{|\Pi_n|},$$

where the supremum is taken over all sequences of parallelepipeds

$$\Pi_n = [a_n^{(1)}, b_n^{(1)}] \times \cdots \times [a_n^{(k)}, b_n^{(k)}] \subset \mathbb{Z}^k, \ n \in \mathbb{N},$$

with $b_n^{(i)} - a_n^{(i)} \to \infty$, $1 \leq i \leq k$.

By the $\mathbb{Z}^k$-version of Furstenberg's correspondence principle (see, for example, Proposition 7.2 in [**32**]), given $E \subset \mathbb{Z}^k$ with $d^*(E) > 0$, there is a probability space $(X, \mathcal{B}, \mu)$, commuting invertible measure preserving transformations $T_1, T_2, \dots, T_k$ of X and $A \in \mathcal{B}$ with $d^*(E) = \mu(A)$ such that for any $\mathbf{n_1}, \mathbf{n_2}, \dots, \mathbf{n_m} \in \mathbb{Z}^k$ one has

$$d^*(E \cap (E - \mathbf{n_1}) \cap (E - \mathbf{n_2}) \cap \cdots \cap (E - \mathbf{n_m})) \geq \mu(A \cap T^{-\mathbf{n_1}} A \cap \cdots \cap T^{-\mathbf{n_m}} A),$$

where for $\mathbf{n} = (n_1, \dots, n_k)$, $T^{\mathbf{n}} = T_1^{n_1} \cdots T_k^{n_k}$.

We see now that Corollary 6.25 together with Furstenberg's correspondence principle implies the following result.

COROLLARY 6.26. *Let $c_1, \cdots, c_k$ be positive non-integers. If $E \subset \mathbb{Z}^k$ with $d^*(E) > 0$, then there exists a prime p such that $([p^{c_1}], \cdots, [p^{c_k}]) \in E - E$. Moreover,*

$$\liminf_{N\to\infty} \frac{|\{p \leq N : ([p^{c_1}], \cdots, [p^{c_k}]) \in E - E\}|}{\pi(N)} \geq d^*(E)^2.$$

DÉMONSTRATION. By a special case of Furstenberg's correspondence principle, given $E \subset \mathbb{Z}^k$ with $d^*(E) > 0$, there exist a probability space $(X, \mathcal{B}, \mu)$, commuting invertible measure preserving transformations $T_1, \dots, T_k$ of X and $A \in \mathcal{B}$ with $d^*(E) = \mu(A)$ such that for any $l_1, l_2, \dots, l_k \in \mathbb{Z}$ one has

$$d^*(E \cap (E - (l_1, l_2, \cdots, l_k)) \geq \mu(A \cap T_1^{-l_1} T_2^{-l_2} \cdots T_k^{-l_k} A).$$

Note that

$$\begin{aligned}|\{p \le N : ([p^{c_1}], \cdots, [p^{c_k}]) \in E - E\}| &\ge |\{p \le N : d^*(E \cap E - ([p^{c_1}], \cdots, [p^{c_k}]) > 0\}| \\ &\ge \sum_{p \le N} d^*(E \cap E - ([p^{c_1}], \cdots, [p^{c_k}])) \\ &\ge \sum_{p \le N} \mu(A \cap T_1^{-[p^{c_1}]} T_2^{-[p^{c_2}]} \cdots T_k^{-[p^{c_k}]} A).\end{aligned}$$

Hence, by Corollary 6.25,

$$\begin{aligned}&\liminf_{N\to\infty} \frac{|\{p \le N : ([p^{c_1}], \cdots, [p^{c_k}]) \in E - E\}|}{\pi(N)} \\ &\quad \ge \lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p \le N} \mu(A \cap T_1^{-[p^{c_1}]} T_2^{-[p^{c_2}]} \cdots T_k^{-[p^{c_k}]} A) \\ &\quad \ge \mu(A)^2 = d^*(E)^2.\end{aligned}$$

□

REMARK 12. It is not hard to see that Theorem 6.22, Corollary 6.25 and Corollary 6.26 remain true if one replaces in the formulations $([p^{c_1}], \cdots, [p^{c_k}])$ by $([(p-h)^{c_1}], \cdots, [(p-h)^{c_k}])$ where h is arbitrarily integer. We will utilize this remark for $h = \pm 1$ in the next section.

4. Application to Nice FC^+ sets

DEFINITION 6.27. *A sequence $(\mathbf{d}_n)_{n\in\mathbb{N}}$ in $\mathbb{Z}^k$ is called ergodic if the following mean ergodic theorem is valid : for any ergodic measure preserving $\mathbb{Z}^k$-action $T = (T^{\mathbf{m}})_{(\mathbf{m}\in\mathbb{Z}^k)}$ on a probability space $(X, \mathcal{B}, \mu)$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f \circ T^{\mathbf{d}_n} = \int f \, d\mu \text{ for any } f \in L^2(\mu).$$

Recall that a subset D of $\mathbb{Z}^k$ is *a set of recurrence* if given any measure preserving $\mathbb{Z}^k$-action $T = (T^{\mathbf{m}})_{(\mathbf{m}\in\mathbb{Z}^k)}$ on a probability space $(X, \mathcal{B}, \mu)$ and any set $A \in \mathcal{B}$ with $\mu(A) > 0$, there exists $\mathbf{d} \in D$ $(\mathbf{d} \neq 0)$ such that

$$\mu(A \cap T^{-\mathbf{d}} A) > 0.$$

DEFINITION 6.28. *Let D be a subset of $\mathbb{Z}^k$. We will write $D = \{\mathbf{d}_n : n \in \mathbb{N}\}$ with the convention that $\mathbf{d}_n$ are pairwise distinct and the sequence $(|\mathbf{d}_n|)$ is non-decreasing. (Here $|\mathbf{d}| = \sup_{1\le i\le k} |d_i|$ for $\mathbf{d} = (d_1, d_2, \cdots, d_k)$.)*

(1) (cf. [24]) A set D is a set of nice recurrence if given any measure preserving $\mathbb{Z}^k$-action $T = (T^{\mathbf{m}})_{(\mathbf{m}\in\mathbb{Z}^k)}$ on a probability space $(X, \mathcal{B}, \mu)$, any set $A \in \mathcal{B}$ with $\mu(A) > 0$ and any $\epsilon > 0$, we have

$$\mu(A \cap T^{-\mathbf{d}} A) \ge \mu^2(A) - \epsilon$$

for infinitely many $\mathbf{d} \in D$.

(2) (cf. [28] and [31]) A set D is an averaging set of recurrence if given any measure preserving $\mathbb{Z}^k$-action $T = (T^{\mathbf{m}})_{(\mathbf{m}\in\mathbb{Z}^k)}$ on a probability space $(X, \mathcal{B}, \mu)$ and any set $A \in \mathcal{B}$ with $\mu(A) > 0$ we have

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-\mathbf{d}_n} A) > 0.$$

DEFINITION 6.29 (cf. [**31**], Definition 1.2.1). *A subset D of $\mathbb{Z}^k\backslash\{0\}$ is a van der Corput set (vdC set) if for any family $(u_{\mathbf{n}})_{\mathbf{n}\in\mathbb{Z}^k}$ of complex numbers of modulus 1 such that*

$$\forall d \in D, \lim_{N_1,\cdots,N_k\to\infty} \frac{1}{N_1 \cdots N_k} \sum_{\mathbf{n}\in\prod_{i=1}^{k}[0,N_i)} u_{\mathbf{n}+\mathbf{d}}\overline{u_{\mathbf{n}}} = 0$$

we have

$$\lim_{N_1,\cdots,N_k\to\infty} \frac{1}{N_1 \cdots N_k} \sum_{\mathbf{n}\in\prod_{i=1}^{k}[0,N_i)} u_{\mathbf{n}} = 0.$$

DEFINITION 6.30 ([**31**]). *An infinite set D of $\mathbb{Z}^k$ is a nice FC^+ set if for any positive finite measure σ on $\mathbb{T}^k$,*

$$\sigma(\{(0, 0, \cdots, 0)\}) \leq \limsup_{|\mathbf{d}|\to\infty, \mathbf{d}\in D} |\hat{\sigma}(\mathbf{d})|.$$

REMARK 13. The following results are obtained in [**31**] for sets in $\mathbb{Z}$ and can be generalized to $\mathbb{Z}^k$.

(1) An ergodic sequence in $\mathbb{Z}^k$ is an averaging set of recurrence and a set of nice recurrence. This can be obtained by using the same argument as in the proof of Corollary 6.25.

(2) A nice FC^+ set in $\mathbb{Z}^k$ is a set of nice recurrence. The proof for $\mathbb{Z}$ was given in [**31**]. Here we add the proof for reader's convenience. Let $T = (T^{\mathbf{n}})_{\mathbf{n}\in\mathbb{Z}^k}$ be a measure preserving $\mathbb{Z}^k$-action on a probability space $(X, \mathcal{B}, \mu)$ and $A \in \mathcal{B}$. Then there exists a positive measure σ on $\mathbb{T}^k$ such that $\hat{\sigma}(\mathbf{n}) = \mu(A \cap T^{-\mathbf{n}} A)$ and $\sigma(\{(0, 0, \cdots, 0)\}) \geq \mu(A)^2$. Then the result follows.

(3) A nice FC^+ set is a vdC set. This is an immediate consequence of the following spectral characterization of vdC sets (see Theorem 1.8 in [**31**]) : A set $D \subset \mathbb{Z}^k$ is vdC if and only if any positive measure σ on $\mathbb{T}^k$ with $\hat{\sigma}(\mathbf{d}) = 0$ for all $\mathbf{d} \in D$ satisfies $\sigma(\{(0, 0, \cdots, 0\}) = 0$.

Next we also obtain the following result, which can be viewed as an extension of Sárkőzy's Theorem. (See [**206**] and [**208**].)

THEOREM 6.31. *If α_i are positive integers and β_i are positive and non-integers, then*

$$D_1 = \{\Big((p-1)^{\alpha_1}, \cdots, (p-1)^{\alpha_k}, [(p-1)^{\beta_1}], \cdots, [(p-1)^{\beta_l}]\Big) \,|\, p \in \mathcal{P}\},$$

and

$$D_2 = \{\Big((p+1)^{\alpha_1}, \cdots, (p+1)^{\alpha_k}, [(p+1)^{\beta_1}], \cdots, [(p+1)^{\beta_l}]\Big) \,|\, p \in \mathcal{P}\}$$

are nice FC^+ sets in $\mathbb{Z}^{k+l}$, and so they are vdC sets and also sets of nice recurrence.

REMARK 14. Recall that a set D of positive integers is a van der Corput set (or vdC set) if given a real sequence $(x_n)_{n\in\mathbb{N}}$, equidistribution mod1 of $(x_{n+d}-x_n)_{n\in\mathbb{N}}$ for all $d\in D$ implies the equidistribution of $(x_n)_{n\in\mathbb{N}}$. Let $\mathcal{P}$ be the set of all prime numbers. It is shown in [**115**] that $\mathcal{P}-h$ is a vdC set if and only if $h=\pm 1$. Since a nice FC^+ set is a vdC set (see section 3.5 in [**31**]), we cannot replace ± 1 by any other integer h on Theorem 6.31.

The following lemma, which tells us how to recognize a nice FC^+ set, will be utilized in the proof of Theorem 6.31.

LEMMA 6.32 (cf. [**31**] Proposition 2.11). *Let $D\subset\mathbb{Z}^k$. For each $q\in\mathbb{N}$, define*

$$D_q := \{\mathbf{d}=(d_1,d_2,\dots,d_k)\in D : q! \text{ divides } d_i \text{ for } 1\le i\le k\}.$$

Suppose that, for every q, there exists a sequence $(\mathbf{d}^{q,n})_{n\in\mathbb{N}}$ in D_q such that

(i) $|\mathbf{d}^{q,n}|$ is non-decreasing and (ii) for any $\mathbf{x}=(x_1,\cdots,x_k)\in\mathbb{R}^k$, if one of x_i is irrational, the sequence $(\mathbf{x}\cdot\mathbf{d}^{q,n})_{n\in\mathbb{N}}$ is uniformly distributed mod1.

Then D is a nice FC^+ set.

DÉMONSTRATION. For simplicity of notation we will confine ourselves to the case $k=1$. In this case we write $d^{q,n}$ for $\mathbf{d}^{q,n}$.

We need to show that, for any positive finite measure σ on $\mathbb{T}$,

$$\sigma(\{0\}) \le \limsup_{d\in D,|d|\to\infty} |\hat{\sigma}(d)|.$$

Given q, define $f_N(x)=\frac{1}{N}\sum_{n=1}^{N} e(d^{q,n}x)$. Let $A_q=\{\frac{a}{q!} : 0\le a\le q!-1, a\in\mathbb{N}\}\subset\mathbb{T}$ and let $B_q=\{r\in\mathbb{T}\cap\mathbb{Q} : r\notin A_q\}$. Then $\lim_{N\to\infty} f_N(x)=0$ if x is irrational and $\lim_{N\to\infty} f_N(x)=1$ if $x\in A_q$. Since B_q is countable, we can choose a sequence N_j such that $\lim_{N_j\to\infty} f_{N_j}(x)$ exists for every $x\in B_q$, thus for every $x\in\mathbb{T}$. Let $f(x):=\lim_{N_j\to\infty} f_{N_j}(x)$. Note that $0\le|f(x)|\le 1$ for all x.

By the dominated convergence theorem,

$$\int_{\mathbb{T}} f(x)\,d\sigma = \lim_{N_j\to\infty}\frac{1}{N_j}\sum_{n=1}^{N_j}\int e(d^{q,n}x)\,d\sigma = \lim_{N_j\to\infty}\frac{1}{N_j}\sum_{n=1}^{N_j}\hat{\sigma}(d^{q,n}). \tag{4.1}$$

Since $f(x)=0$ for $x\in\mathbb{T}\backslash\mathbb{Q}$,

$$\begin{aligned}\left|\int_{\mathbb{T}} f(x)\,d\sigma\right| &= \left|\int_{A_q} f(x)\,d\sigma+\int_{B_q} f(x)\,d\sigma+\int_{\mathbb{T}\backslash\mathbb{Q}} f(x)\,d\sigma\right| \\ &= \left|\int_{A_q} f(x)\,d\sigma+\int_{B_q} f(x)\,d\sigma\right| \ge \int_{A_q} f(x)\,d\sigma-\int_{B_q}|f(x)|\,d\sigma \\ &\ge \sigma(A_q)-\sigma(B_q). \qquad (4.2)\end{aligned}$$

Also we have

$$\begin{aligned}\limsup_{d\in D,|d|\to\infty}|\hat{\sigma}(d)| &\ge \limsup_{n\to\infty}|\hat{\sigma}(d^{q,n})| \\ &\ge \limsup_{N_j\to\infty}\frac{1}{N_j}\sum_{n=1}^{N_j}|\hat{\sigma}(d^{q,n})| \ge \left|\lim_{N_j\to\infty}\frac{1}{N_j}\sum_{n=1}^{N_j}\hat{\sigma}(d^{q,n})\right|. \qquad (4.3)\end{aligned}$$

From equations (4.1), (4.2) and (4.3),

$$\sigma(A_q) - \sigma(B_q) \leq \limsup_{d\in D, |d|\to\infty} |\hat{\sigma}(d)|.$$

By the continuity of the measure, $\lim_{q\to\infty} \sigma(A_q) = \sigma(\mathbb{T}\cap\mathbb{Q})$ and $\lim_{q\to\infty} \sigma(B_q) = 0$. So,

$$\sigma(\{0\}) \leq \sigma(\mathbb{T}\cap\mathbb{Q}) \leq \limsup_{d\in D, |d|\to\infty} |\hat{\sigma}(d)|.$$

□

PROPOSITION 6.33. *Let D_1 and D_2 be as in Theorem 6.31 and $D_i^{(r)} = D_i \bigcap (\bigoplus_{j=1}^{k+l} r\mathbb{Z})$. Then $D_i^{(r)}$ has positive lower relative density in D_i for $i = 1, 2$.*

DÉMONSTRATION. Let us prove this for D_1. Without loss of generality we can assume that all β_i are distinct. Note that if $p \in (1 + r\mathbb{Z}) \bigcap \mathcal{P}$ and $0 \leq \left\{\frac{(p-1)^{\beta_i}}{r}\right\} < \frac{1}{r}$ for $1 \leq i \leq l$, then $\left((p-1)^{\alpha_1}, \cdots, (p-1)^{\alpha_k}, [(p-1)^{\beta_1}], \cdots, [(p-1)^{\beta_l}]\right) \in D_1$. The result follows from Corollary 6.19 :

$$\left(\frac{(p-1)^{\beta_1}}{r}, \cdots, \frac{(p-1)^{\beta_l}}{r}\right)$$

is uniformly distributed mod1 in $\mathbb{T}^l$ along the increasing sequence of primes $p \in 1 + r\mathbb{Z}$. The proof for D_2 is completely analogous. □

PROOF OF THEOREM 6.31. Let us prove that D_1 is a nice FC^+ set.

Denote $D_1 = (\mathbf{d}_n)_{n\in\mathbb{N}}$, where

$$\mathbf{d}_n = \left((p_n-1)^{\alpha_1}, \cdots, (p_n-1)^{\alpha_k}, [(p_n-1)^{\beta_1}], \cdots, [(p_n-1)^{\beta_l}]\right).$$

Enumerate the elements of $D_1^{(q!)}$ by $(\mathbf{d}^{q,n})_{n\in\mathbb{N}}$, where $|\mathbf{d}^{q,n}|$ is non-decreasing. From Lemma 6.32, it is sufficient to show that for any $\mathbf{x} = (x_1, x_2, \cdots, x_{k+l})$, if one of x_i is irrational, $(\mathbf{d}^{q,n} \cdot \mathbf{x})_{n\in\mathbb{N}}$ is u.d. mod1.

For any non-zero integer h, by Lemma 6.20,

$$\begin{aligned}
&\frac{1}{|\{n \leq N : \mathbf{d}_n \in D_1^{(q!)}\}|} \sum_{n\leq N} e(h(\mathbf{d}^{q,n} \cdot \mathbf{x})) \\
&= \frac{1}{|\{n \leq N : \mathbf{d}_n \in D_1^{(q!)}\}|} \sum_{n\leq N} e(h(\mathbf{d}^{q,n} \cdot \mathbf{x})) \frac{1}{(q!)^{k+l}} \sum_{j_1=1}^{q!} \cdots \sum_{j_{k+l}=1}^{q!} e\left(\mathbf{d}_n \cdot \left(\frac{j_1}{q!}, \cdots, \frac{j_{k+l}}{q!}\right)\right) \\
&= \frac{N}{|\{n \leq N : \mathbf{d}_n \in D_1^{(q!)}\}|} \frac{1}{(q!)^{k+l}} \sum_{j_1=1}^{q!} \cdots \sum_{j_{k+l}=1}^{q!} \frac{1}{N} \sum_{n\leq N} e\left(\mathbf{d}_n \cdot \left(h\mathbf{x} + \left(\frac{j_1}{q!}, \cdots, \frac{j_{k+l}}{q!}\right)\right)\right).
\end{aligned}$$

Then the result follows from Proposition 6.21 and Proposition 6.33. The proof for D_2 is completely analogous. □

COROLLARY 6.34. *Let D_1 and D_2 be as in Theorem 6.31. If $E \subset \mathbb{Z}^{k+l}$ with $d^*(E) > 0$, then for any $\epsilon > 0$*

$$R_i(E, \epsilon) := \{\mathbf{d} \in D_i : d^*(E \cap E - \mathbf{d}) \geq d^*(E)^2 - \epsilon\}$$

is infinite for $i = 1, 2$.

We will see in the next section that the sets $R_i(E,\epsilon)$ actually have positive lower relative density.

5. Uniform distribution and sets of recurrence

THEOREM 6.35. *Let D_1 and D_2 be as in Theorem 6.31 and enumerate the elements of D_1 or D_2 as follows (where the sign $-$ corresponds to D_1 and sign $+$ corresponds to D_2) :*

$$\mathbf{d}_n = \left((p_n \pm 1)^{\alpha_1}, \cdots, (p_n \pm 1)^{\alpha_k}, [(p_n \pm 1)^{\beta_1}], \cdots, [(p_n \pm 1)^{\beta_l}]\right).$$

For each $r \in \mathbb{N}$, let $D_i^{(r)} = D_i \cap \bigoplus_{j=1}^{k+l} r\mathbb{Z}$ and enumerate the elements of $D_i^{(r)}$ by $(\mathbf{d}_n^{(r)})$ such that $|\mathbf{d}_n^{(r)}|$ is non-decreasing. Let $(T^{\mathbf{d}})_{\mathbf{d}\in\mathbb{Z}^{k+l}}$ be a measure preserving $\mathbb{Z}^{k+l}$-action on a probability space $(X,\mathcal{B},\mu)$. Then for $A \in \mathcal{B}$ with $\mu(A) > 0$ and $\epsilon > 0$, there exists $r \in \mathbb{N}$ such that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-\mathbf{d}_n^{(r)}} A) \geq \mu(A)^2 - \epsilon. \tag{5.1}$$

Moreover,

(1)

$$\{\mathbf{d} \in D_i : \mu(A \cap T^{-\mathbf{d}} A) \geq \mu^2(A) - \epsilon\} \tag{5.2}$$

has positive lower relative density in D_i for $i = 1, 2$. Hence, D_1 and D_2 are sets of nice recurrence.

(2)

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-\mathbf{d}_n} A) > 0. \tag{5.3}$$

Thus D_1 and D_2 are averaging sets of recurrence.

DÉMONSTRATION. We will prove this result for D_1. (The proof for D_2 is similar.) For $\mathbb{Z}^{k+l}$-action T, there are commuting measure preserving transformations $T_1, \cdots, T_{k+l}$ such that $T^{\mathbf{m}} = T_1^{m_1} \cdots T_{k+l}^{m_{k+l}}$ for $\mathbf{m} = (m_1, m_2, \cdots, m_{k+l})$.

First we will show that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-\mathbf{d}_n} A) \tag{5.4}$$

and

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-\mathbf{d}_n^{(r)}} A) \tag{5.5}$$

exist.

By Theorem 6.24, there exists a measure ν on $\mathbb{T}^{k+l}$ such that

$$\mu(A \cap T^{-\mathbf{n}} A) = \int 1_A \; T^{\mathbf{n}} 1_A \, d\mu = \int_{\mathbb{T}^{k+l}} e(\mathbf{n} \cdot \gamma) \, d\nu(\gamma).$$

Thus, in order to prove that (5.4) and (5.5) exist, it is sufficient to show that for every γ,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}e(\mathbf{d}_n\cdot\gamma)\quad\text{and}\quad\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}e(\mathbf{d}_n^{(r)}\cdot\gamma)$$

exist. Moreover, by Lemma 6.20,

$$\begin{aligned}&\frac{1}{N}\sum_{n=1}^{N}e(\mathbf{d}_n^{(r)}\cdot\gamma)\\&=\frac{1}{N}\sum_{n=1}^{N}e\left(\mathbf{d}_n\cdot\gamma\right)\left(\frac{1}{r}\sum_{j_1=1}^{r}e\left(\frac{(p_n-1)^{\alpha_1}j_1}{r}\right)\right)\cdots\left(\frac{1}{r}\sum_{j_{k+l}=1}^{r}e\left(\frac{[(p_n-1)^{\beta_l}]j_{k+l}}{r}\right)\right)\\&=\frac{1}{r^{k+l}}\sum_{j_1=1}^{r}\cdots\sum_{j_{k+l}=1}^{r}\frac{1}{N}\sum_{n=1}^{N}e\left(\mathbf{d}_n\cdot\left(\gamma+\left(\frac{j_1}{r}+\cdots+\frac{j_{k+l}}{r}\right)\right)\right).\end{aligned}$$

Hence, we only need to show that $\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}e(\mathbf{d}_n\cdot\gamma)$ exists for every γ. From Proposition 6.21, if $\gamma\notin\mathbb{Q}^{k+l}$, $\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}e(\mathbf{d}_n\cdot\gamma)=0$. If $\gamma=(\gamma_1,\gamma_2,\cdots,\gamma_{k+l})\in\mathbb{Q}^{k+l}$, then we can find a common denominator $q\in\mathbb{N}$ for $\gamma_1,\ldots,\gamma_{k+l}$ such that $\gamma_i=\frac{a_i}{q}$ for each i. Then

$$\begin{aligned}&\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}e(\mathbf{d}_n\cdot\gamma)\\&=\lim_{N\to\infty}\frac{1}{\pi(N)}\sum_{p\le N}e\left(\sum_{i=1}^{k}(p-1)^{\alpha_i}\frac{a_i}{q}+\sum_{j=1}^{l}[(p-1)^{\beta_j}]\frac{a_{k+j}}{q}\right)\\&=\lim_{N\to\infty}\frac{1}{\pi(N)}\sum_{\substack{(t,q)=1\\0\le t\le q-1}}\sum_{\substack{p\equiv t \bmod q\\p\le N}}e\left(\sum_{i=1}^{k}(p-1)^{\alpha_i}\frac{a_i}{q}+\sum_{j=1}^{l}[(p-1)^{\beta_j}]\frac{a_{k+j}}{q}\right)\\&=\sum_{\substack{(t,q)=1\\0\le t\le q-1}}e\left(\sum_{i=1}^{k}(t-1)^{\alpha_i}\frac{a_i}{q}\right)\lim_{N\to\infty}\frac{1}{\pi(N)}\sum_{\substack{p\equiv t \bmod q\\p\le N}}e\left(\sum_{j=1}^{l}[(p-1)^{\beta_j}]\frac{a_{k+j}}{q}\right).\end{aligned}$$

We claim that

$$\lim_{N\to\infty}\frac{1}{\pi(N)}\sum_{\substack{p\equiv t \bmod q\\p\le N}}e\left(\sum_{j=1}^{l}[(p-1)^{\beta_j}]\frac{a_{k+j}}{q}\right)$$

exists. Without loss of generality we assume that all β_i are distinct. Then the claim is a consequence of following two facts :

(1) $([(p-1)^{\beta_1}],\cdots,[(p-1)^{\beta_l}])$ is u.d. in $\mathbb{Z}_q^l$ along $p\in t+q\mathbb{Z}$ for $(t,q)=1$, since $\left(\frac{(p-1)^{\beta_1}}{q},\cdots,\frac{(p-1)^{\beta_l}}{q}\right)$ is u.d. mod1 in $\mathbb{T}^l$ along $p\in t+q\mathbb{Z}$ from Corollary 6.19.

(2) $\{p\in\mathcal{P}:p\equiv t \bmod q\}$ has a density $\frac{1}{\phi(q)}$ in $\mathcal{P}$ for $(t,q)=1$.

Now let us show (5.1). Applying Theorem 6.23 to (unitary operators induced by) $T_1, \cdots, T_{k+l}$ we have $1_A = f + g$, where $f \in \mathcal{H}_{rat}$ and $g \in \mathcal{H}_{tot}$. Note that $\mathcal{H}_{rat} = \overline{\bigcup_{q=1}^{\infty} \mathcal{H}_q}$, where $\mathcal{H}_q = \{f : T_i^{q!} f = f \text{ for } i = 1, 2, \ldots, k+l\}$.

For $\epsilon > 0$, there exists $\mathbf{a} = (a_1, \cdots, a_{k+l}) \in \mathbb{Z}^{k+l}$ and $f_{\mathbf{a}} \in \mathcal{H}_{rat}$ such that $T^{\mathbf{a}} f_{\mathbf{a}} = f_{\mathbf{a}}$, $||f_{\mathbf{a}} - f|| < \epsilon/2$ and $\int f_{\mathbf{a}}\, d\mu = \mu(A)$.

Choose r large such that $a_i | r$ for all i. Note that the set of $\{\mathbf{d}_n^{(r)}\}$ has relative positive density in D_1 . Consider

$$\frac{1}{N}\sum_{n=1}^{N} \mu(A \cap T^{-\mathbf{d}_n^{(r)}} A) = \frac{1}{N}\sum_{n=1}^{N} \int f\, T^{\mathbf{d}_n^{(r)}} f\, d\mu + \frac{1}{N}\sum_{n=1}^{N} \int g\, T^{\mathbf{d}_n^{(r)}} g\, d\mu.$$

For $f \in \mathcal{H}_{rat}$,

$$\begin{aligned} \int f\, T^{\mathbf{d}_n^{(r)}} f\, d\mu &= \langle f_{\mathbf{a}}, f_{\mathbf{a}} \rangle + \langle f_{\mathbf{a}}, T^{\mathbf{d}_n^{(r)}}(f - f_{\mathbf{a}}) \rangle + \langle f - f_{\mathbf{a}}, T^{\mathbf{d}_n^{(r)}} f \rangle \\ &\geq \mu^2(A) - \epsilon, \end{aligned}$$

Also note that $(\mathbf{d}_n^{(r)} \cdot \gamma)$ is u.d mod1 for $\gamma \notin (\mathbb{Q}/\mathbb{Z})^{k+l}$. Hence,

$$\frac{1}{N}\sum_{n=1}^{N} \int g\, T^{\mathbf{d}_n^{(r)}} g\, d\mu = \int \frac{1}{N}\sum_{n=1}^{N} e(\mathbf{d}_n^{(r)} \cdot \gamma)\, d\nu(\gamma) \to 0,$$

since $\nu(\mathbb{Q}/\mathbb{Z})^{k+l} = 0$ due to $g \in \mathcal{H}_{tot}$. Then,

$$\frac{1}{N}\sum_{n=1}^{N} \mu(A \cap T^{-\mathbf{d}_n^{(r)}} A) \geq \mu(A)^2 - \epsilon.$$

By Proposition 6.33, $\{\mathbf{d} \in D_1 : \mu(A \cap T^{-\mathbf{d}} A) \geq \mu^2(A) - \epsilon\}$ has positive lower relative density in D_1.

For (5.3), choose ϵ small such that $\mu^2(A) - \epsilon \geq \mu^2(A)/2$. Since $(\mathbf{d}_n^{(r)})$ has positive lower density, say α, we have

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mu(A \cap T^{-\mathbf{d}_n} A) \geq \alpha \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mu(A \cap T^{-\mathbf{d}_n^{(r)}} A) \geq \frac{\alpha}{2}\mu^2(A).$$

□

Via Furstenberg's correspondence principle, one can deduce the following corollary. (See also proof of Corollary 6.26.)

Corollary 6.36. *Let D_1 and D_2 be as in Theorem 6.31. If $E \subset \mathbb{Z}^{k+l}$ with $d^*(E) > 0$, then for any $\epsilon > 0$*

$$\{\mathbf{d} \in D_i : d^*(E \cap E - \mathbf{d}) \geq d^*(E)^2 - \epsilon\}$$

has positive lower relative density in D_i for $i = 1, 2$. Furthermore,

$$\liminf_{N\to\infty} \frac{\left|\{p \leq N : \left((p-1)^{\alpha_1}, \cdots, (p-1)^{\alpha_k}, [(p-1)^{\beta_1}], \cdots, [(p-1)^{\beta_l}]\right) \in E - E\}\right|}{\pi(N)} > 0.$$

$$\liminf_{N\to\infty} \frac{\left|\{p \leq N : \left((p+1)^{\alpha_1}, \cdots, (p+1)^{\alpha_k}, [(p+1)^{\beta_1}], \cdots, [(p+1)^{\beta_l}]\right) \in E - E\}\right|}{\pi(N)} > 0.$$

Acknowledgements

We would like to thank Angelo Nasca for helpful remarks on the preliminary draft of this paper.

Chapitre 7

Forms of differing degrees over number fields

This chapter is joint work with Christopher Frei and appeared in the *Mathematika*, **63** (2017), no. 1, 92 – 123.

1. Introduction

1.1. Main result. Let K be a number field of degree n over $\mathbb{Q}$. We consider a system of polynomials in s variables over the ring of integers $\mathcal{O}_K$ of K and let D be the maximum of their degrees. We assume the polynomials to be ordered by their degrees, that is, for each $d \in \{1, \dots, D\}$, we are given polynomials

$$G_{d,1}, \dots, G_{d,t_d} \in \mathcal{O}_K[x_1, \dots, x_s]$$

of degree d, where $t_d \geq 0$ with $t_D \geq 1$. The total number of polynomials is $T := t_1 + \cdots + t_D$. We are interested in quantitative statements about the common zeros of these polynomials.

To this end, we fix an integral ideal $\mathfrak{n}$ of $\mathcal{O}_K$ and a $\mathbb{Z}$-basis $\omega_1, \dots, \omega_n$ of $\mathfrak{n}$. We will also consider $\omega_1, \dots, \omega_n$ as an $\mathbb{R}$-basis of $V := K \otimes_{\mathbb{Q}} \mathbb{R}$. By a *box $\mathcal{B}$ aligned to the basis*, we mean the set of all $\mathbf{x} = (x_1, \dots, x_s) \in V^s$, where each x_i has the form $x_{i,1}\omega_1 + \cdots + x_{i,n}\omega_n$, such that the coordinates $\mathbf{X} = (x_{i,j})_{i,j} \in \mathbb{R}^{ns}$ lie in a given box $B \subseteq [-1,1]^{ns}$ with sides aligned to the coordinate axes of $\mathbb{R}^{ns}$. Given such a box $\mathcal{B}$, we study the asymptotics of the counting function

$$N(P) := \#\{\mathbf{x} \in \mathfrak{n}^s \cap P\mathcal{B} \,:\, G_{d,i}(\mathbf{x}) = 0 \text{ for all } 1 \leq d \leq D, 1 \leq i \leq t_d\}.$$

This counting function is a classical object of interest and has been investigated in many special cases. To name a few, Birch [**40**] considered forms over $\mathbb{Q}$ with all degrees equal. Skinner [**230**] generalized Birch's result to arbitrary number fields K. Schmidt [**220**] and Browning and Heath-Brown [**48**] considered forms over $\mathbb{Q}$ whose degrees might be different.

The main purpose of this article is to generalize the work of Browning and Heath-Brown [**48**] to arbitrary number fields, just as Skinner did with Birch's work. In addition, we fix an error in Skinner's paper [**230**], see Subsection 1.5. To state the main result, we need to introduce some further notation. Let

$$\Delta := \{1 \leq d \leq D \,:\, t_d \geq 1\}.$$

For each $1 \leq d \leq D$, $1 \leq i \leq t_d$, let $F_{d,i}$ be the leading form of the polynomial $G_{d,i}$, that is, the degree-d-part of $G_{d,i}$. For each degree $d \in \Delta$, we consider the $(t_d \times s)$-Jacobian matrix

$$J_d(\mathbf{x}) := \begin{pmatrix} \nabla F_{d,1}(\mathbf{x}) \\ \vdots \\ \nabla F_{d,t_d}(\mathbf{x}) \end{pmatrix}$$

corresponding to the leading forms of degree d and the affine variety $S_d \subset \mathbb{A}^s_K$ defined by the condition $\operatorname{rank}(J_d(\mathbf{x})) < t_d$. We define B_d to be the dimension of S_d, and $B_d := 0$ if $d \notin \Delta$. In the following, we always assume that $B_d < s$.

Let $\mathcal{D}_0 := 0$ and, for $1 \le d \le D$,

$$\mathcal{D}_d := t_1 + 2t_2 + \cdots + dt_d,$$

so that $\mathcal{D} := \mathcal{D}_D$ is the sum of the degrees of all our polynomials $G_{d,i}$. Write

$$s_d := \sum_{k=d}^{D} \frac{2^{k-1}(k-1)t_k}{s - B_k}. \tag{1.1}$$

Our main theorem is as follows.

THEOREM 7.1. *Assume that*

$$\mathcal{D}_d \left(\frac{2^{d-1}}{s - B_d} + s_{d+1} \right) + s_{d+1} + \sum_{j=d+1}^{D} s_j t_j < 1 \tag{1.2}$$

holds for all $d \in \Delta \cup \{0\}$. Then there is a positive δ, such that

$$N(P) = \mathfrak{S}\mathfrak{J} \cdot P^{n(s-\mathcal{D})} + O(P^{n(s-\mathcal{D})-\delta})$$

for $P \to \infty$. Here, $\mathfrak{S}$ is the usual singular series and $\mathfrak{J}$ is the usual singular integral. The implicit constant in the error term may depend on K, the polynomials $G_{d,i}$, $\mathfrak{n}$, the basis $\omega_1, \dots, \omega_n$, and the box $\mathcal{B}$, but not on P and the positive constant δ depends on K and the systems of forms.

More precisely, the positive constant δ in the exponent of the error term may be chosen depending only on n, T, and the margin by which the inequalities (1.2) are satisfied.

We describe later in this introduction what is meant by the usual singular series and the usual singular integral. Theorem 7.1 is a number field version of [**48**, Theorem 1.2]. One should note that our hypotheses (1.2) are exactly the ones used by Browning and Heath-Brown over $\mathbb{Q}$. In particular, they are independent of the degree n of K.

In the special case where all degrees are equal to D, our hypotheses (1.2) simplify to Skinner's condition

$$s - B_D > t_D(t_D + 1)(D-1)2^{D-1} \tag{1.3}$$

from [**230**], which is the same one as Birch's [**40**].

Our proof relies on Skinner's number field version of the Hardy-Littlewood circle method [**230**, **231**]. There are only very few other applications of the circle method over number fields whose results are even close to the best available results over $\mathbb{Q}$, and in particular do not depend on the degree n of K. See, for example, [**47**, **215**, **230**].

1.2. Consequences. Here we collect some consequences of Theorem 7.1. They are number field versions of the results stated in the introduction of [**48**]. We omit most of the proofs, since they are almost verbatim the same as in [**48**]. The following corollary provides simpler hypotheses that imply those of Theorem 7.1.

COROLLARY 7.2. *Let* $B_{\max} := \max\{B_d : d \in \Delta\}$ *and*

$$u_d := \sum_{k=d}^{D} 2^{k-1}(k-1)t_k \quad \text{for} \quad 1 \le d \le D+1,$$

$$s_0(d) := \mathcal{D}_d(2^{d-1} + u_{d+1}) + u_{d+1} + \sum_{j=d+1}^{D} u_j t_j$$

$$s_0 := \max\{s_0(d) : d \in \Delta \cup \{0\}\}.$$

Then the conclusion of Theorem 7.1 holds whenever $s > B_{\max} + s_0$.

Of course, the explicit bounds for s_0 in case of systems of quadratic forms and a form of higher degree, computed in [**48**, Corollary 1.4], and in case of systems consisting of two forms of differing degrees, computed in [**48**, Corollary 1.5] are still valid in our number field version. Moreover, [**48**, Theorem 1.6] provides us with the bounds

$$s_0 + T - 1 \le \mathcal{D}^2 2^{D-1} \le T^2 D^2 2^{D-1}$$

and

$$s_0 + T - 1 \le (\mathcal{D}-1)2^{\mathcal{D}}. \tag{1.4}$$

By the last inequality, the condition $s > B_{\max} + s_0$ in the corollary is implied by the original condition (1.3) of Birch and Skinner in case of a single form of degree D.

In the following, we will specialize our results to non-singular systems of leading forms. To the system of forms $F_{d,j}(\mathbf{x})$, $1 \le d \le D, 1 \le j \le t_d$, we associate the $(T \times s)$-Jacobian matrix $J(\mathbf{x})$ formed by the partial derivatives of all T forms $F_{d,j}$ with with respect to all s variables x_i. We call the system $(F_{d,j})_{d,j}$ of forms *non-singular*, if $\operatorname{rank} J(\mathbf{x}) = T$ for every nonzero $\mathbf{x} \in \overline{\mathbb{Q}}^s$ satisfying $F_{d,j}(\mathbf{x}) = 0$ for all d, j.

Following [**48**], we define two systems of forms $(F_{d,j})_{d,j}$ and $(\widetilde{F}_{d,j})_{d,j}$, with $F_{d,j}, \widetilde{F}_{d,j} \in \mathcal{O}_K[x_1, \dots, x_s]$ of degree d, to be *equivalent* if for every pair (d, j) we have

$$\widetilde{F}_{d,j} = F_{d,j} - \sum_{i<j} H_{d,i}F_{d,i} - \sum_{e<d}\sum_{i \le t_e} H_{e,i}F_{e,i}, \tag{1.5}$$

with forms $H_{e,i} \in \mathcal{O}_K[x_1, \dots, x_s]$ of degree $d - e$. Moreover, we define a system $(F_{d,j})_{d,j}$ of forms to be *optimal*, if for any choice of (d, i), the sub-system

$$\{F_{d,j} : j \ge i\} \cup \{F_{e,j} : d < e \le D, j \le t_e\}$$

is nonsingular. In [**48**, Section 3] it is shown that every nonsingular system of forms is equivalent to an optimal system, and that every optimal system $(F_{d,i})_{d,i}$ satisfies

$$B_d \le t_d + \cdots + t_D - 1 \quad \text{for all} \quad 1 \le d \le D,$$

and hence in particular $B_{\max} \le T - 1$.

Assume we are given a system of polynomials $(G_{d,j})_{d,j}$, with $G_{d,j} \in \mathcal{O}_K[x_1, \dots, x_s]$ of degree d, with leading forms $(F_{d,j})_{d,j}$, and a system of forms $(\widetilde{F}_{d,j})_{d,j}$ equivalent to $(F_{d,j})_{d,j}$. Applying the expression for $\widetilde{F}_{d,j}$ in (1.5) to the polynomials $G_{d,j}$ instead of the forms $F_{d,j}$, we can easily write down a system of polynomials $(\widetilde{G}_{d,j})_{d,j}$ with leading forms $(\widetilde{F}_{d,j})_{d,j}$, such that the $\widetilde{G}_{d,j}$ generate the same ideal of $\mathcal{O}_K[x_1, \dots, x_s]$ as the $G_{d,j}$. Hence, we may replace every system $(G_{d,j})_{d,j}$ with a non-singular system of leading forms $(F_{d,j})_{d,j}$ by a system $(\widetilde{G}_{d,j})_{d,j}$ with an optimal system of leading forms $(\widetilde{F}_{d,j})_{d,j}$, and having the same common zeros.

Together with (1.4), this allows us to deduce the following generalization of [**48**, Theorem 1.7].

THEOREM 7.3. *Suppose that our system of leading forms $(F_{d,j})_{d,j}$ is non-singular and satisfies $s > (\mathcal{D}-1)2^{\mathcal{D}}$. Then there is a positive δ such that*

$$N(P) = \mathfrak{S}\mathfrak{J} \cdot P^{n(s-\mathcal{D})} + O(P^{n(s-\mathcal{D})-\delta}).$$

Here, $\mathfrak{J} > 0$ whenever the system of equations

$$F_{d,j}(\mathbf{x}) = 0 \quad \text{for} \quad 1 \le d \le D, 1 \le j \le t_d$$

has a nonzero solution in the interior of $\mathcal{B} \subset V$. Moreover, $\mathfrak{S} > 0$ whenever the system of equations

$$G_{d,j}(\mathbf{x}) = 0 \quad \text{for} \quad 1 \le d \le D, 1 \le j \le t_d$$

has a nonzero solution in the completion $n_{\mathfrak{p}}^s \subset (\mathcal{O}_K)_{\mathfrak{p}}^s$ for every prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$.

The conditions under which $\mathfrak{J} > 0$ and $\mathfrak{S} > 0$ follow from well known facts about $\mathfrak{J}$ and $\mathfrak{S}$ and from the fact that, under the hypotheses of Theorem 7.3, the system of leading forms $(F_{d,j})_{d,j}$ defines a smooth complete intersection (see [**48**, Lemma 3.2]). In particular, the singular series $\mathfrak{S}$ has the usual interpretation as a product of local densities.

Theorem 7.3 has far-reaching consequences for smooth projective complete intersections. In fact, every smooth complete intersection $X \subseteq \mathbb{P}_K^{s-1}$ is defined by a non-singular system of forms $(F_{d,i})_{d,i}$, and if X is non-degenerate (not contained in a proper linear subspace of $\mathbb{P}_K^{s-1}$), then $\deg(X) \ge \mathcal{D}$.

It is known that an asymptotic formula as in Theorem 7.3 implies the Hasse principle and weak approximation for X, see [**230**]. We can also say something about the Manin-Peyre conjecture [**83**, **183**]. Let Ω_K be the set of all places of K, and for each $v \in \Omega_K$ let $n_v := [K_v : \mathbb{Q}_v]$ be the local degree at v. Let $|\cdot|_v$ be any norm on K_v^s, coinciding with the usual max-norm if v is non-archimedean. Then

$$H((x_1 : \cdots : x_s)) := \prod_{v \in \Omega_K} |(x_1, \ldots, x_s)|_v^{n_v(s-\mathcal{D})} \tag{1.6}$$

defines an anticanonical height function on the rational points $X(K)$. The proof of [**139**, Theorem 4.8] shows that the conclusion of Theorem 7.3 implies the Manin-Peyre conjecture for X with respect to the height H.

Thus, every smooth and non-degenerate complete intersection $X \subseteq \mathbb{P}_K^{s-1}$ with

$$s > (\deg(X) - 1)2^{\deg X}$$

satisfies the Hasse principle, weak approximation, and the Manin-Peyre conjecture for the anticanonical heights defined above.

Browning and Heath-Brown show in [**48**] that every smooth and geometrically integral variety $X \subseteq \mathbb{P}_{\mathbb{Q}}^m$ satisfying

$$\dim(X) \ge (\deg(X) - 1)2^{\deg(X)} - 1 \tag{1.7}$$

is already a complete intersection. Their arguments hold as well over arbitrary number fields, which provides us with the following nice consequence of Theorem 7.3, generalizing [**48**, Theorem 1.1].

THEOREM 7.4. *Let $X \subseteq \mathbb{P}^m_K$ be a smooth and geometrically integral variety satisfying* (1.7). *Then X satisfies the Hasse principle, weak approximation, and the Manin-Peyre conjecture with respect to the height functions defined in* (1.6).

1.3. The circle method over number fields. Our proof of Theorem 7.1 relies on the Hardy-Littlewood circle method, implemented over the number field K by Skinner [**230**]. We start by reviewing some notation from [**230**].

Let $\Omega_K, \Omega_\infty, \Omega_0$ denote the sets of all places, archimedean places, and non-archimedean places of K, and write K_v for the completion of K at the place v. If $v \in \Omega_\infty$ then we identify K_v with the field $\mathbb{R}$ or $\mathbb{C}$ in the usual way.

We identify $V = K \otimes_{\mathbb{Q}} \mathbb{R}$ with $\prod_{v\in\Omega_\infty} K_v$. This allows us to naturally define the conjugates $x^{(v)} \in K_v$ of $x \in V$ via projection and to extend the norm and trace of K to functions $N : V \to \mathbb{R}$, $\mathrm{Tr} : V \to \mathbb{R}$. On V, we moreover have an $\mathbb{R}$-vector norm defined by

$$|x| := \max\{|x_1|, \dots, |x_n|\} \quad \text{for} \quad x = x_1\omega_1 + \cdots + x_n\omega_n$$

that satisfies $|x| \asymp \max_{v\in\Omega_\infty}\{|x|_v\}$, where $|x|_v$ is the usual absolute value on $K_v \in \{\mathbb{R}, \mathbb{C}\}$. We extend the norm to V^s via $|\mathbf{x}| := \max_{j=1,\dots,s}\{|x_j|\}$ for $\mathbf{x} = (x_1, \dots, x_s)$.

Let

$$R := \{x_1\omega_1 + \cdots + x_n\omega_n \,:\, x_i \in [0, 1)\} \subset V.$$

We normalize the Haar measure on V by $\mathrm{vol}(R) = 1$. Elements of $V^T = \prod_{d=1}^D V^{t_d}$ are written with double indices $\boldsymbol{\alpha} = (\alpha_{d,i})_{\substack{1\le d\le D\\ 1\le i\le t_d}}$. We write $e(x) = e^{2\pi i x}$ for $x \in \mathbb{R}$ and $\Phi(x) = e(\mathrm{Tr}(x))$ for $x \in V$. The circle method is based on the identity

$$N(P) = \int_{\boldsymbol{\alpha}\in R^T} S(\boldsymbol{\alpha})\, \mathrm{d}\boldsymbol{\alpha}, \tag{1.8}$$

where

$$S(\boldsymbol{\alpha}) := \sum_{\mathbf{x}\in\mathfrak{n}^s\cap P\mathcal{B}} \Phi\left(\sum_{d=1}^{D}\sum_{i=1}^{t_D} \alpha_{d,i}G_{d,i}(\mathbf{x})\right). \tag{1.9}$$

We divide R^T into major and minor arcs as follows. Let $\varpi \in (0, 1/3)$ be a fixed constant to be specified in Section 5. For $\gamma \in K$, we have the denominator ideal $\mathfrak{a}_\gamma := \{\beta \in \mathcal{O}_K \,:\, \beta\gamma \in \mathfrak{n}\}$. For $\boldsymbol{\gamma} = (\gamma_{d,i})_{d,i} \in (R \cap K)^T$, let $\mathfrak{a}_{\boldsymbol{\gamma}} := \bigcap_{d,i} \mathfrak{a}_{\gamma_{d,i}}$. The major arc corresponding to $\boldsymbol{\gamma}$ is

$$\mathfrak{M}_{\boldsymbol{\gamma}} := \{\boldsymbol{\alpha} \in R^T \,:\, |\alpha_{d,i} - \gamma_{d,i}| \le P^{-d+\varpi} \text{ for all } 1 \le d \le D, 1 \le i \le t_d\},$$

where the distance $|\alpha_{d,i} - \gamma_{d,i}|$ is to be understood modulo $\mathfrak{n}$. We define the major arcs

$$\mathfrak{M} := \bigcup_{\substack{\boldsymbol{\gamma}\in(R\cap K)^T\\ \mathfrak{N}\mathfrak{a}_{\boldsymbol{\gamma}}\le P^{\varpi}}} \mathfrak{M}_{\boldsymbol{\gamma}}$$

and the minor arcs

$$\mathfrak{m} := R^T \smallsetminus \mathfrak{M}.$$

In Section 4, we show that, under the hypotheses of Theorem 7.1, the contribution of the minor arcs $\mathfrak{m}$ to the integral in (1.8) is absorbed by the error term. In Sections 5 and 6,

we evaluate the contribution of the major arcs $\mathfrak{M}$ as $\mathfrak{S}\mathfrak{J}P^{n(s-\mathcal{D})}+O(P^{n(s-\mathcal{D})-\delta})$, with the singular series

$$\mathfrak{S}=\prod_{\mathfrak{p}}\sum_{j=0}^{\infty}\frac{1}{\mathfrak{N}\mathfrak{p}^{js}}\sum_{\substack{\boldsymbol{\gamma}\in(R\cap K)^T\\ \mathfrak{a}_{\boldsymbol{\gamma}}=\mathfrak{p}^j}}\sum_{\mathbf{x}\in(\mathfrak{n}/\mathfrak{p}^j\mathfrak{n})^s}\Phi\left(\sum_{d=1}^{D}\sum_{i=1}^{t_d}\gamma_{d,i}G_{d,i}(\mathbf{x})\right)$$

and the singular integral

$$\mathfrak{J}=\int_{\boldsymbol{\gamma}\in V^T}\int_{\mathbf{x}\in\mathcal{B}}\Phi\left(\sum_{d=1}^{D}\sum_{i=1}^{t_d}\gamma_{d,i}F_{d,i}(\mathbf{x})\right)\mathrm{d}\mathbf{x}\,\mathrm{d}\boldsymbol{\gamma}.$$

In Sections 2 and 3, we prove the main tool to be used in our estimations, an iterative Weyl-type lemma for the exponential sum $S(\boldsymbol{\alpha})$ that generalizes the version from [**48**] to number fields.

1.4. Further notation. It is sometimes useful to identify V with $\mathbb{R}^n$ via the basis $\omega_1,\dots,\omega_n$. Then $\mathbf{x}\in V^s$ with $x_i=x_{i,1}\omega_1+\cdots+x_{i,n}\omega_n$ is identified with $\mathbf{X}=(x_{i,j})_{i,j}\in\mathbb{R}^{ns}$. The volume on V becomes the usual Lebesgue measure on $\mathbb{R}^n$, and the norm $|\cdot|$ becomes the usual max-norm on $\mathbb{R}^n$, which we will also denote by $|\cdot|$. To each polynomial $G\in V[x_1,\dots,x_s]$, we associate the polynomial

$$G^*(\mathbf{X}):=\mathrm{Tr}(G(\mathbf{x}))\in\mathbb{R}[\mathbf{X}]$$

and the system $G_j^*(\mathbf{X})$, $1\le j\le n$, defined via

$$G_j^*(\mathbf{X}):=\mathrm{Tr}(\omega_jG(\mathbf{x}))\in\mathbb{R}[\mathbf{X}].$$

Then any $\mathbf{x}$ in V^s satisfies $G_{d,i}(\mathbf{x})=0$ for all d,i if and only if $G_{d,i,j}^*(\mathbf{X})=0$ for all d,i,j, and our system of T polynomials in s variables over $\mathcal{O}_K$ is equivalent to a system of nT polynomials in ns variables over $\mathbb{Z}$. In fact, the affine variety defined over $\mathbb{Q}$ by the polynomials $G_{d,i,j}^*(\mathbf{X})$ is the Weil restriction of the K-variety defined by the polynomials $G_{d,i}(\mathbf{x})$. These identifications allow us to write $S(\boldsymbol{\alpha})$ as an exponential sum over $\mathbb{Z}^{ns}$:

$$S(\boldsymbol{\alpha})=\sum_{\mathbf{X}\in\mathbb{Z}^{ns}\cap PB}e\left(\sum_{d=1}^{D}\sum_{i=1}^{t_d}\sum_{j=1}^{n}\alpha_{d,i,j}G_{d,i,j}^*(\mathbf{X})\right),$$

where $\alpha_{d,i}=\alpha_{d,i,1}\omega_1+\cdots+\alpha_{d,i,n}\omega_n$.

We denote the standard basis of the free V-module V^s by $\mathbf{v}_1,\dots,\mathbf{v}_s$, and the standard basis of $\mathbb{R}^{ns}$ by $\mathbf{E}_{ij}$ ($i=1,\dots,s$ and $j=1,\dots,n$). By our identification, we obtain $\mathbf{E}_{ij}=\omega_j\mathbf{v}_i$.

For $\beta\in\mathbb{R}$, we write $\|\beta\|$ for the distance of β to the nearest integer.

For any form $F\in V[x_1,\dots,x_s]$ of degree d, we write $F(\mathbf{x}_1|\cdots|\mathbf{x}_d)$ for the corresponding polar d-multilinear form, normalized by $d!F(\mathbf{x})=F(\mathbf{x}|\cdots|\mathbf{x})$. Similarly, $F^*(\mathbf{X}_1|\cdots|\mathbf{X}_d)$ is the polar d-multilinear form corresponding to F^*. Observe that $F^*(\mathbf{X}_1|\cdots|\mathbf{X}_d)=\mathrm{Tr}(F(\mathbf{x}_1|\cdots|\mathbf{x}_d))$.

1.5. The singular integral. Our main task in Section 6 is to show that the integral

$$\mathfrak{J}(H):=\int_{\substack{\boldsymbol{\gamma}\in V^T\\ |\boldsymbol{\gamma}|\le H}}\int_{\mathbf{x}\in\mathcal{B}}\Phi\left(\sum_{d=1}^{D}\sum_{i=1}^{t_d}\gamma_{d,i}F_{d,i}(\mathbf{x})\right)\mathrm{d}\mathbf{x}\,\mathrm{d}\boldsymbol{\gamma}\tag{1.10}$$

converges absolutely as $H\to\infty$, and that

$$\mathfrak{J}(H)-\mathfrak{J}\ll H^{-\delta}$$

for some positive δ, see Lemma 7.28. In the case where all degrees are equal, i.e. $t_d = 0$ for all $d \neq D$, this is done by Skinner in [**230**, Lemma 9]. For the proof, Skinner suggests to think of our forms $F_{D,i}$ as the forms $F^*_{D,i,j}$ over $\mathbb{Z}$ and to apply Schmidt's Lemma 8.1 from [**220**]. Schmidt's lemma is a formalization of the classical indirect approach to the singular integral, already used by Birch [**40**], where the Weyl-type lemma used in the treatment of the minor arcs is applied once more to bound the inner integral in (1.10). Hence, it depends on a certain hypothesis, called the *restricted hypothesis* by Schmidt. Applied to our forms $F^*_{D,i,j}$, it requires that at least one of the following alternatives hold for some $\Omega > nt_D + 1$ and each $\Delta \in (0, 1]$: Every $\boldsymbol{\alpha} \in R^{t_D}$ satisfies

(i) $|S(\boldsymbol{\alpha})| \leq P^{ns-\Delta\Omega}$, or

(ii) there is $q \in \mathbb{N}$, $q \leq P^{\Delta}$ with $\|q\alpha_{D,i,j}\| \leq P^{-D+\Delta}$ for all $1 \leq i \leq t_D$ and $1 \leq j \leq n$.

Skinner gives no argument why this hypothesis would hold. In fact, we were not able to deduce it from either Skinner's Weyl-type lemma [**230**, Lemma 2], or Birch's Weyl-type lemma [**40**, Lemma 2.5] applied to the $F^*_{D,i,j}$, without replacing the lower bound (1.3) on the number of variables s by the stronger bound

$$s - B_D > t_D(nt_D + 1)(D - 1)2^{D-1}, \tag{1.11}$$

and we see no reason why it should hold otherwise. Let us note that with the stronger assumption (1.11) instead of (1.3), the main theorem of [**230**] would follow directly from the techniques of [**40**] applied to the $G^*_{D,i,j}$.

In Section 6, we apply genuine number field arguments to our treatment of the singular integral, culminating in Lemma 7.28. When all degrees are equal to D, the hypothesis (5.6) of Lemma 7.28 is exactly Skinner's hypothesis (1.3), so we prove [**230**, Lemma 9] as a special case of Lemma 7.28.

2. Exponential sums

For a function $f : V^s \to \mathbb{R}$ and $\mathbf{h} \in V^s$, we write $\Delta_{\mathbf{h}}(f)(\mathbf{x}) := f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})$ for the usual forward difference operator. The following lemma is a number field analog of van der Corput's variant of Weyl differencing.

LEMMA 7.5. *Let $q \in \mathcal{O}_K \setminus \{0\}$ and $H \in \mathbb{Z}$ with $1 \leq H \ll P/|q|$. Let $f : V^s \to \mathbb{R}$ and $\mathcal{I}$ be a box aligned to the basis $\omega_1, \dots, \omega_n$. Then*

$$\left| \sum_{\mathbf{x} \in \mathfrak{n}^s \cap P\mathcal{I}} \Phi(f(\mathbf{x})) \right|^2 \ll \left(\frac{P}{H}\right)^{ns} \sum_{\substack{\mathbf{h} \in \mathfrak{n}^s \\ |\mathbf{h}| < H}} \left| \sum_{\mathbf{x} \in \mathfrak{n}^s \cap P\mathcal{I}'(\mathbf{h})} \Phi(\Delta_{q\mathbf{h}}(f)(\mathbf{x})) \right|,$$

where $\mathcal{I}'(\mathbf{h})$ is a box aligned to the basis $\omega_1, \dots, \omega_n$ that depends on $\mathbf{h}$. The implicit constant depends only on K and s.

DÉMONSTRATION. The proof is analogous to the version over $\mathbb{Q}$ (see the proof of [**48**, Lemma 4.1]). Let $\chi_{P\mathcal{I}}$ be the indicator function of $P\mathcal{I}$. Let $R^* \subset V$ be the set of all $\mathbf{u} =$

$u_1\omega_1 + \cdots + u_n\omega_n \in V$ with $1 \le u_j \le H$ for all $1 \le j \le n$. Then

$$\begin{aligned} H^{ns} \sum_{\mathbf{x}\in\mathfrak{n}^s\cap P\mathcal{I}} \Phi(f(\mathbf{x})) &= \sum_{\mathbf{u}\in(\mathfrak{n}\cap R^*)^s} \sum_{\mathbf{x}\in\mathfrak{n}^s} \Phi(f(\mathbf{x}+q\mathbf{u}))\chi_{P\mathcal{I}}(\mathbf{x}+q\mathbf{u}) \\ &= \sum_{\substack{\mathbf{x}\in\mathfrak{n}^s \\ |\mathbf{x}|\ll P}} \sum_{\mathbf{u}\in(\mathfrak{n}\cap R^*)^s} \Phi(f(\mathbf{x}+q\mathbf{u})\chi_{P\mathcal{I}}(\mathbf{x}+q\mathbf{u}). \end{aligned}$$

Here, we used that $\chi_{P\mathcal{I}}(\mathbf{x}+q\mathbf{u}) \neq 0$ implies $|\mathbf{x}| \le P + |q\mathbf{u}| \ll P + |q|\,|\mathbf{u}| \ll P$. By Cauchy's inequality,

$$\begin{aligned} H^{2ns}\left|\sum_{\mathbf{x}\in\mathfrak{n}^s\cap P\mathcal{I}} \Phi(f(\mathbf{x}))\right|^2 &\ll P^{ns} \sum_{\substack{\mathbf{x}\in\mathfrak{n}^s \\ |\mathbf{x}|\ll P}} \left|\sum_{\mathbf{u}\in(\mathfrak{n}\cap R^*)^s} \Phi(f(\mathbf{x}+q\mathbf{u}))\chi_{P\mathcal{I}}(\mathbf{x}+q\mathbf{u})\right|^2 \\ &= P^{ns} \sum_{\substack{\mathbf{h}\in\mathfrak{n}^s \\ |\mathbf{h}|<H}} n(\mathbf{h}) \sum_{y\in\mathfrak{n}^s} \Phi(f(\mathbf{y}+q\mathbf{h}))\chi_{P\mathcal{I}}(\mathbf{y}+q\mathbf{h})\overline{\Phi(f(\mathbf{y}))}\chi_{P\mathcal{I}}(\mathbf{y}), \end{aligned}$$

where

$$n(\mathbf{h}) = \#\left\{(\mathbf{u},\mathbf{v}) \in (\mathfrak{n}\cap R^*)^{2s} \,:\, \mathbf{h} = \mathbf{v}-\mathbf{u}\right\} \le H^{ns}.$$

Thus,

$$\left|\sum_{\mathbf{x}\in\mathfrak{n}^s\cap P\mathcal{I}} \Phi(f(\mathbf{x}))\right|^2 \ll \left(\frac{P}{H}\right)^{ns} \sum_{\substack{\mathbf{h}\in\mathfrak{n}^s \\ |\mathbf{h}|<H}} \left|\sum_{\mathbf{y}\in\mathfrak{n}^s\cap P\mathcal{I}'(\mathbf{h})} \Phi(f(\mathbf{y}+q\mathbf{h}))\overline{\Phi(f(\mathbf{h}))}\right|,$$

where $\mathcal{I}'(\mathbf{h}) \subseteq \mathcal{I}$ is a box aligned to the basis $\omega_1,\dots,\omega_n$, depending on $\mathbf{h}$. □

Let $f, g \in V[x_1,\dots,x_s]$ with $\deg f \le d$. Suppose that $qg = g_1 + g_2$ with $q \in \mathfrak{n}\setminus\{0\}$, $g_1 \in \mathcal{O}_K[x_1,\dots,x_s]$ and $g_2 \in V[x_1,\dots,x_s]$ such that all coefficients a_j of g_2 of each degree j satisfy

$$|a_j| \ll_j \varphi P^{-j}, \tag{2.1}$$

for some $\varphi \ge |q|$.

Let $\mathcal{B}'$ be a box aligned to the basis $\omega_1,\dots,\omega_n$. We are interested in the estimation of the exponential sum

$$\Sigma := \sum_{\mathbf{x}\in\mathfrak{n}^s\cap P\mathcal{B}'} \Phi(f(\mathbf{x}) + g(\mathbf{x})).$$

In the process, f will be the dominant polynomial whereas g originates from the higher exponents which are already well approximable. We aim for an estimate of the form

$$|\Sigma| = P^{ns}L$$

with small L. Let F be the homogeneous part of f of degree d, and recall that $(qF)^*(\mathbf{X}_1|\cdots|\mathbf{X}_d)$ is the d-multilinear polar form corresponding to the form $(qF)^*(\mathbf{X})$.

Lemma 7.6. *For $M \ge 1$, we have*

$$L^{2^{d-1}} \ll P^{-(d-1)ns}\,(\varphi M)^{(d-1)ns}\,(\log P)^{ns}\mathcal{M}, \tag{2.2}$$

where $\mathcal{M}$ *is the number of all* $(\mathbf{x}_1, \dots, \mathbf{x}_{d-1}) \in (\mathfrak{n}^s)^{d-1}$ *satisfying*

$$|\mathbf{x}_i| \le \frac{P}{\varphi M} \qquad \text{for all } 1 \le i < d, \text{ and}$$

$$\|(qF)^*(\mathbf{X}_1|\cdots|\mathbf{X}_{d-1}|\mathbf{E}_{i,j})\| \le \frac{1}{P\varphi^{d-2}M^{d-1}} \qquad \text{for all } 1 \le i \le s, 1 \le j \le n.$$

Démonstration. This is mostly analogous to the proof of [**48**, Lemma 4.1]. The lemma holds trivially if $\varphi \ge P$. Hence, we assume that $\varphi \le P$.

We start by $d-2$ Weyl differencing steps, that is $d-2$ applications of Lemma 7.5 with $q := 1$, $H := P$, linked by Cauchy's inequality). This yields

$$L^{2^{d-2}} \ll P^{-ns(d-1)} \sum_{\substack{\mathbf{h}_1 \in \mathfrak{n}^s \\ |\mathbf{h}_1|<P}} \cdots \sum_{\substack{\mathbf{h}_{d-2} \in \mathfrak{n}^s \\ |\mathbf{h}_{d-2}|<P}} \left| \sum_{\mathbf{x} \in \mathfrak{n}^s \cap P\mathcal{I}} \Psi(\mathbf{x}) \right|, \tag{2.3}$$

where

$$\Psi(\mathbf{x}) := \Phi\left(\Delta_{\mathbf{h}_1,\dots,\mathbf{h}_{d-2}}(f+g)(\mathbf{x})\right)$$

and $\mathcal{I} \subset \mathcal{B}'$ is a box aligned to the basis $\omega_1, \dots, \omega_n$ that depends on $\mathbf{h}_1, \dots, \mathbf{h}_{d-2}$.

For the $(d-1)$-st differencing step, we choose q as in the setup before our Lemma and

$$H := \left\lfloor \frac{P}{\varphi} \right\rfloor \ge 1.$$

By Lemma 7.5,

$$\left| \sum_{\mathbf{x} \in \mathfrak{n}^s \cap P\mathcal{I}} \Psi(\mathbf{x}) \right|^2 \ll \varphi^{ns} \sum_{\substack{\mathbf{w} \in \mathfrak{n}^s \\ |\mathbf{w}|<H}} \left| \sum_{\mathbf{y} \in \mathfrak{n}^s \cap P\mathcal{I}'} \Psi(\mathbf{y}+q\mathbf{w})\overline{\Psi(\mathbf{y})} \right|,$$

where $\mathcal{I}' \subseteq \mathcal{I}$ is a box aligned to the basis $\omega_1, \dots, \omega_n$, depending on $\mathbf{w}$. Together with Cauchy's inequality and (2.3), this yields

$$L^{2^{d-1}} \ll P^{-nsd}\varphi^{ns} \sum_{\substack{\mathbf{h}_1 \in \mathfrak{n}^s \\ |\mathbf{h}_1|<P}} \cdots \sum_{\substack{\mathbf{h}_{d-2} \in \mathfrak{n}^s \\ |\mathbf{h}_{d-2}|<P}} \sum_{\substack{\mathbf{w} \in \mathfrak{n}^s \\ |\mathbf{w}|<H}} \left| \sum_{\mathbf{y} \in \mathfrak{n}^s \cap P\mathcal{I}'} \Psi(\mathbf{y}+q\mathbf{w})\overline{\Psi(\mathbf{y})} \right|.$$

Note that Ψ implicitly depends on $\mathbf{h}_1, \dots, \mathbf{h}_{d-2}$, and $\mathcal{I}'$ depends on $\mathbf{h}_1, \dots, \mathbf{h}_{d-2}$ and $\mathbf{w}$.

Now we take a closer look at the summands of the innermost sum :

$$\Psi(\mathbf{y}+q\mathbf{w})\overline{\Psi(\mathbf{y})} = \Phi(\Delta_{\mathbf{h}_1,\dots,\mathbf{h}_{d-2},q\mathbf{w}}(f+g)(\mathbf{y})).$$

Recall that F is the leading form of f of degree d. Then by linearity of $F(\mathbf{x}_1|\cdots|\mathbf{x}_d)$ we have that

$$\Delta_{\mathbf{h}_1,\dots,\mathbf{h}_{d-2},q\mathbf{w}}(f)(\mathbf{y}) = qF(\mathbf{h}_1|\cdots|\mathbf{h}_{d-2}|\mathbf{w}|\mathbf{y}) + C.$$

Now we focus on g. Since all coefficients of $\Delta_{\mathbf{h}_1,\dots,\mathbf{h}_{d-2},q\mathbf{w}}(g_1)$ are $\mathcal{O}_K$-multiples of q, we have

$$\Phi(\Delta_{\mathbf{h}_1,\dots,\mathbf{h}_{d-2},q\mathbf{w}}(q^{-1}g_1)(\mathbf{y})) = 1$$

for all $\mathbf{y} \in \mathcal{O}_K$. Since $|\mathbf{h}_i| \le P$ for all $1 \le i \le d-2$, it follows from (2.1) that each coefficient b_j of $\Delta_{\mathbf{h}_1,\dots,\mathbf{h}_{d-2}}(g_2)$ of any degree j satisfies

$$|b_j| \ll_j |\mathbf{h}_1| \cdots |\mathbf{h}_{d-2}|\, \varphi P^{-j-(d-2)}.$$

Since $|q\mathbf{w}| \ll |q|\, H \le \varphi H \ll P$, the coefficients c_j of $\Delta_{\mathbf{h}_1,\dots,\mathbf{h}_{d-2},\text{`}\mathbf{w}'}(q^{-1}g_2)$ of degree j are bounded by

$$|c_j| \ll \left|q^{-1}b_{j+1}qw_\ell\right| \ll |b_{j+1}|\,|\mathbf{w}| \ll_j |\mathbf{h}_1|\cdots|\mathbf{h}_{d-2}|\,\varphi P^{-j-(d-1)}H \ll P^{-j},$$

where b_{j+1} a coefficient of $\Delta_{\mathbf{h}_1,\dots,\mathbf{h}_{d-2}}(g_2)$ of degree $j+1$ and w_ℓ is a component on $\mathbf{w}$.

Write $\mathbf{u} := q\mathbf{w}$. With the tuples $\mathbf{H}_i, \mathbf{U} \in \mathbb{Z}^{sn}$ corresponding to $\mathbf{h}_i, \mathbf{u} \in \mathfrak{n}^s$, we have shown that any j-th order partial derivative of $\Delta_{\mathbf{H}_1,\dots,\mathbf{H}_{d-2},\mathbf{U}}((q^{-1}g_2)^*)$ is $\ll_j P^{-j}$ uniformly on $[-P,P]^{ns}$.

Let $I' \subseteq [-1,1]^{sn}$ be the box in $\mathbb{R}^{sn}$ corresponding to $\mathcal{I}'$. Using the above computations,

$$\sum_{\mathbf{y}\in\mathfrak{n}^s\cap P\mathcal{I}'} \Psi(\mathbf{y}+q\mathbf{w})\overline{\Psi(\mathbf{y})}$$
$$= \sum_{\mathbf{Y}\in\mathbb{Z}^{ns}\cap PI'} e((qF)^*(\mathbf{H}_1|\cdots|\mathbf{H}_{d-2}|\mathbf{W}|\mathbf{Y}) + \mathrm{Tr}(C) + \Delta_{\mathbf{H}_1,\dots,\mathbf{H}_{d-2},\mathbf{U}}((q^{-1}g_2)^*)(\mathbf{Y})).$$

In the same manner as Browning and Heath-Brown, we apply multidimensional partial summation and our uniform bounds for the partial derivatives of $\Delta_{\mathbf{H}_1,\dots,\mathbf{H}_{d-2},\mathbf{U}}((q^{-1}g_2)^*)$ to obtain

$$\sum_{\mathbf{y}\in\mathfrak{n}^s\cap\mathcal{I}'} \Psi(\mathbf{y}+q\mathbf{w})\overline{\Psi(\mathbf{y})} \ll \left|\sum_{\mathbf{Y}\in\mathbb{Z}^{ns}\cap I''} e\left((qF)^*(\mathbf{H}_1|\dots|\mathbf{H}_{d-2}|\mathbf{W}|\mathbf{Y})\right)\right|,$$

with a box $I'' \subseteq I'$ aligned to the coordinate axes. We are now in exactly the same situation as in the proof of [**48**, Lemma 4.1], just in Dimension sn instead of n and with φ instead of $q\varphi$. What remains of the proof is identical to the arguments of [**48**] starting at (4.5), just with tuples $\mathbf{Y}, \mathbf{H}_j, \mathbf{W}$ of sn variables instead of tuples $\mathbf{y}, \mathbf{x}_j, \mathbf{w}$ of n variables. □

3. The iterative argument

Our aim in this section is to find a Weyl-type estimate for the exponential sum $S(\alpha)$ defined in (1.9). To this end, we write

$$|S(\alpha)| = P^{ns}L.$$

We define $Q_{D+1} := 1$ and, for $d \in \Delta$,

$$Q_d := (\log P)^{e(d)}L^{-s_d/n}, \tag{3.1}$$

where $e(d)$ is an explicit but irrelevant exponent which could be computed from the arguments in the proof of Lemma 7.7. For those $1 \le d \le D$ with $d \notin \Delta$ we set $Q_d := Q_k$, where k is the smallest integer bigger than d in Δ. Similarly, we can extend the definition of the exponents $e(d)$ to these values.

For $j \in \Delta$, we consider upper bounds

$$L^{2^{j-1}} \le \left(\frac{Q_{j+1}}{P}\right)^{n(s-B_j)} (\log P)^{ns+1}. \tag{3.2}$$

Let $I_d^{(1)}$ be the set of all $\boldsymbol{\alpha} \in R^T$ such that (3.2) holds for $j = d$ but fails for every $j > d$. Moreover, let $I^{(2)}$ be the set of all $\boldsymbol{\alpha} \in R^T$ such that (3.2) fails for all $j \in \Delta$. We are going to prove the following number field analogue of [**48**, Lemma 6.2].

LEMMA 7.7. *Let $d \in \Delta$ and $P \gg 1$. If $\boldsymbol{\alpha} \in I_d^{(1)}$ then*

$$L^{2^{d-1}+(s-B_d)s_{d+1}} \ll P^{-n(s-B_d)+\epsilon}. \tag{3.3}$$

Moreover, there are $q_j \in \mathfrak{n}$, $\boldsymbol{\nu}_j \in \mathfrak{n}^{t_j}$ for all $d < j \leq D, j \in \Delta$ satisfying

$$q_k \mid q_j \qquad \text{for all } k > j, k \in \Delta \tag{3.4}$$

$$|q_j| \leq Q_j \tag{3.5}$$

$$|q_j \alpha_{j,i} - \nu_{j,i}| \leq Q_j P^{-j} \qquad \text{for all } 1 \leq i \leq t_j. \tag{3.6}$$

If $\boldsymbol{\alpha} \in I^{(2)}$ then there are $q \in \mathfrak{n}$, $\boldsymbol{\nu}_j \in \mathfrak{n}^{t_j}$ for all $j \in \Delta$, satisfying (3.4), (3.5), *and* (3.6).

The idea is to iteratively apply Lemma 7.6. Recall that $\mathbf{v}_j$ denotes the j-th element of the standard basis of V^s. For $d \leq D$, we define the matrix

$$\widehat{J}_d(\mathbf{x}_1, \dots, \mathbf{x}_{d-1}) := (F_{d,i}(\mathbf{x}_1 | \cdots | \mathbf{x}_{d-1} | \mathbf{v}_j))_{\substack{1 \leq i \leq t_d \\ 1 \leq j \leq s}}$$

and the corresponding affine variety $\widehat{S}_d \subseteq (\mathbb{A}_K^s)^{d-1}$ defined by the condition

$$\operatorname{rank}(\widehat{J}_d(\mathbf{x}_1, \dots, \mathbf{x}_{d-1})) < t_d.$$

We need an estimate for the number of integral points on $\widehat{S}_d$ of bounded norm. Let

$$\mathcal{M}_0(P) := \#\{(\mathbf{x}_1, \dots, \mathbf{x}_{d-1}) \in \widehat{S}_d(K) \cap (\mathfrak{n}^s)^{d-1} \,:\, |\mathbf{x}_i| \leq P \text{ for all } 1 \leq i < d\}.$$

LEMMA 7.8. *For $P \geq 1$, we have*

$$\mathcal{M}_0(P) \ll P^{n(B_d + s(d-2))}.$$

DÉMONSTRATION. As in [**48**, Lemma 5.1], using [**230**, Lemma 3] instead of [**40**, Lemma 3.1]. □

The main tool for the proof of Lemma 7.7 is the following iterative argument.

LEMMA 7.9. *Let $|S(\boldsymbol{\alpha})| = P^{ns}L$ and $d \in \Delta$. Furthermore suppose that*

- *either $d = D$, $q = 1$ and $Q = 1$,*
- *or $d < D$, and there exist $Q \geq 1$ and $q \in \mathfrak{n}$ with $|q| \leq Q$, and $\boldsymbol{\nu}_j \in \mathfrak{n}^{t_j}$, such that*

$$|q\alpha_{j,i} - \nu_{j,i}| \leq QP^{-j} \quad \text{for } d < j \leq D \text{ and } 1 \leq i \leq t_j.$$

Then, for P sufficiently large, either

$$L^{2^{d-1}} \leq \left(\frac{Q}{P}\right)^{n(s-B_d)} (\log P)^{ns+1},$$

or there exists $q^* \in \mathfrak{n}$ with

$$|q^*| \leq Q^* := \left(\frac{(\log P)^{ns+1}}{L^{2^{d-1}}}\right)^{\frac{t_d(d-1)}{n(s-B_d)}} (\log P)$$

and $\boldsymbol{\nu}_d \in \mathfrak{n}^{t_d}$, such that

$$|q^* q \alpha_{d,i} - \nu_{d,i}| \leq QQ^* P^{-d} \text{ for } 1 \leq i \leq t_d.$$

DÉMONSTRATION. The key tool is Lemma 7.6. We distinguish the two cases $d = D$ and $d < D$:

— $\mathbf{d = D}$: In this case, we choose $\varphi := 1$, $g = g_1 = g_2 := 0$, and

$$f(\mathbf{x}) := \sum_{j=1}^{D} \sum_{i=1}^{t_j} \alpha_{j,i} G_{j,i}(\mathbf{x})$$

— $\mathbf{d < D}$: Then we let

$$f(\mathbf{x}) := \sum_{j=1}^{d} \sum_{i=1}^{t_j} \alpha_{j,i} G_{j,i}(\mathbf{x}) \quad \text{and} \quad g(\mathbf{x}) := \sum_{j=d+1}^{D} \sum_{i=1}^{t_j} \alpha_{j,i} G_{j,i}(\mathbf{x}).$$

By hypothesis, we have $q\alpha_{j,i} = \nu_{j,i} + \theta_{j,i}$ for $d < j \le D$ and $1 \le i \le t_j$, with $|\theta_{j,i}| \le QP^{-j}$. We write $qg = g_1 + g_2$, where

$$g_1 := \sum_{j=d+1}^{D} \sum_{i=1}^{t_j} \nu_{j,i} G_{j,i}(\mathbf{x}) \quad \text{and} \quad g_2 := \sum_{j=d+1}^{D} \sum_{i=1}^{t_j} \theta_{j,i} G_{j,i}(\mathbf{x}).$$

This allows us to choose $\varphi := Q$.

Now we apply Lemma 7.6 with $M := \max\{1, M_1\}$, where

$$M_1 := \frac{P}{Q} \left(\frac{L^{2^{d-1}}}{(\log P)^{ns+1}} \right)^{\frac{1}{n(s-B_d)}}.$$

If $M_1 \le 1$ then

$$\frac{L^{2^{d-1}}}{(\log P)^{ns+1}} \le \left(\frac{Q}{P} \right)^{n(s-B_d)},$$

as required by the first alternative in the conclusion of the lemma.

Therefore, we may suppose that $M = M_1 > 1$ and consider two cases according to whether all points $(\mathbf{x}_1, \dots, \mathbf{x}_{d-1}) \in (\mathfrak{n}^s)^{d-1}$ counted by $\mathcal{M}$ from Lemma 7.6 are in the affine variety $\widehat{S}_d$ or not.

If they all lie in $\widehat{S}_d$ then an application of Lemma 7.8 implies

$$\mathcal{M} \le \mathcal{M}_0 \left(\frac{P}{QM} \right) \ll \left(\frac{P}{QM} \right)^{nB_d + ns(d-2)}.$$

Hence we have

$$\begin{aligned} L^{2^{d-1}} &\ll P^{-(d-1)sn} (QM)^{(d-1)ns} (\log P)^{ns} \left(\frac{P}{QM} \right)^{nB_d + ns(d-2)} \\ &\ll (QM)^{ns - nB_d} P^{nB_d - ns} (\log P)^{ns} \\ &= \left(\frac{QM}{P} \right)^{n(s-B_d)} (\log P)^{ns}. \end{aligned}$$

Substituting M yields

$$L^{2^{d-1}} \ll L^{2^{d-1}} (\log P)^{-1}$$

which is a contradiction for P sufficiently large.

In the remaining case, we are given a point $(\mathbf{x}_1, \dots, \mathbf{x}_{d-1}) \in (\mathfrak{n}^s)^{d-1}$ with

— $|\mathbf{x}_i| \le \frac{P}{QM}$ for all $1 \le i < d$,

— $\|(qF)^*(\mathbf{X}_1 | \cdots | \mathbf{X}_{d-1} | \mathbf{E}_{p,\ell})\| \le \frac{1}{PQ^{d-2}M^{d-1}}$ for all $1 \le p \le s$, $1 \le \ell \le n$, and

— $\operatorname{rank}(\widehat{J}_d(\mathbf{x}_1,\dots,\mathbf{x}_{d-1})) = t_d$.

Without loss of generality, we assume that the matrix W consisting of the first t_d columns of $\widehat{J}_d(\mathbf{x}_1,\dots,\mathbf{x}_{d-1})$ has full rank. Let $\widetilde{q}^* := \det W$. Then $\widetilde{q}^* \in \mathfrak{n}$ and

$$|\widetilde{q}^*| \ll |\mathbf{x}_1|^{t_d}\cdots|\mathbf{x}_{d-1}|^{t_d} \ll \left(\frac{P}{QM}\right)^{t_d(d-1)}.$$

We set

$$\beta_p^* = \sum_{k=1}^{n}\beta_{p,k}^*\omega_k := qF(\mathbf{x}_1|\mathbf{x}_2|\cdots|\mathbf{x}_{d-1}|\mathbf{v}_p) = \sum_{i=1}^{t_d}\alpha_{d,i}qF_{d,i}(\mathbf{x}_1|\cdots|\mathbf{x}_{d-1}|\mathbf{v}_p).$$

Then

$$\operatorname{Tr}(\omega_\ell\beta_p^*) = \operatorname{Tr}\left(\omega_\ell qF(\mathbf{x}_1|\cdots|\mathbf{x}_{d-1}|\mathbf{v}_p)\right) = (qF)^*(\mathbf{X}_1|\cdots|\mathbf{X}_{d-1}|\mathbf{E}_{p,\ell}),$$

so

$$\left\|\operatorname{Tr}(\omega_\ell\beta_p^*)\right\| \le \frac{1}{PQ^{d-2}M^{d-1}}.$$

Thus, we can write

$$\sum_{k=1}^{n}\beta_{p,k}^*\operatorname{Tr}(\omega_\ell\omega_k) = \operatorname{Tr}(\omega_\ell\beta_p^*) = a_{p,\ell} + d_{p,\ell}, \tag{3.7}$$

with $a_{p,\ell}\in\mathbb{Z}$ and $|d_{p,\ell}| \le (PQ^{d-2}M^{d-1})^{-1}$. With

$$\Omega := (\operatorname{Tr}(\omega_\ell\omega_k))_{k,\ell=1,\dots,n} \qquad B_p^* := (\beta_{p,1}^*,\dots,\beta_{p,n}^*)$$
$$A_p := (a_{p,1},\dots,a_{p,n}) \qquad D_p := (d_{p,1},\dots,d_{p,n}),$$

we can write (3.7) as

$$B_p^*\Omega = A_p + D_p.$$

Therefore,

$$B_p^* = A_p\Omega^{-1} + D_p\Omega^{-1} =: A_p' + D_p',$$

Write $d_p' := d_{p,1}'\omega_1 + \cdots + d_{p,n}'\omega_n$, where $D_p' = (d_{p,1}',\dots,d_{p,n}')$, and define a_p' analogously. Then

$$\left|d_p'\right| \ll \max_k\{|d_{p,k}|\} \le \frac{1}{PQ^{d-2}M^{d-1}}.$$

Let $\boldsymbol{\alpha}_d := (\alpha_{d,1},\dots,\alpha_{d,t_d})$. On the one hand, by our definition of β_p^* we see that

$$\boldsymbol{\alpha}_d\cdot qW = \left(\beta_p^*\right)_{1\le p\le t_d}.$$

On the other hand, since W has full rank there exists $\boldsymbol{\nu}_d\in\mathfrak{n}^{t_d}$ such that

$$\boldsymbol{\nu}_d\cdot W = \det(\Omega)\widetilde{q}^*(a_p')_{1\le p\le t_d}.$$

Subtracting one from the other yields

$$(\det(\Omega)\widetilde{q}^*q\boldsymbol{\alpha}_d - \boldsymbol{\nu}_d)\cdot W = \det(\Omega)\widetilde{q}^*(d_p')_{1\le p\le t_d}.$$

We let $q^* := \widetilde{q}^* \det(\Omega)$ and obtain

$$\begin{aligned} q^* q\boldsymbol{\alpha}_d - \boldsymbol{\nu}_d &= q^*(d'_p)_{1\le p\le t_d} W^{-1} \\ &\ll |\mathbf{x}_1|^{t_d-1} \cdots |\mathbf{x}_{d-1}|^{t_d-1} \max_p\{|d'_p|\} \\ &\ll \left(\frac{P}{QM}\right)^{(t_d-1)(d-1)} \frac{1}{PQ^{d-2}M^{d-1}}. \end{aligned}$$

Furthermore, we have

$$|q^*| \ll |\widetilde{q}^*| \ll \left(\frac{P}{QM}\right)^{t_d(d-1)},$$

and thus

$$|q^*| \le Q^* = \left(\frac{P}{QM}\right)^{t_d(d-1)} (\log P)$$

for large enough P. □

Now we are ready to prove Lemma 7.7.

Proof of Lemma 7.7. We iteratively apply the preceding lemma in order to reduce the degree of f in every step. In the first step with $d = D$, we see that either

$$L^{2^{D-1}} \le P^{n(B_d-s)} (\log P)^{ns+1},$$

and hence $\boldsymbol{\alpha} \in I_D^{(1)}$, or there is a $q_D \le Q_D$, with

$$Q_D = \left((\log P)^{ns+1} L^{-2^{D-1}}\right)^{\frac{(D-1)t_D}{n(s-B_D)}} \log P,$$

and $\boldsymbol{\nu}_D \in \mathfrak{n}^{t_D}$ such that

$$|q\alpha_{D,i} - \nu_{D,i}| \ll QP^{-D} \quad (1 \le i \le t_D).$$

In the second case, then we apply Lemma 7.9 with $d := \max\{\Delta \setminus \{D\}\}$. Then either

$$L^{2^{d-1}} \le \left(\frac{Q_D}{P}\right)^{n(s-B_d)} (\log P)^{ns+1},$$

and thus $\boldsymbol{\alpha} \in I_d^{(1)}$, or there is a $q_d := q_D q^* \le Q_d := Q_D Q^*$ with

$$Q^* = \left(\frac{(\log P)^{ns+1}}{L^{2^{d-1}}}\right)^{\frac{t_d(d-1)}{n(s-B_d)}} \log P,$$

and $\boldsymbol{\nu}_d \in \mathfrak{n}^{t_d}$, such that

$$|q_d\alpha_{d,i} - \nu_{d,i}| \le Q_d P^{-d} \quad (1 \le i \le t_d).$$

Since we also have

$$|q_d\alpha_{D,i} - q^*\nu_{D,i}| \le Q^* Q_D P^{-D} = Q_d P^{-D},$$

so we may apply Lemma 7.9 again with the next lower value of d. Iterating this process we get sequences of q_d and Q_d for decreasing values of $d \in \Delta$. The set of $\boldsymbol{\alpha}$ such that for all $d \in \Delta$ the second case of Lemma 7.9 holds is exactly $I^{(2)}$. □

4. Minor arcs

First, let us consider the integral of $S(\boldsymbol{\alpha})$ over $I_D^{(1)}$.

LEMMA 7.10. *If*

$$\mathcal{D}\frac{2^{D-1}}{s-B_D} < 1, \tag{4.1}$$

then

$$\int_{I_D^{(1)}} |S(\boldsymbol{\alpha})| \, \mathrm{d}\boldsymbol{\alpha} \ll P^{n(s-\mathcal{D})-\delta},$$

for some $\delta > 0$.

DÉMONSTRATION. For $\boldsymbol{\alpha} \in I_D^{(1)}$, we have

$$L^{2^{D-1}} \leq P^{-n(s-B_D)}(\log P)^{ns+1} \leq P^{-n(s-B_D)+\varepsilon}.$$

Therefore the integral can be estimated by

$$\begin{aligned}\int_{I_D^{(1)}} |S(\boldsymbol{\alpha})| \, \mathrm{d}\boldsymbol{\alpha} &\ll \mathrm{vol}(I_D^{(1)}) \sup_{\boldsymbol{\alpha}\in I_D^{(1)}} |S(\boldsymbol{\alpha})| \ll 1 \cdot P^{ns} P^{\frac{-n(s-B_D)}{2^{D-1}}+\epsilon} \\ &= P^{n\left(s-\frac{s-B_D}{2^{D-1}}\right)+\varepsilon} \ll P^{n(s-\mathcal{D})-\delta}.\end{aligned}$$

for a suitable $\delta > 0$, using (4.1), provided that ϵ was chosen small enough. □

We split R^T into dyadic sets as follows. For any $L_0 > 0$, let

$$\mathcal{A}(L_0) := \left\{\boldsymbol{\alpha} \in R^T : \ |S(\boldsymbol{\alpha})| = P^{ns}L \text{ with } L_0 < L \leq 2L_0\right\}.$$

For $I = I_d^{(1)}$, $d < D$, or $I = I^{(2)}$, we write $\mathcal{A}(L_0; I) := I \cap \mathcal{A}(L_0) \cap \mathfrak{m}$ and estimate the integral

$$T(L_0; I) := \int_{\mathcal{A}(L_0;I)} |S(\boldsymbol{\alpha})| \, \mathrm{d}\boldsymbol{\alpha}.$$

We will make use of the following facts.

LEMMA 7.11. *Let* $\epsilon > 0$ *and* $a \in \mathcal{O}_K$ *with* $N(a) \leq H$. *Then the number of* $b \in \mathcal{O}_K$ *with* $b \mid a$ *and* $|b| \leq H$ *is* $\ll_\epsilon H^\epsilon$.

DÉMONSTRATION. There are at most $\ll_\epsilon H^{\epsilon/2}$ ideals of $\mathcal{O}_K$ dividing the principal ideal $a\mathcal{O}_K$. Let $\mathfrak{b}$ be any principal ideal among these divisors. The number of generators of $\mathfrak{b}$ with all conjugates bounded by $\ll H$ is $\ll_\epsilon H^{\epsilon/2}$, which one can see by counting units with bounded conjugates (for example, as in the proof of [**85**, Lemma 7.2]). □

LEMMA 7.12. *There are positive constants* e_0, e_1 *such that all* $\boldsymbol{\alpha} \in R^T \smallsetminus I_D^{(1)}$ *satisfy* $L \gg P^{-e_0}$ *and* $Q_j \ll P^{e_1}$ *for all* j.

DÉMONSTRATION. The lower bound for L follows directly from the definition of $I_D^{(1)}$. The upper bound for Q_j is an immediate consequence of this. □

LEMMA 7.13. *Let* $d \in \Delta$, $d < D$. *If*

$$\mathcal{D}_d\left(\frac{2^{d-1}}{s-B_d} + s_{d+1}\right) + s_{d+1}\sum_{j=d+1}^{D} s_j t_j < 1, \tag{4.2}$$

then $T(L_0; I_d^{(1)}) \ll P^{n(s-\mathcal{D})-\delta}$ *for some* $\delta > 0$.

DÉMONSTRATION. By Lemma 7.7, every $\boldsymbol{\alpha} \in I_d^{(1)}$ satisfies

$$L^{\frac{2^{d-1}}{n(s-B_d)}+\frac{s_{d+1}}{n}} \ll P^{-1+\varepsilon},$$

and there are $q_j \in \mathfrak{n} \smallsetminus \{0\}, \boldsymbol{\nu}_j \in \mathfrak{n}^{t_j}$ for all $j \in \Delta$, $j > d$, satisfying (3.4), (3.5), and (3.6).

Since $q_j R$ is a fundamental domain for the ideal lattice $q_j\mathfrak{n} \subseteq \mathfrak{n}$ in V, there are exactly $|N(q_j)|$ points of $\mathfrak{n}$ in $q_j R$. Hence, for any given q_j, it is enough to consider $\ll |N(q_j)|^{t_j}$ elements $\boldsymbol{\nu}_j$.

Let us estimate the volume of the set of all $(\alpha_{j,i})_{j>d,1\le i\le t_d}$ belonging to a given $q_j, \boldsymbol{\nu}_j$. By (3.6), we see that every coordinate $q_j\alpha_{j,i}$ takes values in a set of volume $\le Q_j^n P^{-jn}$. Since multiplication by q_j is an $\mathbb{R}$-linear transformation on V of determinant $\asymp N(q_j)$, each $\alpha_{i,j}$ belongs to a set of volume $\ll N(q_j)^{-1} Q_j^n P^{-jn}$, and hence the total volume is $\ll \prod_{j=d+1}^{D} (|N(q_j)|^{-1} Q_j^n P^{-jn})^{t_j}$. Due to (3.4), Lemma 7.11 and Lemma 7.12, each choice of q_{d+1} defines $\ll_\epsilon P^\epsilon$ values of $q_{d+2}, \dots, q_D$. Summing over all these q_j and all the corresponding $\boldsymbol{\nu}_j$, we see that

$$\operatorname{vol} \mathcal{A}(L_0; I_d^{(1)}) \ll P^{\varepsilon} Q_{d+1}^n \prod_{j=d+1}^{D} (Q_j P^{-j})^{nt_j} \ll P^{2\varepsilon - n\sum_{j=d+1}^{D} jt_j} L_0^{-s_{d+1}-\sum_{j=d+1}^{D} s_j t_j}.$$

Therefore,

$$T(L_0; I_d^{(1)}) \ll P^{n(s-\mathcal{D}+\mathcal{D}_d)+2\varepsilon} L_0^{1-s_{d+1}-\sum_{j=d+1}^{D} s_j t_j} \ll P^{n(s-\mathcal{D})-\delta},$$

as long as (4.2) holds and ϵ is small enough. □

Finally, we concentrate on the integral over $\mathcal{A}(L_0; I^{(2)})$. In particular, we will make use of the fact that $\mathcal{A}(L_0; I^{(2)}) \subseteq \mathfrak{m}$.

LEMMA 7.14. *Let $d \in \Delta$, $d < D$. If*

$$s_1 + \sum_{j=1}^{D} s_j t_j < 1. \tag{4.3}$$

then $T(L_0; I^{(2)}) \ll P^{n(s-\mathcal{D})-\delta}$ for some $\delta > 0$.

DÉMONSTRATION. For each $\boldsymbol{\alpha} \in I^{(2)}$, we have $q_d \in \mathfrak{n} \smallsetminus \{0\}$ and $\boldsymbol{\nu}_d \in \mathfrak{n}^{t_d}$, $d \in \Delta$, with (3.4), (3.5) and (3.6), and as in the previous lemma it suffices to consider $\ll |N(q_d)|^{t_d}$ tuples $\boldsymbol{\nu}_d$ for each q_d.

Let $\boldsymbol{\gamma} := (q_d^{-1}\nu_{d,i})_{d,i}$. Then it is not hard to see that

$$|\alpha_{d,i} - \gamma_{d,i}| \ll Q_d^n P^{-d} \text{ for all } j, d \text{ and}$$
$$\mathfrak{N}\, \mathfrak{a}_{\boldsymbol{\gamma}} \ll Q_1^n.$$

With $e_{\max} := \max_d\{e(d)\}$, we have

$$Q_d^n \ll L^{-s_1} (\log P)^{ne_{\max}}.$$

Let ϖ be as in the definition of the major arcs, and suppose that $L \ge P^{-\varpi/(2s_1)}$. If P is large enough, we deduce that

$$|\alpha_{d,i} - \gamma_{d,i}| \ll P^{-d+\varpi} \text{ for all } j, d \text{ and}$$
$$\mathfrak{N}\, \mathfrak{a}_{\boldsymbol{\gamma}} \ll P^{\varpi},$$

and hence $\boldsymbol{\alpha} \in \mathfrak{M}$. We conclude that $T(L_0; I^{(2)}) = 0$ unless

$$L_0 \ll P^{-\varpi/(2s_1)}. \tag{4.4}$$

Let us assume that (4.4) holds. As in the proof of Lemma 7.13, we see that

$$\mathrm{vol}(\mathcal{A}(L_0, I^{(2)})) \ll P^{\varepsilon} Q_1^n \prod_{j=1}^{D} (Q_j P^{-j})^{nt_j} \ll P^{2\varepsilon - n\sum_{j=1}^{D} jt_j} L_0^{-s_1 - \sum_{j=1}^{D} s_j t_j}.$$

This implies that

$$T(L_0; I^{(2)}) \ll P^{n(s-\mathcal{D})+2\varepsilon} L_0^{1-s_1-\sum_{j=1}^{D} s_j t_j} \ll P^{n(s-\mathcal{D})-\delta},$$

provided that (4.3) holds and ϵ is small enough. □

The previous lemmata allow us to estimate the integral of $|S(\boldsymbol{\alpha})|$ over $\boldsymbol{\alpha} \in \mathfrak{m}$. Lemma 7.10 gives a sufficient bound for the integral over $\mathfrak{m} \cap I_D^{(1)}$. For $\boldsymbol{\alpha} \in I_d^{(1)}$, $d < D$, or $\boldsymbol{\alpha} \in I^{(2)}$, we have $c_0 P^{-e_0} \le L \le c_1$, with constants c_0, c_1 independent from P. We split this interval in dyadic parts and obtain

$$\int_{\mathfrak{m} \cap I_d^{(1)}} |S(\boldsymbol{\alpha})| \, \mathrm{d}\boldsymbol{\alpha} \ll \sum_{j=0}^{\lceil \log_2(c_0^{-1} c_1 P^{e_0}) \rceil} T(2^j c_0 P^{-e_0}, I_d^{(1)}) \ll P^{n(s-\mathcal{D})-\delta} (\log P)$$

by Lemma 7.13. An analogous argument using Lemma 7.14 bounds the integral over $\mathfrak{m} \cap I^{(2)}$.

5. Major arcs : singular series

We now choose the parameter ϖ in the definition of the major arcs by

$$\varpi := \frac{1}{4 + (n+1)T}.$$

Furthermore recall that $\mathfrak{B} \subset V^s$ is a box aligned to the basis and $B \subseteq [-1,1]^{ns}$ the corresponding box in R^{ns}.

We start by showing that the major arcs are disjoint in pairs provided P is large enough.

Lemma 7.15. *Let $\boldsymbol{\gamma}_1 \neq \boldsymbol{\gamma}_2 \in (R \cap K)^T$ with $\mathfrak{N}\mathfrak{a}_{\boldsymbol{\gamma}_j} \le P^{\varpi}$ for $j \in \{1,2\}$. For $P \gg 1$, we have $\mathfrak{M}_{\boldsymbol{\gamma}_1} \cap \mathfrak{M}_{\boldsymbol{\gamma}_2} = \emptyset$.*

Démonstration. If $\boldsymbol{\alpha} \in \mathfrak{M}_{\boldsymbol{\gamma}_1} \cap \mathfrak{M}_{\boldsymbol{\gamma}_2}$ then, writing $\boldsymbol{\gamma}_j = (\gamma_{j,d,i})_{d,i}$,

$$|\gamma_{1,d,i} - \gamma_{2,d,i}| \le |\gamma_{1,d,i} - \alpha_{d,i}| + |\alpha_{d,i} - \gamma_{2,d,i}| \ll P^{-d+\varpi} \le P^{-1+\varpi}$$

holds for all $1 \le d \le D, 1 \le i \le t_d$. By Minkowski's convex body theorem, there is a nonzero $q \in \mathfrak{a}_{\boldsymbol{\gamma}_1} \cap \mathfrak{a}_{\boldsymbol{\gamma}_2}$ with $|q| \ll P^{2\varpi/n}$. Hence, $q(\gamma_{1,d,i} - \gamma_{2,d,i}) \in \mathfrak{n}$ and $|q(\gamma_{1,d,i} - \gamma_{2,d,i})| \ll P^{-1+(1+2/n)\varpi}$ for all d, i. Since $\varpi < 1/3$, this implies that $\boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_2$ whenever P is large enough. □

For $\boldsymbol{\gamma} \in (R \cap K)^T$, we define

$$\Sigma(\boldsymbol{\gamma}) := \sum_{\mathbf{x} \in (\mathfrak{n}/\mathfrak{a}_{\boldsymbol{\gamma}}\mathfrak{n})^s} \Phi\left(\sum_{d=1}^{D} \sum_{i=1}^{t_d} \gamma_{d,i} G_{d,i}(\mathbf{x}) \right),$$

and for $\boldsymbol{\gamma} \in V^T$, let

$$J(\boldsymbol{\gamma}) := \int_{\mathcal{B}} \Phi\left(\sum_{d=1}^{D}\sum_{i=1}^{t_d} \gamma_{d,i} F_{d,i}(\mathbf{x})\right) \mathrm{d}\mathbf{x}.$$

LEMMA 7.16. *For $\boldsymbol{\gamma} \in (R \cap K)^T$ with $\mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}} \le P^{\varpi}$, let $\boldsymbol{\alpha} \in \mathfrak{M}_{\boldsymbol{\gamma}}$ and write $\alpha_{d,i} = \gamma_{d,i} + \theta_{d,i}$ for all $1 \le d \le D$ and $1 \le i \le t_d$. Then*

$$S(\boldsymbol{\alpha}) = \mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}}^{-s} P^{ns} \Sigma(\boldsymbol{\gamma}) J((\theta_{d,i}P^d)_{d,i}) + O\left(\mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}} \sum_{d=1}^{D}\sum_{i=1}^{t_d} |\theta_{d,i}|\, P^{ns+d-1} + \mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}} P^{ns-1}\right).$$

DÉMONSTRATION. Whenever d, i, j appear as indices of a sum, the sum runs over $1 \le d \le D$, $1 \le i \le t_d$, and $1 \le j \le n$. As usual, we write $\gamma_{d,i} = \gamma_{d,i,1}\omega_1 + \cdots + \gamma_{d,i,n}\omega_n$, and similarly $\theta_{d,i} = \theta_{d,i,1}\omega_1 + \cdots + \theta_{d,i,n}\omega_n$. With these conventions, we have

$$S(\boldsymbol{\alpha}) = \sum_{\mathbf{X} \in \mathbb{Z}^{sn} \cap PB} e\left(\sum_{d,i,j} (\theta_{d,i,j} + \gamma_{d,i,j}) G^*_{d,i,j}(\mathbf{X})\right).$$

Let $N := \mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}} \in (\mathbb{N} \cap \mathfrak{a}_{\boldsymbol{\gamma}})$. Then $N\boldsymbol{\gamma} \in \mathfrak{n}^T$, so in particular $N\gamma_{d,i,j} \in \mathbb{Z}$ for all d, i, j. Applying the standard argument over $\mathbb{Q}$, we see that $S(\boldsymbol{\alpha})$ is the sum of

$$\frac{1}{N^{ns}} \sum_{\mathbf{Y} \in ([0,N-1] \cap \mathbb{Z})^{sn}} e\left(\sum_{d,i,j} \gamma_{d,i,j} G^*_{d,i,j}(\mathbf{Y})\right) \cdot J((\theta_{d,i}P^d)_{d,i}) \cdot P^{ns}$$

and an error term as in the lemma. It remains to show that

$$\frac{1}{N^{ns}} \sum_{\mathbf{Y} \in ([0,N-1] \cap \mathbb{Z})^{sn}} e\left(\sum_{d,i,j} \gamma_{d,i,j} G^*_{d,i,j}(\mathbf{Y})\right) = \frac{1}{N^s} \Sigma(\boldsymbol{\gamma}).$$

This follows from the following observations. Write $\mathbf{y} = (y_1, \ldots, y_s)$, with, as usual, $y_j = y_{j,1}\omega_1 + \cdots + y_{j,n}\omega_n$. If $\mathbf{Y}$ runs through $([0, N-1] \cap \mathbb{Z})^{ns}$ then $\mathbf{y}$ runs through a set of representatives of $(\mathfrak{n}/(N\mathfrak{n}))^s$. Moreover,

$$e\left(\sum_{d,i,j} \gamma_{d,i,j} G^*_{d,i,j}(\mathbf{Y})\right) = \Phi\left(\sum_{d,i} \gamma_{d,i} G_{d,i}(\mathbf{y})\right)$$

depends only on $\mathbf{y}$ modulo $\mathfrak{a}_{\boldsymbol{\gamma}}\mathfrak{n}$, and each coset of $(\mathfrak{n}/(N\mathfrak{n}))^s$ modulo $\mathfrak{a}_{\boldsymbol{\gamma}}\mathfrak{n}$ has $N^{(n-1)s}$ elements. □

For $H > 0$, let

$$\mathfrak{S}(H) := \sum_{\substack{\boldsymbol{\gamma} \in (R \cap K)^T \\ \mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}} \le H}} \frac{\Sigma(\boldsymbol{\gamma})}{\mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}}^s}$$

and

$$\mathfrak{J}(H) := \int_{\substack{\boldsymbol{\gamma} \in V^T \\ |\boldsymbol{\gamma}| \le H}} J(\boldsymbol{\gamma})\, \mathrm{d}\boldsymbol{\gamma}.$$

LEMMA 7.17. *There is a positive constant δ such that, for large enough P,*

$$\int_{\mathfrak{M}} S(\boldsymbol{\alpha})\, \mathrm{d}\boldsymbol{\alpha} = \mathfrak{S}(P^{\varpi})\mathfrak{J}(P^{\varpi}) P^{n(s-\mathcal{D})} + O(P^{n(s-\mathcal{D})-\delta}).$$

Démonstration. By Lemma 7.15, we have

$$\int_{\mathfrak{M}} S(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha} = \sum_{\substack{\boldsymbol{\gamma}\in(R\cap K)^T\\ \mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}}\le P^{\varpi}}} \int_{\mathfrak{M}_{\boldsymbol{\gamma}}} S(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha}.$$

Using Lemma 7.16 and the obvious fact that $\mathrm{vol}(\mathfrak{M}_{\boldsymbol{\gamma}}) = P^{-n\mathcal{D}+nT\varpi}$, it follows that

$$\int_{\mathfrak{M}_{\boldsymbol{\gamma}}} S(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha} = \frac{1}{\mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}}^s} P^{ns}\Sigma(\boldsymbol{\gamma}) \int_{|\theta_{d,i}|\le P^{-d+\varpi}} J((\theta_{d,i}P^d)_{d,i})\,\mathrm{d}\boldsymbol{\theta} + O(P^{n(s-\mathcal{D})-1+(nT+2)\varpi}).$$

After a coordinate change in the integral over $\boldsymbol{\theta}$ and summing over all $\boldsymbol{\gamma}$, we obtain

$$\int_{\mathfrak{M}} S(\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{\alpha} = \mathfrak{S}(P^{\varpi})\mathfrak{J}(P^{\varpi})P^{n(s-\mathcal{D})} + O\left(m(P^{\varpi})\cdot P^{n(s-\mathcal{D})-1+(nT+2)\varpi}\right),$$

where

$$m(P^{\varpi}) := |\{\boldsymbol{\gamma}\in(R\cap K)^T \mid \mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}}\le P^{\varpi}\}| \ll P^{\varpi(T+1)}.$$

Hence, we obtain an error term

$$O(P^{n(s-\mathcal{D})-1+\varpi((n+1)T+3)}) = O(P^{n(s-\mathcal{D})-\delta}). \qquad \square$$

Whenever the respective limit exists, we let

$$\mathfrak{S} := \lim_{H\to\infty}\mathfrak{S}(H) \quad \text{and} \quad \mathfrak{J} := \lim_{H\to\infty}\mathfrak{J}(H).$$

The rest of this section is devoted to the absolute and fast convergence of the singular series $\mathfrak{S}$. The singular integral $\mathfrak{J}$ will be treated in the next section.

We start with an estimate for $\Sigma(\boldsymbol{\gamma})$. Let $\boldsymbol{\gamma}\in(R\cap K)^T$. By definition of $\mathfrak{a}_{\boldsymbol{\gamma}}$, we can write $\gamma_{d,i}\mathcal{O}_K = \frac{\mathfrak{n}\mathfrak{a}_{d,i}}{\mathfrak{a}_{\boldsymbol{\gamma}}}$, where $\mathfrak{a}_{d,i}$ is an ideal of $\mathcal{O}_K$ and $\mathfrak{a}_{\boldsymbol{\gamma}} + \sum_{d=1}^{D}\sum_{i=1}^{t_d}\mathfrak{a}_{d,i} = \mathcal{O}_K$.

Lemma 7.18. *Write $\gamma_{d,i}\mathcal{O}_K = \frac{\mathfrak{n}\mathfrak{a}_{d,i}}{\mathfrak{a}_{\boldsymbol{\gamma}}}$ as above. Then, for $\epsilon>0$,*

$$\Sigma(\boldsymbol{\gamma}) \ll \mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}}^{s+\epsilon}\min_{j\in\Delta}\left\{\frac{\mathfrak{N}(\mathfrak{a}_{\boldsymbol{\gamma}}+\sum_{d=j}^{D}\sum_{i=1}^{t_d}\mathfrak{a}_{d,i})}{\mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}}}\right\}^{1/s_j}.$$

Démonstration. This generalizes [**48**, Lemma 8.2]. The proof is essentially the same. We choose $\boldsymbol{\alpha}=\boldsymbol{\gamma}$, $\boldsymbol{\theta}=\mathbf{0}$ in Lemma 7.16. Clearly, $\boldsymbol{\alpha}=\boldsymbol{\gamma}$ is in $\mathfrak{M}_{\boldsymbol{\gamma}}$ for $P=\mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}}^A$, with any large fixed value of A. Since $J(\mathbf{0})\gg 1$, we obtain

$$\Sigma(\boldsymbol{\gamma}) \ll \frac{\mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}}^s\cdot|S(\boldsymbol{\gamma})|}{P^{ns}} + \frac{\mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}}^{s+1}}{P}.$$

We choose $A>s+1$ to obtain

$$\Sigma(\boldsymbol{\gamma}) \ll 1 + \mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}}^s L, \tag{5.1}$$

with L defined via $|S(\boldsymbol{\alpha})| = P^{ns}L$ as earlier in the paper. Let us apply Lemma 7.7 to estimate L. If $\boldsymbol{\gamma}\in I_d^{(1)}$ for some $d\in\Delta$ then (3.3) and (5.1) show that

$$\Sigma(\boldsymbol{\gamma}) \ll 1 + \frac{N\mathfrak{a}_{\boldsymbol{\gamma}}^s P^{\epsilon}}{P^{n(s-B_d)/(2^{d-1}+(s-B_d)s_{d+1})}} \ll 1$$

if A is chosen large enough. Now let us assume that $\boldsymbol{\alpha}=\boldsymbol{\gamma}\in I^{(2)}$. In this case, Lemma 7.7 yields $q_j\in\mathfrak{n}$ and $\boldsymbol{\nu}_j\in\mathfrak{n}^{t_j}$ for $j\in\Delta$ satisfying (3.4), (3.5), and (3.6), with Q_j given in (3.1).

Assume first that $q_j\gamma_{j,i} \neq \nu_{j,i}$ for some j, i. By Minkowski's convex body theorem, there is a $q \in \mathfrak{a}_{\boldsymbol{\gamma}} \smallsetminus \{0\}$ with $|q| \ll \mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}}^{1/n}$. Then $q(q_j\gamma_{j,i} - \nu_{j,i}) \in \mathfrak{n}$, and so

$$1 \ll |q(q_j\gamma_{j,i} - \nu_{j,i})| \ll \mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}}^{1/n} Q_j P^{-j} \ll \mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}}^{1/n} L^{-s_j/n} P^{-j+\epsilon}.$$

This gives an upper bound for L, and substituting this bound in (5.1) shows that $\Sigma(\boldsymbol{\gamma}) \ll 1$ as long as we have chosen A big enough. Hence, we are left with the case where

$$q_j\gamma_{j,i} = \nu_{j,i} \quad \text{for all} \quad j \in \Delta, 1 \leq i \leq t_j. \tag{5.2}$$

Since $\nu_{j,i} \in \mathfrak{n}$, we find integral ideals $\mathfrak{b}_{j,i}$ such that $\nu_{j,i}\mathcal{O}_K = \mathfrak{n}\mathfrak{b}_{j,i}$. After cancellation, (5.2) gives

$$q_j\mathfrak{a}_{j,i} = \mathfrak{a}_{\boldsymbol{\gamma}}\mathfrak{b}_{j,i} \quad \text{for all} \quad j \in \Delta, 1 \leq i \leq t_j.$$

In the following arguments, we write $\mathfrak{a}^{(j)} := \mathfrak{a}_{j,1} + \cdots + \mathfrak{a}_{j,t_j}$ and $\mathfrak{b}^{(j)} := \mathfrak{b}_{j,1} + \cdots + \mathfrak{b}_{j,t_j}$. Then $q_j\mathfrak{a}^{(j)} = \mathfrak{a}_\gamma\mathfrak{b}^{(j)}$ holds for all $j \in \Delta$. With

$$\mathfrak{d}_j := \mathfrak{a}_{\boldsymbol{\gamma}} + \mathfrak{a}^{(j)} \text{ and } \mathfrak{f}_j := q_j\mathcal{O}_K + \mathfrak{b}^{(j)},$$

we have thus

$$\frac{q_j\mathcal{O}_K}{\mathfrak{f}_j} \cdot \frac{\mathfrak{a}^{(j)}}{\mathfrak{d}_j} = \frac{\mathfrak{a}_{\boldsymbol{\gamma}}}{\mathfrak{d}_j} \cdot \frac{\mathfrak{b}^{(j)}}{\mathfrak{f}_j}.$$

We claim that

$$\frac{\mathfrak{a}_{\boldsymbol{\gamma}}}{\sum_{d>j}\mathfrak{d}_d} \text{ divides } q_j\mathcal{O}_K \text{ for all } j \in \Delta. \tag{5.3}$$

Indeed, since $\frac{q_j\mathcal{O}_K}{\mathfrak{f}_j} + \frac{\mathfrak{b}^{(j)}}{\mathfrak{f}_j} = \mathcal{O}_K$, we see that $\frac{q_j\mathcal{O}_K}{\mathfrak{f}_j} \mid \frac{\mathfrak{a}_{\boldsymbol{\gamma}}}{\mathfrak{d}_j}$. The opposite divisibility follows analogously, and hence

$$\frac{q_j\mathcal{O}_K}{\mathfrak{f}_j} = \frac{\mathfrak{a}_{\boldsymbol{\gamma}}}{\mathfrak{d}_j}. \tag{5.4}$$

Let $k \in \Delta$, $k > j$. Since (5.4) holds for k as well as for j, we see that $q_j\mathfrak{d}_j\mathfrak{f}_k = q_k\mathfrak{d}_k\mathfrak{f}_j$. By (3.4), we can write $q_j = q_k\hat{q}_j$ with $\hat{q}_j \in \mathcal{O}_K$. Substituting this in the above equality, cancelling $q_k\mathcal{O}_K$, and dividing both sides by $\mathfrak{d}_j + \mathfrak{f}_j$ shows that

$$\hat{q}_j\mathfrak{f}_k\frac{\mathfrak{d}_j}{\mathfrak{d}_j + \mathfrak{f}_j} = \mathfrak{d}_k\frac{\mathfrak{f}_j}{\mathfrak{d}_j + \mathfrak{f}_j},$$

and in particular

$$\frac{\mathfrak{d}_j}{\mathfrak{d}_j + \mathfrak{f}_j} \text{ divides } \mathfrak{d}_k \text{ for all } k > j. \tag{5.5}$$

Let $\delta_j := \sum_{d>j}\mathfrak{d}_d$. By (5.4),

$$q_j\mathcal{O}_K = \frac{\mathfrak{a}_{\boldsymbol{\gamma}}}{\delta_j} \cdot \frac{\mathfrak{f}_j\delta_j}{\mathfrak{d}_j} = \frac{\mathfrak{a}_{\boldsymbol{\gamma}}}{\delta_j} \cdot \frac{\mathfrak{f}_j}{\mathfrak{d}_j + \mathfrak{f}_j} \cdot \frac{\delta_j}{\mathfrak{d}_j(\mathfrak{d}_j + \mathfrak{f}_j)^{-1}}.$$

By (5.5), the second and the third factor on the right-hand side are integral ideals, and hence (5.3) holds as claimed. Hence,

$$\frac{\mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}}}{\mathfrak{N}(\mathfrak{a}_{\boldsymbol{\gamma}} + \sum_{d=j}^{D}\sum_{i=1}^{t_d}\mathfrak{a}_{d,i})} = \frac{\mathfrak{N}\,\mathfrak{a}_{\boldsymbol{\gamma}}}{\mathfrak{N}\,\delta_j} \leq \mathfrak{N}(q_j\mathcal{O}_K) \ll |q_j|^n \ll Q_j^n \ll L^{-s_j}(\log P)^{ne(j)}.$$

This gives an upper bound for L which, once substituted into (5.1), proves the lemma. □

LEMMA 7.19. *Let $\mathfrak{a}$ be a fractional ideal of K. Then*

$$|R \cap \mathfrak{a}| \ll \frac{1}{\mathfrak{N}\,\mathfrak{a}} + 1.$$

DÉMONSTRATION. There is a constant c depending only on K and our basis $\omega_1, \dots, \omega_n$ such that $R \cap \mathfrak{a} \subseteq \{x \in \mathfrak{a} \mid \left|x^{(j)}\right| \leq c \text{ for all } v\}$. Here, the $x^{(v)}$ are all the (real and complex) conjugates of x. The result then follows from [**85**, Lemma 7.1]. □

We are now ready to treat our singular series under the hypotheses of Theorem 7.1.

LEMMA 7.20. *Assume that*

$$s_1 + \sum_{j=1}^{D} s_j t_j < 1. \tag{5.6}$$

Then the series defining $\mathfrak{S}$ converges absolutely and there is a positive constant δ such that

$$\mathfrak{S} - \mathfrak{S}(H) \ll H^{-\delta}$$

holds for large enough H.

DÉMONSTRATION. We write

$$A(\mathfrak{a}) := \sum_{\substack{\boldsymbol{\gamma} \in (R \cap K)^T \\ \mathfrak{a}_{\boldsymbol{\gamma}} = \mathfrak{a}}} |\Sigma(\boldsymbol{\gamma})|.$$

For each $\boldsymbol{\gamma} \in (R \cap K)^T$ with $\mathfrak{a}_{\boldsymbol{\gamma}} = \mathfrak{a}$, we write again $\gamma_{d,j}\mathcal{O}_K = \frac{\mathfrak{n}\mathfrak{a}_{d,j}}{\mathfrak{a}}$ and define

$$\mathfrak{d}_j = \mathfrak{a} + \sum_{d=j}^{D} \sum_{i=1}^{t_d} \mathfrak{a}_{d,i}. \tag{5.7}$$

Then for $j_0 := \min \Delta$ we have $\mathfrak{d}_{j_0} = \mathcal{O}_K$. By Lemma 7.18, we see that

$$\Sigma(\boldsymbol{\gamma}) \ll \mathfrak{N}\,\mathfrak{a}^{s+\epsilon/2} \min_{j \in \Delta} \left\{ \frac{\mathfrak{N}\,\mathfrak{d}_j}{\mathfrak{N}\,\mathfrak{a}} \right\}^{1/s_j} \leq \mathfrak{N}\,\mathfrak{a}^{s+\epsilon/2} \prod_{j \in \Delta} \left(\frac{\mathfrak{N}\,\mathfrak{d}_j}{\mathfrak{N}\,\mathfrak{a}} \right)^{\lambda_j/s_j},$$

for any $\lambda_j \geq 0$ with $\sum_{j \in \Delta} \lambda_j = 1$. As in [**48**, Section 8], we choose

$$\lambda_j := \begin{cases} \theta + t_{j_0} s_{j_0} & \text{if } j = j_0 \\ t_j s_j & \text{if } j \in \Delta \smallsetminus \{j_0\}, \end{cases} \tag{5.8}$$

where $\theta = 1 - \sum_{j \in \Delta} t_j s_j \in (s_1, 1)$. Hence,

$$A(\mathfrak{a}) \ll \mathfrak{N}\,\mathfrak{a}^{s+\epsilon/2} \sum_{\substack{\mathfrak{d}_j \mid \mathfrak{a} \\ j \in \Delta}} m((\mathfrak{d}_j)_{j \in \Delta}, \mathfrak{a}) \left(\frac{1}{\mathfrak{N}\,\mathfrak{a}} \right)^{\theta/s_{j_0}} \prod_{j \in \Delta} \left(\frac{\mathfrak{N}\,\mathfrak{d}_j}{\mathfrak{N}\,\mathfrak{a}} \right)^{t_j},$$

where $m((\mathfrak{d}_j)_{j \in \Delta}, \mathfrak{a})$ is the number of all $\boldsymbol{\gamma} \in (R \cap K)^T$ with $\mathfrak{a}_{\boldsymbol{\gamma}} = \mathfrak{a}$ and (5.7). Clearly, any $\boldsymbol{\gamma}$ with $\mathfrak{a}_{\boldsymbol{\gamma}} = \mathfrak{a}$ and (5.7) satisfies

$$\gamma_{j,i} \in \frac{\mathfrak{n}\mathfrak{d}_j}{\mathfrak{a}} \text{ for all } j \in \Delta, 1 \leq i \leq t_j.$$

Using this and Lemma 7.19,

$$m((\mathfrak{d}_j)_{j\in\Delta},\mathfrak{a}) \le \prod_{j\in\Delta}\left|R\cap\frac{\mathfrak{n}\mathfrak{d}_j}{\mathfrak{a}}\right|^{r_j} \ll \prod_{j\in\Delta}\left(\frac{\mathfrak{N}\,\mathfrak{a}}{\mathfrak{N}\,\mathfrak{d}_j}\right)^{t_j}.$$

Hence,

$$A(\mathfrak{a}) \ll \mathfrak{N}\,\mathfrak{a}^{s+\epsilon/2-\theta/s_{j_0}} \sum_{\substack{\mathfrak{d}_j\mid\mathfrak{a}\\ j\in\Delta}} 1 \ll \mathfrak{N}\,\mathfrak{a}^{s-\theta/s_{j_0}+\epsilon}.$$

Since $\theta > s_1 = s_{j_0}$, this shows that $\mathfrak{S}$ converges absolutely and that $\mathfrak{S}-\mathfrak{S}(H) \ll H^{\delta}$ for some appropriate delta. □

6. Major arcs : singular integral

Throughout this section, we will assume (5.6). For $\boldsymbol{\gamma} = (\gamma_{d,i})_{d,i} \in V^T$, we write $\boldsymbol{\gamma}_d := (\gamma_{d,1},\dots,\gamma_{d,t_d})$.

LEMMA 7.21. *Let $a\in[0,n]$ and $\epsilon>0$. For any $\boldsymbol{\gamma}\in V^T$, we have*

$$J(\boldsymbol{\gamma}) \ll 1. \tag{6.1}$$

Assume that $|\boldsymbol{\gamma}|\ge 1$ and let $d\in\Delta$ with $\boldsymbol{\gamma}_d\neq 0$. Then there exists a unit $u_d\in\mathcal{O}_K^{\times}$ such that

(1) $|u_d| \ll |\boldsymbol{\gamma}|^{a/n}$,

(2) $J(\boldsymbol{\gamma}) \ll |\boldsymbol{\gamma}|^{\epsilon}\,|u_d\boldsymbol{\gamma}_d|^{-a/s_d}$.

Moreover, we have

$$J(\boldsymbol{\gamma}) \ll |\boldsymbol{\gamma}|^{\epsilon}\,|\boldsymbol{\gamma}_d|^{-1/s_d}. \tag{6.2}$$

DÉMONSTRATION. It is clear that $J(\boldsymbol{\gamma})\ll 1$ holds for all $\boldsymbol{\gamma}\in V^T$. Let $d\in\Delta$ and assume that $|\boldsymbol{\gamma}|\ge 1$ and $\boldsymbol{\gamma}_d\neq 0$. We apply Lemma 7.16 with $\boldsymbol{\alpha} := (P^{-j}\gamma_{j,i})_{j,i}$ and $P := |\boldsymbol{\gamma}|^A$ for fixed large A. Clearly, $\boldsymbol{\alpha}\in\mathfrak{M}_{\mathbf{0}}$ as soon as $A\ge 1/\varpi$. Since $\Sigma(\mathbf{0})=1$, we obtain

$$J(\boldsymbol{\gamma}) \ll L + |\boldsymbol{\gamma}|\,P^{-1} \ll L + |\boldsymbol{\gamma}|^{-a/s_d}, \tag{6.3}$$

if A was chosen big enough.

If $L\le |\boldsymbol{\gamma}|^{-a/s_d+\epsilon}$ then (6.3) yields

$$J(\boldsymbol{\gamma}) \ll |\boldsymbol{\gamma}|^{\epsilon}\,|\boldsymbol{\gamma}|^{-a/s_d} \le |\boldsymbol{\gamma}|^{\epsilon}\,|\boldsymbol{\gamma}_d|^{-a/s_d},$$

and we can choose $u_d=1$. Therefore, we may assume from now on that

$$L \ge |\boldsymbol{\gamma}|^{-a/s_d+\epsilon}, \tag{6.4}$$

so that

$$J(\boldsymbol{\gamma}) \ll L. \tag{6.5}$$

The remainder of this proof is devoted to the deduction of suitable upper bounds for L. Let us first assume that $\boldsymbol{\alpha}\in I_j^{(1)}$ for some $j\in\Delta$. Then the definition of $I_j^{(1)}$, see (3.2), yields an upper bound

$$L \ll |\boldsymbol{\gamma}|^{-a/s_d+\epsilon} \le |\boldsymbol{\gamma}|^{\epsilon}\,|\boldsymbol{\gamma}_d|^{-a/s_d},$$

provided that we have chosen A big enough to ensure that $A(s-B_j) > a(2^{j-1}+s_{j+1}(s-B_j))/(ns_d)$.

If $\boldsymbol{\alpha} \in I^{(2)}$ then Lemma 7.7 yields $q_d \in \mathfrak{n}$, and $\boldsymbol{\nu}_d \in \mathfrak{n}^{t_d}$ satisfying

$$|q_d| \leq Q_d \tag{6.6}$$

$$|q_d \alpha_{d,i} - \nu_{d,i}| \leq Q_d P^{-d} \qquad \text{for all } 1 \leq i \leq t_d, \tag{6.7}$$

with Q_d defined by (3.1).

Suppose that $\nu_{d,i} \neq 0$ for some $1 \leq i \leq t_d$. Then

$$1 \leq |\nu_{d,i}| \ll Q_d P^{-d} |\gamma_{d,i}| + Q_d P^{-d} \ll Q_d P^{-d} |\boldsymbol{\gamma}| = L^{-s_d/n} (\log P)^{e(d)} P^{-d} |\boldsymbol{\gamma}|.$$

This yields an upper bound

$$L \ll (\log P)^{ne(d)/s_d} P^{-nd/s_d} |\boldsymbol{\gamma}|^{n/s_d} \ll |\boldsymbol{\gamma}|^{\epsilon} |\boldsymbol{\gamma}|^{\frac{n}{s_d}(1-dA)} \ll |\boldsymbol{\gamma}|^{\epsilon} |\boldsymbol{\gamma}|^{-a/s_d} \ll |\boldsymbol{\gamma}|^{\epsilon} |\boldsymbol{\gamma}_d|^{-a/s_d}$$

if A is chosen big enough. We are left with the case where $\boldsymbol{\nu}_d = \mathbf{0}$. Let t be a generator of the principal ideal $q_d \mathcal{O}_K$ with the property that

$$N(q_d)^{1/n} \ll |t|_v \ll N(q_d)^{1/n} \text{ for all } v \in \Omega_\infty, \tag{6.8}$$

and let $u_d := q_d/t \in \mathcal{O}_K^\times$. Thanks to (6.8), (6.6) and (6.4), we obtain

$$|u_d| \asymp N(q_d)^{-1/n} |q_d| \leq Q_d = L^{-s_d/n} (\log P)^{e(d)} \ll |\boldsymbol{\gamma}|^{a/n}.$$

Moreover, due to (6.7), for all $1 \leq i \leq t_d$ we have

$$P^{-d} |u_d \boldsymbol{\gamma}_d| = |u_d \boldsymbol{\alpha}_d| \asymp N(q_d)^{-1} |q_d \boldsymbol{\alpha}_d| \leq Q_d P^{-d} = L^{-s_d/n} (\log P)^{e(d)} P^{-d}, \tag{6.9}$$

so

$$L \ll |\boldsymbol{\gamma}|^{\epsilon} |u_d \boldsymbol{\gamma}_d|^{-n/s_d}.$$

The estimate $J(\boldsymbol{\gamma}) \ll |\boldsymbol{\gamma}|^{\epsilon} |u_d \boldsymbol{\gamma}_d|^{-a/s_d}$ follows immediately from this and (6.5) if $|u_d \boldsymbol{\gamma}_d| \geq 1$, and from the trivial estimate $J(\boldsymbol{\gamma}) \ll 1$ if $|u_d \boldsymbol{\gamma}_d| \leq 1$.

For the proof of (6.2), we proceed as above with $a = 1$, until (6.9). Here, we conclude that

$$P^{-d} |\boldsymbol{\gamma}_d| = |\boldsymbol{\alpha}_d| \ll \left|q_d^{-1}\right| |q_d \boldsymbol{\alpha}_d| \ll Q_d^{n-1} \cdot Q_d P^{-d} = L^{-s_d} (\log P)^{ne(d)} P^{-d},$$

and thus

$$L \ll |\boldsymbol{\gamma}|^{\epsilon} |\boldsymbol{\gamma}_d|^{-1/s_d}.$$

□

We write

$$N(\boldsymbol{\gamma}_d) := \prod_{v \in \Omega_\infty} |\boldsymbol{\gamma}_d|_v^{n_v},$$

where $|\boldsymbol{\gamma}_d|_v := \max\{|\gamma_{d,1}|_v, \dots, |\gamma_{d,t_d}|_v\}$ and $n_v := [K_v : \mathbb{R}]$ is the local degree. Let $b \in (0,1)$ be a constant to be specified later, and for $H \geq 1$ let

$$M(H) := \{\boldsymbol{\gamma} \in V^T : |\boldsymbol{\gamma}| \gg H\},$$

$$M_>(H) := \{\boldsymbol{\gamma} \in V^T : |\boldsymbol{\gamma}| \in (H, 2H] \text{ and there exists } d \text{ with } N(\boldsymbol{\gamma}_d) \geq H^b\},$$

$$M_<(H) := \{\boldsymbol{\gamma} \in V^T : |\boldsymbol{\gamma}| \in (H, 2H] \text{ and for all } d, \text{ we have } N(\boldsymbol{\gamma}_d) \leq H^b\}.$$

Define the integral

$$I(H) := \int_{M(H)} \max_{d \in \Delta} \{|\boldsymbol{\gamma}_d|^{n/s_d}\}^{-1} \, \mathrm{d}\boldsymbol{\gamma}.$$

Lemma 7.22. *There is $\delta > 0$ such that, for $H \geq 1$,*

$$I(H) \ll H^{-\delta}.$$

DÉMONSTRATION. We identify V^T with $\mathbb{R}^{nT}$ using the basis $\omega_1, \dots, \omega_n$ of V. The exponent n/s_d in the definition $I(H)$ is good enough for the arguments given after [**48**, Lemma 8.3] to apply. □

For $\mathbf{u} = (u_d)_{d\in\Delta}$ with $u_d \in \mathcal{O}_K^\times$ for all $d \in \Delta$, let

$$I_>(\mathbf{u}, H) := \int_{M_>(H)} \max_{d\in\Delta}\{|u_d\boldsymbol{\gamma}_d|^{n/s_d}\}^{-1}\,\mathrm{d}\boldsymbol{\gamma}.$$

LEMMA 7.23. *Let δ be as in Lemma 7.22. Then, for $H \geq 1$,*

$$I_>(\mathbf{u}, H) \ll H^{-\delta b/n}.$$

DÉMONSTRATION. Let $\phi : V^T \to V^T$ be the $\mathbb{R}$-linear transformation $(\gamma_{d,i})_{d,i} \mapsto (u_d\gamma_{d,i})_{d,i}$. Since the u_d are all units, we have $\det\phi \asymp 1$. Moreover, let $\boldsymbol{\gamma} \in M_>(H)$ and d such that $N(u_d\boldsymbol{\gamma}_d) = N(\boldsymbol{\gamma}_d) \geq H^b$. Then in particular

$$\max_{v\in\Omega_\infty}\{|u_d\boldsymbol{\gamma}_d|_v\} \geq H^{b/n},$$

and thus $|\phi(\boldsymbol{\gamma})| \gg H^{b/n}$. We have shown that $\phi(M_>(H)) \subseteq M(H^{b/n})$. By Lemma 7.22, we obtain

$$I_>(\mathbf{u}, H) \asymp \int_{\phi(M_>(H))} \max_{d\in\Delta}\{|\boldsymbol{\nu}_d|^{n/s_d}\}^{-1}\,\mathrm{d}\boldsymbol{\nu} \ll \int_{M(H^{b/n})} \max_{d\in\Delta}\{|\boldsymbol{\nu}_d|^{n/s_d}\}^{-1}\,\mathrm{d}\boldsymbol{\nu} = I(H^{b/n}) \ll H^{-b\delta/n}.$$

□

Let

$$I_<(H) := \int_{M_<(H)} \max_{d\in\Delta}\{|\boldsymbol{\gamma}_d|^{1/s_d}\}^{-1}\,\mathrm{d}\boldsymbol{\gamma}.$$

We will prove that $I_<(H) \ll H^{-\delta}$ for some $\delta > 0$.

LEMMA 7.24. *For $A_1, \dots, A_n, B \in (0, \infty)$, let*

$$I(A_1, \dots, A_n, B) := \int_{\substack{0\leq x_i\leq A_i \\ x_1\cdots x_n\leq B}} \mathrm{d}x_1 \cdots \mathrm{d}x_n.$$

Then

$$I(A_1, \dots, A_n, B) \ll B\log\left(\frac{A_1\cdots A_n}{B} + 2\right)^{n-1}.$$

DÉMONSTRATION. Elementary computations using induction. □

LEMMA 7.25. *Let $\theta > 0$. For any small $\epsilon > 0$, and $H \geq 1$, we have*

$$\int_{\substack{\gamma\in V \\ N(\gamma)\leq H^b \\ |\gamma|\geq H}} \frac{1}{(1+|\gamma|)^{1+\theta}}\,\mathrm{d}\gamma \ll H^{-1-\theta+\epsilon+b}.$$

DÉMONSTRATION. For $v \in \Omega_\infty$, let $t_v := |\gamma|_v^{n_v}$, and assume that $\max_v |\gamma|_v = |\gamma|_v$. Passing to polar coordinates at the complex places, we see that the integral in the lemma is

$$\ll \int_{t_w \gg H^{n_w}} \frac{1}{t_w^{(1+\theta)/n_w}} \left(\int_{\substack{t_v, v\neq w \\ t_v \leq t_w^{n_v/n_w} \\ \prod_{v\neq w} t_v \leq H^b/t_w}} \prod_{v\neq w} \mathrm{d}t_v \right) \mathrm{d}t_w.$$

Using the notation of Lemma 7.24, the inner integral is just

$$I((t_w^{n_v/n_w})_{v\neq w}, H^b/t_w) \ll \frac{H^b}{t_w^{1-\epsilon}},$$

and thus the integral in the lemma is

$$\ll H^b \int_{t_w \gg H^{n_w}} \frac{1}{t_w^{(1+\theta)/n_w+1-\epsilon}}\,\mathrm{d}t_w \ll H^{b-1-\theta+n_w\epsilon}.$$

□

LEMMA 7.26. *Let $\theta > 0$. For any small $\epsilon > 0$, we have*

$$\int_{\substack{\gamma\in V\\ N(\gamma)\leq H^b}} \frac{1}{(1+|\gamma|)^{1+\theta}}\,\mathrm{d}\gamma \ll H^b.$$

DÉMONSTRATION. We start as in the proof of Lemma 7.25 and see that the integral is

$$\begin{aligned}
&\ll \int_{t_w=0}^{\infty} \frac{1}{(1+t_w^{1/n_w})^{(1+\theta)}} \cdot I((t_w^{n_v/n_w})_{v\neq w}, H^b/t_w)\,\mathrm{d}t_w.\\
&\ll \int_{t_w=0}^{H^{bn_w/n}} t_w^{\sum_{v\neq w} n_v/n_w}\,\mathrm{d}t_w + H^b \int_{t_w=H^{bn_w/n}}^{\infty} \frac{1}{t_w^{(1+\theta)/n_w+1-\epsilon}}\,\mathrm{d}t_w\\
&\ll H^b + H^{b(1-1/n-\theta/n+\epsilon n_w/n)} \ll H^b.
\end{aligned}$$

□

LEMMA 7.27. *Assume that $b \leq 1/T$. Then there is $\delta > 0$ such that, for $H \geq 1$,*

$$I_<(H) \ll H^{-\delta}.$$

DÉMONSTRATION. Let $d_0 \in \Delta$ and assume that $|\boldsymbol{\gamma}| = |\boldsymbol{\gamma}_{d_0}| \in (H, 2H]$. Since $|\boldsymbol{\gamma}| > H \geq 1$, we have

$$\max_{d\in\Delta}\{|\boldsymbol{\gamma}_d|^{1/s_d}\} \gg \max_{d\in\Delta}\{(1+|\boldsymbol{\gamma}_d|)^{1/s_d}\} \geq \prod_{d\in\Delta}(1+|\boldsymbol{\gamma}_d|)^{\lambda_d/s_d}$$

for any choice of $0 \leq \lambda_d \leq 1$ with $\sum_{d\in\Delta}\lambda_d = 1$. We choose

$$\lambda_d := s_d t_d + \theta/|\Delta|,$$

where $\theta := 1 - \sum_{d\in\Delta} s_d t_d \in (0,1)$. With $\theta_d := \theta/(|\Delta|\, s_d t_d)$, this gives

$$I_<(H) \ll \int_{\substack{\boldsymbol{\gamma}_{d_0}\in V^{t_{d_0}}\\ N(\boldsymbol{\gamma}_{d_0})\leq H^b\\ |\boldsymbol{\gamma}_{d_0}|\geq H}} \frac{1}{(1+|\boldsymbol{\gamma}_{d_0}|)^{t_{d_0}(1+\theta_{d_0})}}\,\mathrm{d}\boldsymbol{\gamma}_{d_0} \prod_{\substack{d\in\Delta\\ d\neq d_0}} \int_{\substack{\boldsymbol{\gamma}_d\in V^{t_d}\\ N(\boldsymbol{\gamma}_d)\leq H^b}} \frac{1}{(1+|\boldsymbol{\gamma}_d|)^{t_d(1+\theta_d)}}\,\mathrm{d}\boldsymbol{\gamma}_d.$$

Further assuming that $|\boldsymbol{\gamma}_{d_0}| = |\boldsymbol{\gamma}_{d_0,i_0}|$, we get

$$I_<(H) \ll \int_{\substack{\gamma_{d_0,i_0}\in V\\ N(\gamma_{d_0,i_0})\leq H^b\\ |\gamma_{d_0,i_0}|\geq H}} \frac{1}{(1+|\gamma_{d_0,i_0}|)^{1+\theta_{d_0}}}\,\mathrm{d}\gamma_{d_0,i_0} \prod_{\substack{d\in\Delta\\ 1\leq i\leq t_d\\ (d,i)\neq(d_0,i_0)}} \int_{\substack{\gamma_{d,i}\in V\\ N(\gamma_{d,i})\leq H^b}} \frac{1}{(1+|\gamma_{d,i}|)^{1+\theta_d}}\,\mathrm{d}\gamma_{d,i}.$$

By Lemma 7.25 and Lemma 7.26, this product is $\ll H^{-1-\theta_{d_0}+\epsilon+Tb} \ll H^{-\theta_{d_0}/2}$ if we choose ϵ small enough. □

With all our auxiliary results in place, we can now proceed to our main task, the estimation of the singular integral $\mathfrak{J}$.

LEMMA 7.28. *Assume* (5.6). *Then the integral defining* $\mathfrak{J}$ *converges absolutely and there is a positive constant* δ *such that*

$$\mathfrak{J} - \mathfrak{J}(H) \ll H^{-\delta}$$

holds for all large enough H.

DÉMONSTRATION. Let us fix $b := 1/T$. We have

$$\begin{aligned}\mathfrak{J} - \mathfrak{J}(H) &\ll \int_{|\boldsymbol{\gamma}|>H} |J(\boldsymbol{\gamma})| \, \mathrm{d}\boldsymbol{\gamma} = \sum_{j=0}^{\infty} \int_{2^j H<|\boldsymbol{\gamma}|\leq 2^{j+1}H} |J(\boldsymbol{\gamma})| \, \mathrm{d}\boldsymbol{\gamma} \\ &\ll \sum_{j=0}^{\infty} \left(\int_{M_<(2^j H)} |J(\boldsymbol{\gamma})| \, \mathrm{d}\boldsymbol{\gamma} + \int_{M_>(2^j H)} |J(\boldsymbol{\gamma})| \, \mathrm{d}\boldsymbol{\gamma} \right).\end{aligned}$$

We first consider the integrals over $M_<(2^j H)$. Here, we estimate $|J(\boldsymbol{\gamma})|$ by (6.2) and obtain

$$\sum_{j=0}^{\infty} \int_{M_<(2^j H)} |J(\boldsymbol{\gamma})| \, \mathrm{d}\boldsymbol{\gamma} \ll \sum_{j=0}^{\infty} (2^j H)^{\epsilon} I_<(2^j H) \ll H^{\epsilon-\delta} \sum_{j=0}^{\infty} 2^{j(\epsilon-\delta)} \ll H^{-\delta/2},$$

by Lemma 7.27, if ϵ was chosen small enough.

For the integrals over $M_>(2^j H)$, we use the estimates from *(1)* and *(2)* in Lemma 7.21 with $a = n$. We obtain

$$\sum_{j=0}^{\infty} \int_{M_>(2^j H)} |J(\boldsymbol{\gamma})| \, \mathrm{d}\boldsymbol{\gamma} \ll \sum_{j=0}^{\infty} (2^j H)^{\epsilon} \sum_{\substack{\mathbf{u}=(u_d)_{d\in\Delta} \\ u_d \in \mathcal{O}_K^{\times} \\ |u_d| \ll 2^j H}} I_>(\mathbf{u}, 2^j H).$$

By Lemma 7.23, we have $I_>(\mathbf{u}, 2^j) \ll (2^j H)^{-\delta}$. Moreover, it is well known that the number of units $u \in \mathcal{O}_K^{\times}$ with $|u| \ll 2^j H$ is is $\ll \log(2^j H)^{|\Omega_\infty|-1}$. Hence, the inner sum in the above expression has $\ll (2^j H)^{\epsilon}$ summands. Altogether, we see that

$$\sum_{j=0}^{\infty} \int_{M_>(2^j H)} |J(\boldsymbol{\gamma})| \, \mathrm{d}\boldsymbol{\gamma} \ll H^{2\epsilon-\delta} \sum_{j=0}^{\infty} 2^{j(2\epsilon-\delta')} \ll H^{-\delta/2},$$

if ϵ was chosen small enough. □

Our Theorem 7.1 is now an immediate consequence of the estimation of the minor arcs in Section 4 and the treatment of the major arcs in Lemma 7.17, Lemma 7.20, and Lemma 7.28.

Acknowledgements

We would like to thank Prof. Tim Browning, Prof. Jörg Brüdern and Dr. Damaris Schindler for helpful discussions, and Prof. Christopher Skinner for useful and encouraging remarks. The first-named author was supported by a Humboldt Research Fellowship for Postdoctoral Researchers of the Alexander von Humboldt Foundation. Major parts of the present work were established when the second-named author was a visitor of the Institut für Algebra, Zahlentheorie und Diskrete Mathematik at Leibniz Universität Hannover. He thanks the institute for its hospitality.

Chapitre 8

Construction of normal numbers via pseudo polynomial prime sequences

This chapter appeared in the *Acta Arithmetica*, **166** (2014), 81 – 99.

1. Introduction

Let $q \geq 2$ be a positive integer. Then every real $\theta \in [0,1)$ admits a unique expansion of the form

$$\theta = \sum_{k \geq 1} a_k q^k \quad (a_k \in \{0, \dots, q-1\})$$

called the q-ary expansion. We denote by $\mathcal{N}(\theta, d_1 \cdots d_\ell, N)$ the number of occurrences of the block $d_1 \cdots d_\ell$ amongst the first N digits, *i.e.*

$$\mathcal{N}(\theta, d_1 \cdots d_\ell, N) := \#\{0 \leq i < n \colon a_{i+1} = d_1, \dots, a_{i+\ell} = d_\ell\}.$$

Then we call a number normal of order ℓ in base q if for each block of length ℓ the frequency of occurrences tends to $q^{-\ell}$. As a qualitative measure of the distance of a number from being normal we introduce for integers N and ℓ the discrepancy of θ by

$$\mathcal{R}_{N,\ell}(\theta) = \sup_{d_1 \dots d_\ell} \left| \frac{\mathcal{N}(\theta, d_1 \cdots d_\ell, N)}{N} - q^{-k} \right|,$$

where the supremum is over all blocks of length ℓ. Then a number θ is normal to base q if for each $\ell \geq 1$ we have that $\mathcal{R}_{N,\ell}(\theta) = o(1)$ for $N \to \infty$. Furthermore we call a number absolutely normal if it is normal in all bases $q \geq 2$.

Borel [**43**] used a slightly different, but equivalent (*cf.* Chapter 4 of [**56**]), definition of normality to show that almost all real numbers are normal with respect to the Lebesgue measure. Despite their omnipresence it is not known whether numbers such as $\log 2$, π, e or $\sqrt{2}$ are normal to any base. The first construction of a normal number is due to Champernowne [**58**] who showed that the number

$$0.1\,2\,3\,4\,5\,6\,7\,8\,9\,10\,11\,12\,13\,14\,15\,16\,17\,18\,19\,20\dots$$

is normal in base 10.

The construction of Champernowne laid the base for a class of normal numbers which are of the form

$$\sigma_q = \sigma_q(f) = 0.\lfloor f(1) \rfloor_q \lfloor f(2) \rfloor_q \lfloor f(3) \rfloor_q \lfloor f(4) \rfloor_q \lfloor f(5) \rfloor_q \lfloor f(6) \rfloor_q \dots,$$

where $\lfloor \cdot \rfloor_q$ denotes the expansion in base q of the integer part. Davenport and Erdős [**63**] showed that $\sigma(f)$ is normal for f being a polynomial such that $f(\mathbb{N}) \subset \mathbb{N}$. This construction was extended by Schiffer [**214**] to polynomials with rational coefficients. Furthermore he showed that for these polynomials the discrepancy $\mathcal{R}_{N,\ell}(\sigma(f)) \ll (\log N)^{-1}$ and that this is best possible. These results where extended by Nakai and Shiokawa [**167**] to polynomials having

real coefficients. Madritsch, Thuswaldner and Tichy [**146**] considered transcendental entire functions of bounded logarithmic order. Nakai and Shiokawa [**166**] used pseudo-polynomial functions, *i.e.* these are function of the form

$$f(x) = \alpha_0 x^{\beta_0} + \alpha_1 x^{\beta_1} + \cdots + \alpha_d x^{\beta_d} \tag{1.1}$$

with $\alpha_0, \beta_0, \alpha_1, \beta_1, \ldots, \alpha_d, \beta_d \in \mathbb{R}$, $\alpha_0 > 0$, $\beta_0 > \beta_1 > \cdots > \beta_d > 0$ and at least one $\beta_i \notin \mathbb{Z}$. Since we often only need the leading term we write $\alpha = \alpha_0$ and $\beta = \beta_0$ for short. They were also able to show that the discrepancy is $\mathcal{O}((\log N)^{-1})$. We refer the interested reader to the books of Kuipers and Niederreiter [**131**], Drmota and Tichy [**66**] or Bugeaud [**56**] for a more complete account on the construction of normal numbers.

The present method of construction by concatenating function values is in strong connection with properties of q-additive functions. We call a function f strictly q-additive, if $f(0) = 0$ and the function operates only on the digits of the q-ary representation, i.e.,

$$f(n) = \sum_{h=0}^{\ell} f(d_h) \quad \text{for} \quad n = \sum_{h=0}^{\ell} d_h q^h.$$

A very simple example of a strictly q-additive function is the sum of digits function s_q, defined by

$$s_q(n) = \sum_{h=0}^{\ell} d_h \quad \text{for} \quad n = \sum_{h=0}^{\ell} d_h q^h.$$

Refining the methods of Nakai and Shiokawa [**166**] the author obtained the following result.

THEOREM 8.1 ([**149**, Theorem 1.1]). *Let $q \geq 2$ be an integer and f be a strictly q-additive function. If p is a pseudo-polynomial as defined in (1.1), then there exists $\eta > 0$ such that*

$$\sum_{n \leq N} f\left(\lfloor p(n) \rfloor\right) = \mu_f N \log_q(p(N)) + NF\left(\log_q(p(N))\right) + \mathcal{O}\left(N^{1-\eta}\right),$$

where

$$\mu_f = \frac{1}{q} \sum_{d=0}^{q-1} f(d)$$

and F is a 1-periodic function depending only on f and p.

In the present paper, however, we are interested in a variant of $\sigma_q(f)$ involving primes. As a first example, Champernowne [**58**] conjectured and later Copeland and Erdős [**60**] proved that the number

$$0.2\,3\,5\,7\,11\,13\,17\,19\,23\,29\,31\,37\,41\,43\,47\,53\,59\,61\,67\ldots$$

is normal in base 10. Similar to the construction above we want to consider the number

$$\tau_q = \tau_q(f) = 0.\lfloor f(2) \rfloor_q \lfloor f(3) \rfloor_q \lfloor f(5) \rfloor_q \lfloor f(7) \rfloor_q \lfloor f(11) \rfloor_q \lfloor f(13) \rfloor_q \ldots,$$

where the arguments of f run through the sequence of primes.

Then the paper of Copeland and Erdős corresponds to the function $f(x) = x$. Nakai and Shiokawa [**168**] showed that the discrepancy for polynomials having rational coefficients is $\mathcal{O}((\log N)^{-1})$. Furthermore Madritsch, Thuswaldner and Tichy [**146**] showed, that transcendental entire functions of bounded logarithmic order yield normal numbers. Finally in a

recent paper Madritsch and Tichy [**147**] considered pseudo-polynomials of the special form αx^β with $\alpha > 0$, $\beta > 1$ and $\beta \notin \mathbb{Z}$.

The aim of the present paper is to extend this last construction to arbitrary pseudo-polynomials. Our first main result is the following

THEOREM 8.2. *Let f be a pseudo-polynomial as in (1.1). Then*

$$\mathcal{R}_N(\tau_q(f)) \ll (\log N)^{-1}.$$

In our second main result we use the connection of this construction of normal numbers with the arithmetic mean of q-additive functions as described above. Known results are due to Shiokawa [**224**] and Madritsch and Tichy [**147**]. Similar results concerning the moments of the sum of digits function over primes have been established by Kátai [**116**].

Let $\pi(x)$ stand for the number of primes less than or equal to x. Then adapting these ideas to our method we obtain the following

THEOREM 8.3. *Let f be a pseudo-polynomial as in (1.1). Then*

$$\sum_{p\le P} s_q(\lfloor f(p)\rfloor) = \frac{q-1}{2}\pi(P)\log_q P^\beta + \mathcal{O}(\pi(P)),$$

where the sum runs over the primes and the implicit $\mathcal{O}$-constant may depend on q and β.

REMARK 15. With simple modifications Theorem 8.3 can be extended to completely q-additive functions replacing s_q.

The proof of the two theorems is divided into four parts. In the following section we rewrite both statements in order to obtain as a common base the central theorem – Theorem 8.4. In Section 3 we start with the proof of this central theorem by using an indicator function and its Fourier series. These series contain exponential sums which we treat by different methods (with respect to the position in the expansion) in Section 4. Finally, in Section 5 we put the estimates together in order to proof the central theorem and therefore our two statements.

2. Preliminaries

Throughout the rest p will always denote a prime. The implicit constant of $\ll$ and $\mathcal{O}$ may depend on the pseudo-polynomial f and on the parameter $\varepsilon > 0$. Furthermore we fix a block $d_1 \cdots d_\ell$ of length ℓ and N, the number of digits we consider.

In the first step we want to know in the expansion of which prime the N-th digit occurs. This can be seen as the translation from the digital world to the world of blocks. To this end let $\ell(m)$ denote the length of the q-ary expansion of an integer m. Then we define an integer P by

$$\sum_{p\le P-1} \ell\left(\lfloor f(p)\rfloor\right) < N \le \sum_{p\le P} \ell\left(\lfloor f(p)\rfloor\right),$$

where the sum runs over all primes. Thus we get the following relation between N and P

$$\begin{aligned} N &= \sum_{p\le P} \ell(\lfloor f(p)\rfloor) + \mathcal{O}(\pi(P)) + \mathcal{O}(\beta\log_q(P)) \\ &= \frac{\beta}{\log q}P + \mathcal{O}\left(\frac{P}{\log P}\right). \end{aligned} \tag{2.1}$$

Here we have used the prime number theorem in the form (*cf.* [**237**, Théorème 4.1])

$$\pi(x) = \operatorname{Li} x + \mathcal{O}\left(\frac{x}{(\log x)^G}\right), \tag{2.2}$$

where G is an arbitrary positive constant and

$$\operatorname{Li} x = \int_2^x \frac{\mathrm{d}t}{\log t}.$$

Now we show that we may neglect the occurrences of the block $d_1 \cdots d_\ell$ between two expansions. We write $\mathcal{N}(f(p))$ for the number of occurrences of this block in the q-ary expansion of $\lfloor f(p) \rfloor$. Then (2.1) implies that

$$\left| \mathcal{N}(\tau_q(f); d_1 \cdots d_\ell; N) - \sum_{p \le P} \mathcal{N}(f(p)) \right| \ll \frac{N}{\log N}. \tag{2.3}$$

In the next step we use the polynomial-like behavior of f. In particular, we collect all the values having the same length of expansion. Let j_0 be a sufficiently large integer. Then for each integer $j \ge j_0$ there exists a P_j such that

$$q^{j-2} \le f(P_j) < q^{j-1} \le f(P_j + 1) < q^j$$

with

$$P_j \asymp q^{\frac{j}{\beta}}.$$

Furthermore we set J to be the greatest length of the q-ary expansions of $f(p)$ over the primes $p \le P$, i.e.,

$$J := \max_{p \le P} \ell(\lfloor f(p) \rfloor) = \log_q(f(P)) + \mathcal{O}(1) \asymp \log P.$$

Now we show that we may suppose that each expansion has the same length (which we reach by adding leading zeroes). For $P_{j-1} < p \le P_j$ we may write $f(p)$ in q-ary expansion, i.e.,

$$f(p) = b_{j-1} q^{j-1} + b_{j-2} q^{j-2} + \cdots + b_1 q + b_0 + b_{-1} q^{-1} + \ldots. \tag{2.4}$$

Then we denote by $\mathcal{N}^*(f(p))$ the number of occurrences of the block $d_1 \cdots d_\ell$ in the string $0 \cdots 0 b_{j-1} b_{j-2} \cdots b_1 b_0$, where we filled up the expansion with leading zeroes such that it has length J. The error of doing so can be estimated by

$$\begin{aligned} 0 &\le \sum_{p \le P} \mathcal{N}^*(f(p)) - \sum_{p \le P} \mathcal{N}(f(p)) \\ &\le \sum_{j=j_0+1}^{J-1} (J - j)\,(\pi(P_{j+1}) - \pi(P_j)) + \mathcal{O}(1) \\ &\le \sum_{j=j_0+2}^{J} \pi(P_j) + \mathcal{O}(1) \ll \sum_{j=j_0+2}^{J} \frac{q^{j/\beta}}{j} \ll \frac{P}{\log P} \ll \frac{N}{\log N}. \end{aligned}$$

In the following three sections we will estimate this sum of indicator functions $\mathcal{N}^*$ in order to prove the following theorem.

THEOREM 8.4. *Let f be a pseudo polynomial as in* (1.1). *Then*

$$\sum_{p\le P} \mathcal{N}^*\left(\lfloor f(p)\rfloor\right) = q^{-\ell}\pi(P)\log_q P^\beta + \mathcal{O}\left(\frac{P}{\log P}\right) \tag{2.5}$$

Using this theorem we can simply deduce our two main results.

PROOF OF THEOREM 8.2. We insert (2.5) into (2.3) and get the desired result. □

PROOF OF THEOREM 8.3. For this proof we have to rewrite the statement. In particular, we use that the sum of digits function counts the number of 1s, 2s, etc. and assigns weights to them, i.e.,

$$s_q(n) = \sum_{d=0}^{q-1} d\cdot \mathcal{N}(n;d).$$

Thus

$$\begin{aligned}\sum_{p\le P} s_q(\left\lfloor p^\beta\right\rfloor) &= \sum_{p\le P}\sum_{d=0}^{q-1} d\cdot\mathcal{N}(p^\beta) = \sum_{p\le P}\sum_{d=0}^{q-1} d\cdot\mathcal{N}^*(p^\beta) + \mathcal{O}\left(\frac{P}{\log P}\right)\\ &= \frac{q-1}{2}\pi(P)\log_q(P^\beta) + \mathcal{O}\left(\frac{P}{\log P}\right)\end{aligned}$$

and the theorem follows. □

In the following sections we will prove Theorem 8.4 in several steps. First we use the "method of little glasses" in order to approximate the indicator function by a Fourier series having smooth coefficients. Then we will apply different methods (depending on the position in the expansion) for the estimation of the exponential sums that appear in the Fourier series. Finally we put everything together and get the desired estimate.

3. Proof of Theorem 8.4, Part I

We want to ease notation by splitting the pseudo-polynomial f into a polynomial and the rest. Then there exists a unique decomposition of the following form :

$$f(x) = g(x) + h(x) \tag{3.1}$$

where $h\in\mathbb{R}[X]$ is a polynomial of degree k (where we set $k=0$ if h is the zero polynomial) and

$$g(x) = \sum_{j=1}^{r}\alpha_j x^{\theta_j}$$

with $r\ge 1$, $\alpha_r\neq 0$, α_j real, $0<\theta_1<\cdots<\theta_r$ and $\theta_j\notin\mathbb{Z}$ for $1\le j\le r$.

Let γ and ρ be two parameter which we will frequently use in the sequel. We suppose that

$$0<\gamma<\rho<\min\left(\frac{1}{4(k+1)},\frac{\theta_r}{2}\right).$$

The aim of this section is to calculate the Fourier transform of $\mathcal{N}^*$. In order to count the occurrences of the block $d_1\cdots d_\ell$ in the q-ary expansion of $\lfloor f(p)\rfloor$ $(2\le p\le P)$ we define the indicator function

$$\mathcal{I}(t) = \begin{cases}1, & \text{if } \sum_{i=1}^{\ell} d_i q^{-i} \le t-\lfloor t\rfloor < \sum_{i=1}^{\ell} d_i q^{-i} + q^{-\ell};\\ 0, & \text{otherwise};\end{cases}$$

which is a 1-periodic function. Indeed, we have

$$\mathcal{I}(q^{-j}f(p)) = 1 \Longleftrightarrow d_1 \cdots d_\ell = b_{j-1} \cdots b_{j-\ell}, \tag{3.2}$$

where $f(p)$ has an expansion as in (2.4). Thus we may write our block counting function as follows

$$\mathcal{N}^*(f(p)) = \sum_{j=\ell}^{J} \mathcal{I}\left(q^{-j}f(p)\right). \tag{3.3}$$

In the following we will use Vinogradov's "method of little glasses" (*cf.* [**247**]). We want to approximate $\mathcal{I}$ from above and from below by two 1-periodic functions having small Fourier coefficients. To this end we will use the following

LEMMA 8.5 ([**247**, Lemma 12]). *Let α, β, Δ be real numbers satisfying*

$$0 < \Delta < \frac{1}{2}, \quad \Delta \le \beta - \alpha \le 1 - \Delta.$$

Then there exists a periodic function $\psi(x)$ with period 1, satisfying

(1) $\psi(x) = 1$ in the interval $\alpha + \frac{1}{2}\Delta \le x \le \beta - \frac{1}{2}\Delta$,

(2) $\psi(x) = 0$ in the interval $\beta + \frac{1}{2}\Delta \le x \le 1 + \alpha - \frac{1}{2}\Delta$,

(3) $0 \le \psi(x) \le 1$ in the remainder of the interval $\alpha - \frac{1}{2}\Delta \le x \le 1 + \alpha - \frac{1}{2}\Delta$,

(4) $\psi(x)$ has a Fourier series expansion of the form

$$\psi(x) = \beta - \alpha + \sum_{\substack{\nu=-\infty \\ \nu\neq 0}}^{\infty} A(\nu)e(\nu x),$$

where

$$|A(\nu)| \ll \min\left(\frac{1}{\nu}, \beta - \alpha, \frac{1}{\nu^2\Delta}\right). \tag{3.4}$$

We note that we could have used Vaaler polynomials [**242**], however, we do not gain anything by doing so as the estimates we get are already best possible. Setting

$$\delta = P^{-\gamma}, \qquad \begin{aligned} \alpha_- &= \sum_{\lambda=1}^{\ell} d_\lambda q^{-\lambda} + (2\delta)^{-1}, \quad \beta_- = \sum_{\lambda=1}^{\ell} d_\lambda q^{-\lambda} + q^{-\ell} - (2\delta)^{-1}, \\ \alpha_+ &= \sum_{\lambda=1}^{\ell} d_\lambda q^{-\lambda} - (2\delta)^{-1}, \quad \beta_+ = \sum_{\lambda=1}^{\ell} d_\lambda q^{-\lambda} + q^{-\ell} + (2\delta)^{-1}. \end{aligned} \tag{3.5}$$

and an application of Lemma 8.5 with $(\alpha, \beta, \delta) = (\alpha_-, \beta_-, \delta)$ and $(\alpha, \beta, \delta) = (\alpha_+, \beta_+, \delta)$, respectively, provides us with two functions $\mathcal{I}_-$ and $\mathcal{I}_+$. By our choice of $(\alpha_\pm, \beta_\pm, \delta)$ it is immediate that

$$\mathcal{I}_-(t) \le \mathcal{I}(t) \le \mathcal{I}_+(t) \qquad (t \in \mathbb{R}). \tag{3.6}$$

Lemma 8.5 also implies that these two functions have Fourier expansions

$$\mathcal{I}_\pm(t) = q^{-\ell} \pm P^{-\gamma} + \sum_{\substack{\nu=-\infty \\ \nu\neq 0}}^{\infty} A_\pm(\nu)e(\nu t) \tag{3.7}$$

satisfying

$$|A_\pm(\nu)| \ll \min(|\nu|^{-1}, P^\gamma |\nu|^{-2}).$$

In a next step we want to replace $\mathcal{I}$ by $\mathcal{I}_+$ in (3.3). For this purpose we observe, using (3.6), and (3.7) that

$$\left|\mathcal{I}(t) - q^{-\ell}\right| \ll P^{-\gamma} + \sum_{\substack{\nu=-\infty \\ \nu\neq 0}}^{\infty} A_\pm(\nu) e(\nu t).$$

Thus setting $t = q^{-j} f(p)$ and summing over $p \leq P$ yields

$$\left|\sum_{p\leq P} \mathcal{I}(q^{-j} f(p)) - \frac{\pi(P)}{q^\ell}\right| \ll \pi(P) P^{-\gamma} + \sum_{\substack{\nu=-\infty \\ \nu\neq 0}}^{\infty} A_\pm(\nu) \sum_{p\leq P} e\left(\frac{\nu}{q^j} f(p)\right). \tag{3.8}$$

Now we consider the coefficients $A_\pm(\nu)$. Noting (3.4) one observes that

$$A_\pm(\nu) \ll \begin{cases} \nu^{-1}, & \text{for } |\nu| \leq P^\gamma; \\ P^\gamma \nu^{-2}, & \text{for } |\nu| > P^\gamma. \end{cases}$$

Estimating all summands with $|\nu| > P^\gamma$ trivially we get

$$\sum_{\substack{\nu=-\infty \\ \nu\neq 0}}^{\infty} A_\pm(\nu) e\left(\frac{\nu}{q^j} f(p)\right) \ll \sum_{\nu=1}^{P^\gamma} \nu^{-1} e\left(\frac{\nu}{q^j} f(p)\right) + P^{-\gamma}.$$

Using this in (3.8) yields

$$\left|\sum_{p\leq P} \mathcal{I}(q^{-j} f(p)) - \frac{\pi(P)}{q^\ell}\right| \ll \pi(P) P^{-\gamma} + \sum_{\nu=1}^{P^\gamma} \nu^{-1} S(P, j, \nu),$$

where we have set

$$S(P, j, \nu) := \sum_{p\leq P} e\left(\frac{\nu}{q^j} f(p)\right). \tag{3.9}$$

4. Exponential sum estimates

In the present section we will focus on the estimation of the sum $S(P, j, \nu)$ for different ranges of j. Since j describes the position within the q-ary expansion of $f(p)$ we will call these ranges the "most significant digits", the "least significant digits" and the "digits in the middle", respectively.

Now, if $\theta_r > k \geq 0$, *i.e* the leading coefficient of f origins from the pseudo polynomial part g, then we consider the two ranges

$$1 \leq q^j \leq P^{\theta_r - \rho} \quad \text{and} \quad P^{\theta_r - \rho} < q^j \leq P^{\theta_r}.$$

For the first one we will apply Proposition 8.8 and for the second one Proposition 8.6.

On the other hand, if $k > \theta_r > 0$, meaning that the leading coefficient of f origins from the polynomial part h, then we have an additional part. In particular, in this case we will consider the three ranges

$$1 \leq q^j \leq P^{\theta_r - \rho}, \quad P^{\theta_r - \rho} < q^j \leq P^{k-1+\rho}, \quad \text{and} \quad P^{k-1+\rho} < q^j \leq P^k.$$

We will, similar to above, treat the first and last range by Proposition 8.8 and Proposition 8.6, respectively. For the middle range we will apply Proposition 8.12. Since $2\rho < \theta_r$, we note that the middle range is empty if $k = 1$.

Since the size of j represents the position of the digit in the expansion (*cf.* (3.2)), we will deal in the following subsection with the "most significant digits", the "least significant digits" and the "digits in the middle", respectively.

4.1. Most significant digits. We start our series of estimates for the exponential sum $S(P, j, \nu)$ for j being in the highest range. In particular, we want to show the following

PROPOSITION 8.6. *Suppose that for some $k \geq 1$ we have $\left|f^{(k)}(x)\right| \geq \Lambda$ for any x on $[a, b]$ with $\Lambda > 0$. Then*

$$S(P, j, \nu) \ll \frac{1}{\log P}\Lambda^{-\frac{1}{k}} + \frac{P}{(\log P)^G}.$$

The main idea of the proof is to use Riemann-Stieltjes integration together with

LEMMA 8.7 ([**114**, Lemma 8.10]). *Let $F\colon [a, b] \to \mathbb{R}$ and suppose that for some $k \geq 1$ we have $\left|F^{(k)}(x)\right| \geq \Lambda$ for any x on $[a, b]$ with $\Lambda > 0$. Then*

$$\left|\int_a^b e(F(x))\mathrm{d}x\right| \leq k2^k\Lambda^{-1/k}.$$

PROOF OF PROPOSITION 8.6. We rewrite the sum into a Riemann-Stieltjes integral :

$$S(P, j, \nu) = \sum_{p \leq P} e\left(\frac{\nu}{q^j}f(p)\right) = \int_2^P e\left(\frac{\nu}{q^j}f(t)\right)\mathrm{d}\pi(t) + \mathcal{O}(1).$$

Then we apply the prime number theorem in the form (2.2) to gain the usual integral back. Thus

$$S(P, j, \nu) = \int_{P(\log P)^{-G}}^P e\left(\frac{\nu}{q^j}f(t)\right)\frac{\mathrm{d}t}{\log t} + \mathcal{O}\left(\frac{P}{(\log P)^G}\right).$$

Now we use the second mean-value theorem to get

$$S(P, j, \nu) \ll \frac{1}{\log P}\sup_{\xi}\left|\int_{P(\log P)^{-G}}^{\xi} e\left(\frac{\nu}{q^j}f(t)\right)\mathrm{d}t\right| + \frac{P}{(\log P)^G}. \tag{4.1}$$

Finally an application of Lemma 8.7 proves the lemma. □

4.2. Least significant digits. Now we turn our attention to the lowest range of j. In particular, the goal is the proof of the following

PROPOSITION 8.8. *Let P and ρ be positive reals and f be a pseudo-polynomial as in* (3.1). *If j is such that*

$$1 \leq q^j \leq P^{\theta_r - \rho} \tag{4.2}$$

holds, then for $1 \leq \nu \leq P^\gamma$ there exists $\eta > 0$ (depending only on f and ρ) such that

$$S(P, j, \nu) \ll (\log P)^8 P^{1-\eta}.$$

Before we launch into the proof we collect some tools that will be necessary in the sequel. A standard idea for estimating exponential sums over the primes is to rewrite them into ordinary exponential sums over the integers having von Mangoldt's function as weights and then to apply Vaughan's identity. We denote by

$$\Lambda(n) = \begin{cases} \log p, & \text{if } n = p^k \text{ for some prime } p \text{ and an integer } k \geq 1\,; \\ 0, & \text{otherwise.} \end{cases}$$

von Mangoldt's function. For the rewriting process we use the following

LEMMA 8.9. *Let g be a function such that $|g(n)| \leq 1$ for all integers n. Then*

$$\left|\sum_{p\leq P} g(p)\right| \ll \frac{1}{\log P} \max_{t\leq P} \left|\sum_{n\leq t} \Lambda(n)g(n)\right| + \sqrt{P}.$$

DÉMONSTRATION. This is Lemma 11 of [**159**]. However, the proof is short and we need some piece later.

We start with a summation by parts yielding

$$\sum_{p\leq P} g(p) = \frac{1}{\log P} \sum_{p\leq x} \log(p)g(p) + \int_2^P \left(\sum_{p\leq t} \log(p)g(p)\right) \frac{\mathrm{d}t}{t(\log t)^2}.$$

Now we cut the integral at $\sqrt{P}$ and use Chebyshev's inequality (*cf.* [**237**, Théorème 1.3]) in the form $\sum_{p\leq t} \log(p) \leq \log(t)\pi(t) \ll t$ for the lower part. Thus

$$\begin{aligned} \left|\sum_{p\leq P} g(p)\right| &\leq \left(\frac{1}{\log P} + \int_{\sqrt{P}}^P \frac{\mathrm{d}t}{t(\log t)^2}\right) \max_{\sqrt{P}<t\leq P} \left|\sum_{p\leq P} \log(p)g(p)\right| + \mathcal{O}(\sqrt{P}) \\ &= \frac{2}{\log P} \max_{\sqrt{P}<t\leq P} \left|\sum_{p\leq t} \log(p)g(p)\right| + \mathcal{O}(\sqrt{P}). \end{aligned}$$

Finally we again use Chebyshev's inequality $\pi(t) \ll t/\log(t)$ to obtain

$$\left|\sum_{n\leq t} \Lambda(n)g(n) - \sum_{p\leq t} \log(p)g(p)\right| \leq \sum_{p\leq \sqrt{t}} \log(p) \sum_{a=2}^{\left\lfloor \frac{\log(t)}{\log(p)} \right\rfloor} 1 \leq \pi(\sqrt{t})\log(t) \ll \sqrt{t}. \tag{4.3}$$

□

In the next step we use Vaughan's identity to subdivide this weighted exponential sum into several sums of Type I and II.

LEMMA 8.10 ([**30**, Lemma 2.3]). *Assume $F(x)$ to be any function defined on the real line, supported on $[P/2, P]$ and bounded by F_0. Let further U, V, Z be any parameters satisfying $3 \leq U < V < Z < P$, $Z \geq 4U^2$, $P \geq 64Z^2U$, $V^3 \geq 32P$ and $Z - \frac{1}{2} \in \mathbb{N}$. Then*

$$\left|\sum_{P/2<n\leq P} \Lambda(n)F(n)\right| \ll K\log P + F_0 + L(\log P)^8,$$

where K and L are defined by

$$K = \max_M \sum_{m=1}^{\infty} d_3(m) \left| \sum_{Z<n\leq M} F(mn) \right|,$$
$$L = \sup \sum_{m=1}^{\infty} d_4(m) \left| \sum_{U<n<V} b(n)F(mn) \right|,$$

where the supremum is taken over all arithmetic functions $b(n)$ satisfying $|b(n)| \leq d_3(n)$.

After subdividing the weighted exponential sum with Vaughan's identity we will use the following lemma in order to estimate the occurring exponential sums.

LEMMA 8.11 ([**30**, Lemma 2.5]). *Let $X, k, q \in \mathbb{N}$ with $k, q \geq 0$ and set $K = 2^k$ and $Q = 2^q$. Let $h(x)$ be a polynomial of degree k with real coefficients. Let $g(x)$ be a real $(q+k+2)$ times continuously differentiable function on $[X/2, X]$ such that $\left|f^{(r)}(x)\right| \asymp FX^{-r}$ $(r = 1, \ldots, q+k+2)$. Then, if $F = o(X^{q+2})$ for F and X large enough, we have*

$$\left| \sum_{X/2<x\leq X} e(g(x)+h(x)) \right| \ll X^{1-\frac{1}{K}} + X\left(\frac{\log^k X}{F}\right)^{\frac{1}{K}} + X\left(\frac{F}{X^{q+2}}\right)^{\frac{1}{(4KQ-2K)}}.$$

Now we have the necessary tools to state the

PROOF OF PROPOSITION 8.8. An application of Lemma 8.9 yields

$$S(P, j, \nu) \ll \frac{1}{\log P} \max \left| \sum_{n\leq P} \Lambda(n) e\left(\frac{\nu}{q^j}(g(n)+h(n))\right) \right| + P^{\frac{1}{2}}.$$

We split the inner sum into $\leq \log P$ sub sums of the form

$$\left| \sum_{X<n\leq 2X} \Lambda(n) e\left(\frac{\nu}{q^j}(g(n)+h(n))\right) \right|$$

with $2X \leq P$ and let S be a typical one of them. We may assume that $X \geq P^{1-\rho}$.

Using Vaughan's identity (Lemma 8.10) with $U = \frac{1}{4}X^{1/5}$, $V = 4X^{1/3}$ and Z the unique number in $\frac{1}{2} + \mathbb{N}$, which is closest to $\frac{1}{4}X^{2/5}$, we obtain

$$S \ll 1 + (\log X)S_1 + (\log X)^8 S_2, \tag{4.4}$$

where

$$S_1 = \sum_{x<\frac{2X}{Z}} d_3(x) \sum_{y>Z,\, \frac{X}{x}<y<\frac{2X}{x}} e\left(\frac{\nu}{q^j}(g(xy)+h(xy))\right)$$
$$S_2 = \sum_{\frac{X}{V}<x\leq\frac{2X}{U}} d_4(x) \sum_{U<y<V,\, \frac{X}{x}<y\leq\frac{2X}{x}} b(y) e\left(\frac{\nu}{q^j}(g(xy)+h(xy))\right)$$

We start with the estimation of S_1. Since $d_3(x) \ll x^{\varepsilon}$ we have for

$$|S_1| \ll X^{\varepsilon} \sum_{x \le \frac{2X}{Z}} \left| \sum_{\substack{\frac{X}{x} < y \frac{2X}{x} \\ y > Z}} e\left(\frac{\nu}{q^j}(g(xy) + h(xy))\right)\right|.$$

For estimating the inner sum we fix x and denote $Y = \frac{X}{x}$. Since $\theta_r \notin \mathbb{Z}$ and $\theta_r > k \ge 0$, we have that

$$\left|\frac{\partial^{\ell} g(xy)}{\partial y^{\ell}}\right| \asymp X^{\theta_r} Y^{-\ell}.$$

Now on the one hand, since $q^j \le P^{\theta_r - \rho}$, we have $\nu q^{-j} X^{\theta_r} \gg X^{\rho}$. On the other hand for $\ell \ge 5(\lfloor \theta_r \rfloor + 1)$ we get

$$\frac{\nu}{q^j} X^{\theta_r} Y^{-\ell} \le P^{\gamma} X^{\theta_r - \frac{2}{5}\ell} \ll X^{-\frac{1}{2}}.$$

Thus an application of Lemma 8.11 yields the following estimate :

$$\begin{aligned} |S_1| &\ll X^{\varepsilon} \sum_{x \le 2X/Z} Y \left[Y^{-\frac{1}{K}} + (\log Y)^k X^{-\frac{\rho}{K}} + X^{-\frac{1}{2}\frac{1}{4K \cdot 8L^5 - 2K}} \right] \\ &\ll X^{1+\varepsilon} (\log X) \left(X^{-\rho} + X^{-\frac{1}{64L^5 - 4}} \right)^{\frac{1}{K}}, \end{aligned} \tag{4.5}$$

where we have used that $\frac{k}{K} < 1$ and $\rho < \frac{1}{3}$.

For the second sum S_2 we start by splitting the interval $(\frac{X}{V}, \frac{2X}{U}]$ into $\le \log X$ subintervals of the form $(X_1, 2X_1]$. Thus

$$|S_2| \le (\log X) X^{\varepsilon} \sum_{X_1 < x \le 2X_1} \left| \sum_{\substack{U < y < V \\ \frac{X}{x} < y \le \frac{2X}{x}}} b(y) e\left(\frac{\nu}{q^j}(g(xy) + h(xy))\right)\right|$$

Now an application of Cauchy's inequality together with $|b(y)| \ll X^{\varepsilon}$ yields

$$\begin{aligned} |S_2|^2 &\le (\log X)^2 X^{2\varepsilon} X_1 \sum_{X_1 < x \le 2X_1} \left| \sum_{\substack{U < y < V \\ \frac{X}{x} < y \le \frac{2X}{x}}} b(y) e\left(\frac{\nu}{q^j}(g(xy) + h(xy))\right)\right|^2 \\ &\ll (\log X)^2 X^{4\varepsilon} X_1 \\ &\quad \times \left(X_1 \frac{X}{X_1} + \left| \sum_{X_1 < x \le 2X_1} \sum_{A < y_1 < y_2 \le B} e\left(\frac{\nu}{q^j}(g(xy_1) - g(xy_2) + h(xy_1) - h(xy_2))\right)\right| \right) \end{aligned}$$

where $A = \max\{U, \frac{X}{x}\}$ and $B = \min\{U, \frac{2X}{x}\}$. Changing the order of summation, we get

$$\begin{aligned} |S_2|^2 &\ll (\log X)^2 X^{4\varepsilon} X_1 \\ &\quad \times \left(X + \sum_{A < y_1 < y_2 \le B} \left| \sum_{X_1 < x \le 2X_1} e\left(\frac{\nu}{q^j}(g(xy_1) - g(xy_2) + h(xy_1) - h(xy_2))\right)\right| \right) \end{aligned}$$

As above we want to apply Lemma 8.11. To this end we fix y_1 and $y_2 \neq y_1$. Similarly to above we get that

$$\left|\frac{\partial^\ell \left(g(xy_1) - g(xy_2) + h(xy_1) - h(xy_2)\right)}{\partial x^\ell}\right| \asymp \frac{|y_1 - y_2|}{y_1} X^{\theta_r} X_1^{-\ell}.$$

Now, on the one hand we have $\frac{\nu}{q^j}\frac{|y_1-y_2|}{y_1} X^{\theta_r} \gg X^\rho$ and on the other hand

$$\frac{\nu}{q^j}\frac{|y_1 - y_2|}{y_1} X^{\theta_r} X_1^{-\ell} \ll X^{\gamma+\theta_r}\left(\frac{X}{V}\right)^{-\ell} \ll X^{\gamma+\theta_r-\frac{2}{3}\ell} \ll X^{-\frac{1}{2}}$$

if $\ell \geq 2\lfloor \theta_r \rfloor + 3$. Thus again an application of Lemma 8.11 yields

$$\begin{aligned} |S_2|^2 &\ll (\log X)^2 X^{4\varepsilon} X_1 \left(X + \sum_{A<y_1<y_2\leq B} X_1 \left(X_1^{-\frac{1}{K}} + X^{-\frac{\rho}{K}} + X^{-\frac{1}{2}\frac{1}{4K\cdot 2L^2-2K}} \right) \right) \\ &\ll (\log X)^2 X^{4\varepsilon} \left(X^{\frac{5}{3}} + X^{2-\frac{\rho}{K}} + X^{2-\frac{1}{16KL^2-4K}} \right). \end{aligned} \tag{4.6}$$

Plugging the two estimates (4.5) and (4.6) into (4.4) proves the proposition. □

4.3. The digits in the middle. Now we are getting more involved in order to consider those j leading to a position between θ_r and k. These sums correspond to the "digits in the middle" in the proof of Theorem 8.4. We want to prove the following

PROPOSITION 8.12. *Let P and ρ be positive reals and f be a pseudo-polynomial as in* (3.1). *If $2\rho < \theta_r < k$ and j is such that*

$$P^{\theta_r-\rho} < q^j \leq P^{k-1+\rho} \tag{4.7}$$

holds, then for $1 \leq \nu \leq P^\gamma$ we have

$$S(P, j, \nu) = \sum_{p\leq P} e\left(\frac{\nu f(p)}{q^j}\right) \ll P^{1-\frac{\rho}{4^k}}.$$

The main idea in this range is to use that the dominant part of f comes from the polynomial h. Therefore after getting rid of the function g we will estimate the sum over the polynomial by the following

LEMMA 8.13. *Let $h \in \mathbb{R}[X]$ be a polynomial of degree $k \geq 2$. Suppose α is the leading coefficient of h and that there are integers a, q such that*

$$|q\alpha - a| < \frac{1}{q} \quad \textit{with} \quad (a, q) = 1.$$

Then we have for any $\varepsilon > 0$ and $H \leq X$

$$\sum_{X<p\leq X+H} \log(p) e(h(p)) \ll H^{1+\varepsilon}\left(\frac{1}{q} + \frac{1}{H^{\frac{1}{2}}} + \frac{q}{H^k}\right)^{4^{1-k}}.$$

DÉMONSTRATION. This is a slight variant of [**105**, Theorem 1], where we sum over an interval of the form $]X, X + H]$ instead of one of the form $]0, X]$. □

Now we have enough tools to state the

Proof of Proposition 8.12. As in the Proof of Proposition 8.8 we start by an application of Lemma 8.9 yielding

$$S(P,j,\nu) \ll \frac{1}{\log P} \max \left| \sum_{n \leq P} \Lambda(n) e\left(\frac{\nu}{q^j}(g(n)+h(n))\right)\right| + P^{\frac{1}{2}}.$$

We split the inner sum into $\leq \log P$ sub sums of the form

$$S := \sum_{X < n \leq X+H} \Lambda(n) e\left(\frac{\nu}{q^j}(g(n)+h(n))\right)$$

with $P^{1-2\rho} \leq X \leq P$ and

$$H = \min\left(P^{1-\theta_r} |\nu|^{-1} q^j, X\right).$$

Now we want to separate the function parts g and h. Therefore we define two functions T and φ by

$$T(x) = \sum_{X < n \leq X+x} \Lambda(n) e\left(\frac{\nu}{q^j} h(n)\right) \quad \text{and} \quad \varphi(x) := e\left(\frac{\nu}{q^j} g\left(X+x\right)\right)$$

Then an application of summation by parts yields

$$\begin{aligned} \sum_{X < n \leq X+H} \Lambda(n) e\left(\frac{\nu}{q^j}(g(n)+h(n))\right) &= \sum_{n=1}^{H} \varphi(n)(T(n)-T(n-1)) \\ &= \sum_{n=1}^{H} T(n)\left(\varphi(n)-\varphi(n+1)\right) + \varphi(H-1)T(H) \\ &\ll |T(H)| + \sum_{n=1}^{H-1} |\varphi(n)-\varphi(n+1)|\,|T(n)| \end{aligned} \tag{4.8}$$

Let α_k be the leading coefficient of P. Then by Diophantine approximation there always exists a rational a/b with $b > 0$, $(a,b) = 1$,

$$1 \leq b \leq H^{k-\rho} \quad \text{and} \quad \left|\frac{\nu\alpha_k}{q^j} - \frac{a}{b}\right| \leq \frac{H^{\rho-k}}{b}.$$

We distinguish three cases according to the size of b.

Case 1. $H^\rho < b$. In this case we may apply Lemma 8.13 together with (4.3) to get

$$T(h) \ll H^{1-\frac{\rho}{4^{k-1}}+\varepsilon}.$$

Case 2. $2 \leq b < H^\rho$. In this case we get that

$$\left|\frac{\nu\alpha_k}{q^j}\right| \geq \left|\frac{a}{b}\right| - \frac{1}{b^2} \geq \frac{1}{2b} \geq \frac{1}{2}H^{-\rho} \geq \frac{1}{2}P^{-\rho}.$$

Since $2\rho < \theta_r$, this contradicts our lower bound $q^j \geq P^{\theta_r-\rho}$.

Case 3. $b = 1$. This case requires a further distinction according to whether $a = 0$ or not.

Case 3.1. $\left|\frac{\nu\alpha_k}{q^j}\right| \geq \frac{1}{2}$. It follows that

$$q^j \leq 2\,|\nu\alpha_k|$$

again contradicting our lower bound $q^j \geq P^{\theta_r-\rho}$.

Case 3.2. $\left|\frac{\nu\alpha_k}{q^j}\right| < \frac{1}{2}$. This implies that $a = 0$ which yields

$$q^j \geq |\nu\alpha_k|\, H^{k-\rho}. \tag{4.9}$$

We distinguish two further cases according to whether $P^{1-\theta_r}\,|\nu|^{-1}\,q^j \leq X$ or not.

Case 3.2.1 $P^{1-\theta_r}\,|\nu|^{-1}\,q^j \leq X$. This implies that $q^j \leq P^{\theta_r}\,|\nu|$ and

$$H = P^{1-\theta_r}\,|\nu|^{-1}\,q^j \geq P^{1-\rho}\,|\nu|^{-1} \geq P^{1-2\rho}.$$

Plugging these estimates into (4.9) gives

$$P^{\theta_r} \geq |\alpha_k|\, P^{(1-2\rho)(k-\rho)}.$$

However, since $4(k+1)\rho < 1$, we have

$$(1-2\rho)(k-\rho) > k - 1 + 2\rho \geq \theta_r$$

yielding a contradiction.

Case 3.2.2 $P^{1-\theta_r}\,|\nu|^{-1}\,q^j > X$. Then $H = X \geq P^{1-2\rho}$ and (4.9) becomes

$$P^{k-1+\rho} \geq |\nu\alpha_k|\, P^{(1-2\rho)(k-\rho)}$$

yielding a similar contradiction as in **Case 3.2.1**.

Therefore **Case 1** is the only possible and we may always apply Lemma 8.13 together with (4.3). Plugging this into (4.8) yields

$$\sum_{X<n\leq X+H} \Lambda(n) e\left(\frac{\nu}{q^j}(g(n)+h(n))\right) \ll H^{1-\frac{\rho}{4^{k-1}}+\varepsilon}\left(1 + \sum_{X<n\leq X+H} |\varphi(n)-\varphi(n+1)|\right)$$

Now by our choice of H together with an application of the mean value theorem we have that

$$\sum_{X\leq n\leq X+H} |\varphi(n)-\varphi(n+1)| \ll H\frac{\nu}{q^j}P^{\theta-1} \ll 1.$$

Thus

$$\sum_{X\leq n\leq X+H} \Lambda(n) e\left(\frac{\nu}{q^j}(g(n)+h(n))\right) \ll H^{1-\frac{\rho}{4^{k-1}}+\varepsilon}.$$

□

5. Proof of Theorem 8.4, Part II

Now we use all the tools from the section above in order to estimate

$$\sum_{j=\ell}^{J}\left|\sum_{p\le P}\mathcal{I}(q^{-j}f(p))-\frac{\pi(P)}{q^{\ell}}\right|\ll\pi(P)H^{-1}J+\sum_{\nu=1}^{H}\nu^{-1}\sum_{j=\ell}^{J}S(P,j,\nu). \tag{5.1}$$

As indicated in the section above, we split the sum over j into two or three parts according to whether $\theta_r > k$ or not. In any case an application of Proposition 8.8 yields for the least significant digits that

$$\sum_{1\le\nu\le P^{\gamma}}\nu^{-1}\sum_{1\le q^j\le P^{\theta_r-\rho}}S(P,j,\nu)\ll(\log P)^9JP^{1-\eta}. \tag{5.2}$$

Now let us suppose that $\theta_r > k$. Then an application of Proposition 8.6 yields

$$\begin{aligned}&\sum_{1\le\nu\le P^{\gamma}}\nu^{-1}\sum_{P^{\theta_r-\rho}<q^j\le P^{\theta_r}}S(P,j,\nu)\\ &\ll\sum_{1\le\nu\le P^{\gamma}}\nu^{-1}\sum_{P^{\theta_r-\rho}<q^j\le P^{\theta_r}}\frac{1}{\log P}\left(\frac{\nu}{q^j}\right)^{-\frac{1}{\lfloor\theta_r\rfloor}}+\frac{P}{(\log P)^{G-2}}\\ &\ll\frac{P}{\log P}.\end{aligned} \tag{5.3}$$

Plugging the estimates (5.2) and (5.3) into (5.1) we get that

$$\sum_{j=\ell}^{J}\left|\sum_{p\le P}\mathcal{I}(q^{-j}f(p))-\frac{\pi(P)}{q^{\ell}}\right|\ll\frac{P}{\log P},$$

which together with (3.3) proves Theorem 8.4 in the case that $\theta_r > k$.

On the other side if $\theta_r < k$, then we consider the two ranges

$$P^{\theta_r-\rho}<q^j\le P^{k-1+\rho}\quad\text{and}\quad P^{k-1+\rho}<q^j\le P^k.$$

For the "digits in the middle" an application of Proposition 8.12 yields

$$\begin{aligned}\sum_{1\le\nu\le P^{\gamma}}\nu^{-1}\sum_{P^{\theta_r-\rho}<q^j\le P^{k-1+\rho}}S(P,j,\nu)&\ll\sum_{1\le\nu\le P^{\gamma}}\nu^{-1}\sum_{P^{\theta_r-\rho}<q^j\le P^{k-1+\rho}}P^{1-\frac{\rho}{4^k}}\\ &\ll(\log P)JP^{1-\frac{\rho}{4^k}}.\end{aligned} \tag{5.4}$$

Finally we consider the most significant digits. By an application of Proposition 8.6 we have

$$\begin{aligned}&\sum_{1\le\nu\le P^{\gamma}}\nu^{-1}\sum_{P^{k-1+\rho}<q^j\le P^{k}}S(P,j,\nu)\\ &\ll\sum_{1\le\nu\le P^{\gamma}}\nu^{-1}\sum_{P^{k-1+\rho}<q^j\le P^{k}}\frac{1}{\log P}\left(\frac{\nu}{q^j}\right)^{-\frac{1}{k}}+\frac{P}{(\log P)^{G-2}}\\ &\ll\frac{P}{\log P}.\end{aligned} \tag{5.5}$$

Plugging the estimates (5.2), (5.4) and (5.5) into (5.1) we get that

$$\sum_{j=\ell}^{J}\left|\sum_{p\le P}\mathcal{I}(q^{-j}f(p))-\frac{\pi(P)}{q^{\ell}}\right|\ll\frac{P}{\log P},$$

which together with (3.3) proves Theorem 8.4 in the case that $\theta_r < k$.

Acknowledgment

The author wants to thank Gérald Tenenbaum for many fruitful discussions and suggestions in connection with the proof of Proposition 8.12.

Chapitre 9

Construction of μ-normal sequences

This chapter is joint work with Bill Mance and appeared in the *Monatshefte für Mathematik*, **179** (2016), 259 – 280.

1. Introduction

Let $q \geq 2$ be a positive integer, then every real $x \in [0,1]$ has a q-adic representation of the form

$$x = \sum_{h \geq 1} d_h q^{-h}$$

with $d_h \in \mathcal{D} := \{0, 1, \ldots, q-1\}$ for $h \geq 1$. We call a number $x \in [0,1]$ *normal* with respect to the base q if for any $k \geq 1$ and any word $\mathbf{b} = b_1 \ldots b_k$ of length k the frequency of occurrences of this word tends to the expected one, namely q^{-k}. Furthermore, we call a number *absolutely normal* if it is normal in every base $q \geq 2$.

In 1909 Borel [**43**] showed that almost all real numbers (with respect to Lebesgue measure) are absolutely normal. This motivated people to look for a concrete example of such a number. It took more than 20 years until 1933 when Champernowne [**58**] provided the first explicit construction by showing that the number

$$0.1\,2\,3\,4\,5\,6\,7\,8\,9\,10\,11\,12\,13\,14\,15\,16\,\ldots$$

is normal to base 10. This construction can be seen as concatenation of function values in a given numeration system. In case of the Champernowne sequence we have the identity together with the decimal system. The constructions were generalized on the one hand by using polynomials or polynomials over primes and on the other hand different numeration systems such as β-expansion, canonical number systems, continued fraction expansion etc. were used.

Let us start with the refinements for the standard q-adic numeration system. Besicovitch [**39**] constructed a normal number by using the sequence of squares. This was extended by Davenport and Erdős [**63**] to polynomials with integer coefficients. Schiffer [**214**] further extended this to polynomials with rational integers and Nakai and Shiokawa [**166**, **167**] used real polynomials and functions of the form $\alpha_1 x^{\beta_1} + \cdots + \alpha_k x^{\beta_k}$ with $\alpha_i > 0$ and $\beta_1 > \beta_2 > \ldots > \beta_k > 0$, respectively. Finally Madritsch *et al.* [**146**] used transcendental entire functions with a certain growth condition.

Another approach was based on sequences over the primes. Copeland and Erdős [**60**] started using the sequence of primes, Nakai and Shiokawa [**168**] used integer polynomials evaluated at the primes and Madritsch and Tichy [**147**] used sequences of the form αp^{β} with $\alpha > 1$ and $\beta \notin \mathbb{Z}$. Finally Madritsch [**150**] considered sequences $\alpha_1 p^{\beta_1} + \cdots + \alpha_k p^{\beta_k}$ where at least one $\beta_i \notin \mathbb{Z}$.

If we look at constructions with different underlying numeration systems then often only the Champernowne construction or a similar one has been considered. For Bernoulli shifts

and continued fractions normal numbers were already investigated by Postnikov and Pyateckiĭ [**188**, **189**], see also Postnikov [**187**]. Furthermore normal sequences for Markov shifts and intrinsically ergodic subshifts were constructed by Smorodinsky and Weiss [**233**]. A different construction for continued fractions is due to Adler *et al.* [**2**]. Generalizations to β-shifts are due to Bertrand-Mathis [**35**], Bertrand-Mathis and Volkmann [**38**] and Ito and Shiokawa [**113**]. However, the latter construct only a sequence with the right frequency of words, which is not an admissible β-expansion. This non-admissibility origins from the fact that not all blocks are allowed to appear in a β-expansion. We can circumvent this by padding zeros in between if the concatenation of two words would contain a forbidden block. More generally if the underlying language allows to fill the gap with a suitable word, then we say that the shift fulfills the specification property (see below for the definition). As a main application of the results of this paper we have the β-expansions in mind and therefore suppose that this property is fulfilled.

Another interesting property of the Champernowne sequence, is the following. For any symbolic dynamical system fulfilling the specification property Bertrand-Mathis [**35**] showed that the Champernowne sequence is generic for the maximal measure. In the present paper we want to modify this sequence and, in particular, its usage of blocks, such that it yields a generic sequence for any given shift invariant measure, not necessarily the maximal one.

We divide this paper into four parts. In the following section we define our main playground together with the results. Then, in Section 3, we modify the Champernowne construction such that it yields a word whose distribution of blocks is generic for any given shift invariant measure μ (this might not be the maximal one). This treatment of different shift invariant measures and, in particular, this modification of Champernowne construction is motivated by recent constructions by Altomare and Mance [**14**] and Mance [**155**, **156**]. In Section 4 we prove the main theorem. This theorem describes a property (which we will call μ-goodness) of a sequence of blocks together with a sequence of repetitions yielding a generic sequence for the measure μ. Finally in Section 5 we put together the modified Champernowne construction from Section 3 and our main theorem to construct generic sequences for different measures in different dynamical systems such as the q-adics, β-expansions, Lüroth series or continued fraction expansion.

2. Definitions and statement of results

Our notation and definitions mainly follow the articles of Bertrand-Mathis and Volkmann [**35**, **38**] as well as the book of Lind and Marcus [**138**]. Let A be a fixed (possibly infinite) alphabet. We denote by A^+ the semigroup generated by A under concatenation. Let ε denote the empty word and $A^* = A^+ \cup \{\varepsilon\}$. A subset $\mathcal{L} \subset A^*$ is called a *language*. The length of a word $\omega = a_1 a_2 \dots a_k$ with $a_i \in A$ for $1 \le i \le k$ is denoted by $|\omega| = k$ and we write A^k for the set of words of length k (over A). Let $\omega \in A^k$. For $\ell \ge 1$ an integer we recursively define $\omega^1 = \omega$ and $\omega^\ell = \omega\omega^{\ell-1}$.

Let $k \ge 1$ be an integer and ν be a probability measure on A^k. Then we call ν shift-invariant, if for all words $\mathbf{b} \in A^{k-1}$

$$\sum_{d \in A} \nu(d\mathbf{b}) = \sum_{d \in A} \nu(\mathbf{b}d)$$

holds.

Now we switch to sequences. We denote by $A^{\mathbb{N}}$ the set of sequences over the alphabet A. For a sequence $\omega = a_1 a_2 a_3 \dots \in A^{\mathbb{N}}$ and a positive integer k, we denote by $\omega|_k = a_1 a_2 \dots a_k$

the truncation of ω to the first k letters. For a language $\mathcal{L} \subset A^*$ let $W^\infty = W^\infty(\mathcal{L}) \subset A^{\mathbb{N}}$ be the set of sequences $\omega = (a_i)_{i\geq 1}$ over A such that every i and k with $1 \leq i < k$, the word $a_i a_{i+1} \cdots a_k$ is in $\mathcal{L}$.

We consider the discrete topology on A and the corresponding product topology on $A^{\mathbb{N}}$. Let $\omega = (a_i)_{i\geq 1} \in A^{\mathbb{N}}$, then we define the shift operator $T\colon A^{\mathbb{N}} \to A^{\mathbb{N}}$ as the mapping $(T(\omega))_i = a_{i+1}$ for $i \geq 1$. Let $\mathfrak{B}$ be the σ-algebra generated by all cylinder sets of $A^{\mathbb{N}}$

$$c(\mathbf{w}) = [\mathbf{w}] = \{\omega \in A^{\mathbb{N}} \colon \omega|_{|\mathbf{w}|} = \mathbf{w}\}$$

for some word $\mathbf{w} \in A^*$. Let μ be a probability measure on $\mathfrak{B}$. We write $\mu(\mathbf{w})$ for $\mu(c(\mathbf{w}))$. Furthermore the measure μ is called *T-invariant* if for all $B \in \mathfrak{B}$ we have that $\mu(T^{-1}B) = \mu(B)$. Let I be the set of all T-invariant probability measures μ on $\mathfrak{B}$. Then we associate with each language $\mathcal{L}$ the symbolic dynamical system

$$S_{\mathcal{L}} = S = (W^\infty(\mathcal{L}), \mathfrak{B}, T, I).$$

Note that W^∞ is invariant under T, *i.e.* $\forall \omega \in W^\infty \colon T\omega \in W^\infty$, and closed with respect to this topology.

Before we continue, we want to present the application we have in mind. Let $\beta > 1$ be a real and $A_\beta := \{0, 1, \ldots, \lceil \beta \rceil - 1\}$. Then every number $x \in [0, 1)$ admits a greedy β-expansion given by the following algorithm due to Rényi [**194**] : set $x_0 = x$, and for $n \geq 1$, let $d_n = \lfloor \beta x_{n-1} \rfloor$ and $x_n = \beta x_{n-1} - d_n$. Then

$$x = \sum_{k\geq 1} d_k \beta^{-k},$$

where $d_k \in A_\beta$. For $\beta > 1$ and $x \in [0, 1)$ we denote by $\mathrm{d}_\beta(x) = d_1 d_2 d_3 \ldots \in A_\beta^{\mathbb{N}}$ the greedy β-expansion of x. We note that for $\beta \in \mathbb{Z}$ this yields the β-adic expansion from Section 1.

Let D_β be the set of all greedy β-expansions in $[0, 1)$. We call a finite word (resp. a sequence) β-admissible if it is a factor of an element (resp. an element) of D_β. Not every number is β-admissible and the β-expansion of 1 plays a central role in the characterization of all β-admissible sequences. In particular, let $\mathrm{d}_\beta(1) = b_1 b_2 \ldots$ be the greedy β-expansion of 1. Since the expansion might be finite we define the quasi-greedy expansion $\mathrm{d}^*_\beta(1)$ by

$$\mathrm{d}^*_\beta(1) = \begin{cases} (b_1 b_2 \ldots b_{t-1}(b_t - 1))^\ell & \text{if } \mathrm{d}_\beta(1) = b_1 b_2 \ldots b_t \text{ is finite,} \\ \mathrm{d}_\beta(1) & \text{otherwise.} \end{cases}$$

Then Parry [**177**] gave the following characterization of the language associated with β.

LEMMA 9.1. *Let $\beta > 1$ be a real number, and let σ be a sequence over $A_\beta = \{0, 1, \ldots, \lceil \beta \rceil - 1\}$. Then σ belongs to D_β if and only if for all $k \geq 0$*

$$T^k(\sigma) \prec_{lex} d^*_\beta(1),$$

where T is the shift and $\prec_{lex}$ denotes the lexicographic ordering.

A real number $\beta > 1$ is a Parry number if the greedy beta-expansion of 1, $d_\beta(1)$, is eventually periodic. The present paper focuses only in β-expansions of real numbers where β is a Parry number.

Let $\varphi = \frac{1+\sqrt{5}}{2}$. Then $d_\varphi(1) = 11$ and therefore the factor 11 is forbidden in the greedy φ-expansion. In our construction we want to guarantee that this is also satisfied if we concatenate two words. In particular, the two words 1001 and 1010 are φ-admissible, however,

their concatenation 10011010 is not. We can circumvent this by padding 0 in between the two words, yielding 1001 0 1010, which is φ-admissible.

More generally we have the following concept. We say that a language $\mathcal{L}$ fulfills the *specification property* if there exists a positive integer j such that for any two words $\mathbf{a}, \mathbf{b} \in \mathcal{L}$ there exists a word $\mathbf{u} \in \mathcal{L}$ with $|\mathbf{u}| \leq j$ such that $\mathbf{aub} \in \mathcal{L}$. Then we generalize concatenation as follows. For any pair of finite words $\mathbf{a}$ and $\mathbf{b}$ we fix a $\mathbf{u_{a,b}}$ with $|\mathbf{u_{a,b}}| \leq j$ such that $\mathbf{au_{a,b}b} \in \mathcal{L}$. Then for $\mathbf{a}, \mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathcal{L}$ and $n \in \mathbb{N}$ we write

$$\mathbf{a}_1 \odot \mathbf{a}_2 \odot \cdots \odot \mathbf{a}_m := \mathbf{a}_1\, \mathbf{u}_{\mathbf{a}_1,\mathbf{a}_2}\, \mathbf{a}_2\, \mathbf{u}_{\mathbf{a}_2,\mathbf{a}_3}\, \mathbf{a}_3 \cdots \mathbf{a}_{m-1}\, \mathbf{u}_{\mathbf{a}_{m-1},\mathbf{a}_m}\, \mathbf{a}_m$$

for short and recursively define

$$\mathbf{a}^{\odot 1} = \mathbf{a} \quad \text{and} \quad \mathbf{a}^{\odot n} = \mathbf{a} \odot \mathbf{a}^{\odot(n-1)} \quad \text{for } n \geq 2.$$

Let $\mu \in I$ be a T-invariant measure. A word $\mathbf{b} \in \mathcal{L}$ is *μ-admissible* if $\mu(\mathbf{b}) \neq 0$. Let $\mathcal{D}_\mu \subset \mathcal{L}$ denote the set of μ-admissible words and let $\mathcal{D}_{\mu,k}$ denote the set of μ-admissible words of length k. Given a word $\mathbf{b} \in A^k$ and a sequence $\omega = a_1a_2a_3 \ldots \in A^{\mathbb{N}}$ we let $N_n(\mathbf{b}, \omega)$ denote the number of times the word $\mathbf{b}$ occurs starting in position no greater than n in the word ω, *i.e.*

$$N_n(\mathbf{b}, \omega) = \#\{0 \leq i < n \colon a_{i+1}a_{i+2} \cdots a_{i+k} = \mathbf{b}\}.$$

For $\mathbf{w} \in A^k$ we write $N(\mathbf{b}, \mathbf{w})$ in place of $N_{|\mathbf{w}|-|\mathbf{b}|+1}(\mathbf{b}, \mathbf{w})$.

The concept of normal sequences is in close connection with the concept of generic ones. We call a sequence $\omega = a_1a_2a_3 \ldots \in A^{\mathbb{N}}$ *associated to* μ if there exists an infinite subset $F \subset \mathbb{N}$ such that for any finite word $\mathbf{b} \in A^*$

$$\lim_{\substack{n\to\infty \\ n\in F}} \frac{N_n(\mathbf{b}, \omega)}{n} = \mu(\mathbf{b}).$$

A sequence ω is *generic for* μ if for every finite word $\mathbf{b} \in A^*$,

$$\lim_{\substack{n\to\infty \\ n\in \mathbb{N}}} \frac{N_n(\mathbf{b}, \omega)}{n} = \mu(\mathbf{b}).$$

Using the definition of a sequence being associated to a measure, we can say that ω is generic if ω is associated to μ for F equal to the whole set of natural numbers. Moreover we note that if a sequence is generic for μ then μ is the only associated measure.

For the definition of normality we start with the following generalization of the concept of (ε, k)-normality originally due to Besicovitch [**39**]. Let ϵ be a positive real less than 1 and k be positive integer. A word $\mathbf{w}$ is called *(ϵ, k, μ)-normal* if for all $t \leq k$ and words $\mathbf{b}$ in $\mathcal{D}_{\mu,t}$, we have

$$\mu(\mathbf{b})|\mathbf{w}|(1-\epsilon) \leq N(\mathbf{b}, \mathbf{w}) \leq \mu(\mathbf{b})|\mathbf{w}|(1+\epsilon).$$

A sequence $\omega \in A^{\mathbb{N}}$ is called *μ-normal of order k* if for every μ-admissible word $\mathbf{b} \in \mathcal{D}_{\mu,k}$ we have

$$\lim_{n\to\infty} \frac{N_n(\mathbf{b}, \omega)}{n} = \mu(\mathbf{b}).$$

We denote by $\mathcal{N}_{\mu,k}$ the set of all μ-normal sequences of order k. Furthermore we call ω *μ-normal* if it is μ-normal of all orders $k \geq 1$, *i.e.* $\omega \in \mathcal{N}_\mu := \bigcap_{k=1}^{\infty} \mathcal{N}_{\mu,k}$. Clearly any μ-normal sequence $\omega \in A^{\mathbb{N}}$ is also generic for μ and vice versa.

We will use Landau's $\mathcal{O}$-notation to describe the asymptotic behaviour of functions. In particular, we say that $f = o(g)$, if for every $\varepsilon > 0$ there exists x_0 such that $|f(x)| \leq \varepsilon\, |g(x)|$ for $x \geq x_0$ or equivalently $\lim_{x\to\infty} |f(x)| \,/\, |g(x)| = 0$.

Let $(k_i)_{i\geq 1}$ be a sequence of positive integers. For $i \geq 1$ let $\nu_i \colon A^{k_i} \to [0,1]$ be a shift-invariant probability measure. Then we call $(\nu_i)_{i\geq 1}$ an approximation scheme for μ, if $(\nu_i)_{i\geq 1}$ converges weakly to $\mu \in I$ (written $\nu_i \to \mu$), such that $\mathcal{D}_{\nu_i} \subset \mathcal{D}_\mu$ for all $i \geq 1$ and such that $\nu_i(\mathbf{b})$ is eventually non-increasing in i. We note that a version of our main theorem without the condition $\mathcal{D}_{\nu_i} \subset \mathcal{D}_\mu$ is still true, but every example we will consider has this property.

Furthermore let $(\mathbf{w}_i)_{i\geq 1}$ be a sequence of finite words and $(\ell_i)_{i\geq 1}$ be a non-decreasing sequence of positive integers. Then we call the sequence of pairs $(\mathbf{w}_i, \ell_i)_{i\geq 1}$ μ-good with respect to the approximation scheme $(\nu_i)_{i\geq 1}$ if each $\mathbf{w}_i$ is $(\varepsilon_i, k_i, \nu_i)$-normal satisfying

$$\frac{1}{\varepsilon_{i-1} - \varepsilon_i} = o(|\mathbf{w}_i|); \tag{2.1}$$

$$\frac{\ell_{i-1}}{\ell_i} \cdot \frac{|\mathbf{w}_{i-1}|}{|\mathbf{w}_i|} = o(i^{-1}); \tag{2.2}$$

$$\frac{1}{\ell_i} \cdot \frac{|\mathbf{w}_{i+1}|}{|\mathbf{w}_i|} = o(1). \tag{2.3}$$

Now we are able to state our main theorem.

THEOREM 9.2. *Let A be an alphabet, $\mathcal{L}$ be a set of finite words over A and $W^\infty(\mathcal{L})$ the set of sequences generated by $\mathcal{L}$. Let T be a shift of $A^{\mathbb{N}}$ and μ be a shift invariant probability measure on $W^\infty(\mathcal{L})$. Let $(k_i)_{i\geq 1}$ be a sequence of positive integers and for $i \geq 1$ let $\nu_i \colon A^{k_i} \to [0,1]$ be a shift-invariant probability measure on A^{k_i} such that $(\nu_i)_{i\geq 1}$ is an approximation scheme for μ. Let $(\mathbf{w}_i)_{i\geq 1}$ be a sequence of finite words and let $(\ell_i)_{i\geq 1}$ be a non-decreasing sequence of positive integers. Suppose that $(\mathbf{w}_i, \ell_i)_{i\geq 1}$ is μ-good with respect to $(\nu_i)_{i\geq 1}$, then for each integer $k \in [1, \limsup_{i\to\infty} k_i]$, the sequence $\omega = \mathbf{w}_1^{\odot \ell_1} \odot \mathbf{w}_2^{\odot \ell_2} \odot \cdots$ is μ-normal of order k. Moreover, if $\limsup_{i\to\infty} k_i = \infty$, then ω is μ-normal*

We start presenting a construction of a μ-good sequence of pairs of words and integers $(\mathbf{w}_i, \ell_i)_{i\geq 1}$ for a given approximation scheme $(\nu_i)_{i\geq 1}$ in Section 3. Then in Section 4 we prove Main Theorem 9.2. Finally, we apply our main theorem together with our constructed sequence of pairs $(\mathbf{w}_i, \ell_i)_{i\geq 1}$ to different number systems.

3. The construction

Let $A^\ell = \{\mathbf{p}_1, \ldots, \mathbf{p}_{b^\ell}\}$ be an arbitrary ordering of the set of all possible words of length ℓ of the alphabet $A = \{0, 1, \ldots, b-1\}$ of digits in base b.

Recall that $D_{\mu,k}$ denotes the set of μ-admissible words of length k, and let $m_k = \min\{\mu(\mathbf{b}) : \mathbf{b} \in \mathcal{D}_{\mu,k}\}$ for $k \geq 1$ and M be an arbitrary large constant such that $M \geq \frac{1}{m_\ell}$.

We will define a word $p_{b,\ell,M}$ that will contain each of the words in A^ℓ, with the appropriate multiplicity. Since it has to belong to the language, we use padding. Thus

$$\mathbf{p}_{b,\ell,M} := \mathbf{p}_1^{\odot \lceil M\mu(\mathbf{p}_1)\rceil} \odot \mathbf{p}_2^{\odot \lceil M\mu(\mathbf{p}_2)\rceil} \odot \cdots \odot \mathbf{p}_{b^\ell}^{\odot \lceil M\mu(\mathbf{p}_{b^\ell})\rceil}.$$

In the following we show the (ε, k)-normality of $\mathbf{p}_{b,\ell,M}$ for $k \leq \ell$. Thus it suffices to find an ε such that for all words $\mathbf{b}$ of length $k \leq \ell$ we have

$$(1-\varepsilon)\mu(\mathbf{b}) \leq \frac{N(\mathbf{b}, \mathbf{p}_{b,\ell,M})}{|\mathbf{p}_{b,\ell,M}|} \leq (1+\varepsilon)\mu(\mathbf{b}) \tag{3.1}$$

To this end we need lower and upper bounds for the length of $\mathbf{p}_{b,\ell,M}$ as well as lower and upper bounds for the number of occurrences of a fixed word $\mathbf{b}$ within $\mathbf{p}_{b,\ell,M}$.

Let j be the maximum size of the padding given by the specification property. Starting with the estimation of the length of $\mathbf{p}_{b,\ell,M}$ we get as upper bound

$$|\mathbf{p}_{b,\ell,M}| \leq \sum_{i=1}^{b^\ell} \lceil M\mu(\mathbf{p}_i)\rceil\,(j+\ell) \leq M(j+\ell)\sum_{i=1}^{b^\ell}\mu(\mathbf{p}_i) + (j+\ell)b^\ell = (j+\ell)\left(M+b^\ell\right). \tag{3.2}$$

On the other hand we obtain as lower bound

$$|\mathbf{p}_{b,\ell,M}| \geq \sum_{i=1}^{b^\ell} \lceil M\mu(\mathbf{p}_i)\rceil\,\ell \geq M\ell\sum_{i=1}^{b^\ell}\mu(\mathbf{p}_i) = M\ell. \tag{3.3}$$

Now we provide upper and lower bounds for the number of occurrences of a word $\mathbf{b}$ of length k in $\mathbf{p}_{b,\ell,M}$.

For the lower bound we only count the possible occurrences within a $\mathbf{p}_i$. If there is an occurrence then we can write $\mathbf{p}_i$ as $\mathbf{c}_1\mathbf{b}\mathbf{c}_2$ with possible empty $\mathbf{c}_1$ or $\mathbf{c}_2$. Since the word $\mathbf{b}$ is fixed and all possible words of length ℓ occur in $\mathbf{p}_{b,\ell,M}$, we let $\mathbf{c}_1$ and $\mathbf{c}_2$ vary over all possible words. Thus

$$\begin{aligned} N(\mathbf{b},\mathbf{p}_{b,\ell,M}) &\geq \sum_{m=0}^{\ell-k}\sum_{|\mathbf{c}_1|=m}\sum_{|\mathbf{c}_2|=\ell-k-m}\lceil M\mu(\mathbf{c}_1\mathbf{b}\mathbf{c}_2)\rceil \\ &\geq M\sum_{m=0}^{\ell-k}\sum_{|\mathbf{c}_1|=m}\sum_{|\mathbf{c}_2|=\ell-k-m}\mu(\mathbf{c}_1\mathbf{b}\mathbf{c}_2) \\ &= M\sum_{m=0}^{\ell-k}\sum_{|\mathbf{c}_1|=m}\sum_{|\mathbf{c}_2|=\ell-k-m-1}\sum_{d=0}^{b-1}\mu(\mathbf{c}_1\mathbf{b}\mathbf{c}_2 d) \\ &= \cdots = M\sum_{m=0}^{\ell-k}\mu(\mathbf{b}) = (\ell-k+1)M\mu(\mathbf{b}), \end{aligned} \tag{3.4}$$

where we have used the shift invariance of μ, *i.e.* $\sum_{d=0}^{b-1}\mu(d\mathbf{a}) = \sum_{d=0}^{b-1}\mu(\mathbf{a}d) = \mu(\mathbf{a})$.

For the upper bound we have to consider several different possibilities : The word $\mathbf{b}$ can occour

(1) within $\mathbf{p}_i$,

(2) between two similar words $\mathbf{p}_i \odot \mathbf{p}_i$ or

(3) between two different words $\mathbf{p}_i \odot \mathbf{p}_{i+1}$.

If the word $\mathbf{b}$ is completely within $\mathbf{p}_i$, then we again have that $\mathbf{p}_i = \mathbf{c}_1\mathbf{b}\mathbf{c}_2$ with possible empty $\mathbf{c}_1$ or $\mathbf{c}_2$. By using similar means as above we get that

$$\sum_{\mathbf{c}_1,\mathbf{c}_2}\lceil M\mu(\mathbf{c}_1\mathbf{b}\mathbf{c}_2)\rceil \leq \sum_{\mathbf{c}_1,\mathbf{c}_2}(M\mu(\mathbf{c}_1\mathbf{b}\mathbf{c}_2)+1) = \cdots = (\ell-k+1)\left(M\mu(\mathbf{b})+b^{\ell-k}\right),$$

Now we turn our attention to the number of occurrences in $\mathbf{p}_{b,\ell,M}$ between two consecutive words. First we assume that these words are equal. Let $n = |\mathbf{p}_i \odot \mathbf{p}_i|$ be the length of the

resulting word. Then $\mathbf{p}_i \odot \mathbf{p}_i = \mathbf{c}_1\mathbf{b}\mathbf{c}_2$ with $\ell - k + 1 \le |\mathbf{c}_1| \le n - \ell - 1$. Thus similar to above we get that there are

$$\begin{aligned}
&\sum_{m=\ell-k+1}^{n-\ell-1} \sum_{|\mathbf{c}_1|=m} \sum_{|\mathbf{c}_2|=n-k-m} \lceil M\mu(\mathbf{c}_1\mathbf{b}\mathbf{c}_2)\rceil \\
&\le M \sum_{m=\ell-k+1}^{n-\ell-1} \sum_{|\mathbf{c}_1|=m} \sum_{|\mathbf{c}_2'|=n-k-m-1} \sum_{d=0}^{b-1} \mu(\mathbf{c}_1\mathbf{b}\mathbf{c}_2'd) + \sum_{m=\ell-k+1}^{n-\ell-1} b^{n-k} \\
&= \cdots = M \sum_{m=\ell-k+1}^{n-\ell-1} \mu(\mathbf{b}) + (n - 2\ell + 2k - 2)b^{n-k} \\
&= (n - 2\ell + k - 1)\left(M\mu(\mathbf{b}) + b^{n-k}\right) \\
&\le (j + k - 1)\left(M\mu(\mathbf{b}) + b^{2\ell+j-k}\right)
\end{aligned}$$

occurrences between two identical words.

Finally, we trivially estimate the number of occurrences between two different words by their total amount, which is $\le (j + k - 1)b^\ell$.

Combining these three bounds and using $k \le \ell$ we get as upper bound for the number of occurrences

$$\begin{aligned}
&N(\mathbf{b}, \mathbf{p}_{b,\ell,M}) \\
&\le (\ell - k + 1)\left(M\mu(\mathbf{b}) + b^{\ell-k}\right) + (j + k - 1)\left(M\mu(\mathbf{b}) + b^{2\ell+j-k}\right) + (j + k - 1)b^\ell \\
&\le (\ell + j)\left(M\mu(\mathbf{b}) + b^{2\ell+j-k}\right).
\end{aligned} \tag{3.5}$$

Now we calculate ε such that (3.1) holds. Using our lower bound for the number of occurrences in (3.4) together with our upper bound for the length in (3.2) we get that

$$\frac{N(\mathbf{b}, \mathbf{p}_{b,\ell,M})}{|\mathbf{p}_{b,\ell,M}|} \ge \frac{(\ell - k + 1)M\mu(\mathbf{b})}{(\ell + j)\,(M + b^\ell)} \ge \mu(\mathbf{b})\left(1 - \frac{j + k - 1}{\ell + j}\right)\left(1 - \frac{b^\ell}{M + b^\ell}\right)$$

which implies for ε the upper bound

$$\varepsilon \le \frac{j + k - 1}{\ell + j} + \frac{b^\ell}{M + b^\ell}.$$

On the other side an application of the upper bound for the number of occurrences in (3.5) together with the lower bound for the length in (3.3) yields

$$\frac{N(\mathbf{b}, \mathbf{p}_{b,\ell,M})}{|\mathbf{p}_{b,\ell,M}|} \le \mu(\mathbf{b})\left(1 + \frac{j}{\ell}\right)\left(1 + \frac{1}{m_k}\frac{b^{2\ell+j-k}}{M}\right).$$

Putting these together we get that $\mathbf{p}_{b,\ell,M}$ is (ε, k, μ)-normal for

$$k \le \ell \quad \text{and} \quad \varepsilon \le \max\left(\frac{j + k - 1}{\ell + j} + \frac{b^\ell}{M + b^\ell}, \frac{j}{\ell} + \frac{1}{m_k}\frac{b^{2\ell+j-k}}{M}\right). \tag{3.6}$$

A more careful control of the available words and their distribution, would lead to a reduction in the number of copies ℓ_i for special cases (*cf.* Vandehey [**243**]).

4. Proof of Main Theorem 9.2

In our proof we will use a classical counting argument. We could use a variant of the "hot spot lemma" (*cf.* Moshchevitin and Shkredov [**163**] and Shkredov [**225**]). However, on the one hand, since their results are for the full shift over finite and infinite alphabets, we need to develop a variant of the "hot spot lemma" for dynamical systems satisfying the specification property. On the other hand, since our proof follows along similar lines to the proof of Main Theorem 1.15 in [**155**], we will only include those parts that differ significantly and omit the proofs, which are similar to proofs of lemmas in [**155**].

Throughout this section, we will fix a sequence $W = ((\mathbf{w}_i, \ell_i))_{i=1}^{\infty}$ that is μ-good for the approximation scheme $(\nu_i)_{i\geq 1}$. Suppose that every $\mathbf{w}_i$ is (ϵ_i, k_i, ν_i)-normal. Then we define the set of supported lengths $R(W) = [1, \limsup_{i\to\infty} k_i] \cap \mathbb{N}$.

Set $\omega = \mathbf{w}_1^{\odot\ell_1} \odot \mathbf{w}_2^{\odot\ell_2} \odot \cdots$ to be the constructed infinite word and denote by σ_k the kth word,

$$\sigma_k = \mathbf{w}_k^{\odot\ell_k} \mathbf{u}_{\mathbf{w}_k,\mathbf{w}_{k+1}}.$$

Let L_i be the length of the concatenation up to the ith word, *i.e.*

$$L_i = \sum_{k=1}^{i} |\sigma_k| = \sum_{k=1}^{i} \left(\ell_k|\mathbf{w}_k| + (\ell_k - 1)|\mathbf{u}_{\mathbf{w}_k,\mathbf{w}_k}| + |\mathbf{u}_{\mathbf{w}_k,\mathbf{w}_{k+1}}|\right).$$

For a given n, the letter $i = i(n)$ will always be understood to be the positive integer that satisfies $L_i < n \leq L_{i+1}$, *i.e.* position n lies in the word σ_{i+1}. Let $m = n - L_i$, then we consider $\sigma_{i+1}|m$. Let x be the largest integer such that there is a word $\mathbf{v}$ for which

$$\sigma_{i+1}|m = (\mathbf{w}_{i+1}\mathbf{u}_{\mathbf{w}_{i+1},\mathbf{w}_{i+1}})^x \mathbf{v} \tag{4.1}$$

Then m can be written in the form

$$m = x(|\mathbf{w}_{i+1}| + \left|\mathbf{u}_{\mathbf{w}_{i+1},\mathbf{w}_{i+1}}\right|) + y$$

with $y = |\mathbf{v}|$. We have that x and y satisfy

$$0 \leq x < \ell_{i+1} \text{ and } 0 \leq y < |\mathbf{w}_{i+1}| + j,$$

where j is the bound from the specification property.

Thus, we can write the first n digits of ω as concatenation of the complete words $\sigma_1, \ldots, \sigma_i$, the x repetitions of the word $\mathbf{w}_{i+1}$ and the rest $\mathbf{v}$, *i.e.*

$$\omega|_n = \sigma_1\sigma_2\cdots\sigma_i \mathbf{w}_{i+1}^{\odot x} \odot \mathbf{v}. \tag{4.2}$$

For a word $\mathbf{b}$, let

$$\phi_n(\mathbf{b}) = \sum_{k=1}^{i} |\sigma_k|\, \nu_k(\mathbf{b}) + m\nu_{i+1}(\mathbf{b}). \tag{4.3}$$

Since $(\mathbf{w}_i, \ell_i)_{i=1}^{\infty}$ is μ-good, we have that $\lim_{n\to\infty} \frac{\phi_n(\mathbf{b})}{n} = \mu(\mathbf{b})$. Therefore ω is μ-normal if and only if

$$\lim_{n\to\infty} \frac{N_n(\mathbf{b}, \omega)}{\phi_n(\mathbf{b})} = 1 \tag{4.4}$$

for all words $\mathbf{b} \in \mathcal{D}_\mu$.

For a given word $\mathbf{b}$ of supported length $k \in R(W)$, the following lemma provides us with upper and lower bounds for $N_n(\mathbf{b}, \omega)$.

LEMMA 9.3. *If $k \le k_i$ and $\mathbf{b} \in \mathcal{D}_{\nu_i,k}$, then*

$$N_n(\mathbf{b},\omega) \le L_{i-1} + ((1+\epsilon_i)\nu_i(\mathbf{b})|\mathbf{w}_i| + k + j)\,\ell_i + ((1+\epsilon_{i+1})\nu_{i+1}(\mathbf{b})|\mathbf{w}_{i+1}| + k + j)x + y$$

and

$$N_n(\mathbf{b},\omega) \ge (1-\epsilon_i)\nu_i(\mathbf{b})|\mathbf{w}_i|\ell_i + (1-\epsilon_{i+1})\nu_{i+1}(\mathbf{b})|\mathbf{w}_{i+1}|.$$

DÉMONSTRATION. By (4.2) we have that

$$N_n(\mathbf{b},\omega) = N(\mathbf{b}, \sigma_1\cdots\sigma_{i-1}\mathbf{w}_i^{\odot\ell_i} \odot \mathbf{w}_{i+1}^{\odot x} \odot \mathbf{v}).$$

On the one hand, we have that

$$\begin{aligned} N_n(\mathbf{b},\omega) &\le L_{i-1} + \ell_i(N(\mathbf{b},\mathbf{w}_i) + k + j) + x(N(\mathbf{b},\mathbf{w}_{i+1}) + k + j) + y \\ &\le L_{i-1} + \ell_i((1+\epsilon_i)\nu_i(\mathbf{b})|\mathbf{w}_i| + k + j) + x((1+\epsilon_{i+1})\nu_{i+1}(\mathbf{b})|\mathbf{w}_{i+1}| + k + j) + y. \end{aligned}$$

On the other hand

$$\begin{aligned} N_n(\mathbf{b},\omega) &\ge \ell_i N(\mathbf{b},\mathbf{w}_i) + xN(\mathbf{b},\mathbf{w}_{i+1}) \\ &\ge (1-\epsilon_i)\nu_i(\mathbf{b})\ell_i|\mathbf{w}_i| + (1-\epsilon_{i+1})\nu_{i+1}(\mathbf{b})|\mathbf{w}_{i+1}|. \end{aligned}$$

□

Now we estimate $N_n(\mathbf{b},\omega)/\phi_n(\mathbf{b})$ from above and below. On the one hand using the upper bound for $N_n(\mathbf{b},\omega)$ in Lemma 9.3 and the definition of $\phi_n(\mathbf{b})$ in (4.3) yields

$$\begin{aligned} &\frac{N_n(\mathbf{b},\omega)}{\phi_n(\mathbf{b})} - 1 \\ &\le \frac{L_{i-1} + (\epsilon_i\nu_i(\mathbf{b})|\mathbf{w}_i| + (k+j))\ell_i + (\epsilon_{i+1}\nu_{i+1}(\mathbf{b})|\mathbf{w}_{i+1}| + (k+j))x + y}{\phi_{L_i}(\mathbf{b}) + \nu_{i+1}(\mathbf{b})\left(\left|\mathbf{w}_{i+1}\mathbf{u}_{\mathbf{w}_{i+1},\mathbf{w}_{i+1}}\right| x + y\right)} =: g_{i,\mathbf{b}}(x,y) \end{aligned} \tag{4.5}$$

On the other hand combining the lower bound for $N_n(\mathbf{b},\omega)$ in Lemma 9.3 and the definition of $\phi_n(\mathbf{b})$ in (4.3) yields

$$\begin{aligned} &\frac{N_n(\mathbf{b},\omega)}{\phi_n(\mathbf{b})} - 1 \\ &\ge -\frac{\phi_{L_{i-1}}(\mathbf{b}) + \epsilon_i\nu_i(\mathbf{b})\ell_i|\mathbf{w}_i| + \nu_{i+1}(\mathbf{b})\left(\epsilon_{i+1}|\mathbf{w}_{i+1}| + \left|\mathbf{u}_{\mathbf{w}_{i+1},\mathbf{w}_{i+1}}\right|\right)x + \nu_{i+1}(\mathbf{b})y}{\phi_{L_i}(\mathbf{b}) + \nu_{i+1}(\mathbf{b})\left(\left|\mathbf{w}_{i+1}\mathbf{u}_{\mathbf{w}_{i+1},\mathbf{w}_{i+1}}\right| x + y\right)} =: -f_{i,\mathbf{b}}(x,y) \end{aligned} \tag{4.6}$$

Therefore

$$\left|\frac{N_n(\mathbf{b},\omega)}{\phi_n(\mathbf{b})} - 1\right| < \max\left(f_{i,\mathbf{b}}(x,y), g_{i,\mathbf{b}}(x,y)\right).$$

However, since the numerator of $g_{i,\mathbf{b}}(x,y)$ is clearly greater than the numerator of $f_{i,\mathbf{b}}(x,y)$ and their denominators are the same we deduce the following.

LEMMA 9.4. *For any i let $k \in R(W)$, $k \le k_i$, and $\mathbf{b} \in \mathcal{D}_{\nu_i,k}$. Then*

$$\left|\frac{N_n(\mathbf{b},\omega)}{\phi_n(\mathbf{b})} - 1\right| < g_{i,\mathbf{b}}(x,y). \tag{4.7}$$

We are looking for a good upper bound for $g_{i,\mathbf{b}}(x,y)$ where (x,y) ranges over values in $\{0,1,\ldots,\ell_{i+1}\} \times \{0,1,\ldots,|\mathbf{w}_{i+1}|-1\}$.

LEMMA 9.5. *If $k \in R(W)$, $\varepsilon_i < 1/2$, $\ell_i > 0$, $\mathbf{b} \in \mathcal{D}_{\nu_i,k}$,*

$$|\mathbf{w}_i| > 2 \cdot (k+j) + 2\frac{L_{i-1}\nu_{i+1}(\mathbf{b}) - \phi_{L_{i-1}}}{\ell_i \nu_i(\mathbf{b})}, \quad |\mathbf{w}_{i+1}| > \frac{k+j}{\nu_{i+1}(\mathbf{b})(\epsilon_i - \epsilon_{i+1})},$$

and

$$(x,y) \in \{0,1,\dots,\ell_{i+1}\} \times \{0,1,\dots,|\mathbf{w}_{i+1}| + j - 1\},$$

then

$$g_{i,\mathbf{b}}(x,y) < g_{i,\mathbf{b}}(0,|\mathbf{w}_{i+1}|+j) = \frac{(L_{i-1} + \epsilon_i \nu_i(\mathbf{b})\ell_i|\mathbf{w}_i| + (k+j)\ell_i) + |\mathbf{w}_{i+1}| + j}{\phi_{L_i}(\mathbf{b}) + \nu_{i+1}(\mathbf{b})\,(|\mathbf{w}_{i+1}| + j)}. \tag{4.8}$$

DÉMONSTRATION. We note that $g_{i,\mathbf{b}}(x,y)$ is a rational function of x and y of the form

$$g_{i,\mathbf{b}}(x,y) = \frac{C + Dx + Ey}{F + Gx + Hy}$$

where

$$\begin{aligned} &C = L_{i-1} + \epsilon_i \nu_i(\mathbf{b})\ell_i|\mathbf{w}_i| + (k+j)\ell_i, && D = \epsilon_{i+1}\nu_{i+1}(\mathbf{b})|\mathbf{w}_{i+1}| + (k+j), && E = 1, \\ &F = \phi_{L_i}(\mathbf{b}), && G = \nu_{i+1}(\mathbf{b})\left|\mathbf{w}_{i+1}\mathbf{u}_{\mathbf{w}_{i+1},\mathbf{w}_{i+1}}\right|, \text{ and} && H = \nu_{i+1}(\mathbf{b}). \end{aligned}$$

We will show that if we fix y, then $g_{i,\mathbf{b}}(x,y)$ is a decreasing function of x and if we fix x, then $g_{i,\mathbf{b}}(x,y)$ is an increasing function of y. To see this, we compute the partial derivatives :

$$\begin{aligned} \frac{\partial g_{i,\mathbf{b}}}{\partial x}(x,y) &= \frac{D(F+Gx+Hy) - G(C+Dx+Ey)}{(F+Gx+Hy)^2} = \frac{D(F+Hy) - G(C+Ey)}{(F+Gx+Hy)^2}; \\ \frac{\partial g_{i,\mathbf{b}}}{\partial y}(x,y) &= \frac{E(F+Gx+Hy) - H(C+Dx+Ey)}{(F+Gx+Hy)^2} = \frac{E(F+Gx) - H(C+Dx)}{(F+Gx+Hy)^2}. \end{aligned} \tag{4.9}$$

Thus, the sign of $\frac{\partial g_{i,\mathbf{b}}}{\partial x}(x,y)$ does not depend on x and the sign of $\frac{\partial g_{i,\mathbf{b}}}{\partial y}(x,y)$ does not depend on y. We will first show that $g_{i,\mathbf{b}}(x,y)$ is an increasing function of y by verifying that

$$E(F+Gx) > H(C+Dx). \tag{4.10}$$

Therefore we check that $EF > HC$ and $EGx > HDx$. To see that the first inequality holds we need to show that

$$EF \geq \phi_{L_{i-1}}(\mathbf{b}) + \nu_i(\mathbf{b})|\mathbf{w}_i|\ell_i > \nu_{i+1}(\mathbf{b})\,(L_{i-1} + \epsilon_i\nu_i(\mathbf{b})\ell_i|\mathbf{w}_i| + (k+j)\ell_i) = HC.$$

The first inequality is clear from the definition and the second one equals

$$|\mathbf{w}_i| > \frac{\nu_{i+1}(\mathbf{b})}{\nu_i(\mathbf{b})} \cdot \frac{k+j}{1 - \nu_{i+1}(\mathbf{b})\epsilon_i} + \frac{L_{i-1}\nu_{i+1}(\mathbf{b}) - \phi_{L_{i-1}}}{\ell_i\nu_i(\mathbf{b})(1 - \nu_{i+1}(\mathbf{b})\varepsilon_i)}. \tag{4.11}$$

Since $\epsilon_i < 1/2$, we know that $(1 - \nu_{i+1}(\mathbf{b})\epsilon_i)^{-1} < 2$. Additionally, since $\nu_i(\mathbf{b})$ is eventually non-increasing we have for sufficiently large i that $\nu_{i+1}(\mathbf{b}) \leq \nu_i(\mathbf{b})$. Therefore,

$$\frac{\nu_{i+1}(\mathbf{b})}{\nu_i(\mathbf{b})} \cdot \frac{k+j}{1 - \nu_{i+1}(\mathbf{b})\epsilon_i} + \frac{L_{i-1}\nu_{i+1}(\mathbf{b}) - \phi_{L_{i-1}}}{\ell_i\nu_i(\mathbf{b})(1 - \nu_{i+1}(\mathbf{b})\varepsilon_i)} < 2 \cdot (k+j) + 2\frac{L_{i-1}\nu_{i+1}(\mathbf{b}) - \phi_{L_{i-1}}}{\ell_i\nu_i(\mathbf{b})}.$$

The verification of $EGx > HDx$ is equivalent to the verification of

$$|\mathbf{w}_{i+1}|x \geq (\epsilon_{i+1}\nu_{i+1}(\mathbf{b})|\mathbf{w}_{i+1}| + (k+j))x. \tag{4.12}$$

Clearly, (4.12) is true if $x = 0$. If $x > 0$ we can rewrite (4.12) as

$$|\mathbf{w}_{i+1}| \geq \frac{1}{1 - \nu_{i+1}(\mathbf{b})\epsilon_{i+1}} \cdot (k + j).$$

Similar to (4.11), $(1-\nu_{i+1}(\mathbf{b})\epsilon_{i+1})^{-1}(k+j) \leq 2(k+j) < |\mathbf{w}_i| \leq |\mathbf{w}_{i+1}|$. Thus (4.10) is satisfied and $g_{i,\mathbf{b}}(x, y)$ is an increasing function of y.

To show that $\frac{\partial g_{i,\mathbf{b}}}{\partial x}(x, y) < 0$ we proceed as follows : because the sign of $\frac{\partial g_{i,\mathbf{b}}}{\partial x}(x, y)$ does not depend on x, we will know that $g_{i,\mathbf{b}}(x, y)$ is decreasing in x if for each y

$$\lim_{x\to\infty} g_{i,\mathbf{b}}(x, y) < g_{i,\mathbf{b}}(0, y).$$

Since $g_{i,\mathbf{b}}(x, y)$ is an increasing function of y, we know for all y that $g_{i,\mathbf{b}}(0, 0) < g_{i,\mathbf{b}}(0, y)$. Hence, it is enough to show that

$$\lim_{x\to\infty} g_{i,\mathbf{b}}(x, y) < g_{i,\mathbf{b}}(0, 0).$$

Since $\lim_{x\to\infty} g_{i,\mathbf{b}}(x, y) = D/G$ and $g_{i,\mathbf{b}}(0, 0) = C/F$, it is sufficient to show that $CG > DF$, where C, D, F and G are as in (4.9).

Since $0 \leq |u_{\mathbf{a},\mathbf{b}}| \leq j$, it suffices for checking $CG > DF$ that

$$\begin{aligned} &L_{i-1}\nu_{i+1}(\mathbf{b})|\mathbf{w}_{i+1}| > (\epsilon_{i+1}\nu_{i+1}(\mathbf{b})|\mathbf{w}_{i+1}| + (k + j))\,\phi_{L_{i-1}}(\mathbf{b}) \quad \text{and} \\ &\epsilon_i\nu_{i+1}(\mathbf{b})|\mathbf{w}_{i+1}| > \epsilon_{i+1}\nu_{i+1}(\mathbf{b})|\mathbf{w}_{i+1}| + (k + j), \end{aligned} \tag{4.13}$$

Since $L_{i-1} > \phi_{L_{i-1}}(\mathbf{b})$, in order to prove the first inequality of (4.13), it is enough to show that

$$\nu_{i+1}(\mathbf{b})|\mathbf{w}_{i+1}| > \epsilon_{i+1}\nu_{i+1}(\mathbf{b})|\mathbf{w}_{i+1}| + (k + j),$$

which is equivalent to

$$|\mathbf{w}_{i+1}| > \frac{k + j}{\nu_{i+1}(\mathbf{b})(1 - \epsilon_{i+1})}.$$

But $\epsilon_i < 1/2$, so

$$\frac{k + j}{\nu_{i+1}(\mathbf{b})(1 - \epsilon_{i+1})} < \frac{k + j}{\nu_{i+1}(\mathbf{b})(\epsilon_i - \epsilon_{i+1})} < |\mathbf{w}_{i+1}|.$$

To verify the second inequality of (4.13) we note that this is equivalent to

$$|\mathbf{w}_{i+1}| > \frac{k + j}{\nu_{i+1}(\mathbf{b})(\epsilon_i - \epsilon_{i+1})},$$

which is given in the hypotheses.

So, we may conclude that $g_{i,\mathbf{b}}(x, y)$ is a decreasing function of x and an increasing function of y. Since $x \geq 0$ and $y < |\mathbf{w}_{i+1}| + j$, we achive the given upper bound by setting $x = 0$ and $y = |\mathbf{w}_{i+1}| + j$. □

Now we use this upper bound in order to show that $g_{i,\mathbf{b}}$ converges to the constant zero function. Here we will use the requirements for μ-good sequences (2.2) and (2.3).

LEMMA 9.6. *Let* $\mathbf{b} \in \mathcal{D}_{\mu,k}$ *for* $k \in R(W)$. *Then* $\lim_{i\to\infty} g_{i,\mathbf{b}} = 0$.

DÉMONSTRATION. An application of Lemma 9.5 yields

$$\begin{aligned} g_{i,\mathbf{b}}(x,y) \le g_{i,\mathbf{b}}(0,|\mathbf{w}_{i+1}|+j) &= \frac{(L_{i-1}+\epsilon_i\nu_i(\mathbf{b})\ell_i|\mathbf{w}_i|+(k+j)\ell_i)+|\mathbf{w}_{i+1}|+j}{\phi_{L_i}(\mathbf{b})+\nu_{i+1}(\mathbf{b})\left(|\mathbf{w}_{i+1}|+j\right)} \\ &\le \frac{\sum_{k=1}^{\ell-2}\ell_k\left(\mathbf{w}_k+j\right)+\ell_{i-1}\left(\mathbf{w}_{i-1}+j\right)+\epsilon_i\nu_i(\mathbf{b})\ell_i|\mathbf{w}_i|+(k+j)\ell_i)+|\mathbf{w}_{i+1}|+j}{\ell_i\nu_i(\mathbf{b})|\mathbf{w}_i|} \\ &= \frac{\sum_{k=1}^{\ell-2}\ell_k\left(|\mathbf{w}_k|+j\right)}{\ell_i\nu_i(\mathbf{b})|\mathbf{w}_i|}+\frac{\ell_{i-1}\left(|\mathbf{w}_{i-1}|+j\right)}{\ell_i\nu_i(\mathbf{b})|\mathbf{w}_i|}+\epsilon_i+\frac{k+j}{\nu_i(\mathbf{b})|\mathbf{w}_i|}+\frac{|\mathbf{w}_{i+1}|+j}{\ell_i\nu_i(\mathbf{b})|\mathbf{w}_i|} \end{aligned}$$

Now we focus on these five terms. For the first one we apply (2.2) to get

$$\frac{\sum_{k=1}^{\ell-2}\ell_k\left(|\mathbf{w}_k|+j\right)}{\ell_i\nu_i(\mathbf{b})|\mathbf{w}_i|} < \frac{i\ell_{i-2}(\mathbf{w}_{i-2}+j)}{\ell_i\nu_i(\mathbf{b})|\mathbf{w}_i|} < \left(\frac{\ell_{i-2}(|\mathbf{w}_{i-2}|+j)}{\ell_{i-1}(|\mathbf{w}_{i-1}|+j)}\right)\left(\frac{i\ell_{i-1}(|\mathbf{w}_{i-1}|+j)}{\ell_i\nu_i(\mathbf{b})|\mathbf{w}_i|}\right) = o(1).$$

An application of (2.3) yields for the second one that

$$\frac{\ell_{i-1}\left(|\mathbf{w}_{i-1}|+j\right)}{\ell_i\nu_i(\mathbf{b})|\mathbf{w}_i|} = o(1). \qquad \square$$

The third, fourth and fifth clearly are $o(1)$ and the lemma follows.

Now we have all the tools needed for the proof of our Main Theorem.

PROOF OF MAIN THEOREM 9.2. Let $\mathbf{b} \in \mathcal{D}_{\mu,k}$ for $k \in R(W)$. Since $(\epsilon_{i-1}-\epsilon_i)^{-1} = o(|\mathbf{w}_i|)$, there exists n large enough so that $|\mathbf{w}_i|$ and $|\mathbf{w}_{i+1}|$ satisfy the hypotheses of Lemma 9.5.

Since $\lim_{n\to\infty} i(n) = \infty$, we conclude by applying Lemma 9.6 in (4.7) that

$$\lim_{n\to\infty}\left|\frac{N_n(\mathbf{b},\omega)}{\phi_n(\mathbf{b})}-1\right| = 0$$

which implies that

$$\lim_{n\to\infty}\frac{N_n(\mathbf{b},\omega)}{n} = \mu(\mathbf{b}).$$

On the contrary let $\mathbf{b} \in A^k \setminus \mathcal{D}_{\mu,k}$. Since

$$\begin{aligned} 1 &= \lim_{n\to\infty}\sum_{\mathbf{b}'\in A^k}\frac{N_n(\mathbf{b}',\omega)}{n} \\ &= \sum_{\mathbf{b}'\in\mathcal{D}_{\mu,k}}\lim_{n\to\infty}\frac{N_n(\mathbf{b}',\omega)}{n}+\sum_{\mathbf{b}'\in A^k\setminus\mathcal{D}_{\mu,k}}\lim_{n\to\infty}\frac{N_n(\mathbf{b}',\omega)}{n} \\ &= \sum_{\mathbf{b}'\in\mathcal{D}_{\mu,k}}\mu(\mathbf{b}')+\sum_{\mathbf{b}'\in A^k\setminus\mathcal{D}_{\mu,k}}\lim_{n\to\infty}\frac{N_n(\mathbf{b}',\omega)}{n} \\ &= 1+\sum_{\mathbf{b}'\in A^k\setminus\mathcal{D}_{\mu,k}}\lim_{n\to\infty}\frac{N_n(\mathbf{b}',\omega)}{n} \end{aligned}$$

and $N_n(\mathbf{b}',\omega) \ge 0$ we get that

$$\lim_{n\to\infty}\frac{N_n(\mathbf{b},\omega)}{n} = 0 = \mu(\mathbf{b}).$$

Therefore combining the two limits from above we get for $\mathbf{b} \in A^k$ that

$$\lim_{n\to\infty} \frac{N_n(\mathbf{b},\omega)}{n} = \mu(\mathbf{b}),$$

which implies that $\omega \in \mathcal{N}_{\mu,k}$. □

5. Applications

In the following subsections we show different numeration systems in which our construction provides normal numbers. In particular, we consider the q-ary expansions, Lüroth series expansions, β-expansions and continued fraction expansions. We only have restrictions on the concatenation in the case of β-expansions; all other examples are in the full-shift. It is easy to combine our construction for β-expansions and continued fractions in order to get constructions for α-continued fractions (*cf.* Nakada [**165**]) or Rosen-continued fractions [**200**], which have an infinite digit set with restrictions on the concatenation of words.

The main ingredient in all our constructions is the following lemma which is a combination of our results in Section 3 and Main Theorem 9.2.

LEMMA 9.7. *Let μ be a shift-invariant probability measure and let $(\nu_i)_{i\geq 1}$ be an approximation scheme for μ. Suppose that $q_i \geq 2$, M_i and ℓ_i are sequences of positive integers such that*

$$M_i \geq (\min\{\mu(\mathbf{b}) : \mathbf{b} \in \mathcal{D}_{\nu_i,i}\})^{-1} \quad \text{and} \quad q_i^{2i} = o(M_i) \tag{5.1}$$

and $(\mathbf{p}_{q_i,i,M_i}, \ell_i)$ is μ-good for the approximation scheme $(\nu_i)_{i\geq 1}$. Then the sequence $\omega = \mathbf{w}_1^{\odot\ell_1} \odot \mathbf{w}_2^{\odot\ell_2} \odot \cdots$ is μ-normal.

5.1. Normal in base q. Let $A = \{0, 1, \ldots, q-1\}$. In this example we take as language the full-shift A^* and therefore we do not have any restrictions on the concatenation, *i.e.* $j = 0$. Let

$$\nu(t) = \begin{cases} \frac{1}{q} & \text{if } 0 \leq t \leq q-1 \\ 0 & \text{if } t \geq q. \end{cases}$$

For every $i \in \mathbb{N}$ and $\mathbf{b} = b_1 \ldots b_i$, define $\nu_i(\mathbf{b}) = \prod_{t=1}^{i} \nu(b_t)$. Clearly for $\mathbf{b} \in A^*$ we have $\mu(\mathbf{b}) = q^{-|\mathbf{b}|}$ and $\nu_i \to \mu$.

Let $q_i = q$, $M_i = q^{2i} \log i$, $\ell_i = i^{2i}$, and put $\mathbf{w}_i = \mathbf{p}_{q,i,M_i}$, so $iq^{2i} \log i \leq |\mathbf{w}_i| \leq iq^{2i} \log i + iq^i$. A short computation shows that (2.1), (2.2), (2.3), and (5.1) hold with $\epsilon_i = 1/\sqrt{i}$. Thus, by Lemma 9.7, the number whose digits of its q-ary expansion are formed by $\omega = \mathbf{w}_1^{\odot\ell_1} \odot \mathbf{w}_2^{\odot\ell_2} \odot \cdots$ is normal in base q.

5.2. Arbitrary measures. Let $A = \mathbb{N} \cup \{0\}$ and let μ be a shift-invariant measure on $A^{\mathbb{N}}$. We first need to define a sequence of measures (ν_i) that converges weakly to μ. Consider a word $\mathbf{b} = b_1 \ldots b_i$. If there is an index n such that $b_n > i$, then let $\nu_i(\mathbf{b}) = 0$. Let $S = \{n : b_n = i\}$. If $S = \emptyset$, then let $\nu_i(\mathbf{b}) = \mu(\mathbf{b})$. If $S \neq \emptyset$, then let

$$\nu_i(\mathbf{b}) = \sum_{\mathbf{b}'} \mu(\mathbf{b}'),$$

where the sum is over all words $\mathbf{b}' = b'_1 \ldots b'_k$ such that for each index n in S, $b'_n \geq i$. Set

$$M_i = \left\lceil \max\left(i^{2i} \log i, (\inf\{\mu(\mathbf{b}) : \mathbf{b} \in \mathcal{D}_{\nu_i,i}\})^{-1} \right) \right\rceil,$$

$\mathbf{w}_i = \mathbf{p}_{i,i,M_i}$, $j = 0$, $\ell_1 = 1$, and

$$\ell_i = \left\lceil \log i \cdot \max\left(\frac{M_{i+1} + (i+1)^{i+1}}{M_i}, \left(\frac{M_{i-1} + (i-1)^{i-1}}{M_i} \right) \cdot i\ell_{i-1} \right) \right\rceil \text{ for } i > 1.$$

We note that $iM_i \le |\mathbf{w}_i| \le i(M_i + i^i)$, so

$$\begin{aligned} \frac{\ell_{i-1}}{\ell_i} \cdot \frac{|\mathbf{w}_{i-1}|}{|\mathbf{w}_i|} \cdot i &\le \frac{\ell_{i-1}}{\ell_i} \cdot \frac{(i-1)\left(M_{i-1} + (i-1)^{i-1}\right)}{iM_i} \cdot i \\ &< \frac{\ell_{i-1}}{\left(\frac{M_{i-1}+(i-1)^{i-1}}{M_i}\right) \cdot i\ell_{i-1} \cdot \log i} \frac{M_{i-1} + (i-1)^{i-1}}{M_i} \cdot i = \frac{1}{\log i} \to 0 \end{aligned}$$

and

$$\begin{aligned} \frac{1}{\ell_i} \cdot \frac{|\mathbf{w}_{i+1}|}{|\mathbf{w}_i|} &\le \frac{1}{\ell_i} \cdot \frac{(i+1)\left(M_{i+1} + (i+1)^{i+1}\right)}{iM_i} \\ &\le \frac{1}{\frac{M_{i+1}+(i+1)^{i+1}}{M_i} \cdot \log i} \cdot \frac{1+1}{1} \cdot \frac{M_{i+1} + (i+1)^{i+1}}{M_i} = \frac{2}{\log i} \to 0. \end{aligned}$$

Therefore, conditions (2.1), (2.2), (2.3), and (5.1) hold with $\epsilon_i = 1/\sqrt{i}$. Thus, by Lemma 9.7, the infinite word $\omega = \mathbf{w}_1^{\odot \ell_1} \odot \mathbf{w}_2^{\odot \ell_2} \odot \cdots\cdots$ is μ-normal.

5.3. Lüroth series expansions. This example may be modified to construct normal numbers with respect to *Generalized Lüroth series expansions* (see [**62**] for a definition of these expansions.) Put

$$\nu_i(t) = \begin{cases} 0 & t = 0, 1 \\ \frac{1}{t(t-1)} & 2 \le t \le i+1 \\ \frac{1}{i+1} & t = i+2 \\ 0 & t > i+2 \end{cases}$$

and

$$\mu(t) = \begin{cases} 0 & i = 0, 1 \\ \frac{1}{t(t-1)} & t \ge 2 \end{cases}$$

For $\mathbf{b} = b_1 \dots b_i$, define $\nu_i(\mathbf{b}) = \prod_{t=1}^{i} \nu_i(b_t)$ and $\mu(\mathbf{b}) = \prod_{t=1}^{i} \mu(b_t)$. Clearly, $\nu_i \to \mu$. Next, we let $j = 0$, $q_i = i + 2$, $M_i = \max(3!^2, i^{2i} \log i)$, $\ell_i = \left\lfloor i^2 \log i \right\rfloor$, and $\mathbf{w}_i = \mathbf{p}_{i+2,i,M_i}$. Note that for all $i \ge 1$

$$M_i \ge (i+1)!^2 > \left(\min\{\mu(\mathbf{b}) : \mathbf{b} \in \mathcal{D}_{\nu_i,i}\}\right)^{-1}.$$

Since conditions (2.1), (2.2), (2.3), and (5.1) hold, we deduce by an application of Lemma 9.7, that the number whose digits of its Lüroth series expansions are formed by $\mathbf{w}_1^{\odot \ell_1} \odot \mathbf{w}_2^{\odot \ell_2} \odot \cdots$ is normal with respect to the Lüroth series expansions.

This construction has been partially improved (by lowering the number of repetitions) in a recent paper by Vandehey [**243**].

5.4. Unfair coin. We note that already Postnikov and Pyateckiĭ [**189**] used the Champernowne word for such a construction. However, since it is an easy application of Lemma 9.7 we state this example here for completeness.

Let $p \in (0,1), p \neq 1/2$. Here, we consider measures ν_i where

$$\nu_i(t) = \begin{cases} p & \text{if } t = 0 \\ 1-p & \text{if } t = 1 \\ 0 & \text{if } t > 1 \end{cases}.$$

For $\mathbf{b} = b_1 \dots b_i$, let $\nu_i(\mathbf{b}) = \prod_{t=1}^{i} \nu_i(b_t)$ and $\mu = \nu_1$. Set

$$M_i = \left(\frac{1}{\min(p, 1-p)} \right)^{2i},$$

$j = 0$, $\ell_i = i^{2i}$, and put $\mathbf{w}_i = \mathbf{p}_{2,i,M_i}$. Then $\mathbf{w}_i$ is $(1/\sqrt{i}, \sqrt{i}, \nu_i)$-normal and using Lemma 9.7 we get that $\omega = \mathbf{w}_1^{\odot \ell_1} \odot \mathbf{w}_2^{\odot \ell_2} \odot$ is μ-normal.

5.5. β-expansions. Since the paddinge size depends on the expansion of 1 we denote by $\mathrm{d}_\beta(1) = b_1 \dots b_t (b_{t+1} \dots b_{t+p})^\ell$ the β-expansion of 1. If 1 has a finite expansion then we set $p = 0$. We are looking for the longest possible sequence of zeroes occurring in the expansion of 1. As one easily checks, the longest occurs if $b_1 = \dots = b_{t+p-1} = 0$ and $b_{t+p} \neq 0$. Thus we can set the padding size j to be

$$j = t + p.$$

We wish to minimize the length of a cylinder set defined by a word of length ℓ. Define

$$\phi_\beta(\ell) = \begin{cases} 1 & \text{if } 1 \leq \ell \leq t \\ r & \text{if } t + (r-2)p \leq \ell \leq t + (r-1)p \end{cases}.$$

Then the length of this interval is at least $\beta^{-(t+\phi_\beta(\ell)p)}$. We use the fact that $\mu_\beta(I) \geq (1 - 1/\beta)\lambda(I)$ and put

$$M_i = \max\left(\frac{\beta^{t+\phi_\beta(i)p}}{1 - \frac{1}{\beta}}, \lceil \beta \rceil^{2i} \log i \right).$$

Put $\mathbf{w}_i = \mathbf{p}_{\lceil \beta \rceil, i, M_i}$ and $q_i = \lceil \beta \rceil$. Note that $\lim_{i \to \infty} \frac{\phi(i)}{i/p} = 1$, so for large i

$$(i+j) \lceil \beta \rceil^{2i} \log i \leq |\mathbf{w}_i| \leq (i+j) \left(\lceil \beta \rceil^{2i} \log i + \lceil \beta \rceil^{i} \right)$$

Thus, for large i

$$|\mathbf{w}_i| \approx i \lceil \beta \rceil^{2i} \log i.$$

Put $\ell_i = i^{2i}$ and the computation follows the same lines as above.

5.6. Continued fraction expansions. For a word $\mathbf{b} = b_1 \dots b_i$, let $\Delta_{\mathbf{b}}$ be the set of all real numbers in $(0,1)$ whose first i digits of it's continued fraction expansions are equal to $\mathbf{b}$. Put

$$\mu(\mathbf{b}) = \frac{1}{\log 2} \int_{\Delta_{\mathbf{b}}} \frac{dx}{1+x}.$$

If there is an index n such that $b_n > i$, then let $\nu_i(\mathbf{b}) = 0$. Let $S = \{n : b_n = i\}$. For $i < 8$, set $\nu_i(\mathbf{b}) = \mu(\mathbf{b})$. For $i \geq 8$, if $S = \emptyset$, then let $\nu_i(\mathbf{b}) = \mu(\mathbf{b})$. If $S \neq \emptyset$, then let

$$\nu_i(\mathbf{b}) = \sum_{\mathbf{b}'} \mu(\mathbf{b}'),$$

where the sum is over all words $\mathbf{b}' = b'_1 \dots b'_i$ such that for each index n in S, $b'_n \geq i$.

Put $m_i = \min_{\mathbf{b}\in\mathcal{D}_{\nu_i},|\mathbf{b}|=i} \nu_i(\mathbf{b})$. We wish to find a lower bound for m_i. If $\mathbf{b} = b_1 \dots b_k$, then let

$$\frac{p_k}{q_k} = \frac{1}{b_1 + \cfrac{1}{b_2 + \ddots + \cfrac{1}{b_k}}}.$$

It is well known that $\lambda(\Delta_{\mathbf{b}}) = \frac{1}{q_k(q_k+q_{k-1})}$ and $\mu(\mathbf{b}) > \frac{1}{2\log 2}\lambda(\Delta_{\mathbf{b}})$.

Thus, we may find a lower bound for m_i by minimizing $(q_i(q_i + q_{i-1}))^{-1}$ for words $\mathbf{b}$ in $\mathcal{D}_{\nu_i}$. The minimum will occur for $\mathbf{b} = ii \dots i$. It is known that $q_n = iq_{n-1} + q_{n-2}$ if we set $q_0 = 1$ and $q_1 = i$. Set

$$r_1 = \frac{i + \sqrt{i^2+4}}{2}, \quad r_2 = \frac{i - \sqrt{i^2+4}}{2}.$$

Then

$$q_n = \frac{r_1^{n+1} - r_2^{n+1}}{\sqrt{i^2+4}}.$$

Thus,

$$\frac{1}{q_i(q_i+q_{i-1})} = \frac{i^2+4}{(r_1^{i+1} - r_2^{i+1})((r_1^{i+1} + r_1^i) - (r_2^{i+1} - r_2^i))} > \frac{\log 2}{i^{2i}} \text{ for } i \geq 8.$$

Thus, $m_i > \frac{1}{2\log 2}\left(\frac{\log 2}{i^{2i}}\right) = \frac{1}{2}i^{-2i}$. Let $M_i = 2i^{2i}\log i$, $j = 0$, $\mathbf{w}_i = \mathbf{p}_{i+1,i,M_i}$. Set $\ell_i = 0$ for $i < 8$ and $\ell_i = \lfloor i^2 \log i \rfloor$ for $i \geq 8$. Then for $i \geq 9$

$$\frac{\ell_{i-1}}{\ell_i}\frac{|\mathbf{w}_{i-1}|}{|\mathbf{w}_i|}i < \frac{2(i-1)^{2i-1} + i^{i-1}}{2i^{2i}} = \left(1 - \frac{1}{i}\right)^{2i}\frac{1}{i-1} + \frac{1}{2i^{i+1}} \to 0$$

and

$$\frac{|\mathbf{w}_{i+1}|}{\ell_i|\mathbf{w}_i|} \leq \frac{2(i+1)^{2i+3} + (i+2)^{i+1}}{i^2\log i \cdot 2i^{2i+1}} = \left(1 + \frac{1}{i}\right)^{2i}\frac{(i+1)^3}{i^3\log i} + o(i^{-i}) \to 0.$$

By Lemma 9.7 the number whose digits of its continued fraction expansions are formed by $\mathbf{w}_1^{\odot\ell_1} \odot \mathbf{w}_2^{\odot\ell_2} \odot \cdots$ is normal with respect to the continued fraction expansions.

Acknowledgment

Research of the second author is partially supported by the U.S. NSF grant DMS-0943870.

Parts of this research work were done when the authors were visiting the Department of Analysis and Computational Number Theory at Graz University of Technology. Their stay was supported by FWF project P26114. The authors thank the institution for its hospitality.

The authors thank the anonymous referee, who read very carefully the manuscript and his/her suggestions improve considerably the presentation of the results.

Chapitre 10

Computable Absolutely Pisot Normal Numbers

This chapter is joint work with Adrian-Maria Scheerer and Robert Tichy and will appear in the *Acta Arithmetica*.

1. Introduction

In this paper we are interested in simultaneous normality to several bases. In particular, we analyze the order of convergence to normality of an absolutely normal number generated by an algorithm of Becher, Heiber and Slaman (Section 2) and are concerned with normality to non-integer bases. We give an algorithmic construction of a real number that is normal to each base from a given sequence of Pisot numbers (Section 3 and Section 4).

1.1. Normality to a single base. A real number $x \in [0,1)$ is called *simply normal to base* b, $b \geq 2$ an integer, if in its b-ary expansion

$$x = \sum_{n \geq 1} a_n b^{-n}, \ a_n \in \{0, \dots, b-1\}$$

every digit $d \in \{0, 1, \dots b-1\}$ appears with the expected frequency $\frac{1}{b}$. The number x is called *normal to base* b if each of x, bx, b^2x, ... is simply normal to every base b, b^2, b^3, ... This is equivalent (see e.g. [**56**, Chapter 4]) to the property that all digital blocks of arbitrary length k appear with the expected frequency, i.e. if for all $k \geq 1$ and all $d \in \{0, \dots, b-1\}^k$,

$$\lim_{N \to \infty} \frac{1}{N} |\{1 \leq n \leq N : (a_n, \dots, a_{n+k-1}) = d\}| = \frac{1}{b^k}. \tag{1.1}$$

Furthermore, Pillai [**185**] showed that x is normal to base b if and only if it is simply normal to every base b, b^2, b^3, ...

Normal numbers were introduced by Borel [**43**] in 1909. He showed that almost all real numbers (with respect to Lebesgue measure) are simply normal to all bases $b \geq 2$, thus *absolutely normal* (see Section 1.3). It is a long standing open problem to show that important real numbers such as $\sqrt{2}$, $\ln 2, e, \pi, \dots$ are normal, for instance in decimal expansion. There has only been little progress in this direction in the last decades, see e.g. [**18**].

However, specifically constructed examples of normal numbers are known. Champernowne in 1933 [**58**] showed that the real number constructed by concatenating the expansions in base 10 of the positive integers, i.e.

$$0,1\,2\,3\,4\,5\,6\,7\,8\,9\,10\,11\dots,$$

is normal to base 10. This construction has been extended in various directions (*cf.* Erdős and Davenport [**63**], Schiffer [**214**], Nakai and Shiokawa [**166**], Madritsch, Thuswaldner and Tichy [**146**], Scheerer [**210**]).

1.2. Discrepancy of normal numbers. The *discrepancy* of a sequence $(x_n)_{n\geq 1}$ of real numbers is defined as

$$D_N(x_n) = \sup_J \left| \frac{1}{N} |\{1 \leq n \leq N : x_n \bmod 1 \in J\}| - \lambda(J) \right|,$$

where the supremum is extended over subintervals $J \subseteq [0,1)$ and where λ denotes the Lebesgue measure. A sequence is *uniformly distributed modulo* 1 if its discrepancy tends to zero as $N \to \infty$.

It is known [**251**] that x is normal to base b if and only if the sequence $(b^n x)_{n\geq 1}$ is uniformly distributed modulo 1. Hence x is normal to base b if and only if $D_N(b^n x) \to 0$ as $N \to \infty$. It is thus a natural quantitative measure for the normality of x to base b to consider the discrepancy of the sequence $(b^n x)_{n\geq 1}$.

Answering a question of Erdős, in 1975 Philipp [**184**] showed a law of the iterated logarithm for discrepancies of lacunary sequences which implies $D_N(b^n x) = O(\sqrt{\log\log N/N})$ almost everywhere. Recently, Fukuyama [**86**] was able to determine

$$\limsup_{N\to\infty} \frac{D_N(b^n x)\sqrt{N}}{\sqrt{\log\log N}} = c(b) \quad \text{a.e.},$$

for some explicit positive constant $c(b)$. Schmidt [**219**] showed that there is an absolute constant $c > 0$ such that for any sequence $(x_n)_{n\geq 1}$ of real numbers $D_N(x_n) \geq c\frac{\log N}{N}$ holds for infinitely many N. Schiffer [**214**] showed that the discrepancies of constructions of normal numbers in the spirit of Champernowne satisfy upper bounds of order $O(\frac{1}{\log N})$. Levin [**136**] constructed for any integer $b \geq 2$ a real number α such that $D_N(b^n\alpha) = O(\frac{(\log N)^2}{N})$. It is an open question whether there exist an integer $b \geq 2$ and a real number x with optimal discrepancy bound $D_N(b^n x) = O(\frac{\log N}{N})$.

1.3. Absolute normality and order of convergence. A number x is called *absolutely normal* if it is normal to any integer base $b \geq 2$. Since normality to base b is equivalent to simple normality to all bases b^n, $n \geq 1$, absolute normality is equivalent to simple normality to all bases $b \geq 2$.

Since most constructions of numbers normal to a single base b are concatenations of the b-ary expansions of $f(n)$, $n \geq 1$, where f is a positive-integer-valued increasing function, they essentially depend on the choice of the base b. Therefore they cannot be used for producing absolutely normal numbers.

All known examples of absolutely normal numbers have been established in the form of algorithms [1] that output the digits of this number to some base one after the other. In 1917, Lebesgue [**133**] and Sierpinski [**228**] developed a method for the determination of an absolutely normal number. In Sierpinski's method, for instance, this number is given as the infimum of a certain set. This construction was made computable by Becher and Figueira [**20**] who gave a recursive formulation of Sierpinski's construction. Other algorithms for constructing absolutely normal numbers are due to Turing [**241**] (see also Becher, Figueira and Picchi [**21**]), Schmidt [**218**] (see also Scheerer [**209**]) and Levin [**135**] (see also Alvarez and Becher [**15**]).

There seems to be a trade-off between the complexity of the algorithms and the speed of convergence of the corresponding discrepancies. The discrepancies satisfy upper bounds of the order $O(N^{-1/6})$ (Sierpinski), $O(N^{-1/16})$ (Turing), $O((\log N)^{-1})$ (Schmidt) and $O(N^{-1/2}(\log N)^3)$

1. With the exception of Chaitin's constant, which is absolutely normal but not computable [**57**].

(Levin). All algorithms, except the one due to Schmidt, need double exponential many mathematical operations to output the first N digits of the produced absolutely normal number. Schmidt's algorithm requires exponentially many mathematical operations (see [**15**, **21**, **209**] and the original articles [**135**, **218**, **228**, **241**]).

In light of Philipp's result [**184**], all known constructions of absolutely normal numbers satisfy orders of convergence to normality larger than that for almost all real numbers. It is unknown whether there exists a real number x such that for every integer base $b \geq 2$, $D_N(b^n x) = O(\frac{(\log N)^\delta}{N})$ for some $\delta \geq 1$.

In Section 2 we are interested in another construction of an absolutely normal number which is due to Becher, Heiber and Slaman [**22**]. They established an algorithm which computes the digits of an absolutely normal number in polynomial time. We show (Theorem 10.8) that the corresponding discrepancy is slightly worse than $O(\frac{1}{\log N})$, and that at a small loss of computational speed the discrepancy can in fact be $O(\frac{1}{\log N})$.

1.4. Normality to non-integer bases. Section 3 of the present article treats normality in a context where the underlying base is not necessarily integer. Let $\beta > 1$ be a real number. Expansions of real numbers to base β, so-called *β-expansions*, were introduced and studied by Rényi [**194**] and Parry [**177**] and later by many authors from an arithmetic and ergodic-theoretic point of view.

In the theory of β-expansions it is natural to consider *Pisot numbers* β, i.e. real algebraic integers $\beta > 1$, such that all its conjugates lie inside the (open) unit disc. A real number x is called *normal to base* β, or β-normal, if the sequence $(\beta^n x)_{n\geq 1}$ is uniformly distributed modulo 1 with respect to the unique entropy maximizing measure for the underlying transformation $x \mapsto \beta x \bmod 1$ (see Section 3.1). A real number is called *absolutely Pisot normal* if it is normal to all bases that are Pisot numbers. Since there are only countably many Pisot numbers, the Birkhoff ergodic theorem implies that almost all real numbers are in fact absolutely Pisot normal.

The main result of Section 3 is an algorithm that computes an absolutely Pisot normal number. More generally, for a sequence $(\beta_j)_{j\geq 1}$ of Pisot numbers, we construct a real number x that is normal to each of the bases β_j, $j \geq 1$ (Section 3.3 and Theorem 10.14). Bearing in mind that the set of computable real numbers is countable, we thus show that there is in fact a computable real number that is β_j-normal for each $j \geq 1$.

Our algorithm constructs in each step a sequence of finitely many nested intervals, corresponding to the first finitely many bases considered. This is also the essential idea of the construction of an absolutely normal number by Becher, Heiber and Slaman [**22**]. We need to establish lower and upper bounds for the length of β-adic subintervals in a given interval to control the number of specified digits when changing the base. However, the equivalence (absolute normality) $\Leftrightarrow$ (simple normality to all bases) does not hold for non-integer expansions. Instead, we argue with the concept of *(ε, k)-normality* as introduced by Besicovitch [**39**] and studied in the case of Pisot numbers by Bertrand-Mathis and Volkmann [**38**].

Our algorithm should be compared to the one due to Levin [**135**]. While his construction is not restricted to Pisot numbers, it uses exponential sums and is as such not realizable only with elementary operations. The algorithm we present in Section 3 is completely elementary.

In Section 4 we give explicit estimates of all constants that appear in our algorithm. We use a theorem on large deviations for a sum of dependent random variables to give an estimate for the measure of the set of non-(ϵ, k)-normal numbers of length n (Proposition 10.17). Our approach gives all implied constants explicitly, and as such makes a consequence of the ineffective Shannon-McMillan-Breimann theorem effective. The results of this section might be of independent interest.

1.5. Notation. For a real number x, we denote by $\lfloor x \rfloor$ the largest integer not exceeding x. The fractional part of x is denoted as $\{x\}$, hence $x = \lfloor x \rfloor + \{x\}$. We put $\lceil x \rceil = -\lfloor -x \rfloor$. Two functions f and g are $f = O(g)$ or equivalently $f \ll g$ if there is a x_0 and a positive constant C such that $f(x) \leq Cg(x)$ for all $x \geq x_0$. We mean $\lim_{x\to\infty} f(x)/g(x) = 1$ when we say $f \sim g$ and $g \neq 0$.

When we speak of *words*, we mean finite or infinite sequences of symbols (called letters) of a certain (specified) set, the alphabet. *Blocks* are finite words. The concatenation of two blocks $u = u_1 \dots u_k$ and $v_1 \dots v_l$ is the block $u_1 \dots u_k v_1 \dots v_l$ and is denoted by uv or $u * v$. If u_i for $i \leq m$ are blocks, $*_{i<m} u_i$ is their concatenation in increasing order of i. The *length* of the block $u = u_1 \dots u_k$ is denoted by $\|u\|$ and is in this case equal to k.

We denote by λ the Lebesgue measure.

For a finite set, $|\cdot|$ means its number of elements.

Mathematical operations include addition, subtraction, multiplication, division, comparison, exponentiation and logarithm. *Elementary operations* take a fixed amount of time. The cost of mathematical operations depends on the digits of the input or on the desired precision of the output. Addition or subtraction of two n-digit numbers takes $O(n)$ elementary operations, multiplication or division of two n-digit numbers takes $O(n^2)$ elementary operations, and to compute the first n digits of exp and log takes $O(n^{5/2})$ elementary operations. These estimates are crude but sufficient for our purposes.

The complexity of a computable function f is the time it takes to compute the first N values $f(i)$, $1 \leq i \leq N$. The algorithm we analyze outputs the digits of a real number X to some base. By the complexity of the algorithm we mean the time it takes to output the first N digits of X to some base.

2. Discrepancy

In this section, we analyze the speed of convergence to normality of the absolutely normal number produced by the algorithm by Becher, Heiber and Slaman in [**22**]. We follow the notation and terminology therein.

2.1. The Algorithm.

Notation. A *t-sequence* is a nested sequence of intervals $\mathbf{I} = (I_2, \dots, I_t)$, such that I_2 is dyadic and for each base $2 \leq b \leq t-1$, I_{b+1} is a $(b+1)$-adic subinterval of I_b such that $\lambda(I_{b+1}) \geq \lambda(I_b)/2(b+1)$.

Let $x_b(\mathbf{I})$ be the block in base b such that $0.x_b(\mathbf{I})$ is the representation of the left endpoint of I_b in base b. In each step i, the algorithm computes a sequence $\mathbf{I}_i = (I_{i,2}, \dots, I_{i,t_i})$ of nested intervals $I_{i,2} \supset \dots \supset I_{i,t_i}$. If $b \leq t_i$, let $x_b(\mathbf{I}_i) = x_{i,b}$ be the base b representation of the left endpoint of $I_{i,b}$ and let $u_{i+1,b} = u_b(\mathbf{I}_{i+1})$ be such that $x_{i+1,b} = x_{i,b} * u_{i+1,b}$.

If u is a block of digits to base b, the *simple discrepancy* of u in base b is defined as $D(u, b) = \max_{0\leq d<b} |N_d(u)/\|u\| - 1/b|$ where $N_d(u)$ is the number of times the digit d appears in the block u.

Let $k = k(\epsilon, \delta, t)$ be the function

$$k(\epsilon, \delta, t) = \max(\lceil 6/\epsilon \rceil, \lceil -\log(\delta/(2t))6/\epsilon^2 \rceil) + 1.$$

From Lemma 4.1 and 4.2 of [**22**] we further have a function $h = h(t, \epsilon)$ that counts the number of mathematical operations needed to carry out one step of the algorithm. See also Lemma 10.5.

Input. A computable non-decreasing unbounded function $f : \mathbb{N} \to \mathbb{R}$ such that $f(1)$ is known and satisfies $f(1) > h(2, 1)$.

First step. Set $t_1 = 2$, $\epsilon_1 = \frac{1}{2}$, $k_1 = 1$ and $\mathbf{I}_1 = (I_{1,2})$ with $I_{1,2} = [0, 1)$.

Step $i + 1$ for $i \geq 1$. Given are values $t_i = v$, $\epsilon_i = \frac{1}{v}$ and a t_i-sequence $\mathbf{I}_i$.

We want to assign values to t_{i+1}, ϵ_{i+1}. If $i + 1$ is a power of 2, then we carry out the following procedure.

— We spend at most i computational steps on computing the first values $f(1), f(2), \ldots$ of f, obtaining $f(m)$, where m, $1 \leq m \leq i$, is the largest integer such that $f(m)$ has been computed.

— We put $\delta = (8t_i 2^{t_i+v+1} t_i!(v + 1)!)^{-1}$.

— We try to compute $k(\frac{1}{v+1}, \delta, v + 1)$ and $h(v + 1, \frac{1}{v+1})$ in i steps each. If we succeed in computing these values, and if additionally

$$h(v + 1, \frac{1}{v + 1}) < f(m) \tag{2.1}$$

and for each $b \leq t_i$

$$\frac{\lceil \log_2(v + 1) \rceil k(1/(v + 1), \delta, v + 1) + \lceil -\log_2(\delta) \rceil}{\|x_{i,b}\|} < \frac{1}{v + 1}, \tag{2.2}$$

then we define $t_{i+1} = v + 1$ and $\epsilon_{i+1} = \frac{1}{v+1}$. Otherwise, we let $t_{i+1} = t_i = v$, $\epsilon_{i+1} = \epsilon_i = \frac{1}{v}$.

If $i + 1$ is no power of 2, then define $t_{i+1} = t_i = v$, $\epsilon_{i+1} = \epsilon_i = \frac{1}{v}$.

Furthermore, we compute $\delta_{i+1} = (8t_i 2^{t_i+t_{i+1}} t_i! t_{i+1}!)^{-1}$ and

$$k_{i+1} = \max(\lceil 6/\epsilon_{i+1} \rceil, \lceil -\log(\delta_{i+1}/(2t_i))6/\epsilon_{i+1}^2 \rceil) + 1.$$

Then we find a t_{i+1}-sequence $\mathbf{I}_{i+1}$ by means of the following steps.

— We let L be a dyadic subinterval of I_{i,t_i} such that $\lambda(L) \geq \lambda(I_{i,t_i})/4$.

— For each dyadic subinterval J_2 of L of measure $2^{-\lceil \log_2 t_i \rceil k_{i+1}} \lambda(L)$, we find $\mathbf{J} = (J_2, J_3, \ldots, J_{t_{i+1}})$, a t_{i+1}-sequence starting with J_2.

— Finally we choose $\mathbf{I}_{i+1}$ to be the leftmost of the t_{i+1} sequences $\mathbf{J}$ considered above such that for each $b \leq t_i$, $D(u_b(\mathbf{J}), b) \leq \epsilon_{i+1}$.

Output. Let X be the unique real number in the intersection of the intervals of the sequences $\mathbf{I}_i$. In base b we have $X = \lim_{i\to\infty} 0.x_{i,b} = 0.*_{i\geq 1} u_{i,b}$. It is the content of Theorem 3.9 in [**22**] that X is absolutely normal.

2.2. Speed of convergence to normality. In this section we estimate the discrepancy $D_N(b^n X)$ for integer $b \geq 2$. Two factors play a role : How many digits in each step are computed, and how rapidly ϵ_i decays to zero. By virtue of the algorithm, at least one digit is added in each step, and ϵ_i can decay at most as fast as $O(\frac{1}{\log i})$. As can be expected from the algorithm, the discrepancy depends both on growth and complexity of f.

It was shown in [**22**] that to output the first N digits of X, the algorithm requires time $O(N^2 f(N))$.

We begin our analysis by first showing that in each step of the algorithm not too many digits are attached.

LEMMA 10.1 (Lemma 3.3 in [**22**]). *For an interval I and a base b, there is a b-adic subinterval I_b such that $\lambda(I_b) \geq \lambda(I)/(2b)$.*

LEMMA 10.2. *If i is large enough, then $1 \leq \|u_{i,b}\| \ll (\log i)^A$ for $A > 3$. Thus $i \ll \|x_{i,b}\| \ll i(\log i)^A$.*

DÉMONSTRATION. We assume the base b to be fixed and i large enough such that $t_{i+1} \geq b$. In step $i+1$ we have the following sequence of nested subintervals :

$$I_{i,b} \supset \ldots \supset I_{i,t_i} \supset L \supset I_{i+1,2} \supset \ldots \supset I_{i+1,b}. \tag{2.3}$$

By Lemma 10.1, and the choice of $I_{i+1,2}$, we know the following lower bounds on the measures of the intervals in (2.3). We have $\lambda(I_{i,t_i}) \geq \lambda(I_{i,b})/(2^{t_i-b}t_i!/b!)$, $\lambda(L) \geq \lambda(I_{i,t_i})/4$, $\lambda(I_{i+1,2}) = 2^{-\lceil \log_2 t_i \rceil k_{i+1}}\lambda(L)$ and $\lambda(I_{i+1,b}) \geq \lambda(I_{i+1,2})/(2^b b!)$. Combining inequalities yields $\lambda(I_{i+1,b}) \geq \lambda(I_{i,b})/(2^{2+t_i}2^{\lceil \log_2 t_i \rceil k_{i+1}}t_i!)$. Hence in stage $i+1$ we are adding at most $O(t_i + (\log t_i)k_{i+1} + \log t_i!)$ many digits in base b. The way the algorithm is designed only allows for $t_i = O(\log i)$. The growth of k_{i+1} can be analyzed and is $O(t_i^3 \log t_i)$. Hence in stage $i+1$ at most $O(t_i + (\log t_i)k_{i+1} + \log t_i!) = O((\log i)^A)$ digits are added to the b-ary expansion of X, where $A > 3$ to accommodate all double-log factors.

The lower bound on the number of digits added comes from the fact that by the choice of $I_{i+1,2}$, $I_{i+1,b}$ is strictly smaller than $I_{i,b}$, so at least one digit is added in each stage. □

Next, we investigate the conditions involving k and h that are responsible for how fast $t_i \to \infty$ and $\epsilon_i \to 0$ with step i of the algorithm. We start by showing that condition (2.2) on k always holds, provided i is large enough. This involves estimating the growth as well as the complexity of k.

Recall that $k(\epsilon, \delta, t) = \max(\lceil 6/\epsilon \rceil, \lceil -\log(\delta/(2t))6/\epsilon^2 \rceil) + 1$.

LEMMA 10.3. *Let $v \geq 2$ be an integer and $\delta = (8v2^{2v+1}v!(v+1)!)^{-1}$. Then the growth of $k(\frac{1}{v+1}, \delta, v+1)$ is $O(v^3 \log v)$. Furthermore, $k(\frac{1}{v+1}, \delta, v+1)$ can be computed in $O(v^2(\log v)^2)$ elementary operations.*

DÉMONSTRATION. We have for the growth

$$\begin{aligned} k(\frac{1}{v+1}, \delta, v+1) &= \max(\lceil 6(v+1) \rceil, \lceil \log(2(v+1)8v2^{2v+1}v!(v+1)!)6(v+1)^2 \rceil) + 1 \\ &\leq 6(v+1)^2 \left(\log(16v(v+1)) + (2v+1)\log 2 + \log v! + \log(v+1)!\right) + 2 \\ &= O(v^2(\log v + v + v \log v)) \\ &= O(v^3 \log v). \end{aligned}$$

Since in the expression for k we are rounding, the most relevant part is the computation of the significant digits of $\log(16v(v+1)2^{2v+1}v!(v+1)!)$. The argument of this expression is computable with $O(v^2(\log v)^2)$ elementary operations and has $O(v\log v)$ many digits. We only need to compute $O(\log v)$ many digits of the logarithm, which takes another $O((\log v)^{5/2})$ elementary operations. In total this are $O(v^2(\log v)^2)$ many elementary operations. □

COROLLARY 10.4. *For i to be large enough, condition (2.2) on k is always satisfied, i.e. for each $b \le t_i$*

$$\frac{\lceil \log(v+1)\rceil k(1/(v+1),\delta,v+1)+\lceil -\log(\delta)\rceil}{\|x_{i,b}\|} < \frac{1}{v+1}$$

where v is such that $t_i = v = 1/\epsilon_i$.

DÉMONSTRATION. This is a consequence of $k(1/(v+1),\delta,v+1) = O(v^3\log v)$, $\log(1/\delta) = O(v\log v)$, $\|x_{i,b}\| \gg i$ and $v = t_i = O(\log i)$ by the way the algorithm is designed. □

Now we investigate condition (2.1) on h involving f. The function h counts the number of mathematical operations needed to carry out one step of the algorithm. We want to know an upper bound for the growth of h.

LEMMA 10.5. *With $t_i = \frac{1}{\epsilon_i} = O(\log i)$ we have*

$$h(t_i,\epsilon_i) = O(i^{\log^4 i}).$$

This upper bound for h can be computed with i elementary operations, provided i is large enough.

DÉMONSTRATION. The function h decomposes as $h = h_*(h_1 g + h_2 + h_3 + h_4)h_0$ as can be seen from the proof of Lemma 4.2 in [**22**]. Here :

- g (from Lemma 4.1 in [**22**]), is the minimum number of digits sufficient to represent all the endpoints of the intervals that we are working with in one step (squared). We know from Lemma 10.2 that $g = O(i^2(\log i)^{2A})$ for $A > 3$.
- It takes $h_1 g$ many mathematical operations to find a t_{i+1}-sequence for each J_2. We have $h_1 = t_{i+1}$.
- h_2 is the number of mathematical operations needed to compute the base b representation $u_b(\mathbf{J})$ for each $2 \le b \le t_i$. We have $h_2 \le \lceil \log_2 t_i\rceil k_{i+1}$.
- h_3 counts the number of mathematical operations needed to compute thresholds of the form $(1/b+\epsilon_{i+1})\|u_b(\mathbf{J})\|$. We have $h_3 = t_i$.
- h_4 comes from counting occurrences of digits in $u_b(\mathbf{J})$ and comparing with the previously computed thresholds. We have $h_4 \ll t_i(\lceil \log_2 t_i\rceil k_{i+1})^2$.
- h_* is the maximum number of iterations it takes to find a suitable t_{i+1}-sequence. There are $2^{\lceil \log_2 t_i\rceil k_{i+1}}$ many different subintervals J_2 of L, hence $h_* = 2^{\lceil \log_2 t_i\rceil k_{i+1}}$. With $k_{i+1} = O(\log^4 i)$ we obtain $h_* = O(i^{\log^4 i})$.
- Finally, the function h_0 is the number of elementary operations needed to carry out each mathematical operation in one step of the algorithm. Since all values that appear in the calculations of one step of the algorithm are at most exponential in t_i which is at most of order $\log i$, and because the number of elementary operations involved depends only on the number of digits of the numbers involved, h_0 is at most of order $\mathrm{poly}(\log i)$.

These bounds can be seen from Lemma 4.1 and Lemma 4.2 in [**22**]. Combining them gives $h = O(i^{\log^4 i})$.

Remark that, when t_i is bounded by a slower growing function in i such as $\log\log i$, then the significant term in h comes from g and is a power of i. Otherwise h_* is the significant term.

For the complexity of the upper bound for h, note that $i^{\log^4 i}$ can be computed in a power of $\log i$ many elementary operations, so certainly with i elementary operations when i is large enough. □

Lemma 10.5 has the following two immediate corollaries for the speed of convergence to normality of Becher, Heiber and Slaman's algorithm.

PROPOSITION 10.6. *Becher, Heiber, Slaman's algorithm achieves discrepancy of* $D_N(b^n X) = O(\frac{1}{\log N})$ *for* f *computable in real-time with growth* $f \gg i^{\log^4 i}$. *In this case, the complexity is* $O(i^{2+\log^4 i})$.

PROPOSITION 10.7. *If* f *is a polynomial in* i *of degree* d, *then the complexity of* X *is* $O(N^{d+2})$ *but the discrepancy of* $(b^n X)_{n\geq 0}$ *is* $D_N(b^n X) = O_d(\frac{1}{(\log N)^{1/5}})$.

The subscript in O_d indicates that the implied constant might depend on the degree d of f.

DÉMONSTRATION. These corollaries follow by observing that the complexity of f is such that f is for large enough i computed up to the actual value $f(i)$ (i.e. $m = i$) and that either the condition on h, (2.1), is satisfied, hence the discrepancy is optimal, or that condition (2.1) is only satisfied for $e^{(\log i)^{1/5}}$ of the values that it is checked for. □

In a similar manner, using Lemma 10.5, one can show quantitatively how growth and complexity of f influence the discrepancy (and the complexity) of Becher, Heiber, Slaman's algorithm. This can be done for example by measuring complexity and growth of f in the following (crude) way. We denote by $\log_{(k)}$ and $\exp_{(k)}$ the k times iterated logarithm or exponential where $\exp_{(k)} = \log_{(-k)}$, and $\exp_{(0)} = \log_{(0)} = id$. Let c be the integer such that in i elementary operations f can be computed up to a value $f(m)$ with $m \sim \log_{(c)} i$. Let g be the integer such that f grows as $f \sim \exp_{(g)} i$. We allow $g \in \mathbb{Z}$ but c is non-negative.

THEOREM 10.8. *Assume* f *is such that the integers* c *and* g *above can be defined. Then Becher, Heiber, Slaman's algorithm computes an absolutely normal number* X *such that for any base* $b \geq 2$,

$$D_N(b^n X) = O\left(\frac{1}{(\log_{(1-g+c)} N)^{1/5}}\right) \tag{2.4}$$

if $1 - g + c > 0$, *and*

$$D_N(b^n X) = O\left(\frac{1}{\log N}\right) \tag{2.5}$$

otherwise.

DÉMONSTRATION. We have $h \ll \max(\text{poly}(i), e^{t_i^5})$ and $t_i \ll \log i$ by the way the algorithm is defined. t_i only increases if i is a power of two and if $h \leq f(m)$. The latter condition is satisfied for all i large enough if $g - c \geq 1$, and for all i (that are powers of two) that satisfy $i \ll \exp((\exp_{g-c-1}(i))^{1/5})$. With $1/t_i = \epsilon_i$ this gives in this case an upper bound for the discrepancy of order $1/(\log_{(1-g+c)} N)^{1/5}$. □

3. Absolutely Pisot Normal Numbers

In this section, we give an algorithmic construction of a real number that is normal to each base from a given sequence of Pisot numbers. For more information about β-expansions and β-normal numbers see for example the book [**56**]. We have partly followed the notation in [**38**].

3.1. β-expansions of real numbers. Let $\beta > 1$ be a real number. Then each real number $x \in [0, 1)$ has a representation of the form

$$x = \sum_{i=1}^{\infty} \epsilon_i \beta^{-i}, \tag{3.1}$$

with integer digits $0 \le \epsilon_i < \beta$. One way to obtain such a representation is the following. Let T_β be the *β-transformation* $T_\beta : [0, 1) \to [0, 1)$, $x \mapsto \beta x \pmod 1$. Then $\epsilon_i = \lfloor \beta T_\beta^{i-1}(x) \rfloor$ for $i \ge 1$.

Rényi [**194**] showed that there is a unique probability measure μ_β on $[0, 1)$ that is equivalent to the Lebesgue measure and such that μ_β is invariant and ergodic with respect to T_β and has maximum entropy. The measure μ_β satisfies $(1 - \frac{1}{\beta})\lambda \le \mu_\beta \le \frac{\beta}{\beta-1}\lambda$.

Let $c(d)$ be the *cylinder set* corresponding to the block d, i.e. the set of all real numbers in the unit interval whose first $\|d\|$ digits coincide with d. A *β-adic interval* is a cylinder set $c(d)$ for some d.

Let W^∞ be the set of right-infinite words $\omega = \omega_1\omega_2 \dots$ with digits $0 \le \omega_i < \beta$ that appear as the β-expansions of real numbers in the unit interval. Let $\mathcal{L}_n$ be the set of all finite subwords of length n of words $\omega \in W^\infty$ and let $W = \bigcup_{n\ge1} \mathcal{L}_n$. We call the words in W *admissible*.

We have $\beta^n \le |\mathcal{L}_n| \le \frac{\beta}{\beta-1}\beta^n$ for the number of elements of $\mathcal{L}_n$.

For an infinite word $\omega = \omega_1\omega_2 \dots \in W^\infty$ and a block $d = d_1d_2 \dots d_k$ of digits $0 \le d_i < \beta$ we denote by $N_d(\omega, n)$ the number of (possibly overlapping) occurrences of d within the first n letters of ω. If the word ω is finite, we write $N_d(\omega)$ for $N_d(\omega, \|\omega\|)$.

An infinite word $\omega \in W^\infty$ is called *μ_β-normal* if for all $d \in \mathcal{L}_k$,

$$\lim_{n\to\infty} \frac{1}{n} N_d(\omega, n) = \mu_\beta(c(d)).$$

A real number $x \in [0, 1)$ is called *normal to base β* or *β-normal*, if the infinite word $\epsilon_1\epsilon_2 \dots$ defined by its β-expansion (3.1) is μ_β-normal.

For fixed $\epsilon > 0$ and positive integers k, n, a word $\omega \in \mathcal{L}_n$ is called *(ϵ, k)-normal* if for all $d \in \mathcal{L}_k$

$$\mu_\beta(c(d))(1 - \epsilon)\|\omega\| < N_d(\omega) < \mu_\beta(c(d))(1 + \epsilon)\|\omega\|.$$

The set of all (ϵ, k)-normal words in $\mathcal{L}_n$ will be denoted by $E_n(\epsilon, k, \beta)$ and its complement by $E_n^c(\epsilon, k, \beta)$. If β is understood from the context, we simply write $E_n(\epsilon, k)$ and $E_n^c(\epsilon, k)$.

A *Pisot number* β is a real algebraic integer $\beta > 1$ such that all its conjugates have absolute value less than 1, and as usual we include all positive integers $b \ge 2$ in this definition. All Pisot numbers smaller than the golden mean were found by Dufresnoy and Pisot [**69**]. In particular, they showed that the smallest one is the positive root of $x^3 - x - 1$ (called the plastic number) which is approximately $1.32471 > \sqrt[3]{2}$.

3.2. Preliminaries.

LEMMA 10.9 ([**38**, Lemma 3]). *Let $\beta > 1$ be Pisot. For every $\varepsilon > 0$ and positive integer k there exist $\eta = \eta(\varepsilon, k)$, $0 < \eta < 1$, $C = C(\varepsilon, k) > 0$ and $n_0 = n_0(\epsilon, k)$ such that for the number of non-(ϵ, k)-normal words of length n*

$$|E_n^c(\epsilon, k)| < C\,|\mathcal{L}_n|^{1-\eta}$$

holds for all $n \geq n_0$.

In Section 4.2 we give explicit estimates for n_0, C and η.

The following Lemma contains the underlying idea of our construction.

LEMMA 10.10 ([**38**, Lemma 4]). *Let $a_1, a_2, \ldots$ be a sequence of finite words $a_n \in W$ such that $a = a_1 a_2 \ldots \in W^\infty$ and $\|a_n\| \to \infty$ as $n \to \infty$. Suppose that for any $\epsilon > 0$ and any positive integer k there exists an integer $n_0(\epsilon, k)$ such that all a_n with $n \geq n_0(\epsilon, k)$ are (ϵ, k)-normal. If*

$$n = o\left(\|a_1 a_2 \ldots a_n\|\right) \quad \text{and} \quad \|a_{n+1}\| = o(\|a_1 a_2 \ldots a_n\|), \tag{3.2}$$

then the infinite word $a = a_1 a_2 \ldots$ is μ_β-normal.

DÉMONSTRATION. Let $\epsilon > 0$ and $d \in \mathcal{L}_k$. It suffices to show that, as $N \to \infty$,

$$\mu_\beta(c(d))(1-\epsilon)N < N_d(a, N) < \mu_\beta(c(d))(1+\epsilon)N.$$

We have $N_d(a, N) = N_d(a_1 a_2 \ldots a_n, N)$, where n is such that $\|a_1 a_2 \ldots a_{n-1}\| < N \leq \|a_1 \ldots a_n\|$. Then, for N large enough,

$$\begin{aligned} N_d(a_1 \ldots a_n, N) &\leq N_d(a_1 \ldots a_{n_0(\epsilon,k)}) + n(k-1) + N_d(a_{n_0+1}) + \ldots + N_d(a_n) \\ &\leq \text{const}(\epsilon, k) + n(k-1) + \sum_{i=n_0+1}^{n} \mu_\beta(c(d))(1+\epsilon)\|a_i\|. \end{aligned}$$

Dividing by N gives the desired result, assuming conditions (3.2). The calculation for the lower bound for $N_d(a, N)$ is similar. □

LEMMA 10.11. *Let $\beta > 1$ be Pisot. There exists $M \geq 0$ such that for all $n \geq 1$ and all $d \in L_n$ the Lebesgue measure of the cylinder set $c(d)$ satisfies*

$$\beta^{-(M+1)}\beta^{-n} \leq \lambda(c(d)) \leq \beta^{-n}. \tag{3.3}$$

DÉMONSTRATION. This is Proposition 2.6 of [**137**]. □

Following the argument in [**137**], one can take M to be the size of the largest block of consecutive zeros in the modified β-expansion of 1 (see Section 4.1). We give an explicit upper bound on M in Proposition 10.15.

We wish to control the lengths when changing the base. The following is an analogue to Lemma 3.3 in [**22**]; see also Lemma 10.1.

LEMMA 10.12. *Let β be Pisot and M as above. For any interval I there is a β-adic subinterval I_β of I such that $\lambda(I_\beta) \geq \lambda(I)/2\beta^{M+4}$.*

DÉMONSTRATION. We can assume $\lambda(I) > 0$. Let m be the smallest integer such that $\beta^{-m} < \lambda(I)$. Thus $\lambda(I)/\beta \leq \beta^{-m} < \lambda(I)$. If there exists an interval of order m in I, then let I_β be this β-adic interval and we have $\lambda(I_\beta) \geq \lambda(I)/\beta$.

Otherwise there must be a word $a \in \mathcal{L}_m$ such that $\pi(a) \in I$ but neither $\pi(a^-)$ nor $\pi(a^+)$ is in I, where a^- and a^+ are the lexicographically previous or next elements of a of the same length and where $\pi(a)$ is the real number in the unit interval whose β-expansion starts with a. Then by Lemma 10.11 we have that $\lambda(I) < 2\beta^{-m}$. Since $\beta^{-m} < \lambda(I)$ and the smallest Pisot number is bigger than $2^{1/3}$, we get that $2\beta^{-m-3} < \lambda(I)$. Thus there must be a β-adic interval I_β of order $m+3$ in I and we have

$$\lambda(I_\beta) \geq \frac{1}{\beta^{M+1+m+3}} = \frac{1}{2\beta^{M+4}} \cdot \frac{2}{\beta^m} > \frac{\lambda(I)}{2\beta^{M+4}}.$$

□

3.3. The Algorithm.

Notation. Let $(\beta_j)_{j\geq 1}$ be a sequence of Pisot numbers. Let t be a positive integer. A *t-sequence* is a sequence of intervals $\mathbf{I} = (I_1, \ldots, I_t)$ such that for $1 \leq j \leq t$, I_j is β_j-adic, such that for $1 \leq j \leq t-1$, $I_{j+1} \subset I_j$, and such that $\lambda(I_{j+1}) \geq \lambda(I_j)/2\beta_{j+1}^{M_{\beta_{j+1}}+4}$. If we have two β-adic intervals $J \subset I$ then $u_\beta(J)$ means the block of digits that is added to the base β expansion of the numbers in I to obtain the β-expansion of numbers in J. The notation $u_j(\mathbf{J})$ for a t-sequence $\mathbf{J}$ shall mean $u_{\beta_j}(J_j)$. We denoted the dependence on β_j of all appearing constants M, n_0, C and of $\mathcal{L}_n$ explicitly with an β_j.

Input. Given are values $\epsilon_1 = 1$, $k_1 = 1$, $t_1 = 1$ and a sequence $(\beta_j)_{j\geq 1}$ of Pisot numbers β_j.

First step. Let $\mathbf{I}_1$ be a t_1-sequence such that $\mathbf{I}_1 = (I_{1,t_1})$, with $I_{1,t_1} = [0,1)$.

For each $i \in \mathbb{N}$, let $t_i = \lceil \log i \rceil$. Replace the original sequence of bases $(\beta_j)_{j\geq 1}$ by suitably repeating the β_j such that for each i the conditions

$$\max_{1\leq j\leq t_i} \beta_j \leq \beta_1 i, \tag{3.4}$$

$$\max_{1\leq j\leq t_i} M_{\beta_j} \leq (M_{\beta_1}+1)(1+\log i), \tag{3.5}$$

$$\sum_{1\leq j\leq t_i} (M_{\beta_j}+4)\log\beta_j \leq (M_{\beta_1}+4)\log\beta_1(1+\log i) \tag{3.6}$$

are satisfied.

Step $i+1$ for $i \geq 1$. From step i, we have a t_i-sequence $\mathbf{I}_i$ of nested intervals $I_{i,1} \supset \ldots \supset I_{i,t_i}$ where each $I_{i,j}$ is β_j-adic.

Let

$$t_{i+1} = \lceil \log(i+1) \rceil, \quad \epsilon_{i+1} = \frac{1}{t_{i+1}}, \quad k_{i+1} = t_{i+1},$$

$$\delta_{i+1} = \frac{1}{2}\frac{1}{2\beta_1^{M_{\beta_1}+4}}\frac{1}{t_i}\frac{1}{2^{t_i}\prod_{j\leq t_i}\beta_j^{M_{\beta_j}+4}}\frac{1}{2^{t_{i+1}}\prod_{j\leq t_{i+1}}\beta_j^{M_{\beta_j}+4}}.$$

Choose n_{i+1} to be the least integer such that

$$n_{i+1} \geq \max_{j\leq t_{i+1}} \left(n_{\beta_j}(\epsilon_{i+1}, k_{i+1})\right), \tag{3.7}$$

and such that for all $1 \leq j \leq t_{i+1}$

$$\lambda(E^c_{n_{i+1}}(\epsilon_{i+1}, k_{i+1}, \beta_j)) < \delta_{i+1}. \tag{3.8}$$

Furthermore, let

$$v_i = \left\lceil \max_{j=1,\ldots,t_i} \frac{\log \beta_j}{\log \beta_1} \right\rceil.$$

Then we perform the following steps.

— Take L to be a β_1-adic interval of I_{i,t_i} of length $\lambda(L) \geq \lambda(I_{i,t_i}) 2^{-1} \beta_1^{-(M_{\beta_1}+4)}$.

— For each β_1-adic sub-interval J_1 of L with $u_1(J_1) = v_i n_{i+1}$ find a t_{i+1}-sequence $\mathbf{J} = (J_1, \ldots, J_{t_{i+1}})$.

— Choose the "leftmost" of the t_{i+1}-sequences $\mathbf{J}$ such that $u_j(\mathbf{J})$ is $(\varepsilon_{i+1}, k_{i+1})$-normal for $1 \leq j \leq t_i$.

Output. The unique real number X in the intersection of all $I_{i,j}$.

We need to show that the algorithm is well-defined and that the produced number is in fact β_j-normal for all $j \geq 1$.

PROPOSITION 10.13. *The real number X computed by this algorithm is well-defined.*

DÉMONSTRATION. We have to show that in each step $i+1$ there exists at least one t_{i+1}-sequence $\mathbf{J}$. Let $\mathcal{S}$ be the union of the intervals $J_{t_{i+1}}$ over the $|\mathcal{L}^{\beta_1}_{v_i n_{i+1}}|$ many t_{i+1}-sequences $\mathbf{J}$. By definition of the interval L we have that $\lambda(L) \geq \lambda(I_{i,t_i}) 2^{-1} \beta_1^{-(M_{\beta_1}+4)}$. Furthermore for each sequence we have that $\lambda(J_{t_{i+1}}) \geq 2^{-t_{i+1}} \prod_{j=1}^{t_{i+1}} \beta_j^{-(M_{\beta_j}+4)} \lambda(J_1)$. Since the sub-intervals $J_1 \subset L$ form a partition of L we have that $\lambda(\mathcal{S}) \geq 2^{-t_{i+1}} \prod_{j=1}^{t_{i+1}} \beta_j^{-(M_{\beta_j}+4)} \lambda(L)$. Combining these inequalities yields

$$\lambda(\mathcal{S}) \geq 2^{-t_i - t_{i+1} - 1} \prod_{j=1}^{t_i} \beta_j^{-(M_{\beta_j}+4)} \prod_{j=1}^{t_{i+1}} \beta_j^{-(M_{\beta_j}+4)} \lambda(I_{i,1}).$$

Now we calculate the measure of the set $\mathcal{N}$ of non-suitable intervals and show that it is less than $\lambda(\mathcal{S})$. For the length of the added word we have $\|u_1(\mathbf{J})\| \geq v_i n_{i+1}$ and for each $2 \leq j \leq t_{i+1}$ we have $\|u_j(\mathbf{J})\| \geq n_{i+1}$. By the choice of n_{i+1}, the subsets of $I_{i,j}$, where $u_j(\mathbf{J})$ is not $(\varepsilon_{i+1}, k_{i+1})$-normal, have Lebesgue measure less than $\delta_{i+1}\lambda(I_{i,j})$, and hence less than $\delta_{i+1}\lambda(I_{i,1})$. Since we consider t_i many bases, we obtain $\lambda(\mathcal{N}) < t_i \delta_{i+1} \lambda(I_{i,1})$.

Combining the estimates of $\mathcal{N}$ and $\mathcal{S}$ we obtain $\lambda(\mathcal{N}) < \lambda(\mathcal{S})$. Since $\mathcal{N} \subset \mathcal{S}$ there must be a t_{i+1}-sequence $\mathbf{J}$ such that $u_j(\mathbf{J})$ is $(\epsilon_{i+1}, k_{i+1})$-normal for each $1 \leq j \leq t_i$. □

THEOREM 10.14. *Let $(\beta_j)_{j \geq 1}$ be a sequence of Pisot numbers. Then the real number X generated by this algorithm is β_j-normal for each $j \geq 1$.*

DÉMONSTRATION. We need to verify the growth and normality assumptions of Lemma 10.10 on the words that correspond to the digits added in each considered base in each step of the algorithm.

To find bounds for the number of added digits in step $i+1$ in base β_j, for $j \leq t_i$, consider the chain of intervals

$$I_{i,j} \supset \ldots \supset I_{i,t_i} \supset L \supset J_1 \supset \ldots \supset J_j$$

which is considered in step $i+1$. We find a lower bound on the Lebesgue measure of J_j in the form of

$$\lambda(J_j) \geq \frac{1}{2^{t_i}} \frac{1}{\beta_1^{M_1+1}} \frac{1}{\beta_1^{v_i n_{i+1}}} \prod_{l=1}^{t_i} \frac{1}{\beta_l^{M_{\beta_l}+4}} \cdot \lambda(I_{i,j}).$$

Thus, Lemma 10.11 implies for the number $\|u_j^{(i+1)}(\mathbf{J})\|$ of digits added in base β_j, $j \leq t_i$, in step $i+1$ of the algorithm, that

$$\frac{\log\left(\frac{1}{F_{i+1}} \frac{1}{\beta_j^{M_{\beta_j}+1}}\right)}{\log \beta_j} \leq \|u_j^{(i+1)}(\mathbf{J})\| \leq \frac{\log \frac{1}{F_{i+1}}}{\log \beta_j}$$

where $F_{i+1} = 2^{-t_i} \beta_1^{-(M_1+1)} \beta_1^{-v_i n_{i+1}} \prod_{l=1}^{t_i} \beta_l^{-M_{\beta_l}-4}$.

Hence $\|u_j^{(i+1)}(\mathbf{J})\| \sim \log 1/F_{i+1}$ with implied constants only depending on β_j. We thus need to show that

$$\log 1/F_{i+1} = t_i \log 2 + (v_i n_{i+1} + M_1 + 1) \log \beta_1 + \sum_{l=1}^{t_i} (M_l + 4) \log \beta_l$$

satisfies assumptions (3.2) of Lemma 10.10.

We now look at the growth of n_{i+1}. In light of Proposition 10.17, condition (3.7) requires

$$n_{i+1} \geq M_{\beta_j} + k_{i+1} \tag{3.9}$$

for all $1 \leq j \leq t_{i+1}$. We have $\epsilon_{i+1} = 1/t_{i+1} \to 0$ and $k_{i+1} = t_{i+1} \to \infty$ as $t_{i+1} \to \infty$. Thus also n_{i+1} tends to infinity at least logarithmically in i.

Since $\lambda \leq \frac{\beta}{\beta-1}\mu_\beta$, $\beta^k \leq |\mathcal{L}_k| \leq \frac{\beta}{\beta-1}\beta^k$, and because of Proposition 10.17, condition (3.8) on n_{i+1} is satisfied, if for all $j \leq t_i$,

$$4\left(\frac{\beta_j}{\beta_j - 1}\right)^2 \beta_j^k \beta_j^{n_{i+1}\eta(\epsilon_{i+1}, k_{i+1})} < \delta_{i+1}.$$

With η from equation (4.3), this translates into the requirement that for every $j \leq t_i$,

$$n_{i+1} \geq \frac{(M_{\beta_j}+1)\log\beta_j + \log\frac{\beta_j}{\beta_j - 1}}{\epsilon_{i+1} \min(\frac{\epsilon_{i+1}\beta_j^{k_{i+1}}}{16}, \frac{3}{4})} \left(\log\left(4\left(\frac{\beta_j}{\beta_j-1}\right)^2 \beta_j^{k_{i+1}}\right) + \log\frac{1}{\delta_{i+1}}\right), \tag{3.10}$$

where

$$\begin{aligned}\log\frac{1}{\delta_{i+1}} =& 2\log 2 + \log t_i + (t_i + t_{i+1})\log 2 + (M_{\beta_1}+4)\log\beta_1 \\ &+ 2\sum_{1\leq j\leq t_i} (M_{\beta_j}+4)\log\beta_j + \sum_{t_i < j \leq t_{i+1}} (M_{\beta_j}+4)\log\beta_j\end{aligned}$$

(where the last sum is empty if $t_i = t_{i+1}$).

Properties (3.4) and (3.6) on the sequence $(\beta_j)_{j\geq 1}$ imply

$$\max_{1\leq j\leq t_i}\left((M_{\beta_j}+1)\log\beta_j + \log\frac{\beta_j}{\beta_j-1}\right) \leq \left((M_{\beta_1}+4)\log\beta_1 + \log\frac{\beta_1}{\sqrt[3]{2}-1} + 1\right)(1+\log i). \tag{3.11}$$

Conditions (3.4) - (3.6) can be achieved by suitably repeating the bases β_j. All conditions are satisfied in step 1, and the process of repeating the bases is possible computably.

Properties (3.4) - (3.6) and (3.11), together with $t_{i+1} = k_{i+1} = 1/\epsilon_{i+1} \sim \log i$, imply that for i large enough

$$n_{i+1} \geq O\left(\frac{\log i}{1/\log i}\left(\log i + (\log i)^2 + \log\log i + \log i + \log i\right)\right) = O\left((\log i)^4\right)$$

where the implied constant only depends on β_1. Hence n_{i+1} grows at least as $O(\log i)$ and at most as $O((\log i)^4)$, where the implied constants depend only on β_1. Thus $\log 1/F_{i+1}$ and hence also $\|u_j^{(i+1)}(\mathbf{J})\|$ growths at least as $O(\log i)$ and at most as $O((\log i)^4)$, where again the implied constants only depend on β_1. Thus $\|u_j^{(i+1)}(\mathbf{J})\|$ satisfies conditions (3.2) of Proposition 10.10. Hence the number X produced by this algorithm is β_j-normal for every $j \geq 1$. □

Remark. The choices of how t_i, ϵ_i and k_i change with the step i of the algorithm and the conditions on the sequence of bases $(\beta_j)_{j\geq 1}$ are rather arbitrary. There is a lot of freedom to optimize for other quantities, such as done in Becher, Heiber, Slaman [**22**] where computational speed is optimized. This is not taken into account here. However, in a similar way the corresponding discrepancy estimates can be worked out for absolutely Pisot normal numbers. This is rather technical and leads to an upper bound of the order $\frac{1}{(\log N)^c}$ with some explicit small positive constant.

Remark. Following these lines, an extension of Becher, Heiber, Slaman's algorithm to a countable set of real bases that are β-numbers is possible, provided these bases are bounded away from 1 and such there is a uniform bound on the length of the periodic part in their orbit of 1.

A β-number is a real number β such that the orbit of 1 under T_β is finite. Pisot numbers are β-numbers. It is not known under which conditions Salem numbers are or are not β-numbers (a *Salem number* is a real algebraic integer $\beta > 1$ such that all its conjugates have absolute values at most equal to one, with equality in at least one case). Salem numbers of degree 4 are β-numbers, but there is computational and heuristic evidence that higher degree Salem numbers exist that are no β-numbers, see for example [**45**].

Note that β-numbers satisfy the specification property - one can always use a block of zeros to make the concatenation of two admissible blocks admissible. This is because admissible words can be characterized as precisely the subwords of the lexicographic largest word in the β-shift. Since the orbit of 1 is finite, this word will be eventually periodic and hence the lengths of subwords consisting of only zeros is bounded. Thus Lemma 3 in [**38**] on the number of (ϵ, k)-normal admissible words is valid and can be used as an existence criterion for a t_i sequence $\mathbf{J}$ in each step of the algorithm.

Note also that β-numbers also satisfy Proposition 2.6 of [**137**] needed to control the decay of the length of subintervals. However, we are looking for a lower bound for the measure of cylinder intervals of the form (3.3) that is uniform for all bases β under consideration. This can be achieved by requiring that there is a uniform bound on the length of the period of the orbit of 1 under T_β for each β under consideration.

When adapting the proof of Lemma 10.12 to β-numbers, we moreover need to require that the set of β-numbers under consideration is bounded away from 1, as above with the plastic number.

4. Explicit Estimates for β-expansions

In this section we make explicit the constants in Lemma 10.9 using large deviation estimates for certain dependent random variables. This requires us to provide an upper bound for the length of the largest block of zeros appearing in the modified β-expansion of 1 for a Pisot number β.

4.1. Number of zeros in the expansion of 1. Let β be a Pisot number and denote by $d_\beta(1) = 0.\epsilon_1\epsilon_2\ldots$ the β-expansion of 1, i.e. $\epsilon_1 = \lfloor\beta\rfloor$ and $\epsilon_i = \lfloor\beta T_\beta^{i-1}(1)\rfloor$ for $i \geq 1$. Let $d_\beta^*(1)$ be the modified β-expansion of 1, i.e. $d_\beta^*(1) = d_\beta(1)$ if the sequence $\epsilon_1\epsilon_2\ldots$ does not end with infinitely many zeros, and $d_\beta^*(1) = 0.(\epsilon_1\epsilon_2\ldots\epsilon_{n-1}(\epsilon_n - 1))^\omega$ when $d_\beta(1)$ ends in infinitely many zeros and ϵ_n is the last non-zero digit. It is known that $d_\beta^*(1)$ is purely periodic or eventually periodic if β is Pisot. We reprove this fact here and give an explicit upper bound for the preperiod length v and period length p and take $v + p$ as a trivial upper bound for the size of the largest block of zeros in $d_\beta^*(1)$. Note that $d_\beta^*(1)$ is (eventually) periodic if the orbit of 1 under T_β is finite, and that the number of distinct elements in this orbit is precisely $v + p$.

PROPOSITION 10.15. *Let β be a Pisot number of degree d with r real conjugates $\beta = \beta_1, \beta_2, \ldots, \beta_r$ and $2s$ complex conjugates $\beta_{r+1}, \ldots, \beta_d$. Then the orbit of 1 under the map T_β, i.e. the set*

$$\{T_\beta^k(1) \mid k \geq 0\},$$

is finite and its number of elements is bounded by

$$M = d!\det(B)^{-1}2^{r+s-1}\pi^s C^{r+2s-1} + d \tag{4.1}$$

where

$$B = \begin{pmatrix} 1 & \beta & \ldots & \beta^{d-1} \\ 1 & \beta_2 & \ldots & \beta_2^{d-1} \\ \vdots & \vdots & & \vdots \\ 1 & \beta_d & \ldots & \beta_d^{d-1} \end{pmatrix} \tag{4.2}$$

and where

$$C = 1 + \frac{\lfloor\beta\rfloor}{1-\eta}$$

with $\eta = \max_{2\leq j\leq d}|\beta_j| < 1$.

DÉMONSTRATION. For $k \geq 0$, $T_\beta^k(1)$ is an element of $\mathbb{Z}[\beta]$, hence there is a unique representation $T_\beta^k(1) = p_0^{(k)} + p_1^{(k)}\beta + \ldots + p_{d-1}^{(k)}\beta^{d-1}$ with $p_i^{(k)} \in \mathbb{Z}$. Denote by σ_j, $1 \leq j \leq d$, the j-th conjugation, ordered such that the first r are real, and $\sigma_{r+i} = \bar{\sigma}_{r+s+i}$ for $1 \leq i \leq s$. We have

$$T_\beta^k(1) = \beta^k\left(1 - \sum_{l=1}^{k}\epsilon_l\beta^{-l}\right)$$

hence for $2 \leq j \leq d$

$$|\sigma_j(T_\beta^k(1))| \leq 1 + \frac{\lfloor\beta\rfloor}{1-\eta}$$

where $\eta = \max_{2\leq j\leq d}|\beta_j| < 1$.

Note that

$$B\begin{pmatrix} p_0^{(k)} \\ p_1^{(k)} \\ \vdots \\ p_{d-1}^{(k)} \end{pmatrix} = \begin{pmatrix} T_\beta^k(1) \\ \sigma_2(T_\beta^k(1)) \\ \vdots \\ \sigma_d(T_\beta^k(1)) \end{pmatrix}$$

where B is as in (4.2) and has determinant $\det B = \prod_{1\le i<j\le d}(\beta_j - \beta_i) \neq 0$. Now, since the vector of $T_\beta^k(1)$ and its conjugates can be canonically embedded in a compact convex set in $\mathbb{R}^{r+2s}$ of volume $2^{r+s-1}\pi^s C^{r+2s-1}$, we can count the $\mathbb{Z}^d$-lattice points in a compact convex set in $\mathbb{R}^d$ of volume $\det(B)^{-1}2^{r+s-1}\pi^s C^{r+2s-1}$. By loosing a factor of 2, we can make this set additionally centrally symmetric if we allow $T_\beta^k(1)$ (formally) to take on values in the interval $[-1,1]$. Then we can use a result by Blichfeldt [**41**] and bound the number of $\mathbb{Z}^d$-lattice points in $B^{-1}Y$ by

$$|B^{-1}Y \cap \mathbb{Z}^d| \le d!\det(B)^{-1}2^{r+s-1}\pi^s C^{r+2s-1} + d$$

with $C = 1 + \frac{\lfloor\beta\rfloor}{1-\eta}$ and hence obtain an upper bound for the number of distinct points in the orbit of 1 under T_β which is also a trivial upper bound for the maximum number of consecutive zeros in the modified β-expansion of 1 as explained above. □

4.2. Number of not (ϵ, k)-normal numbers. Let β be a Pisot number and let $\mathcal{L}_n$ be the set of all admissible words of length n. Fix $\epsilon > 0$ and a positive integer k. We wish to find explicit estimates for the number of non-(ϵ, k)-normal words of length n for fixed $\epsilon > 0$ and k such as given in Lemma 10.9 (Lemma 3 in [**38**]). The method in [**38**] uses methods of ergodic theory and the authors are not aware of a method to make the implied constants explicit. Therefore we use a probabilistic approach by viewing the digits to base β as random variables and using a variant of Hoeffding's inequality for dependent random variables to bound the tail distribution of their sum. This approach automatically gives all involved constants explicitly. We use the following Lemma due to Siegel (Theorem 5 in [**226**]).

LEMMA 10.16. *Let $X = X_1 + X_2 + \ldots + X_l$ be the sum of l possibly dependent random variables. Suppose that X_i, for $i = 1, 2, \ldots, l$, is the sum of n_i mutually independent random variables having values in the interval $[0, 1]$. Let $\mathbb{E}[X_i] = n_i p_i$. Then for $a \ge 0$*

$$\mathbb{P}(X - \mathbb{E}[X] \ge a) < \exp\left(-\frac{a^2}{8(\sum_i \sqrt{p_i(1-p_i)n_i})^2}\right) + \exp\left(-\frac{3a}{4\sum_i(1-p_i)^2}\right).$$

PROPOSITION 10.17. *Let β be a Pisot number. The μ_β-measure of the set of not (ϵ, k)-normal words of length n satisfies*

$$\mu_\beta(E_n^c(\epsilon, k)) \le 4|\mathcal{L}_k||\mathcal{L}_n|^{-\eta}$$

for $n \ge M + k$ with $\eta > 0$ as in equation (4.3) and M as in equation (4.1).

DÉMONSTRATION. Let $d \in \mathcal{L}_k$ and for $n \ge M+k$, let $X_1, \ldots, X_{M+1} : \mathcal{L}_n \to \mathbb{R}$ be random variables where $X_i(\omega)$ denotes the number of occurrences of the word d in $\omega = \omega_1 \ldots \omega_n$ at positions

$$\omega_{i+j(M+1)}\omega_{i+j(M+1)+1} \cdots \omega_{i+j(M+1)+k-1}$$

for $0 \le j \le \lfloor\frac{n-k+1-i}{M+1}\rfloor$. The X_i are dependent, but each is a sum of $n_i = \lfloor\frac{n-k+1-i}{M+1}\rfloor + 1$ independent identically distributed random variables $Y_j^{(i)}$ that take value one if and only if the

word d appears in ω starting at digit $\omega_{i+j(M+1)}$ and zero otherwise. Let $X = X_1 + \ldots + X_{M+1}$. We have $\mathbb{E}[X] = n\mu_\beta(c(d))$ and $\mathbb{E}[X_i] = n_i\mu_\beta(c(d))$. Denote by $\bar{E}_n(\epsilon, k)$ the set of words of length n for which there is a subword d of length k that appears more often than $n(\mu_\beta(c(d))+\epsilon)$ times and let $\bar{E}_n(\epsilon, d)$ be the set of words of length n for which the subword d appears more often than $n(\mu_\beta(c(d)) + \epsilon)$ times. We apply Lemma 10.16 with $l = M + 1$, n_i as above, $p_i = \mu_\beta(c(d))$ and $a = n\epsilon$ and obtain

$$\begin{aligned}\mu_\beta(\{\omega \in \mathcal{L}_n \mid X > n(\mu_\beta(c(d)) + \epsilon)\}) &= \mu_\beta(\bar{E}_n(\epsilon, d)) \\ &< \exp\left(-\frac{(n\epsilon)^2}{8\mu_\beta(c(d))(1-\mu_\beta(c(d)))(M+1)^2(\lfloor\frac{n-k}{M+1}\rfloor + 1)}\right) \\ &\quad + \exp\left(-\frac{3n\epsilon}{4(M+1)(1-\mu_\beta(c(d)))^2}\right).\end{aligned}$$

Note that by slight abuse of notation we write $\mu_\beta(\omega)$ and mean $\mu_\beta(c(\omega))$ for a finite word ω. Using $(1-\beta^{-1})\beta^{-(M+1)}\beta^{-k} \le \mu_\beta(c(d)) \le \beta^{-k}$ and $n \ge M+1$, this is

$$< \exp\left(-\frac{\epsilon^2 n}{16(M+1)\beta^{-k}}\right) + \exp\left(-\frac{3\epsilon n}{4(M+1)}\right) < 2\exp\left(-\frac{\epsilon n}{M+1}\min(\frac{\epsilon}{16\beta^{-k}}, \frac{3}{4})\right).$$

Finally, since $\mu_\beta(\bar{E}_n(\epsilon, k)) \le \sum_{d\in\mathcal{L}_k} \mu_\beta(\bar{E}_n(\epsilon, d))$ and using that $\beta^n \le |\mathcal{L}_n| \le \frac{\beta}{\beta-1}\beta^n$ we obtain

$$\begin{aligned}\mu_\beta(\bar{E}_n(\epsilon, k)) &\le |\mathcal{L}_k| 2\exp\left(-\frac{\epsilon n}{M+1}\min(\frac{\epsilon}{16\beta^{-k}}, \frac{3}{4})\right) \\ &\le 2|\mathcal{L}_k||\mathcal{L}_n|^{-\eta}\end{aligned}$$

with

$$\eta = \frac{\epsilon \min(\frac{\epsilon}{16\beta^{-k}}, \frac{3}{4})}{\log(\frac{\beta}{\beta-1}) + (M+1)\log\beta} > 0. \tag{4.3}$$

Using the same argument with $Y = n - X$ gives a symmetrical upper bound for the number of words ω of length n in which the word d appears less than $n\mu_\beta(c(d)) - \epsilon n$ times. Thus we obtain an upper bound for the number of not (ϵ, k)-normal words of length n of the form

$$4|\mathcal{L}_k||\mathcal{L}_n|^{-\eta}$$

for $n \ge M + k$ with η as in (4.3). □

Corollary 10.18. *Let β be a Pisot number. The number of not (ϵ, k)-normal words of length n satisfies*

$$|E_n^c(\epsilon, k)| \le C|\mathcal{L}_n|^{1-\eta}$$

for $n \ge M + k$ with $\eta > 0$ as in equation (4.3), M as in equation (4.1), and where $C = 4|\mathcal{L}_k|\beta^{M+1}\frac{\beta}{\beta-1}$.

Démonstration. Since the Parry measure μ_β satisfies

$$\left(1 - \frac{1}{\beta}\right)\lambda \le \mu_\beta \le \frac{\beta}{\beta-1}\lambda$$

with respect to the Lebesgue measure λ, and due to the bounds on the Lebesgue measure of β-adic cylinder intervals from Lemma 10.11, the bound from Proposition 10.17 on the μ_β measure of the set of non-(ϵ, k)-normal words of length n implies for the number of such words

$$|E_n^c(\epsilon, k)| \leq C|\mathcal{L}_n|^{1-\eta}, \tag{4.4}$$

where $C = C(\beta, k) = 4|\mathcal{L}_k|\beta^M \frac{\beta}{\beta-1}$ and $\eta = \eta(\beta, \epsilon, k)$ as given in equation (4.3) and where we used that $\beta^n \leq |\mathcal{L}_n| \leq \frac{\beta}{\beta-1}\beta^n$. □

Acknowledgements. We are grateful to an anonymous referee for several suggestions improving the paper. For the realization of the present paper the first author received support from the Conseil Régional de Lorraine. Parts of this research work were done when the first author was visiting the Department of Mathematics of Graz University of Technology. The author thanks the institution for their hospitality. The second author was supported by the Austrian Science Fund (FWF) : I 1751-N26 ; W1230, Doctoral Program "Discrete Mathematics" ; and SFB F 5510-N26. He would like to thank Karma Dajani and Bing Li for some interesting discussions on β-expansions. All authors thank the Erwin-Schrödinger-Institute (ESI) for its hospitality during the workshop *Normal Numbers : Arithmetic, Computational and Probabilistic Aspects*, November 14–18, 2016.

Chapitre 11

Non-normal numbers in dynamical systems fulfilling the specification property

This chapter is joint work with Izabela Petrykiewicz and appeared in *Discrete and Continuous Dynamical Systems - Series A*, Vol. 34 (2014), 4751 – 4764.

1. Introduction

Let $N \geq 2$ be an integer, called the base, and $\Sigma := \{0, 1, \ldots, N-1\}$, called the set of digits. Then for every $x \in [0, 1)$ we denote by

$$x = \sum_{h=1}^{\infty} d_h(x) N^{-h},$$

where $d_h(x) \in \Sigma$ for all $h \geq 1$, the unique non-terminating N-ary expansion of x. For every positive integer n and a block of digits $\mathbf{b} = b_1 \ldots b_k \in \Sigma^k$ we write

$$\Pi(x, \mathbf{b}, n) := \frac{|\{0 \leq i < n : d_{i+1}(x) = b_1, \ldots, d_{i+k}(x) = b_k\}|}{n}$$

for the frequency of the block $\mathbf{b}$ among the first n digits of the N-ary expansion of x. Furthermore, let

$$\Pi_k(x, n) := (\Pi(x, \mathbf{b}, n))_{\mathbf{b} \in \Sigma^k}$$

be the vector of frequencies of all blocks $\mathbf{b}$ of length k.

Now we call a number k-normal if for every block $\mathbf{b} \in \Sigma^k$ of digits of length k, the limit of the frequency $\Pi(x, \mathbf{b}, n)$ exists and equals N^{-k}. A number is called normal with respect to base N if it is k-normal for all $k \geq 1$. Furthermore, a number is called absolutely normal if it is normal to any base $N \geq 2$.

On the one hand, it is a classical result due to Borel [**43**] that Lebesgue almost all numbers are absolutely normal. So the set of normal numbers is large from a measure theoretical viewpoint.

On the other hand, it suffices for a number to be not normal if the limit of the frequency vector is not the uniform one. First results concerning the Hausdorff dimension or the Baire category of non-normal numbers were obtained by Šalát [**204**] and Volkmann [**249**]. Stronger variants of non-normal numbers were of recent interest. In particular, Albeverio *et al.* [**9**, **10**] considered the fractal structure of essentially non-normal numbers and their variants. The theory of multifractal divergence points lead to the investigation of extremely non-normal numbers by Olsen [**172**, **173**] and Olsen and Winter [**175**]. The important result for our considerations is that both essentially and extremely non-normal numbers are large from a topological point of view.

2. Definitions and statement of result

We start with the definition of a dynamical system. Let M be a compact metric space and $\phi : M \to M$ be a continuous map. Then we call the pair (M, ϕ) a (topological) dynamical system.

The second ingredient is the definition of a topological partition. Let M be a metric space and let $\mathcal{P} = \{P_0, \dots, P_{N-1}\}$ be a finite collection of disjoint open sets. Then we call $\mathcal{P}$ a topological partition (of M) if M is the union of the closures $\overline{P_i}$ for $i = 0, \dots, N-1$, *i.e.*

$$M = \overline{P_0} \cup \dots \cup \overline{P_{N-1}}.$$

Suppose now that a dynamical system (M, ϕ) and a topological partition $\mathcal{P} = \{P_0, \dots, P_{N-1}\}$ of M are given. We want to consider the symbolic dynamical system behind. Therefore, let $\Sigma = \{0, \dots, N-1\}$ be the alphabet corresponding to the topological partition $\mathcal{P}$. Furthermore, define

$$\Sigma^k = \{0, \dots, N-1\}^k, \quad \Sigma^* = \bigcup_{k \geq 1} \Sigma^k \cup \{\epsilon\}, \quad \text{and} \quad \Sigma^{\mathbb{N}} = \{0, \dots, N-1\}^{\mathbb{N}}$$

to be the set of words of length k, the set of finite and the set of infinite words over Σ, respectively, where ϵ is the empty word. For an infinite word $\omega = a_1 a_2 a_3 \ldots \in \Sigma^{\mathbb{N}}$ and a positive integer n, let $\omega|n = a_1 a_2 \dots a_n$ denote the truncation of ω to the n-th place. Finally, for $\omega \in \Sigma^*$ we denote by $[\omega]$ the cylinder set of all infinite words starting with the same letters as ω, *i.e.*

$$[\omega] := \{\gamma \in \Sigma^{\mathbb{N}} : \gamma | \, |\omega| = \omega\}.$$

Now we want to describe the shift space that is generated by our topological partition. Therefore, we call $\omega = a_1 a_2 \dots a_n \in \Sigma^n$ allowed for $(\mathcal{P}, \phi)$ if

$$\bigcap_{k=1}^{n} \phi^{-k} (P_{a_k}) \neq \emptyset.$$

Let $\mathcal{L}_{\mathcal{P},\phi}$ be the set of allowed words. Then $\mathcal{L}_{\mathcal{P},\phi}$ is a language and there is a unique shift space $X_{\mathcal{P},\phi} \subseteq \Sigma^{\mathbb{N}}$, whose language is $\mathcal{L}_{\mathcal{P},\phi}$. We call $X_{\mathcal{P},\phi} \subseteq \Sigma^{\mathbb{N}}$ the one-sided symbolic dynamical system corresponding to $(\mathcal{P}, \phi)$.

Furthermore, we split the language up corresponding to the length of the words. For $k \geq 1$ we denote by

$$\mathcal{L}_k = \{\omega \in \mathcal{L}_{\mathcal{P},\phi} \colon \ |\omega| = k\}.$$

Then we have that $\mathcal{L}_{\mathcal{P},\phi} = \bigcup_{k=1}^{\infty} \mathcal{L}_k$.

Finally, for each $\omega = a_1 a_2 a_3 \ldots \in X_{\mathcal{P},\phi}$ and $n \geq 0$ we denote by $D_n(\omega)$ the cylinder set of order n corresponding to ω in M, *i.e.*

$$D_n(\omega) := \bigcap_{k=0}^{n} \phi^{-k}(P_{a_k}) \subseteq M.$$

After providing all the ingredients necessary for the statement of our result we want to link this concept with the N-ary representations of Section 1.

Example 11.1. *Let $M = \mathbb{R}/\mathbb{Z}$ be the circle and $\phi : M \to M$ be defined by $\phi(x) = Nx \pmod 1$. We divide M into N subintervals $P_0, \dots, P_{N-1}$ of the form $P_i = (i/N, (i+1)/N)$ and let $\Sigma = \{0, \dots, N-1\}$. Then the underlying system is the N-ary representation. Furthermore, it is easy to verify that the language $\mathcal{L}_{\mathcal{P},\phi}(x)$ is the set of all words over Σ, so that the one-sided symbolic dynamical system $X_{\mathcal{P},\phi}$ is the full one-sided N-shift $\Sigma^{\mathbb{N}}$.*

Our second example will be the main motivation for this paper. In particular, we will consider β-expansions, where $\beta > 1$ is not necessarily an integer. These systems are of special interest, since the underlying symbolic dynamical system is not the full-shift. The first authors investigating these number systems were Parry [**177**] and Renyi [**194**]. For a more modern account on these number systems we refer the interested reader to the book of Dajani and Kraaikamp [**62**].

EXAMPLE 11.2. *Let $\beta > 1$ be a real number and $\phi\colon [0,1) \to [0,1)$ be the transformation given by*

$$\phi(x) = \beta x \mod 1.$$

The sets

$$P_i := \left(\frac{i}{\beta}, \frac{i+1}{\beta}\right) \quad (i = 0, \ldots, \lfloor \beta \rfloor - 1)$$

and

$$P_{\lfloor \beta \rfloor + 1} := \left(\frac{\lfloor \beta \rfloor}{\beta}, 1\right)$$

*together with ϕ form a number system partition of M. The corresponding language is called the β-shift (cf. [**62, 177, 194**]).*

Before extending the notions of normal and non-normal numbers we want to investigate the properties of the β-shift in more detail. We say that a language $\mathcal{L}$ fulfills the specification property if there exists a positive integer $j \geq 0$ such that we can concatenate any two words $\mathbf{a}$ and $\mathbf{b}$ by padding a word of length less than j in between, *i.e.* if, for every pair $\mathbf{a}, \mathbf{b} \in \mathcal{L}$, there exists a word $\mathbf{u} \in \mathcal{L}$ with $|\mathbf{u}| \leq j$ such that $\mathbf{aub} \in \mathcal{L}$. Furthermore, we call the language connected of order j if this padding word can always be chosen of length j. Note that the β-shift fulfills this property.

Suppose for the rest of the paper that (M, ϕ) is a number system partition, together with a dynamical system $X_{\mathcal{P},\phi}$ that fulfills the specification property with a parameter j. Since the partition $\mathcal{P}$ and the transformation ϕ are fixed, we may write $X = X_{\mathcal{P},\phi}$ and $\mathcal{L} = \mathcal{L}_{\mathcal{P},\phi}$ for short.

In order to extend the definition of normal and thus non-normal numbers to M we need that the expansion is unique. Therefore, we suppose that $\bigcap_{n=0}^{\infty} \overline{D_n(\omega)}$ consists of exactly one point. This motivates the definition of the map $\pi_{\mathcal{P},\phi} : X \to M$ by

$$\{\pi_{\mathcal{P},\phi}(\omega)\} = \bigcap_{n=0}^{\infty} \overline{D_n(\omega)}.$$

However, the converse need not be true. In particular, we consider Example 11.2 with $\beta = \frac{1+\sqrt{5}}{2}$ (the golden mean). Then clearly $\beta^2 - \beta - 1 = 0$. Now on the one hand, every word in X is mapped to a unique real number. However, if we consider expansion of $\frac{1}{\beta}$, which lies between the two intervals P_0 and P_1, then, since

$$\frac{1}{\beta} = \frac{1}{1 - \frac{1}{\beta^2}},$$

we get that $010101\ldots$ and $100000\ldots$ are possible expansions of $\frac{1}{\beta}$. Similarly we get that $101010\ldots$ and $010000\ldots$ are possible expansion of $\frac{1}{\beta^2}$.

However, one observes that these ambiguities originate from the intersections of two partitions $\overline{P_i} \cap \overline{P_j}$ for $i \neq j$. Thus we concentrate on the inner points, which somehow correspond to the irrational numbers in the above case of the decimal expansion. Let

$$U = \bigcup_{i=0}^{N-1} P_i,$$

which is an open and dense ($\overline{U} = M$) set. Then for each $n \geq 1$ the set

$$U_n = \bigcap_{k=0}^{n-1} \phi^{-k}(U),$$

is open and dense in M. Thus by the Baire Category Theorem, the set

$$U_\infty = \bigcap_{n=0}^{\infty} U_n \tag{2.1}$$

is dense. Since $M \setminus U_\infty$ is the countable union of nowhere dense sets it suffices to show that a set is residual in U_∞ in order to show that it is in fact residual in M. Furthermore, for $x \in U_\infty$ we may call ω **the** symbolic expansion of x if $\pi_{\mathcal{P},\phi}(\omega) = x$. Thus in the following we will silently suppose that $x \in U_\infty$.

After defining the environment we want to pull over the definitions of normal and non-normal numbers to the symbolic dynamical system. To this end let $\mathbf{b} \in \Sigma^k$ be a block of letters of length k and $\omega = a_1 a_2 a_3 \ldots \in X$ be the symbolic representation of an element. Then we write

$$\mathrm{P}(\omega, \mathbf{b}, n) = \frac{|\{0 \leq i < n : a_{i+1} = b_1, \ldots, a_{i+k} = b_k\}|}{n}$$

for the frequency of the block $\mathbf{b}$ among the first n letters of ω. In the same manner as above let

$$\mathrm{P}_k(\omega, n) = (\mathrm{P}(\omega, \mathbf{b}, n))_{\mathbf{b} \in \Sigma^k}$$

be the vector of all frequencies of blocks $\mathbf{b}$ of length k among the first n letters of ω.

Let μ be a given ϕ-invariant probability measure on X and $\omega \in X$. Then we call the measure μ associated to ω if there exists a infinite sub-sequence F of $\mathbb{N}$ such that for any block $\mathbf{b} \in \Sigma^k$

$$\lim_{\substack{n \to \infty \\ n \in F}} P(\omega, \mathbf{b}, n) = \mu([\mathbf{b}]).$$

Furthermore, we call ω a generic point for μ if we can take $F = \mathbb{N}$: then μ is the only measure associated with ω. If μ is the maximal measure, then we call ω normal. Finally, for a ϕ-invariant probability measure on X we define its entropy by

$$H(\mu) = \lim_{N \to \infty} -\frac{1}{N} \sum_{a_1, \ldots, a_n \in A^n} \mu([a_1, \ldots, a_n]) \log(\mu([a_1, \ldots, a_n])).$$

The existence of such an invariant measure for the β-shift was independently proven by Gelfond [**90**] and Parry [**177**]. Bertrand-Mathis [**35**] constructed such an invariant measure by generalizing the construction of Champernowne for any dynamical system fulfilling the specification property. She also showed that this measure is ergodic, strongly mixing, its entropy is $\log \beta$ and it is generic for the maximal measure. An application of Birkhoff's ergodic theorem yields that almost all numbers $\omega \in X$ are normal (*cf.* Chapter 3.1.2 of [**62**]).

Normal sequences for β-shifts where constructed by Ito and Shiokawa [**113**], however, these expansions provide no admissible numbers. Furthermore, Bertrand-Mathis and Volkmann [**38**] constructed normal numbers on connected dynamical systems.

We note that we equivalently could have defined the measure-theoretic dynamical system with respect to M instead of X. However, since the definition of essentially and extremely non-normal numbers does not depend on this, we will not consider this in the following.

As already mentioned above, the aim of the present paper is to show that the non-normal numbers are a large set in the topological sense. Sigmund [**229**] showed, that for any dynamical system fulfilling the specification property, the set of non-normal numbers is residual. However, in the present paper we want to show that even smaller sets, namely the essentially and extremely non-normal numbers, are also residual.

We start by defining the simplex of all probability vectors Δ_k by

$$\Delta_k = \left\{ (p_{\mathbf{i}})_{\mathbf{i}\in\mathcal{L}_k} : p_{\mathbf{i}} \geq 0, \sum_{\mathbf{i}\in\mathcal{L}_k} p_{\mathbf{i}} = 1 \right\}.$$

Let $\|\cdot\|_1$ denote the 1-norm then $(\Delta_k, \|\cdot\|_1)$ is a metric space. On the one hand, we clearly have that any vector $\mathrm{P}_k(\omega, n)$ of frequencies of blocks of digits of length k belongs to Δ_k. On the other hand, if we assume for example that the word 11 is forbidden in the expansion. Then the maximum frequency for the single letter 0 is 1 and for 1 is $\frac{1}{2}$. Therefore, the probability vector $(0, 1)$ cannot be reached.

Let $\mathrm{A}_k(\omega)$ be the set of accumulation points of the sequence $(\mathrm{P}_k(\omega, n))_n$ with respect to $\|\cdot\|_1$, *i.e.* for $\omega \in X$ we set

$$\mathrm{A}_k(\omega) := \{\mathbf{p} \in \Delta_k : \mathbf{p} \text{ is an accumulation point of } (\mathrm{P}_k(\omega, n))_n\}.$$

Then we define S_k as union of all possible accumulation points, *i.e.*

$$\mathrm{S}_k := \bigcup_{\omega\in X} \mathrm{A}_k(\omega).$$

We note that in the case of N-ary expansions this definition leads to the shift invariant probability vectors (*cf.* Theorem 0 of Olsen [**174**]).

We call a number $\omega \in X$ essentially non-normal if for all $i \in \Sigma$ the limit

$$\lim_{n\to\infty} \mathrm{P}(\omega, i, n)$$

does not exist. For the case of N-ary expansions Albeverio *et al.* [**9, 10**] could prove the following

THEOREM ([**9, 10**, Theorem 1]). *Let $(\mathcal{P}, \phi)$ be the N-ary representation of Example 11.1. Then the set of essentially non-normal numbers is residual.*

This result has been generalized to Markov partitions whose underlying language is the full shift by the first author [**144**]. Our first results is the following generalization.

THEOREM 11.3. *Let $\mathcal{P} = \{P_0, \ldots, P_{N-1}\}$ be a number system partition for (M, ϕ). Suppose that*

- *$\bigcap_{n=0}^{\infty} \overline{D_n(\omega)}$ consists of exactly one point;*
- *$X_{\mathcal{P},\phi}$ fulfills the specification property;*
- *for all $i \in \Sigma$ there exist $\mathbf{q}_{i,1} = (q_{1,1}, \ldots q_{1,N-1}), \mathbf{q}_{i,2} = (q_{2,1}, \ldots q_{2,N-1}) \in S_1$ such that $|q_{1,i} - q_{2,i}| > 0$.*

Then the set of essentially non-normal numbers is residual.

REMARK 16. The requirement that for each digit we need at least two possible distributions is sufficient in order to prevent that the underlying language is too simple. For example, we want to exclude the case of the shift over the alphabet $\{0,1\}$ with forbidden words 00 and 11.

A different concept of non-normal numbers are those being arbitrarily close to any given configuration. In particular, we want to generalize the idea of extremely non-normal numbers and their Cesàro variants to the setting of number system partitions.

For any infinite word $\omega \in X$ we clearly have $\mathrm{A}_k(\omega) \subset S_k$. On the other hand, we call $\omega \in X$ extremely non-k-normal if the set of accumulation points of the sequence $(\mathrm{P}_k(\omega,n))_n$ (with respect to $\|\cdot\|_1$) equals S_k, *i.e.* $\mathrm{A}_k(\omega) = \mathrm{S}_k$. Furthermore, we call a number extremly non-normal if it is extremely non-k-normal for all $k \geq 1$.

The set of extremely non-normal numbers for the N-ary represenation has been considered by Olsen [**174**].

THEOREM ([**174**, Theorem 1]). *Let $(\mathcal{P},\phi)$ be the N-ary expansion of Example 11.1. Then the set of extremely non-normal numbers is residual in M.*

This result was generalized to iterated function systems by Baek and Olsen [**17**] and to finite Markov partitions by the first author [**144**]. Furthermore, number systems with infinite set of digits like the continued fraction expansion or Lüroth expansion were considered by Olsen [**170**], Šalát [**203**], respectively. Finally, Šalát [**205**] considered the Hausdorff dimension of sets with digital restrictions with respect to the Cantor series expansion.

We want to extend this notion to the Cesàro averages of the frequencies. To this end for a fixed block $b_1 \dots b_k \in \Sigma^k$ let

$$\mathrm{P}^{(0)}(\omega,\mathbf{b},n) = \mathrm{P}(\omega,\mathbf{b},n).$$

For $r \geq 1$ we recursively define

$$\mathrm{P}^{(r)}(\omega,\mathbf{b},n) = \frac{\sum_{j=1}^{n} \mathrm{P}^{(r-1)}(\omega,\mathbf{b},j)}{n}$$

to be the rth iterated Cesàro average of the frequency of the block of digits $\mathbf{b}$ under the first n digits. Furthermore, we define by

$$\mathrm{P}_k^{(r)}(\omega,n) := \left(\mathrm{P}^{(r)}(\omega,\mathbf{b},n)\right)_{\mathbf{b}\in\Sigma^k}$$

the vector of rth iterated Cesàro averages. As above, we are interested in the accumulation points. Thus similar to above let $\mathrm{A}_k^{(r)}(\omega)$ denote the set of accumulation points of the sequence $(\mathrm{P}_k^{(r)}(\omega,n))_n$ with respect to $\|\cdot\|_1$,*i.e.*

$$\mathrm{A}_k^{(r)}(\omega) := \left\{\mathbf{p} \in \Delta_k : \mathbf{p} \text{ is an accumulation point of } (\mathrm{P}_k^{(r)}(\omega,n))_n\right\}.$$

Now we call a number rth iterated Cesàro extremely non-k-normal if the set of accumulation points is the full set, *i.e.* $\mathrm{A}_k^{(r)} = \mathrm{S}_k$.

For $r \geq 1$ and $k \geq 1$ we denote by $\mathbb{E}_k^{(r)}$ the set of rth iterated Cesàro extremely non-k-normal numbers of M. Furthermore, for $r \geq 1$ we denote by $\mathbb{E}^{(r)}$ the set of rth iterated Cesàro

extremely non-normal numbers and by $\mathbb{E}$ the set of completely Cesàro extremely non-normal numbers, *i.e.*

$$\mathbb{E} = \bigcap_k \mathbb{E}_k^{(r)} \quad \text{and} \quad \mathbb{E} = \bigcap_r \mathbb{E}^{(r)} = \bigcap_{r,k} \mathbb{E}_k^{(r)}.$$

As above, this has already been considered for the case of the N-ary expansion by Hyde *et al.* [**110**].

THEOREM ([**110**, Theorem 1.1]). *Let $(\mathcal{P}, \phi)$ be the N-ary representation of Example 11.1. Then for all $r \geq 1$ the set $\mathbb{E}_1^{(r)}$ is residual.*

In the context of extremely non-normal numbers our result is the following.

THEOREM 11.4. *Let k, r and N be positive integers. Furthermore, let $\mathcal{P} = \{P_0, \ldots, P_{N-1}\}$ be a number system partition for (M, ϕ). Suppose that $\mathcal{L}_{\mathcal{P},\phi}$ fulfills the specification property. Then the set $\mathbb{E}_k^{(r)}$ is residual.*

Since the set of non-normal numbers is a countable intersection of sets $\mathbb{E}_k^{(r)}$ we get the following

COROLLARY 11.5. *Let N be a positive integer and $\mathcal{P} = \{P_0, \ldots, P_{N-1}\}$ be a number system partition for (M, ϕ). Suppose that $\mathcal{L}_{\mathcal{P},\phi}$ fulfills the specification property. Then the sets $\mathbb{E}^{(r)}$ and $\mathbb{E}$ are residual.*

3. Proof of Theorem 11.3

Before we start proving Theorem11.3, we will construct sets Z_n which we will use in order to "measure" the distance between the proportion of occurrences of blocks and $\mathbf{q}$. Let $k \in \mathbb{N}$ and $\mathbf{q} \in S_k$ be fixed. For $n \geq 1$ let

$$Z_n = Z_n(\mathbf{q}, k) = \left\{ \omega \in \bigcup_{\ell \geq kn|\mathcal{L}_k|} \mathcal{L}_\ell \,|\, \|\mathrm{P}_k(\omega) - \mathbf{q}\|_1 \leq \frac{1}{n} \right\}.$$

The main idea consists now in the construction of a word having the desired frequencies. In particular, for a given word ω we want to show that we can add a word from Z_n to get a word whose frequency vector is sufficiently near to $\mathbf{q}$. Therefore, we first need that Z_n is not empty.

THEOREM 11.6. *Let $\mathbf{q} \in \mathrm{S}_k$. Then*

$$\dim\{\omega \in X \colon \lim_{n\to\infty} \mathrm{P}_k(\omega, n) = \mathbf{q}\} = \inf_{\mathbf{q}\in \mathrm{S}_k} H(\mathbf{q}).$$

DÉMONSTRATION. This is essentially Theorem 6 of [**171**] (see also Theorem 7.1 of [**176**]).

In our considerations we have two main differences. On the one hand, we consider dynamical system fulfilling the specification property, whereas Olsen [**171**] investigates subshifts of finite type modelled by a directed and strongly connected multigraph. However, in his proof he never uses the finitude of the set of exceptions. This means that they stay true if we replace the subshift of finite type by one fulfilling the specification property.

Another difference is that we have a number system partition, whereas Olsen [**171**] analyses a graph directed self-conformal iterated function system satisfying the Strong Open Set Condition. In the iterated function system, we have first the functions and the partition and in our case we have first a partition and then the restricted function. Therefore, this

changes only the point of view. Furthermore, the Strong Open Set Condition is satisfied by the topological partition, which we use.

After adapting these differences the proof runs along the same lines. □

LEMMA 11.7. *For all $n \geq 1$, $\mathbf{q} \in \mathrm{S}_k$ and $k \in \mathbb{N}$ we have $Z_n(\mathbf{q}, k) \neq \emptyset$.*

DÉMONSTRATION. If follows from Theorem 11.6 that

$$\dim\{\omega \in X\colon \lim_{n\to\infty} \mathrm{P}_k(\omega, n) = \mathbf{q}\} = \inf_{\mathbf{q}\in \mathrm{S}_k} H(\mathbf{q}) > 0,$$

which implies that $\{\omega \in X\colon \lim_{n\to\infty} \mathrm{P}_k(\omega, n) = \mathbf{q}\} \neq \emptyset$. Thus we chose a $\omega \in X$ such that $\lim_{n\to\infty} \mathrm{P}_k(\omega, n) = \mathbf{q}$. Then for sufficiently large ℓ the truncated word $\omega \mid \ell$ lies in Z_n. □

Since we may not put any two words together, we use the specification property do define a modified concatenation. For any pair of finite words $\mathbf{a}$ and $\mathbf{b}$ we fix a $\mathbf{u}_{\mathbf{a},\mathbf{b}}$ with $|\mathbf{u}_{\mathbf{a},\mathbf{b}}| \leq j$ such that $\mathbf{a}\mathbf{u}_{\mathbf{a},\mathbf{b}}\mathbf{b} \in \mathcal{L}$. Then for $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathcal{L}$ and $n \in \mathbb{N}$ we write

$$\mathbf{a}_1 \odot \mathbf{a}_2 \odot \cdots \odot \mathbf{a}_m := \mathbf{a}_1\mathbf{u}_{\mathbf{a}_1,\mathbf{a}_2}\mathbf{a}_2\mathbf{u}_{\mathbf{a}_2,\mathbf{a}_3}\mathbf{a}_3 \cdots \mathbf{a}_{m-1}\mathbf{u}_{\mathbf{a}_{m-1},\mathbf{a}_m}\mathbf{a}_m$$

and

$$\mathbf{a}^{\odot n} := \underbrace{\mathbf{a} \odot \mathbf{a} \odot \cdots \odot \mathbf{a}}_{n \text{ times}}.$$

Then we have the following result.

LEMMA 11.8. *Let $k \in \mathbb{N}$, and let $\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_m \in S_k$. For every $\varepsilon > 0$, and every $\omega_0 \in \mathcal{L}_{\mathcal{P},\phi}$, there exists $\omega \in \mathcal{L}_{\mathcal{P},\phi}$ whose prefix is ω_0, and $n_1 < n_2 < ... < n_m$ such that for all $1 \leq i \leq m$ we have*

$$\|\mathrm{P}_k(\omega, n_i) - \mathbf{q}_i\|_1 \leq \varepsilon. \tag{3.1}$$

DÉMONSTRATION. Let $\varepsilon > 0$ and $\omega_0 \in \mathcal{L}_{\mathcal{P},\phi}$ be given. We will define $\omega = \omega_o \odot \omega_1 \odot \omega_2 \odot ... \odot \omega_m$ recursively. For $i \geq 1$, let

$$l_i \geq \frac{2(|\omega_0 \odot \omega_1 \odot ... \odot \omega_{i-1}| + k + j - 1)|\mathcal{L}_k|}{\varepsilon}.$$

Choose any $\omega_i \in Z_{l_i}(\mathbf{q}_i, k)$. We now show that (3.1) is satisfied with $\omega = \omega_0 \odot \omega_1 \odot ... \odot \omega_i$ and $n_i = |\omega_0 \odot \omega_1 \odot ... \odot \omega_i|$.
Let $\mathbf{q}_i = (q_{\mathbf{b_1}}, ..., q_{\mathbf{b}_{|\mathcal{L}_\mathbf{k}|}})$ We have

$$\begin{aligned}
\|\mathrm{P}_k(\omega_0 \odot ... \odot \omega_i, n_i) - \mathbf{q}_i\|_1 &= \sum_{\mathbf{b}\in\mathcal{L}_k} |\mathrm{P}(\omega_0 \odot ... \odot \omega_i, \mathbf{b}) - q_\mathbf{b}| \\
&\leq \sum_{\mathbf{b}\in\mathcal{L}_k} |\mathrm{P}(\omega_0 \odot ... \odot \omega_i, \mathbf{b}) - \mathrm{P}(\omega_i, \mathbf{b})| + \sum_{\mathbf{b}\in\mathcal{L}_k} |\mathrm{P}(\omega_i, \mathbf{b}) - q_\mathbf{b}| \\
&\leq \sum_{\mathbf{b}\in\mathcal{L}_k} |\mathrm{P}(\omega_0 \odot ... \odot \omega_i, \mathbf{b}) - \mathrm{P}(\omega_i, \mathbf{b})| + \frac{|\mathcal{L}_k|}{l_i} \\
&\leq \sum_{\mathbf{b}\in\mathcal{L}_k} |\mathrm{P}(\omega_0 \odot ... \odot \omega_i, \mathbf{b}) - \mathrm{P}(\omega_i, \mathbf{b})| + \frac{\varepsilon}{2},
\end{aligned}$$

by Lemma 11.7 and our choice of l_i. Now consider $|\mathrm{P}(\omega_0 \odot ... \odot \omega_i, \mathbf{b}) - \mathrm{P}(\omega_i, \mathbf{b})|$. The block $\mathbf{b}$ can occur in $\omega_0 \odot \omega_1 \odot ... \odot \omega_{i-1}$, ω_i, and between these two words, hence we have

$$\begin{aligned}
&|\mathrm{P}(\omega_0 \odot ... \odot \omega_i, \mathbf{b}) - \mathrm{P}(\omega_i, \mathbf{b})| \\
&\qquad = \max\left(\mathrm{P}(\omega_0 \odot ... \odot \omega_i, \mathbf{b}) - \mathrm{P}(\omega_i, \mathrm{P}(\omega_i, \mathbf{b}) - \mathbf{b}), \mathrm{P}(\omega_0 \odot ... \odot \omega_i, \mathbf{b})\right) \\
&\qquad \leq \max\left(\frac{\mathrm{P}(\omega_0 \odot ... \odot \omega_{i-1}, \mathbf{b}) + \mathrm{P}(\omega_i, \mathbf{b}) + (k+j-1)}{|\omega_0 \odot ... \odot \omega_i|} - \frac{\mathrm{P}(\omega_i, \mathbf{b})}{|\omega_i|},\right. \\
&\qquad\quad \left.\frac{\mathrm{P}(\omega_i, \mathbf{b})}{|\omega_i|} - \frac{\mathrm{P}(\omega_0 \odot ... \odot \omega_{i-1}, \mathbf{b}) + \mathrm{P}(\omega_i, \mathbf{b})}{|\omega_0 \odot ... \odot \omega_i|}\right) \\
&\qquad \leq \max\left(\frac{|\omega_0 \odot ... \odot \omega_{i-1}| + (k+j-1)}{|\omega_i|}, \frac{\omega_0 \odot ... \odot \omega_{i-1}| + j}{|\omega_i|}\right) \\
&\qquad \leq \frac{|\omega_0 \odot ... \odot \omega_{i-1}| + (k+j-1)}{|l_i|} \leq \frac{\varepsilon}{2|\mathcal{L}_k|},
\end{aligned}$$

which follows by the choice of l_i.

This implies that

$$\|\mathrm{P}_k(\omega_0 \odot ... \odot \omega_i, n_i) - \mathbf{q}_i\|_1 \leq \varepsilon$$

completing the proof of the lemma. □

We have now all the ingredients needed to prove Theorem 11.3.

Proof of Theorem 11.3. Let $\mathbf{q}_1, ..., \mathbf{q}_m \in S_1$ be such that for all $i \in \Sigma$ there exists $\mathbf{q}_{1,i} = (q_{i,1}, ...q_{i,N-1}), \mathbf{q}_{2,i} = (q'_{i,1}, ...q'_{i,N-1}) \in \{\mathbf{q}_1, ..., \mathbf{q}_m\}$ with $|q_{i,i} - q'_{i,i}| > 0$. Let $0 < \varepsilon < \frac{\min_{i\in\Sigma} |q_{i,i} - q'_{i,i}|}{2}$. For each $\omega \in \mathcal{L}_{\mathcal{P},\phi}$ let $\omega \odot u_{\omega,\varepsilon}$ be a word described by Lemma 11.8. Then for all $n \in \mathbb{N}$, we define sets C_n as follows :

$$C_n = \{\omega \odot u_{\omega,\varepsilon}\alpha_1\alpha_2... \in X_{\mathcal{P},\phi} | |\omega| = n, \text{ and } \alpha_i \in \Sigma\}.$$

Let I_n be the interior of C_n. Let $D_n = \cup_{k=n}^{\infty} I_k$, and $F = \cap_{n=1}^{\infty} D_n$. It is clear that D_n is open and dense in $X_{\mathcal{P},\phi}$. Since F is a countable intersection of open and dense sets, it is residual. We now need to show that if $w \in F$, then w is essentially non-normal. Let $w \in F$. Then there exists $(n_k)_k \subseteq \mathbb{N}$ such that $w \in C_{n_k}$. Lemma 11.8 then implies that for each digit of i, the sequence $(\mathrm{P}(\omega, i, n))_n$ does not converge. □

4. Proof of Theorem 11.4

Now we draw our attention to the case of extremely non-normal numbers and their Cesàro variants. Let $k \in \mathbb{N}$ and $\mathbf{q} \in S_k$ be fixed throughout the rest of this section. We consider how many copies of elements in Z_n we have to add in order to get the desired properties.

Lemma 11.9. *Let $\mathbf{q} \in \mathrm{S}_k$ and $n, t \in \mathbb{N}$ be positive integers. Furthermore, let $\omega = \omega_1 \dots \omega_t \in \mathcal{L}_t$ be a word of length t. Then, for any $\gamma \in Z_n(\mathbf{q}, k)$ and any*

$$\ell \geq R := t\left(1 + \frac{|\gamma|}{k}\right). \tag{4.1}$$

we get that

$$\left\|\mathrm{P}_k(\omega \odot \gamma^{\odot\ell}) - \mathbf{q}\right\| \leq \frac{4}{n}.$$

DÉMONSTRATION. We set $s := |\gamma|$, $\sigma := \omega \odot \gamma^{\odot \ell}$ and $L := |\sigma|$. For a fixed block $\mathbf{i}$ an occurrence can happen in ω, in γ or somewhere in between. Thus for every $\mathbf{i} \in \Sigma^k$ we clearly have that

$$\frac{\ell s}{L}\mathrm{P}(\gamma, \mathbf{i}) \leq \mathrm{P}(\sigma, \mathbf{i}) \leq \frac{\ell s \mathrm{P}(\gamma, \mathbf{i})}{L} + \frac{t + \ell(k+j-1)}{L}.$$

Now we concentrate on the occurrences inside the copies of γ and show that we may neglect all other occurrences, *i.e.*

$$\|\mathrm{P}_k(\sigma) - \mathbf{q}\| \leq \left\|\mathrm{P}_k(\sigma) - \frac{\ell s}{L}\mathrm{P}_k(\gamma)\right\| + \left\|\frac{\ell s}{L}\mathrm{P}_k(\gamma) - \mathbf{q}\right\|.$$

We will estimate both parts separately. For the first one we get that

$$\begin{aligned}\left\|\mathrm{P}_k(\sigma) - \frac{\ell s}{L}\mathrm{P}(\gamma, \mathbf{i})\right\| &= \sum_{\mathbf{i} \in \Sigma^k} \left|\mathrm{P}(\sigma, \mathbf{i}) - \frac{\ell s}{L}\mathrm{P}(\gamma, \mathbf{i})\right| \leq \sum_{\mathbf{i} \in \Sigma^k} \frac{t + \ell(k+j)}{L} \\ &\leq |\mathcal{L}_k| \, \frac{t + \ell(k+j)}{\ell n (k+j) \, |\mathcal{L}_k|} = \frac{t}{\ell n (k+j)} + \frac{1}{n}.\end{aligned}$$

where we have used that $L \geq \ell s \geq \ell n (k+j) \, |\mathcal{L}_k|$.

For the second part we get that

$$\begin{aligned}\left\|\frac{\ell s}{L}\mathrm{P}_k(\gamma) - \mathbf{q}\right\| &\leq \left\|\frac{\ell s}{L}\mathrm{P}_k(\gamma) - \mathrm{P}_k(\gamma)\right\| + \|\mathrm{P}_k(\gamma) - \mathbf{q}\| \\ &\leq \ell s \left|\frac{1}{L} - \frac{1}{\ell s}\right| + \frac{1}{n} \\ &\leq \frac{t + \ell j}{L} + \frac{1}{n} \leq \frac{t + \ell j}{\ell n (k+j) \, |\mathcal{L}_k|} + \frac{1}{n}.\end{aligned}$$

Putting these together yields

$$\|\mathrm{P}_k(\sigma) - \mathbf{q}\| \leq \frac{2}{n} + \frac{t}{\ell n (k+j)} + \frac{t + \ell j}{\ell n (k+j) \, |\mathcal{L}_k|}.$$

By our assumptions on the size of ℓ in (4.1) this proves the lemma. □

As in the papers of Olsen [**170**, **174**] our main idea is to construct a residual set $E \subset \mathbb{E}_k^{(r)}$. But before we start we want to ease up notation. To this end we recursively define the function $\varphi_1(x) = 2^x$ and $\varphi_m(x) = \varphi_1(\varphi_{m-1}(x))$ for $m \geq 2$. Furthermore, we set $\mathbb{D} = (\mathbb{Q}^N \cap \mathrm{S}_k)$. Since $\mathbb{D}$ is countable and dense in S_k we may concentrate on the probability vectors $\mathbf{q} \in \mathbb{D}$.

Now we say that a sequence $(\mathbf{x}_n)_n$ in $\mathbb{R}^{N^k}$ has property P if for all $\mathbf{q} \in \mathbb{D}$, $m \in \mathbb{N}$, $i \in \mathbb{N}$, and $\varepsilon > 0$, there exists a $j \in \mathbb{N}$ satisfying :

(1) $j \geq i$,
(2) $j/2^j < \varepsilon$,
(3) if $j < n < \varphi_m(j)$ then $\|\mathbf{x}_n - \mathbf{q}\| < \varepsilon$.

Then we define our set E to consist of all frequency vectors having property P, *i.e.*

$$E = \{x \in U_\infty : \ (\mathrm{P}_k^{(1)}(x; n))_{n=1}^\infty \text{ has property } P\}.$$

We will proceed in three steps showing that

(1) E is residual,
(2) if $(\mathrm{P}^{(r)}(x; n))_{n=1}^\infty$ has property P, then also $(\mathrm{P}^{(r+1)}(x; n))_{n=1}^\infty$ has property P, and

(3) $E \subseteq \mathbb{E}_k^{(r)}$.

LEMMA 11.10. *The set E is residual.*

DÉMONSTRATION. For fixed $h, m, i \in \mathbb{N}$ and $\mathbf{q} \in \mathbb{D}$, we say that a sequence $(\mathbf{x}_n)_n$ in $\mathbb{R}^{N^k}$ has property $P_{h,m,\mathbf{q},i}$ if for every $\varepsilon > 1/h$, there exists $j \in \mathbb{N}$ satisfying :

(1) $j \geq i$,

(2) $j/2^j < \varepsilon$,

(3) if $j < n < \varphi_m(2^j)$, then $\|\mathbf{x}_n - \mathbf{q}\| < \varepsilon$.

Now let $E_{h,m,\mathbf{q},i}$ be the set of all points whose frequency vector satisfies property $P_{h,m,\mathbf{q},i}$, *i.e.*

$$E_{h,m,\mathbf{q},i} := \left\{ x \in U_\infty : \left(\mathrm{P}_k^{(1)}(x;n) \right)_{n=1}^{\infty} \text{ has property } P_{h,m,\mathbf{q},i} \right\}.$$

Obviously we have that

$$E = \bigcap_{h\in\mathbb{N}} \bigcap_{m\in\mathbb{N}} \bigcap_{\mathbf{q}\in\mathbb{D}} \bigcap_{i\in\mathbb{N}} E_{h,m,\mathbf{q},i}.$$

Thus it remains to show, that $E_{h,m,\mathbf{q},i}$ is open and dense.

(1) $E_{h,m,\mathbf{q},i}$ **is open.** Let $x \in E_{h,m,\mathbf{q},i}$, then there exists a $j \in \mathbb{N}$ such that $j \geq i$, $j/2^j < 1/h$, and if $j < n < \varphi_m(2^j)$, then

$$\left\| \mathrm{P}_k^{(1)}(x;n) - \mathbf{q} \right\|_1 < 1/h.$$

Let $\omega \in X$ be such that $x = \pi(\omega)$ and set $t := \varphi_m(2^j)$. Since $D_t(\omega)$ is open, there exists a $\delta > 0$ such that the ball $B(x,\delta) \subseteq D_t(\omega)$. Furthermore, since all $y \in D_t(\omega)$ have their first $\varphi_m(2^j)$ digits the same as x, we get that

$$B(x,\delta) \subseteq D_t(\omega) \subseteq E_{h,m,\mathbf{q},i}.$$

(2) $E_{h,m,\mathbf{q},i}$ **is dense.** Let $x \in U_\infty$ and $\delta > 0$. We must find $y \in B(x,\delta) \cap E_{h,m,\mathbf{q},i}$.

Let $\omega \in X$ be such that $x = \pi(\omega)$. Since $\overline{D_t(\omega)} \to 0$ and $x \in D_t(\omega)$ there exists a t such that $D_t(\omega) \subset B(x,\delta)$. Let $\sigma = \omega|t$ be the first t digits of x.

Now, an application of Lemma 11.7 yields that there exists a finite word γ such that

$$\|\mathrm{P}_k(\gamma) - \mathbf{q}\| \leq \frac{1}{6h}.$$

Let $\varepsilon \geq \frac{1}{h}$ and L be as in the statement of Lemma 11.9. Then we choose j such that

$$\frac{j}{2^j} < \varepsilon \quad \text{and} \quad j \geq \max(L, i).$$

An application of Lemma 11.9 then gives us that

$$\|\mathrm{P}_k(\sigma\gamma^*|j) - \mathbf{q}\| \leq \frac{6}{n} = \frac{1}{h} \leq \varepsilon.$$

Thus we choose $y \in D_j(\sigma\gamma^*)$. Then on the one hand, $y \in D_j(\sigma\gamma^*) \subset D_t(\omega) \subset B(x,\delta)$ and on the other hand, $y \in D_j(\sigma\gamma^*) \subset E_{h,m,\mathbf{q},i}$

It follows that E is the countable intersection of open and dense sets and therefore E is residual in U_∞. □

LEMMA 11.11. *Let $\omega \in X_{\mathcal{P},\phi}$. If $(\mathrm{P}^{(r)}(\omega,n))_{n=1}^{\infty}$ has property P, then also $(\mathrm{P}^{(r+1)}(\omega,n))_{n=1}^{\infty}$ has property P.*

This is Lemma 2.2 of [**110**]. However, the proof is short so we present it here for completeness.

DÉMONSTRATION. Let $\omega \in X_{\mathcal{P},\phi}$ be such that $(\mathrm{P}_k^{(r)}(\omega;n))_{n=1}^{\infty}$ has property P, and fix $\varepsilon > 0, \mathbf{q} \in \mathrm{S}_k$, $i \in \mathbb{N}$ and $m \in \mathbb{N}$. Since $(\mathrm{P}_k^{(r)}(\omega,n))_{n=1}^{\infty}$ has property P, there exists $j' \in \mathbb{N}$ with $j' \geq i$, $j'/2^{j'} < \varepsilon/3$, and such that for $j' < n < \varphi_{m+1}(2^{j'})$ we have that $\left\|\mathrm{P}_k^{(r)}(\omega,n) - \mathbf{q}\right\| < \varepsilon/3$.

We set $j = 2^{j'}$ and show that $(\mathrm{P}_k^{(r+1)}(\omega,n))_{n=1}^{\infty}$ has property P with this j. For all $j < n < \varphi_m(2^j)$ (*i.e.* $2^{j'} < n < \varphi_{m+1}(2^{j'})$), we have

$$\begin{aligned}
\left\|\mathrm{P}_k^{(r+1)}(\omega,n) - \mathbf{q}\right\| &= \left\| \frac{\mathrm{P}_k^{(r)}(\omega,1) + \mathrm{P}_k^{(r)}(\omega,2) + \cdots + \mathrm{P}_k^{(r)}(\omega,n)}{n} - \mathbf{q} \right\| \\
&= \left\| \frac{\mathrm{P}_k^{(r)}(\omega,1) + \mathrm{P}_k^{(r)}(\omega,2) + \cdots + \mathrm{P}_k^{(r)}(\omega,j')}{n} \right. \\
&\quad \left. + \frac{\mathrm{P}_k^{(r)}(\omega,j'+1) + \mathrm{P}_k^{(r)}(\omega,2) + \cdots + \mathrm{P}_k^{(r)}(\omega,n) - (n-j')\mathbf{q}}{n} - \frac{j'\mathbf{q}}{n} \right\| \\
&\leq \frac{\left\|\mathrm{P}_k^{(r)}(\omega,1) + \mathrm{P}_k^{(r)}(\omega,2) + \cdots + \mathrm{P}_k^{(r)}(\omega,j')\right\|}{n} \\
&\quad + \frac{\left\|\mathrm{P}_k^{(r)}(\omega,j'+1) - \mathbf{q}\right\| + \cdots + \left\|\mathrm{P}_k^{(r)}(\omega,n) - \mathbf{q}\right\|}{n} - \frac{\|j'\mathbf{q}\|}{n} \\
&\leq \frac{j'}{n} + \frac{\varepsilon}{3}\frac{n-j'}{n} + \frac{j'}{n} \leq \frac{j'}{2^{j'}} + \frac{\varepsilon}{3} + \frac{j'}{2^{j'}} \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.
\end{aligned}$$

□

LEMMA 11.12. *The set E is a subset of $\mathbb{E}_k^{(r)}$.*

DÉMONSTRATION. We will show, that for any $x \in E$ we also have $x \in \mathbb{E}_k^{(r)}$. To this end, let $x \in E$ and $\omega \in X_{\mathcal{P},\phi}$ be the symbolic expansion of x. Since $(\mathrm{P}_k^{(1)}(\omega,n))_n$ has property P, by Lemma 11.11 we get that $(\mathrm{P}_k^{(r)}(\omega,n)$ has property P.

Thus it suffices to show that $\mathbf{p}$ is an accumulation point of $(\mathrm{P}_k^{(r)}(\omega,n))_n$ for any $\mathbf{p} \in \mathrm{S}_k$. Therefore, we fix $h \in \mathbb{N}$ and find a $\mathbf{q} \in \mathbb{D}$ such that

$$\|\mathbf{p} - \mathbf{q}\| < \frac{1}{h}.$$

Since $(\mathrm{P}_k^{(r)}(\omega,n))_n$ has property P for any $m \in \mathbb{N}$ we find $j \in \mathbb{N}$ with $j \geq h$ and such that if $j < n < \varphi_m(2^j)$ then $\left\|\mathrm{P}_k^{(r)}(\omega,n) - \mathbf{q}\right\| < \frac{1}{h}$. Hence let n_h be any integer with $j < n_h < \varphi_m(2^j)$, then

$$\left\|\mathrm{P}_k^{(r)}(\omega,n_h) - \mathbf{q}\right\| < \frac{1}{h}.$$

Thus each n_h in the sequence $(n_h)_h$ satisfies

$$\left\|\mathbf{p} - \mathrm{P}_k^{(r)}(\omega, n_h)\right\| \leq \|\mathbf{p} - \mathbf{q}\| + \left\|\mathrm{P}_k^{(r)}(\omega, n_h) - \mathbf{q}\right\| < \frac{2}{h}.$$

Since $n_h > h$ we may extract an increasing sub-sequence $(n_{h_u})_u$ such that $\mathrm{P}_k^{(r)}(\omega, n_{h_u}) \to \mathbf{p}$ for $u \to \infty$. Thus $\mathbf{p}$ is an accumulation point of $\mathrm{P}_k^{(r)}(\omega, n)$, which proves the lemma. □

Proof of Theorem 11.4. Since by Lemma 11.10 E is residual in U_∞ and by Lemma 11.12 E is a subset of $\mathbb{E}_k^{(r)}$ we get that $\mathbb{E}_k^{(r)}$ is residual in U_∞. Again we note that $M \setminus U_\infty$ is the countable union of nowhere dense sets and therefore $\mathbb{E}_k^{(r)}$ is also residual in M. □

Bibliographie

[1] B. Adamczewski and J. Bell, *An analogue of Cobham's theorem for fractals*, Trans. Amer. Math. Soc. **363** (2011), no. 8, 4421–4442. MR2792994

[2] R. Adler, M. Keane, and M. Smorodinsky, *A construction of a normal number for the continued fraction transformation*, J. Number Theory **13** (1981), no. 1, 95–105. MR602450 (82k :10070)

[3] S. Akiyama, T. Borbély, H. Brunotte, A. Pethő, and J. M. Thuswaldner, *Generalized radix representations and dynamical systems. I*, Acta Math. Hungar. **108** (2005), no. 3, 207–238. MR2162561 (2006i :37023)

[4] S. Akiyama, H. Brunotte, and A. Pethő, *Cubic CNS polynomials, notes on a conjecture of W. J. Gilbert*, J. Math. Anal. Appl. **281** (2003), no. 1, 402–415. MR1980100 (2004j :11009)

[5] S. Akiyama, H. Brunotte, A. Pethő, and J. M. Thuswaldner, *Generalized radix representations and dynamical systems. II*, Acta Arith. **121** (2006), no. 1, 21–61. MR2216302 (2007h :11085)

[6] S. Akiyama and H. Rao, *New criteria for canonical number systems*, Acta Arith. **111** (2004), no. 1, 5–25. MR2038059 (2005d :11007)

[7] S. Akiyama, H. Brunotte, and A. Pethő, *Reducible cubic CNS polynomials*, Period. Math. Hungar. **55** (2007), no. 2, 177–183. MR2375040 (2008m :11058)

[8] S. Akiyama and A. Pethő, *On canonical number systems*, Theoret. Comput. Sci. **270** (2002), no. 1-2, 921–933. MR1871104 (2002k :11009)

[9] S. Albeverio, M. Pratsiovytyi, and G. Torbin, *Singular probability distributions and fractal properties of sets of real numbers defined by the asymptotic frequencies of their s-adic digits*, Ukraïn. Mat. Zh. **57** (2005), no. 9, 1163–1170. MR2216038 (2006k :11156)

[10] ———, *Topological and fractal properties of real numbers which are not normal*, Bull. Sci. Math. **129** (2005), no. 8, 615–630. MR2166730 (2006g :28018)

[11] J.-P. Allouche, E. Cateland, W. J. Gilbert, H.-O. Peitgen, J. O. Shallit, and G. Skordev, *Automatic maps in exotic numeration systems*, Theory Comput. Syst. **30** (1997), no. 3, 285–331. MR1432196

[12] J.-P. Allouche, M. Mendès France, and J. Peyrière, *Automatic Dirichlet series*, J. Number Theory **81** (2000), no. 2, 359–373. MR1752260

[13] J.-P. Allouche, J. Shallit, and G. Skordev, *Self-generating sets, integers with missing blocks, and substitutions*, Discrete Math. **292** (2005), no. 1-3, 1–15. MR2131083

[14] C. Altomare and B. Mance, *Cantor series constructions contrasting two notions of normality*, Monatsh. Math. **164** (2011), no. 1, 1–22. MR2827169 (2012i :11077)

[15] N. Álvarez and V. Becher, *M. Levin's construction of absolutely normal numbers with very low discrepancy*, Math. Comp. **86** (2017), no. 308, 2927–2946. MR3667031

[16] G. E. Andrews, *The theory of partitions*, Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1976. Encyclopedia of Mathematics and its Applications, Vol. 2. MR0557013 (58 #27738)

[17] I.-S. Baek and L. Olsen, *Baire category and extremely non-normal points of invariant sets of IFS's*, Discrete Contin. Dyn. Syst. **27** (2010), no. 3, 935–943. MR2629566 (2011i :11112)

[18] D. H. Bailey and J. Borwein, *Pi Day is upon us again and we still do not know if pi is normal*, Amer. Math. Monthly **121** (2014), no. 3, 191–206. MR3168990

[19] N. L. Bassily and I. Kátai, *Distribution of the values of q-additive functions on polynomial sequences*, Acta Math. Hungar. **68** (1995), no. 4, 353–361. MR1333478 (96c :11112)

[20] V. Becher and S. Figueira, *An example of a computable absolutely normal number*, Theoret. Comput. Sci. **270** (2002), no. 1-2, 947–958. MR1871106 (2002m :11070)

[21] V. Becher, S. Figueira, and R. Picchi, *Turing's unpublished algorithm for normal numbers*, Theoret. Comput. Sci. **377** (2007), no. 1-3, 126–138. MR2323391

[22] V. Becher, P. A. Heiber, and T. A. Slaman, *A polynomial-time algorithm for computing absolutely normal numbers*, Inform. and Comput. **232** (2013), 1–9. MR3132518

[23] V. Bergelson, *Sets of recurrence of $\mathbf{Z}^m$-actions and properties of sets of differences in $\mathbf{Z}^m$*, J. London Math. Soc. (2) **31** (1985), no. 2, 295–304. MR809951

[24] ———, *A density statement generalizing Schur's theorem*, J. Combin. Theory Ser. A **43** (1986), no. 2, 338–343. MR867659

[25] ———, *Ergodic Ramsey theory*, Logic and combinatorics (Arcata, Calif., 1985), 1987, pp. 63–87. MR891243

[26] ———, *Ergodic Ramsey theory—an update*, Ergodic theory of $\mathbf{Z}^d$ actions (Warwick, 1993–1994), 1996, pp. 1–61. MR1411215

[27] V. Bergelson, H. Furstenberg, and R. McCutcheon, *IP-sets and polynomial recurrence*, Ergodic Theory Dynam. Systems **16** (1996), no. 5, 963–974. MR1417769

[28] V. Bergelson and I. J. Hå land, *Sets of recurrence and generalized polynomials*, Convergence in ergodic theory and probability (Columbus, OH, 1993), 1996, pp. 91–110. MR1412598

[29] V. Bergelson and I. J. Hå land Knutson, *Weak mixing implies weak mixing of higher orders along tempered functions*, Ergodic Theory Dynam. Systems **29** (2009), no. 5, 1375–1416. MR2545011

[30] V. Bergelson, G. Kolesnik, M. Madritsch, Y. Son, and R. Tichy, *Uniform distribution of prime powers and sets of recurrence and van der Corput sets in $\mathbb{Z}^k$*, Israel J. Math. **201** (2014), no. 2, 729–760. MR3265301

[31] V. Bergelson and E. Lesigne, *Van der Corput sets in $\mathbb{Z}^d$*, Colloq. Math. **110** (2008), no. 1, 1–49. MR2353898 (2008j :11089)

[32] V. Bergelson and R. McCutcheon, *An ergodic IP polynomial Szemerédi theorem*, Mem. Amer. Math. Soc. **146** (2000), no. 695, viii+106. MR1692634

[33] V. Berthé and M. Rigo (eds.), *Combinatorics, automata and number theory*, Encyclopedia of Mathematics and its Applications, vol. 135, Cambridge University Press, Cambridge, 2010. MR2742574

[34] A. Bertrand-Mathis, *Ensembles intersectifs et récurrence de Poincaré*, Israel J. Math. **55** (1986), no. 2, 184–198. MR868179 (87m :11071)

[35] ———, *Points génériques de Champernowne sur certains systèmes codes ; application aux θ-shifts*, Ergodic Theory Dynam. Systems **8** (1988), no. 1, 35–51. MR939059 (89d :94032)

[36] ———, *Comment écrire les nombres entiers dans une base qui n'est pas entière*, Acta Math. Hungar. **54** (1989), no. 3-4, 237–241. MR1029085

[37] ———, *Comment écrire les nombres relatifs dans une base qui n'est pas entière*, Unif. Distrib. Theory **9** (2014), no. 2, 135–156. MR3430815

[38] A. Bertrand-Mathis and B. Volkmann, *On (ϵ, k)-normal words in connecting dynamical systems*, Monatsh. Math. **107** (1989), no. 4, 267–279. MR1012458 (90m :11115)

[39] A. S. Besicovitch, *The asymptotic distribution of the numerals in the decimal representation of the squares of the natural numbers*, Math. Z. **39** (1935), no. 1, 146–156. MR1545494

[40] B. J. Birch, *Forms in many variables*, Proc. Roy. Soc. Ser. A **265** (1961/1962), 245–263. MR0150129 (27 #132)

[41] H. F. Blichfeldt, *Notes on geometry of numbers*, Bull. Amer. Math. Soc. **27** (1921), no. 4, 150–153.

[42] B. Boigelot and J. Brusten, *A generalization of Cobham's theorem to automata over real numbers*, Theoret. Comput. Sci. **410** (2009), no. 18, 1694–1703. MR2508527 (2010f :68133)

[43] E. Borel, *Les probabilités dénombrables et leurs applications arithmétiques.*, Palermo Rend. **27** (1909), 247–271 (French).

[44] J. Bourgain, *Ruzsa's problem on sets of recurrence*, Israel J. Math. **59** (1987), no. 2, 150–166. MR920079 (89d :11012)

[45] D. W. Boyd, *On the beta expansion for Salem numbers of degree* 6, Math. Comp. **65** (1996), no. 214, 861–875, *S*29–*S*31. MR1333306

[46] A. Bremner, *On power bases in cyclotomic number fields*, J. Number Theory **28** (1988), no. 3, 288–298. MR932377 (89d :11092)

[47] T. D. Browning and P. Vishe, *Cubic hypersurfaces and a version of the circle method for number fields*, Duke Math. J. **163** (2014), no. 10, 1825–1883. MR3229043

[48] T. Browning and R. Heath-Brown, *Forms in many variables and differing degrees*, J. Eur. Math. Soc. (JEMS) **19** (2017), no. 2, 357–394. MR3605019

[49] H. Brunotte, A. Huszti, and A. Pethő, *Bases of canonical number systems in quartic algebraic number fields*, J. Théor. Nombres Bordeaux **18** (2006), no. 3, 537–557. MR2330426 (2008g :11179)

[50] H. Brunotte, *On trinomial bases of radix representations of algebraic integers*, Acta Sci. Math. (Szeged) **67** (2001), no. 3-4, 521–527. MR1876451 (2002j :11005)

[51] ———, *Characterization of CNS trinomials*, Acta Sci. Math. (Szeged) **68** (2002), no. 3-4, 673–679. MR1954540 (2003k :11157)

[52] ———, *On cubic CNS polynomials with three real roots*, Acta Sci. Math. (Szeged) **70** (2004), no. 3-4, 495–504. MR2107523 (2005h :11055)

[53] ———, *Symmetric CNS trinomials*, Integers **9** (2009), A19, 201–214. MR2534909 (2010g :11039)

[54] ———, *A unified proof of two classical theorems on CNS polynomials*, Integers **12** (2012), no. 4, 709–721. MR2988542

[55] ———, *Unusual CNS polynomials*, Math. Pannon. **24** (2013), no. 1, 125–137. MR3234910

[56] Y. Bugeaud, *Distribution modulo one and Diophantine approximation*, Cambridge Tracts in Mathematics, vol. 193, Cambridge University Press, Cambridge, 2012. MR2953186

[57] G. J. Chaitin, *A theory of program size formally identical to information theory*, J. Assoc. Comput. Mach. **22** (1975), 329–340. MR0411829

[58] D. G. Champernowne, *The construction of decimals normal in the scale of ten*, J. Lond. Math. Soc. **8** (1933), 254–260 (English).

[59] A. Cobham, *On the base-dependence of sets of numbers recognizable by finite automata*, Math. Systems Theory **3** (1969), 186–192. MR0250789

[60] A. H. Copeland and P. Erdős, *Note on normal numbers*, Bull. Amer. Math. Soc. **52** (1946), 857–860. MR0017743 (8,194b)

[61] J. H. Curtiss, *A note on the theory of moment generating functions*, Ann. Math. Statistics **13** (1942), 430–433. MR0007577

[62] K. Dajani and C. Kraaikamp, *Ergodic theory of numbers*, Carus Mathematical Monographs, vol. 29, Mathematical Association of America, Washington, DC, 2002. MR1917322 (2003f :37014)

[63] H. Davenport and P. Erdős, *Note on normal decimals*, Canadian J. Math. **4** (1952), 58–63. MR0047084 (13,825g)

[64] H. Delange, *Sur la fonction sommatoire de la fonction "somme des chiffres"*, Enseignement Math. (2) **21** (1975), no. 1, 31–47. MR0379414 (52 #319)

[65] M. Drmota and G. Gutenbrunner, *The joint distribution of Q-additive functions on polynomials over finite fields*, J. Théor. Nombres Bordeaux **17** (2005), no. 1, 125–150. MR2152215 (2006c :11145)

[66] M. Drmota and R. F. Tichy, *Sequences, discrepancies and applications*, Lecture Notes in Mathematics, vol. 1651, Springer-Verlag, Berlin, 1997. MR1470456 (98j :11057)

[67] M. Drmota, *The joint distribution of q-additive functions*, Acta Arith. **100** (2001), no. 1, 17–39. MR1864623 (2002h :11093)

[68] M. Drmota and W. Steiner, *The Zeckendorf expansion of polynomial sequences*, J. Théor. Nombres Bordeaux **14** (2002), no. 2, 439–475. MR2040687 (2005c :11014)

[69] J. Dufresnoy and Ch. Pisot, *Etude de certaines fonctions méromorphes bornées sur le cercle unité. Application à un ensemble fermé d'entiers algébriques*, Ann. Sci. Ecole Norm. Sup. (3) **72** (1955), 69–92. MR0072902 (17,349d)

[70] J. M. Dumont, P. J. Grabner, and A. Thomas, *Distribution of the digits in the expansions of rational integers in algebraic bases*, Acta Sci. Math. (Szeged) **65** (1999), no. 3-4, 469–492. MR1737265 (2001f :11132)

[71] F. Durand, *Cobham's theorem for substitutions*, J. Eur. Math. Soc. (JEMS) **13** (2011), no. 6, 1799–1814. MR2835330 (2012i :68133)

[72] F. Durand and M. Rigo, *On Cobham's theorem*, Automata : from Mathematics to Applications, July 2011.

[73] S. Eilenberg, *Automata, languages, and machines. Vol. A*, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York, 1974. Pure and Applied Mathematics, Vol. 58. MR0530382

[74] P. D. T. A. Elliott, *Probabilistic number theory. I*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science], vol. 239, Springer-Verlag, New York, 1979. Mean-value theorems. MR551361 (82h :10002a)

[75] P. Erdős and J. Lehner, *The distribution of the number of summands in the partitions of a positive integer*, Duke Math. J. **8** (1941), 335–345. MR0004841 (3,69a)

[76] P. Erdős, C. Mauduit, and A. Sárközy, *On arithmetic properties of integers with missing digits. I. Distribution in residue classes*, J. Number Theory **70** (1998), no. 2, 99–120. MR1625049 (99e :11127)

[77] ———, *On arithmetic properties of integers with missing digits. II. Prime factors*, Discrete Math. **200** (1999), no. 1-3, 149–164. Paul Erdős memorial collection. MR1692287 (2000d :11103)

[78] P. Erdős and M. Szalay, *On the statistical theory of partitions*, Topics in classical number theory, Vol. I, II (Budapest, 1981), 1984, pp. 397–450. MR781149 (86f :11075)

[79] P. Flajolet, X. Gourdon, and P. Dumas, *Mellin transforms and asymptotics : harmonic sums*, Theoret. Comput. Sci. **144** (1995), no. 1-2, 3–58. Special volume on mathematical analysis of algorithms. MR1337752 (96h :68093)

[80] P. Flajolet, P. Grabner, P. Kirschenhofer, H. Prodinger, and R. F. Tichy, *Mellin transforms and asymptotics : digital sums*, Theoret. Comput. Sci. **123** (1994), no. 2, 291–314. MR1256203 (94m :11090)

[81] P. Flajolet and R. Sedgewick, *Analytic combinatorics*, Cambridge University Press, Cambridge, 2009. MR2483235 (2010h :05005)

[82] A. S. Fraenkel, *Systems of numeration*, Amer. Math. Monthly **92** (1985), no. 2, 105–114. MR777556 (86d :11016)

[83] J. Franke, Yu. I. Manin, and Yu. Tschinkel, *Rational points of bounded height on Fano varieties*, Invent. Math. **95** (1989), no. 2, 421–435. MR974910 (89m :11060)

[84] C. Frei and M. Madritsch, *Forms of differing degrees over number fields*, Mathematika **63** (2017), no. 1, 92–123. MR3610007

[85] C. Frei and M. Pieropan, *O-minimality on twisted universal torsors and Manin's conjecture over number fields*, Ann. Sci. Éc. Norm. Supér. (4) **49** (2016), no. 4, 757–811. MR3552013

[86] K. Fukuyama, *Metric discrepancy results for alternating geometric progressions*, Monatsh. Math. **171** (2013), no. 1, 33–63. MR3066814

[87] H. Furstenberg, *Recurrence in ergodic theory and combinatorial number theory*, Princeton University Press, Princeton, N.J., 1981. M. B. Porter Lectures. MR603625

[88] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256. MR0498471 (58 #16583)

[89] I. Gaál and L. Robertson, *Power integral bases in prime-power cyclotomic fields*, J. Number Theory **120** (2006), no. 2, 372–384. MR2257552

[90] A. O. Gelfond, *A common property of number systems*, Izv. Akad. Nauk SSSR. Ser. Mat. **23** (1959), 809–814. MR0109817 (22 #702)

[91] ———, *Sur les nombres qui ont des propriétés additives et multiplicatives données*, Acta Arith. **13** (1967/1968), 259–265. MR0220693 (36 #3745)

[92] W. J. Gilbert, *Radix representations of quadratic fields*, J. Math. Anal. Appl. **83** (1981), no. 1, 264–274. MR632342 (83m :12005)

[93] B. Gittenberger and J. M. Thuswaldner, *The moments of the sum-of-digits function in number fields*, Canad. Math. Bull. **42** (1999), no. 1, 68–77. MR1695870 (2000f :11096)

[94] ———, *Asymptotic normality of b-additive functions on polynomial sequences in the Gaussian number field*, J. Number Theory **84** (2000), no. 2, 317–341. MR1796518 (2001k :11144)

[95] W. M. Y. Goh and E. Schmutz, *The number of distinct part sizes in a random integer partition*, J. Combin. Theory Ser. A **69** (1995), no. 1, 149–158. MR1309156

[96] P. J. Grabner, P. Kirschenhofer, H. Prodinger, and R. F. Tichy, *On the moments of the sum-of-digits function*, Applications of Fibonacci numbers, Vol. 5 (St. Andrews, 1992), 1993, pp. 263–271. MR1271366 (95d :11123)

[97] S. W. Graham and G. Kolesnik, *van der Corput's method of exponential sums*, London Mathematical Society Lecture Note Series, vol. 126, Cambridge University Press, Cambridge, 1991. MR1145488 (92k :11082)

[98] K. Gröchenig and A. Haas, *Self-similar lattice tilings*, J. Fourier Anal. Appl. **1** (1994), no. 2, 131–170. MR1348740 (96j :52037)

[99] V. Grünwald, *Intorno all' aritmetica dei sistemi numerici a base negativa con particolare riguardo al sistema numerico a base negativo-decimale per lo studio delle sue analogie coll' aritmetica ordinaria (decimale)*., Battaglini G. **23** (1885), 203–221 (Italian).

[100] K. Győry, *Sur les polynômes à coefficients entiers et de discriminant donné. III*, Publ. Math. Debrecen **23** (1976), no. 1-2, 141–165. MR0437491 (55 #10419c)

[101] G. Hansel and T. Safer, *Vers un théorème de Cobham pour les entiers de Gauss*, Bull. Belg. Math. Soc. Simon Stevin **10** (2003), no. suppl., 723–735. MR2073023 (2005c :68236)

[102] G. H. Hardy, *Divergent series*, Éditions Jacques Gabay, Sceaux, 1992. With a preface by J. E. Littlewood and a note by L. S. Bosanquet, Reprint of the revised (1963) edition. MR1188874 (93g :01100)

[103] G. H. Hardy and S. Ramanujan, *Asymptotic formulae in combinatory analysis*, Proc. London Math. Soc. **17** (1918), 75–115.

[104] K. G. Hare, S. Laishram, and T. Stoll, *Stolarsky's conjecture and the sum of digits of polynomial values*, Proc. Amer. Math. Soc. **139** (2011), no. 1, 39–49. MR2729069 (2011j :11012)

[105] G. Harman, *Trigonometric sums over primes. I*, Mathematika **28** (1981), no. 2, 249–254 (1982). MR645105 (83j :10045)

[106] D. R. Heath-Brown, *Cubic forms in ten variables*, Proc. London Math. Soc. (3) **47** (1983), no. 2, 225–257. MR703978

[107] ———, *The Pjateckiĭ-Šapiro prime number theorem*, J. Number Theory **16** (1983), no. 2, 242–266. MR698168 (84j :10053)

[108] L.-K. Hua, *Introduction to number theory*, Springer-Verlag, Berlin, 1982. Translated from the Chinese by Peter Shiu. MR665428 (83f :10001)

[109] H.-K. Hwang, *Limit theorems for the number of summands in integer partitions*, J. Combin. Theory Ser. A **96** (2001), no. 1, 89–126. MR1855788 (2002f :11138)

[110] J. Hyde, V. Laschos, L. Olsen, I. Petrykiewicz, and A. Shaw, *Iterated Cesàro averages, frequencies of digits, and Baire category*, Acta Arith. **144** (2010), no. 3, 287–293. MR2672291 (2011f :11089)

[111] G. Ifrah, *Histoire universelle des chiffres*, Robert Laffont, Paris, 1994. 2nd edition.

[112] A. E. Ingham, *A Tauberian theorem for partitions*, Ann. of Math. (2) **42** (1941), 1075–1090. MR0005522 (3,166a)

[113] S. Ito and I. Shiokawa, *A construction of β-normal sequences*, J. Math. Soc. Japan **27** (1975), 20–23. MR0357361 (50 #9829)

[114] H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004. MR2061214 (2005h :11005)

[115] T. Kamae and M. Mendès France, *Van der Corput's difference theorem*, Israel J. Math. **31** (1978), no. 3-4, 335–342. MR516154 (80a :10070)

[116] I. Kátai, *On the sum of digits of primes*, Acta Math. Acad. Sci. Hungar. **30** (1977), no. 1–2, 169–173. MR0472747 (57 #12437)

[117] I. Kátai and I. Környei, *On number systems in algebraic number fields*, Publ. Math. Debrecen **41** (1992), no. 3-4, 289–294. MR1189110 (93m :11107)

[118] I. Kátai and B. Kovács, *Kanonische Zahlensysteme in der Theorie der quadratischen algebraischen Zahlen*, Acta Sci. Math. (Szeged) **42** (1980), no. 1-2, 99–107. MR576942 (81i :12002)

[119] ———, *Canonical number systems in imaginary quadratic fields*, Acta Math. Acad. Sci. Hungar. **37** (1981), no. 1-3, 159–164. MR616887 (83a :12005)

[120] I. Kátai and J. Szabó, *Canonical number systems for complex integers*, Acta Sci. Math. (Szeged) **37** (1975), no. 3-4, 255–260. MR0389759 (52 #10590)

[121] A. J. Kempner, *Anormal systems of numeration.*, Am. Math. Mon. **43** (1936), 610–617 (English).

[122] S. I. Khmelnik, *Specialized digital computer for operations with complex numbers*, Questions of Radio Electronics **XII** (1964), no. 2. in Russian.

[123] P. Kirschenhofer and J. M. Thuswaldner, *Shift radix systems—a survey*, Numeration and substitution 2012, 2014, pp. 1–59. MR3330559

[124] D. E. Knuth, *The art of computer programming. Vol. 2*, Second, Addison-Wesley Publishing Co., Reading, Mass., 1981. Seminumerical algorithms, Addison-Wesley Series in Computer Science and Information Processing. MR633878 (83i :68003)

[125] D. Knuth, *Negafibonacci numbers and the hyperbolic plane*, December 15, 2013. Paper presented at the annual meeting of the Mathematical Association of America, The Fairmont Hotel, San Jose.

[126] D. E. Knuth, *An imaginary number system*, Comm. ACM **3** (1960), 245–247. MR0127508

[127] ———, *The art of computer programming. Vol. 2 : Seminumerical algorithms*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont, 1969. MR0286318

[128] B. Kovács, *Canonical number systems in algebraic number fields*, Acta Math. Acad. Sci. Hungar. **37** (1981), no. 4, 405–407. MR619892 (82j :12014)

[129] B. Kovács and A. Pethő, *Number systems in integral domains, especially in orders of algebraic number fields*, Acta Sci. Math. (Szeged) **55** (1991), no. 3-4, 287–299. MR1152592 (92m :11116)

[130] ———, *On a representation of algebraic integers*, Studia Sci. Math. Hungar. **27** (1992), no. 1-2, 169–172. MR1207568 (94d :11080)

[131] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley-Interscience [John Wiley & Sons], New York, 1974. Pure and Applied Mathematics. MR0419394 (54 #7415)

[132] S. Lang, *Introduction to transcendental numbers*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1966. MR0214547

[133] H. Lebesgue, *Sur certaines démonstrations d'existence*, Bull. Soc. Math. France **45** (1917), 132–144. MR1504765

[134] H. W. Lenstra Jr., *Solving the Pell equation*, Notices Amer. Math. Soc. **49** (2002), no. 2, 182–192. MR1875156

[135] M. B. Levin, *Absolutely normal numbers*, Vestnik Moskov. Univ. Ser. I Mat. Mekh. **1** (1979), 31–37, 87. MR525299

[136] ———, *On the discrepancy estimate of normal numbers*, Acta Arith. **88** (1999), no. 2, 99–111. MR1700240

[137] B. Li and J. Wu, *Beta-expansion and continued fraction expansion*, J. Math. Anal. Appl. **339** (2008), no. 2, 1322–1331. MR2377089 (2008m :11148)

[138] D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding*, Cambridge University Press, Cambridge, 1995. MR1369092 (97a :58050)

[139] D. Loughran, *Rational points of bounded height and the Weil restriction*, Israel J. Math. **210** (2015), no. 1, 47–79. MR3430268

[140] M. Madritsch and T. Stoll, *On simultaneous digital expansions of polynomial values*, Acta Math. Hungar. **143** (2014), no. 1, 192–200. MR3215614

[141] M. G. Madritsch, *A note on normal numbers in matrix number systems*, Math. Pannon. **18** (2007), no. 2, 219–227. MR2363115

[142] ———, *Generating normal numbers over Gaussian integers*, Acta Arith. **135** (2008), no. 1, 63–90. MR2453524

[143] ———, *Asymptotic normality of b-additive functions on polynomial sequences in number systems*, Ramanujan J. **21** (2010), no. 2, 181–210. MR2593247

[144] ———, *Non-normal numbers with respect to Markov partitions*, Discrete Contin. Dyn. Syst. **34** (2014), no. 2, 663–676.

[145] M. G. Madritsch and J. M. Thuswaldner, *Additive functions for number systems in function fields*, Finite Fields Appl. **16** (2010), no. 3, 204–229. MR2610710

[146] M. G. Madritsch, J. M. Thuswaldner, and R. F. Tichy, *Normality of numbers generated by the values of entire functions*, J. Number Theory **128** (2008), no. 5, 1127–1145.

[147] M. G. Madritsch and R. F. Tichy, *Construction of normal numbers via generalized prime power sequences*, J. Integer Seq. **16** (2013), no. 2, Article 13.2.12, 17. MR3032395

[148] M. G. Madritsch and V. Ziegler, *On multiplicatively independent bases in cyclotomic number fields*, Acta Math. Hungar. **146** (2015), no. 1, 224–239. MR3348190

[149] M. G. Madritsch, *The summatory function of q-additive functions on pseudo-polynomial sequences*, J. Théor. Nombres Bordeaux **24** (2012), no. 1, 153–171. MR2914904

[150] ———, *Construction of normal numbers via pseudo-polynomial prime sequences*, Acta Arith. **166** (2014), no. 1, 81–100. MR3273499

[151] M. G. Madritsch and A. Pethő, *Asymptotic normality of additive functions on polynomial sequences in canonical number systems*, J. Number Theory **131** (2011), no. 9, 1553–1574. MR2802135 (2012f :11140)

[152] ———, *The moments of b-additive functions in canonical number systems*, Acta Sci. Math. (Szeged) **78** (2012), no. 3-4, 403–418. MR3052471

[153] M. G. Madritsch and R. F. Tichy, *Dynamical systems and uniform distribution of sequences*, From arithmetic to zeta-functions, 2016, pp. 263–276. MR3642360

[154] M. G. Madritsch and V. Ziegler, *An infinite family of multiplicatively independent bases of number systems in cyclotomic number fields*, Acta Sci. Math. (Szeged) **81** (2015), no. 1-2, 33–44. MR3381872

[155] B. Mance, *Construction of normal numbers with respect to the Q-Cantor series expansion for certain Q*, Acta Arith. **148** (2011), no. 2, 135–152. MR2786161 (2012c :11153)

[156] ———, *Cantor series constructions of sets of normal numbers*, Acta Arith. **156** (2012), no. 3, 223–245. MR2999070

[157] C. Mauduit and A. Sárközy, *On the arithmetic structure of sets characterized by sum of digits properties*, J. Number Theory **61** (1996), no. 1, 25–38. MR1418316 (97g :11107)

[158] ———, *On the arithmetic structure of the integers whose sum of digits is fixed*, Acta Arith. **81** (1997), no. 2, 145–173. MR1456239 (99a :11096)

[159] C. Mauduit and J. Rivat, *Sur un problème de Gelfond : la somme des chiffres des nombres premiers*, Ann. of Math. (2) **171** (2010), no. 3, 1591–1646. MR2680394 (2011j :11137)

[160] R. McCutcheon, *Three results in recurrence*, Ergodic theory and its connections with harmonic analysis (Alexandria, 1993), 1995, pp. 349–358. MR1325710

[161] G. Meinardus, *Asymptotische Aussagen über Partitionen*, Math. Z. **59** (1954), 388–398. MR0062781 (16,17e)

[162] H. L. Montgomery, *Ten lectures on the interface between analytic number theory and harmonic analysis*, CBMS Regional Conference Series in Mathematics, vol. 84, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1994. MR1297543 (96i :11002)

[163] N. G. Moshchevitin and I. D. Shkredov, *On the Pyatetskiĭ-Shapiro criterion for normality*, Mat. Zametki **73** (2003), no. 4, 577–589. MR1991904 (2005e :37013)

[164] W. Müller, J. M. Thuswaldner, and R. F. Tichy, *Fractal properties of number systems*, Period. Math. Hungar. **42** (2001), no. 1-2, 51–68. MR1832694 (2002k :11130)

[165] H. Nakada, *Metrical theory for a class of continued fraction transformations and their natural extensions*, Tokyo J. Math. **4** (1981), no. 2, 399–426. MR646050 (83k :10095)

[166] Y. Nakai and I. Shiokawa, *A class of normal numbers*, Japan. J. Math. (N.S.) **16** (1990), no. 1, 17–29. MR1064444 (91g :11081)

[167] ———, *Discrepancy estimates for a class of normal numbers*, Acta Arith. **62** (1992), no. 3, 271–284. MR1197421 (94a :11113)

[168] ———, *Normality of numbers generated by the values of polynomials at primes*, Acta Arith. **81** (1997), no. 4, 345–356. MR1472814 (98h :11098)

[169] M. B. Nathanson, *Additive number theory*, Graduate Texts in Mathematics, vol. 164, Springer-Verlag, New York, 1996. The classical bases. MR1395371 (97e :11004)

[170] L. Olsen, *Extremely non-normal continued fractions*, Acta Arith. **108** (2003), no. 2, 191–202. MR1974522 (2004f :11080)

[171] ———, *Multifractal analysis of divergence points of deformed measure theoretical Birkhoff averages*, J. Math. Pures Appl. (9) **82** (2003), no. 12, 1591–1649. MR2025314 (2004k :37036)

[172] ———, *Applications of multifractal divergence points to sets of numbers defined by their N-adic expansion*, Math. Proc. Cambridge Philos. Soc. **136** (2004), no. 1, 139–165. MR2034019 (2004j :11090)

[173] ———, *Applications of multifractal divergence points to some sets of d-tuples of numbers defined by their N-adic expansion*, Bull. Sci. Math. **128** (2004), no. 4, 265–289. MR2052170 (2005a :28019)

[174] ———, *Extremely non-normal numbers*, Math. Proc. Cambridge Philos. Soc. **137** (2004), no. 1, 43–53. MR2075041 (2005f :11156)

[175] L. Olsen and S. Winter, *Normal and non-normal points of self-similar sets and divergence points of self-similar measures*, J. London Math. Soc. (2) **67** (2003), no. 1, 103–122. MR1942414 (2003i :28009)

[176] ———, *Multifractal analysis of divergence points of deformed measure theoretical Birkhoff averages. II. Non-linearity, divergence points and Banach space valued spectra*, Bull. Sci. Math. **131** (2007), no. 6, 518–558. MR2351308 (2010b :28023)

[177] W. Parry, *On the β-expansions of real numbers*, Acta Math. Acad. Sci. Hungar. **11** (1960), 401–416. MR0142719 (26 #288)

[178] W. Penney, *A "binary" system for complex numbers*, J. ACM **12** (April 1965), no. 2, 247–248.

[179] A. Pethő, *On a polynomial transformation and its application to the construction of a public key cryptosystem*, Computational number theory (Debrecen, 1989), 1991, pp. 31–43. MR1151853 (93e :94011)

[180] A. Pethő, *Notes on CNS polynomials and integral interpolation*, More sets, graphs and numbers, 2006, pp. 301–315. MR2223397 (2007d :11027)

[181] A. Pethő and R. F. Tichy, *S-unit equations, linear recurrences and digit expansions*, Publ. Math. Debrecen **42** (1993), no. 1-2, 145–154. MR1208858 (94a :11013)

[182] A. Pethő, *Algebraische algorithmen*, Friedr. Vieweg & Sohn, Braunschweig, 1999. MR1711312 (2000k :68180)

[183] E. Peyre, *Hauteurs et mesures de Tamagawa sur les variétés de Fano*, Duke Math. J. **79** (1995), no. 1, 101–218. MR1340296 (96h :11062)

[184] W. Philipp, *Limit theorems for lacunary series and uniform distribution* mod 1, Acta Arith. **26** (1974/75), no. 3, 241–251. MR0379420

[185] S. S. Pillai, *On normal numbers*, Proc. Indian Acad. Sci., Sect. A. **12** (1940), 179–184. MR0002324

[186] J. Pintz, W. L. Steiger, and E. Szemerédi, *On sets of natural numbers whose difference set contains no squares*, J. London Math. Soc. (2) **37** (1988), no. 2, 219–231. MR928519

[187] A. G. Postnikov, *Arithmetic modeling of random processes*, Trudy Mat. Inst. Steklov. **57** (1960), 84. MR0148639 (26 #6146)

[188] A. G. Postnikov and I. I. Pyateckiĭ, *A Markov-sequence of symbols and a normal continued fraction*, Izv. Akad. Nauk SSSR. Ser. Mat. **21** (1957), 729–746. MR0101857 (21 #664)

[189] ———, *On Bernoulli-normal sequences of symbols*, Izv. Akad. Nauk SSSR. Ser. Mat. **21** (1957), 501–514. MR0101856 (21 #663)

[190] H. Rademacher, *On the Partition Function p(n)*, Proc. London Math. Soc. **S2-43** (1937), no. 4, 241. MR1575213

[191] D. Ralaivaosaona, *On the number of summands in a random prime partition*, Monatsh. Math. **166** (2012), no. 3-4, 505–524. MR2925152

[192] K. Ramachandra, *Contributions to the theory of transcendental numbers. I, II*, Acta Arith. 14 (1967/68), 65-72 ; ibid. **14** (1967/1968), 73–88. MR0224566

[193] G. Ranieri, *Générateurs de l'anneau des entiers d'une extension cyclotomique*, J. Number Theory **128** (2008), no. 6, 1576–1586. MR2419179

[194] A. Rényi, *Representations for real numbers and their ergodic properties*, Acta Math. Acad. Sci. Hungar **8** (1957), 477–493. MR0097374 (20 #3843)

[195] G. Rhin, *Sur la répartition modulo* 1 *des suites* $f(p)$, Acta Arith. **23** (1973), 217–248. MR0323731

[196] L. Robertson and R. Russell, *A hybrid Gröbner bases approach to computing power integral bases*, Acta Math. Hungar. **147** (2015), no. 2, 427–437. MR3420587

[197] L. Robertson, *Power bases for cyclotomic integer rings*, J. Number Theory **69** (1998), no. 1, 98–118. MR1611089

[198] ———, *Power bases for 2-power cyclotomic fields*, J. Number Theory **88** (2001), no. 1, 196–209. MR1825999

[199] ———, *Monogeneity in cyclotomic fields*, Int. J. Number Theory **6** (2010), no. 7, 1589–1607. MR2740723

[200] D. Rosen, *A class of continued fractions associated with certain properly discontinuous groups*, Duke Math. J. **21** (1954), 549–563. MR0065632 (16,458d)

[201] K. F. Roth and G. Szekeres, *Some asymptotic formulae in the theory of partitions*, Quart. J. Math., Oxford Ser. (2) **5** (1954), 241–259. MR0067913 (16,797b)

[202] J. Sakarovitch, *Elements of automata theory*, Cambridge University Press, Cambridge, 2009. Translated from the 2003 French original by Reuben Thomas. MR2567276

[203] T. Šalát, *Zur metrischen Theorie der Lürothschen Entwicklungen der reellen Zahlen*, Czechoslovak Math. J. **18 (93)** (1968), 489–522. MR0229605 (37 #5179)

[204] T. Šalát, *A remark on normal numbers*, Rev. Roumaine Math. Pures Appl. **11** (1966), 53–56. MR0201386 (34 #1270)

[205] ———, *Über die Cantorschen Reihen*, Czechoslovak Math. J. **18 (93)** (1968), 25–56. MR0223305 (36 #6353)

[206] A. Sárközy, *On difference sets of sequences of integers. I*, Acta Math. Acad. Sci. Hungar. **31** (1978), no. 1–2, 125–149. MR0466059 (57 #5942)

[207] A. Sárközy, *On difference sets of sequences of integers. II*, Ann. Univ. Sci. Budapest. Eötvös Sect. Math. **21** (1978), 45–53 (1979). MR536201 (80j :10062a)

[208] ———, *On difference sets of sequences of integers. III*, Acta Math. Acad. Sci. Hungar. **31** (1978), no. 3-4, 355–386. MR487031 (80j :10062b)

[209] A.-M. Scheerer, *Computable absolutely normal numbers and discrepancies*, Math. Comp. **86** (2017), no. 308, 2911–2926. MR3667030

[210] ———, *Normality in Pisot numeration systems*, Ergodic Theory Dynam. Systems **37** (2017), no. 2, 664–672. MR3614043

[211] K. Scheicher, P. Surer, J. M. Thuswaldner, and C. E. van de Woestijne, *Digit systems over commutative rings*, Int. J. Number Theory **10** (2014), no. 6, 1459–1483.

[212] K. Scheicher and J. M. Thuswaldner, *Canonical number systems, counting automata and fractals*, Math. Proc. Cambridge Philos. Soc. **133** (2002), no. 1, 163–182. MR1900260 (2003b :11070)

[213] ———, *On the characterization of canonical number systems*, Osaka J. Math. **41** (2004), no. 2, 327–351. MR2069090 (2005c :11013)

[214] J. Schiffer, *Discrepancy of normal numbers*, Acta Arith. **47** (1986), no. 2, 175–186. MR867496 (88d :11072)

[215] D. Schindler and A. Skorobogatov, *Norms as products of linear polynomials*, J. Lond. Math. Soc. (2) **89** (2014), no. 2, 559–580. MR3188633

[216] H. P. Schlickewei, *Linear equations in integers with bounded sum of digits*, J. Number Theory **35** (1990), no. 3, 335–344. MR1062338

[217] W. M. Schmidt, *On normal numbers*, Pacific J. Math. **10** (1960), 661–672. MR0117212 (22 #7994)

[218] ———, *über die Normalität von Zahlen zu verschiedenen Basen*, Acta Arith. **7** (1961/1962), 299–309. MR0140482

[219] ———, *Irregularities of distribution. VII*, Acta Arith. **21** (1972), 45–50. MR0319933

[220] ———, *The density of integer points on homogeneous varieties*, Acta Math. **154** (1985), no. 3-4, 243–296. MR781588 (86h :11027)

[221] E. Schmutz, *Part sizes of random integer partitions*, Indian J. Pure Appl. Math. **25** (1994), no. 6, 567–575. MR1285219

[222] P. Schreiber, *A note on the cattle problem of Archimedes*, Historia Math. **20** (1993), no. 3, 304–306. MR1238181

[223] H. G. Senge and E. G. Straus, PV*-numbers and sets of multiplicity*, Period. Math. Hungar. **3** (1973), 93–100. Collection of articles dedicated to the memory of Alfréd Rényi, II. MR0340185

[224] I. Shiokawa, *On the sum of digits of prime numbers*, Proc. Japan Acad. **50** (1974), 551–554. MR0369238 (51 #5473)

[225] I. D. Shkredov, *On the Pyatetskiĭ-Shapiro normality criterion for continued fractions*, Fundam. Prikl. Mat. **16** (2010), no. 6, 177–188. MR2825526

[226] A. Siegel, *Toward a usable theory of chernoff bounds for heterogeneous and partially dependent random variables*. (1992).

[227] C. L. Siegel, *Generalization of Waring's problem to algebraic number fields*, Amer. J. Math. **66** (1944), 122–136. MR0009778 (5,200c)

[228] W. Sierpinski, *Démonstration élémentaire du théorème de M. Borel sur les nombres absolument normaux et détermination effective d'une tel nombre*, Bull. Soc. Math. France **45** (1917), 125–132. MR1504764

[229] K. Sigmund, *On dynamical systems with the specification property*, Trans. Amer. Math. Soc. **190** (1974), 285–299. MR0352411 (50 #4898)

[230] C. M. Skinner, *Forms over number fields and weak approximation*, Compositio Math. **106** (1997), no. 1, 11–29. MR1446148 (98b :14021)

[231] C. M. Skinner, *Rational points on nonsingular cubic hypersurfaces*, Duke Math. J. **75** (1994), no. 2, 409–466. MR1290198 (95d :11138)

[232] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, Published electronically at `http://oeis.org/` (2010).

[233] M. Smorodinsky and B. Weiss, *Normal sequences for Markov shifts and intrinsically ergodic subshifts*, Israel J. Math. **59** (1987), no. 2, 225–233. MR920084 (89b :28006)

[234] S. Srinivasan and R. F. Tichy, *Uniform distribution of prime power sequences*, Anz. Österreich. Akad. Wiss. Math.-Natur. Kl. **130** (1993), 33–36 (1994). MR1294872 (95h :11071)

[235] C. L. Stewart, *On the representation of an integer in two different bases*, J. Reine Angew. Math. **319** (1980), 63–72. MR586115

[236] K. B. Stolarsky, *The binary digits of a power*, Proc. Amer. Math. Soc. **71** (1978), no. 1, 1–5. MR495823 (80b :10020)

[237] G. Tenenbaum, *Introduction à la théorie analytique et probabiliste des nombres*, Second, Cours Spécialisés [Specialized Courses], vol. 1, Société Mathématique de France, Paris, 1995. MR1366197 (97e :11005a)

[238] J. M. Thuswaldner, *The sum of digits function in number fields*, Bull. London Math. Soc. **30** (1998), no. 1, 37–45. MR1479034 (98i :11088)

[239] ———, *The sum of digits function in number fields : distribution in residue classes*, J. Number Theory **74** (1999), no. 1, 111–125. MR1670564 (99m :11134)

[240] J. M. Thuswaldner and R. F. Tichy, *Waring's problem with digital restrictions*, Israel J. Math. **149** (2005), 317–344. Probability in mathematics. MR2191219 (2006k :11190)

[241] A. M. Turing, *A note on normal numbers*, Collected works of a.m. turing, 1992, pp. 117–119.

[242] J. D. Vaaler, *Some extremal functions in Fourier analysis*, Bull. Amer. Math. Soc. (N.S.) **12** (1985), no. 2, 183–216. MR776471 (86g :42005)

[243] J. Vandehey, *A simpler normal number construction for simple Lüroth series*, J. Integer Seq. **17** (2014), no. 6, Article 14.6.1, 18. MR3209785

[244] R. C. Vaughan, *An elementary method in prime number theory*, Acta Arith. **37** (1980), 111–115. MR598869

[245] ———, *The Hardy-Littlewood method*, Second, Cambridge Tracts in Mathematics, vol. 125, Cambridge University Press, Cambridge, 1997. MR1435742 (98a :11133)

[246] I. M. Vinogradov, *Selected works*, Springer-Verlag, Berlin, 1985. With a biography by K. K. Mardzhanishvili, Translated from the Russian by Naidu Psv [P. S. V. Naidu], Translation edited by Yu. A. Bakhturin. MR807530 (87a :01042)

[247] ———, *The method of trigonometrical sums in the theory of numbers*, Dover Publications Inc., Mineola, NY, 2004. Translated from the Russian, revised and annotated by K. F. Roth and Anne Davenport, Reprint of the 1954 translation. MR2104806 (2005f :11172)

[248] K. Vogel, *Vorgriechische mathematik i : Vorgeschichte und Ägypten*, Mathematische Studienhefte, Hermann Schödel Verlag KG, Hannover, 1958.

[249] B. Volkmann, *On non-normal numbers*, Compositio Math. **16** (1964), 186–190 (1964). MR0174548 (30 #4749)

[250] M. Waldschmidt, *Diophantine approximation on linear algebraic groups*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 326, Springer-Verlag, Berlin, 2000. Transcendence properties of the exponential function in several variables. MR1756786

[251] D. D. Wall, *NORMAL NUMBERS*, ProQuest LLC, Ann Arbor, MI, 1950. Thesis (Ph.D.)–University of California, Berkeley. MR2937990

[252] H. Weyl, *Über die Gleichverteilung von Zahlen mod. Eins.*, Math. Ann. **77** (1916), 313–352 (German).

[253] H. C. Williams, R. A. German, and C. R. Zarnke, *Solution of the Cattle Problem of Archimedes*, Math. Comp. **19** (1965), no. 92, 671–674.

[254] E. Zeckendorf, *Représentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas*, Bull. Soc. Roy. Sci. Liège **41** (1972), 179–182. MR0308032 (46 #7147)