Université de Montréal

Irrégularités dans la distribution des nombres premiers et des suites plus générales dans les progressions arithmétiques

par

Daniel Fiorilli

Département de mathématiques et de statistique
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de
Philosophiæ Doctor (Ph.D.)
en Mathématiques

août 2011

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée

Irrégularités dans la distribution des nombres premiers et des suites plus générales dans les progressions arithmétiques

présentée par

Daniel Fiorilli

a été évaluée par un jury composé des personnes suivantes :

Matilde Lalín
(président-rapporteur)

Andrew Granville
(directeur de recherche)

Chantal David
(membre du jury)

John Friedlander
(examinateur externe)

Pierre McKenzie
(représentant du doyen de la FESP)

Thèse acceptée le:
25 août 2011

SOMMAIRE

Le sujet principal de cette thèse est la distribution des nombres premiers dans les progressions arithmétiques, c'est-à-dire des nombres premiers de la forme $qn + a$, avec a et q des entiers fixés et $n = 1, 2, 3, \ldots$ La thèse porte aussi sur la comparaison de différentes suites arithmétiques par rapport à leur comportement dans les progressions arithmétiques. Elle est divisée en quatre chapitres et contient trois articles.

Le premier chapitre est une invitation à la théorie analytique des nombres, suivie d'une revue des outils qui seront utilisés plus tard. Cette introduction comporte aussi certains résultats de recherche, que nous avons cru bon d'inclure au fil du texte.

Le deuxième chapitre contient l'article *Inequities in the Shanks-Rényi prime number race : an asymptotic formula for the densities*, qui est le fruit de recherche conjointe avec le professeur Greg Martin. Le but de cet article est d'étudier un phénomène appelé le «Biais de Chebyshev», qui s'observe dans les «courses de nombres premiers» (voir [**32**]). Chebyshev [**75**] a observé qu'il semble y avoir plus de premiers de la forme $4n + 3$ que de la forme $4n + 1$. De manière plus générale, Rubinstein et Sarnak [**68**] ont montré l'existence d'une quantité $\delta(q; a, b)$, qui désigne la probabilité d'avoir plus de premiers de la forme $qn+a$ que de la forme $qn + b$. Dans cet article nous prouvons une formule asymptotique pour $\delta(q; a, b)$ qui peut être d'un ordre de précision arbitraire (en terme de puissance négative de q). Nous présentons aussi des résultats numériques qui supportent nos formules.

Le troisième chapitre contient l'article *Residue classes containing an unexpected number of primes*. Le but est de fixer un entier $a \neq 0$ et ensuite d'étudier la répartition des premiers de la forme $qn + a$, en moyenne sur q. Nous montrons que l'entier a fixé au départ a une grande influence sur cette répartition, et qu'il existe en fait certaines progressions arithmétiques contenant moins de premiers que d'autres. Ce phénomène est plutôt surprenant, compte tenu du théorème des premiers dans les progressions arithmétiques qui stipule que les premiers sont équidistribués dans les classes d'équivalence $\bmod q$.

Le quatrième chapitre contient l'article *The influence of the first term of an arithmetic progression*. Dans cet article on s'intéresse à des irrégularités similaires à celles observées au troisième chapitre, mais pour des suites arithmétiques plus générales. En effet, nous étudions des suites telles que les entiers s'exprimant comme la somme de deux carrés, les valeurs d'une forme quadratique binaire, les k-tuplets de premiers et les entiers sans petit facteur premier. Nous démontrons que dans chacun de ces exemples, ainsi que dans une grande classe de suites arithmétiques, il existe des irrégularités dans les progressions arithmétiques $a \bmod q$, avec a fixé et en moyenne sur q.

Mots clés : Théorie analytique des nombres, nombres premiers dans les progressions arithmétiques, fonctions L de Dirichlet, zéros de fonctions L, courses de nombres premiers, applications du grand crible, suites arithmétiques.

SUMMARY

The main subject of this thesis is the distribution of primes in arithmetic progressions, that is of primes of the form $qn + a$, with a and q fixed, and $n = 1, 2, 3, \dots$ The thesis also compares different arithmetic sequences, according to their behaviour over arithmetic progressions. It is divided in four chapters and contains three articles.

The first chapter is an invitation to the subject of analytic number theory, which is followed by a review of the various number-theoretic tools to be used in the following chapters. This introduction also contains some research results, which we found adequate to include.

The second chapter consists of the article *Inequities in the Shanks-Rényi prime number race : an asymptotic formula for the densities*, which is joint work with Professor Greg Martin. The goal of this article is to study «Chebyshev's Bias», a phenomenon appearing in «prime number races» (see [**32**]). Chebyshev [**75**] was the first to observe that there tends to be more primes of the form $4n + 3$ than of the form $4n + 1$. More generally, Rubinstein and Sarnak [**68**] showed the existence of the quantity $\delta(q; a, b)$, which stands for the probability of having more primes of the form $qn + a$ than of the form $qn + b$. In this paper, we establish an asymptotic series for $\delta(q; a, b)$ which is precise to an arbitrary order of precision (in terms of negative powers of q). We also provide many numerical results supporting our formulas.

The third chapter consists of the article *Residue classes containing an unexpected number of primes*. We fix an integer $a \neq 0$ and study the distribution of the primes of the form $qn + a$, on average over q. We show that the choice of a has a significant influence on this distribution, and that some arithmetic progressions contain, on average over q, fewer primes than typical arithmetic progressions. This phenomenon is quite surprising since in light of the prime number theorem for arithmetic progressions, the primes are equidistributed in the residue classes $\bmod q$.

The fourth chapter consists of the article *The influence of the first term of an arithmetic progression*. In this article we are interested in studying more general arithmetic sequences and finding irregularities similar to those observed in chapter three. Examples of such sequences are the integers which can be written as the sum of two squares, values of binary quadratic forms, prime k-tuples and integers free of small prime factors. We show that a broad class of arithmetic sequences exhibits such irregularities over the arithmetic progressions $a \bmod q$, with a fixed and on average over q.

Key words : Analytic number theory, Primes in arithmetic progressions, Dirichlet L-functions, Zeros of L-functions, Prime number races, Applications of large sieve, Arithmetic sequences.

TABLE DES MATIÈRES

LISTE DES FIGURES

LISTE DES TABLEAUX

REMERCIEMENTS

J'aimerais tout d'abord remercier mon directeur Andrew Granville pour m'avoir guidé tout au long du doctorat. Andrew m'a appris la rigueur dans l'écriture, la présentation et la recherche en général et je lui en suis très reconnaissant. C'est aussi lui qui m'a convaincu d'étudier la théorie analytique des nombres, sujet qui me passionne toujours.

J'aimerais aussi remercier Greg Martin avec qui j'ai travaillé pendant les deux premières années et demies de mon doctorat, et qui m'a invité à Vancouver pour compléter le travail. Je remercie aussi Michel Balazard et Michael A. Tsfasman qui m'ont si gentiment accueilli au laboratoire Poncelet de Moscou lors du semestre thématique en théorie des nombres.

Je voudrais remercier John Friedlander pour s'être intéressé à mon travail et avoir soulevé des points importants lors de mon exposé au *Canadian Number Theory Association XI Meeting*. Je lui suis très reconnaissant pour le rapport détaillé et précis qu'il a accepté de produire en tant qu'examinateur externe.

J'aimerais remercier tous mes collègues au département de mathématiques et de statistique de l'Université de Montréal. Je remercie particulièrement l'équipe de théorie des nombres pour de nombreuses conversations enrichissantes desquelles j'ai beaucoup appris.

Je remercie le Conseil de Recherche en Sciences Naturelles et en Génie du Canada et la Faculté des Études Supérieures et Postdoctorales de l'Université de Montréal pour m'avoir soutenu financièrement tout au long du doctorat.

Finalement, j'offre mes plus sincères remerciements à toute ma famille qui m'a soutenue tout au cours de cette longue épreuve. Un merci spécial à ma famille rapprochée et à ma copine.

Chapitre 1

INTRODUCTION ET NOTIONS PRÉLIMINAIRES

1.1. Introduction

La théorie des nombres est un domaine qui peut sembler à la fois très élémentaire, de par l'énoncé de certains de ses problèmes fondamentaux, et extrêmement technique, de par leur démonstration. Un exemple illustrant cette idée est la conjecture de Goldbach, qui stipule que tout entier pair supérieur à 2 peut s'écrire comme la somme de deux nombres premiers. Cette conjecture n'a pas été démontrée jusqu'ici, toutefois une version plus faible de la conjecture a été démontrée en 1937 par Vinogradov.

Théorème 1.1.1 (Vinogradov)**.** *Il existe un entier N_0 tel que tout nombre impair $n > N_0$ peut être écrit comme somme de trois nombres premiers.*

Naturellement, on se demande quelle est la taille de N_0. Vinogradov n'a pas donné de constante explicite, toutefois il a récemment été démontré [**57**] qu'on peut prendre $N_0 = e^{3100}$. Un tel nombre suggère que la preuve de ce théorème doit nécessiter des outils avancés. En effet, la preuve de Vinogradov repose sur la méthode du cercle, méthode développée au début du vingtième siècle par Hardy et Littlewood. Vinogradov étant un expert des sommes exponentielles, il a fait usage de ses bornes révolutionnaires pour résoudre le problème. En jettant un coup d'oeil à la preuve, on note avec surprise la présence d'outils d'analyse complexe.

L'utilisation d'outils d'analyse complexe pour attaquer des problèmes purement arithmétiques est loin d'être unique à ce problème ; en fait une grande partie de la théorie des nombres repose sur de tels outils. Un autre exemple frappant est celui du grand théorème de Fermat qui a finalement cédé après trois siècles d'essais infructueux. Le problème porte sur les solutions entières à l'équation $x^n + y^n = z^n$, avec $n \geq 3$ et $xyz \neq 0$. Wiles a prouvé que de telles solutions non-triviales n'existaient pas, et sa preuve repose sur une multitude d'outils de théorie des nombres algébrique et analytique, d'analyse complexe, de théorie des représentations, etc. Un des points cruciaux de la preuve est de démontrer que toute courbe elliptique rationnelle est modulaire, ce qui donne un prolongement analytique à une certaine «fonction zêta». Les «fonctions zêta», communément appelées fonctions L, apparaissent naturellement dans une grande quantité de problèmes en théorie des nombres, et nous allons nous consacrer à leur étude tout au long de la prochaine section.

1.2. Notes historiques sur la fonction zêta de Riemann

La fonction

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}, \tag{1.2.1}$$

appelée «fonction zêta de Riemann», est centrale en théorie des nombres. Le premier mathématicien à l'étudier a été Euler. Un des problème célèbres de l'époque était le problème de Bâle, qui consistait à calculer la valeur de

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \dots$$

Le problème, posé en 1644, a été résolu par Euler en 1735, qui montra que cette série infinie égale $\pi^2/6$. Euler montra plus généralement que pour tout entier $k \geq 1$,

$$\zeta(2k) = \sum_{n=1}^{\infty} \frac{1}{n^{2k}} = \frac{|B_{2k}|(2\pi)^{2k}}{2 \cdot (2k)!}, \tag{1.2.2}$$

où B_n désigne le n-ième nombre de Bernoulli[1]. Euler montra aussi l'identité

$$\zeta(s) = \prod_{p \text{ premier}} \left(1 - \frac{1}{p^s}\right)^{-1},$$

appelée «produit eulérien», qui s'avérera plus tard fondamentale dans l'étude de la distribution des nombres premiers. Euler conclut rapidement une nouvelle preuve qu'il y a une infinité de nombres premiers, mais plus encore, il prouva avec cette identité que la série

$$\sum_{p \text{ premier}} \frac{1}{p}$$

diverge. Un autre calcul étonnant d'Euler était celui de la valeur de la fonction zêta aux entiers négatifs :

$$\zeta(1-k) = -\frac{B_k}{k}, \tag{1.2.3}$$

pour $k \geq 1$ (et donc $\zeta(-2n) = 0$ pour $n \geq 1$). Dans un article datant de 1768 **[21]**, Euler utilisa sa formule de sommation pour créer un lien formidable entre (1.2.2) et (1.2.3). Il montra la formule suivante, pour $n = 2, 3, 4, \ldots$:

$$\frac{1 - 2^{n-1} + 3^{n-1} - 4^{n-1} + 5^{n-1} - 6^{n-1} + \ldots}{1 - 2^{-n} + 3^{-n} - 4^{-n} + 5^{-n} - 6^{-n} + \ldots} = -\frac{(n-1)!(2^n-1)}{(2^{n-1}-1)\pi^n} \cos\left(\frac{n\pi}{2}\right), \tag{1.2.4}$$

et conjectura que cette formule reste valide quel que soit n réel. Pour supporter sa conjecture, il en donna la preuve dans les cas particuliers $n = \frac{1}{2}$ et $n = \frac{3}{2}$, pour lesquels il utilisa sa généralisation des factorielles, la fonction gamma. Non seulement[2] il avait une formule pour $\zeta(1-k)$, mais il venait de découvrir l'équation fonctionnelle de $\zeta(s)$! « Ou bien je ferai voir, qu'en connaissant la somme de la première série pour un exposant quelconque m, on en peut toujours déterminer la somme de l'autre série pour l'exposant $n = m + 1$.» Pour justifier ses manipulations avec les séries divergentes, Euler écrit : «Mais j'ai déjà remarqué dans une autre occasion, qu'il faut donner au mot de *somme* une

[1] Les nombres de Bernoulli B_n sont définis par la fonctions génératrice suivante :

$$\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} B_n \frac{x^n}{n!}.$$

[2] On peut facilement passer de la somme alternée $1 - 2^{-n} + 3^{-n} - 4^{-n} \ldots$ à la somme originale $1 + 2^{-n} + 3^{-n} + 4^{-n} \ldots$ en multipliant par $\left(1 - 2^{1-n}\right)^{-1}$.

signification plus étendue, et entendre par là une fraction, ou autre expression analytique, laquelle étant développée selon les principes de l'analyse produise la même série dont on cherche la somme. » Le concept sous-jacent de prolongement analytique n'allait toutefois pas être disponible avant le développement rigoureux de l'analyse complexe.

Le prochain acteur principal dans l'histoire de la fonction zêta est Riemann, qui eut l'idée révolutionnaire d'étudier $\zeta(s)$ avec un argument complexe s. Cette idée qui semble farfelue est en fait extrêmement puissante dans l'étude de la distribution des nombres premiers. Le triomphe de Riemann est sans doute sa formule explicite, qui montre que l'étude des nombres premiers est équivalente à l'étude des zéros complexes de $\zeta(s)$. Riemann montra tout d'abord que la définition (1.2.1), valide pour $\mathrm{Re}(s) > 1$, pouvait être étendue à tout le plan complexe, à l'exception d'un pôle simple en $s = 1$. Il démontra l'équation fonctionnelle conjecturée par Euler (comparer avec (1.2.4)) :

$$\zeta(s) = 2^s \pi^{s-1} \sin\left(\frac{\pi s}{2}\right) \Gamma(1-s)\zeta(1-s),$$

où $\Gamma(s)$ est la fonction gamma d'Euler (voir le chapitre II.0 de **[76]**). Une version plus moderne de la formule explicite de Riemann est la suivante :

$$\psi(x) = x - \sum_{\rho : \zeta(\rho)=0} \frac{x^\rho}{\rho} - \log(2\pi),$$

où $\psi(x) := \sum_{\substack{p^k \leq x \\ k \geq 1}} \log p$ est la quantité (pondérée) de puissances de nombres premiers dans l'intervalle $[1, x]$. Après quelques calculs numériques, Riemann s'aperçut que les seuls zéros ρ dans la région $0 \leq \mathrm{Re}(\rho) \leq 1$, appelés zéros non-triviaux, se situaient sur la droite $\mathrm{Re}\,\rho = \frac{1}{2}$, et il déduisit de cette hypothèse l'estimation miraculeuse

$$|\psi(x) - x| \leq C\sqrt{x}\log^2 x,$$

avec $C > 0$ une constante. L'hypothèse de Riemann n'est toujours pas résolue, et elle est considérée par certains comme le plus important problème ouvert des mathématiques. Les mathématiciens après Riemann ont toutefois pu démontrer des résultats plus faibles, qui ont été suffisants pour achever la démonstration du théorème des nombres premiers. Ce théorème, conjecturé par

Gauss en 1792, stipule que

$$\lim_{x \to \infty} \frac{\psi(x)}{x} = 1.$$

Le théorème a finalement été démontré indépendamment par Hadamard et de la Vallée Poussin en 1896. Les deux mathématiciens ont achevé la preuve en montrant que $\zeta(s)$ ne possédait pas de zéros sur la droite $\mathrm{Re}(s) = 1$.

1.3. Notions d'analyse

1.3.1. Formules de Perron et transformée de Mellin

Dans cette section nous allons voir comment déduire la formule explicite de Riemann. Mentionnons tout d'abord la formule de Perron, qui est un outil analytique permettant d'étudier les sommes tronquées.

Proposition 1.3.1 (Formule de Perron)**.**

$$\frac{1}{2\pi i} \int_{\mathrm{Re}(s)=2} \left(\frac{x}{n}\right)^s \frac{ds}{s} = \begin{cases} 1 & \textit{si}\ n < x \\ \frac{1}{2} & \textit{si}\ n = x \\ 0 & \textit{si}\ n > x. \end{cases}$$

Voyons maintenant comment on peut déduire la formule explicite. Notons tout d'abord que si on prend la dérivée logarithmique du produit eulérien de $\zeta(s)$, on obtient la formule suivante :

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{\substack{p \\ k \geq 1}} \frac{\log p}{p^{ks}}.$$

Donc en utilisant la formule de Perron, on a pour $x \notin \mathbb{N}$ que

$$\begin{aligned} \psi(x) = \sum_{\substack{p^k \leq x \\ k \geq 1}} \log p &= \sum_{\substack{p \\ k \geq 1}} \log p \frac{1}{2\pi i} \int_{\mathrm{Re}(s)=2} \left(\frac{x}{p^k}\right)^s \frac{ds}{s} \\ &= \frac{1}{2\pi i} \int_{\mathrm{Re}(s)=2} \left(\sum_{\substack{p \\ k \geq 1}} \frac{\log p}{p^{ks}} \right) x^s \frac{ds}{s} \\ &= \frac{1}{2\pi i} \int_{\mathrm{Re}(s)=2} \left(-\frac{\zeta'(s)}{\zeta(s)} \right) x^s \frac{ds}{s}. \qquad (1.3.1) \end{aligned}$$

Dans la prochaine étape nous allons utiliser le théorème des résidus. Ce théorème affirme que l'intégrale sur un parcours fermé d'une fonction méromorphe est égal à $2\pi i$ fois la somme des résidus à l'intérieur du parcours. Nous sommes en présence d'un parcours infini, donc le théorème ne s'applique pas directement, mais les détails techniques requis pour rendre l'argument rigoureux nuiraient à l'exposition. Ainsi on déplace le parcours $\mathrm{Re}(s) = 2$ vers la gauche jusqu'à la droite $\mathrm{Re}(s) = -1$. Les pôles de la dérivée logarithmique $\frac{\zeta'(s)}{\zeta(s)}$, tous simples, se trouvent exactement aux emplacements où $\zeta(s)$ elle-même possède un zéro ou un pôle. On peut aussi montrer que le résidu à ces pôles vaut exactement l'ordre de $\zeta(s)$. La fonction $\zeta(s)$ possède un pôle, en $s = 1$, qui est simple. Nous avons déjà vu que $\zeta(-2n) = 0$ pour $n \geq 1$; ce sont les zéros triviaux de $\zeta(s)$. En fait, aucun autre zéro n'existe hors de la bande critique $0 \leq \mathrm{Re}(s) \leq 1$. Ainsi (1.3.1) devient

$$\psi(x) = x - \frac{\zeta'(0)}{\zeta(0)} - \sum_{\substack{\rho:\zeta(\rho)=0 \\ 0\leq \mathrm{Re}(\rho)\leq 1}} \frac{x^\rho}{\rho} + \frac{1}{2\pi i}\int_{\mathrm{Re}(s)=-1} \left(-\frac{\zeta'(s)}{\zeta(s)}\right) x^s \frac{ds}{s}.$$

Ici, le terme x provient du pôle simple de $\frac{\zeta'(s)}{\zeta(s)}$ en $s = 1$, le terme $-\frac{\zeta'(0)}{\zeta(0)}$ provient du pôle simple de $\frac{x^s}{s}$ en $s = 0$, et les termes $-\frac{x^\rho}{\rho}$ proviennent des pôles simples de $\frac{\zeta'(s)}{\zeta(s)}$ aux zéros de $\zeta(s)$ (noter que dans cette somme les zéros sont comptés avec multiplicité). La dernière étape est de déplacer le parcours d'intégration encore plus loin vers la gauche, et d'utiliser les bornes sur $\frac{\zeta'(s)}{\zeta(s)}$ pour remarquer que l'intégrale tend vers zéro à mesure que le parcours se déplace vers la gauche. On obtient à la limite la formule explicite :

$$\psi(x) = x - \frac{\zeta'(0)}{\zeta(0)} - \sum_{\rho:\zeta(\rho)=0} \frac{x^\rho}{\rho}. \tag{1.3.2}$$

Il existe plusieurs variantes de la formule de Perron. Pour chacune, nous donnerons au moins un exemple d'application.

Proposition 1.3.2. *Pour* $k \geq 1$ *un entier,*

$$\frac{1}{2\pi i}\int_{\mathrm{Re}(s)=2} \left(\frac{x}{n}\right)^s \frac{ds}{s(s+1)\cdots(s+k)} = \begin{cases} \frac{1}{k!}\left(1-\frac{n}{x}\right)^k & \textit{si } n \leq x \\ 0 & \textit{si } n > x. \end{cases}$$

(Voir le lemme 3.6.8.)

Proposition 1.3.3. *Pour* $k \geq 2$ *un entier,*

$$\frac{1}{2\pi i}\int_{\mathrm{Re}(s)=2}\left(\frac{x}{n}\right)^s \frac{ds}{s^k} = \begin{cases} \frac{1}{(k-1)!}\log(x/n)^{k-1} & \textit{si } n \leq x \\ 0 & \textit{si } n > x. \end{cases}$$

(Voir (1.6.3).*)*

Proposition 1.3.4.

$$\frac{1}{2\pi i}\int_{\mathrm{Re}(s)=2}\left(\frac{x}{n}\right)^s \Gamma(s)ds = e^{-n/x}.$$

(Voir (1.6.4)*,* (1.6.5) *ou* (2.5.8).*)*

L'avantage de ces formules est que l'intégrale converge absolument, ce qui simplifie beaucoup les calculs en pratique. Les trois formules sont des exemples d'application de la transformée de Mellin (avec $y = n/x$).

Théorème 1.3.1 (Inversion de Mellin)**.** *Définissons pour* $\phi : \mathbb{R} \to \mathbb{R}$ *continue la transformée de Mellin*

$$\mathcal{M}\phi(s) := \int_0^\infty y^{s-1}\phi(y)dy.$$

Définissons aussi pour une fonction complexe $f(s)$*, holomorphe dans la bande* $a < \mathrm{Re}(s) < b$*, la transformée inverse de Mellin :*

$$\mathcal{M}^{-1}f(y) := \frac{1}{2\pi i}\int_{\mathrm{Re}(s)=c} y^{-s}f(s)ds$$

avec $a < c < b$*. Alors sous certaines conditions de convergence on a*

$$\mathcal{M}\phi = f \Leftrightarrow \mathcal{M}^{-1}f = \phi.$$

À l'aide de ce théorème on peut montrer sous certaines conditions l'identité générale

$$\sum_n a_n\phi(n) = \frac{1}{2\pi i}\int_{\mathrm{Re}(s)=c}\left(\sum_n \frac{a_n}{n^s}\right)\mathcal{M}\phi(s)ds.$$

Donnons une dernière variante de Perron, avec sa preuve.

Proposition 1.3.5.

$$\frac{1}{2\pi i}\int_{\mathrm{Re}(s)=2}\left(\frac{x}{n}\right)^s \frac{ds}{s(s-1)^{1/2}} = \begin{cases} \frac{1}{\sqrt{\pi}}\int_1^{x/n}\frac{dt}{\sqrt{\log t}} & \textit{si } n \leq x \\ 0 & \textit{si } n > x. \end{cases} \tag{1.3.3}$$

(Voir (1.8.2).*)*

Noter[3] que pour y proche de 1, $\int_1^y \frac{dt}{\sqrt{\log t}} = 2(y-1)^{1/2} + O((y-1)^{3/2})$.

DÉMONSTRATION. Par le théorème 1.3.1, il suffit de prouver que pour $\text{Re}(s) > 1$,

$$\frac{\sqrt{\pi}}{s(s-1)^{1/2}} = \int_1^\infty u^{-s-1} \int_1^u \frac{dt}{\sqrt{\log t}} du. \tag{1.3.4}$$

Soit $s > 1$ on nombre réel. On a

$$\int_1^\infty u^{-s-1} \int_1^u \frac{dt}{\sqrt{\log t}} du = \int_1^\infty \frac{1}{\sqrt{\log t}} \int_t^\infty u^{-s-1} du dt = \frac{1}{s} \int_1^\infty \frac{t^{-s} dt}{\sqrt{\log t}}$$
$$= \frac{1}{s(s-1)^{1/2}} \int_0^\infty w^{-1/2} e^{-w} dw = \frac{\sqrt{\pi}}{s(s-1)^{1/2}},$$

d'où (1.3.4) découle par l'unicité du prolongement analytique. □

1.3.2. Densités logarithmiques

Pour un ensemble $\mathcal{A} \subset \mathbb{N}$ on peut s'intéresser à la densité naturelle définie par

$$\delta_0(A) := \lim_{N\to\infty} \frac{\#(\mathcal{A} \cap [1, N])}{N} = \lim_{N\to\infty} \frac{1}{N} \sum_{\substack{n\in\mathcal{A} \\ n\leq N}} 1,$$

si cette limite existe. Dans certains problèmes toutefois, il est avantageux d'ajouter des poids et de considérer plutôt la densité logarithmique :

$$\delta_1(A) := \lim_{N\to\infty} \frac{1}{\log N} \sum_{\substack{n\in\mathcal{A} \\ n\leq N}} \frac{1}{n}.$$

Proposition 1.3.6. *Supposons que* $\delta_0(A)$ *existe. Alors* $\delta_1(A)$ *existe aussi et* $\delta_1(A) = \delta_0(A)$. *(Toutefois, l'inverse n'est pas nécessairement vrai.)*

PREUVE. Supposons que $\delta_0(A) = \delta$ existe. Alors[4]

$$S(t) := \sum_{\substack{n\in\mathcal{A} \\ n\leq N}} 1 = \delta N + o(N).$$

[3] On dit que $A = O(B)$ s'il existe une constante $c > 0$ telle que $|A| \leq cB$.

[4] On dit que $f(N) = o(g(N))$ si $\lim_{N\to\infty} \frac{f(N)}{g(N)} = 0$.

Ainsi,

$$\begin{aligned}\sum_{\substack{n\in\mathcal{A}\\ n\leq N}}\frac{1}{n}&=\int_{1^-}^{N}\frac{dS(t)}{t}=\frac{S(N)}{N}+\int_{1^-}^{N}\frac{S(t)}{t^2}dt\\&=O(1)+\int_{1^-}^{N}\frac{\delta}{t}dt+\int_{1^-}^{N}\frac{S(t)-\delta t}{t^2}dt\\&=\delta\log N+O\left(\int_{1}^{\log N}\frac{1}{t}+\int_{\log N}^{N}\frac{S(t)-\delta t}{t^2}dt\right)\\&=\delta\log N+o(\log N).\end{aligned}$$

□

L'avantage d'utiliser la densité logarithmique est qu'elle existe dans certains problèmes où la densité naturelle n'existe pas.

On peut aussi itérer le processus en définissant

$$\delta_k(A):=\lim_{N\to\infty}\frac{1}{\log_k N}\sum_{\substack{n\in\mathcal{A}\\ n\leq N}}\frac{1}{n\log n\log\log n\cdots\log_{k-1}n},$$

où $\log_k n=\log\log\cdots\log n$ désigne la k-ième itérée de la fonction log. Il suit que si $\delta_k(A)$ existe, alors $\delta_l(A)$ existe pour chaque $l>k$, et $\delta_l(A)=\delta_k(A)$.

D'autre part, on serait tenté d'étudier des densités de la forme

$$\delta_\nu(A):=\lim_{N\to\infty}\frac{1}{N^{1-\nu}/(1-\nu)}\sum_{\substack{n\in\mathcal{A}\\ n\leq N}}\frac{1}{n^\nu},$$

avec $0<\nu<1$. Il est toutefois inutile de s'intéresser à de telles densités, car $\delta_\nu(A)$ existe ssi $\delta_0(A)$ existe, et dans ce cas, $\delta_\nu(A)=\delta_0(A)$.

1.4. Notions de probabilités

1.4.1. Espérance et variance

Considérons une variable aléatoire continue X ayant comme fonction de densité $f_X(x)$. L'espérance de X est donnée par

$$\mathbb{E}(X):=\int_{\mathbb{R}}xf_X(x)dx.$$

L'espérance est additive, c'est-à-dire que pour des variables aléatoires X, Y on a $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$. Si X et Y sont indépendantes, on a en plus que $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. La variance est définie par

$$\mathrm{Var}(X) := \mathbb{E}((X - \mathbb{E}(x))^2) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2.$$

1.4.2. Moments et cumulants

On définit le n-ième moment de X par

$$m_n(X) := \int_{\mathbb{R}} x^n f_X(x) dx = \mathbb{E}(X^n).$$

On définit aussi le n-ième moment centré par

$$M_n(X) := \mathbb{E}((X - \mathbb{E}(X))^n) = m_n(X - \mathbb{E}(X)).$$

La connaissance de tous les moments détermine la distribution de X. Si $X \sim N(0,1)$ est distribuée selon la loi normale centrée réduite, c'est-à-dire que $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, alors on a

$$m_n(X) = M_n(X) = \begin{cases} (n-1)\cdot(n-3)\cdots 3\cdot 1 & \text{si } n \text{ est pair} \\ 0 & \text{si } n \text{ est impair.} \end{cases}$$

Nous savons que l'espérance est additive. Aussi, si X et Y sont indépendantes, alors on a $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$. Ce n'est toutefois pas le cas des moments supérieurs. Pour remédier à cette situation, nous allons définir les cumulants, qui eux aussi déterminent la distribution de X.

Definition 1.4.1. *Les cumulants $\kappa_n(X)$ sont définis par la fonction génératrice*

$$g_X(t) := \log \mathbb{E}(e^{itX}) = \sum_{n=1}^{\infty} \frac{\kappa_n(X) i^n t^n}{n!}.$$

La transformée de Fourier

$$\hat{X}(t) := \mathbb{E}(e^{itX}) = \int_{\mathbb{R}} e^{itx} f_X(x) dx,$$

aussi appelée fonction caractéristique de X, est parfois très utile pour comprendre la distribution de X. Comme $g_X(t) = \log \hat{X}(t)$, les cumulants sont en réalité les coefficients du logarithme de la fonction caractéristique (divisés par

i^n). Ils ont le grand avantage d'avoir la propriété d'additivité suivante : si les variables aléatoires X et Y sont indépendantes, alors

$$\kappa_n(X+Y) = \kappa_n(X) + \kappa_n(Y).$$

Ce fait découle directement de la définition des cumulants et de la propriété suivante : pour X, Y indépendantes, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. Les cumulants de la loi normale $X \sim N(\mu, \sigma^2)$ sont donnés par

$$\kappa_1(X) = \mu, \qquad \kappa_2(X) = \frac{\sigma^2}{2}, \qquad 0 = \kappa_3(X) = \kappa_4(X) = \kappa_5(X) = \dots$$

1.5. Notions de théorie des nombres analytique

1.5.1. Caractères multiplicatifs et fonctions L de Dirichlet

Soit $q \geq 1$ un entier. Un caractère de Dirichlet $\bmod q$ est une fonction complètement multiplicative (c'est-à-dire un homomorphisme)

$$\chi : (\mathbb{Z}/q\mathbb{Z})^\times \longrightarrow \mathbb{C} \setminus \{0\},$$

c'est-à-dire que $\chi(mn) = \chi(m)\chi(n)$. La propriété d'être complètement multiplicatif s'avérera fondamentale plus tard pour donner un produit eulérien à la fonction zêta associée à χ. On montre facilement que $\chi(1) = 1$. Le caractère $\chi_0 \equiv 1$ est appelé caractère principal. Définissons la fonction d'Euler

$$\phi(q) := q \prod_{p|q} \left(1 - \frac{1}{p}\right). \tag{1.5.1}$$

Le théorème d'Euler stipule que peu importe l'entier a tel que $(a, q) = 1$, on a $a^{\phi(q)} \equiv 1 \bmod q$. En utilisant ce théorème, on montre facilement que $\chi(n)^{\phi(q)} = 1$, donc $\chi(n)$ est une racine $\phi(q)$-ième de l'unité. Cette observation nous permet de déduire que l'ensemble de tous caractères $\chi \bmod q$ forme un groupe d'ordre $\phi(q)$. On peut étendre la définition de χ à tout $\mathbb{Z}$ en posant

$$\chi(n) := \begin{cases} \chi(n \bmod q) & \text{si } (n, q) = 1 \\ 0 & \text{si } (n, q) > 1. \end{cases}$$

Le caractère χ est donc périodique de période q, mais cette période n'est pas nécessairement minimale, car il pourrait s'avérer que pour un certain entier

$q_0 \mid q$ on ait $\chi(n+q_0) = \chi(q_0)$, pour tout n avec $(n, q) = 1$. La période minimale (qui est clairement unique), aussi appelée conducteur de χ, est notée q^*. Si $q^* = q$, alors χ est dit primitif. On dit qu'un caractère $\chi_1 \bmod d$ induit un autre caractère $\chi \bmod q$ si $d \mid q$ et $\chi(n) = \chi_1(n)$ pour $(n, q) = 1$. On peut prouver que si $\chi \bmod q$ n'est pas primitif, alors il existe un unique caractère $\chi^* \bmod q^*$ qui induit χ.

L'utilité des caractères réside dans leur habilité d'isoler les différentes progressions arithmétiques $\bmod q$. En effet, nous avons les relations d'orthogonalités suivantes :

$$\frac{1}{\phi(q)} \sum_{\chi \bmod q} \chi(a) = \begin{cases} 1 & \text{si } a \equiv 1 \bmod q \\ 0 & \text{sinon,} \end{cases}$$

$$\frac{1}{\phi(q)} \sum_{a \bmod q} \chi(a) = \begin{cases} 1 & \text{si } \chi = \chi_0 \\ 0 & \text{sinon.} \end{cases}$$

Ainsi, pour calculer la quantité de nombres premiers $\equiv a \bmod q$ dans l'intervalle $[1, x]$, on prend

$$\begin{aligned} \psi(x; q, a) := \sum_{\substack{p^e \leq x \\ p^e \equiv a \bmod q}} \log p &= \sum_{p^e \leq x} \log p \frac{1}{\phi(q)} \sum_{\chi \bmod q} \chi(p^e a^{-1}) \\ &= \frac{1}{\phi(q)} \sum_{\chi \bmod q} \chi(a^{-1}) \sum_{p^e \leq x} \chi(p^e) \log p, \end{aligned}$$

et donc il suffit d'étudier la quantité

$$\psi(x, \chi) := \sum_{p^e \leq x} \chi(p^e) \log p.$$

De manière analogue à $\psi(x)$, il existe une formule explicite pour $\psi(x, \chi)$ en terme des zéros complexes d'une fonction zêta. Cette fonction est

$$L(s, \chi) := \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}.$$

Par multiplicativité des caractères de Dirichlet, on obtient le produit eulérien

$$L(s, \chi) = \prod_p \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}.$$

La formule explicite en question est, pour $\chi \neq \chi_0$ tel que $\chi(-1) = -1$,

$$\psi(x,\chi) = -\sum_{\rho_\chi : L(\rho_\chi,\chi)=0} \frac{x^{\rho_\chi}}{\rho_\chi} - \frac{L'(0,\chi)}{L(0,\chi)}.$$

Si $\chi(-1) = 1$, alors $L(0,\chi) = 0$ et la formule devient

$$\psi(x,\chi) = -\sum_{\substack{\rho_\chi : L(\rho_\chi,\chi)=0,\\ \rho_\chi \neq 0}} \frac{x^{\rho_\chi}}{\rho_\chi} - \log x - b(\chi),$$

où $b(\chi)$ est le résidu en $s = 0$ de $\frac{L'(s,\chi)}{sL(s,\chi)}$. En utilisant ces formules, on arrive à prouver le théorème des nombres premiers dans les progressions arithmétiques.

Théorème 1.5.1. *Soit* a, q *des entiers fixés tels que* $(a,q) = 1$. *On a l'asymptotique*

$$\psi(x;q,a) \sim \frac{x}{\phi(q)}. \tag{1.5.2}$$

L'asymptotique (1.5.2) est valide pour des valeurs fixées de a et q, toutefois il est naturel de se demander si elle reste vraie uniformément, c'est-à-dire quand a et q varient avec x. Une réponse affirmative existe dans un intervalle limité pour q.

Théorème 1.5.2 (Siegel-Walfisz). *Soit* $A > 0$ *un nombre réel fixé. Il existe une constante positive* $c = c(A)$ *telle que, uniformément pour* $1 \leq a < q \leq (\log x)^A$ *avec* $(a,q) = 1$,

$$\psi(x;q,a) = \frac{x}{\phi(q)} + O\left(\frac{x}{e^{c\sqrt{\log x}}}\right). \tag{1.5.3}$$

Ici, $O(f(x))$ désigne une quantité bornée en valeur absolue par $Cf(x)$, avec $C > 0$ une constante réelle. La borne sur q dans le théorème 1.5.2 provient de la localisation des zéros des fonctions $L(s,\chi)$. Si l'hypothèse de Riemann généralisée est vraie, c'est-à-dire que tous les zéros des fonctions $L(s,\chi)$ dans la région $0 < \operatorname{Re} s < 1$ sont sur la droite $\operatorname{Re} s = \frac{1}{2}$, alors l'asymptotique (1.5.2) est valide pour $q \leq x^{1/2-\epsilon}$. Comme pour la fonction $\zeta(s)$, il est possible de trouver une région libre de zéros de $L(s,\chi)$ proche de la droite $\operatorname{Re}(s) = 1$, à au plus une exception près. Cette exception, si elle existe, est un zéro réel appelé zéro de Siegel, et force le caractère χ à être réel.

Proposition 1.5.1. *Il existe une constante* C *telle* $L(s,\chi)$ *possède au plus un seul zéro dans la région*

$$\mathrm{Re}(s) \geq 1 - \frac{C}{\log(q(|t|+1))}.$$

Si un tel zéro existe, il est réel et χ *est un caractère réel.*

Définition 1.5.1. *Soit* χ *un caractère réel* $\bmod q$. *Un zéro réel* β_χ *de la fonction* $L(s,\chi)$ *est dit de Siegel par rapport à la constante* C *si l'inégalité suivante est respectée :*

$$\beta_\chi > 1 - \frac{C}{\log q}.$$

L'existence de zéros proche de $s = 1$ distortionnerait grandement la distribution des nombres premiers dans les progressions arithmétiques, donc un grand effort a été fait pour exclure une telle situation. Le meilleur résultat en ce sens est celui de Siegel.

Théorème 1.5.3 (Siegel)**.** *Si un zéro de Siegel* β_χ *existe, alors pour chaque* $\epsilon > 0$ *il existe une constante* C_ϵ *telle que*

$$\beta_\chi < 1 - \frac{C_\epsilon}{q^\epsilon}.$$

Le théorème de Siegel est un exemple de résultat «ineffectif», en ce sens qu'il est impossible de calculer la constante C_ϵ, de par la nature même de la preuve. Le théorème 1.5.2 est aussi ineffectif, car sa preuve repose sur le théorème de Siegel.

1.5.2. Méthode de Selberg-Delange

Soit $f(n) : \mathbb{N} \to \mathbb{C}$ une fonction multiplicative, c'est-à-dire que $f(mn) = f(m)f(n)$ pour tous les entiers m, n tels que $(m,n) = 1$. Il est souvent nécessaire de donner des bonnes estimations pour la somme

$$\sum_{n \leq x} f(n),$$

ou une version pondérée de celle-ci. Comme nous avons vu à la section 1.3.1, nous pouvons réécrire cette somme sous la forme d'une intégrale complexe à l'aide de la formule de Perron :

$$\sum_{n\leq x} f(n) = \frac{1}{2\pi i}\int_{\mathrm{Re}(s)=2}\left(\sum_{n=1}^{\infty}\frac{f(n)}{n^s}\right)x^s\frac{ds}{s}.$$

Il est donc primordial d'étudier les propriétés analytiques de la fonction

$$Z_f(s) := \sum_{n=1}^{\infty}\frac{f(n)}{n^s}.$$

Dans le cas de la distribution des nombres premiers, nous avons pu utiliser le prolongement analytique de $\frac{\zeta'(s)}{\zeta(s)}$, valide dans tout le plan complexe, et les bornes sur cette fonction nous ont permis de déplacer le parcours d'intégration arbitrairement loin vers la gauche. Nous n'avons pas toujours cette chance dans d'autres problèmes, et notre but sera déplacer le parcours d'intégration le plus loin possible.

Dans le cas où $Z_f(s)$ ressemble beaucoup à une puissance complexe z de $\zeta(s)$, la méthode de Selberg-Delange permet de donner un prolongement analytique à $Z_f(s)$ dans un domaine limité. La méthode générale est d'étudier les propriétés du produit $Z_f(s)\zeta(s)^{-z}$, qui est souvent défini dans un plus grand domaine que $Z_f(s)$.

Commençons par l'exemple particulier $f(n) := \frac{1}{\phi(n)}$, qui sera fondamental au chapitre 3. En utilisant le développement en produit eulérien, on obtient

$$\begin{aligned}
Z_f(s) &= \sum_{n=1}^{\infty}\frac{1}{\phi(n)n^s}\\
&= \prod_p\left(1+\frac{1}{(p-1)p^s}+\frac{1}{(p-1)p^{2s+1}}+\frac{1}{(p-1)p^{3s+2}}+\dots\right)\\
&= \prod_p\left(1+\frac{1}{\left(1-\frac{1}{p}\right)p^{s+1}}\left(1-\frac{1}{p^{s+1}}\right)^{-1}\right)\\
&= \prod_p\left(1-\frac{1}{p^{s+1}}\right)^{-1}\left(1-\frac{1}{p^{s+1}}+\frac{1}{\left(1-\frac{1}{p}\right)p^{s+1}}\right)\\
&= \zeta(s+1)\prod_p\left(1+\frac{1}{(p-1)p^{s+1}}\right).
\end{aligned}$$

Noter que la fonction $Z_1(s) := \prod_p \left(1 + \frac{1}{(p-1)p^{s+1}}\right)$ converge pour $\mathrm{Re}(s) > -1$. Montrons comment un peut donner un prolongement analytique à $Z_f(s)$ dans la région $\mathrm{Re}(s) > -\frac{3}{2}$. L'idée est basée sur le fait que $Z_1(s)$ ressemble beaucoup à $\zeta(s+2)$, donc pour en détecter la différence on définit

$$\begin{aligned} Z_2(s) &:= \zeta(s+2)^{-1} Z_1(s) \\ &= \prod_p \left(1 - \frac{1}{p^{s+2}}\right)\left(1 + \frac{1}{(p-1)p^{s+1}}\right) \\ &= \prod_p \left(1 + \frac{1}{(p-1)p^{s+2}} - \frac{1}{(p-1)p^{2s+3}}\right), \end{aligned}$$

et comme ce produit converge pour $\mathrm{Re}(s) > -\frac{3}{2}$, on obtient que

$$Z_f(s) = \zeta(s+1)\zeta(s+2)Z_2(s)$$

est définie dans cette même région. Montrons maintenant comment utiliser cette technique pour obtenir une estimation précise de la somme pondérée suivante :

$$\sum_{n \le x} \frac{1}{\phi(n)}\left(1 - \frac{n}{x}\right) = \frac{1}{2\pi i}\int_{\mathrm{Re}(s)=2} Z_f(s)\frac{x^s ds}{s(s+1)} \tag{1.5.4}$$

(par la proposition 1.3.2). La fonction $Z_f(s)$ possède un pôle double en $s = 0$, donc en déplaçant le parcours d'intégration vers la gauche et en calculant le résidu à ce point on obtient que (1.5.4) est

$$= C_1 \log x + C_3 + \frac{1}{2\pi i}\int_{\mathrm{Re}(s)=-\frac{1}{2}} Z_f(s)x^s \frac{ds}{s(s+1)},$$

où

$$C_1 := \frac{\zeta(2)\zeta(3)}{\zeta(6)}, \qquad C_3 := \frac{\zeta(2)\zeta(3)}{\zeta(6)}\left(\gamma - 1 - \sum_p \frac{\log p}{p^2 - p + 1}\right).$$

Toutefois, le prolongement analytique que nous avons obtenu nous permet d'aller encore plus loin vers la gauche et d'obtenir que (1.5.4) est

$$= C_1 \log x + C_3 + C_4 \frac{\log x}{x} + \frac{C_5}{x} + \frac{1}{2\pi i}\int_{\mathrm{Re}(s)=-\frac{11}{8}+\epsilon} Z_f(s)x^s \frac{ds}{s(s+1)},$$

où

$$C_4 := \frac{1}{2}, \qquad C_5 := \frac{1}{2}\left(\log 2\pi + \gamma + \sum_p \frac{\log p}{p(p-1)} + 1\right).$$

Nous pouvons utiliser les bornes de croissance verticale de $\zeta(s)$ pour montrer que cette dernière intégrale est $\ll_\epsilon x^{-\frac{11}{8}+\epsilon}$, donc finalement on obtient l'estimation

$$\sum_{n\leq x}\frac{1}{\phi(n)}\left(1-\frac{n}{x}\right)=C_1\log x+C_2+C_3\frac{\log x}{x}+\frac{C_4}{x}+O_\epsilon\left(\frac{1}{x^{\frac{11}{8}+\epsilon}}\right). \qquad (1.5.5)$$

L'exposant $\frac{11}{8}$ peut être amélioré à $\frac{743}{538}$ en utilisant un résultat de sous-convexité de la fonction $\zeta(s)$ [**41**]. Pour plus de détails, voir le lemme 3.6.8.

Un autre exemple important est la fonction $f_z(n):=z^{\omega(n)}$, où z est un nombre complexe et $\omega(n)$ dénote le nombre de diviseurs premiers de n. C'est pour cet exemple particulier que Selberg inventa la méthode [**71**]. Il réussit à simplifier considérablement la preuve d'un grand résultat de Sathe ([**69**],[**70**]), qui consiste à donner une asymptotique pour $\pi_k(x)$ qui est uniforme pour $k\leq c\log\log x$, avec $c>1$. Ici, $\pi_k(x)$ désigne le nombre d'entiers inférieurs à x possédant exactement k facteurs premiers. Le résultat de Sathe est que si $z:=(k-1)/\log\log x$, alors nous avons uniformément pour $z\leq 2-\epsilon$ que

$$\pi_k(x)\sim C_z\frac{x}{\log x}\frac{(\log\log x)^{k-1}}{(k-1)!},$$

où

$$C_z:=\frac{1}{\Gamma(z+1)}\prod_p\left(1+\frac{z}{p-1}\right)\left(1-\frac{1}{p}\right)^z.$$

Pour une version générale de la méthode de Selberg-Delange, voir [**76**].

1.6. COURSES DE NOMBRES PREMIERS

Pour une très belle vulgarisation du sujet, référons le lecteur à l'article de Granville et Martin [**32**].

1.6.1. Changements de signe

Nous avons vu que l'hypothèse de Riemann, qui affirme que tous les zéros non-triviaux de $\zeta(s)$ sont situés sur la droite $\mathrm{Re}(s)=\frac{1}{2}$, implique[5]

$$|\psi(x)-x|\ll x^{\frac{1}{2}}(\log x)^2.$$

[5] Nous utilisons la notation de Vinogradov : $f(x)\ll g(x)\Leftrightarrow|f(x)|\leq Cg(x)$, où C est une constante, et similairement pour $f(x)\gg g(x)$.

Cette dernière estimation est en réalité une condition équivalente à l'hypothèse de Riemann, et une autre condition équivalente est

$$|\pi(x) - \mathrm{Li}(x)| \ll x^{\frac{1}{2}} \log x,$$

où $\pi(x) := \#\{p \leq x\}$ et $\mathrm{Li}(x) := \int_2^x \frac{dt}{\log t}$. Le théorème des nombres premiers nous donne que

$$|\pi(x) - \mathrm{Li}(x)| \ll \frac{x}{e^{C\sqrt{\log x}}},$$

donc $E(x) := \pi(x) - \mathrm{Li}(x)$ est un terme d'erreur, ainsi il est naturel d'étudier son comportement. Au lieu de seulement s'intéresser à sa taille en valeur absolue, nous allons tenter de faire une étude qualitative. Intéressons-nous aux changements de signe de $E(x)$. On peut montrer à l'aide d'un ordinateur que $\mathrm{Li}(x) > \pi(x)$ pour $x < 10^{23}$, donc on serait tenter de conjecturer que cette inégalité reste valide pour tout x. Or, Littlewood **[56]** a prouvé en 1914 que $E(x)$ possède une infinité de changements de signes. L'outil fondamental de sa preuve est la formule explicite (1.3.2). L'idée est d'utiliser le principe des tiroirs pour synchroniser les fréquences $\gamma_i \log x$, où γ_i sont les parties imaginaires des zéros non-triviaux de $\zeta(s)$. En utilisant cette idée, Littlewood démontra que[6]

$$\psi(x) - x = \Omega_{\pm}\left(x^{\frac{1}{2}} \log\log\log x\right),$$

d'où on peut déduire en faisant une sommation par parties que

$$\pi(x) - \mathrm{Li}(x) = \Omega_{\pm}\left(x^{\frac{1}{2}} \frac{\log\log\log x}{\log x}\right). \tag{1.6.1}$$

Noter que le $\log\log\log x$ était vraiment essentiel, car en enlevant les puissances de premiers de $\psi(x)$, on fait une erreur qui est d'ordre $x^{\frac{1}{2}}$. Plusieurs auteurs se sont intéressés au premier changement de signe de $E(x)$, que nous allons noter x_0, et la première borne pour x_0 fut donnée par Skewes :

$$x_0 < 10^{10^{10^{963}}}.$$

Cette borne fut beaucoup améliorée au fil du temps, Bays et Hudson on démontré en 2000 **[7]** que

$$x_0 < 1.398244 \cdot 10^{316}.$$

[6] Pour $g(x) > 0$, on dit que $E(x) = \Omega_{\pm}(g(x))$ si $\limsup \frac{E(x)}{g(x)} > 0$ et $\liminf \frac{E(x)}{g(x)} < 0$.

Pour ce qui est de $E_0(x) := \psi(x) - x$, on peut en détecter les premiers changements de signe avant même $x = 10$. Nous allons voir plus tard que la différence fondamentale entre $E(x)$ et $E_0(x)$ est que $E(x)$ est biaisée, en fait sous certaines hypothèses, la «probabilité» que $E(x) > 0$ est de $0.00000026\ldots$, tandis qu'elle est d'exactement $\frac{1}{2}$ pour $E_0(x)$. Nous allons utiliser respectivement la notation $W(T), V(T)$ pour le nombre de changements de signes de $E(x), E_0(x)$ dans l'intervalle $x \in (0, T]$. Nous avons vu que $W(2 \cdot 10^{316}) \geq 1$ et $V(10) \geq 1$. Le comportement asymptotique de $V(T), W(T)$ quand $T \to \infty$ reste encore un mystère. Des bornes inférieures ont été données par Kaczorowski, qui prouva (**[45]**,**[46]**,**[49]**) que

$$\liminf_{T\to\infty} \frac{V(T)}{\log T} > \frac{\gamma_0}{\pi}, \qquad \liminf_{T\to\infty} \frac{W(T)}{\log T} > 0,$$

où

$$\gamma_0 = 14.134725\ldots \tag{1.6.2}$$

est la partie imaginaire du premier zéro non-trivial de $\zeta(s)$. Toutefois, des calculs numériques **[62]** suggèrent qu'en réalité[7],

$$V(T) \asymp \sqrt{T}.$$

Cet ordre de grandeur est supporté par le modèle probabiliste de marche aléatoire. Voyons maintenant pourquoi la preuve de **[45]** est limitée à une borne d'ordre $\log T$. Définissons pour $k \in \mathbb{Z}_{\geq 0}$

$$\psi_k(x) := \sum_{n \leq x} \Lambda(n) \frac{\left(\log \frac{x}{n}\right)^k}{k!}.$$

Ici, $\Lambda(n)$ est la fonction de von-Mangoldt définie par

$$\Lambda(n) := \begin{cases} \log p & \text{si } n = p^e \\ 0 & \text{sinon.} \end{cases}$$

En posant

$$E_k(x) := \psi_k(x) - x,$$

[7] On dit que $f(x) \asymp g(x)$ si $f(x) \gg g(x)$ et $f(x) \ll g(x)$.

on trouve que $E_{k+1}(x) = \int_0^x E_k(t)\frac{dt}{t}$. En effet,

$$\begin{aligned}\int_0^x (\psi_k(t) - t)\,\frac{dt}{t} &= \int_1^x \sum_{n\leq t} \Lambda(n)\frac{\left(\log\frac{t}{n}\right)^k}{k!}\frac{dt}{t} - \int_0^x dt\\ &= \sum_{n\leq x}\Lambda(n)\int_n^x \frac{\left(\log\frac{t}{n}\right)^k}{k!}\frac{dt}{t} - x\\ &= \sum_{n\leq x}\Lambda(n)\int_0^{\log\frac{x}{n}} \frac{u^k}{k!}du - x\\ &= \sum_{n\leq x}\Lambda(n)\frac{\left(\log\frac{x}{n}\right)^{k+1}}{(k+1)!} - x.\end{aligned}$$

Ayant établi ceci, on obtient par le théorème de Rolle que si $V_k(T)$ désigne le nombre de changement de signes de $E_k(t)$ dans l'intervalle $t \in (0, T]$, alors $V_{k+1}(T) - 1 \leq V_k(T)$ pour $k \geq 1$. Une version généralisée de ce théorème nous permet d'obtenir l'inégalité $V_1(T) - 1 \leq V_0(T)$. Cet argument est à la base de **[45]** (et de la plupart des résultats similaires), et son utilité réside dans le fait que par la formule de Perron (1.3.3), on peut représenter $E_k(x)$ par une somme qui converge de plus en plus rapidement. En effet, on peut montrer une formule explicite pour $\psi_k(x)$:

$$\psi_k(x) = x - \sum_{\rho} \frac{x^\rho}{\rho^{k+1}} + \sum_{i=0}^{k} a_{k-i}\frac{(\log x)^i}{i!}, \tag{1.6.3}$$

où $a_j := -\left(\frac{\zeta'}{\zeta}\right)^j(0)$. Nous montrons maintenant (sous l'hypothèse de Riemann) que l'argument de lissage utilisé dans **[45]** ne pourra jamais donner une meilleure borne inférieure que $V(T) \gg \log T$.

Proposition 1.6.1. *Supposons l'hypothèse de Riemann. Pour* $k \geq 5$,

$$V_k(T) = \frac{\gamma_0}{2\pi}\log T + O(1).$$

DÉMONSTRATION. Soit $x_k > 1$ assez grand tel que pour $x \geq x_k$,

$$\left|\sum_{i=0}^{k} a_{k-i}\frac{(\log x)^i}{i!}\right| + \sum_{n=1}^{\infty}\frac{x^{-2n}}{(2n)^{k+1}} < 15^{-k-3}\sqrt{x}.$$

Étudions pour $t > \log x_k$ la fonction

$$f(t) := \frac{\psi_k(e^t) - e^t}{e^{t/2}}.$$

Par (1.6.3), on a, en notant par γ les parties imaginaires des zéros non-triviaux de $\zeta(s)$ (voir (1.6.2) pour la définition de γ_0),

$$\begin{aligned} f(t) &= -\sum_{\gamma} \frac{e^{i\gamma t}}{\left(\frac{1}{2}+i\gamma\right)^{k+1}} + \overline{O}(15^{-k-3}) \\ &= -\operatorname{Re}\left(\frac{e^{i\gamma_0 t}}{\left(\frac{1}{2}+i\gamma_0\right)^{k+1}}\right) + \overline{O}\left(15^{-k-3} + \sum_{\gamma\neq\gamma_0} \frac{1}{|\frac{1}{2}+i\gamma|^{k+1}}\right) \\ &= \frac{-\cos(\gamma_0 t - \phi_k) + \overline{O}\left(0.3\right)}{\left(\frac{1}{4}+\gamma_0^2\right)^{\frac{k+1}{2}}}, \end{aligned}$$

où $\phi_k := \arctan(2(k+1)\gamma_0)$. Ici, $\overline{O}(A)$ désigne une quantité qui est bornée par A en valeur absolue. Nous avons utilisé les estimations de Backlund[8] **[2]** combinées à un calcul avec le logiciel *Mathematica* pour obtenir le dernier terme d'erreur. On conclut tout de suite (par le théorème des accroissements finis) que

$$V_k(T) \geq \frac{\gamma_0}{2\pi}\log T + O(1).$$

Pour la borne supérieure, on revient à la formule explicite. En modifiant au besoin la valeur de x_k et en utilisant encore une fois le résultat de Backlund, on obtient que

$$\begin{aligned} f'(t) &= -\sum_{\gamma} \frac{i\gamma e^{i\gamma t}}{\left(\frac{1}{2}+i\gamma\right)^{k+1}} + \overline{O}(15^{-k-3}) \\ &= \frac{\gamma_0\left(-\cos(\gamma_0 t + \frac{\pi}{2} - \phi_k) + \overline{O}\left(0.3\right)\right)}{\left(\frac{1}{4}+\gamma_0^2\right)^{\frac{k+1}{2}}}. \end{aligned}$$

En séparant l'intervalle $[0, T]$ en sous-intervalles de la forme

$$I_l^- = \left[\frac{\frac{6l-1}{3}\pi + \phi_k}{\gamma_0}, \frac{\frac{6l+1}{3}\pi + \phi_k}{\gamma_0}\right], I_l^+ = \left[\frac{\frac{6l+2}{3}\pi + \phi_k}{\gamma_0}, \frac{\frac{6l+4}{3}\pi + \phi_k}{\gamma_0}\right],$$

$$J_l^- = \left[\frac{\frac{6l+4}{3}\pi + \phi_k}{\gamma_0}, \frac{\frac{6l+5}{3}\pi + \phi_k}{\gamma_0}\right] \text{ et } J_l^+ = \left[\frac{\frac{6l+1}{3}\pi + \phi_k}{\gamma_0}, \frac{\frac{6l+2}{3}\pi + \phi_k}{\gamma_0}\right],$$

[8] Backlund prouva que si $N(T)$ désigne le nombre de zéros non-triviaux ρ de $\zeta(s)$ avec $|\operatorname{Im}(\rho)| \leq T$, alors pour $T \geq 2$, $|N(T) - \frac{T}{2\pi}\log\frac{T}{2\pi e} + \frac{7}{8}| < 0.137\log T + 0.443\log\log T + 4.35$.

on obtient que $f(t) > 0$ pour $t \in I_l^+$ et $f(t) < 0$ pour $t \in I_l^-$, donc les zéros de f sont hors de ces intervalles. Toutefois, $f'(t) > 0$ pour $t \in J_l^+$ et $f'(t) < 0$ pour $t \in J_l^-$, donc f est injective dans ces intervalles, et ainsi possède un seul zéro par intervalle. On conclut que

$$V_k(T) = \frac{\gamma_0}{\pi} \log T + O(1).$$

□

Il est à noter que Kaczorowski [**47**] prouva un résultat similaire avec un poids beaucoup plus lisse. En définissant $V_\Gamma(T)$ comme le nombre de changements de signe de la fonction

$$E_\Gamma(x) := \sum_{n \geq 1} (\Lambda(n) - 1) e^{-n/x} \tag{1.6.4}$$

pour $x \in (0, T]$, il prouva sous l'hypothèse de Riemann que

$$V_\Gamma(T) \sim \frac{\gamma_0}{\pi} \log T.$$

Il réussit à aller plus loin [**48**] en montrant sans aucune hypothèse que

$$V_\Gamma(T) \geq \left(\frac{\gamma_0}{\pi} + o(1) \right) \log T, \qquad V_\Gamma(T) = o((\log T)^2).$$

Plus récemment, Kaczorowski et Wiertelak [**50**] ont montré que pour $H(T)$ arbitraire, il existe $\gg \frac{\log T}{H(T)}$ changements de signes[9] de $\psi(x) - x$ dans l'intervalle $(0, T]$ qui sont de taille $\gg \sqrt{x} \log\log H(T)$. En prenant $H(T) := e^{(\log\log T)^\epsilon}$, on obtient qu'il existe $\gg (\log T) e^{-(\log\log T)^\epsilon}$ oscillations «à la Littlewood», c'est-à-dire de taille $\sqrt{x} \log\log\log x$.

En conclusion, la méthode qu'on dispose pour étudier $V(T)$ détecte seulement les «grandes oscillations», et on pourrait suspecter que la majorité des oscillations sont très petites. Plus précisément, la méthode de Kaczorowski permet d'étudier les oscillations de taille $\asymp \sqrt{x}$, qui sont aussi présentes pour $E_k(x)$. Nous avons vu que $E_k(x)$ possède $\frac{\gamma_0}{\pi} \log T + O(1)$ changements de signe dans l'intervalle $(0, T]$ pour $k \geq 5$, et ces changements de signe correspondent à de grandes oscillations de $E(x)$, ce qui explique pourquoi il est si difficile

[9] On entend ici qu'il existe une suite $0 < x_1 < x_2 < \cdots < x_{\lfloor \frac{c \log T}{H(T)} \rfloor}$ telle que $\psi(x_{2i}) - x_{2i} > c\sqrt{x_{2i}} \log\ \log H(T)$ et $\psi(x_{2i+1}) - x_{2i+1} < -c\sqrt{x_{2i+1}} \log\ \log H(T)$.

d'améliorer la borne $V(T) \geq \frac{\gamma_0}{\pi} \log T + O(1)$. Par ailleurs, le modèle probabiliste confirme ce phénomène. En effet, avec probabilité 1, l'ensemble des zéros d'un mouvement brownien ne contient pas de points isolés, et ces zéros correspondent à de très petites oscillations.

1.6.2. Biais de Tchebychev

Commençons par un extrait d'une lettre **[75]** de Tchebychev adressée à Fuss en 1853 : «En cherchant l'expression limitative des fonctions qui déterminent la totalité des nombres premiers de la forme $4n+1$ et de ceux de la forme $4n+3$, pris au-dessous d'une limite très grande, je suis parvenu à reconnaître que ces deux fonctions diffèrent notablement entre elles par leurs seconds termes, dont la valeur, pour les nombres $4n+3$, est plus grande que celle pour les nombres $4n+1$; ainsi, si de la totalité des nombres premiers de la forme $4n+3$, on retranche celle des nombres premiers de la forme $4n+1$, et que l'on divise ensuite cette différence par la quantité $\frac{\sqrt{x}}{\log x}$, on trouvera plusieurs valeurs de x telles, que ce quotient s'approchera de l'unité aussi près qu'on le voudra. Cette différence dans la répartition des nombres premiers de la forme $4n+1$ et $4n+3$, se manifeste clairement dans plusieurs cas. Par exemple, 1) à mesure que c s'approche de zéro, la valeur de la série

$$e^{-3c} - e^{-5c} + e^{-7c} + e^{-11c} - e^{-13c} - e^{-17c} + e^{-19c} + e^{-23c} + \dots \qquad (1.6.5)$$

s'approche de $+\infty$; 2) la série

$$f(3) - f(5) + f(7) + f(11) - f(13) - f(17) + f(19) + f(23) + \dots \qquad (1.6.6)$$

où $f(x)$ est une fonction constamment décroissante, ne peut être convergente, à moins que la limite du produit $x^{\frac{1}{2}} f(x)$, pour $x = \infty$, ne soit zéro.»

La normalisation par $\sqrt{x}/\log x$ est justifiée par l'hypothèse de Riemann, car sous (une version généralisée de) cette hypothèse, on a[10]

$$\pi(x;4,3)-\pi(x;4,1)\ll\sqrt{x}\log x. \tag{1.6.7}$$

De plus, on peut prouver que (1.6.5) tend bien vers $+\infty$ en supposant l'hypothèse de Riemann généralisée. Or, certaines configurations des zéros de $\zeta(s)$ hors de la droite $\mathrm{Re}(s)=\frac{1}{2}$ pourraient forcer (1.6.5) à osciller indéfiniment (à être $\Omega_{\pm}(x^{\theta})$, où $\frac{1}{2}<\theta<1$, pour être plus précis). La preuve de ces deux affirmations se fait en utilisant la formule de Perron (voir la proposition (1.3.4)) pour donner une formule explicite avec le poids $e^{-n/x}$, qu'on transforme ensuite avec une sommation par parties pour obtenir (1.6.5). De même, l'hypothèse de Riemann est primordiale pour l'étude de (1.6.6). Toutefois, l'article de Riemann date de 1859, donc six ans après la lettre de Tchebyshev, ainsi on peut croire que ce dernier suspectait la borne (1.6.7) bien avant Riemann. Tchebyshev n'a pas publié la preuve de ses résultats, et il semble que la marge de sa lettre était trop petite pour l'inclure, car on peut montrer que la véracité de ses affirmations est équivalente à l'hypothèse de Riemann pour $L(s,\left(\frac{-4}{\cdot}\right))$.

Le résultat de Littlewood (1.6.1) s'applique aussi à $\pi(x;4,3)-\pi(x;4,1)$, et donc cette fonction possède une infinité de changements de signes. Toutefois, l'observation de Tchebychev était valide en ce sens que l'inégalité $\pi(x;4,3)>\pi(x;4,1)$ est respectée pour une grande majorité de valeurs de x. Les premiers à tenter de rendre cette dernière observation rigoureuse furent Knapowski et Turan, qui conjecturèrent **[54]** que la densité naturelle de l'ensemble $\{n:\pi(n;4,3)>\pi(n;4,1)\}$ est nulle, c'est-à-dire que

$$\lim_{N\to\infty}\frac{1}{N}\#\{n\leq N:\pi(n;4,1)>\pi(n;4,3)\}=0.$$

Cette conjecture s'avéra fausse **[51]**, et non seulement la limite ne vaut pas zéro, mais elle n'existe pas **[52]**. La non-existence de cette limite pose un problème, et

[10] Ici, $\pi(x;q,a)$ désigne la quantité de premiers $p\leq x$ tels que $p\equiv a \bmod q$. De même,

$$\psi(x;q,a):=\sum_{\substack{n\leq x\\ n\equiv a \bmod q}}\Lambda(n).$$

pour y remédier certains auteurs ont commencé à utilisé une mesure différente, qui est plus naturelle pour ce problème. On définit la densité logarithmique

$$\delta(q;a,b) := \lim_{N\to\infty} \frac{1}{\log N} \sum_{\substack{n\leq N:\\ \pi(x;q,a)>\pi(x;q,b)}} \frac{1}{n},$$

si cette limite existe. L'existence de $\delta(q;a,b)$ fut prouvée sous deux conditions par Rubinstein et Sarnak [**68**]. La première condition est l'hypothèse de Riemann généralisée (HRG), c'est-à-dire que pour chaque caractère χ mod q, les zéros non-triviaux de la fonction $L(s,\chi)$ se situent sur la droite $\text{Re}(s) = \frac{1}{2}$. La deuxième condition (IL) stipule que pour q fixé, les parties imaginaires de tous les zéros de $L(s,\chi)$ avec χ mod q sont linéairement indépendantes sur $\mathbb{Q}$. En particulier, IL implique que $L(s,\chi)$ n'a pas de zéro réel dans la bande critique. Rubinstein et Sarnak tirèrent plusieurs autres conséquences de ces deux hypothèses.

Théorème 1.6.1 (Rubinstein et Sarnak)**.** *Supposons HRG et IL. Alors les densités logarithmiques* $\delta(q;a,b)$ *existent,* $0 < \delta(q;a,b) < 1$ *et* $\delta(q;a,b) + \delta(q;b,a) = 1$. *De plus,*

$$\begin{aligned}
\delta(q;a,b) = \frac{1}{2} \quad &\textit{ssi} \quad \begin{cases} a \equiv \square, b \equiv \square (\operatorname{mod} q) \textit{ ou} \\ a \not\equiv \square, b \not\equiv \square (\operatorname{mod} q), \end{cases}\\
\delta(q;a,b) < \frac{1}{2} \quad &\textit{ssi} \quad a \equiv \square, b \not\equiv \square (\operatorname{mod} q),\\
\delta(q;a,b) > \frac{1}{2} \quad &\textit{ssi} \quad a \not\equiv \square, b \equiv \square (\operatorname{mod} q).
\end{aligned}$$

Finalement,

$$\max_{\substack{a,b:\\ (ab,q)=1}} \left| \delta(q;a,b) - \frac{1}{2} \right| \to 0 \textit{ quand } q \to \infty. \tag{1.6.8}$$

Il est à remarquer que l'affirmation $\delta(q;a,b) + \delta(q;b,a) = 1$ n'est pas triviale, car elle est équivalente à l'affirmation que la densité logarithmique de l'ensemble $\{n : \pi(n;q,a) = \pi(n;q,b)\}$ soit zéro.

Au chapitre 2, nous raffinerons (1.6.8) en donnant une formule asymptotique pour $\delta(q;a,b)$ qui possède un terme d'erreur borné par une puissance arbitrairement grande de q^{-1}. Cette formule[11], valide sous les hypothèses HRG et IL, est la suivante, dans le cas où a est un résidu quadratique $\bmod q$ et b ne l'est pas[12] :

$$\delta(q;a,b) = \frac{1}{2} + \frac{\rho(q)}{\sqrt{2\pi V(q;a,b)}} \sum_{\ell=0}^{K} \frac{1}{V(q;a,b)^{\ell}} \sum_{j=0}^{\ell} \rho(q)^{2j} s_{q;a,b}(\ell,j) + O_K\left(\frac{\rho(q)^{2K+3}}{V(q;a,b)^{K+3/2}}\right).$$

Ici, les $s_{q;a,b}(\ell,j)$ (voir la définition 2.2.8) sont des quantités bornées par une fonction de ℓ, et $V(q;a,b) \sim 2\phi(q)\log q$ est une certaine variance que nous allons étudier en détail au chapitre 2.3. En prenant $K=0$, on obtient

$$\delta(q;a,b) = \frac{1}{2} + \frac{\rho(q)}{\sqrt{2\pi V(q;a,b)}} + O\left(\frac{\rho(q)^3}{V(q;a,b)^{3/2}}\right). \tag{1.6.9}$$

On peut développer $V(q;a,b)$ et $s_{q;a,b}(\ell,j)$ en termes élémentaires, on y dédiera la section 2.3.4.

À partir de la formule (1.6.9), nous pourrons détecter la délicate influence qu'ont a et b sur $\delta(q;a,b)$, ce qui nous permettra entre autres de faire des «courses de courses de nombres premiers», c'est à dire d'étudier les inégalités de la forme $\delta(q;a,b) > \delta(q;a',b')$, avec a, a', b, b' fixés et $q \to \infty$.

Nous allons ensuite valider nos résultats théoriques par de nombreux calculs numériques, en donnant par exemple au tableaux 2.4 et 2.5 la liste des $\delta(q;a,b)$ avec $q = 101$ et $q = 420$. Ces tableaux nous permettront d'étudier l'effet individuel de chaque terme arithmétique qui apparaît dans nos formules pour $V(q;a,b)$. Nous démontrerons aussi l'utilité computationnelle de notre formule asymptotique, en calculant au tableau 2.1 l'ensemble de toutes les valeurs de $\delta(q;a,b)$ avec $q \leq 1000$, et au tableau 2.8 la liste des 120 courses les plus biaisées, c'est-à-dire la liste des 120 plus grandes valeurs de $\delta(q;a,b)$.

[11] On note par $\rho(q)$ le nombre de racines carrées mod q. On a que $\rho(q) \in \{2^{\omega(q)}, 2^{\omega(q)-1}, 2^{\omega(q)+1}\}$.

[12] On peut facilement en déduire une formule pour tous les cas, car $\delta(q;b,a) = 1 - \delta(q;a,b)$.

1.7. Distribution des nombres premiers dans les grandes progressions arithmétiques

Comme nous avons remarqué dans la section 1.5.1, le plus grand niveau d'uniformité connu de l'asymptotique (1.5.2) est $q \leq (\log x)^B$. En fait, le théorème 1.5.2 donne, pour $A > 0$ fixé, l'estimation

$$\max_{\substack{1\leq a<q\leq(\log x)^B \\ (a,q)=1}} \left|\psi(x;q,a) - \frac{x}{\phi(q)}\right| \ll \frac{x}{(\log x)^A}.$$

1.7.1. Résultats en moyenne

Il est en général plus facile le borner la moyenne d'une suite que de borner ses termes individuellement. Pour cette raison, certains auteurs ont commencé à étudié la quantité

$$\sum_{q\leq Q} \left|\psi(x;q,a) - \frac{x}{\phi(q)}\right|,$$

qui est souvent suffisante pour résoudre des problèmes concrets en théorie des nombres. Des résultats spectaculaires ont été découverts au fil des années ; Bombieri et Vinogradov ont même réussi à démontrer un résultat en moyenne de la force de l'hypothèse de Riemann généralisée. Le premier résultat dans cette direction est dû à Barban **[3]**. Voici la version forte du théorème de Bombieri-Vinogradov.

Théorème 1.7.1 (Bombieri-Vinogradov). *Soit* $A > 0$ *un nombre réel fixé et* $B = B(A) := A + 5$. *On a*

$$\sum_{q\leq x^{\frac{1}{2}}/(\log x)^B} \max_{y\leq x} \max_{(a,q)=1} \left|\psi(y;q,a) - \frac{y}{\phi(q)}\right| \ll \frac{x}{(\log x)^A}.$$

Parmi les applications frappantes de ce théorème, on trouve la solution de deux pages de Halberstam au problème des diviseurs de Titchmarsh[13] **[34]**,

[13] Il s'agit de prouver que pour $a \neq 0$ fixé et $\tau(n) := \sum_{d|n} 1$, on a l'asymptotique suivante : $\sum_{a<p\leq x} \tau(p-a) \sim \frac{\zeta(2)\zeta(3)}{\zeta(6)} \frac{\phi(a)}{a} \prod_{p|a} \left(1 - \frac{1}{p^2-p+1}\right) x$.

l'existence de nombres premiers très proches l'un de l'autre[14] **[29]** et le théorème de Chen[15] **[14]**.

On croit que le théorème 1.7.1 reste valide avec des valeurs beaucoup plus grandes de q.

Conjecture 1.7.1 (Elliot,Halberstam)**.** *Soit $\epsilon > 0$ et $A > 0$ des nombres réels fixés. On a*

$$\sum_{q \leq x^{1-\epsilon}} \max_{(a,q)=1} \left| \psi(x;q,a) - \frac{x}{\phi(q)} \right| \ll \frac{x}{(\log x)^A}.$$

Goldston, Pintz et Yildirim ont démontré que la conjecture d'Elliot-Halberstam impliquait l'existence d'une infinité de paires p_1, p_2 de nombres premiers tel que $|p_1 - p_2| \leq 16$.

Plusieurs auteurs ont tenté d'améliorer le théorème de Bombieri-Vinogradov, et de trouver ses limitations. Une idée pour une telle amélioration est de fixer l'entier a et d'enlever $\max_{(a,q)=1}$. Une autre est de remplacer les valeurs absolues par un poids $\lambda(q)$. Parmi les résultats positifs dans cette direction, on trouve les suivants, qui vont de plus en plus loin en q.

Théorème 1.7.2 (Bombieri, Friedlander, Iwaniec **[11]**)**.** *Soit $a \neq 0$, $x \geq y \geq 3$, et $Q^2 \leq xy$. On a*

$$\sum_{\substack{Q \leq q < 2Q \\ (q,a)=1}} \left| \psi(x;q,a) - \frac{x}{\phi(q)} \right| \ll x \left(\frac{\log y}{\log x} \right)^2 (\log \log x)^B$$

où B est une certaine constante réelle.

Ce théorème nous permet d'aller jusqu'à $Q = x^{\frac{1}{2} + o\left(\frac{1}{(\log \log x)^B}\right)}$. La plus grande valeur de Q connue fut obtenue peu de temps après par les mêmes auteurs, qui prouvèrent qu'on pouvait aller jusqu'à $Q = x^{\frac{1}{2}+o(1)}$, peu importe la nature du $o(1)$.

Théorème 1.7.3 (Bombieri, Friedlander, Iwaniec **[12]**)**.** *Soit $a \neq 0$ un entier et $A > 0$, $2 \leq Q \leq x^{3/4}$ des nombres réels. Soit $\mathcal{Q}$ l'ensemble des entiers q copremiers*

[14] Goldston, Pintz et Yildirim on démontré que si p_n dénote le n-ième nombre premier, alors $\liminf_{n \to \infty} \frac{p_{n+1} - p_n}{\log p_n} = 0$.

[15] Chen a démontré qu'il existe N_0 tel que tout entier pair $n > N_0$ peut s'écrire de la forme $n = p + P_2$, avec p premier et P_2 le produit d'au plus deux premiers.

avec a *dans l'intervalle* $Q' < q \leq Q$. *Alors,*

$$\sum_{q\in\mathcal{Q}} \left|\pi(x;q,a) - \frac{\pi(x)}{\phi(q)}\right|$$
$$\leq \left\{ K\left(\theta - \frac{1}{2}\right)^2 \frac{x}{\log x} + O_A\left(\frac{x(\log\log x)^2}{(\log x)^3}\right)\right\} \sum_{q\in\mathcal{Q}} \frac{1}{\phi(q)} + O_{a,A}\left(\frac{x}{(\log x)^A}\right),$$

où $\theta := \frac{\log Q}{\log x}$ *et* K *est une constante absolue.*

Théorème 1.7.4 (Bombieri, Friedlander, Iwaniec [**10**]). *Soit* $a \neq 0$, $A > 0$, $\epsilon > 0$ *et* $Q = x^{4/7-\epsilon}$. *Si* $\lambda(q)$ *est une fonction bien-factorisable*[16] *de niveau* Q, *alors on a*

$$\sum_{\substack{q\leq Q\\(q,a)=1}} \lambda(q)\left(\psi(x;q,a) - \frac{x}{\phi(q)}\right) \ll \frac{x}{(\log x)^A}.$$

Théorème 1.7.5 (Bombieri, Friedlander, Granville, Iwaniec [**10**] & [**26**]). *Soit* $a \neq 0$ *et* $A > 0$ *fixés. On a*

$$\sum_{\substack{Q\leq q<2Q\\(q,a)=1}} \left(\psi(x;q,a) - \frac{x}{\phi(q)}\right) \ll Q\log(x/Q) + \frac{x}{(\log x)^A}.$$

Ce théorème nous permet d'aller jusqu'à $Q = x/(\log x)^B$, soit encore plus loin que dans la conjecture d'Elliot-Halberstam !

Parmi les résultats négatifs, on trouve celui de Friedlander-Granville.

Théorème 1.7.6 (Friedlander, Granville [**25**]). *Soit* $B > 1$ *un nombre réel fixé. Il existe des valeurs arbitrairement grandes de* a *et* x *telles que*

$$\sum_{q<x/(\log x)^B} \left|\psi(x;q,a) - \frac{x}{\phi(q)}\right| \gg x.$$

1.7.2. Résultats sur la variance

Une autre approche qui s'avéra fructueuse fut de prendre une moyenne sur les résidus $a \bmod q$ du *carré* du terme d'erreur

$$\psi(x;q,a) - \frac{x}{\phi(q)}, \tag{1.7.1}$$

[16] Pour la définition de fonction bien-factorisable, voir [**10**].

ce qui revient à en étudier la variance. Barban [**4**], et peu après Davenport & Halberstam [**17**] prouvèrent une borne supérieure (de force équivalente à l'hypothèse de Riemann généralisée) pour cette variance.

Théorème 1.7.7 (Barban-Davenport-Halberstam)**.** *Soit* $A > 0$ *fixé. Il existe* $B = B(A)$ *tel que*

$$\sum_{q \leq x/(\log x)^B} \sum_{\substack{a \bmod q: \\ (a,q)=1}} \left(\psi(x; q, a) - \frac{x}{\phi(q)} \right)^2 \ll \frac{x^2}{(\log x)^A}. \tag{1.7.2}$$

Ce théorème fut amélioré par Gallagher [**28**], qui montra qu'on pouvait prendre $B(A) = A + 1$. Les deux ingrédients principaux de la preuve de ces résultats sont le grand crible et le théorème de Siegel-Walfisz. Par la suite, Montgomery [**61**] et Hooley [**39**] donnèrent une asymptotique au terme de gauche de (1.7.2), et par le fait même montrèrent que le résultat de Gallagher est optimal.

Théorème 1.7.8 (Hooley-Montgomery)**.** *Soit* $A > 0$ *fixé. On a*

$$\sum_{q \leq Q} \sum_{\substack{a \bmod q: \\ (a,q)=1}} \left(\psi(x; q, a) - \frac{x}{\phi(q)} \right)^2 = Qx \log Q + D_2 Qx + O\left(Q^{\frac{5}{4}} x^{\frac{3}{4}} + \frac{x^2}{(\log x)^A} \right),$$

où $D_2 := -\frac{1}{2} \left(\log 2\pi + \gamma + \sum_p \frac{\log p}{p(p-1)} \right)$.

L'idée principale de Hooley est assez simple. Si pour $n \leq x$ on a $n \equiv a \bmod q$ avec $q > Q$, alors $n = a + qr$, donc $n \equiv a \bmod r$ avec $r < (x-a)/Q$. Cette méthode d'«interchangement des diviseurs» s'avère fort puissante, car elle permet de transformer les progressions arithmétiques de modules $q \sim x/(\log x)^A$ en progressions de module $q \ll (\log x)^A$, pour lesquelles nous pouvons appliquer le théorème de Siegel-Walfisz. La méthode sera fondamentale aux chapitres 3 et 4.

1.7.3. Résultats sur la moyenne

Les résultats du chapitre 3 portent sur la moyenne de (1.7.1). Nous donnons ici une version qui implique seulement les nombres premiers, donc nous remplacerons[17] $\psi(x;q,a)$ par $\theta(x;q,a)$. Fixons un entier $a \neq 0$, et intéressons-nous à la moyenne suivante :

$$\nu(a,M) := \frac{1}{\frac{\phi(a)}{a}\frac{x}{M}} \sum_{\substack{q \leq \frac{x}{M} \\ (q,a)=1}} \left(\theta(x;q,a) - \frac{\theta(x,\chi_0)}{\phi(q)} - \vartheta(a) \right).$$

Nous avons divisé par $\frac{\phi(a)}{a}\frac{x}{M}$ car c'est la quantité (approximative) d'entiers $q \leq \frac{x}{M}$ tels que $(q,a)=1$. La quantité $\theta(x;q,a) - \frac{\theta(x,\chi_0)}{\phi(q)}$ donne l'écart entre la quantité (pondérée) de premiers dans la classe $a \bmod q$ et la quantité moyenne (pondérée) de premiers dans les classes $b \bmod q$, avec $(b,q)=1$. Nous avons soustrait $\vartheta(a)$ de $\theta(x;q,a)$, car si $a=p$, alors la classe d'équivalence $a \bmod q$ possède le premier p pour tout q.

La quantité $\nu(a,M)$ mesure la prépondérance de la classe d'équivalence $a \bmod q$ par rapport aux autres classes $\bmod\, q$, en moyenne sur q. Nous allons prouver au chapitre 3 que cette prépondérence dépend grandement de la valeur de a. En effet, nous démontrerons que pour $M = M(x) \leq \log^B x$, où $B>0$ est fixé,

$$\nu(a,M) = \begin{cases} -\frac{1}{2}\log M - C_5 + O_\epsilon\left(\frac{1}{M^{\frac{205}{538}-\epsilon}}\right) & \text{si } a = \pm 1 \\ -\frac{1}{2}\log p + O_\epsilon\left(\frac{1}{M^{\frac{205}{538}-\epsilon}}\right) & \text{si } a = \pm p^e \\ O_\epsilon\left(\frac{1}{M^{\frac{205}{538}-\epsilon}}\right) & \text{si } \omega(a) \geq 2, \end{cases} \tag{1.7.3}$$

avec[18]

$$C_5 := \frac{1}{2}\left(\log 2\pi + \gamma + \sum_p \frac{\log p}{p(p-1)} + 1 \right).$$

On conclut que les classes d'équivalence ±1 mod q possèdent beaucoup moins de premiers que les classes $\pm p^e \bmod q$, qui eux-mêmes possèdent beaucoup

[17] Nous utilisons les notations $\vartheta(a) := \log a$ si a est premier et $\vartheta(a) := 0$ sinon, $\theta(x;q,a) := \sum_{\substack{n \leq x \\ n \equiv a \bmod q}} \vartheta(n)$, et $\theta(x,\chi) := \sum_{n\leq x} \vartheta(n)\chi(n)$. De plus, nous dénoterons par $\omega(n)$ le nombre de facteurs premiers distincts de n.

[18] On dénote par $\gamma := \lim_{n\to\infty}\left(\sum_{k\leq n}\frac{1}{k} - \log n\right)$ la constante d'Euler-Mascheroni.

moins de premiers que les classes $a \bmod q$ avec $\omega(a) \geq 2$. Il semble contradictoire que certaines classes possèdent moins de premiers que la moyenne, mais aucune n'en contienne plus. Ce «paradoxe» peut s'expliquer par le fait que la proportion d'entiers n ayant au plus un facteur premier est nulle, donc «presque toutes» les classes d'équivalence ne possèdent pas de prépondérance de premiers.

1.8. Quelques suites arithmétiques

Le but de cette section est d'étudier quelques exemples de suites arithmétiques, en exposant certains résultats sur leur distribution dans les progressions arithmétiques. Nous verrons que pour chacune de ces suites, il existe un résultat analogue à (1.7.3), c'est-à-dire que certaines classes d'équivalence possèdent une prépondérance (presque toujours négative) d'éléments de la suite. Chacune de ces suites donne un exemple d'application des résultats plus généraux du chapitre 4. Nous allons voir à la section 4.3.1 comment donner un cadre général (basé sur **[33]**) pour étudier de telles suites dans les progressions arithmétiques.

1.8.1. Entiers représentables comme la somme de deux carrés

L'étude des entiers s'exprimant comme la somme de deux carrés remonte à l'antiquité. Parmi les résultats de cette époque on trouve l'identité de Diophante, qui montre que le produit de deux nombres de cette forme s'exprime lui aussi comme somme de deux carrés :

$$(a^2+b^2)(c^2+d^2) = (ac-bd)^2 + (ad+bc)^2.$$

Les mathématiciens de la Renaissance réussirent à donner une caractérisation complète de ces entiers.

Théorème 1.8.1 (Fermat)**.** *Un premier* p *s'exprime comme la somme de deux carrés ssi* $p = 2$ *ou* $p \equiv 1 \bmod 4$.

On peut déduire par multiplicativité une version plus générale.

Théorème 1.8.2. *Un entier positif* n *peut s'exprimer comme somme de deux carrés ssi pour chaque* $p \mid n$ *avec* $p \equiv 3 \bmod 4$, p *divise* n *à une puissance paire, c'est-à-dire que* $p^{2\nu} \parallel n$*, avec* $\nu \in \mathbb{N}$.

À l'aide de cette caractérisation, on peut démontrer un résultat de Landau sur la proportion d'entiers $n \leq x$ qui s'expriment comme somme de deux carrés. Définissons

$$b(n) := \begin{cases} 1 & \text{si } \exists x, y \in \mathbb{Z} \text{ tels que } n = x^2 + y^2, \\ 0 & \text{sinon.} \end{cases}$$

Théorème 1.8.3 (Landau)**.**

$$\sum_{n \leq x} b(n) = \frac{Bx}{\sqrt{\log x}} \left(1 + O\left(\frac{1}{\log x}\right)\right), \tag{1.8.1}$$

où $B := \frac{1}{\sqrt{2}} \prod_{p \equiv 3 \bmod 4} \left(1 - \frac{1}{p^2}\right)^{-1}$.

Le terme d'erreur de (1.8.1) est vraiment d'ordre $\frac{1}{\log x}$, en fait on peut prouver (voir l'exercice 240 de [**76**]) que pour tout entier $K \geq 1$ fixé,

$$\sum_{n \leq x} b(n) = \sum_{i=0}^{K} B_i \frac{x}{(\log x)^{i+\frac{1}{2}}} + O\left(\frac{x}{(\log x)^{K+\frac{3}{2}}}\right),$$

avec $B_i \in \mathbb{R}$ des constantes. Ce résultat s'explique par le fait que la «fonction zêta» (nous utilisons la notation $\chi_d := \left(\frac{4d}{\cdot}\right)$)

$$\sum_{n=1}^{\infty} \frac{b(n)}{n^s} = (1 - 2^{-s})^{-\frac{1}{2}} \zeta(s)^{\frac{1}{2}} L(s, \chi_{-1})^{\frac{1}{2}} \prod_{p \equiv 3 \bmod 4} \left(1 - \frac{1}{p^{2s}}\right)^{-\frac{1}{2}},$$

définie pour $\mathrm{Re}(s) > \frac{1}{2}$ (nous avons utilisé la méthode de Selberg-Delange), possède une singularité essentielle en $s = 1$ de type $(s-1)^{-\frac{1}{2}}$. Il est possible en utilisant la proposition 1.3.5 d'enlever cette singularité essentielle, et ainsi d'obtenir le résultat suivant :

$$\sum_{n \leq x} b(n) h(x/n) = Bx + O\left(x e^{-C\sqrt{\log x}}\right), \tag{1.8.2}$$

avec B et $C > 0$ des constantes et $h(y) := \int_1^y \frac{dt}{\sqrt{\log t}}$.

La suite $\{b(n)\}_{n \geq 1}$ a aussi été étudiée dans les progressions arithmétiques, et Prachar prouva en 1953 que si a et q sont des entiers copremiers avec $a \equiv$

$1 \bmod (4,q)$, alors

$$\sum_{\substack{n\leq x\\ n\equiv a \bmod q}} b(n) \sim \frac{B_q x}{\sqrt{\log x}}, \tag{1.8.3}$$

où

$$B_q := B\frac{(4,q)}{(2,q)q}\prod_{\substack{p\mid q\\ p\equiv 3 \bmod 4}}\left(1+\frac{1}{p}\right).$$

La question de l'uniformité en q de l'asymptotique (1.8.3) s'est naturellement posée par la suite, et le meilleur résultat à date est dû à Iwaniec [**43**], qui montra en utilisant le crible semi-dimensionnel que l'asymptotique (1.8.3) est valide dès que $q \leq x^{o(1)}$.

Théorème 1.8.4 (Iwaniec)**.** *Pour* a *et* q *des entiers copremiers avec* $a \equiv 1 \bmod (4,q)$,

$$\sum_{\substack{n\leq x\\ n\equiv a \bmod q}} b(n) = \frac{B_q x}{\sqrt{\log x}}\left(1+O\left(\left(\frac{\log q}{\log x}\right)^{\frac{1}{5}}\right)\right).$$

Il n'est pas difficile de généraliser ce résultat à tous les entiers a et q, avec $2 \nmid a$. En effet, on peut montrer que pour de tels entiers,

$$\sum_{\substack{n\leq x\\ n\equiv a \bmod q}} b(n) = \mathbf{g}_a(q)\sum_{n\leq x} b(n)\left(1+O\left(\left(\frac{\log q}{\log x}\right)^{\frac{1}{5}}\right)\right),$$

où $\mathbf{g}_a(q)$ est une fonction multiplicative que nous allons définir sur les puissances de premiers. Pour $p \neq 2$ tel que $p^f \parallel a$, $f \geq 0$, on définit

$$\mathbf{g}_a(p^e) := \frac{1}{p^e}\times\begin{cases} 1 & \text{si } p\equiv 1 \bmod 4\\ 1 & \text{si } p\equiv 3 \bmod 4, e\leq f, 2\mid e\\ \frac{1}{p} & \text{si } p\equiv 3 \bmod 4, e\leq f, 2\nmid e\\ 1+\frac{1}{p} & \text{si } p\equiv 3 \bmod 4, e> f, 2\mid f\\ 0 & \text{si } p\equiv 3 \bmod 4, e> f, 2\nmid f. \end{cases} \tag{1.8.4}$$

De plus, $\mathbf{g}_a(2) := \frac{1}{2}$, et pour $e \geq 2$, $\mathbf{g}_a(2^e) := \frac{1+(-1)^{\frac{a-1}{2}}}{2^{e+2}}$.

Au chapitre 4, nous donnerons l'analogue suivant de (1.7.3). Fixons $a \equiv 1 \bmod 4$ et $0 < \lambda < 1/5$. Alors dans l'intervalle $1 \leq M(x) \leq (\log x)^{\lambda}$, nous avons

$$\frac{1}{x/2M}\sum_{\frac{x}{2M}<q\leq\frac{x}{M}}\left(\sum_{\substack{n\leq x\\ n\equiv a \bmod q}} b(n) - b(a) - \mathbf{g}_a(q)\sum_{n\leq x} b(n)\right)$$
$$\sim -\left(\frac{\log M}{\log x}\right)^{\frac{1}{2}}\frac{(-4)^{-l_a-1}(2l_a+2)!}{(4l_a^2-1)(l_a+1)!\pi}\prod_{\substack{p^f\|a:\\ p\equiv 3 \bmod 4,\\ f \text{ impair}}}\frac{\log(p^{\frac{f+1}{2}})}{\log M}, \quad (1.8.5)$$

où $l_a := \#\{p^f \| a : p \equiv 3 \bmod 4, 2 \nmid f\}$ est le nombre de facteurs premiers de a intervenant à une puissance impaire, qui sont congrus à 3 modulo 4.

Noter que (1.8.5) est $o((\log x)^{-1/2})$ si et seulement si $b(|a|) = 0$. Dans le cas $a \equiv 3 \bmod 4$, alors on peut montrer que (1.8.5) est toujours $o((\log x)^{-1/2})$. Finalement, si $a = \square + \square$, alors le terme de droite de (1.8.5) devient $-\frac{1}{2\pi}\left(\frac{\log M}{\log x}\right)^{\frac{1}{2}}$.

1.8.2. Valeurs d'une forme quadratique binaire, avec multiplicité

Fort de ses découvertes sur les premiers s'exprimant comme somme de deux carrés, Fermat conjectura que $p = x^2 + 2y^2$ ssi $p \equiv 1$ ou $3 \bmod 8$, et $p = x^2 + 3y^2$ ssi $p = 3$ ou $p \equiv 1 \bmod 3$. Nous nous intéresserons plus généralement à

$$Q(x, y) = \alpha x^2 + \beta xy + \gamma y^2$$

une forme quadratique définie positive[19], avec $\alpha, \beta, \gamma \in \mathbb{Z}$ et $(\alpha, \beta, \gamma) = 1$. Nous allons compter les entiers représentés par $Q(x, y)$ avec multiplicité en définissant

$$r_Q(n) := \#\{(x, y) \in \mathbb{Z}_{\geq 0} : Q(x, y) = n\}.$$

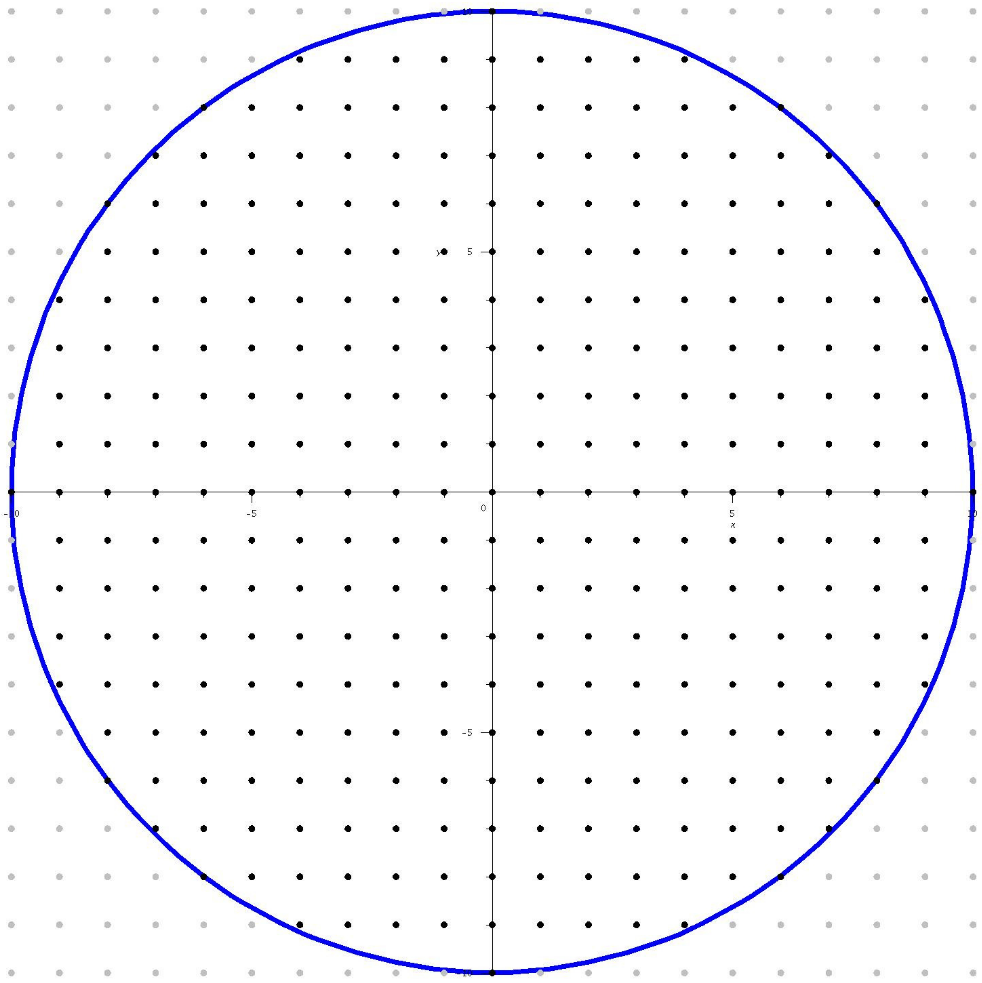
Gauss donna un magnifique argument pour compter la valeur moyenne de $r_Q(n)$ pour $n \leq N$. Son idée fut d'interpréter la quantité $\sum_{n\leq N} r_Q(n)$ comme étant le nombre de points à coordonnés entières $(a, b) \in \mathcal{R}_N$, où

$$\mathcal{R}_N := \{(x, y) \in \mathbb{R}^2_{\geq 0} : Q(x, y) \leq N\}.$$

En observant la figure 1.1 (dans laquelle on considère les quatre cadrans à la fois), on voit que la quantité qui nous intéresse est très bien approximée par

[19] Une condition nécessaire et suffisante est d'exiger que $\Delta := \beta^2 - 4\alpha\gamma < 0$ et $\alpha > 0$.

FIG. 1.1. Le problème du cercle



l'aire de $\mathcal{R}_N$, avec un terme d'erreur de l'ordre de son périmètre. On en déduit l'estimation de Gauss :

$$\sum_{n \leq N} r_Q(n) = A_Q N + O(\sqrt{N}),$$

avec $A_Q :=$ Aire$(\mathcal{R}_1)$.

Le cas particulier $Q(x, y) := x^2 + y^2$ (voir la figure 1.1) est appelé «problème du cercle». Le problème consiste à trouver l'ordre de grandeur exact de

$$E(N) := \#\{(m, n) : m^2 + n^2 \leq N\} - \pi N.$$

La meilleure borne supérieure est celle de Huxley **[42]**, qui prouva que $E(N) \ll_\epsilon N^{\frac{131}{416}+\epsilon}$. D'autre part, Hardy **[35]** et Landau ont prouvé que $E(N) = \Omega_\pm(N^{\frac{1}{4}})$, ce qui a été amélioré par la suite par de nombreux auteurs, les meilleurs résultats étant ceux de Corrádi et Kátai **[15]**, Hafner **[38]** et Soundararajan **[73]**. Soundararajan montra que

$$E(N) = \Omega((N \log N)^{\frac{1}{4}} (\log \log N)^{\frac{3}{4}(2^{1/3}-1)} (\log \log \log N)^{-\frac{5}{8}}).$$

Revenons au cas général, en s'intéressant à la distribution de $r_Q(n)$ dans les progressions arithmétiques. On peut se convaincre avec un argument intuitif de l'asymptotique suivante :

$$\sum_{\substack{n \leq N \\ n \equiv a \bmod q}} r_Q(n) \sim \frac{R_a(q)}{q^2} \sum_{n \leq N} r_Q(n),$$

où

$$R_a(q) := \#\{1 \leq x, y \leq q : Q(x,y) \equiv a \bmod q\}.$$

Plaksin a démontré **[65]** que cette asymptotique est valide uniformément pour $q \leq x^{\frac{2}{3}-\epsilon}$.

Théorème 1.8.5 (Plaksin)**.**

$$\sum_{\substack{n \leq N \\ n \equiv a \bmod q}} r_Q(n) = \frac{R_a(q)}{q^2} \sum_{n \leq N} r_Q(n) + E(x,q),$$

où

$$E(x,q) \ll_{a,\epsilon} \begin{cases} (x/q)^{\frac{3}{4}+\epsilon} & \textit{si } q \leq x^{\frac{1}{3}} \\ x^{\frac{2}{3}+\epsilon} q^{-\frac{1}{2}} & \textit{si } x^{\frac{1}{3}} < q \leq x^{\frac{2}{3}}. \end{cases}$$

Gauss développa grandement la théorie des formes quadratiques. Il remarqua que les formes quadratiques devaient être étudiées modulo les transformations de $SL_2(\mathbb{Z})$. Par exemple, les deux formes quadratiques

$$Q_1(x,y) = x^2 + 2y^2, \qquad Q_2(x,y) = 3x^2 + 10xy + 9y^2$$

représentent exactement les mêmes entiers, car $Q_2(x,y) = Q_1(x+y, x+2y)$, et $Q_1(x,y) = Q_2(2x-y, y-x)$. En général, nous dirons que Q_1 et Q_2 sont équivalentes s'il existe

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$$

telle que $Q_2(x,y) = Q_1(ax+by, cx+dy)$. Q_1 et Q_2 représentent alors les mêmes entiers, et on peut voir avec un calcul qu'ils ont le même discriminant. Gauss démontra que toute forme quadratique ést équivalente à une unique forme $Q_0(x,y) = \alpha_0 x^2 + \beta_0 xy + \gamma_0 y^2$ avec la propriété suivante :

$$-\alpha_0 < \beta_0 \leq \alpha_0 < \gamma_0 \qquad \text{ou} \qquad 0 \leq \beta_0 \leq \alpha_0 = \gamma_0. \tag{1.8.6}$$

Si $\Delta < 0$ est fixé, il n'est pas difficile de montrer qu'il existe seulement un nombre fini de valeurs de α, β, γ qui satisfont (1.8.6) telles que $\beta^2 - 4\alpha\gamma = \Delta$, donc il existe un nombre fini de classes d'équivalence de formes quadratiques de discriminant Δ. Ce nombre, noté $h(\Delta)$, est appelé nombre de classes.

Pour un entier n, on définit $r_\Delta(n)$ comme étant le nombre de représentations distinctes[20] de n par l'ensemble des $h(\Delta)$ formes non-équivalentes de discriminant Δ. Il existe une formule remarquable pour $r_\Delta(n)$:

$$r_\Delta(n) = \sum_{d|n} \left(\frac{\Delta}{d}\right). \tag{1.8.7}$$

On peut exploiter la multiplicativité en n :

$$\begin{aligned} r_\Delta(n) &= \prod_{p^f \| n} \left(1 + \left(\frac{\Delta}{p}\right) + \cdots + \left(\frac{\Delta}{p^f}\right)\right) \\ &= \prod_{\substack{p^f \| n: \\ \left(\frac{\Delta}{p}\right)=1}} (f+1) \prod_{\substack{p^f \| n: \\ \left(\frac{\Delta}{p}\right)=-1 \\ f \text{ impair}}} 0 \end{aligned}$$

[20] Ici, on compte les représentations dans les quatre quadrants modulo les automorphismes de $Q(x,y)$ par des éléments de $SL_2(\mathbb{Z})$. Pour $\Delta \notin \{-3,-4\}$, il existe un seul automorphisme non-trivial : $(x,y) \mapsto (-x,-y)$, donc chaque paire de représentations $n = Q(a,b) = Q(-a,-b)$ est comptée une seule fois. Pour $Q(x,y) = x^2+y^2$, on a aussi l'automorphisme $(x,y) \mapsto (-y,x)$, donc on compte seulement une fois le quadruplet de solutions $a^2+b^2 = (-a)^2 + (-b)^2 = (-b)^2 + a^2 = b^2 + (-a)^2$. Pour $Q(x,y) = x^2+xy+y^2$, on a l'automorphisme $(x,y) \mapsto (x+y,-x)$, donc on compte seulement une fois le sextuplet de solutions $Q(a,b) = Q(a+b,-a) = Q(b,-a-b) = Q(-a,-b) = Q(-a-b,a) = Q(-b,a+b)$. L'utilité d'un tel compte est qu'il permet d'utiliser la théorie des anneaux pour obtenir la formule (1.8.7).

(un produit vide vaut toujours 1). Pour $\Delta < -1$, nous noterons par $w(\Delta)$ le nombre d'automorphismes des formes de discriminant Δ, donc

$$w(\Delta) = \begin{cases} 2 & \text{si } \Delta < -4, \\ 4 & \text{si } \Delta = -4, \\ 6 & \text{si } \Delta = -3. \end{cases}$$

Dirichlet utilisa (1.8.7) pour obtenir sa formule du nombre de classes :

$$L\left(1, \left(\frac{\Delta}{\cdot}\right)\right) = \frac{2\pi h(\Delta)}{w(\Delta)\sqrt{-\Delta}}.$$

Voyons maintenant l'analogue de (1.7.3) pour la suite $r_Q(n)$. Fixons $Q(x,y)$ telle que $\Delta \equiv 1, 5, 9, 12, 13 \bmod 16$, et définissons $\rho_a(q) := R_a(q)/q$. Soit a un entier fixé tel que $(a, 2\Delta) = 1$. Alors, dans l'intervalle $M = M(x) \leq x^\lambda$, avec $0 < \lambda < \frac{1}{12}$ fixé, nous avons

$$\frac{1}{x/M} \sum_{q \leq \frac{x}{M}} \left(\sum_{\substack{n \leq x \\ n \equiv a \bmod q}} r_Q(n) - \frac{\rho_a(q)}{q} \sum_{n \leq x} r_Q(n) - r_Q(a) \right)$$
$$= -C_Q \rho_a(4\Delta) r_\Delta(|a|) + O_\epsilon\left(\frac{1}{M^{1/3-\epsilon}}\right),$$

où

$$C_Q := \frac{A_Q}{2L\left(1, \left(\frac{\Delta}{\cdot}\right)\right)} \qquad \left(= \frac{w(\Delta)\sqrt{|\Delta|}}{4\pi h(\Delta)} A_Q\right).$$

On peut voir que $\rho_a(4\Delta)$ est soit zéro, ou égal à $2^{\omega(2\Delta)}$, $2^{\omega(2\Delta)-2}$ ou $3 \cdot 2^{\omega(2\Delta)-2}$, selon la nature de $Q(x,y)$. Ainsi, si $\rho_a(4\Delta) > 0$, alors ce nombre est indépendant de a.

On conclut que la prépondérance est nulle si $\rho_a(4\Delta) = 0$ ou si aucune forme de discriminant Δ ne représente $|a|$. Dans le cas contraire, la prépondérance est négative et proportionnelle au nombre de telles représentations.

1.8.3. k-tuplets premiers

Un des grands problèmes ouverts de la théorie des nombres est la conjecture des premiers jumeaux. On croit qu'il existe une infinité de premiers p tels que $p+2$ est aussi premier. Cette conjecture n'est toujours pas démontrée jusqu'à maintenant, mais certains résultats partiels sont connus. Le crible combinatoire de Brun lui permit de prouver que la somme

$$\sum_{p:p+2 \text{ est premier}} \frac{1}{p}$$

est convergente. Ce résultat découle de la borne supérieure suivante :

$$\pi_2(x) := \#\{p \leq x : p+2 \text{ premier}\} \ll \frac{x}{(\log x)^2}.$$

On croit que $x/(\log x)^2$ est le bon ordre de grandeur pour $\pi_2(x)$, en fait il existe une conjecture précise à cet effet.

Conjecture 1.8.1 (Hardy-Littlewood)**.**

$$\pi_2(x) \sim 2C_2 \frac{x}{(\log x)^2},$$

où $C_2 := \prod_{p>2}\left(1 - \frac{1}{(p-1)^2}\right)$ *est la constante des premiers jumeaux.*

La conjecture de Hardy et Littlewood est basée sur leur méthode du cercle (introduite par Hardy et Ramanujan), et permet de prédire (et dans plusieurs cas important de prouver) de nombreux résultats. La fameux article de Hardy-Littlewood [**36**] contient de nombreuses autres conjectures (environ 15), dont les suivantes, qui sont encore aujourd'hui des problèmes ouverts.

Conjecture 1.8.2 (Hardy-Littlewood)**.**

$$\#\{p, q \textit{ premiers} : p+q = 2n\} \sim 2C_2 \frac{2n}{(\log 2n)^2} \prod_{\substack{p|n \\ p\neq 2}} \frac{p-1}{p-2} \qquad (n \to \infty),$$

$$\#\{p \leq x : p+2k \textit{ est premier}\} \sim 2C_2 \frac{x}{(\log x)^2} \prod_{\substack{p|k \\ p\neq 2}} \frac{p-1}{p-2} \qquad (x \to \infty),$$

$$\#\{n \leq x : n^2+1 \textit{ est premier}\} \sim \frac{x}{2\log x} \prod_{p\neq 2} \left(1 - \left(\frac{-1}{p}\right)\frac{1}{p-1}\right).$$

On croit aussi qu'il existe une infinité de premiers de Sophie Germain, c'est-à-dire de premiers p tels que $2p+1$ est aussi premier. De plus, on pense qu'il existe une infinité de triplets de premiers des formes suivantes : $(p, p+2, p+6)$, $(p, 2p+1, p+2)$, $(p, 2p+1, 3p+8)$,... Toutefois, n, $n+2$ et $n+4$ ne peuvent être tous premiers (pour $n > 3$), car au moins un des trois est divisible par

3. Ce dernier triplet est un exemple de triplet non-admissible, c'est-à-dire qui ne peut admettre simultanément trois valeurs premières, sauf pour un nombre fini de valeurs de n.

En général, on considère $\mathcal{H} := \{\mathcal{L}_1, \ldots \mathcal{L}_k\}$ un k-tuplet de formes linéaires $\mathcal{L}_i = a_i n + b_i$ avec $a_i, b_i \in \mathbb{Z}$ et $a_i \geq 1$. On définit aussi

$$\mathcal{P}(n; \mathcal{H}) := (a_1 n + b_1)(a_2 n + b_2) \cdots (a_k n + b_k), \tag{1.8.8}$$

et le produit singulier

$$\mathfrak{S}(\mathcal{H}) := \prod_p \left(1 - \frac{\nu_{\mathcal{H}}(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}.$$

(La terminologie de produit singulier provient de l'article de Hardy-Littlewood.) On dit que $\mathcal{H}$ est admissible si pour chaque premier p,

$$\nu_{\mathcal{H}}(p) := \#\{a \bmod p : \mathcal{P}(a; \mathcal{H}) \equiv 0 \bmod p\} < p.$$

On peut voir que si $\nu_{\mathcal{H}}(p) = p$, alors en réduisant tous les $a_i n + b_i$ mod p, on obtient l'ensemble de tous les résidus, donc au moins un des entiers $a_i n + b_i$ est divisible par p (et ainsi $a_i n + b_i$ n'est pas premier pour $n > p + |b_i|$). Ceci explique pourquoi nous considérons seulement les k-tuplets admissibles.

Avec cette notation, nous pouvons formuler la conjecture générale des k-tuplets de Hardy et Littlewood

Conjecture 1.8.3 (Hardy-Littlewood). *Si $\mathcal{H}$ est admissible, alors pour $x \to \infty$ on a l'asymptotique*

$$\sum_{n \leq x} \Lambda(a_1 n + b_1)\Lambda(a_2 n + b_2) \cdots \Lambda(a_k n + b_k) \sim \mathfrak{S}(\mathcal{H}) x.$$

Les constantes précises (comme $\mathfrak{S}(\mathcal{H})$) qui apparaissent dans les conjectures de Hardy et Littlewood proviennent toutes de la méthode du cercle. Voyons maintenant comment obtenir une telle constante dans le cas des premiers jumeaux. En utilisant la formule (pour $l \in \mathbb{Z}$)

$$\int_0^1 e^{2\pi i l \alpha}\, d\alpha = \begin{cases} 1 & \text{si } l = 0 \\ 0 & \text{sinon,} \end{cases} \tag{1.8.9}$$

on obtient l'identité

$$\psi_2(x) := \sum_{n\le x}\Lambda(n)\Lambda(n+2) = \sum_{m,n\le x}\Lambda(m)\Lambda(n)\int_0^1 e^{2\pi i(m-n-2)\alpha}d\alpha + O(\log x)$$
$$= \int_0^1 e^{-4\pi i\alpha}\left|\sum_{n\le x}\Lambda(n)e^{2\pi i n\alpha}\right|^2 d\alpha + O(\log x), \tag{1.8.10}$$

qui explique l'appellation «méthode du cercle»[21]. Il est donc primordial de comprendre la fonction

$$f_x(\alpha) := \sum_{n\le x}\Lambda(n)e^{2\pi i n\alpha}. \tag{1.8.11}$$

Hardy et Littlewood ont réalisé que $f_x(\alpha)$ est grande quand α est proche d'un nombre rationnel. La prochaine étape est donc de séparer le domaine d'intégration $[0,1]$ en «arcs mineurs», où nous supposons que la contribution à l'intégrale est négligeable, et en «arcs majeurs», qui constituent le terme principal. Plus formellement, on définit

$$\mathfrak{M}_{P,Q} := \left\{\alpha\in[0,1] : \left|\alpha-\frac{a}{q}\right| < \frac{1}{Q} \text{ pour } a, q \text{ avec } 1\le a\le q, q\le P\right\}$$

et $\mathfrak{m}_{P,Q} := [0,1]\setminus\mathfrak{M}_{P,Q}$. Les conjectures de Littlewood sont basées sur la supposition que pour P et Q bien choisis, la contribution de l'intégrale sur $\mathfrak{m}_{P,Q}$ dans (1.8.10) est négligeable. Ensuite, il est possible d'estimer l'intégrale sur $\mathfrak{M}_{P,Q}$ avec précision. Pour ce faire, on suit les arguments de Davenport (chapitre 26 de [**16**]), et on choisit $P := (\log x)^A$, $Q := x/(\log x)^A$, avec A assez grand. En posant $\beta := \alpha - a/q$ et utilisant les propriétés des sommes de Gauss[22], on montre que pour α proche de a/q,

$$f_x(\alpha) = \sum_{\substack{n\le x\\(n,q)=1}}\Lambda(n)e^{2\pi i n\alpha} + O(\log q\log x)$$
$$= \frac{1}{\phi(q)}\sum_{\chi \bmod q}\tau(\overline{\chi})\chi(a)\sum_{n\le x}\chi(n)\Lambda(n)e^{2\pi i n\beta} + O((\log x)^2), \tag{1.8.12}$$

[21] Originalement, Hardy et Littlewood travaillaient avec la variable $z = e^{2\pi i\alpha}$, donc leur intégrale était sur le cercle unité. Dans ce cas, l'identité (1.8.10) découle du théorème des résidus.

[22] Pour un caractère χ mod q, la somme de Gauss est définie par $\tau(\chi) := \sum_{a=1}^{q}\chi(a)e^{\frac{2\pi i a}{q}}$. Ces sommes permettent de faire un pont entre caractères additifs et multiplicatifs. Pour leurs propriétés, voir par exemple [**8**].

donc on se retrouve avec une somme qui ressemble beaucoup à $\psi(x,\chi) := \sum_{n\leq x}\chi(n)\Lambda(n)$. Si $\chi \neq \chi_0$, le théorème de Siegel-Walfisz nous donne la borne $\psi(x,\chi) \ll x/e^{C\sqrt{\log x}}$ (car $q \leq P = (\log x)^A$), ce qui donne (après une sommation par parties) une contribution $\ll (1+|\beta|x)x/e^{-C\sqrt{\log x}}$ à (1.8.12). Comme $\tau(\chi_0) = \mu(q)$, on obtient (en utilisant une deuxième sommation par parties)

$$\begin{aligned} f_x(\alpha) &= \frac{\mu(q)}{\phi(q)} \sum_{\substack{n\leq x\\(n,q)=1}} \Lambda(n)e^{2\pi i n\beta} + O((1+|\beta|x)x/e^{C\sqrt{\log x}}) \\ &= \frac{\mu(q)}{\phi(q)} \sum_{n\leq x} e^{2\pi i n\beta} + O((1+|\beta|x)x/e^{C\sqrt{\log x}}). \end{aligned}$$

Noter que pour $\alpha \in \mathfrak{M}_{P,Q}$, ce terme d'erreur est $\ll x/e^{C_1\sqrt{\log x}}$. Donc au total, en posant $E := x/e^{C_2\sqrt{\log x}}$,

$$\begin{aligned} &\int_{\mathfrak{M}_{P,Q}} e^{-4\pi i\alpha}|f_x(\alpha)|^2\,d\alpha \\ &\quad= \sum_{\substack{1\leq a\leq q\leq P\\(a,q)=1}} \frac{\mu^2(q)}{\phi^2(q)} \int_{\frac{a}{q}-\frac{1}{Q}}^{\frac{a}{q}+\frac{1}{Q}} e^{-4\pi i\alpha}\left|\sum_{n\leq x} e^{2\pi i n\beta}\right|^2 d\alpha + O(x/e^{C_2\sqrt{\log x}}) \\ &\quad= \sum_{q\leq P} \frac{\mu^2(q)}{\phi^2(q)} \sum_{\substack{1\leq a\leq q\\(a,q)=1}} e^{\frac{-4\pi i a}{q}} \int_{-\frac{1}{Q}}^{\frac{1}{Q}} e^{-4\pi i\beta}\left|\sum_{n\leq x} e^{2\pi i n\beta}\right|^2 d\beta + O(x/e^{C_2\sqrt{\log x}}) \\ &\quad= I_Q(x)\sum_{q\leq P} \frac{|\mu(q)|}{\phi(q)}\frac{\mu(q/(2,q))}{\phi(q/(2,q))} + O(x/e^{C_2\sqrt{\log x}}), \end{aligned} \qquad (1.8.13)$$

par l'évaluation classique des sommes de Ramanujan[23], où

$$\begin{aligned} I_Q(x) :&= \int_{-\frac{1}{Q}}^{\frac{1}{Q}} e^{-4\pi i\beta}\left|\sum_{n\leq x} e^{2\pi i n\beta}\right|^2 d\beta \\ &= \int_0^1 e^{-4\pi i\beta}\left|\sum_{n\leq x} e^{2\pi i n\beta}\right|^2 d\beta + O(x/(\log x)^A). \end{aligned}$$

En réutilisant l'identité (1.8.9), on voit que cette dernière intégrale est égale au nombre d'entiers $n, m \leq x$ tels que $n = m+2$, donc à $x + O(1)$. En substituant

[23] Les sommes de Ramanujan sont définies par $c_q(n) := \sum_{\substack{1\leq a\leq q\\(a,q)=1}} e^{\frac{2\pi i a n}{q}}$. En utilisant la formule d'inversion de Möbius, on peut montrer que $c_q(n) = \frac{\mu(q/(n,q))\phi(q)}{\phi(q/(n,q))}$.

dans (1.8.13), on trouve

$$\begin{aligned}\int_{\mathfrak{M}_{P,Q}} e^{-4\pi i\alpha}|f_x(\alpha)|^2\,d\alpha &= x\sum_{q\leq P}\frac{|\mu(q)|}{\phi(q)}\frac{\mu(q/(2,q))}{\phi(q/(2,q))}+O(x/(\log x)^A)\\ &= x\sum_{q=1}^{\infty}\frac{|\mu(q)|}{\phi(q)}\frac{\mu(q/(2,q))}{\phi(q/(2,q))}+O(x/(\log x)^{A-1})\\ &= 2x\prod_{p\neq 2}\left(1-\frac{1}{(p-1)^2}\right)+O(x/(\log x)^{A-1}).\end{aligned}$$

On conclut que si la contribution des arcs mineurs est négligeable, alors

$$\psi_2(x)\sim 2x\prod_{p\neq 2}\left(1-\frac{1}{(p-1)^2}\right).$$

Il est en général très difficile de borner la contribution des arcs mineurs. Dans le cas du problème des trois premiers toutefois, Vinogradov réussit à démontrer que cette contribution est bien négligeable en utilisant ses estimations sur les sommes exponentielles pour borner (1.8.11). Il en déduisit le théorème 1.1.1.

Retournons maintenant aux progressions arithmétiques. Définissons

$$\mathbf{a}(n) := \Lambda(a_1n+b_1)\Lambda(a_2n+b_2)\cdots\Lambda(a_kn+b_k),$$

$$\mathcal{A}(x;q,a) := \sum_{\substack{n\leq x\\ n\equiv a \bmod q}}\mathbf{a}(n), \qquad \mathcal{A}(x):=\mathcal{A}(x;1,0).$$

Au chapitre 4, nous démontrerons l'analogue suivant de (1.7.3) dans le cas des k-tuplets de nombres premiers. Définissons $\gamma(q) := \prod_{p|q}\left(1-\frac{\nu_{\mathcal{H}}(p)}{p}\right)$. Alors, en supposant une version uniforme de la conjecture de Hardy-Littlewood, on a pour $M = M(x)\leq \log x$ et $\mathcal{P}_a := \mathcal{P}(a,\mathcal{H})$ (voir (1.8.8)) que

$$\frac{1}{\frac{\phi(\mathcal{P}_a)}{\mathcal{P}_a}\frac{x}{2M}}\sum_{\substack{\frac{x}{2M}<q\leq\frac{x}{M}:\\ (q,\mathcal{P}_a)=1}}\left(\mathcal{A}(x;q,a)-\mathbf{a}(a)-\frac{\mathcal{A}(x)}{q\gamma(q)}\right)\text{ est}$$

$$\begin{cases}\sim -\dfrac{(\log M)^{k-\omega(\mathcal{P}_a)}}{2(k-\omega(\mathcal{P}_a))!}\displaystyle\prod_{p|\mathcal{P}_a}\frac{p-\nu_{\mathcal{H}}(p)}{p-1}\log p & \text{si } \omega(\mathcal{P}_a)\leq k,\\ = O\left(\frac{1}{M^{\delta_k}}\right) & \text{sinon,}\end{cases}$$

où $\delta_k>0$ est un nombre réel. Noter qu'en prenant $k=1$, on retrouve (1.7.3).

1.8.4. Entiers sans petit facteur premier

Le crible d'Ératosthène est basé sur le fait que tout nombre composé n est divisible par un premier $p \leq \sqrt{n}$. Ainsi, si de la table des entiers entre 1 et x on crible les multiples de tous les premiers $p \leq \sqrt{x}$, le résultat sera la liste des nombres premiers entre $\sqrt{x}$ et x. En utilisant le principe d'inclusion-exclusion, on obtient que la quantité d'entiers non-criblés est de[24]

$$\begin{aligned}\pi(x)-\pi(\sqrt{x}) &= \lfloor x \rfloor - \sum_{p\leq\sqrt{x}} \left\lfloor \frac{x}{p} \right\rfloor + \sum_{p_1,p_2\leq\sqrt{x}} \left\lfloor \frac{x}{p_1p_2} \right\rfloor - \sum_{p_1,p_2,p_3\leq\sqrt{x}} \left\lfloor \frac{x}{p_1p_2p_3} \right\rfloor + \dots \\ &\approx x - \sum_{p\leq\sqrt{x}} \frac{x}{p} + \sum_{p_1,p_2\leq\sqrt{x}} \frac{x}{p_1p_2} - \sum_{p_1,p_2,p_3\leq\sqrt{x}} \frac{x}{p_1p_2p_3} + \dots \\ &= x \prod_{p\leq\sqrt{x}} \left(1-\frac{1}{p}\right). \qquad (1.8.14)\end{aligned}$$

Cette approximation n'est pas exacte, car on sait par le théorème de Mertens[25] que (1.8.14) est asymptotique à $2e^{-\gamma}x/\log x$, et le théorème des nombres premiers donne $\pi(x) \sim x/\log x$. Le problème provient du fait qu'on a criblé $\sqrt{x}$ premiers, et ceci est beaucoup trop. Plutôt que de s'intéresser à $\pi(x)$ directement, on s'intéresse à la quantité d'entiers n'ayant pas de facteurs premiers $< y$, donc à

$$\Phi(x,y) := \#\{n \leq x : p \mid n \Rightarrow p \geq y\}.$$

Pour des valeurs modérées de y (en fonction de x), un raffinement de l'argument précédent permet d'obtenir l'asymptotique

$$\Phi(x,y) \sim x \prod_{p\leq y} \left(1-\frac{1}{p}\right);$$

il s'agit du lemme fondamental du crible combinatoire (voir (4.6.8), où nous utilisé la notation $\mathcal{A}(x,y) := \Phi(x,y)$, ou **[18]**). Pour des valeurs plus grandes de y, Buchstab **[13]** utilisa le théorème des nombres premiers pour montrer que pour $u > 1$ fixé,

$$\Phi(x, x^{\frac{1}{u}}) \sim \omega(u) \frac{x}{\log x^{\frac{1}{u}}}.$$

[24] On utilise la notation $\lfloor x \rfloor$ pour le plus petit entier inférieur ou égal à x.

[25] Le théorème de Mertens est l'estimation $\prod_{p\leq y}\left(1-\frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\log y}$.

La fonction $\omega(u)$, maintenant appelée fonction de Buchstab, est l'unique solution à l'équation différentielle aux différences

$$(u\omega(u))' = \omega(u-1) \qquad (u > 2)$$

avec la condition initiale $u\omega(u) = 1$, pour $1 \leq u \leq 2$. Cette fonction est continue, et elle tend vers $e^{-\gamma}$ quand $u \to \infty$. Pour une étude plus approfondie de $\Phi(x,y)$ et du problème analogue des entiers friables, voir [**31**] ou [**76**].

Étudions maintenant les entiers sans petit facteur premier dans les progressions arithmétiques, en définissant

$$\Phi(x,y;q,a) := \#\{n \leq x : n \equiv a \bmod q, p \mid n \Rightarrow p \geq y\}.$$

Comme dans le cas des nombres premiers, si a et q sont des entiers tels que $(a,q) = 1$, alors on s'attend à l'asymptotique

$$\Phi(x,y;q,a) \sim \frac{\Phi(x,y)}{\phi(q)}. \tag{1.8.15}$$

Une version plus générale du lemme fondamental du crible nous permet d'obtenir l'asymptotique (1.8.15), uniformément pour q et y tels que q est y-friable, $y < x/q$ et $y \leq x^{o(1)}$ (voir [**79**]). Il existe aussi un théorème de Bombieri-Vinogradov [**78**] pour les entiers sans petit facteur premier.

Théorème 1.8.6 (Wolke)**.** *Peu importe* $A > 0$*, il existe* $B = B(A)$ *tel que si* $Q \leq x^{\frac{1}{2}}/\log^B x$*, nous avons uniformément pour* $y \leq \sqrt{x}$*,*

$$\sum_{q \leq Q} \max_{(a,q)=1} \max_{z \leq x} \left| \Phi(z,y;q,a) - \frac{1}{\phi(q)} \sum_{\substack{n \leq z \\ (n,q)=1 \\ p|n \Rightarrow p \geq y}} 1 \right| \ll \frac{x}{\log^A x}. \tag{1.8.16}$$

Voici l'analogue de (1.7.3) que nous démontrerons dans ce cas. Définissons

$$\nu_y(a,M) := \frac{1}{\frac{x}{2M}\frac{\phi(a)}{a}} \sum_{\substack{\frac{x}{2M} < q \leq \frac{x}{M} \\ (q,a)=1}} \left(\Phi(x,y;q,a) - \epsilon_{a=\pm 1} - \frac{\Phi(x,y)}{q\gamma_y(q)} \right),$$

où $\epsilon_{a=\pm 1}$ vaut 1 si $a = \pm 1$, et 0 sinon. Soit $a \neq 0$ fixé, $\delta > 0$ et $M = M(x) \leq (\log x)^{1-\delta}$. Alors pour $y \leq e^{(\log M)^{\frac{1}{2}-\delta}}$ avec $y \to \infty$,

$$\nu_y(a,M) = \begin{cases} -\frac{1}{2} + o(1) & \text{si } a = \pm 1 \\ o(1) & \text{sinon,} \end{cases}$$

et pour $(\log x)^{\log\log\log x} \leq y \leq \sqrt{x}$,

$$\nu_y(a, M) = \frac{\mathcal{A}(x,y)}{x} \times \begin{cases} (-\frac{1}{2} + o(1)) \log M & \text{si } a = \pm 1 \\ -\frac{\log p}{2} + o(1) & \text{si } a = \pm p^k \\ o(1) & \text{sinon.} \end{cases}$$

En prenant $y = \sqrt{x}$, on retrouve une version non-pondérée de (1.7.3).

Chapitre 2

INEQUITIES IN THE SHANKS-RÉNYI PRIME NUMBER RACE : AN ASYMPTOTIC FORMULA FOR THE DENSITIES

Authors : Daniel Fiorilli and Greg Martin

Abstract : Chebyshev was the first to observe a bias in the distribution of primes in residue classes. The general phenomenon is that if a is a nonsquare (mod q) and b is a square (mod q), then there tend to be more primes congruent to a (mod q) than b (mod q) in initial intervals of the positive integers ; more succinctly, there is a tendency for $\pi(x; q, a)$ to exceed $\pi(x; q, b)$. Rubinstein and Sarnak defined $\delta(q; a, b)$ to be the logarithmic density of the set of positive real numbers x for which this inequality holds ; intuitively, $\delta(q; a, b)$ is the "probability" that $\pi(x; q, a) > \pi(x; q, b)$ when x is "chosen randomly". In this paper, we establish an asymptotic series for $\delta(q; a, b)$ that can be instantiated with an error term smaller than any negative power of q. This asymptotic formula is written in terms of a variance $V(q; a, b)$ that is originally defined as an infinite sum over all nontrivial zeros of Dirichlet L-functions corresponding to characters (mod q) ; we show how $V(q; a, b)$ can be evaluated exactly as a finite expression. In addition to providing the exact rate at which $\delta(q; a, b)$ converges to $\frac{1}{2}$ as q grows, these evaluations allow us to compare the various density values $\delta(q; a, b)$ as a and b vary modulo q ; by analyzing the resulting formulas, we can explain and predict which of these densities will be larger or smaller, based on arithmetic properties of the residue classes a and

b (mod q). For example, we show that if a is a prime power and a' is not, then $\delta(q; a, 1) < \delta(q; a', 1)$ for all but finitely many moduli q for which both a and a' are nonsquares. Finally, we establish rigorous numerical bounds for these densities $\delta(q; a, b)$ and report on extensive calculations of them, including for example the determination of all 117 density values that exceed $\frac{9}{10}$.

2.1. Introduction

We have known for over a century now that the prime numbers are asymptotically evenly distributed among the reduced residue classes modulo any fixed positive integer q. In other words, if $\pi(x; q, a)$ denotes the number of primes not exceeding x that are congruent to a (mod q), then

$$\lim_{x\to\infty} \frac{\pi(x; q, a)}{\pi(x; q, b)} = 1$$

for any integers a and b that are relatively prime to q. However, this information by itself is not enough to tell us about the distribution of values of the difference $\pi(x; q, a) - \pi(x; q, b)$, in particular whether this difference must necessarily take both positive and negative values. Several authors—notably Chebyshev in 1853 and Shanks **[72]** in 1959—observed that $\pi(x; 4, 3)$ has an extremely strong tendency to be greater than $\pi(x; 4, 1)$, and similar biases exist for other moduli as well. The general phenomenon is that $\pi(x; q, a)$ tends to exceed $\pi(x; q, b)$ when a is a nonsquare modulo q and b is a square modulo q.

In 1994, Rubinstein and Sarnak **[68]** developed a framework for studying these questions that has proven to be quite fruitful. Define $\delta(q; a, b)$ to be the logarithmic density of the set of real numbers $x \ge 1$ satisfying $\pi(x; q, a) > \pi(x; q, b)$. (Recall that the logarithmic density of a set S of positive real numbers is

$$\lim_{X\to\infty} \left(\frac{1}{\log X} \int\limits_{\substack{1\le x\le X \\ x\in S}} \frac{dx}{x} \right),$$

or equivalently the natural density of the set $\{\log x \colon x \in S\}$.) Rubinstein and Sarnak investigated these densities under the following two hypotheses :

- The Generalized Riemann Hypothesis (GRH) : all nontrivial zeros of Dirichlet L-functions have real part equal to $\frac{1}{2}$

- A linear independence hypothesis (LI) : the nonnegative imaginary parts of these nontrivial zeros are linearly independent over the rationals

Under these hypotheses, they proved that the limit defining $\delta(q;a,b)$ always exists and is strictly between 0 and 1. Among other things, they also proved that $\delta(q;a,b)$ tends to $\frac{1}{2}$ as q tends to infinity, uniformly for all pairs a,b of distinct reduced residues (mod q).

In the present paper, we examine these densities $\delta(q;a,b)$ more closely. We are particularly interested in a quantitative statement of the rate at which $\delta(q;a,b)$ approaches $\frac{1}{2}$. In addition, computations show that for a fixed modulus q, the densities $\delta(q;a,b)$ vary as a and b range over nonsquares and squares modulo q, respectively. We are also interested in determining which pairs a,b (mod q) give rise to larger or smaller values of $\delta(q;a,b)$, and especially in giving criteria that depend as directly as possible on a and b rather than on analytic data such as the zeros of Dirichlet L-functions.

Our first theorem, which is proved in Section 2.2.4, exhibits an asymptotic series for $\delta(q;a,b)$:

Theorem 2.1.1. *Assume GRH and LI. Let q be a positive integer, and let $\rho(q)$ be the function defined in Definition 2.1.1. Let a and b be reduced residues* (mod q) *such that a is a nonsquare* (mod q) *and b is a square* (mod q)*, and let $V(q;a,b)$ be the variance defined in Definition 2.1.2. Then for any nonnegative integer K,*

$$\delta(q;a,b) = \frac{1}{2} + \frac{\rho(q)}{\sqrt{2\pi V(q;a,b)}} \sum_{\ell=0}^{K} \frac{1}{V(q;a,b)^{\ell}} \sum_{j=0}^{\ell} \rho(q)^{2j} s_{q;a,b}(\ell,j) + O_K\biggl(\frac{\rho(q)^{2K+3}}{V(q;a,b)^{K+3/2}}\biggr), \quad (2.1.1)$$

where the real numbers $s_{q;a,b}(\ell,j)$, which are bounded in absolute value by a function of ℓ uniformly in q, a, b, and j, are defined in Definition 2.2.8. In particular, $s_{q;a,b}(0,0)=1$, so that

$$\delta(q;a,b) = \frac{1}{2} + \frac{\rho(q)}{\sqrt{2\pi V(q;a,b)}} + O\biggl(\frac{\rho(q)^3}{V(q;a,b)^{3/2}}\biggr). \quad (2.1.2)$$

We will see in Proposition 2.3.4 that $V(q;a,b) \sim 2\phi(q)\log q$, and so the error term in equation (2.1.1) is $\ll_{K,\varepsilon} 1/q^{K+3/2-\varepsilon}$.

The assumption that a is a nonsquare (mod q) and b is a square (mod q) is natural in this context, reflecting the bias observed by Chebyshev. Rubinstein and Sarnak showed (assuming GRH and LI) that $\delta(q;b,a)+\delta(q;a,b)=1$; therefore if a is a square (mod q) and b is a nonsquare (mod q), the right-hand sides of the asymptotic formulas (2.1.1) and (2.1.2) become $\frac12-\cdots$ instead of $\frac12+\cdots$. Rubinstein and Sarnak also showed that $\delta(q;b,a)=\delta(q;a,b)=\frac12$ if a and b are both squares or both nonsquares (mod q).

The definitions of $\rho(q)$ and of $V(q;a,b)$ are as follows :

Definition 2.1.1. *As usual, $\omega(q)$ denotes the number of distinct prime factors of q. Define $\rho(q)$ to be the number of real characters* (mod q)*, or equivalently the index of the subgroup of squares in the full multiplicative group* (mod q)*, or equivalently still the number of solutions of $x^2\equiv1$* (mod q)*. An exercise in elementary number theory shows that*

$$\rho(q)=\begin{cases}2^{\omega(q)}, & \text{if } 2\nmid q,\\ 2^{\omega(q)-1}, & \text{if } 2\mid q \text{ but } 4\nmid q,\\ 2^{\omega(q)}, & \text{if } 4\mid q \text{ but } 8\nmid q,\\ 2^{\omega(q)+1}, & \text{if } 8\mid q,\end{cases}$$

which implies that $\rho(q)\ll_\varepsilon q^\varepsilon$ for every $\varepsilon>0$. ♢

Definition 2.1.2. *For any Dirichlet character χ* (mod q)*, define*

$$b(\chi)=\sum_{\substack{\gamma\in\mathbb{R}\\ L(1/2+i\gamma,\chi)=0}}\frac{1}{\frac14+\gamma^2}.$$

We adopt the convention throughout this paper that the zeros are listed with multiplicity in all such sums (though note that the hypothesis LI, when in force, implies that all such zeros are simple). For any reduced residues a and b (mod q)*, define*

$$V(q;a,b)=\sum_{\chi\,(\mathrm{mod}\ q)}|\chi(b)-\chi(a)|^2 b(\chi).$$

We will see in Proposition 2.2.3 that $V(q;a,b)$ is the variance of a particular distribution associated with the difference $\pi(x;q,a)-\pi(x;q,b)$. ♢

As the asymptotic series in Theorem 2.1.1 depends crucially on the variance $V(q;a,b)$, we next give a formula for it (established in Section 2.3.2) that involves only a finite number of easily computed quantities :

Theorem 2.1.2. *Assume GRH. For any pair* a, b *of distinct reduced residues modulo* q,

$$V(q;a,b) = 2\phi(q)\big(\mathcal{L}(q) + K_q(a-b) + \iota_q(-ab^{-1})\log 2\big) + 2M^*(q;a,b),$$

where the functions $\mathcal{L}$, K_q, *and* ι_q *are defined in Definition 2.1.3 and the quantity* $M^*(q;a,b)$ *is defined in Definition 2.1.4.*

The definitions of these three arithmetic functions and of the analytic quantity M^* are as follows :

Definition 2.1.3. *As usual,* $\phi(q)$ *denotes Euler's totient function, and* $\Lambda(q)$ *denotes the von Mangoldt function, which takes the value* $\log p$ *if* q *is a power of the prime* p *and 0 otherwise. For any positive integer* q, *define*

$$\mathcal{L}(q) = \log q - \sum_{p \mid q} \frac{\log p}{p-1} + \frac{\Lambda(q)}{\phi(q)} - (\gamma_0 + \log 2\pi),$$

where $\gamma_0 = \lim_{x\to\infty} \big(\sum_{n\le x} \frac{1}{n} - \log x\big)$ *is Euler's constant ; it can be easily shown that* $\mathcal{L}(q)$ *is positive when* $q \ge 43$. *Note that* $\mathcal{L}(q) = \log(q/2\pi e^{\gamma_0})$ *when* q *is prime and that* $\mathcal{L}(q) = \log q + O(\log\log q)$ *for any integer* $q \ge 3$. *Also let*

$$\iota_q(n) = \begin{cases} 1, & \textit{if } n \equiv 1 \pmod{q}, \\ 0, & \textit{if } n \not\equiv 1 \pmod{q} \end{cases}$$

denote the characteristic function of the integers that are congruent to 1 (mod q). *Finally, define*

$$K_q(n) = \frac{\Lambda(q/(q,n))}{\phi(q/(q,n))} - \frac{\Lambda(q)}{\phi(q)}.$$

Note that these last two functions depend only on the residue class of n *modulo* q. *For this reason, in expressions such as* $\iota_q(n^{-1})$ *or* $K_q(n^{-1})$, *the argument* n^{-1} *is to be interpreted as an integer that is the multiplicative inverse of* n (mod q). *In addition, note that* $K_q(n) \ge 0$, *since the only way that the second term can contribute is if* q *is a prime power, in which case the first term contributes at least as much. On the* *other hand* K_q *is bounded above, since if* q *is a power of the prime* p *then* $K_q(n) \le (\log p)/(p-1) \le \log 2$. *Note also that* $K_q(n) = 0$ *when* $(n,q) = 1$. ♢

Definition 2.1.4. *As usual,* $L(s,\chi) = \sum_{n=1}^{\infty} \chi(n) n^{-s}$ *denotes the* L*-function associated to the Dirichlet character* χ. *Given such a character* $\chi \pmod q$*, let* q^* *denote its conductor (that is, the smallest integer* d *such that* χ *is induced by a character modulo* d*), and let* χ^* *be the unique character modulo* q^* *that induces* χ. *Now define*

$$M^*(q;a,b) = \sum_{\substack{\chi \pmod q \\ \chi \ne \chi_0}} |\chi(a) - \chi(b)|^2 \frac{L'(1,\chi^*)}{L(1,\chi^*)}$$

and

$$M(q;a,b) = \sum_{\substack{\chi \pmod q \\ \chi \ne \chi_0}} |\chi(a) - \chi(b)|^2 \frac{L'(1,\chi)}{L(1,\chi)}.$$

♢

The formula for $V(q;a,b)$ in Theorem 2.1.2 is exact and hence well suited for computations. For theoretical purposes, however, we need a better understanding of $M^*(q;a,b)$, which our next theorem (proved in Section 2.3.3) provides :

Theorem 2.1.3. *Assume GRH. For any pair* a, b *of distinct reduced residues modulo* q*, let* r_1 *and* r_2 *denote the least positive residues of* ab^{-1} *and* $ba^{-1} \pmod q$*, and let the quantity* $H(q;a,b)$ *be defined in Definition 2.1.5. Then*

$$M^*(q;a,b) = \phi(q)\left(\frac{\Lambda(r_1)}{r_1} + \frac{\Lambda(r_2)}{r_2} + H(q;a,b) + O\left(\frac{\log^2 q}{q}\right)\right),$$

where the implied constant is absolute.

(The unexpected appearance of the specific integers r_1 and r_2, in a formula for a quantity depending upon entire residue classes (mod q), is due to the approximation of infinite series by their first terms—see Proposition 2.3.7.) The quantity $H(q;a,b)$ is usually quite small, unless there is an extreme coincidence in the locations of a and b relative to the prime divisors of q, which would be reflected in a small value of the quantity $e(q;p,r)$ defined as follows :

Definition 2.1.5. *Given an integer* q *and a prime* p*, let* $\nu \ge 0$ *be the integer such that* $p^\nu \parallel q$ *(that is,* $p^\nu \mid q$ *but* $p^{\nu+1} \nmid q$*). For any reduced residue* $r \pmod q$*, define*

$e(q;p,r) = \min\{e \geq 1\colon p^e \equiv r^{-1} \pmod{q/p^\nu}\}$, *and define*

$$h(q;p,r) = \frac{1}{\phi(p^\nu)} \frac{\log p}{p^{e(q;p,r)}}.$$

When r *is not in the multiplicative subgroup generated by* $p \pmod{q/p^\nu}$, *we make the convention that* $e(q;p,r) = \infty$ *and* $h(q;p,r) = 0$. *Finally, for any integers* a *and* b, *define*

$$H(q;a,b) = \sum_{p \mid q} \big(h(q;p,ab^{-1}) + h(q;p,ba^{-1}) \big).$$

Note that if $q = p^\nu$ *is a prime power, then* $h(q;p,r) = (\log p)/p^\nu(p-1)$ *is independent of* r, *which implies that* $H(q;a,b) \ll (\log q)/q$ *when* q *is a prime power.* $\diamondsuit$

The extremely small relative error in Theorem 2.1.1 implies that the formula given therein is useful even for moderate values of q. The following corollary of the above theorems, the proof of which is given in Section 2.4.1, is useful only for large q due to a worse error term. It has the advantage, however, of isolating the fine-scale dependence of $\delta(q;a,b)$ on the residue classes a and b from its primary dependence on the modulus q :

Corollary 2.1.4. *Assume GRH and LI. Let* $q \geq 43$ *be an integer. Let* a *and* b *be reduced residues* $\pmod q$ *such that* a *is a nonsquare* $\pmod q$ *and* b *is a square* $\pmod q$, *and let* r_1 *and* r_2 *denote the least positive residues of* ab^{-1} *and* ba^{-1} $\pmod q$. *Then*

$$\delta(q;a,b) = \frac{1}{2} + \frac{\rho(q)}{2\sqrt{\pi\phi(q)\mathcal{L}(q)}} \bigg(1 - \frac{\Delta(q;a,b)}{2\mathcal{L}(q)} + O\Big(\frac{1}{\log^2 q}\Big) \bigg), \tag{2.1.3}$$

where

$$\Delta(q;a,b) = K_q(a-b) + \iota_q(-ab^{-1}) \log 2 + \frac{\Lambda(r_1)}{r_1} + \frac{\Lambda(r_2)}{r_2} + H(q;a,b) \tag{2.1.4}$$

(here, the functions $\mathcal{L}$, K_q, *and* ι_q *are defined in Definition 2.1.3, and* H *is defined in Definition 2.1.5). Moreover,* $\Delta(q;a,b)$ *is nonnegative and bounded above by an absolute constant.*

Armed with this knowledge of the delicate dependence of $\delta(q;a,b)$ on the residue classes a and b, we are actually able to “race races”, that is, investigate inequalities between various values of $\delta(q;a,b)$ as q increases. We remark that Feuerverger and Martin [**22**, Theorem 2(b)] showed that $\delta(q;a,b) =$

$\delta(q; ab^{-1}, 1)$ for any square $b \pmod{q}$, and so it often suffices to consider only the densities $\delta(q; a, 1)$. Some surprising inequalities come to light when we fix the residue class a and allow the modulus q to vary (among moduli relatively prime to a for which a is a nonsquare). Our next theorem, which is a special case of Corollary 2.4.2 derived in Section 2.4.2, demonstrates some of these inequalities :

Theorem 2.1.5. *Assume GRH and LI.*

- *For any integer $a \neq -1$, we have $\delta(q; -1, 1) < \delta(q; a, 1)$ for all but finitely many integers q with $(q, a) = 1$ such that both -1 and a are nonsquares* $\pmod{q}$.
- *If a is a prime power and $a' \neq -1$ is an integer that is not a prime power, then $\delta(q; a, 1) < \delta(q; a', 1)$ for all but finitely many integers q with $(q, aa') = 1$ such that both a and a' are nonsquares* $\pmod{q}$.
- *If a and a' are prime powers with $\Lambda(a)/a > \Lambda(a')/a'$, then $\delta(q; a, 1) < \delta(q; a', 1)$ for all but finitely many integers q with $(q, aa') = 1$ such that both a and a' are nonsquares* $\pmod{q}$.

Finally, these results have computational utility as well. A formula [**22**, equation (2-57)] for calculating the value of $\delta(q; a, b)$ is known. However, this formula requires knowledge of a large number of zeros of all Dirichlet L-functions associated to characters $\pmod{q}$ even to estimate via numerical integration ; therefore it becomes unwieldy to use the formula when q becomes large. On the other hand, the asymptotic series in Theorem 2.1.1 can be made completely effective, and the calculation of $V(q; a, b)$ is painless thanks to Theorem 2.1.2. Therefore the densities $\delta(q; a, b)$ can be individually calculated, and collectively bounded, for large q.

For example, the values of $\delta(q; a, b)$ for all moduli up to 1000 are plotted in Figure 2.1. The modulus q is given on the horizontal axis ; the vertical line segment plotted for each q extends between the maximal and minimal values of $\delta(q; a, b)$, as a runs over all nonsquares $\pmod{q}$ and b runs over all squares $\pmod{q}$. (Of course both a and b should be relatively prime to q. We also omit moduli of the form $q \equiv 2 \pmod{4}$, since the distribution of primes into residue classes modulo such q is the same as their distribution into residue classes modulo $q/2$.)

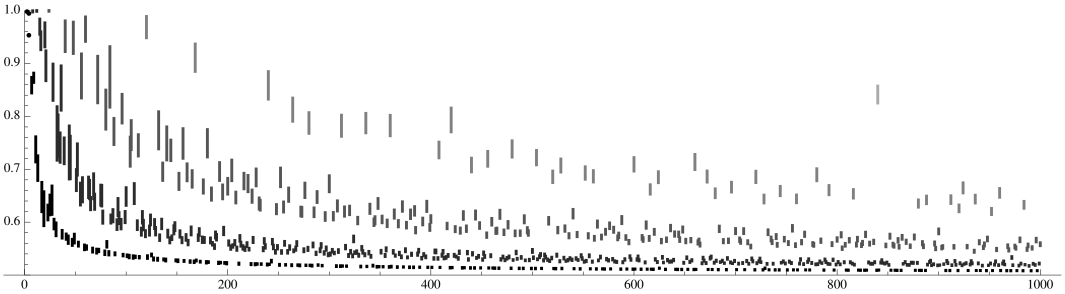
The values shown in Figure 2.1 organize themselves into several bands; each band corresponds to a constant value of $\rho(q)$, the effect of which on the density $\delta(q; a, b)$ can be clearly seen in the second term on the right-hand side of equation (2.1.3). For example, the lowest (and darkest) band corresponds to moduli q for which $\rho(q) = 2$, meaning odd primes and their powers (as well as $q = 4$); the second-lowest band corresponds to those moduli for which $\rho(q) = 4$, consisting essentially of numbers with two distinct prime factors; and so on, with the first modulus $q = 840$ for which $\rho(q) = 32$ (the segment closest to the upper right-hand corner of the graph) hinting at the beginning of a fifth such band. Each band decays roughly at a rate of $1/\sqrt{q \log q}$, as is also evident from the aforementioned term of equation (2.1.3).

To give one further example of these computations, which we describe in Section 2.5.4, we are able to find the largest values of $\delta(q; a, b)$ that ever occur. (All decimals listed in this paper are rounded off in the last decimal place.)

Theorem 2.1.6. *Assume GRH and LI. The ten largest values of* $\delta(q; a, b)$ *are given in Table 2.1.*

Our approach expands upon the seminal work of Rubinstein and Sarnak **[68]**, who introduced a random variable whose distribution encapsulates the information needed to understand $\pi(x; q, a) - \pi(x; q, b)$. We discuss these random variables, formulas and estimates for their characteristic functions (that is, Fourier transforms), and the subsequent derivation of the asymptotic series from Theorem 2.1.1 in Section 2.2. In Section 2.3 we demonstrate how to

FIG. 2.1. All densities $\delta(q; a, b)$ with $q \leq 1000$



TAB. 2.1. The top 10 most unfair prime number races

q	a	b	$\delta(q;a,b)$
24	5	1	0.999988
24	11	1	0.999983
12	11	1	0.999977
24	23	1	0.999889
24	7	1	0.999834
24	19	1	0.999719
8	3	1	0.999569
12	5	1	0.999206
24	17	1	0.999125
3	2	1	0.999063

transform the variance $V(q;a,b)$ from an infinite sum into a finite expression; we can even calculate it extremely precisely using only arithmetic (rather than analytic) information. We also show how the same techniques can be used to establish a central limit theorem for the aforementioned distributions, and we outline how modifications of our arguments can address the two-way race between all nonresidues and all residues (mod q). We investigate the fine-scale effect of the particular residue classes a and b upon the density $\delta(q;a,b)$ in Section 2.4; we also show how a similar analysis can explain a "mirror image" phenomenon noticed by Bays and Hudson [**5**]. Finally, Section 2.5 is devoted to explicit estimates and a description of our computations of the densities and the resulting conclusions, including Theorem 2.1.6.

Acknowledgments

The authors thank Brian Conrey and K. Soundararajan for suggesting proofs of Lemma 2.2.1(c) and Proposition 2.3.5, respectively, that were superior to our original proofs. We also thank Andrew Granville for indicating how to improve the error term in Proposition 2.3.6, as well as Robert Rumely and Michael

Rubinstein for providing lists of zeros of Dirichlet L-functions and the appropriate software to compute these zeros, which are needed for the calculations of the densities in Section 2.5. Finally, we express our gratitude to our advisors past and present, Andrew Granville, Hugh Montgomery, and Trevor Wooley, both for their advice about this paper and for their guidance in general. Le premier auteur est titulaire d'une bourse doctorale du Conseil de recherches en sciences naturelles et en génie du Canada. The second author was supported in part by grants from the Natural Sciences and Engineering Research Council of Canada.

2.2. The asymptotic series for the density $\delta(q; a, b)$

The ultimate goal of this section is to prove Theorem 2.1.1. We begin in Section 2.2.1 by describing a random variable whose distribution is the same as the limiting logarithmic distribution of a suitably normalized version of $\pi(x; q, a) - \pi(x; q, b)$, as well as calculating its variance. This approach is the direct descendant of that of Rubinstein and Sarnak **[68]**; one of our main innovations is the exact evaluation of the variance $V(q; a, b)$ in a form that does not involve the zeros of Dirichlet L-functions. In Section 2.2.2 we derive the formula for the characteristic function (Fourier transform) of that random variable; this formula is already known, but our derivation is slightly different and allows us to write the characteristic function in a convenient form (see Proposition 2.2.5). We then use our knowledge of the characteristic function to write the density $\delta(q; a, b)$ as the truncation of an infinite integral in Section 2.2.3, where the error terms are explicitly bounded using knowledge of the counting function $N(T, \chi)$ of zeros of Dirichlet L-functions. Finally, we derive the asymptotic series from Theorem 2.1.1 from this truncated integral formula in Section 2.2.4.

2.2.1. Distributions and random variables

We begin by describing random variables related to the counting functions of primes in arithmetic progressions. As is typical when considering primes in arithmetic progressions, we first consider expressions built out of Dirichlet characters.

Definition 2.2.1. *For any Dirichlet character χ such that GRH holds for $L(s,\chi)$, define*

$$E(x,\chi) = \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \frac{x^{i\gamma}}{\frac{1}{2}+i\gamma}.$$

This sum does not converge absolutely, but (thanks to GRH and the functional equation for Dirichlet L-functions) it does converge conditionally when interpreted as the limit of $\sum_{|\gamma|<T}$ as T tends to infinity. All untruncated sums over zeros of Dirichlet L-functions in this paper should be similarly interpreted. ♢

Definition 2.2.2. *For any real number γ, let Z_γ denote a random variable that is uniformly distributed on the unit circle, and let X_γ denote the random variable that is the real part of Z_γ. We stipulate that the collection $\{Z_\gamma\}_{\gamma\geq 0}$ is independent and that $Z_{-\gamma} = \overline{Z_\gamma}$; this implies that the collection $\{X_\gamma\}_{\gamma\geq 0}$ is also independent and that $X_{-\gamma} = X_\gamma$.* ♢

By the limiting logarithmic distribution of a real-valued function $f(t)$, we mean the measure $d\nu$ having the property that the limiting logarithmic density of the set of positive real numbers such that $f(t)$ lies between α and β is $\int_\alpha^\beta d\nu$ for any interval (α, β).

Proposition 2.2.1. *Assume LI. Let $\{c_\chi\colon \chi \pmod q\}$ be a collection of complex numbers, indexed by the Dirichlet characters $\pmod q$, satisfying $c_{\bar\chi} = \overline{c_\chi}$. The limiting logarithmic distribution of the function*

$$\sum_{\chi \pmod q} c_\chi E(x,\chi)$$

is the same as the distribution of the random variable

$$2\sum_{\chi \pmod q} |c_\chi| \sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} \frac{X_\gamma}{\sqrt{\frac{1}{4}+\gamma^2}}.$$

Proof. We have

$$\begin{aligned}\sum_{\chi \pmod q} c_\chi E(x,\chi) &= \lim_{T\to\infty} \sum_{\chi \pmod q} c_\chi \sum_{\substack{|\gamma|<T \\ L(1/2+i\gamma,\chi)=0}} \frac{x^{i\gamma}}{\frac12 + i\gamma} \\ &= \lim_{T\to\infty} \sum_{\chi \pmod q} c_\chi \Bigg(\sum_{\substack{0<\gamma<T \\ L(1/2+i\gamma,\chi)=0}} \frac{x^{i\gamma}}{\frac12 + i\gamma} + \sum_{\substack{-T<\gamma<0 \\ L(1/2+i\gamma,\chi)=0}} \frac{x^{i\gamma}}{\frac12 + i\gamma} \Bigg).\end{aligned}$$

(The assumption of LI precludes the possibility that $\gamma = 0$.) By the functional equation, the zeros of $L(s,\chi)$ below the real axis correspond to those of $L(s,\bar\chi)$ above the real axis. Therefore

$$\begin{aligned}&\sum_{\chi \pmod q} c_\chi E(x,\chi) \\ &\quad= \lim_{T\to\infty} \sum_{\chi \pmod q} c_\chi \Bigg(\sum_{\substack{0<\gamma<T \\ L(1/2+i\gamma,\chi)=0}} \frac{x^{i\gamma}}{\frac12 + i\gamma} + \sum_{\substack{0<\gamma<T \\ L(1/2+i\gamma,\bar\chi)=0}} \frac{x^{-i\gamma}}{\frac12 - i\gamma} \Bigg) \qquad (2.2.1) \\ &\quad= \lim_{T\to\infty} \Bigg(\sum_{\chi \pmod q} c_\chi \sum_{\substack{0<\gamma<T \\ L(1/2+i\gamma,\chi)=0}} \frac{x^{i\gamma}}{\frac12 + i\gamma} + \sum_{\chi \pmod q} \overline{c_{\bar\chi} \sum_{\substack{0<\gamma<T \\ L(1/2+i\gamma,\bar\chi)=0}} \frac{x^{i\gamma}}{\frac12 + i\gamma}} \Bigg).\end{aligned}$$

Reindexing this last sum by replacing $\bar\chi$ by χ, we obtain

$$\begin{aligned}&\sum_{\chi \pmod q} c_\chi E(x,\chi) \\ &\quad= \lim_{T\to\infty} \Bigg(\sum_{\chi \pmod q} c_\chi \sum_{\substack{0<\gamma<T \\ L(1/2+i\gamma,\chi)=0}} \frac{x^{i\gamma}}{\frac12 + i\gamma} + \sum_{\chi \pmod q} \overline{c_\chi \sum_{\substack{0<\gamma<T \\ L(1/2+i\gamma,\chi)=0}} \frac{x^{i\gamma}}{\frac12 + i\gamma}} \Bigg) \\ &\quad= \lim_{T\to\infty} 2\,\mathrm{Re} \Bigg(\sum_{\chi \pmod q} c_\chi \sum_{\substack{0<\gamma<T \\ L(1/2+i\gamma,\chi)=0}} \frac{x^{i\gamma}}{\frac12 + i\gamma} \Bigg) \qquad (2.2.2) \\ &\quad= 2 \lim_{T\to\infty} \sum_{\chi \pmod q} |c_\chi| \,\mathrm{Re} \Bigg(\sum_{\substack{0<\gamma<T \\ L(1/2+i\gamma,\chi)=0}} \frac{e^{i\gamma \log x}\theta_{\chi,\gamma}}{\sqrt{\frac14 + \gamma^2}} \Bigg),\end{aligned}$$

where $\theta_{\chi,\gamma} = c_\chi|\frac12 + i\gamma|/|c_\chi|(\frac12 + i\gamma)$ is a complex number of modulus 1. The quantity $e^{i\gamma \log x}\theta_{\chi,\gamma}$ is uniformly distributed (as a function of $\log x$) on the unit circle as x tends to infinity, and hence its limiting logarithmic distribution is the same as the distribution of Z_γ. Since the various γ in each inner sum are linearly independent over the rationals by LI, the tuple $(e^{i\gamma\log x}\theta_{\chi,\gamma})_{0<\gamma<T}$ is uniformly distributed in the $N(T,\chi)$-dimensional torus by Kronecker's theorem. Therefore the limiting logarithmic distribution of the sum

$$\sum_{\substack{0<\gamma<T\\ L(1/2+i\gamma,\chi)=0}} \frac{e^{i\gamma\log x}\theta_{\chi,\gamma}}{\sqrt{\frac14+\gamma^2}}$$

is the same as the distribution of the random variable

$$\sum_{\substack{0<\gamma<T\\ L(1/2+i\gamma,\chi)=0}} \frac{Z_\gamma}{\sqrt{\frac14+\gamma^2}}.$$

Finally, the work of Rubinstein and Sarnak [**68**, Section 3.1] shows that the limiting logarithmic distribution of

$$\sum_{\chi\,(\mathrm{mod}\ q)} c_\chi E(x,\chi) = 2\lim_{T\to\infty} \sum_{\chi\,(\mathrm{mod}\ q)} |c_\chi|\,\mathrm{Re}\Bigg(\sum_{\substack{0<\gamma<T\\ L(1/2+i\gamma,\chi)=0}} \frac{e^{i\gamma\log x}\theta_{\chi,\gamma}}{\sqrt{\frac14+\gamma^2}}\Bigg)$$

is the same as the distribution of the random variable

$$\begin{aligned}\sum_{\chi\,(\mathrm{mod}\ q)} c_\chi E(x,\chi) &= 2\lim_{T\to\infty} \sum_{\chi\,(\mathrm{mod}\ q)} |c_\chi| \sum_{\substack{0<\gamma<T\\ L(1/2+i\gamma,\chi)=0}} \frac{X_\gamma}{\sqrt{\frac14+\gamma^2}} \\ &= 2\sum_{\chi\,(\mathrm{mod}\ q)} |c_\chi| \sum_{\substack{\gamma>0\\ L(1/2+i\gamma,\chi)=0}} \frac{X_\gamma}{\sqrt{\frac14+\gamma^2}},\end{aligned}$$

the convergence of this last limit being ensured by the fact that the X_γ are bounded and that each of the sums

$$\sum_{\substack{\gamma>0\\ L(1/2+i\gamma,\chi)=0}} \left(\frac{1}{\sqrt{\frac14+\gamma^2}}\right)^2 \le b(\chi)$$

is finite. This establishes the lemma. □

We shall have further occasion to change the indexing of sums, between over all γ and over only positive γ, in the same manner as in equations (2.2.1) and (2.2.2); henceforth we shall justify such changes "by the functional equation for Dirichlet L-functions" and omit the intermediate steps.

Definition 2.2.3. *For any relative prime integers* q *and* a*, define*

$$c(q;a) = -1 + \#\{x \ (\mathrm{mod}\ q)\colon x^2 \equiv a \ (\mathrm{mod}\ q)\}.$$

Note that $c(q;a)$ *takes only the values* -1 *and* $\rho(q)-1$*. Now, with* X_γ *as defined in Definition 2.2.2, define the random variable*

$$X_{q;a,b} = c(q,b) - c(q,a) + 2 \sum_{\chi \,(\mathrm{mod}\ q)} |\chi(b) - \chi(a)| \sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} \frac{X_\gamma}{\sqrt{\frac{1}{4}+\gamma^2}}.$$

Note that the expectation of the random variable $X_{q;a,b}$ *is either* $\pm\rho(q)$ *or 0, depending on the values of* $c(q,a)$ *and* $c(q,b)$*.* ♢

Definition 2.2.4. *With* $\pi(x;q,a) = \#\{p \le x\colon p \text{ prime}, p \equiv a \ (\mathrm{mod}\ q)\}$ *denoting the counting function of primes in the arithmetic progression* $a \ (\mathrm{mod}\ q)$*, define the normalized error term*

$$E(x;q,a) = \frac{\log x}{\sqrt{x}}\big(\phi(q)\pi(x;q,a) - \pi(x)\big).$$

♢

The next proposition characterizes the limiting logarithmic distribution of the difference of two of these normalized counting functions.

Proposition 2.2.2. *Assume GRH and LI. Let* a *and* b *be reduced residues modulo* q*. The limiting logarithmic distribution of* $E(x;q,a) - E(x;q,b)$ *is the same as the distribution of the random variable* $X_{q;a,b}$ *defined in Definition 2.2.3.*

Remark 2.2.1. *Since* $\delta(q;a,b)$ *is defined to be the logarithmic density of those real numbers* x *for which* $\pi(x;q,a) > \pi(x;q,b)$*, or equivalently for which* $E(x;q,a) > E(x;q,b)$*, we see that* $\delta(q;a,b)$ *equals the probability that* $X_{q;a,b}$ *is greater than 0. However, we never use this fact directly in the present paper, instead quoting from* **[22]** *a consequence of that fact in equation* (2.2.10) *below.*

PROOF. As is customary, define

$$\psi(x,\chi) = \sum_{n \le x} \chi(n)\Lambda(n)$$

A consequence of the explicit formula for $\psi(x,\chi)$ that arises from the analytic proof of the prime number theorem for arithmetic progressions ([**60**, Corollary

12.11] combined with [**60**, (12.12)]) is that for $\chi \neq \chi_0$,

$$\psi(x,\chi) = -\sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \frac{x^{1/2+i\gamma}}{\frac{1}{2}+i\gamma} + O(\log q \cdot \log x)$$

under the assumption of GRH. We also know [**68**, Lemma 2.1] that

$$E(x;q,a) = -c(q,a) + \sum_{\substack{\chi \pmod q \\ \chi\neq\chi_0}} \bar{\chi}(a)\frac{\psi(x,\chi)}{\sqrt{x}} + O_q\left(\frac{1}{\log x}\right). \tag{2.2.3}$$

Combining these last two equations with Definition 2.2.1 for $E(x,\chi)$, we obtain

$$E(x;q,a) = -c(q,a) - \sum_{\substack{\chi \pmod q \\ \chi\neq\chi_0}} \bar{\chi}(a)E(x,\chi) + O_q\left(\frac{1}{\log x}\right).$$

We therefore see that

$$\begin{aligned} E(x;q,a) - E(x;q,b) = c(q,b) - c(q,a) + \sum_{\chi \pmod q} (\bar{\chi}(b)-\bar{\chi}(a))E(x,\chi) \\ + O_q\left(\frac{1}{\log x}\right) \end{aligned}$$

(where we have added in the $\chi = \chi_0$ term for convenience). The error term tends to zero as x grows and thus doesn't affect the limiting distribution, and the constant $c(q,b)-c(q,a)$ is independent of x. Therefore, by Proposition 2.2.1, the limiting logarithmic distribution of $E(x;q,a) - E(x;q,b)$ is the same as the distribution of the random variable

$$c(q,b) - c(q,a) + 2\sum_{\chi \pmod q} |\bar{\chi}(b)-\bar{\chi}(a)| \sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} \frac{X_\gamma}{\sqrt{\frac{1}{4}+\gamma^2}}.$$

Since $|\bar{\chi}(b)-\bar{\chi}(a)| = |\chi(b)-\chi(a)|$, this last expression is exactly the random variable $X_{q;a,b}$ as claimed. □

To conclude this section, we calculate the variance of the random variable $X_{q;a,b}$.

Proposition 2.2.3. *Assume LI. Let* $\{c_\chi\colon \chi \pmod q\}$ *be a collection of complex numbers satisfying* $c_{\bar{\chi}} = \overline{c_\chi}$*. For any constant* μ*, the variance of the random variable*

$$\mu + 2\sum_{\chi \pmod q} c_\chi \sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} \frac{X_\gamma}{\sqrt{\frac{1}{4}+\gamma^2}} \tag{2.2.4}$$

equals $\sum_{\chi \pmod q} |c_\chi|^2 b(\chi)$, *where* $b(\chi)$ *was defined in Definition 2.1.2. In particular, the variance of the random variable* $X_{q;a,b}$ *defined in Definition 2.2.3 is equal to the quantity* $V(q;a,b)$ *defined in Definition 2.1.2.*

PROOF. The random variables $\{X_\gamma \colon \gamma > 0\}$ form an independent collection by definition; it is important to note that no single variable X_γ can correspond to multiple characters χ, due to the assumption of LI. The variance of the sum (2.2.4) is therefore simply the sum of the individual variances, that is,

$$\sigma^2\Bigg(2\sum_{\chi \pmod q} |c_\chi| \sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} \frac{X_\gamma}{\sqrt{\frac14+\gamma^2}}\Bigg) = 4\sum_{\chi \pmod q} |c_\chi|^2 \sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} \frac{\sigma^2(X_\gamma)}{\frac14+\gamma^2}.$$

The variance of any X_γ is $\frac12$, and so this last expression equals

$$\begin{aligned} 2\sum_{\chi \pmod q} |c_\chi|^2 &\sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} \frac{1}{\frac14+\gamma^2} \\ &= \sum_{\chi \pmod q} |c_\chi|^2 \sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} \frac{1}{\frac14+\gamma^2} + \sum_{\chi \pmod q} |c_\chi|^2 \sum_{\substack{\gamma<0 \\ L(1/2+i\gamma,\bar\chi)=0}} \frac{1}{\frac14+\gamma^2} \\ &= \sum_{\chi \pmod q} |c_\chi|^2 \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \frac{1}{\frac14+\gamma^2} = \sum_{\chi \pmod q} |c_\chi|^2 b(\chi) \end{aligned}$$

by the functional equation for Dirichlet L-functions. The fact that $V(q;a,b)$ is the variance of $X_{q;a,b}$ now follows directly from their definitions. □

2.2.2. Calculating the characteristic function

The characteristic function $\hat{X}_{q;a,b}(z)$ of the random variable $X_{q;a,b}$ will be extremely important to our analysis of the density $\delta(q;a,b)$. To derive the formula for this characteristic function, we begin by setting down some relevant facts about the standard Bessel function J_0 of order zero. Specifically, we collect in the following lemma some useful information about the power series coefficients λ_n for

$$\log J_0(z) = \sum_{n=0}^{\infty} \lambda_n z^n, \tag{2.2.5}$$

which is valid for $|z| \le \frac{12}{5}$ since J_0 has no zeros in a disk of radius slightly larger than $\frac{12}{5}$ centered at the origin.

Lemma 2.2.1. *Let the coefficients λ_n be defined in equation* (2.2.5). *Then :*

(1) $\lambda_n \ll \left(\frac{5}{12}\right)^n$ *uniformly for* $n \ge 0$ *;*

(2) $\lambda_0 = 0$ *and* $\lambda_{2m-1} = 0$ *for every* $m \ge 1$ *;*

(3) $\lambda_{2m} < 0$ *for every* $m \ge 1$ *;*

(4) λ_n *is a rational number for every* $n \ge 0$.

PROOF. The fact that $\log J_0$ is analytic in a disk of radius slightly larger than $\frac{12}{5}$ centered at the origin immediately implies part (a). Part (b) follows from the fact that J_0 is an even function with $J_0(0) = 1$. Next, J_0 has the product expansion [**77**, Section 15.41, equation (3)]

$$J_0(z) = \prod_{k=1}^{\infty} \left(1 - \frac{z^2}{z_k^2}\right),$$

where the z_k are the positive zeros of J_0. Taking logarithms of both sides and expanding each summand in a power series (valid for $|z| \le \frac{12}{5}$ as before) gives

$$\log J_0(z) = \sum_{k=1}^{\infty} \log\left(1 - \frac{z^2}{z_k^2}\right) = -\sum_{n=1}^{\infty} \frac{z^{2n}}{n} \sum_{k=1}^{\infty} \frac{1}{z_k^{2n}},$$

which shows that $\lambda_{2n} = -n^{-1} \sum_{k=1}^{\infty} z_k^{-2n}$ is negative, establishing part (c). Finally, the Bessel function $J_0(z) = \sum_{m=0}^{\infty} \left(-\frac{1}{4}\right)^m z^{2m}/(m!)^2$ itself has a power series with rational coefficients, as does $\log(1+z)$; therefore the composition $\log(1 + (J_0(z) - 1))$ also has rational coefficients, establishing part (d). □

Definition 2.2.5. *Let λ_n be defined in equation* (2.2.5). *For any distinct reduced residues* a *and* b (mod q), *define*

$$W_n(q; a, b) = \frac{2^{2n}|\lambda_{2n}|}{V(q; a, b)} \sum_{\chi \,(\mathrm{mod}\ q)} |\chi(a) - \chi(b)|^{2n} \sum_{\substack{\gamma > 0 \\ L(1/2+i\gamma,\chi)=0}} \frac{1}{(1/4 + \gamma^2)^n}, \quad (2.2.6)$$

where $V(q; a, b)$ *was defined in Definition 2.1.2, so that* $W_1(q; a, b) = \frac{1}{2}$ *for example.* ♢

In fact, $W_n(q;a,b)V(q;a,b)$ is (up to a constant factor depending on n) the $2n$th cumulant of $X(q;a,b)$, which explains why it will appear in the lower terms of the asymptotic formula. We have normalized by $V(q;a,b)$ so that the $W_n(q;a,b)$ depend upon q, a, and b in a bounded way :

Proposition 2.2.4. *We have* $W_n(q;a,b) \ll \left(\frac{10}{3}\right)^{2n}$ *uniformly for all integers* q *and all reduced residues* a *and* b (mod q).

PROOF. From Definition 2.2.5 and Lemma 2.2.1(a), we see that

$$\begin{aligned} W_n(q;a,b) &\ll \frac{2^{2n}}{V(q;a,b)}\left(\tfrac{5}{12}\right)^{2n} \sum_{\chi \,(\mathrm{mod}\ q)} |\chi(a)-\chi(b)|^{2n} \sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} \frac{1}{(1/4+\gamma^2)^n} \\ &\ll \frac{(5/6)^{2n}}{V(q;a,b)} \sum_{\chi \,(\mathrm{mod}\ q)} 2^{2n-2}|\chi(b)-\chi(a)|^2 \sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} \frac{4^{n-1}}{1/4+\gamma^2} \\ &= \left(\tfrac{5}{6}\right)^{2n} 2^{2n-2} 4^{n-1} \ll \left(\tfrac{10}{3}\right)^{2n}, \end{aligned}$$

as claimed. □

The following functions are necessary to write down the formula for the characteristic function $\hat{X}_{q;a,b}$.

Definition 2.2.6. *For any Dirichlet character* χ, *define*

$$F(z,\chi) = \prod_{\substack{\gamma>0 \\ L(\frac{1}{2}+i\gamma,\chi)=0}} J_0\left(\frac{2z}{\sqrt{\frac{1}{4}+\gamma^2}}\right).$$

Then define

$$\Phi_{q;a,b}(z) = \prod_{\chi \,(\mathrm{mod}\ q)} F\big(|\chi(a)-\chi(b)|z,\chi\big)$$

for any reduced residues a *and* b (mod q). *Note that* $|F(x,\chi)| \le 1$ *for all real numbers* x, *since the same is true of* J_0. ♢

The quantity $W_n(q;a,b)$ owes its existence to the following convenient expansion :

Proposition 2.2.5. *For any reduced residue classes* a *and* b (mod q),

$$\Phi_{q;a,b}(z) = \exp\left(-V(q;a,b)\sum_{m=1}^{\infty} W_m(q;a,b)z^{2m}\right)$$

for $|z| < \frac{3}{10}$. *In particular,*

$$\Phi_{q;a,b}(z) = e^{-V(q;a,b)z^2/2}\big(1 + O(V(q;a,b)z^4)\big)$$

for $|z| \le \min\{V(q;a,b)^{-1/4}, \frac{1}{4}\}$.

PROOF. Taking logarithms of both sides of the definition of $\Phi_{q;a,b}(z)$ in Definition 2.2.6 yields

$$\log \Phi_{q;a,b}(z) = \sum_{\chi \pmod q} \sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} \log J_0\bigg(\frac{2|\chi(a)-\chi(b)|z}{\sqrt{\frac{1}{4}+\gamma^2}}\bigg).$$

Since $|z| < \frac{3}{10}$, the argument of the logarithm of J_0 is at most $2 \cdot 2 \cdot \frac{3}{10}/\frac{1}{2} = \frac{12}{5}$, and so the power series expansion (2.2.5) converges absolutely, giving

$$\log \Phi_{q;a,b}(z) = \sum_{\chi \pmod q} \sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} \sum_{n=0}^{\infty} \lambda_n \bigg(\frac{2|\chi(a)-\chi(b)|z}{\sqrt{\frac{1}{4}+\gamma^2}}\bigg)^n.$$

By Lemma 2.2.1(b) only the terms $n = 2m$ with $m \ge 1$ survive, and by Lemma 2.2.1(c) we may replace λ_{2m} by $-|\lambda_{2m}|$. We thus obtain

$$\begin{aligned}\log \Phi_{q;a,b}(z) &= -\sum_{m=0}^{\infty} z^{2m} \cdot |\lambda_{2m}| 2^{2m} \sum_{\chi \pmod q} |\chi(a)-\chi(b)|^{2m} \sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} \frac{1}{(\frac{1}{4}+\gamma^2)^m} \\ &= -\sum_{m=1}^{\infty} V(q;a,b) W_m(q;a,b) z^{2m}\end{aligned}$$

for $|z| < \frac{3}{10}$, by Definition 2.2.5 for $W_m(q;a,b)$. This establishes the first assertion of the proposition.

By Proposition 2.2.4, we also have

$$\sum_{m=2}^{\infty} W_m(q;a,b) z^{2m} \ll \sum_{m=2}^{\infty} \big(\tfrac{10}{3}\big)^{2m} z^{2m} = \frac{(10/3)^4 z^4}{1-100z^2/9} \ll z^4 \tag{2.2.7}$$

uniformly for $|z| \le \frac{1}{4}$, say. Therefore by the first assertion of the proposition,

$$\begin{aligned}\Phi_{q;a,b}(z) &= \exp\left(-V(q;a,b)W_1(q;a,b)z^2\right)\times \\ &\qquad \exp\left(-V(q;a,b)\sum_{m=2}^{\infty} W_m(q;a,b)z^{2m}\right) \\ &= e^{-V(q;a,b)z^2/2}\exp\left(O(V(q;a,b)z^4)\right) \\ &= e^{-V(q;a,b)z^2/2}\big(1+O(V(q;a,b)z^4)\big)\end{aligned}$$

as long as $V(q;a,b)z^4 \le 1$. This establishes the second assertion of the proposition. □

All the tools are now in place to calculate the characteristic function $\hat{X}_{q;a,b}(z) = \mathbb{E}\big(e^{izX_{q;a,b}}\big)$.

Proposition 2.2.6. *For any reduced residue classes* a *and* b (mod q),

$$\hat{X}_{q;a,b}(z) = e^{iz(c(q,b)-c(q,a))}\Phi_{q;a,b}(z).$$

In particular,

$$\log \hat{X}_{q;a,b}(z) = i\big(c(q,a)-c(q,b)\big)z - \tfrac{1}{2}V(q;a,b)z^2 + O\big(V(q;a,b)z^4\big)$$

for $|z| \le \frac{1}{4}$.

Remark 2.2.2. *The first assertion of the proposition was shown by Feuerverger and Martin* [**22**] *by a slightly different method. Unfortunately an* i *in the exponential factor of* [**22**, equation (2-21)] *is missing, an omission that is repeated in the statement of* [**22**, Theorem 4].

PROOF. For a random variable X, define the cumulant-generating function

$$g_X(t) = \log \hat{X}(t) = \log \mathbb{E}(e^{itX})$$

to be the logarithm of the characteristic function of X. It is easy to see that $g_{\alpha X}(t) = g_X(\alpha t)$ for any constant α. Moreover, if X and Y are independent random variables, then $\mathbb{E}(e^{itX}e^{itY}) = \mathbb{E}(e^{itX})\mathbb{E}(e^{itY})$ and so $g_{X+Y}(t) = g_X(t) + g_Y(t)$. Note that if the random variable C is constant with value c, then $g_C(t) = itc$.

We can also calculate $g_{X_\gamma}(t)$ where X_γ was defined in Definition 2.2.2. Indeed, if Θ is a random variable uniformly distributed on the interval $[-\pi,\pi]$, then $Z_\gamma = e^{i\Theta}$ and thus $X_\gamma = \cos\Theta$, whence

$$g_{X_\gamma}(t) = \log\mathbb{E}\big(e^{it\cos\Theta}\big) = \log\bigg(\int_{-\pi}^{\pi} e^{it\cos\theta}\,\frac{d\theta}{2\pi}\bigg) = \log J_0(t),$$

where J_0 is the Bessel function of order zero [**1**, 9.1.21].

From Definition 2.2.3, the above observations yield

$$g_{X_{q;a,b}}(t) = it(c(q,b)-c(q,a)) + \sum_{\chi \ (\mathrm{mod}\ q)} \sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} g_{X_\gamma}\bigg(\frac{2|\chi(a)-\chi(b)|}{\sqrt{1/4+\gamma^2}}t\bigg);$$

in other words,

$$\begin{aligned}\log\hat{X}_{q;a,b}(t) &= it(c(q,b)-c(q,a)) + \sum_{\chi \ (\mathrm{mod}\ q)} \sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} \log J_0\bigg(\frac{2|\chi(a)-\chi(b)|}{\sqrt{1/4+\gamma^2}}t\bigg) \\ &= it(c(q,b)-c(q,a)) + \log\Phi_{q;a,b}(x) \qquad (2.2.8)\end{aligned}$$

according to Definition 2.2.6. Exponentiating both sides establishes the first assertion of the proposition. To establish the second assertion, we combine equation (2.2.8) with Proposition 2.2.5 to see that for $|z| \le \frac14$,

$$\begin{aligned}\log\hat{X}_{q;a,b}(t) &= it(c(q,b)-c(q,a)) - V(q;a,b)\sum_{m=1}^{\infty} W_m(q;a,b)z^{2m} \\ &= it(c(q,b)-c(q,a)) - \tfrac12 V(q;a,b)z^2 + O(V(q;a,b)z^4)\end{aligned}$$

by the estimate (2.2.7) and the fact that $W_1(q;a,b) = \frac12$. □

2.2.3. Bounds for the characteristic function

A formula (namely equation (2.2.10) below) is known that relates $\delta(q;a,b)$ to an integral involving $\Phi_{q;a,b}$. Using this formula to obtain explicit estimates for $\delta(q;a,b)$ requires explicit estimates upon $\Phi_{q;a,b}$; our first estimate shows that this function takes its largest values near 0.

Proposition 2.2.7. *Let $0 \le \kappa \le \frac{5}{24}$. For any reduced residue classes a and b (mod q), we have $|\Phi_{q;a,b}(t)| \le |\Phi_{q;a,b}(\kappa)|$ for all $t \ge \kappa$.*

PROOF. From Definition 2.2.6, it suffices to show that for any real number $\gamma > 0$,

$$\left|J_0\left(\frac{2|\chi(a)-\chi(b)|t}{\sqrt{1/4+\gamma^2}}\right)\right| \le \left|J_0\left(\frac{2|\chi(a)-\chi(b)|\kappa}{\sqrt{1/4+\gamma^2}}\right)\right| \tag{2.2.9}$$

for all $t \ge \kappa$. We use the facts that J_0 is a positive, decreasing function on the interval $[0, \frac{5}{3}]$ and that $J_0\big(\frac{5}{3}\big) \ge |J_0(x)|$ for all $x \ge \frac{5}{3}$. Since

$$0 \le \frac{2|\chi(a)-\chi(b)|\kappa}{\sqrt{1/4+\gamma^2}} \le \frac{2\cdot 2\cdot 5/24}{\sqrt{1/4}} = \frac{5}{3},$$

we see that J_0 is positive and decreasing on the interval

$$\left[\frac{2|\chi(a)-\chi(b)|\kappa}{\sqrt{1/4+\gamma^2}}, \frac{5}{3}\right].$$

Together with $J_0\big(\frac{5}{3}\big) \ge |J_0(x)|$ for all $x \ge \frac{5}{3}$, this establishes equation (2.2.9) and hence the lemma. □

Let $N(T,\chi)$ denote, as usual, the number of nontrivial zeros of $L(s,\chi)$ having imaginary part at most T in absolute value. Since the function $\Phi_{q;a,b}$ is a product indexed by these nontrivial zeros, we need to establish the following explicit estimates for $N(T,\chi)$. Although exact values for the constants in the results of this section are not needed for proving Theorem 2.1.1, they will become necessary in Section 2.5 when we explicitly calculate values and bounds for $\delta(q;a,b)$.

Proposition 2.2.8. *Let the nonprincipal character* $\chi \pmod q$ *be induced by* $\chi^* \pmod{q^*}$. *For any real number* $T \ge 1$,

$$N(T,\chi) \le \frac{T}{\pi}\log\frac{q^*T}{2\pi e} + 0.68884\log\frac{q^*T}{2\pi e} + 10.6035.$$

For $T \ge 100$,

$$N(T,\chi) \ge \frac{44T}{45\pi}\log\frac{q^*T}{2\pi e} - 10.551.$$

PROOF. We cite the following result of McCurley [**59**, Theorem 2.1] : for $T \ge 1$ and $\eta \in (0, 0.5]$,

$$\left|N(T,\chi) - \frac{T}{\pi}\log\frac{q^*T}{2\pi e}\right| < C_1\log q^*T + C_2,$$

with $C_1 = \frac{1+2\eta}{\pi \log 2}$ and $C_2 = .3058 - .268\eta + 4\frac{\log\zeta(1+\eta)}{\log 2} - 2\frac{\log\zeta(2+2\eta)}{\log 2} + \frac{2}{\pi}\frac{\log\zeta(\frac{3}{2}+2\eta)}{\log 2}$. (McCurley states his result for primitive nonprincipal characters, but since $L(s,\chi)$ and $L(s,\chi^*)$ have the same zeros inside the critical strip, the result holds for any nonprincipal character.) Taking $\eta = 0.25$, we obtain

$$\left| N(T,\chi) - \frac{T}{\pi}\log\frac{q^*T}{2\pi e} \right| < 0.68884\log q^*T + 8.64865 < 0.68884\log\frac{q^*T}{2\pi e} + 10.6035.$$

This inequality establishes the first assertion of the proposition. The inequality also implies that

$$N(T,\chi) > \frac{44T}{45\pi}\log\frac{q^*T}{2\pi e} + \left(\left(\frac{T}{45\pi} - .68884 \right) \log\frac{q^*T}{2\pi e} - 10.6035 \right);$$

the second assertion of the proposition follows upon calculating that the expression in parentheses is at least -10.551 when $T \ge 100$ (we know that $q^* \ge 3$ as there are no nonprincipal primitive characters modulo 1 or 2). □

The next two results establish an exponentially decreasing upper bound for $\Phi_{q;a,b}(t)$ when t is large.

Lemma 2.2.2. *For any nonprincipal character* $\chi \pmod q$, *we have* $|F(x,\chi)F(x,\bar\chi)| \le e^{-0.2725x}$ *for* $x \ge 200$.

PROOF. First note that

$$F(x,\bar\chi) = \prod_{\substack{\gamma>0 \\ L(1/2+i\gamma,\bar\chi)=0}} J_0\left(\frac{2x}{\sqrt{1/4+\gamma^2}}\right) = \prod_{\substack{\gamma<0 \\ L(1/2+i\gamma,\chi)=0}} J_0\left(\frac{2x}{\sqrt{1/4+(-\gamma)^2}}\right)$$

by the identity $L(s,\bar\chi) = \overline{L(\bar s,\chi)}$, and therefore

$$F(x,\chi)F(x,\bar\chi) = \prod_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} J_0\left(\frac{2x}{\sqrt{1/4+\gamma^2}}\right).$$

Using the bound [**68**, equation (4.5)]

$$|J_0(z)| \le \min\left\{1, \sqrt{\tfrac{2}{\pi|x|}}\right\},$$

we see that for $x \ge 1$,

$$|F(x,\chi)F(x,\bar\chi)| \le \prod_{\substack{-x/2<\gamma<x/2 \\ L(1/2+i\gamma,\chi)=0}} \left| J_0\left(\frac{2x}{\sqrt{1/4+\gamma^2}}\right)\right| \le \prod_{\substack{|\gamma|<x/2 \\ L(1/2+i\gamma,\chi)=0}} \frac{(1/4+\gamma^2)^{1/4}}{\sqrt{\pi x}}.$$

When $x \ge 1$ and $|\gamma| < x/2$, the factor $(1/4+\gamma^2)^{1/4}(\pi x)^{-1/2}$ never exceeds $1/2$. Therefore

$$|F(x,\chi)F(x,\bar\chi)| \le 2^{-N(x/2,\chi)} = \exp\big(-(\log 2)N(x/2,\chi)\big).$$

By Proposition 2.2.8, we thus have for $x \ge 200$

$$\begin{aligned}|F(x,\chi)F(x,\bar\chi)| &\le 2^{10.558}\exp\left(-\frac{22\log 2}{45\pi}x\log\frac{q^*x}{4\pi e}\right)\\ &\le \exp\left(-0.107866x\log\frac{3x}{4\pi e}+7.3183\right) \le e^{-0.2725x},\end{aligned}$$

as claimed. □

Proposition 2.2.9. *For any distinct reduced residue classes* a *and* $b \pmod q$ *such that* $(ab,q)=1$, *we have* $|\Phi_{q;a,b}(t)| \le e^{-0.0454\phi(q)t}$ *for* $t \ge 200$.

PROOF. We begin by noting that the orthogonality relations for Dirichlet characters imply that $\sum_{\chi \pmod q}|\chi(a)-\chi(b)|^2 = 2\phi(q)$ (as we show in Proposition 2.3.1 below). On the other hand, if S is the set of characters $\chi \pmod q$ such that $|\chi(a)-\chi(b)| \ge 1$, then

$$\sum_{\chi \pmod q}|\chi(a)-\chi(b)|^2 \le \sum_{\substack{\chi \pmod q\\ \chi\notin S}} 1 + \sum_{\chi\in S} 4 = \phi(q) - \#S + 4\#S.$$

Combining these two inequalities shows that $2\phi(q) \le \phi(q)+3\#S$, or equivalently $\#S \ge \frac13\phi(q)$. Note that clearly $\chi_0 \notin S$.

From Definition 2.2.6, we have

$$\begin{aligned}|\Phi_{q;a,b}(t)|^2 &= \prod_{\chi \pmod q}|F(|\chi(a)-\chi(b)|t,\chi)|^2\\ &= \prod_{\chi \pmod q}\big|F(|\chi(a)-\chi(b)|t,\chi)F(|\chi(a)-\chi(b)|t,\bar\chi)\big|,\end{aligned}$$

since every character appears once as χ and once as $\bar\chi$ in the product on the right-hand side. Since $|F(x,\chi)| \le 1$ for all real numbers x, we can restrict the product on the right-hand side to those characters $\chi \in S$ and still have a valid upper bound. For any $\chi \in S$, Lemma 2.2.2 gives us $|F(|\chi(a)-\chi(b)|t,\chi)F(|\chi(a)-\chi(b)|t,\bar\chi)| \le e^{-0.2725|\chi(a)-\chi(b)|t} \le e^{-0.2725t}$ for $t \ge 200$, whence

$$\begin{aligned}|\Phi_{q;a,b}(t)|^2 &\le \prod_{\chi\in S} \big|F(|\chi(a)-\chi(b)|t,\chi)F(|\chi(a)-\chi(b)|t,\bar\chi)\big| \\ &\le (e^{-0.2725t})^{\#S} \le (e^{-0.0454\phi(q)t})^2,\end{aligned}$$

which is equivalent to the assertion of the proposition. □

At this point we can establish the required formula for $\delta(q;a,b)$, in terms of a truncated integral involving $\Phi_{q;a,b}$, with an explicit error term. To more easily record the explicit bounds for error terms, we employ a variant of the O-notation : we write $A = \overline{O}(B)$ if $|A| \le B$ (as opposed to a constant times B) for all values of the parameters under consideration.

Proposition 2.2.10. *Assume GRH and LI. Let* a *and* b *be reduced residues* (mod q) *such that* a *is a nonsquare* (mod q) *and* b *is a square* (mod q). *If* $V(q;a,b) \ge 531$, *then*

$$\begin{aligned}\delta(q;a,b) = \frac{1}{2} + \frac{1}{2\pi}\int_{-V(q;a,b)^{-1/4}}^{V(q;a,b)^{-1/4}} &\frac{\sin\rho(q)x}{x}\Phi_{q;a,b}(x)\,dx \\ &+ \overline{O}\bigg(0.03506\frac{e^{-9.08\phi(q)}}{\phi(q)} + 63.67\rho(q)e^{-V(q;a,b)^{1/2}/2}\bigg),\end{aligned}$$

PROOF. Our starting point is the formula of Feuerverger and Martin [**22**, equation (2.57)], which is valid under the assumptions of GRH and LI :

$$\delta(q;a,b) = \frac{1}{2} - \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\sin((c(q,a)-c(q,b))x)}{x}\Phi_{q;a,b}(x)\,dx. \tag{2.2.10}$$

In the case where a is a nonsquare modulo q and b is a square modulo q, the constant $c(q,a)-c(q,b)$ equals $-\rho(q)$, so that

$$\delta(q;a,b) = \frac{1}{2} + \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\sin\rho(q)x}{x}\Phi_{q;a,b}(x)\,dx.$$

The part of the integral where $x \ge 200$ can be bounded using Proposition 2.2.9 :

$$\bigg|\frac{1}{2\pi}\int_{200}^{\infty} \frac{\sin\rho(q)x}{x}\Phi_{q;a,b}(x)\,dx\bigg| \le \frac{1}{400\pi}\int_{200}^{\infty} e^{-0.0454\phi(q)x}\,dx < \frac{0.01753e^{-9.08\phi(q)}}{\phi(q)}.$$

The part where $x \le -200$ is bounded by the same amount, and so

$$\delta(q;a,b) = \frac{1}{2} + \frac{1}{2\pi}\int_{-200}^{200} \frac{\sin\rho(q)x}{x}\Phi_{q;a,b}(x)\,dx + \overline{O}\biggl(0.03506\frac{e^{-9.08\phi(q)}}{\phi(q)}\biggr). \tag{2.2.11}$$

We now consider the part of the integral where $V(q;a,b)^{-1/4} \le x \le 200$. The hypothesis that $V(q;a,b) \ge 531$ implies that $V(q;a,b)^{-1/4} < \frac{5}{24}$, which allows us to make two simplifications. First, by Proposition 2.2.7, we know that $|\Phi_{q;a,b}(x)| \le \Phi_{q;a,b}(V(q;a,b)^{-1/4})$ for all x in the range under consideration. Second, by Proposition 2.2.5 we have

$$\Phi_{q;a,b}(x) = \exp\left(-V(q;a,b)\sum_{m=1}^{\infty} W_m(q;a,b)x^{2m}\right) \le e^{-V(q;a,b)x^2/2}$$

for all real numbers $|x| < \frac{3}{10}$, since $W_1(q;a,b) = \frac{1}{2}$ and all the $W_m(q;a,b)$ are nonnegative by Definition 2.2.5. Since $\frac{5}{24} < \frac{3}{10}$, we see that $|\Phi_{q;a,b}(x)| \le e^{-V(q;a,b)^{1/2}/2}$ for all x in the range under consideration. Noting also that

$$\bigl|\sin(\rho(q)x)/x\bigr| \le \rho(q)$$

for all real numbers x, we conclude that

$$\begin{aligned}\left|\int_{V(q;a,b)^{-1/4}}^{200} \frac{\sin\rho(q)x}{x}\Phi(x)dx\right| &\le \rho(q)\int_{V(q;a,b)^{-1/4}}^{200} e^{-V(q;a,b)^{1/2}/2}dx \\ &\le 200\rho(q)e^{-V(q;a,b)^{1/2}/2}.\end{aligned}$$

The part of the integral where $-200 \le x \le -V(q;a,b)^{-1/4}$ is bounded by the same amount, and thus equation (2.2.11) becomes

$$\begin{aligned}\delta(q;a,b) = \frac{1}{2} + \frac{1}{2\pi}&\int_{-V(q;a,b)^{-1/4}}^{V(q;a,b)^{-1/4}} \frac{\sin\rho(q)x}{x}\Phi_{q;a,b}(x)\,dx \\ &+ \overline{O}\biggl(0.03506\frac{e^{-9.08\phi(q)}}{\phi(q)} + \frac{200}{\pi}\rho(q)e^{-V(q;a,b)^{1/2}/2}\biggr),\end{aligned}$$

which establishes the proposition. □

2.2.4. Derivation of the asymptotic series

In this section we give the proof of Theorem 2.1.1. Our first step is to transform the conclusion of Proposition 2.2.10, which was phrased with a mind towards the explicit calculations in Section 2.5, into a form more convenient for our present purposes :

Lemma 2.2.3. *Assume GRH and LI. For any reduced residues* a *and* b (mod q) *such that* a *is a nonsquare* (mod q) *and* b *is a square* (mod q), *and for any fixed* $J > 0$,

$$\delta(q;a,b) = \frac{1}{2} + \frac{\rho(q)}{2\pi\sqrt{V(q;a,b)}}\times$$
$$\int_{-V(q;a,b)^{1/4}}^{V(q;a,b)^{1/4}} \frac{\sin\big(\rho(q)y/\sqrt{V(q;a,b)}\big)}{\rho(q)y/\sqrt{V(q;a,b)}}\Phi_{q;a,b}\bigg(\frac{y}{\sqrt{V(q;a,b)}}\bigg)\,dy + O_J\big(V(q;a,b)^{-J}\big).$$

PROOF. We make the change of variables $x = y/\sqrt{V(q;a,b)}$ in Proposition 2.2.10, obtaining

$$\delta(q;a,b) = \frac{1}{2} +$$
$$\frac{1}{2\pi}\int_{-V(q;a,b)^{1/4}}^{V(q;a,b)^{1/4}} \frac{\sin\big(\rho(q)y/\sqrt{V(q;a,b)}\big)}{y/\sqrt{V(q;a,b)}}\Phi_{q;a,b}\bigg(\frac{y}{\sqrt{V(q;a,b)}}\bigg)\frac{dy}{\sqrt{V(q;a,b)}}$$
$$+\,\overline{O}\bigg(0.06217\frac{e^{-5.12\phi(q)}}{\phi(q)} + 63.67\rho(q)e^{-V(q;a,b)^{1/2}/2}\bigg),$$

the main terms of which are exactly what we want. The lemma then follows from the estimates

$$e^{-5.12\phi(q)} \ll_J V(q;a,b)^{-J} \quad \text{and} \quad \rho(q)e^{-V(q;a,b)^{1/2}/2} \ll_J V(q;a,b)^{-J}$$

for any fixed constant J : these estimates hold because $V(q;a,b) \sim 2\phi(q)\log q$ by Proposition 2.3.4, while the standard lower bound $\phi(q) \gg q/\log\log q$ follows from equation (2.5.19). □

We will soon be expanding most of the integrand in Lemma 2.2.3 into a power series ; the following definition and lemma treat the integrals that so arise.

Definition 2.2.7. *For any nonnegative integer* k, *define* $(2k-1)!! = (2k-1)(2k-3)\cdots 3\cdot 1$, *where we make the convention that* $(-1)!! = 1$. *Also, for any nonnegative* *integer* k *and any positive real number* B*, define*

$$M_k(B) = \int_{-B}^{B} y^{2k} e^{-y^2/2}\, dy.$$

◊

Lemma 2.2.4. *Let* J *and* B *be positive real numbers. For any nonnegative integer* k*, we have* $M_k(B) = (2k-1)!!\sqrt{2\pi} + O_{k,J}\big(B^{-J}\big)$.

PROOF. We proceed by induction on k. In the case $k = 0$, we have

$$\begin{aligned} M_0(B) = \int_{-B}^{B} e^{-y^2/2}\, dy &= \int_{-\infty}^{\infty} e^{-y^2/2}\, dy - 2\int_{B}^{\infty} e^{-y^2/2}\, dy \\ &= \sqrt{2\pi} + O\bigg(\int_{B}^{\infty} e^{-By/2}\, dy \bigg) \\ &= \sqrt{2\pi} + O\bigg(\frac{2}{B} e^{-B^2/2} \bigg) = \sqrt{2\pi} + O_J(B^{-J}) \end{aligned}$$

as required. On the other hand, for $k \geq 1$ we can use integration by parts to obtain

$$\begin{aligned} M_k(B) &= \int_{-B}^{B} y^{2k-1} \cdot y e^{-y^2/2}\, dy \\ &= -y^{2k-1} e^{-y^2/2} \Big|_{-B}^{B} + (2k-1) \int_{-B}^{B} y^{2k-2} e^{-y^2/2}\, dy \\ &= O\big(B^{2k-1} e^{-B^2/2}\big) + (2k-1) M_{k-1}(B). \end{aligned}$$

Since the error term $B^{2k-1}e^{-B^2/2}$ is indeed $O_{k,J}\big(B^{-J}\big)$, the lemma follows from the inductive hypothesis for $M_{k-1}(B)$. □

The following familiar power series expansions can be truncated with reasonable error terms :

Lemma 2.2.5. *Let* K *be a nonnegative integer and* $C > 1$ *a real number. Uniformly for* $|z| \leq C$*, we have the series expansions*

$$e^z = \sum_{j=0}^{K} \frac{z^j}{j!} + O_{C,K}(|z|^{K+1});$$

$$\frac{\sin z}{z} = \sum_{j=0}^{K} (-1)^j \frac{z^{2j}}{(2j+1)!} + O_{C,K}(|z|^{2(K+1)}).$$

PROOF. The Taylor series for e^z, valid for all complex numbers z, can be written as

$$e^z = \sum_{j=0}^{K} \frac{z^j}{j!} + z^{K+1} \sum_{j=0}^{\infty} \frac{z^j}{(j+K+1)!}.$$

The function $\sum_{j=0}^{\infty} z^j/(j+K+1)!$ converges for all complex numbers z and hence represents an entire function ; in particular, it is continuous and hence bounded in the disc $|z| \le C$. This establishes the first assertion of the lemma, and the second assertion is proved in a similar fashion. □

Everything we need to prove Theorem 2.1.1 is now in place, once we give the definition of the constants $s_{q;a,b}(\ell, j)$ that appear in its statement :

Definition 2.2.8. *For any reduced residues* a *and* b (mod q), *and any positive integers* $j \le \ell$, *define*

$$s_{q;a,b}(\ell, j) = \frac{(-1)^j}{(2j+1)!} \sum_{i_2+2i_3+\cdots+\ell i_{\ell+1}=\ell-j} \!\!\!\cdots\!\!\! \sum \big(2(\ell+i_2+\cdots+i_{\ell+1})-1\big)!! \times \prod_{k=2}^{\ell+1} \frac{(-W_k(q;a,b))^{i_k}}{i_k!},$$

where the indices $i_2, \dots, i_{\ell+1}$ *take all nonnegative integer values that satisfy the constraint* $i_2 + 2i_3 + \cdots + \ell i_{\ell+1} = \ell - j$. *Note that* $s_{q;a,b}(0,0) = 1$ *always. Since* $W_k(q;a,b) \ll \left(\frac{10}{3}\right)^k$ *by Proposition 2.2.4, we see that* $s_{q;a,b}(\ell, j)$ *is bounded in absolute value by some (combinatorially complicated) function of* ℓ *uniformly in* q, a, *and* b *(and uniformly in* j *as well, since there are only finitely many possibilities* $\{0, 1, \dots, \ell\}$ *for* j*).* ♢

PROOF OF THEOREM 2.1.1. To lighten the notation in this proof, we temporarily write ρ for $\rho(q)$, δ for $\delta(q;a,b)$, V for $V(q;a,b)$, and W_k for $W_k(q;a,b)$. We also allow all O-constants to depend on K. Since δ is bounded, the theorem is trivially true when V is bounded, since the error term is at least as large as any other term in that case ; therefore we may assume that V is sufficiently large. For later usage in this proof, we note that $\rho \ll V^{1/4}$, which follows amply from the bound $\rho \ll_\varepsilon q^\varepsilon$ mentioned in Definition 2.1.1 and the asymptotic formula $V \sim 2\phi(q) \log q$ proved in Proposition 2.3.4.

We begin by noting that from Proposition 2.2.5,

$$\Phi_{q;a,b}(x) = \exp\left(-V\sum_{k=1}^{\infty} W_k x^{2k}\right) = \exp\left(-V\sum_{k=1}^{K+1} W_k x^{2k} + O(Vx^{2(K+2)})\right)$$
$$= \exp\left(-V\sum_{k=1}^{K+1} W_k x^{2k}\right)\left(1 + O(Vx^{2(K+2)})\right) \tag{2.2.12}$$

uniformly for all $|x| \le \min(\frac{1}{4}, V^{-1/4})$, where the second equality follows from the upper bound given in Proposition 2.2.4. Inserting this formula into the expression for $\delta(q;a,b)$ from Lemma 2.2.3, applied with $J = K+2$, gives

$$\delta = \frac{1}{2} + \frac{\rho}{2\pi\sqrt{V}}\int_{-V^{1/4}}^{V^{1/4}} \frac{\sin(\rho y/\sqrt{V})}{\rho y/\sqrt{V}} \exp\left(-\sum_{k=1}^{K+1}\frac{W_k y^{2k}}{V^{k-1}}\right)\left(1 + O\!\left(\frac{y^{2(K+2)}}{V^{K+1}}\right)\right) dy$$
$$+ O\big(V^{-K-2}\big).$$

This use of equation (2.2.12) is justified because the argument $y/\sqrt{V}$ of $\Phi_{q;a,b}$ in the integral in Lemma 2.2.3 is at most $V^{1/4}/\sqrt{V} \le \frac{1}{4}$, by the assumption that V is sufficiently large. To simplify the error term in the integral, we ignore all of the factors in the integrand (which are bounded by 1 in absolute value) except for the $k = 1$ term, in which $W_1 = \frac{1}{2}$, to derive the upper bound

$$\int_{-V^{1/4}}^{V^{1/4}} \frac{\sin(\rho y/\sqrt{V})}{\rho y/\sqrt{V}} \exp\left(-\sum_{k=1}^{K+1}\frac{W_k y^{2k}}{V^{k-1}}\right)\frac{y^{2(K+2)}}{V^{K+1}}\, dy$$
$$\ll \frac{1}{V^{K+1}}\int_{-\infty}^{\infty} e^{-y^2/2} y^{2K+4}\, dy \ll_K \frac{1}{V^{K+1}}.$$

Therefore

$$\delta = \frac{1}{2} + \frac{\rho}{2\pi\sqrt{V}}\int_{-V^{1/4}}^{V^{1/4}} \frac{\sin(\rho y/\sqrt{V})}{\rho y/\sqrt{V}} \exp\left(-\sum_{k=1}^{K+1}\frac{W_k y^{2k}}{V^{k-1}}\right) dy + O\!\left(\frac{\rho}{V^{K+3/2}}\right). \tag{2.2.13}$$

The integrand in equation (2.2.13) is the product of $K+2$ functions, namely $K+1$ exponential factors and a factor involving the function $(\sin z)/z$. Our plan is to keep the first exponential function as it is and expand the other factors into their power series at the origin. Note that the argument of the kth exponential factor is at most $W_k V^{1-k/2}$ in absolute value, which is bounded (by a constant depending on K) for all $k \geq 2$ by Proposition 2.2.4. Similarly, the argument of the function $(\sin z)/z$ is bounded by $\rho V^{1/4}/\sqrt{V} \ll 1$. Therefore the expansion of all of these factors, excepting the exponential factor corresponding to $k = 1$, into their power series is legitimate in the range of integration.

Specifically, we have the two identities

$$\sum_{j=0}^{K} \frac{(-1)^j}{(2j+1)!} \frac{(\rho y)^{2j}}{V^j} = \frac{\sin(\rho y/\sqrt{V})}{\rho y/\sqrt{V}} + O\bigg(\frac{(\rho y)^{2(K+1)}}{V^{K+1}} \bigg);$$

$$\begin{aligned} \sum_{i_k=0}^{K} \frac{(-1)^{i_k}}{i_k!} \left(\frac{W_k y^{2k}}{V^{k-1}} \right)^{i_k} &= \exp\left(-\frac{W_k y^{2k}}{V^{k-1}} \right) + O\bigg(\left(\frac{W_k y^{2k}}{V^{k-1}} \right)^{K+1} \bigg) \\ &= \exp\left(-\frac{W_k y^{2k}}{V^{k-1}} \right) + O\bigg(\frac{y^{2k(K+1)}}{V^{K+1}} \bigg), \end{aligned}$$

where the error terms are justified by Lemma 2.2.5 ; in the last equality we have used Proposition 2.2.4 to ignore the contribution of the factor W_k to the error term (since the O-constant may depend on K). From these identities, we deduce that

$$\begin{aligned} &\bigg(\sum_{j=0}^{K} \frac{(-1)^j}{(2j+1)!} \frac{(\rho y)^{2j}}{V^j} \bigg) e^{-y^2/2} \prod_{k=2}^{K+1} \bigg(\sum_{i_k=0}^{K} \frac{(-1)^{i_k}}{i_k!} \left(\frac{W_k y^{2k}}{V^{k-1}} \right)^{i_k} \bigg) \\ &= \bigg(\frac{\sin(\rho y/\sqrt{V})}{\rho y/\sqrt{V}} + O\bigg(\frac{(\rho y)^{2(K+1)}}{V^{K+1}} \bigg) \bigg) e^{-y^2/2} \times \\ &\qquad \prod_{k=2}^{K+1} \bigg(\exp\left(-\frac{W_k y^{2k}}{V^{k-1}} \right) + O\bigg(\left(\frac{W_k y^{2k}}{V^{k-1}} \right)^{K+1} \bigg) \bigg) \\ &= \frac{\sin(\rho y/\sqrt{V})}{\rho y/\sqrt{V}} \prod_{k=1}^{K+1} \exp\left(-\frac{W_k y^{2k}}{V^{k-1}} \right) + O\bigg(y^{(K+2)(K+1)^2} e^{-y^2/2} \frac{\rho^{2K+2}}{V^{K+1}} \bigg). \end{aligned}$$

(The computation of the error term is simplified by the fact that all the main terms on the right-hand side are at most 1 in absolute value, so that we need only figure out the largest powers of y and ρ, and the smallest power of V, that can be obtained by the cross terms.)

Substituting this identity into equation (2.2.13) yields

$$\begin{aligned}
\delta = \frac{1}{2} + {} & \\
\frac{\rho}{2\pi\sqrt{V}} & \int_{-V^{1/4}}^{V^{1/4}} \left(\sum_{j=0}^{K} \frac{(-1)^j}{(2j+1)!} \frac{(\rho y)^{2j}}{V^j} \right) e^{-y^2/2} \prod_{k=2}^{K+1} \left(\sum_{i_k=0}^{K} \frac{(-1)^{i_k}}{i_k!} \left(\frac{W_k y^{2k}}{V^{k-1}} \right)^{i_k} \right) dy \\
& \qquad + O\left(\frac{\rho}{\sqrt{V}} \int_{-\infty}^{\infty} y^{(K+2)(K+1)^2} e^{-y^2/2} \frac{\rho^{2K+2}}{V^{K+1}} \, dy + \frac{\rho}{V^{K+3/2}} \right) \\
= \frac{1}{2} + {} & \frac{\rho}{2\pi\sqrt{V}} \sum_{j=0}^{K} \sum_{i_2=0}^{K} \cdots \sum_{i_{K+1}=0}^{K} \left(\frac{(-1)^j}{(2j+1)!} \frac{\rho^{2j}}{V^j} \right. \\
& \quad \left. \times \prod_{k=2}^{K+1} \frac{1}{i_k!} \left(\frac{-W_k}{V^{k-1}} \right)^{i_k} M_{j+2i_2+\cdots+(K+1)i_{K+1}} \big(V^{1/4} \big) \right) + O\left(\frac{\rho^{2K+3}}{V^{K+3/2}} \right),
\end{aligned}$$

where M was defined in Definition 2.2.7. Invoking Lemma 2.2.4 and then collecting the summands according to the power $\ell = j + i_1 + 2i_2 + \cdots + Ki_{K+1}$ of V in the denominator, we obtain

$$\begin{aligned}
\delta = \frac{1}{2} + {} & \frac{\rho}{\sqrt{2\pi V}} \sum_{j=0}^{K} \sum_{i_2=0}^{K} \cdots \sum_{i_{K+1}=0}^{K} \left(\frac{(-1)^j}{(2j+1)!} \frac{\rho^{2j}}{V^j} \prod_{k=2}^{K+1} \frac{1}{i_k!} \left(\frac{-W_k}{V^{k-1}} \right)^{i_k} \right. \\
& \quad \left. \times \Big((2(j + 2i_2 + \cdots + (K+1)i_{K+1}) - 1)!! + O\big(V^{-(K+1)} \big) \Big) \right) + O\left(\frac{\rho^{2K+3}}{V^{K+3/2}} \right) \\
= \frac{1}{2} + {} & \frac{\rho}{\sqrt{2\pi V}} \sum_{\ell=0}^{K(1+K(K+1)/2)} \frac{1}{V^\ell} \sum_{j=0}^{K} \frac{(-1)^j \rho^{2j}}{(2j+1)!} \sum_{\substack{i_2=0 \\ i_2+2i_3+\cdots+Ki_{K+1}=\ell-j}}^{K} \cdots \sum_{i_{K+1}=0}^{K} \left(\prod_{k=2}^{K+1} \frac{(-W_k)^{i_k}}{i_k!} \right. \\
& \quad \left. \times \big(2(\ell + i_2 + \cdots + i_{K+1}) - 1 \big)!! \right) + O\left(\frac{\rho^{2K+3}}{V^{K+3/2}} \right), \qquad (2.2.14)
\end{aligned}$$

where we have subsumed the first error term into the second with the help of Proposition 2.2.4.

The proof of Theorem 2.1.1 is actually now complete, although it takes a moment to recognize it. For $0 \le \ell \le K$, the values of j that contribute to the sum are $0 \le j \le \ell$, since $\ell - j$ must be a sum of nonnegative numbers due to the condition of summation of the inner sum. In particular, all possible values of j and the i_k are represented in the sum, and the upper bound of K for these variables is unnecessary. We therefore see that the coefficient of $\rho^{2j} V^{-\ell}$ on the right-hand side of equation (2.2.14) matches Definition 2.2.8 for $s_{q;a,b}(\ell, j)$. On the other hand, for each of the finitely many larger values of ℓ, the ℓth summand is bounded above by $\rho^{2K}V^{-K-1}$ times some constant depending only on K (again we have used Proposition 2.2.4 to bound the quantities W_k uniformly), which is smaller than the indicated error term once the leading factor $\rho/\sqrt{2\pi V}$ is taken into account. □

2.3. Analysis of the variance $V(q;a,b)$

In this section we prove Theorems 2.1.2 and 2.1.3, as well as discussing related results to which our methods apply. We begin by establishing some arithmetic identities involving Dirichlet characters and their conductors in Section 2.3.1. Using these identities and a classical formula for $b(\chi)$, we complete the proof of Theorem 2.1.2 in Section 2.3.2. The linear combination of values $\frac{L'}{L}(1,\chi)$ that defines $M^*(q;a,b)$ can be converted into an asymptotic formula involving the von Mangoldt Λ-function, as we show in Section 2.3.3, and in this way we establish Theorem 2.1.3.

Our analysis to this point has the interesting consequence that the densities $\delta(q;a,b)$ can be evaluated extremely precisely using only arithmetic content, that is, arithmetic on rational numbers (including multiplicative functions of integers) and logarithms of integers; we explain this consequence in Section 2.3.4. Next, we show in Section 2.3.5 that the limiting logarithmic distributions of the differences $E(x;q,a)-E(x;q,b)$ obey a central limit theorem as q tends to infinity. Finally, we explain in Section 2.3.6 how our analysis can be modified to apply to the race between the aggregate counting functions $\pi(x;q,N) = \#\{p \le x\colon p$ is a quadratic nonresidue (mod q)$\}$ and $\pi(x;q,R) = \#\{p \le x\colon p$ is a quadratic residue (mod q)$\}$.

2.3.1. Arithmetic sums over characters

We begin by establishing some preliminary arithmetic identities that will be needed in later proofs.

Proposition 2.3.1. *Let* a *and* b *be distinct reduced residue classes* (mod q). *Then*

$$\sum_{\chi \,(\mathrm{mod}\ q)} |\chi(a)-\chi(b)|^2 = 2\phi(q),$$

while for any reduced residue $c \not\equiv 1 \pmod q$ *we have*

$$\sum_{\chi \,(\mathrm{mod}\ q)} |\chi(a)-\chi(b)|^2\chi(c) = -\phi(q)\big(\iota_q(cab^{-1}) + \iota_q(cba^{-1})\big),$$

where ι_q *is defined in Definition 2.1.3.*

PROOF. These sums are easy to evaluate using the orthogonality relation [**60**, Corollary 4.5]

$$\sum_{\chi \,(\mathrm{mod}\ q)} \chi(m) = \begin{cases} \phi(q), & \text{if } m \equiv 1 \pmod q \\ 0, & \text{if } m \not\equiv 1 \pmod q \end{cases} = \phi(q)\iota_q(m). \tag{2.3.1}$$

We have

$$\begin{aligned} \sum_{\chi \,(\mathrm{mod}\ q)} |\chi(a)-\chi(b)|^2 &= \sum_{\chi \,(\mathrm{mod}\ q)} (2-\chi(a)\overline{\chi(b)} - \chi(b)\overline{\chi(a)}) \\ &= \sum_{\chi \,(\mathrm{mod}\ q)} 2 - \sum_{\chi \,(\mathrm{mod}\ q)} \chi(ab^{-1}) - \sum_{\chi \,(\mathrm{mod}\ q)} \chi(ba^{-1}) \\ &= 2\phi(q) + 0 + 0, \end{aligned}$$

since $a \not\equiv b \pmod q$. Similarly,

$$\begin{aligned} \sum_{\chi \,(\mathrm{mod}\ q)} |\chi(a)-\chi(b)|^2\chi(c) &= \sum_{\chi \,(\mathrm{mod}\ q)} (2-\chi(a)\overline{\chi(b)} - \chi(b)\overline{\chi(a)})\chi(c) \\ &= \sum_{\chi \,(\mathrm{mod}\ q)} 2\chi(c) - \sum_{\chi \,(\mathrm{mod}\ q)} \chi(cab^{-1}) - \sum_{\chi \,(\mathrm{mod}\ q)} \chi(cba^{-1}) \\ &= 0 - \phi(q)\big(\iota_q(cab^{-1}) + \iota_q(cba^{-1})\big). \end{aligned}$$

□

The results in the next two lemmas were discovered independently by Vorhauer (see [**60**, Section 9.1, problem 8]).

Lemma 2.3.1. *For any positive integer* q*, we have*

$$\sum_{d|q} \Lambda(q/d)\phi(d) = \phi(q) \sum_{p|q} \frac{\log p}{p-1},$$

while for any proper divisor s *of* q *we have*

$$\sum_{d|s} \Lambda(q/d)\phi(d) = \phi(q)\frac{\Lambda(q/s)}{\phi(q/s)}.$$

PROOF. For the first identity, we group together the contributions from the divisors d such that q/d is a power of a particular prime factor p of q. If $p^r \parallel q$, write $q = mp^r$, so that $p \nmid m$. We get a contribution to the sum only when $d = mp^{r-k}$ for some $1 \le k \le r$. Therefore

$$\sum_{d|q} \Lambda(q/d)\phi(d) = \sum_{p^r \parallel q} \sum_{k=1}^{r} \Lambda(p^k)\phi(mp^{r-k}) = \sum_{p^r \parallel q} \phi(m) \log p \sum_{k=1}^{r} \phi(p^{r-k}).$$

Since $\sum_{a|b} \phi(a) = b$ for any positive integer b, the inner sum is exactly p^{r-1}. Noting that $\phi(m) = \phi(q)/\phi(p^r)$ since $p \nmid n$, we obtain

$$\sum_{d|q} \Lambda(q/d)\phi(d) = \sum_{p^r \parallel q} \frac{\phi(q)}{\phi(p^r)} p^{r-1} \log p = \phi(q) \sum_{p|q} \frac{\log p}{p-1}$$

as claimed.

We turn now to the second identity. If q/s has at least two distinct prime factors, then so will q/d for every divisor d of s, and hence all of the $\Lambda(q/d)$ terms will be 0. Therefore the entire sum equals 0, which is consistent with the claimed identity as $R_q(s) = 0$ as well in this case. Therefore we need only consider the case where q/s equals a prime power p^t.

Again write $q = mp^r$ with $p \nmid m$. Since $s = q/p^t = mp^{t-r}$, the only terms that contribute to the sum are $d = mp^{r-k}$ for $t \le k \le r$. By a similar calculation as before,

$$\begin{aligned}\sum_{d|s} \Lambda(q/d)\phi(d) &= \sum_{k=t}^{r} \Lambda(p^k)\phi(mp^{r-k}) = \phi(m) \log p \sum_{k=t}^{r} \phi(p^{r-k}) \\ &= \frac{\phi(q)}{\phi(p^r)} p^{r-t} \log p = \phi(q)\frac{\log p}{p^{t-1}(p-1)} = \phi(q)\frac{\Lambda(q/s)}{\phi(q/s)},\end{aligned}$$

since $q/s = p^t$. This establishes the second identity. □

Recall that χ^* denotes the primitive character that induces χ and that q^* denotes the conductor of χ^*.

Proposition 2.3.2. *For any positive integer* q*,*

$$\sum_{\chi \,(\mathrm{mod}\ q)} \log q^* = \phi(q)\left(\log q - \sum_{p|q} \frac{\log p}{p-1}\right),$$

while if $a \not\equiv 1 \pmod q$ *is a reduced residue,*

$$\sum_{\chi \,(\mathrm{mod}\ q)} \chi(a) \log q^* = -\phi(q)\frac{\Lambda(q/(q,a-1))}{\phi(q/(q,a-1))}.$$

PROOF. First we show that

$$\sum_{\chi \,(\mathrm{mod}\ q)} \chi(a) \log q^* = \log q \sum_{\chi \,(\mathrm{mod}\ q)} \chi(a) - \sum_{d|q} \Lambda(q/d) \sum_{\chi \,(\mathrm{mod}\ d)} \chi(a) \qquad (2.3.2)$$

for any reduced residue $a \pmod q$. Given a character $\chi \pmod q$ and a divisor d of q, the character χ is induced by a character $(\mathrm{mod}\ d)$ if and only if d is a multiple of q^*. Therefore

$$\sum_{d|q} \Lambda(q/d) \sum_{\chi \,(\mathrm{mod}\ d)} \chi(a) = \sum_{\chi \,(\mathrm{mod}\ q)} \chi(a) \sum_{\substack{d|q \\ q^*|d}} \Lambda(q/d).$$

Making the change of variables $c = q/d$, this identity becomes

$$\begin{aligned} \sum_{d|q} \Lambda(q/d) \sum_{\chi \,(\mathrm{mod}\ d)} \chi(a) &= \sum_{\chi \,(\mathrm{mod}\ q)} \chi(a) \sum_{c|q/q^*} \Lambda(c) \\ &= \sum_{\chi \,(\mathrm{mod}\ q)} \chi(a) \log \tfrac{q}{q^*} = \log q \sum_{\chi \,(\mathrm{mod}\ q)} \chi(a) - \sum_{\chi \,(\mathrm{mod}\ q)} \chi(a) \log q^*, \end{aligned}$$

which verifies equation (2.3.2).

If $a \equiv 1 \pmod q$, then equation (2.3.2) becomes

$$\begin{aligned} \sum_{\chi \,(\mathrm{mod}\ q)} \log q^* &= \log q \sum_{\chi \,(\mathrm{mod}\ q)} 1 - \sum_{d|q} \Lambda(q/d) \sum_{\chi \,(\mathrm{mod}\ d)} 1 \\ &= \phi(q) \log q - \sum_{d|q} \Lambda(q/d)\phi(d) = \phi(q) \log q - \phi(q) \sum_{p|q} \frac{\log p}{p-1} \end{aligned}$$

by Lemma 2.3.1, establishing the first assertion of the lemma. If on the other hand $a \not\equiv 1 \pmod q$, then applying the orthogonality relation (2.3.1) to equation (2.3.2) yields

$$\begin{aligned}\sum_{\chi \,(\mathrm{mod}\ q)} \chi(a) \log q^* &= 0 - \sum_{d\mid q} \Lambda(q/d)\phi(d)\iota_d(a) \\ &= - \sum_{d\mid(q,a-1)} \Lambda(q/d)\phi(d) = -\phi(q)\frac{\Lambda(q/(q,a-1))}{\phi(q/(q,a-1))}\end{aligned}$$

by Lemma 2.3.1 again, establishing the second assertion of the lemma. □

Finally we record a proposition that involves values of both primitive characters and characters induced by them.

Proposition 2.3.3. *Let p be a prime and e a positive integer, and let r be a reduced residue* (mod q). *If $p \nmid q$, then*

$$\sum_{\chi \,(\mathrm{mod}\ q)} \chi(r)\big(\chi^*(p^e) - \chi(p^e)\big) = 0.$$

On the other hand, if $p \mid q$ then

$$\sum_{\chi \,(\mathrm{mod}\ q)} \chi(r)\big(\chi^*(p^e) - \chi(p^e)\big) = \begin{cases} \phi\big(q/p^\nu\big), & \textit{if } rp^e \equiv 1 \ (\mathrm{mod}\ q/p^\nu), \\ 0, & \textit{otherwise}, \end{cases}$$

where $\nu \geq 1$ is the integer such that $p^\nu \parallel q$.

PROOF. The first assertion is trivial : if $p \nmid q$ then $\chi^*(p^e) = \chi(p^e)$ for every character χ (mod q). If $p \mid q$, then $\chi(p^e) = 0$ for every χ, and so

$$\sum_{\chi \,(\mathrm{mod}\ q)} \chi(r)\big(\chi^*(p^e) - \chi(p^e)\big) = \sum_{\chi \,(\mathrm{mod}\ q)} \chi(r)\chi^*(p^e) = \sum_{\chi \,(\mathrm{mod}\ q)} \chi^*(rp^e)$$

since $\chi(r) = \chi^*(r)$ for every χ (mod q) due to the hypothesis that $(r,q) = 1$. Also, we have $\chi^*(p^e) = 0$ for any character χ such that $p \mid q^*$, and so

$$\sum_{\chi \,(\mathrm{mod}\ q)} \chi^*(rp^e) = \sum_{\substack{\chi \,(\mathrm{mod}\ q) \\ q^* \mid q/p^\nu}} \chi^*(rp^e) = \sum_{\chi \,(\mathrm{mod}\ q/p^\nu)} \chi(rp^e),$$

since $(p^e, q/p^\nu) = 1$. The second assertion now follows from the orthogonality relation (2.3.1). □

2.3.2. A formula for the variance

Recall that $b(\chi)$ was defined in Definition 2.1.2; we record a classical formula for $b(\chi)$ in the next lemma, after which we will be able to prove Theorem 2.1.2.

Lemma 2.3.2. *Assume GRH. Let* $q \geq 3$*, and let* χ *be any nonprincipal character modulo* q*. Then*

$$b(\chi) = \log \frac{q^*}{\pi} - \gamma_0 - (1 + \chi(-1)) \log 2 + 2 \operatorname{Re} \frac{L'(1,\chi^*)}{L(1,\chi^*)}.$$

Proof. Since the zeros of $L(s,\chi)$ and $L(s,\chi^*)$ on the line $\operatorname{Re} z = \frac{1}{2}$ are identical, it suffices to show that for any primitive character χ modulo q,

$$\sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \frac{1}{\frac{1}{4}+\gamma^2} = \log \frac{q}{\pi} - \gamma_0 - (1 + \chi(-1)) \log 2 + 2 \operatorname{Re} \frac{L'(1,\chi)}{L(1,\chi)}.$$

There is a certain constant $B(\chi)$ that appears in the Hadamard product formula for $L(s,\chi)$. One classical formula related to it [**60**, equation (10.38)] is

$$\operatorname{Re} B(\chi) = -\sum_{\substack{\rho\in\mathbb{C} \\ 0<\operatorname{Re}\rho<1 \\ L(\rho,\chi)=0}} \operatorname{Re} \frac{1}{\rho}. \tag{2.3.3}$$

We can relate $B(\chi)$ to $b(\chi)$ under GRH by rewriting the previous equation as

$$-2 \operatorname{Re} B(\chi) = \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \operatorname{Re} \left(\frac{2}{\frac{1}{2}+i\gamma} \right) = \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \operatorname{Re} \left(\frac{1-2i\gamma}{\frac{1}{4}+\gamma^2} \right) = b(\chi). \tag{2.3.4}$$

On the other hand, Vorhauer showed in 2006 (see [**60**, equation (10.39)]) that

$$B(\chi) = -\frac{1}{2} \log \frac{q}{\pi} - \frac{L'}{L}(1,\bar{\chi}) + \frac{\gamma_0}{2} + \frac{1+\chi(-1)}{2} \log 2.$$

Taking real parts (which renders moot the difference between $\bar{\chi}$ and χ) and comparing to equation (2.3.4) establishes the lemma. □

Proof of Theorem 2.1.2. We begin by applying Lemma 2.3.2 to Definition 2.1.2 for $V(q;a,b)$, which yields

$$
\begin{aligned}
&V(q;a,b)\\
&\quad= \sum_{\substack{\chi \pmod q\\ \chi\ne\chi_0}} |\chi(a)-\chi(b)|^2\Big(\log\frac{q^*}{\pi}-\gamma_0-(1+\chi(-1))\log 2+2\operatorname{Re}\frac{L'(1,\chi^*)}{L(1,\chi^*)}\Big)\\
&\quad= \sum_{\chi \pmod q} |\chi(a)-\chi(b)|^2\log q^* - (\gamma_0+\log 2\pi)\sum_{\chi \pmod q}|\chi(a)-\chi(b)|^2\\
&\qquad\quad -\log 2\sum_{\chi \pmod q}|\chi(a)-\chi(b)|^2\chi(-1)+2M^*(q;a,b), \qquad (2.3.5)
\end{aligned}
$$

recalling Definition 2.1.4 for $M^*(q;a,b)$. We are permitted to reinclude the principal character χ_0 in the three sums on the right-hand side, since the coefficient $|\chi_0(a)-\chi_0(b)|^2$ always equals 0.

The second and third terms on the right-hand side of equation (2.3.5) are easy to evaluate using Proposition 2.3.1 : we have

$$
-(\gamma_0+\log 2\pi)\sum_{\chi \pmod q}|\chi(a)-\chi(b)|^2 = -2(\gamma_0+\log 2\pi)\phi(q) \qquad (2.3.6)
$$

and

$$
\begin{aligned}
-\log 2\sum_{\chi \pmod q}|\chi(a)-\chi(b)|^2\chi(-1) &= (\log 2)\phi(q)\big(\iota_q(-ab^{-1})+\iota_q(-ba^{-1})\big)\\
&= (2\log 2)\phi(q)\iota_q(-ab^{-1}). \qquad (2.3.7)
\end{aligned}
$$

The first sum on the right-hand side of equation (2.3.5) can be evaluated using Proposition 2.3.2 :

$$
\begin{aligned}
\sum_{\chi \pmod q}|\chi(a)-\chi(b)|^2\log q^* &= \sum_{\chi \pmod q}(2-\chi(ab^{-1})-\chi(ba^{-1}))\log q^*\\
&= 2\phi(q)\Big(\log q-\sum_{p\mid q}\frac{\log p}{p-1}\Big)+\phi(q)\frac{\Lambda(q/(q,ab^{-1}-1))}{\phi(q/(q,ab^{-1}-1))}\\
&\quad+\phi(q)\frac{\Lambda(q/(q,ba^{-1}-1))}{\phi(q/(q,ba^{-1}-1))}.
\end{aligned}
$$

Since $(q, mn) = (q, n)$ for any integer m that is relatively prime to q, we see that $(q, ab^{-1} - 1) = (q, a - b) = (q, b - a) = (q, ba^{-1} - 1)$, and therefore

$$\sum_{\chi \,(\mathrm{mod}\ q)} |\chi(a) - \chi(b)|^2 \log q^* = 2\phi(q)\bigg(\log q - \sum_{p \mid q} \frac{\log p}{p-1} + \frac{\Lambda(q/(q, a-b))}{\phi(q/(q, a-b))}\bigg). \tag{2.3.8}$$

Substituting the evaluations (2.3.6), (2.3.7), and (2.3.8) into equation (2.3.5), we obtain

$$\begin{aligned} V(q; a, b) = 2\phi(q)\bigg(\log q - \sum_{p \mid q} \frac{\log p}{p-1} + \frac{\Lambda(q/(q, a-b))}{\phi(q/(q, a-b))}\bigg) \\ - 2(\gamma_0 + \log 2\pi)\phi(q) + (2\log 2)\phi(q)\iota_q(-ab^{-1}) + 2M^*(q; a, b) \\ = 2\phi(q)\big(\mathcal{L}(q) + K_q(a-b) + \iota_q(-ab^{-1})\log 2\big) + 2M^*(q; a, b), \end{aligned}$$

where $\mathcal{L}(q)$ and $K_q(n)$ were defined in Definition 2.1.3. This establishes the theorem. □

Theorem 2.1.2 has the following asymptotic formula as a corollary :

Proposition 2.3.4. *Assuming GRH, we have*

$$V(q; a, b) = 2\phi(q)\log q + O(\phi(q)\log\log q).$$

PROOF. First note that the function $(\log t)/(t-1)$ is decreasing for $t > 1$. Consequently, $\Lambda(q)/\phi(q)$ is bounded by $\log 2$. Also, letting p_j denote the jth prime, we see that

$$\sum_{p \mid q} \frac{\log p}{p-1} \le \sum_{j=1}^{\omega(q)} \frac{\log p_j}{p_j - 1} \ll \log p_{\omega(q)} \ll \log \omega(q) \ll \log\log q,$$

where the final inequality uses the trivial bound $\omega(q) \le (\log q)/(\log 2)$. From Definition 2.1.3, we conclude that $\mathcal{L}(q) = \log q + O(\log\log q)$. Next, $K_q(a-b)$ is bounded by $\log 2$ as above, and $\iota_q(ab^{-1})$ is of course bounded as well. Finally, on GRH we know that $L'(1,\chi^*)/L(1,\chi^*) \ll \log\log q^* \le \log\log q$ (either see [55], or take $y = \log^2 q$ in Proposition 2.3.5), which immediately implies that $M^*(q; a, b) \ll \phi(q)\log\log q$ by Definition 2.1.4. The proposition now follows from Theorem 2.1.2. □

2.3.3. Evaluation of the analytic term $M^*(q;a,b)$

The goal of this section is a proof of Theorem 2.1.3. We start by examining more closely, in the next two lemmas, the relationship between the quantities $M^*(q;a,b)$ and $M(q;a,b)$ defined in Definition 2.1.4. Recall that $e(q;p,r)$ was defined in Definition 2.1.5.

Lemma 2.3.3. *If* $p^\nu \parallel q$, *then* $\displaystyle\sum_{\substack{e\ge1 \\ rp^e\equiv1 \ (\mathrm{mod}\ q/p^\nu)}} \frac{1}{p^e} = \frac{1}{p^{e(q;p,r)}\big(1-p^{-e(q;p,1)}\big)}.$

PROOF. If r is not in the multiplicative subgroup (mod q/p^ν) generated by p, then the left-hand side is clearly zero, while the right-hand side is zero by the convention that $e(q;p,r)=\infty$ in this case. Otherwise, the positive integers e for which $rp^e \equiv 1$ (mod q/p^ν) are precisely the ones of the form $e(q;p,r)+ke(q;p,1)$ for $k\ge0$, since $e(q;p,r)$ is the first such integer and $e(q;p,1)$ is the order of p (mod q/p^ν). Therefore we obtain the geometric series

$$\sum_{\substack{e\ge1 \\ rp^e\equiv1 \ (\mathrm{mod}\ q/p^\nu)}} \frac{1}{p^e} = \sum_{k=0}^{\infty} \frac{1}{p^{e(q;p,r)+ke(q;p,1)}} = \frac{1}{p^{e(q;p,r)}\big(1-p^{-e(q;p,1)}\big)}$$

as claimed. □

Definition 2.3.1. *If* $p^\nu \parallel q$, *define*

$$h_0(q;p,r) = \frac{1}{\phi(p^\nu)} \frac{\log p}{p^{e(q;p,r)}\big(1-p^{-e(q;p,1)}\big)}$$

and

$$H_0(q;a,b) = \sum_{p\mid q} \big(h_0(q;p,ab^{-1}) + h_0(q;p,ba^{-1}) - 2h_0(q;p,1)\big).$$

We will see later in this section, in the proof of Theorem 2.1.3, that h_0 *and* H_0 *are very close to the functions* h *and* H *also defined in Definition 2.1.5. Notice that if* q *is prime, then* $h_0(q;q,r) = (\log q)/q(q-1)$ *independent of* r *and thus* $H(q;a,b)=0$ *for any* a *and* b. ♢

The next lemma could be proved under a hypothesis much weaker than GRH, but this is irrelevant to our present purposes.

Lemma 2.3.4. *Assume GRH. If* a *and* b *are reduced residues* $\pmod q$*, then*

$$M^*(q;a,b) = M(q;a,b) + \phi(q)H_0(q;a,b),$$

where $M^*(q;a,b)$ *and* $M(q;a,b)$ *are defined in Definition 2.1.4.*

Proof. We begin with the identity

$$\frac{L'(1,\chi)}{L(1,\chi)} = -\lim_{y\to\infty} \sum_{p\le y} \sum_{e=1}^{\infty} \frac{\chi(p^e)\log p}{p^e}.$$

This identity follows from the fact that the Euler product of $L(s,\chi)$ converges uniformly for $\mathrm{Re}(s) \ge 1/2+\varepsilon$; this is implied by the estimate $\sum_{p\le x}\chi(p) \ll_q x^{1/2}\log^2 x$ which itself is a consequence of GRH.

Therefore

$$\begin{aligned}
&M^*(q;a,b) - M(q;a,b) \\
&= \sum_{\substack{\chi \pmod q \\ \chi\ne\chi_0}} \big|\chi(a)-\chi(b)\big|^2 \bigg(\frac{L'(1,\chi^*)}{L(1,\chi^*)} - \frac{L'(1,\chi)}{L(1,\chi)} \bigg) \\
&= -\sum_{\substack{\chi \pmod q \\ \chi\ne\chi_0}} \big|\chi(a)-\chi(b)\big|^2 \lim_{y\to\infty} \sum_{p\le y} \log p \sum_{e=1}^{\infty} \frac{\chi^*(p^e)-\chi(p^e)}{p^e} \\
&= \lim_{y\to\infty} \sum_{p\le y} \log p \sum_{e=1}^{\infty} \frac{1}{p^e} \sum_{\chi \pmod q} \big(\chi(ab^{-1}) + \chi(ba^{-1}) - 2\big)\big(\chi^*(p^e)-\chi(p^e)\big),
\end{aligned}$$

where the inserted term involving χ_0 is always zero. Proposition 2.3.3 tells us that the inner sum vanishes except possibly when the prime p divides q; invoking that proposition three times, we see that

$$\begin{aligned}
M^*(q;a,b) - M(q;a,b) &= \sum_{p^\nu \| q} \phi(q/p^\nu)\log p \\
&\times \Bigg(\sum_{\substack{e\ge1 \\ ab^{-1}p^e\equiv1 \pmod{q/p^\nu}}} \frac{1}{p^e} + \sum_{\substack{e\ge1 \\ ba^{-1}p^e\equiv1 \pmod{q/p^\nu}}} \frac{1}{p^e} - 2 \sum_{\substack{e\ge1 \\ p^e\equiv1 \pmod{q/p^\nu}}} \frac{1}{p^e} \Bigg).
\end{aligned}$$

We can evaluate these inner sums using Lemma 2.3.3 : by comparison with Definition 2.3.1,

$$\begin{aligned} M^*(q;a,b) - M(q;a,b) &= \phi(q) \sum_{p^\nu \| q} \frac{\log p}{\phi(p^\nu)} \bigg(\frac{1}{p^{e(q;p,ab^{-1})}(1-p^{-e(q;p,1)})} \\ &\quad + \frac{1}{p^{e(q;p,ba^{-1})}(1-p^{-e(q;p,1)})} - 2\frac{1}{p^{e(q;p,1)}(1-p^{-e(q;p,1)})} \bigg) \\ &= \phi(q) H_0(q;a,b), \end{aligned}$$

which establishes the lemma. □

We will need the following three propositions, with explicit constants given, when we undertake our calculations and estimations of $\delta(q;a,b)$. Because the need for explicit constants makes their derivations rather lengthy, we will defer the proofs of the first two propositions until Section 2.5.2 and derive only the third one in this section.

Proposition 2.3.5. *Assume GRH. Let χ be a nonprincipal character* (mod q). *For any positive real number y,*

$$\frac{L'(1,\chi)}{L(1,\chi)} = -\sum_{n=1}^{\infty} \frac{\chi(n)\Lambda(n)}{n} e^{-n/y} + \overline{O}\bigg(\frac{14.27 \log q + 16.25}{y^{1/2}} + \frac{16.1 \log q + 17.83}{y^{3/4}} \bigg).$$

Proposition 2.3.6. *If $1 \le a < q$, then*

$$\sum_{n \equiv a \,(\mathrm{mod}\ q)} \frac{\Lambda(n)}{n} e^{-n/q^2} = \frac{\Lambda(a)}{a} + \overline{O}\bigg(\frac{2\log^2 q}{q} + \frac{3.935 \log q}{q} \bigg).$$

Assuming these propositions for the moment, we can derive the following explicit estimate for $M^*(q;a,b)$, after which we will be able to finish the proof of Theorem 2.1.3.

Proposition 2.3.7. *Assume GRH. For any pair a, b of distinct reduced residues modulo q, let r_1 and r_2 denote the least positive residues of ab^{-1} and ba^{-1}* (mod q). *Then for $q \ge 150$,*

$$M^*(q;a,b) = \phi(q)\bigg(\frac{\Lambda(r_1)}{r_1} + \frac{\Lambda(r_2)}{r_2} + H_0(q;a,b) \bigg) + \overline{O}\bigg(\frac{23.619\phi(q)\log^2 q}{q} \bigg).$$

Proof. The bulk of the proof is devoted to understanding $M(q;a,b)$. From Proposition 2.3.5, we have

$$\begin{aligned} M(q;a,b) &= \sum_{\chi \,(\mathrm{mod}\ q)} \bigl|\chi(a)-\chi(b)\bigr|^2 \frac{L'(1,\chi)}{L(1,\chi)} \\ &= \sum_{\chi \,(\mathrm{mod}\ q)} \bigl(2-\chi(ba^{-1})-\chi(ab^{-1})\bigr)\biggl(-\sum_{n=1}^{\infty} \frac{\Lambda(n)\chi(n)}{n} e^{-n/y} \\ &\qquad + \overline{O}\biggl(\frac{14.27\log q+10.6}{y^{1/2}} + \frac{16.1\log q+13.1}{y^{3/4}}\biggr)\biggr) \\ &= \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n} e^{-n/y} \sum_{\chi \,(\mathrm{mod}\ q)} \bigl(\chi(ba^{-1}n)+\chi(ab^{-1}n)-2\chi(n)\bigr) \\ &\qquad + 4\phi(q)\overline{O}\biggl(\frac{14.27\log q+16.25}{y^{1/2}} + \frac{16.1\log q+17.83}{y^{3/4}}\biggr)\biggr), \end{aligned} \tag{2.3.9}$$

and using the orthogonality relations in Proposition 2.3.1, we see that

$$M(q;a,b) = \phi(q)\biggl(\sum_{n\equiv ab^{-1}\ (\mathrm{mod}\ q)} \frac{\Lambda(n)}{n}e^{-n/y} + \sum_{n\equiv ba^{-1}\ (\mathrm{mod}\ q)} \frac{\Lambda(n)}{n}e^{-n/y}$$
$$-2\sum_{n\equiv 1\ (\mathrm{mod}\ q)} \frac{\Lambda(n)}{n}e^{-n/y}\biggr) + 4\phi(q)\overline{O}\biggl(\frac{14.27\log q+16.25}{y^{1/2}} + \frac{16.1\log q+17.83}{y^{3/4}}\biggr)\biggr).$$

At this point we choose $y = q^2$. We calculate that $(14.27\log q + 16.25)/q + (16.1\log q + 17.83)/q^{3/2} < 3.816(\log^2 q)/q$ for $q \ge 150$, and so

$$\begin{aligned} M(q;a,b) = \phi(q)\biggl(&\sum_{n\equiv ab^{-1}\ (\mathrm{mod}\ q)} \frac{\Lambda(n)}{n}e^{-n/q^2} + \sum_{n\equiv ba^{-1}\ (\mathrm{mod}\ q)} \frac{\Lambda(n)}{n}e^{-n/q^2} \\ &- 2\sum_{n\equiv 1\ (\mathrm{mod}\ q)} \frac{\Lambda(n)}{n}e^{-n/q^2}\biggr) + \overline{O}\biggl(\frac{15.263\phi(q)\log^2 q}{q}\biggr). \end{aligned}$$

Let r_1 and r_2 denote the least positive residues of ab^{-1} and ba^{-1} (mod q). Using Proposition 2.3.6 three times, we see that

$$\begin{aligned} M(q;a,b) &= \phi(q)\biggl(\frac{\Lambda(r_1)}{r_1} + \frac{\Lambda(r_2)}{r_2} - 2\frac{\Lambda(1)}{1} + \overline{O}\biggl(3\Bigl(\frac{2\log^2 q}{q} + \frac{3.935\log q}{q}\Bigr)\biggr)\biggr) \\ &\qquad + \overline{O}\biggl(\frac{15.263\phi(q)\log^2 q}{q}\biggr) \\ &= \phi(q)\biggl(\frac{\Lambda(r_1)}{r_1} + \frac{\Lambda(r_2)}{r_2}\biggr) + \overline{O}\biggl(\frac{36.619\phi(q)\log^2 q}{q}\biggr) \end{aligned}$$

for $q \geq 150$. With this understanding of $M(q;a,b)$, the proposition now follows for $M^*(q;a,b)$ by Lemma 2.3.4. □

Proof of Theorem 2.1.3. Since Proposition 2.3.7 tells us that

$$M^*(q;a,b) = \phi(q)\bigg(\frac{\Lambda(r_1)}{r_1} + \frac{\Lambda(r_2)}{r_2} + H_0(q;a,b) + O\bigg(\frac{\log^2 q}{q}\bigg)\bigg),$$

all we need to do to prove the theorem is to show that

$$H_0(q;a,b) = H(q;a,b) + O\bigg(\frac{\log^2 q}{q}\bigg).$$

The key observation is that $p^{e(q;p,1)} \equiv 1 \pmod{q/p^\nu}$ and $p^{e(q;p,1)} \geq p^1 > 1$, and so $p^{e(q;p,1)} > q/p^\nu$. Therefore by Definitions 2.1.5 and 2.3.1, we have $h_0(q;p,r) = h(q;p,r)(1+O(p^\nu/q))$ and $h(q;p,1) \ll (\log p)/\phi(p^\nu)(q/p^\nu) \ll (\log p)/q$. We see that

$$\begin{aligned} H_0(q;a,b) &= \sum_{p^\nu \,\|\, q} \big(h_0(q;p,ab^{-1}) + h_0(q;p,ba^{-1}) - 2h_0(q;p,1)\big) \\ &= \sum_{p^\nu \,\|\, q} \Big(\big(h(q;p,ab^{-1}) + h(q;p,ba^{-1})\big)\big(1+O\big(\tfrac{p^\nu}{q}\big)\big) + O\big(\tfrac{\log p}{q}\big)\Big). \end{aligned}$$

It is certainly true that $h(q;p,r) \ll (\log p)/\phi(p^\nu) \ll (\log p)/p^\nu$, and so the previous equation becomes

$$H_0(q;a,b) = H(q;a,b) + O\bigg(\sum_{p^\nu \,\|\, q}\bigg(\frac{\log p}{p^\nu}\frac{p^\nu}{q} + \frac{\log p}{q}\bigg)\bigg) = H(q;a,b) + O\bigg(\frac{\log q}{q}\bigg),$$

which establishes the theorem. □

2.3.4. Estimates in terms of arithmetic information only

The purpose of this section is to show that the densities $\delta(q;a,b)$ can be calculated extremely precisely using only "arithmetic information". For the purposes of this section, "arithmetic information" means finite expressions composed of elementary arithmetic operations involving only integers, logarithms of integers, values of the Riemann zeta function at positive integers, and the constants π and γ_0. (In fact, all of these quantities themselves can in principal be calculated arbitrarily precisely using only elementary arithmetic operations on integers.) The point is that "arithmetic information" excludes integrals and such quantities as Dirichlet characters and L-functions, Bessel functions, and trigonometric functions. The formula we can derive, with only arithmetic information in the main term, has an error term of the form $O_A(q^{-A})$ for any constant $A > 0$ we care to specify in advance.

To begin, we note that letting y tend to infinity in equation (2.3.9) leads to the heuristic statement

$$\begin{aligned} M&(q;a,b) \\ &= \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n} \sum_{\chi \pmod q} \big(\chi(ba^{-1}n) + \chi(ab^{-1}n) - 2\chi(n)\big) \\ &= \phi(q) \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n} \big(\iota_q(ba^{-1}n) + \iota_q(ab^{-1}n) - 2\iota_q(n)\big) \\ &\text{“}{=}\text{''}\ \phi(q)\bigg(\sum_{n\equiv ab^{-1} \pmod q} \frac{\Lambda(n)}{n} + \sum_{n\equiv ba^{-1} \pmod q} \frac{\Lambda(n)}{n} - 2 \sum_{n\equiv 1 \pmod q} \frac{\Lambda(n)}{n}\bigg), \end{aligned}$$

where the "=" warns that the sums on the right-hand side do not individually converge. In fact, using a different approach based on the explicit formula, one can obtain

$$\begin{aligned} M(q;a,b) = \phi(q)\bigg(&\sum_{\substack{1\le n\le y \\ n\equiv ab^{-1} \pmod q}} \frac{\Lambda(n)}{n} + \sum_{\substack{1\le n\le y \\ n\equiv ba^{-1} \pmod q}} \frac{\Lambda(n)}{n} \\ &- 2 \sum_{\substack{1\le n\le y \\ n\equiv 1 \pmod q}} \frac{\Lambda(n)}{n}\bigg) + O\bigg(\frac{\phi(q)\log^2 qy}{\sqrt{y}}\bigg). \end{aligned} \tag{2.3.10}$$

In light of Theorem 2.1.2 in conjunction with Lemma 2.3.4, we see that we can get an arbitrarily good approximation to $V(q;a,b)$ using only arithmetic information.

By Theorem 2.1.1, we see we can thus obtain an extremely precise approximation for $\delta(q;a,b)$ as long as we can calculate the coefficients $s_{q;a,b}(\ell,j)$ defined in Definition 2.2.8. Inspecting that definition reveals that it suffices to be able to calculate $W_m(q;a,b)$ (or equivalently $W_m(q;a,b)V(q;a,b)$) arbitrarily precisely using only arithmetic content. With the next several lemmas, we describe how such a calculation can be made.

Lemma 2.3.5. *Let* n *be a positive integer, and set* $\ell = \lfloor \frac{n}{2} \rfloor$. *There exist rational numbers* $C_{n,1}, \ldots, C_{n,\ell}$ *such that*

$$\frac{1}{(1/4+t^2)^n} = 2\operatorname{Re}\left(\frac{1}{(1/2-it)^n}\right) + \frac{C_{n,1}}{(1/4+t^2)^{n-1}} + \frac{C_{n,2}}{(1/4+t^2)^{n-2}} + \ldots + \frac{C_{n,\ell}}{(1/4+t^2)^{n-\ell}}$$

for any complex number t.

Proof. Since

$$2\operatorname{Re}\frac{1}{(1/2-it)^n} = \frac{1}{(1/2-it)^n} + \frac{1}{(1/2+it)^n} = \frac{(1/2+it)^n + (1/2-it)^n}{(1/4+t^2)^n},$$

it suffices to show that

$$\frac{(1/2+it)^n + (1/2-it)^n}{(1/4+t^2)^n} = \frac{C_{n,0}}{(1/4+t^2)^n} + \frac{-C_{n,1}}{(1/4+t^2)^{n-1}} + \cdots + \frac{-C_{n,\ell}}{(1/4+t^2)^{n-\ell}}, \tag{2.3.11}$$

where each $C_{n,m}$ is a rational number and $C_{n,0} = 1$. In fact, we need only show that this identity holds for some rational number $C_{n,0}$, since multiplying both sides by $(1/4+t^2)^n$ and taking the limit as t tends to $i/2$ proves that $C_{n,0}$ must equal 1.

Using the binomial theorem,

$$\begin{aligned}(1/2+it)^n + (1/2-it)^n &= \sum_{k=0}^{n} \tbinom{n}{k}\left(\tfrac{1}{2}\right)^{n-k}\left((it)^k + (-it)^k\right) \\ &= \sum_{j=0}^{\ell} \tbinom{n}{2j}\left(\tfrac{1}{2}\right)^{n-2j}\left(2(-1)^j t^{2j}\right) \\ &= 2\sum_{j=0}^{\ell} \tbinom{n}{2j}\left(\tfrac{1}{2}\right)^{n-2j}(-1)^j\left((\tfrac{1}{4}+t^2) - \tfrac{1}{4}\right)^j \\ &= 2\sum_{j=0}^{\ell} \tbinom{n}{2j}\left(\tfrac{1}{2}\right)^{n-2j}(-1)^j \sum_{m=0}^{j} \tbinom{j}{m}(\tfrac{1}{4}+t^2)^m\left(-\tfrac{1}{4}\right)^{j-m},\end{aligned}$$

which is a linear combination of the expressions $(1/4+t^2)^m$, for $0 \le m \le \ell$, with rational coefficients not depending on t. Dividing both sides by $(1/4+t^2)^n$ establishes equation (2.3.11) for suitable rational numbers $C_{n,m}$ and hence the lemma. □

For the rest of this section, we say that a quantity is a *fixed $\mathbb{Q}$-linear combination* of certain elements if the coefficients of this linear combination are rational numbers that are independent of q, a, b and χ (but may depend on n and j where appropriate). Our methods allow the exact calculation of these rational coefficients, but the point of this section would be obscured by the bookkeeping required to record them.

Definition 2.3.2. *As usual,* $\Gamma(z)$ *denotes Euler's Gamma function. For any positive integer* n *and any Dirichlet character* $\chi \pmod q$*, define*

$$b_n(\chi) = \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \frac{1}{(\frac14+\gamma^2)^n},$$

so that $b_1(\chi) = b(\chi)$ *for example.* ♢

Lemma 2.3.6. *Assume GRH. Let* n *be a positive integer, and let* χ *be a primitive character* $\pmod q$*. Then* $b_n(\chi)$ *is a fixed* $\mathbb{Q}$*-linear combination of the quantities*

$$\begin{aligned}\bigg\{\log\frac{q}{\pi}, \left[\frac{d}{ds}\log\Gamma(s)\right]_{s=(1+\xi)/2}, \dots, \left[\frac{d^n}{ds^n}\log\Gamma(s)\right]_{s=(1+\xi)/2}, \\ \operatorname{Re}\left[\frac{d}{ds}\log L(s,\chi)\right]_{s=1}, \dots, \operatorname{Re}\left[\frac{d^n}{ds^n}\log L(s,\chi)\right]_{s=1}\bigg\},\end{aligned} \tag{2.3.12}$$

where $\xi = 0$ *if* $\chi(-1) = 1$ *and* $\xi = 1$ *if* $\chi(-1) = -1$.

Remark 2.3.1. *Since the critical zeros of* $L(s,\chi)$ *and* $L(s,\chi^*)$ *are identical, the lemma holds for any nonprincipal character* χ *if, in the set* (2.3.12)*, we replace* q *by* q^* *and* $L(s,\chi)$ *by* $L(s,\chi^*)$.

PROOF. For primitive characters χ, Lemma 2.3.2 tells us that

$$\begin{aligned}b(\chi) &= \log\frac{q}{\pi} - \gamma_0 - (1+\chi(-1))\log 2 + 2\operatorname{Re}\frac{L'(1,\chi)}{L(1,\chi)} \\ &= \log\frac{q}{\pi} + \left[\frac{\Gamma'(s)}{\Gamma(s)}\right]_{s=(1+\xi)/2} + 2\operatorname{Re}\frac{L'(1,\chi)}{L(1,\chi)},\end{aligned}$$

which establishes the lemma for $n = 1$. We proceed by induction on n. By Lemma 2.3.5, we see that

$$\begin{aligned} b_n(\chi) &= \sum_{\substack{\gamma\in\mathbb{R}\\ L(1/2+i\gamma,\chi)=0}} \frac{1}{(1/4+\gamma^2)^n} \\ &= \sum_{\substack{\gamma\in\mathbb{R}\\ L(1/2+i\gamma,\chi)=0}} \bigg(2\,\mathrm{Re}\,\frac{1}{(1/2-i\gamma)^n} + \frac{C_{n,1}}{(1/4+\gamma^2)^{n-1}} + \frac{C_{n,2}}{(1/4+\gamma^2)^{n-2}} + \dots \\ &\qquad\qquad + \frac{C_{n,\ell}}{(1/4+\gamma^2)^{n-\ell}} \bigg) \\ &= C_{n,1}b_{n-1}(\chi) + \dots + C_{n,\ell}b_{n-\ell}(\chi) + 2 \sum_{\substack{\gamma\in\mathbb{R}\\ L(1/2+i\gamma,\chi)=0}} \mathrm{Re}\,\frac{1}{(1/2-i\gamma)^n} \end{aligned} \tag{2.3.13}$$

(where $\ell = \lfloor \frac{n}{2} \rfloor$). By the induction hypothesis, each term of the form $C_{n,m}\times b_{n-m}(\chi)$ is a fixed $\mathbb{Q}$-linear combination of the elements of the set (2.3.12); therefore all that remains is to show that the sum on the right-hand side of equation (2.3.13) is also a fixed $\mathbb{Q}$-linear combination of these elements.

Consider the known formula [**60**, equation (10.37)]

$$\frac{d}{ds}\log L(s,\chi) = B(\chi) - \frac{d}{ds}\log\Gamma\Big(\frac{s+\xi}{2}\Big) - \frac{1}{2}\log\frac{q}{\pi} + \sum_{\rho}\Big(\frac{1}{s-\rho}+\frac{1}{\rho}\Big),$$

where $\sum_\rho$ denotes a sum over all nontrivial zeros of $L(s,\chi)$ and $B(\chi)$ is a constant (alluded to in the proof of Lemma 2.3.2). If we differentiate this formula $n-1$ times with respect to s, we obtain

$$\frac{d^n}{ds^n}\log L(s,\chi) = -\frac{d^n}{ds^n}\log\Gamma\Big(\frac{s+\xi}{2}\Big) + \sum_{\rho}\frac{(-1)^{n-1}(n-1)!}{(s-\rho)^n}.$$

Setting $s = 1$ and taking real parts, and using GRH, we conclude that

$$\begin{aligned} \sum_{\substack{\gamma\in\mathbb{R}\\ L(1/2+i\gamma,\chi)=0}} \mathrm{Re}\frac{1}{(1/2-i\gamma)^n} &= \sum_{\rho}\mathrm{Re}\,\frac{1}{(1-\rho)^n} \\ &= \frac{(-1)^{n-1}}{(n-1)!}\bigg[\mathrm{Re}\,\frac{d^n}{ds^n}\log L(s,\chi) + \frac{d^n}{ds^n}\log\Gamma\Big(\frac{s+\xi}{2}\Big)\bigg]_{s=1}, \end{aligned}$$

which is a fixed $\mathbb{Q}$-linear combination of the elements of the set (2.3.12) as desired. (Although $\left[\frac{d^n}{ds^n}\log\Gamma(s)\right]_{s=(1+\xi)/2}$ and $\left[\frac{d^n}{ds^n}\log\Gamma\left(\frac{s+\xi}{2}\right)\right]_{s=1}$ differ by a factor of 2^n, this does not invalidate the conclusion.) □

The following three definitions, which generalize earlier notation, will be important in our analysis of the higher-order terms $W_n(q;a,b)V(q;a,b)$.

Definition 2.3.3. *For any positive integers* q *and* n*, define*

$$\mathcal{L}_n(q) = \sum_{|i|\le n} (-1)^i \binom{2n}{n+i} \bigg(\iota_q(a^i b^{-i}) \bigg(\log\frac{q}{\pi} - \sum_{p\mid q} \frac{\log p}{p-1} \bigg) - \big(1-\iota_q(a^i b^{-i})\big) \frac{\Lambda(q/(q,a^i-b^i))}{\phi(q/(q,a^i-b^i))} \bigg).$$

♢

Definition 2.3.4. *Let* χ *be a Dirichlet character* (mod q)*, and let* a *and* b *be integers. For any positive integers* $j\le n$*, define*

$$\mathcal{M}^*_{n,j}(q;a,b) = \frac{1}{\phi(q)} \sum_{\substack{\chi \ (\mathrm{mod}\ q) \\ \chi\ne\chi_0}} |\chi(a)-\chi(b)|^{2n} \left[\frac{d^j}{ds^j}\log L(s,\chi^*)\right]_{s=1}$$

and

$$\mathcal{M}_{n,j}(q;a,b) = \frac{1}{\phi(q)} \sum_{\substack{\chi \ (\mathrm{mod}\ q) \\ \chi\ne\chi_0}} |\chi(a)-\chi(b)|^{2n} \left[\frac{d^j}{ds^j}\log L(s,\chi)\right]_{s=1},$$

so that $\mathcal{M}^*_{1,1}(q;a,b) = M^*(q;a,b)/\phi(q)$ *and* $\mathcal{M}_{1,1}(q;a,b) = M(q;a,b)/\phi(q)$ *for example. One can use Lemma 2.3.7 and Perron's formula to show that*

$$\mathcal{M}_{n,j}(q;a,b) = (-1)^j \sum_{|i|\le j} (-1)^i \binom{2j}{j+i} \sum_{\substack{n\le y \\ n\equiv a^i b^{-i}\ (\mathrm{mod}\ q)}} \frac{\Lambda(n)\log^{j-1} n}{n} + O_j\bigg(\frac{\log^{j+1} qy}{\sqrt{y}}\bigg),$$

in analogy with equation (2.3.10). ♢

Definition 2.3.5. *For any distinct reduced residue classes* a *and* b (mod q)*, define*

$$H_{n,j}(q;a,b) = (-1)^j \sum_{p^\nu \| q} \frac{(\log p)^j}{\phi(p^\nu)} \sum_{|i|\le j} (-1)^i \binom{2n}{n+i} \sum_{\substack{e\ge 1 \\ a^i b^{-i} p^e \equiv 1\ (\mathrm{mod}\ q/p^\nu)}} \frac{e^{j-1}}{p^e}$$

for any integers $1 \le j \le n$*. Notice that the inner sum is*

$$\sum_{\substack{e\ge1 \\ a^ib^{-i}p^e\equiv1 \,(\mathrm{mod}\ q/p^\nu)}} \frac{e^{j-1}}{p^e} = \sum_{\substack{e\ge1 \\ e\equiv e(q;p,a^ib^{-i}) \,(\mathrm{mod}\ e(q;p,1))}} \frac{e^{j-1}}{p^e},$$

where $e(q;p,r)$ *is defined in Definition 2.1.5. It turns out that the identity*

$$\sum_{\substack{e\ge1 \\ e\equiv r\,(\mathrm{mod}\ s)}} \frac{e^m}{p^e} = \frac{1}{p^r(1-p^{-s})}\sum_{g=0}^{m}\binom{m}{g}s^g r^{m-g}\sum_{\ell=0}^{g}\left\{ {g \atop \ell} \right\}\frac{\ell!}{(p^s-1)^\ell}$$

(in which $\left\{ {g \atop \ell} \right\}$ *denotes the Stirling number of the second kind) is valid for any positive integers* m*,* p*,* r*, and* s *such that* $r \le s$ *(as one can see by expanding* $(sk+r)^m$ *by the binomial theorem and then invoking the identity* [**30**, (equation 7.46)]*). Consequently, we see that* $H_{n,j}(q;a,b)$ *is a rational linear combination of the elements of the set* $\{(\log p)^j \colon p \mid q\}$ *(although the rational coefficients depend upon* q*,* a*, and* b*).* ♢

Once we determine how to expand the coefficient $|\chi(a)-\chi(b)|^{2n}$ as a linear combination of individual values of χ, we can establish Proposition 2.3.8 which describes how the cumulant $W_n(q;a,b)V(q;a,b)$ can be evaluated in terms of the arithmetic information already defined.

Lemma 2.3.7. *Let* χ *be a Dirichlet character* (mod q)*, and let* a *and* b *be reduced residues* (mod q)*. For any nonnegative integer* n*, we have*

$$|\chi(a)-\chi(b)|^{2n} = \sum_{|i|\le n}(-1)^i\binom{2n}{n+i}\chi(a^ib^{-i}).$$

PROOF. The algebraic identity

$$\left(2-t-t^{-1}\right)^n = \sum_{|i|\le n}(-1)^i\binom{2n}{n+i}t^i$$

can be verified by a straightforward induction on n. Since

$$|\chi(a)-\chi(b)|^2 = \big(\chi(a)-\chi(b)\big)\overline{\big(\chi(a)-\chi(b)\big)} = 2-\chi(ab^{-1})-\chi(ab^{-1})^{-1},$$

the lemma follows immediately. □

Proposition 2.3.8. *Assume GRH. Let* a *and* b *be reduced residues* (mod q). *For any positive integer* n*, the expression* $W_n(q;a,b)V(q;a,b)/\phi(q)$ *can be written as a fixed* $\mathbb{Q}$*-linear combination of elements in the set*

$$\{\mathcal{L}_n(q)\} \cup \big\{\iota_q\big(a^i b^{-i}\big)\log 2,\ \iota_q\big(-a^i b^{-i}\big)\log 2,\ \iota_q\big(a^i b^{-i}\big)\gamma_0 \colon |i| \le n\big\}$$
$$\cup \big\{\iota_q\big(a^i b^{-i}\big)\zeta(j),\ \iota_q\big(-a^i b^{-i}\big)\zeta(j) \colon |i| \le n,\ 2 \le j \le n\big\}$$
$$\cup \big\{H_{n,j}(q;a,b),\ \mathcal{M}_{n,j}(q;a,b) \colon 1 \le j \le n\big\}. \quad (2.3.14)$$

PROOF. From the definitions (2.2.5) and (2.3.2) of $W_n(q;a,b)$ and $b_n(\chi)$, we have

$$\frac{W_n(q;a,b)V(q;a,b)}{\phi(q)} = \frac{2^{2n}|\lambda_{2n}|}{\phi(q)} \sum_{\chi \,(\mathrm{mod}\ q)} |\chi(a)-\chi(b)|^{2n} \sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} \frac{1}{(1/4+\gamma^2)^n}$$
$$= 2^{2n-1}|\lambda_{2n}| \cdot \frac{1}{\phi(q)} \sum_{\chi \,(\mathrm{mod}\ q)} |\chi(a)-\chi(b)|^{2n} b_n(\chi).$$

Lemma 2.2.1(d) tells us that the numbers λ_{2n} are rational. Therefore by Lemma 2.3.6, it suffices to establish that three types of expressions, corresponding to the three types of quantities in the set (2.3.12), are fixed $\mathbb{Q}$-linear combinations of elements of the set (2.3.14).

Type 1 : $\displaystyle\frac{1}{\phi(q)} \sum_{\chi \,(\mathrm{mod}\ q)} |\chi(a)-\chi(b)|^{2n} \log \frac{q^*}{\pi}$.

Note that Proposition 2.3.2 can be rewritten in the form

$$\frac{1}{\phi(q)} \sum_{\chi \,(\mathrm{mod}\ q)} \chi(a)\log q^* = \iota_q(a)\bigg(\log q - \sum_{p \mid q} \frac{\log p}{p-1}\bigg)$$
$$- \big(1-\iota_q(a)\big)\frac{\Lambda(q/(q,a-1))}{\phi(q/(q,a-1))}. \quad (2.3.15)$$

By Lemma 2.3.7 and the orthogonality relation (2.3.1), we have

$$\frac{1}{\phi(q)} \sum_{\chi \,(\mathrm{mod}\ q)} |\chi(a)-\chi(b)|^{2n}\chi(c) = \sum_{|i| \le n} (-1)^i \binom{2n}{n+i} \iota_q(a^i b^{-i} c). \quad (2.3.16)$$

Therefore, using equation (2.3.15) and Proposition 2.3.1, we get

$$
\begin{aligned}
&\frac{1}{\phi(q)}\sum_{\chi\,(\mathrm{mod}\ q)}|\chi(a)-\chi(b)|^{2n}\log\frac{q^*}{\pi}\\
&=\frac{1}{\phi(q)}\sum_{\chi\,(\mathrm{mod}\ q)}|\chi(a)-\chi(b)|^{2n}\log q^*-\frac{1}{\phi(q)}\sum_{\chi\,(\mathrm{mod}\ q)}|\chi(a)-\chi(b)|^{2n}\log\pi\\
&=\sum_{|i|\le n}(-1)^i\binom{2n}{n+i}\Bigg(\iota_q(a^ib^{-i})\Bigg(\log q-\sum_{p|q}\frac{\log p}{p-1}\Bigg)\\
&\qquad-(1-\iota_q(a^ib^{-i}))\frac{\Lambda(q/(q,a^ib^{-i}-1))}{\phi(q/(q,a^ib^{-i}-1))}-\iota_q(a^ib^{-i})\log\pi\Bigg)=\mathcal{L}_n(q),
\end{aligned}
$$

since $(q,a^ib^{-i}-1)=(q,a^i-b^i)$.

Type 2 : $\displaystyle\frac{1}{\phi(q)}\sum_{\chi\,(\mathrm{mod}\ q)}|\chi(a)-\chi(b)|^{2n}\left[\frac{d^j}{ds^j}\log\Gamma(s)\right]_{s=(1+\xi)/2}$ for some $1\le j\le n$.

The following identities hold for $j\ge 2$ (see [**1**, equations 6.4.2 and 6.4.4]) :

$$
\begin{aligned}
\left[\frac{d^j}{ds^j}\log\Gamma(s)\right]_{s=1}&=(-1)^j(j-1)!\zeta(j);\\
\left[\frac{d^j}{ds^j}\log\Gamma(s)\right]_{s=1/2}&=(-1)^j(j-1)!\zeta(j)(2^j-1).
\end{aligned}
$$

Because $\xi=0$ when $\chi(-1)=1$ and $\xi=1$ when $\chi(-1)=-1$, we may thus write

$$
\left[\frac{d^j}{ds^j}\log\Gamma(s)\right]_{s=(1+\xi)/2}=(-1)^j(j-1)!\zeta(j)\big(2^{j-1}+\chi(-1)(2^{j-1}-1)\big),
$$

whence by equation (2.3.16),

$$
\begin{aligned}
&\frac{1}{\phi(q)}\sum_{\chi\,(\mathrm{mod}\ q)}|\chi(a)-\chi(b)|^{2n}\left[\frac{d^j}{ds^j}\log\Gamma(s)\right]_{s=(1+\xi)/2}\\
&=\frac{1}{\phi(q)}\sum_{\chi\,(\mathrm{mod}\ q)}|\chi(a)-\chi(b)|^{2n}(-1)^j(j-1)!\zeta(j)\big(2^{j-1}+\chi(-1)(2^{j-1}-1)\big)\\
&=(-1)^j(j-1)!\zeta(j)\Bigg(2^{j-1}\sum_{|i|\le n}(-1)^i\binom{2n}{n+i}\iota_q(a^ib^{-i})\\
&\qquad+(2^{j-1}-1)\sum_{|i|\le n}(-1)^i\binom{2n}{n+i}\iota_q(-a^ib^{-i})\Bigg),
\end{aligned}
$$

which is a linear combination of the desired type. The case $j = 1$ can be handled similarly using the identity

$$\left[\frac{d}{ds}\log\Gamma(s)\right]_{s=(1+\xi)/2} = -\gamma_0 - \big(1+\chi(-1)\big)\log 2.$$

Type 3 : $\displaystyle\frac{1}{\phi(q)}\sum_{\chi\,(\mathrm{mod}\ q)} |\chi(a)-\chi(b)|^{2n}\,\mathrm{Re}\left[\frac{d^j}{ds^j}\log L(s,\chi^*)\right]_{s=1}$ for some $1 \le j \le n$.

The expression in question is exactly $\mathcal{M}^*_{n,j}(q;a,b)$, and so it suffices to show that $\mathcal{M}^*_{n,j}(q;a,b) = \mathcal{M}_{n,j}(q;a,b) + H_{n,j}(q;a,b)$. Note that the identity

$$\frac{d^j}{ds^j}\log L(s,\chi) = \frac{d^{j-1}}{ds^{j-1}}\bigg(-\sum_{n=1}^{\infty}\frac{\Lambda(n)\chi(n)}{n^s}\bigg) = (-1)^j\sum_{n=1}^{\infty}\frac{\Lambda(n)(\log n)^{j-1}\chi(n)}{n^s}$$

implies

$$\left[\frac{d^j}{ds^j}\log L(s,\chi)\right]_{s=1} = (-1)^j\sum_p(\log p)^j\sum_{e=1}^{\infty}\frac{e^{j-1}}{p^e}\chi(p^e).$$

The proof of Lemma 2.3.4 can then be adapted to obtain the equation

$$\begin{aligned}
&\mathcal{M}^*_{n,j}(q;a,b) - \mathcal{M}_{n,j}(q;a,b)\\
&\quad= \frac{(-1)^j}{\phi(q)}\sum_{p\mid q}(\log p)^j\sum_{e=1}^{\infty}\frac{e^{j-1}}{p^e}\sum_{\chi\,(\mathrm{mod}\ q)}\big|\chi(a)-\chi(b)\big|^{2n}\chi^*(p^e)\\
&\quad= \frac{(-1)^j}{\phi(q)}\sum_{p\mid q}(\log p)^j\sum_{e=1}^{\infty}\frac{e^{j-1}}{p^e}\sum_{|i|\le n}(-1)^i\binom{2n}{n+i}\sum_{\chi\,(\mathrm{mod}\ q)}\chi(a^ib^{-i})\chi^*(p^e)
\end{aligned}$$

by Lemma 2.3.7. Evaluating the inner sum by Proposition 2.3.3 shows that this last expression is precisely the definition of $H_{n,j}(q;a,b)$, as desired. □

As described at the beginning of this section, Proposition 2.3.14 is exactly what we need to justify the assertion that we can calculate $\delta(q;a,b)$, using only arithmetic information, to within an error of the form $O_A(q^{-A})$. That some small primes in arithmetic progressions (mod q) enter the calculations is not surprising ; interestingly, though, the arithmetic progressions involved are the residue classes a^jb^{-j} for $|j| \le n$, rather than the residue classes a and b themselves !

To give a better flavor of the form these approximations take, we end this section by explicitly giving such a formula with an error term better than

$O(q^{-5/2+\varepsilon})$ for any $\varepsilon > 0$. Taking $K = 1$ in Theorem 2.1.1 gives the formula

$$\delta(q;a,b) = \frac{1}{2} + \frac{\rho(q)}{\sqrt{2\pi V(q;a,b)}}\left(1 - \frac{\rho(q)^2}{6V(q;a,b)} - \frac{3W_2(q;a,b)}{V(q;a,b)}\right) + O\left(\frac{\rho(q)^5}{V(q;a,b)^{5/2}}\right). \quad (2.3.17)$$

Going through the above proofs, one can laboriously work out that

$$\begin{aligned}\frac{W_2(q;a,b)V(q;a,b)}{\phi(q)} &= \tfrac{1}{4}\mathcal{L}_2(q) \\ &- \tfrac{1}{4\phi(q)} \sum_{\chi \,(\mathrm{mod}\ q)} \big|\chi(a) - \chi(b)\big|^4 \big\{\big(\gamma_0 + \log 2 + \tfrac{1}{2}\zeta(2)\big) + \chi(-1)\big(\log 2 + \tfrac{1}{4}\zeta(2)\big)\big\} \\ &\quad + \tfrac{1}{2}\big(\mathcal{M}_{2,1}(q;a,b) + H_{2,1}(q;a,b)\big) - \tfrac{1}{4}\big(\mathcal{M}_{2,2}(q;a,b) + H_{2,2}(q;a,b)\big),\end{aligned}$$

to which Lemma 2.3.7 can be applied with $n = 2$. Combining these two expressions and expanding $V(q;a,b)$ as described after equation (2.3.10) results in the following formula :

Proposition 2.3.9. *Assume GRH and LI. Suppose* a *and* b *are reduced residues* (mod q) *such that* a *is a nonsquare and* b *is a square* (mod q). *Then*

$$\begin{aligned}\delta(q;a,b) = \frac{1}{2} &+ \frac{\rho(q)}{2\sqrt{\pi\phi(q)(\tilde{\mathcal{L}}(q;a,b) + \tilde{\mathcal{R}}(q;a,b))}}\Bigg(1 - \frac{\rho(q)^2}{12\phi(q)\tilde{\mathcal{L}}(q;a,b)} \\ &- \frac{3}{16\phi(q)\tilde{\mathcal{L}}(q;a,b)^2}\Big\{\mathcal{L}_2(q) - (6 + 2\iota_q(a^2b^{-2}))\big(\gamma_0 + \log 2 + \tfrac{1}{2}\zeta(2)\big) \\ &\qquad - (2\iota_q(-a^2b^{-2}) - 8\iota_q(-ab^{-1}))\big(\log 2 + \tfrac{1}{4}\zeta(2)\big) \\ &\qquad + 2\mathcal{F}_1(q;a,b) + 2H_{2,1}(q;a,b) - \mathcal{F}_2(q;a,b) - H_{2,2}(q;a,b)\Big\}\Bigg) \\ &+ O\left(\frac{\rho(q)^5\sqrt{\log q}}{\phi(q)^{5/2}}\right),\end{aligned}$$

where $\mathcal{L}_2(q)$ is defined in Definition 2.3.3 and $H_{2,j}(q;a,b)$ is defined in Definition 2.3.5, and

$$\begin{aligned}
\tilde{\mathcal{L}}(q;a,b) &= \mathcal{L}(q) + K_q(a-b) + \iota_q(-ab^{-1})\log 2 + H_0(q;a,b) \\
&\qquad\qquad\qquad\qquad + \frac{\Lambda(ab^{-1})}{ab^{-1}} + \frac{\Lambda(ba^{-1})}{ba^{-1}} \\
\tilde{\mathcal{R}}(q;a,b) &= \sum_{\substack{q\le n\le q^4 \\ n\equiv ab^{-1} \pmod q}} \frac{\Lambda(n)}{n} + \sum_{\substack{q\le n\le q^4 \\ n\equiv ba^{-1} \pmod q}} \frac{\Lambda(n)}{n} - 2\sum_{\substack{q\le n\le q^4 \\ n\equiv 1 \pmod q}} \frac{\Lambda(n)}{n} \\
\mathcal{F}_1(q;a,b) &= \frac{\Lambda(a^2b^{-2})}{a^2b^{-2}} - 4\frac{\Lambda(ab^{-1})}{ab^{-1}} - 4\frac{\Lambda(ba^{-1})}{ba^{-1}} + \frac{\Lambda(b^2a^{-2})}{b^2a^{-2}} \\
\mathcal{F}_2(q;a,b) &= \frac{\Lambda(a^2b^{-2})\log(a^2b^{-2})}{a^2b^{-2}} - 4\frac{\Lambda(ab^{-1})\log(ab^{-1})}{ab^{-1}} \\
&\qquad\qquad - 4\frac{\Lambda(ba^{-1})\log(ba^{-1})}{ba^{-1}} + \frac{\Lambda(b^2a^{-2})\log(b^2a^{-2})}{b^2a^{-2}}.
\end{aligned}$$

In all these definitions, expressions such as a^2b^{-2} refer to the smallest positive integer congruent to $a^2b^{-2} \pmod q$.

2.3.5. A central limit theorem

In this section we prove a central limit theorem for the functions

$$E(x;q,a) - E(x;q,b) = \phi(q)(\pi(x;q,a) - \pi(x;q,b))x^{-1/2}\log x.$$

The technique we use is certainly not without precedent. Hooley [**40**] and Rubinstein and Sarnak [**68**] both prove central limit theorems for similar normalized error terms under the same hypotheses GRH and LI (though each with different acronyms).

Theorem 2.3.10. *Assume GRH and LI. As q tends to infinity, the limiting logarithmic distributions of the functions*

$$\frac{E(x;q,a) - E(x;q,b)}{\sqrt{2\phi(q)\log q}} \tag{2.3.18}$$

converge in measure to the standard normal distribution of mean 0 and variance 1, uniformly for all pairs a, b of distinct reduced residues modulo q.

We remark that this result can in fact be derived from Rubinstein and Sarnak's 2-dimensional central limit theorem [**68**, Section 3.2] for

$\big(E(x;q,a), E(x;q,b)\big)$, although this implication is not made explicit in their paper. In general, let

$$\phi_{X,Y}(s,t) = \int_0^\infty \int_0^\infty \exp\big(i(sx+ty)\big) f_{X,Y}(x,y)\,dx\,dy$$

denote the joint characteristic function of a pair (X,Y) of real-valued random variables, where $f_{X,Y}(x,y)$ is the joint density function of the pair. Then the characteristic function of the real-valued random variable $X-Y$ is

$$\begin{aligned}\phi_{X-Y}(t) &= \mathbb{E}\big(\exp(it(X-Y))\big) \\ &= \int_0^\infty \int_0^\infty \exp(it(x-y)) f_{X,Y}(x,y)\,dx\,dy = \phi_{X,Y}(t,-t).\end{aligned}$$

The derivation of Theorem 2.3.10 from Rubinstein and Sarnak's 2-dimensional central limit theorem then follows by taking X and Y to be the random variables having the same limiting distributions as $E(x;q,a)$ and $E(x;q,b)$, respectively (which implies that $X-Y = X_{q;a,b}$).

On the other hand, we note that our analysis of the variances of these distributions has the benefit of providing a better quantitative statement of the convergence of our limiting distributions to the Gaussian distribution : see equation (2.3.20) below.

PROOF OF THEOREM 2.3.10. Since the Fourier transform of the limiting logarithmic distribution of $E(x;q,a)-E(x;q,b)$ is $\hat{X}_{q;a,b}(\eta)$, the Fourier transform of the limiting logarithmic distribution of the quotient (2.3.18) is $\hat{X}_{q;a,b}(\eta/\sqrt{2\phi(q)\log q})$. A theorem of Lévy from 1925 [**66**, Section 4.2, Theorem 4], the Continuity Theorem for characteristic functions, asserts that all we need to show is that

$$\lim_{q\to\infty} \hat{X}_{q;a,b}\Big(\frac{\eta}{\sqrt{2\phi(q)\log q}}\Big) = e^{-\eta^2/2} \tag{2.3.19}$$

for every fixed real number η. Because the right-hand side is continuous at $\eta = 0$, it is automatically the characteristic function of the measure to which the limiting logarithmic distributions of the quotients (2.3.18) converge in distribution, according to Lévy's theorem.

When q is large enough in terms of η, we have $|\eta/\sqrt{2\phi(q)\log q}| \le \frac{1}{4}$. For such q, Proposition 2.2.6 implies that

$$\begin{aligned}
\log \hat{X}_{q;a,b}\Big(&\frac{\eta}{\sqrt{2\phi(q)\log q}}\Big) \\
&= \frac{V(q;a,b)}{2\phi(q)\log q}\frac{\eta^2}{2} + O\Big(\frac{(c(q,a)-c(q,b))|\eta|}{\sqrt{\phi(q)\log q}} + \frac{V(q;a,b)\eta^4}{(\phi(q)\log q)^2}\Big) \\
&= -\frac{\eta^2}{2} + O\Big(\frac{\eta^2\log\log q}{\log q} + \frac{|\eta|\rho(q)}{\sqrt{\phi(q)\log q}} + \frac{\eta^4}{\phi(q)\log q}\Big) \qquad (2.3.20)
\end{aligned}$$

using the asymptotic formula for $V(q;a,b)$ given in Proposition 2.3.4. Since η is fixed, this is enough to verify (2.3.19), which establishes the theorem. □

2.3.6. Racing quadratic nonresidues against quadratic residues

This section is devoted to understanding the effect of low-lying zeros of Dirichlet L-functions on prime number races between quadratic residues and quadratic nonresidues. This phenomenon has already been studied by many authors—see for instance [**6**]. Let q be an odd prime, and define $\pi(x;q,N) = \#\{p \le x\colon p \text{ is a quadratic nonresidue (mod } q)\}$ and $\pi(x;q,R) = \#\{p \le x\colon p \text{ is a quadratic residue (mod } q)\}$. Each of $\pi(x;q,N)$ and $\pi(x;q,R)$ is asymptotic to $\pi(x)/2$, but Chebyshev's bias predicts that the difference $\pi(x;q,N) - \pi(x;q,R)$, or equivalently the normalized difference

$$E(x;N,R) = \frac{\log x}{\sqrt{x}}\big(\pi(x;q,N) - \pi(x;q,R)\big),$$

is more often positive than negative.

Our methods lead to an asymptotic formula for $\delta(q;N,R)$, the logarithmic density of the set of real numbers $x \ge 1$ satisfying $\pi(x;q,N) > \pi(x;q,R)$, that explains the effect of low-lying zeros in a straightfoward and quantitative way. We sketch this application now.

First, define the random variable

$$X_{q;N,R} = 2 + 2\sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi_1)=0}} \frac{X_\gamma}{\sqrt{\frac{1}{4}+\gamma^2}},$$

where χ_1 is the unique quadratic character (mod q). Under GRH and LI, the distribution of $X_{q;N,R}$ is the same as the limiting distribution of the normalized error term $E(x;N,R)$. The methods of Section 2.3 then lead to an asymptotic formula analogous to equation (2.1.2) :

$$\delta(q;N,R) = \frac{1}{2} + \sqrt{\frac{2}{\pi V(q;N,R)}} + O\left(\frac{1}{V(q;N,R)^{3/2}}\right), \tag{2.3.21}$$

where

$$V(q;N,R) = b(\chi_1) = \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi_1)=0}} \frac{1}{\frac{1}{4}+\gamma^2}.$$

To simplify the discussion, we explore only the effect of the lowest zero (the zero closest to the real axis) on the size of $V(q;N,R)$.

By the classical formula for the zero-counting function $N(T,\chi)$, the average height of the lowest zero of $L(s,\chi_1)$ is $2\pi/\log q$. Suppose we have a lower-than-average zero, say at height $c\cdot 2\pi/\log q$ for some $0<c<1$. Then we get a higher-than-average contribution to the variance of size

$$\frac{1}{1/4+(c\cdot 2\pi/\log q)^2} - \frac{1}{1/4+(2\pi/\log q)^2}.$$

Since the variance $V(q;N,R) = b(\chi_1)$ is asymptotically $\log q$ by Lemma 2.3.2, this increases the variance by roughly a percentage t given by

$$t \sim \frac{1}{\log q}\left(\frac{1}{1/4+(c\cdot 2\pi/\log q)^2} - \frac{1}{1/4+(2\pi/\log q)^2}\right). \tag{2.3.22}$$

Therefore, given any two of the three parameters

- how low the lowest zero is (in terms of the percentage c of the average),
- how large a contribution we see to the variance (in terms of the percentage t), and
- the size of the modulus q,

we can determine the range for the third parameter from equation (2.3.22).

For example, as c tends to 0, the right-hand side of equation (2.3.22) is asymptotically

$$\frac{64\pi^2}{(\log^2 q + 16\pi^2)\log q}.$$

So if we want to see an increase in variance of 10%, an approximation for the range of q for which this might be possible is given by setting $64\pi^2/(\log^2 q + 16\pi^2)\log q = 0.1$ and solving for q, which gives $\log q = 15.66$ or about $q =$

6,300,000. This assumes that c tends to 0—in other words, that $L(s,\chi_1)$ has an extremely low zero. However, even taking $c = \frac{1}{3}$ on the right-hand side of equation (2.3.22) and setting the resulting expression equal to 0.1 yields about $q = 1{,}600{,}000$. In other words, having a zero that's only a third as high as the average zero, for example, will give a "noticeable" (at least 10%) lift to the variance up to roughly $q = 1{,}600{,}000$.

It turns out that unusually low zeros of this sort are not particularly rare. The Katz-Sarnak model predicts that the proportion of L-functions in the family $\{L(s,\chi)\colon \chi \text{ primitive of order 2}\}$ having a zero as low as $c \cdot 2\pi/\log q$ is asymptotically $2\pi^2c^3/9$ as c tends to 0. Continuing with our example value $c = \frac{1}{3}$, we see that roughly 8% of the moduli less than 1,600,000 will have a 10% lift in the variance $V(q;N,R)$ coming from the lowest-lying zero.

Well-known examples of L-functions having low-lying zeros are the $L(s,\chi_1)$ corresponding to prime moduli q for which the class number $h(-q)$ equals 1, as explained in **[6]** with the Chowla–Selberg formula for $q = 163$. For this modulus, the imaginary part of the lowest-lying zero is $0.202901\ldots = 0.16449\ldots \cdot 2\pi/\log 163$. According to our approximations, this low-lying zero increases the variance by roughly $t = 56\%$; considering this increased variance in equation (2.3.21) explains why the value of $\delta(163;N,R)$ is exceptionally low. The actual value of $\delta(163;N,R)$, along with some neighboring values, are shown in Table 2.2.

TAB. 2.2. Values of $\delta(q;N,R)$ for $q = 163$ and nearby primes

q	$\delta(q;N,R)$
151	0.745487
157	0.750767
163	0.590585
167	0.780096
173	0.659642

Other Dirichlet L-functions having low-lying zeros are the $L(s,\chi_1)$ corresponding to prime moduli q for which the class number $h(-q)$ is relatively small ; a good summary of the first few class numbers is given in [**6**, Table VI].

Notice that in principle, racing quadratic residues against quadratic nonresidues makes sense for any modulus q for which $\rho(q) = 2$, which includes powers of odd primes and twice these powers. However, being a quadratic residue modulo a prime q is exactly equivalent to being a quadratic residue modulo any power of q, and also (for odd numbers) exactly equivalent to being a quadratic residue modulo twice a power of q. Therefore $\delta(q; N, R) = \delta(q^k; N, R) = \delta(2q^k; N, R)$ for every odd prime q. The only other modulus for which $\rho(q) = 2$ is $q = 4$, which has been previously studied : Rubinstein and Sarnak [**68**] calculated that $\delta(4; N, R) = \delta(4; 3, 1) \approx 0.9959$.

2.4. FINE-SCALE DIFFERENCES AMONG RACES TO THE SAME MODULUS

In this section we probe the effect that the specific choice of residue classes a and b has on the density $\delta(q; a, b)$. We begin by proving Corollary 2.1.4, which isolates the quantitative influence of $\delta(q; a, b)$ on a and b from its dependence on q, in Section 2.4.1. We then dissect the relevant influence, namely the function $\Delta(q; a, b)$, showing how particular arithmetic properties of the residue classes a and b predictably affect the density ; three tables of computational data are included to illustrate these conclusions. In Section 2.4.2 we develop this theme even further, proving Theorem 2.4.1 and hence its implication Theorem 2.1.5, which establishes a lasting “meta-bias” among these densities. Finally, in Section 2.4.3 we apply our techniques to the seemingly unrelated “mirror image phenomenon” observed by Bays and Hudson, explaining its existence with a similar analysis.

2.4.1. The impact of the residue classes a and b

The work of the previous sections has provided us with all the tools we need to establish Corollary 2.1.4.

PROOF OF COROLLARY 2.1.4. We begin by showing that the function

$$\Delta(q;a,b) = K_q(a-b) + \iota_q(-ab^{-1})\log 2 + \frac{\Lambda(r_1)}{r_1} + \frac{\Lambda(r_2)}{r_2} + H(q;a,b)$$

defined in equation (2.1.4) is bounded above by an absolute constant (the fact that it is nonnegative is immediate from the definitions of its constituent parts). It has already been remarked in Definition 2.1.3 that K_q is uniformly bounded, as is ι_q. We also have $\Lambda(r)/r \le (\log r)/r$, and this function is decreasing for $r \ge 3$, so the third and fourth terms are each uniformly bounded as well. Finally, from Definition 2.1.5, we see that

$$h(q;p,r) = \frac{1}{\phi(p^\nu)}\frac{\log p}{p^{e(q;p,r)}} \le \frac{1}{p-1}\frac{\log p}{p^1},$$

and so $H(q;a,b) < \sum_p 2(\log p)/p(p-1)$ is uniformly bounded by a convergent sum as well.

We now turn to the main assertion of the corollary. By Theorems 2.1.2 and 2.1.3, we have

$$\begin{aligned} V(q;a,b) &= 2\phi(q)\big(\mathcal{L}(q) + K_q(a-b) + \iota_q(-ab^{-1})\log 2\big) + 2M^*(q;a,b) \\ &= 2\phi(q)\bigg(\mathcal{L}(q) + K_q(a-b) + \iota_q(-ab^{-1})\log 2 + \frac{\Lambda(r_1)}{r_1} + \frac{\Lambda(r_2)}{r_2} \\ &\qquad\qquad + H(q;a,b) + O\Big(\frac{\log^2 q}{q}\Big)\bigg) \\ &= 2\phi(q)\bigg(\mathcal{L}(q) + \Delta(q;a,b) + O\Big(\frac{\log^2 q}{q}\Big)\bigg) \\ &= 2\phi(q)\mathcal{L}(q)\bigg(1 + \frac{\Delta(q;a,b)}{\mathcal{L}(q)} + O\Big(\frac{\log q}{q}\Big)\bigg). \end{aligned}$$

Since $\Delta(q;a,b)$ is bounded while $\mathcal{L}(q) \sim \log q$, we see that $V(q) \sim 2\phi(q)\log q$; moreover, the power series expansion of $(1+t)^{-1/2}$ around $t=0$ implies that

$$\begin{aligned} V(q;a,b)^{-1/2} &= \big(2\phi(q)\mathcal{L}(q)\big)^{-1/2}\bigg(1 - \frac{\Delta(q;a,b)}{2\mathcal{L}(q)} + O\Big(\frac{\Delta(q;a,b)^2}{\mathcal{L}(q)^2} + \frac{\log q}{q}\Big)\bigg) \\ &= \big(2\phi(q)\mathcal{L}(q)\big)^{-1/2}\bigg(1 - \frac{\Delta(q;a,b)}{2\mathcal{L}(q)} + O\Big(\frac{1}{\log^2 q}\Big)\bigg). \end{aligned}$$

(Recall that we are assuming that $q \geq 43$, which is enough to ensure that $\mathcal{L}(q)$ is positive.) Together with the last assertion of Theorem 2.1.1, this formula implies that

$$\delta(q;a,b) = \frac{1}{2} + \frac{\rho(q)}{2\sqrt{\pi\phi(q)\mathcal{L}(q)}}\biggl(1 - \frac{\Delta(q;a,b)}{2\mathcal{L}(q)} + O\Bigl(\frac{1}{\log^2 q}\Bigr)\biggr) + O\biggl(\frac{\rho(q)^3}{V(q;a,b)^{3/2}}\biggr).$$

Since the last error term is $\ll_\varepsilon q^\varepsilon/(\phi(q)\log q)^{3/2}$, it can be subsumed into the first error term, and the proof of the corollary is complete. □

Corollary 2.1.4 tells us that larger values of $\Delta(q;a,b)$ lead to smaller values of the density $\delta(q;a,b)$. Computations of the values of $\delta(q;a,b)$ (using methods described in Section 2.5.4) illustrate this relationship nicely. Since $\delta(q;a,b) = \delta(q;ab^{-1},1)$ when b is a square (mod q), we restrict our attention to densities of the form $\delta(q;a,1)$.

We begin by investigating a prime modulus q, noting that

$$\Delta(q;a,1) = \iota_q(-a)\log 2 + \frac{\Lambda(a)}{a} + \frac{\Lambda(a^{-1})}{a^{-1}} + \frac{2\log q}{q(q-1)}$$

when q is prime (here a^{-1} denotes the smallest positive integer that is a multiplicative inverse of a (mod q)). Therefore we obtain the largest value of $\Delta(q;a,b)$ when $a \equiv -1$ (mod q), and the next largest values are when a is a small prime, so that the $\Lambda(a)/a$ term is large. (These next large values also occur when a^{-1} is a small prime, and in fact we already know that $\delta(q;a,1) = \delta(q;a^{-1},1)$. When q is large, it is impossible for both a and a^{-1} to be small.) Notice that $\Lambda(a)/a$ is generally decreasing on primes a, except that $\Lambda(3)/3 > \Lambda(2)/2$. Therefore the second, third, and fourth-largest values of $\Delta(q;a,1)$ will occur for a congruent to 3, 2, and 5 (mod q), respectively.

This effect is quite visible in the calculated data. We use the prime modulus $q = 163$ as an example, since the smallest 12 primes, as well as -1, are all nonsquares (mod 163). Table 2.3 lists the values of all densities of the form $\delta(163,a,1)$ (remembering that $\delta(q;a,1) = \delta(q;a^{-1},1)$ and that the value of any $\delta(q;a,b)$ is equal to one of these). Even though the relationship between

TAB. 2.3. The densities $\delta(q;a,1)$ computed for $q = 163$

q	a	a^{-1}	$\delta(q;a,1)$	q	a	a^{-1}	$\delta(q;a,1)$
163	162	162	0.524032	163	30	125	0.526809
163	3	109	0.525168	163	76	148	0.526815
163	2	82	0.525370	163	92	101	0.526829
163	5	98	0.525428	163	86	127	0.526869
163	7	70	0.525664	163	128	149	0.526879
163	11	89	0.525744	163	129	139	0.526879
163	13	138	0.526079	163	80	108	0.526894
163	17	48	0.526083	163	114	153	0.526898
163	19	103	0.526090	163	117	124	0.526900
163	23	78	0.526213	163	20	106	0.526906
163	31	142	0.526378	163	42	66	0.526912
163	67	73	0.526437	163	28	99	0.526914
163	37	141	0.526510	163	44	63	0.526925
163	29	45	0.526532	163	12	68	0.526931
163	27	157	0.526578	163	72	120	0.526941
163	32	107	0.526586	163	112	147	0.526975
163	59	105	0.526620	163	110	123	0.526981
163	8	102	0.526638	163	122	159	0.526996
163	79	130	0.526682	163	50	75	0.526997
163	94	137	0.526746	163	52	116	0.527002
163	18	154	0.526768	163			

$\Delta(q;a,1)$ and $\delta(q;a,1)$ given in Corollary 2.1.4 involves an error term, the data is striking. The smallest ten values of $\delta(q;a,1)$ are exactly in the order predicted by our analysis of $\Delta(q;a,1)$: the smallest is $a = 162 \equiv -1 \pmod{163}$, then $a = 3$ and $a = 2$, then the seven next smallest primes in order. (This ordering, which is clearly related to Theorem 2.1.5, will be seen again in Figure 2.2.)

One can also probe more closely the effect of the term $M(q;a,1)$ upon the density $\delta(q;a,1)$. Equation (2.3.10) can be rewritten as the approximation

$$\frac{M(q;a,1)}{\phi(q)} + 2 \sum_{\substack{n\le y \\ n\equiv 1 \pmod q}} \frac{\Lambda(n)}{n} \approx \sum_{\substack{n\le y \\ n\equiv a \pmod q}} \frac{\Lambda(n)}{n} + \sum_{\substack{n\le y \\ n\equiv a^{-1} \pmod q}} \frac{\Lambda(n)}{n} \qquad (2.4.1)$$

TAB. 2.4. The effect of medium-sized prime powers on the densities $\delta(q; a, 1)$, illustrated with $q = 101$

a	a^{-1}	First four prime powers				RHS of (2.4.1)	$\delta(101, a, 1)$
7	29	7	29	433	512	0.563304	0.534839
2	51	2	103	709	859	0.554043	0.534928
3	34	3	337	811	1013	0.528385	0.535103
11	46	11	349	617	1021	0.383090	0.536123
8	38	8	109	139	311	0.332888	0.536499
53	61	53	61	263	457	0.329038	0.536522
12	59	59	113	463	719	0.276048	0.536955
67	98	67	199	269	401	0.271567	0.536993
41	69	41	243	271	647	0.268766	0.537013
28	83	83	331	487	937	0.235130	0.537284
15	27	27	128	229	419	0.235035	0.537293
66	75	167	277	479	571	0.230291	0.537340
18	73	73	523	881	1129	0.215281	0.537463
50	99	151	353	503	757	0.211209	0.537500
55	90	191	257	661	797	0.205833	0.537537
42	89	89	547	1153	1301	0.202289	0.537586
44	62	163	347	751	769	0.199652	0.537607
72	94	173	397	577	599	0.196417	0.537623
32	60	32	739	941	1171	0.191447	0.537660
26	35	127	439	641	733	0.190601	0.537688
39	57	241	443	461	1049	0.187848	0.537708
40	48	149	343	1151	1361	0.178698	0.537780
10	91	293	313	919	1303	0.180422	0.537792
74	86	389	983	1399	1601	0.165153	0.537900
63	93	467	1103	1709	2083	0.146466	0.538067

(where we are ignoring the exact form of the error term). Taking $y = q$ recovers the approximation $M(q; a, 1) \approx \phi(q)\big(\Lambda(a)/a + \Lambda(a^{-1})/a^{-1}\big)$ used in the definition of $\Delta(q; a, b)$, but taking y larger would result in a better approximation.

We examine this effect on the calculated densities for the medium-sized prime modulus $q = 101$. In Table 2.4, the second group of columns records the

Tab. 2.5. The densities $\delta(q; a, 1)$ computed for $q = 420$, together with the values of $K_q(a-1) = \Lambda(q/(q, a-1))/\phi(q/(q, a-1))$

q	a	a^{-1}	$(q, a-1)$	$K_q(a-1)$	$\delta(q; a, 1)$	q	a	a^{-1}	$(q, a-1)$	$K_q(a-1)$	$\delta(q; a, 1)$
420	211	211	210	$\log 2$	0.770742	420	113	197	28	0	0.807031
420	419	419	2	0	0.772085	420	149	389	4	0	0.807209
420	281	281	140	$(\log 3)/2$	0.779470	420	103	367	6	0	0.807284
420	253	337	84	$(\log 5)/4$	0.788271	420	223	307	6	0	0.807302
420	61	241	60	$(\log 7)/6$	0.788920	420	83	167	2	0	0.807505
420	181	181	60	$(\log 7)/6$	0.789192	420	151	331	30	0	0.809031
420	17	173	4	0	0.795603	420	59	299	2	0	0.809639
420	47	143	2	0	0.796173	420	137	233	4	0	0.809647
420	29	29	28	0	0.796943	420	139	139	6	0	0.810290
420	13	97	12	0	0.797669	420	73	397	12	0	0.811004
420	187	283	6	0	0.797855	420	157	313	12	0	0.811197
420	53	317	4	0	0.798207	420	251	251	10	0	0.811557
420	11	191	10	0	0.798316	420	349	349	12	0	0.811706
420	107	263	2	0	0.798691	420	323	407	14	0	0.811752
420	41	41	20	0	0.800067	420	179	359	2	0	0.811765
420	19	199	6	0	0.800937	420	229	409	12	0	0.811776
420	43	127	42	0	0.801609	420	131	311	10	0	0.811913
420	23	347	2	0	0.802681	420	277	373	12	0	0.812052
420	37	193	12	0	0.803757	420	239	239	14	0	0.812215
420	79	319	6	0	0.804798	420	247	403	6	0	0.812215
420	89	269	4	0	0.804836	420	227	383	2	0	0.812777
420	101	341	20	0	0.805089	420	221	401	20	0	0.813594
420	71	71	70	0	0.805123	420	293	377	4	0	0.813793
420	67	163	6	0	0.805196	420	379	379	42	0	0.813818
420	31	271	30	0	0.806076	420	209	209	4	0	0.815037
420	257	353	4	0	0.806638	420	391	391	30	0	0.815604

first four prime powers that are congruent to a or a^{-1} (mod 101). The second-to-last column gives the value of the right-hand side of equation (2.4.1), computed at $y = 10^6$. Note that smaller prime powers in the second group of columns give large contributions to this second-to-last column, a trend that can be visually confirmed. Finally, the last column lists the values of the densities $\delta(q; a, b)$, according to which the rows have been sorted in ascending order. The correlation between larger values of the second-to-last column and smaller values of $\delta(q; a, b)$ is almost perfect (the adjacent entries $a = 40$ and $a = 10$ being the only exception) : the existence of smaller primes and prime powers in the residue classes a and a^{-1} (mod 101) really does contribute positively to the variance $V(q; a, 1)$ and hence decreases the density $\delta(q; a, 1)$. (Note that the effect of the term $\iota_{101}(-a) \log 2$ is not present here, since 101 is a prime congruent to 1 (mod 4) and hence -1 is not a nonsquare.)

Finally we investigate a highly composite modulus q to witness the effect of the term $K_q(a-1) = \Lambda(q/(q,a-1))/\phi(q/(q,a-1)) - \Lambda(q)/\phi(q)$ on the size of $\Delta(q;a,1)$. This expression vanishes unless $a-1$ has such a large factor in common with q that the quotient $q/(q,a-1)$ is a prime power. Therefore we see a larger value of $\Delta(q;a,1)$, and hence expect to see a smaller value of $\delta(q;a,1)$, when $q/(q,a-1)$ is a small prime, for example when $a = \frac{q}{2}+1$.

Table 2.5 confirms this observation with the modulus $q = 420$. Of the six smallest densities $\delta(420;a,1)$, five of them correspond to the residue classes a (and their inverses) for which $q/(q,a-1)$ is a prime power ; the sixth corresponds to $a \equiv -1 \pmod{420}$, echoing the effect already seen for $q = 163$. Moreover, the ordering of these first six densities are exactly as predicted : even the battle for smallest density between $a \equiv -1 \pmod{420}$ and $a = 420/2-1$ is appropriate, since both residue classes cause an increase in $\Delta(420;a,1)$ of size exactly $\log 2$. (Since 420 is divisible by the four smallest primes, the largest effect that the $\Lambda(a)/a$ term could have on $\Delta(q;a,b)$ is $(\log 11)/11$, and so these effects are not nearly as large.) The magnitude of this effect is quite significant : note that the difference between the first and seventh-smallest values of $\delta(420;a,1)$ (from $a = 211$ to $a = 17$) is larger than the spread of the largest 46 values (from $a = 17$ to $a = 391$).

2.4.2. The predictability of the relative sizes of densities

The specificity of our asymptotic formulas to this point suggests comparing, for fixed integers a_1 and a_2, the densities $\delta(q;a_1,1)$ and $\delta(q;a_2,1)$ as q runs through all moduli for which both a_1 and a_2 are nonsquares. (We have already seen that every density is equal to one of the form $\delta(q;a,1)$.) Theorem 2.1.5, which we will derive shortly from Corollary 2.4.2, is a statement about exactly this sort of comparison.

In fact we can investigate even more general families of race games : fix rwo rational numbers r and s, and consider the family of densities $\delta(q;r+sq,1)$ as q varies. We need $r+sq$ to be an integer and relatively prime to q for this density to be sensible ; we further desire $r+sq$ to be a nonsquare (mod q), or else $\delta(q; r + sq, 1)$ simply equals $\frac{1}{2}$. Therefore, we define the set of qualified moduli

$$\begin{aligned} Q(r,s) = \{q \in \mathbb{N} \colon r + sq \in \mathbb{Z},\ (r + sq, q) = 1; \\ \text{there are no solutions to } x^2 \equiv r + sq \pmod{q}\}. \end{aligned}$$

(Note that translating s by an integer does not change the residue class of $r + sq \pmod{q}$, so one could restrict s to the interval $[0, 1)$ without losing generality if desired.)

It turns out that every pair (r, s) of rational numbers can be assigned a "rating" $R(r, s)$ that dictates how the densities in the family $\delta(q; r + sq, 1)$ compare to other densities in similar families.

Definition 2.4.1. *Define a rating function* $R(r, s)$ *as follows :*

- *Suppose that the denominator of* s *is a prime power* p^k *(*$k \geq 1$*).*
  - *If* r *is a power* p^j *of the same prime, then* $R(r, s) = (\log p)/\phi(p^{j+k})$.
  - *If* $r = 1$ *or* $r = 1/p^j$ *for some* $1 \leq j < k$*, then* $R(r, s) = (\log p)/\phi(p^k)$.
  - *If* $r = 1/p^k$*, then* $R(r, s) = (\log p)/p^k$.
  - *Otherwise* $R(r, s) = 0$.
- *Suppose that* s *is an integer.*
  - *If* $r = -1$*, then* $R(r, s) = \log 2$.
  - *If* r *is a prime power* p^j $(j \geq 1)$*, then* $R(r, s) = (\log p)/p^j$.
  - *Otherwise* $R(r, s) = 0$.
- $R(r, s) = 0$ *for all other values of* s. ◊

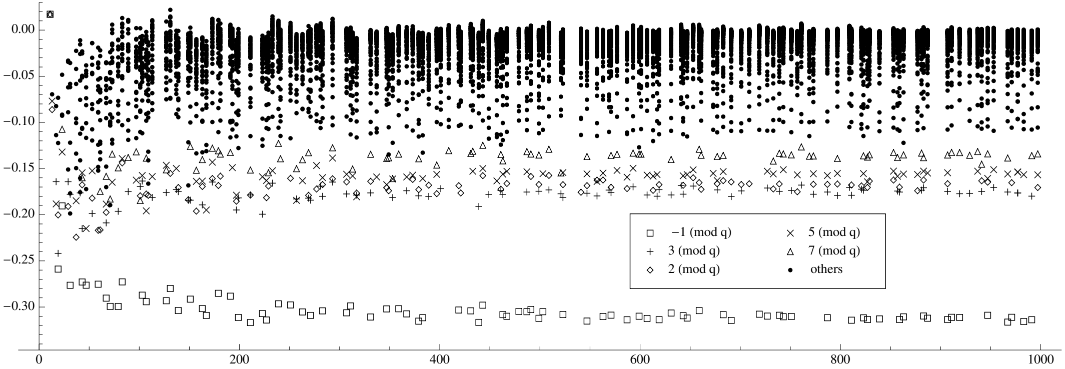
Theorem 2.4.1. *Let* $\Delta(q; a, b)$ *be defined as in equation* (2.1.4)*. For fixed rational numbers* r *and* s*,*

$$\Delta(q; r + sq, 1) = R(r, s) + O_{r,s}\left(\frac{\log q}{q}\right)$$

as q *tends to infinity within the set* $Q(r, s)$.

We will be able to prove this theorem at the end of the section ; first, however, we note an interesting corollary.

FIG. 2.2. Normalized densities $\delta(q; a, 1)$ for primes q, using the normalization (2.4.3)



Corollary 2.4.2. *Assume GRH and LI. If* r_1, s_1, r_2, s_2 *are rational numbers such that* $R(r_1, s_1) > R(r_2, s_2)$*, then*

$\delta(q; r_1{+}s_1q, 1) < \delta(q; r_2{+}s_2q, 1)$ *for all but finitely many* $q \in Q(r_1, s_1) \cap Q(r_2, s_2)$.

PROOF. We may assume that $q \geq 43$. Inserting the conclusion of Theorem 2.4.1 into the formula for $\delta(q; a, b)$ in Corollary 2.1.4, we obtain

$$\delta(q; r + sq, 1) = \frac{1}{2} + \frac{\rho(q)}{2\sqrt{\pi\phi(q)\mathcal{L}(q)}}\left(1 - \frac{R(r, s)}{2\mathcal{L}(q)} + O\left(\frac{1}{\log^2 q}\right)\right) \tag{2.4.2}$$

for any $q \in Q(r, s)$. Therefore for all $q \in Q(r_1, s_1) \cap Q(r_2, s_2)$,

$$\begin{aligned}\delta(q; r_1 + s_1q, 1) &- \delta(q; r_2 + s_2q, 1) \\ &= \left(\frac{-R(r_1, s_1) + R(r_2, s_2)}{2\mathcal{L}(q)} + O\left(\frac{1}{\log^2 q}\right)\right)\frac{\rho(q)}{2\sqrt{\pi\phi(q)\mathcal{L}(q)}}.\end{aligned}$$

Since the constant $-R(r_1, s_1) + R(r_2, s_2)$ is negative by hypothesis, we see that $\delta(q; r_1{+}s_1q, 1) - \delta(q; r_2{+}s_2q, 1)$ is negative when q is sufficiently large in terms of r_1, s_1, r_2, and s_2. □

Notice, from the part of Definition 2.4.1 where s is an integer, that Theorem 2.1.5 is precisely the special case of Corollary 2.4.2 where $s_1 = s_2 = 0$. Therefore we have reduced Theorem 2.1.5 to proving Theorem 2.4.1.

Theorem 2.1.5 itself is illustrated in Figure 2.2, using the computed densities for prime moduli to most clearly observe the relevant phenomenon. For each prime q up to 1000, and for every nonsquare a (mod q), the point

$$\bigg(q, \frac{2\sqrt{\pi\phi(q)\mathcal{L}(q)^3}}{\rho(q)}\big(\delta(q;a,1)-\tfrac12\big)-\mathcal{L}(q)\bigg) \\ = \bigg(q, \sqrt{\pi(q-1)}\bigg(\log\frac{q}{2\pi e^{\gamma_0}}\bigg)^{3/2}\big(\delta(q;a,1)-\tfrac12\big)-\log\frac{q}{2\pi e^{\gamma_0}}\bigg) \qquad (2.4.3)$$

has been plotted ; the values corresponding to certain residue classes have been emphasized with the listed symbols. The motivation for the seemingly strange (though order-preserving) normalization in the second coordinate is equation (2.4.2), which shows that the value in the second coordinate is $-R(a,0)/2 + O(1/\log q)$. In other words, on the vertical axis the value 0 corresponds to $\delta(q;a,b)$ being exactly the "default" value $\frac12 + \rho(q)/2\sqrt{\pi\phi(q)\mathcal{L}(q)}$, the value -0.05 corresponds to $\delta(q;a,b)$ being less than the default value by $0.05\rho(q)/2\sqrt{\pi\phi(q)\mathcal{L}(q)^3}$, and so on. We clearly see in Figure 2.2 the normalized values corresponding to $\delta(q;-1,1)$, $\delta(q;3,1)$, $\delta(q;2,1)$, and so on sorting themselves out into rows converging on the values $-\frac12\log 2$, $-\frac16\log 3$, $-\frac14\log 2$, and so on.

We need to establish several lemmas before we can prove Theorem 2.4.1.

The recurring theme in the following analysis is that solutions to linear congruences (mod q) with fixed coefficients must be at least a constant times q in size, save for specific exceptions that can be catalogued.

Lemma 2.4.1. *Let r and s be rational numbers. If $r \ne -1$ or if s is not an integer, then there are only finitely many positive integers q such that $r+sq$ is an integer and $r+sq \equiv -1$ (mod q).*

PROOF. Write $r = \frac{a}{b}$ and $s = \frac{c}{d}$. The congruence $\frac{a}{b} + \frac{c}{d}q \equiv -1$ (mod q) implies that $\frac{ad}{b} + cq \equiv -d$ (mod q), which means that q must divide $\frac{ad}{b} + d$. This only happens for finitely many q unless $\frac{ad}{b}+d = 0$, which is equivalent (since $d \ne 0$) to $\frac{a}{b} = -1$. In this case the congruence is $\frac{c}{d}q \equiv 0$ (mod q), which can happen only if $\frac{c}{d}$ is an integer. □

Lemma 2.4.2. *Let r and s be rational numbers. Suppose that q is a positive integer such that $r+sq$ is an integer. If $r \ne 1$, then $K_q(r+sq-1) \ll_{r,s} (\log q)/q$.*

PROOF. We first note that $\Lambda(t)/\phi(t) \ll (\log t)/t$ for all positive integers t : if $\Lambda(t)$ is nonzero, then t is a prime power, which means $\phi(t) \geq t/2$. Therefore it suffices to show that $(q, r+sq-1)$ is bounded, since then $q/(q, r+sq-1) \gg_{r,s} q$ and consequently $K_q(r+sq-1) \ll_{r,s} (\log q)/q$ since $(\log t)/t$ is decreasing for $t \geq 3$. But writing $r = \frac{a}{b}$ and $s = \frac{c}{d}$, we have

$$\left(q, \tfrac{a}{b} + \tfrac{c}{d}q - 1\right) \mid (q, d(a-b) + bcq) = (q, d(a-b)) \mid d(a-b).$$

Since $r \neq 1$, we see that $d(a-b)$ is nonzero, and hence $(q, r+sq-1) \leq d|a-b| \ll_{r,s} 1$ as required. □

Lemma 2.4.3. *Let r and s be rational numbers. Assume that r is not a positive integer or s is not an integer. If q and y are positive integers such that $r+sq$ is an integer and $y \equiv r+sq \pmod q$, then $y \gg_{r,s} q$.*

PROOF. Suppose first that s is not an integer, and write $s = c/d$ where $d > 1$. Then s is at least $1/d$ away from the nearest integer, so that sq is at least q/d away from the nearest multiple of q. Since $y = r + sq - mq$ for some integer m, we have $y \geq |sq - mq| - |r| \geq q/d - |r| \gg_{r,s} q$ when q is sufficiently large in terms of r and s.

On the other hand, if s is an integer, then r must also be an integer. If r is nonpositive, then the least integer y congruent to $r \pmod q$ is $q - |r| \gg_r q$ when q is sufficiently large in terms of r. □

Lemma 2.4.4. *Let r and s be rational numbers. Assume that either r is not the reciprocal of a positive integer or that $\frac{s}{r}$ is not an integer. Suppose that positive integers q and y are given such that $r+sq$ is an integer and $(r+sq)y \equiv 1 \pmod q$. Then $y \gg_{r,s} q$.*

PROOF. Write $r = \frac{a}{b}$ and $s = \frac{c}{d}$ with $(a,b) = (c,d) = 1$ and $b, d > 0$. We may assume that $q > 2d^2$, for if $q \leq 2d^2$ then $y \geq 1 \geq \frac{q}{2d^2} \gg_s q$. Note that $a \neq 0$, since $0 + \frac{c}{d}q = c\frac{q}{d}$ cannot be invertible modulo q when $q > d$. The assumption that $r+sq$ is an integer implies that $d(r+sq) = \frac{ad}{b} + cq$ is also an integer ; since $(a, b) = 1$, this implies that $b \mid d$. Therefore we may write $d = b\delta$ for some integer δ. Similarly, it must be true that $b(r + sq) = a + \frac{cq}{\delta}$ is an integer ; since $(c, \delta) \mid (c, d) = 1$, this implies that q is a multiple of δ.

Case 1 : Suppose first that $\delta = 1$. If $a = 1$, then $r = \frac{1}{b}$ would be the reciprocal of a positive integer and $\frac{s}{r} = \frac{c/b}{1/b}$ would be an integer, contrary to assumption ; therefore $a \neq 1$. The condition $(r + sq)y \equiv 1 \pmod{q}$, when multiplied by b, becomes $ay \equiv b \pmod{q}$. Now if $a = -1$, then the congruence in question is equivalent to $y \equiv -b \pmod{q}$; since $b > 0$, this implies that $y \geq q - b \gg_r q$ as desired. Therefore for the rest of Case 1, we can assume that $|a| > 1$.

Since any common factor of a and q would consequently be a factor of b as well, but $(a, b) = 1$, we must have $(a, q) = 1$. Thus we may choose u such that $uq \equiv -1 \pmod{a}$, so that $y_0 = b(uq + 1)/a$ is an integer. We see by direct calculation that y_0 is a solution to $ay \equiv b \pmod{q}$, and all other solutions differ from this one by a multiple of $q/(b, q)$, which is certainly a multiple of $\frac{q}{b}$. In other words, $y = q(\frac{bu}{a} + \frac{z}{b}) + \frac{b}{a}$ for some integer z. If $\frac{bu}{a} + \frac{z}{b} = 0$ then $-z = b(\frac{bu}{a} + \frac{z}{b}) - z = \frac{b^2 u}{a}$ would be an integer, but this is impossible since both b and u are relatively prime to a (here we use $|a| \neq 1$). Therefore $\left|\frac{bu}{a} + \frac{z}{b}\right| \geq \frac{1}{|a|b}$, and so $y \geq \frac{q}{|a|b} - \frac{b}{|a|}$; since $q > 2d^2 = 2b^2$, this gives $y \geq \frac{q}{2|a|b} \gg_{r,s} q$.

Case 2 : Suppose now that $\delta > 1$. The condition $(\frac{a}{b} + \frac{c}{d}q)y \equiv 1 \pmod{q}$ forces $(y, q) = 1$ and so $(y, \delta) = 1$ as well. Multiplying the condition by b yields $ay + cy\frac{q}{\delta} \equiv b \pmod{q}$, which we write as $\frac{cyq}{\delta} - qm = b - ay$ for some integer m. But notice that $(cy, \delta) = 1$, so that $\frac{cy}{\delta}$ is at least $\frac{1}{\delta}$ away from every integer (here we use $\delta > 1$) ; therefore $\frac{cyq}{\delta}$ is at least $\frac{q}{\delta}$ away from the nearest multiple of q. Therefore $\frac{q}{\delta} \leq \left|\frac{cyq}{\delta} - qm\right| = |b - ay| \leq b + |a|y$, and hence $y \geq (q - b\delta)/|a|\delta$; since $q > 2d = 2b\delta$, this gives $y \geq \frac{q}{2|a|\delta} \gg_{r,s} q$. □

Corollary 2.4.3. *Let r and s be rational numbers, and let q be a positive integer such that $r + sq$ is an integer.*

(1) Assume that r is not a positive integer or s is not an integer. Suppose that y is a positive integer such that $y \equiv r + sq \pmod{q}$. Then $\Lambda(y)/y \ll_{r,s} (\log q)/q$.

(2) *Assume that either r is not the reciprocal of a positive integer or that $\frac{s}{r}$ is not an integer. Suppose that y is a positive integer such that $(r+sq)y \equiv 1 \pmod q$. Then $\Lambda(y)/y \ll_{r,s} (\log q)/q$.*

PROOF. Since $\Lambda(y)/y \leq (\log y)/y$, which is a decreasing function for $y \geq 3$, this follows from Lemmas 2.4.3 and 2.4.4. □

Lemma 2.4.5. *Let r and s be rational numbers. Let q be a positive integer such that $r+sq$ is an integer, and let p be a prime such that $p^\nu \parallel q$ with $\nu \geq 1$.*

(a) *Suppose that e is a positive integer such that $p^e \equiv r+sq \pmod{q/p^\nu}$. Then either $p^e = r$ or $p^e \gg_{r,s} q/p^\nu$.*

(b) *Suppose that e is a positive integer such that $p^e(r+sq) \equiv 1 \pmod{q/p^\nu}$. Then either $p^e = 1/r$ or $p^e \gg_{r,s} q/p^\nu$.*

Notice that if $p^e = r$ in (a) then sp^ν is an integer ; also, if $p^e = 1/r$ in (b) then $sp^{e+\nu}$ is an integer. In both cases, it is necessary that the denominator of s be a power of p as well.

PROOF. We may assume that q/p^ν is sufficiently large in terms of r and s, for otherwise any positive integer is $\gg_{r,s} q/p^\nu$. We have two cases to examine.

(a) We are assuming that $p^e \equiv r+sq \pmod{q/p^\nu}$. Suppose first that sp^ν is an integer. Then sq is an integer multiple of q/p^ν, and so $p^e \equiv r \pmod{q/p^\nu}$. This means that either $p^e = r$ or $p^e \geq q/p^\nu + r \gg_r q/p^\nu$, since q/p^ν is sufficiently large in terms of r.

On the other hand, suppose that sp^ν is not an integer. Then

$$p^e \equiv r+sq = r+(sp^\nu)q/p^\nu \equiv r+(sp^\nu - \lfloor sp^\nu \rfloor)q/p^\nu \pmod{q/p^\nu}.$$

If the denominator of s is d, then the difference $sp^\nu - \lfloor sp^\nu \rfloor$ is at least $\frac{1}{d}$, and therefore $p^e \geq q/dp^\nu + r \gg_{r,s} q/p^\nu$ as well, since q/p^ν is sufficiently large in terms of r and s.

(b) We are assuming that $p^e(r+sq) \equiv 1 \pmod{q/p^\nu}$. We apply Lemma 2.4.4 with q/p^ν in place of q and with $y = p^e$, which yields the desired lower bound $p^e \gg_{r,s} q/p^\nu$ unless r is the reciprocal of a positive integer and $\frac{s}{r}$ is an integer. In this case, multiplying the assumed congruence by the integer $1/r$ gives $p^e(1+\frac{s}{r}q) \equiv 1/r \pmod{q/p^\nu}$, which implies $p^e \equiv 1/r \pmod{q/p^\nu}$ since $\frac{s}{r}$ is an integer. Therefore, since q/p^ν is sufficiently large in terms of r, either $p^e = 1/r$ or $p^e \geq q/p^\nu + 1/r > q/p^\nu$.

□

The next two lemmas involve the functions $h(q;p,r)$ and $H(q;a,b)$ that were defined in Definition 2.1.5. Since we are dealing with rational numbers, we make the following clarification : when we say "power of p", we mean p^k for some *positive* integer k (so p^2 and p^1 are powers of p, but neither 1 nor p^{-1} is).

Lemma 2.4.6. *Let r and s be rational numbers, and suppose that q is a positive integer such that $r+sq$ is an integer that is relatively prime to q. Let p be a prime dividing q, and choose $\nu \geq 1$ such that $p^\nu \parallel q$.*

(a) If both r and the denominator of s are powers of p (note that if the denominator of s equals p^k, these conditions imply $\nu = k$), then

$$h(q;p,(r+sq)^{-1}) = \frac{\log p}{r\phi(p^\nu)} + O_{r,s}\left(\frac{\log p}{q}\right);$$

otherwise $h(q;p,(r+sq)^{-1}) \ll_{r,s} (\log p)/q$.

(b) If both $1/r$ and the denominator of s are powers of p (note that if $r = 1/p^j$ and the denominator of s equals p^k, these conditions imply $\nu = k - j$), then

$$h(q;p,r+sq) = \frac{r\log p}{\phi(p^\nu)} + O_{r,s}\left(\frac{\log p}{q}\right),$$

otherwise $h(q;p,r+sq) \ll_{r,s} (\log p)/q$.

PROOF. (a) Assume $p^e \equiv r+sq \pmod{q/p^\nu}$. By Lemma 2.4.5, we have that either $r = p^e$ (which implies that the denominator of s is a power of p), or else $h(q;p,(r+sq)^{-1}) \ll_{r,s} (\log p)/q$. So we only need to compute $h(q;p,(r+sq)^{-1})$ in the case where r is any power of p (say $r = p^e$) and where s has a denominator which is a power of p (say $s = c/p^z$, where $z \leq \nu$ since q is a multiple of the denominator of s).

In this case the congruence $p^e \equiv r + sq \pmod{q/p^\nu}$ is satisfied. Furthermore, e is the minimal such positive integer if q is sufficiently large in terms of r and s. If e is minimal we have $h(q; p, (r + sq)^{-1}) = (\log p)/\phi(p^\nu)p^e = (\log p)/r\phi(p^\nu)$ by definition; if e is not minimal we have $h(q; p, (r+sq)^{-1}) \ll_{r,s} (\log p)/q$ since there are only finitely many possible values of q. In both cases, the proposition is established (the "main term" $(\log p)/r\phi(p^\nu)$ is actually dominated by the error term in the latter case).

(b) Assume $p^e(r+sq) \equiv 1 \pmod{q/p^\nu}$. By Lemma 2.4.5, we have that either $1/r = p^e$ (which implies that the denominator of s is a power of p), or else $h(q; p, r+sq) \ll_{r,s} (\log p)/q$. So we only need to compute $h(q; p, r+sq)$ in the case where $1/r$ is any power of p (say $1/r = p^j$) and where s has a denominator which is a power of p (say $s = c/p^k$, where $k - j = \nu > 0$).

In this case the congruence $p^j(r + sq) \equiv 1 \pmod{q/p^\nu}$ is satisfied (since $p^j sq \equiv 0 \pmod{q/p^\nu}$). We can rewrite this congruence as $p^j \equiv 1/r \pmod{q/p^\nu}$. As above, either j is the minimal such positive integer, in which case $h(q; p, r + sq) = (\log p)/\phi(p^\nu)p^e = (r \log p)/\phi(p^\nu)$ by definition, or else q is bounded in terms of r and s, in which case $h(q; p, r+sq) \ll_{r,s} (\log p)/q$. In both cases, the proposition is established.

□

Corollary 2.4.4. *Let r and s be rational numbers, and suppose that q is a positive integer such that $r + sq$ is an integer that is relatively prime to q.*

(a) Suppose both r and the denominator of s are powers of the same prime p. Then

$$H(q; r + sq, 1) = \frac{\log p}{\phi(p^{j+k})} + O_{r,s}\left(\frac{\log q}{q}\right),$$

where $r = p^j$ and the denominator of s is p^k.

(b) Suppose both $1/r$ and the denominator of s are powers of the same prime p, with $1/r < s$. Then

$$H(q; r + sq, 1) = \frac{\log p}{\phi(p^k)} + O_{r,s}\left(\frac{\log q}{q}\right),$$

where the denominator of s *is* p^k.

(c) *If neither of the above sets of conditions holds, then* $H(q; r + sq, 1) \ll_{r,s} (\log q)/q$.

PROOF. We sum the conclusion of Lemma 2.4.6 over all prime divisors p of q (and, according to Definition 2.1.5, over both residue classes $r + sq$ and $(r + sq)^{-1}$ for each prime divisor). For each such p there is a contribution of $O_{r,s}\big((\log p)/q\big)$ from error terms, and the sum of all these terms is $\ll_{r,s} \frac{1}{q}\sum_{p|q} \log p \le (\log q)/q$. The only remaining task is to consider the possible main terms.

If $r = p^j$ and the denominator p^k of s are powers of the same prime p, then this prime p must divide any q for which $r + sq$ is an integer; hence by Lemma 2.4.6, we have $p^k \parallel q$ and the term $h(q; p, (r + sq)^{-1})$ contributes $(\log p)/r\phi(p^\nu) = (\log p)/\phi(p^{j+k})$ to $H(q; r + sq, 1)$. Similarly, if $r = 1/p^j$ and the denominator p^k of s are powers of the same prime p with $j < k$, then this prime p must divide any q for which $r + sq$ is an integer (this would be false if $j = k$); hence by Lemma 2.4.6, we have $p^{k-j} \parallel q$ and so the term $h(q; p, r + sq)$ contributes $r(\log p)/\phi(p^\nu) = (\log p)/\phi(p^k)$ to $H(q; r + sq, 1)$. For other pairs (r, s), no main term appears, and so the corollary is established. □

PROOF OF THEOREM 2.4.1. From the definition (2.1.4) of $\Delta(q; a, b)$, we have

$$\Delta(q; r+sq, 1) = \iota_q(-(r+sq)) \log 2 + K_q(r+sq-1) + \frac{\Lambda(r_1)}{r_1} + \frac{\Lambda(r_2)}{r_2} + H(q; r+sq, 1),$$

where r_1 and r_2 are the least positive integers congruent to $r+sq$ and $(r+sq)^{-1}$, respectively, modulo q. The results in this section allow us to analyze each term individually :

- If $r = -1$ and s is an integer, then $\iota_q(-(r + sq)) \log 2 = \log 2$. Otherwise, $\iota_q(-(r + sq)) \log 2 = 0$ for all but finitely many (depending on r and s) integers q by Lemma 2.4.1, whence in particular $\iota_q(-(r + sq)) \log 2 \ll_{r,s} (\log q)/q$.

- If $r = 1$, then $(r + sq - 1, q) = (sq, q) = q/d$ where d is the denominator of s, and so $K_q(r + sq - 1) = \Lambda(d)/\phi(d)$ by Definition 2.1.3. Otherwise, $K_q(r + sq - 1) \ll_{r,s} (\log q)/q$ by Lemma 2.4.2 ; this bound also holds if the denominator d of s is not a prime power, since then $\Lambda(d)/\phi(d) = 0$.
- If r is a positive integer and s is an integer, then $r_1 = r$ for all but finitely many q, in which case $\Lambda(r_1)/r_1 = \Lambda(r)/r$. Otherwise $\Lambda(r_1)/r_1 \ll_{r,s} (\log q)/q$ by Corollary 2.4.3 ; this bound also holds if r is not a prime power, since then $\Lambda(r)/r = 0$.

  Similarly, if $r = 1/b$ is the reciprocal of a positive integer and $\frac{s}{r} = bs$ is an integer, then $b(r + sq) = 1 + (bs)q \equiv 1 \pmod{q}$; moreover, b will be the smallest positive integer (for all but finitely many q) such that $b(r+sq) \equiv 1 \pmod{q}$, and so $\Lambda(r_2)/r_2 = \Lambda(b)/b$. Otherwise $\Lambda(r_2)/r_2 \ll_{r,s} (\log q)/q$ by Corollary 2.4.3 ; this bound also holds if the reciprocal b of r is not a prime power, since then $\Lambda(b)/b = 0$. Note also that if b is a prime power, then the denominator of s must be the same prime power, since bs and $r + sq$ are both integers.
- Corollary 2.4.4 tells us exactly when we have a contribution from $H(q; r + sq, 1)$ other than the error term $O_{r,s}\big((\log q)/q\big)$: the denominator of s must be a prime power, and r must be either a power of the same prime or else the reciprocal of a smaller power of the same prime.

In summary, there are six situations in which there is a contribution to $\Delta(q; r + sq, 1)$ beyond the error term $O_{r,s}\big((\log q)/q\big)$: four situations when the denominator of s is a prime power and two situations when s is an integer. All six situations are disjoint, and the contribution to $\Delta(q; r+sq, 1)$ in each situation is exactly $R(r, s)$ as defined in Definition 2.4.1. This establishes the theorem. □

2.4.3. The Bays–Hudson "mirror image phenomenon"

In 1983, Bays and Hudson [5] published their observations of some curious phenomena in the prime number race among the reduced residue classes modulo 11. They graphed normalized error terms corresponding to $\pi(x; 11, 1)$, …, $\pi(x; 11, 10)$, much like the functions $E(x; 11, a)$ discussed in this paper, and from the graph they saw that the terms corresponding to the nonsquare residue classes tended to be positive, while the terms corresponding to the square residue classes tended to be negative, as Chebyshev's bias predicts. Unexpectedly, however, they noticed [**5**, Figure 1]) that the graph corresponding to $\pi(x; 11, 1)$ had a tendency to look like a mirror image of the graph corresponding to $\pi(x; 11, 10)$, and similarly for the other pairs $\pi(x; 11, a)$ and $\pi(x; 11, 11 - a)$. They deemed this observation the "additive inverse phenomenon"; we use the physically suggestive name "mirror image phenomenon".

This prompted them to graph the various normalized error terms corresponding to the sums $\pi(x; 11, a) + \pi(x; 11, b)$ where a is a nonsquare (mod 11) and b is a square (mod 11); all such normalized sums have the same mean value. They witnessed a noticeable difference between the cases $a + b = 11$, when the graph corresponding to the sum was typically quite close to the average value (as in [**5**, Figure 2]), and all other cases which tended to result in more spread-out graphs.

The ideas of the current paper can be used to explain this phenomenon. We consider more generally the limiting logarithmic distributions of the sums of error terms $E(x; q, a) + E(x; q, b)$, where a is a nonsquare (mod q) and b is a square (mod q). The methods of Section 2.2.1 are easily modified to show (under the usual assumptions of GRH and LI) that this distribution has variance

$$V^+(q; a, b) = \sum_{\substack{\chi \pmod q \\ \chi \neq \chi_0}} |\chi(a) + \chi(b)|^2 b(\chi). \tag{2.4.4}$$

Following the method of proof of Theorem 2.1.2, one can show that for any modulus q and any pair a, b of reduced residues modulo q, we have

$$\begin{aligned} V^+(q; a, b) = 2\phi(q)\biggl(\log q - \sum_{p \mid q} \frac{\log p}{p-1} - \frac{\Lambda(q)}{\phi(q)} - (\gamma_0 + \log 2\pi) \\ - K_q(a-b) - \iota_q(-ab^{-1}) \log 2 \biggr) + 2M^+(q; a, b) - 4b(\chi_0), \end{aligned} \tag{2.4.5}$$

where

$$M^+(q; a, b) = \sum_{\substack{\chi \pmod q \\ \chi \neq \chi_0}} |\chi(a) + \chi(b)|^2 \frac{L'(1, \chi^*)}{L(1, \chi^*)}.$$

In particular, we note the term $-\iota_q(-ab^{-1})\log 2$; many of the other terms vanish or simplify in the special case that q is prime. We also note that the primary contribution to $M^+(q;a,b)$ is the expression $-\Lambda(r_1)/r_1 - \Lambda(r_2)/r_2$, where r_1 and r_2 are the least positive residues of ab^{-1} and ba^{-1} (mod q). Both of these expressions are familiar to us from our analysis of $V(q;a,b)$, although their signs are negative in the current setting rather than positive as before.

We see that the variance $V^+(q;a,b)$ of this distribution $E(x;q,a)+E(x;q,b)$ is somewhat smaller than the typical size if there is a small prime congruent to ab^{-1} or ba^{-1} (mod q); more importantly, it is smallest of all if $-ab^{-1} \equiv 1$ (mod q), which is precisely the situation $a+b=q$. In other words, we see very explicitly that the cases where $a+b=q$ yield distributions with smaller-than-normal variance, as observed for $q=11$ by Bays and Hudson. In particular, our theory predicts that for any prime $q \equiv 3$ (mod 4) (so that exactly one of a and $-a$ is a square), the graphs of $E(x;q,a)$ and $E(x;q,q-a)$ will tend to resemble mirror images of each other, more so than the graphs of two functions $E(x;q,a)$ and $E(x;q,b)$ where a and b are unrelated. On the other hand, the contribution of the ι_q term is in a secondary main term, and so the theory predicts that this mirror-image tendency becomes weaker as q grows larger.

We can use the numerical data in the case $q=11$, computed first by Bays and Hudson, to confirm our theoretical evaluation of these variances. We computed the values of each of the twenty-five functions $E(x;11,a)+E(x;11,b)$, where a is a square and b a nonsquare (mod q), on 400 logarithmically equally spaced points spanning the interval $[10^3, 10^7]$. We then computed the variance of our sample points for each function, in order to compare them with the theoretical variance given in equation (2.4.5), which we computed numerically. It is evident from equation (2.4.4) that multiplying both a and b by the same factor does not change $V^+(q;a,b)$, and therefore there are only three distinct values for these theoretical variances : the functions $E(x;q,a)+E(x;q,b)$ where $a+b=11$ all give the same variance, as do the functions where $ab^{-1}\equiv 2$ or $ab^{-1} \equiv 2^{-1} \equiv 6$ (mod 11), and the functions where $ab^{-1}\equiv 7$ or $ab^{-1}\equiv 7^{-1}\equiv$

TAB. 2.6. Observed and theoretical variances for $E(x;11,a)+E(x;11,b)$

Set of functions $E(x;11,a)+E(x;11,b)$	Average variance calculated from sampled data	Theoretical variance
$a+b=11$	5.60	5.31
$\{ab^{-1}, ba^{-1}\} \equiv \{2,6\} \pmod{11}$	7.10	6.82
$\{ab^{-1}, ba^{-1}\} \equiv \{7,8\} \pmod{11}$	9.59	9.06

8 (mod 11). Table 2.6 summarizes our calculations, where the middle column reports the mean of the variances calculated for the functions in each set.

Looking directly at the definition (2.4.4) of $V^+(q;a,b)$, we see that when $a \equiv -b \pmod{q}$, the only characters that contribute to the sum are the even characters, since we have $\chi(a)+\chi(b)=\chi(a)+\chi(-1)\chi(a)=0$ when $\chi(-1)=-1$. As seen earlier in Lemma 2.3.2, the quantity $b(\chi)$ is smaller for even characters than for odd characters, which is another way to express the explanation of the Bays–Hudson observations.

2.5. EXPLICIT BOUNDS AND COMPUTATIONS

We concern ourselves with explicit numerical bounds and computations of the densities $\delta(q;a,b)$ in this final section. We begin in Section 2.5.1 by establishing auxiliary bounds for $\Gamma(z)$, for $\frac{L'}{L}(s,\chi)$, and for the number of zeros of $L(s,\chi)$ near a given height. In Section 2.5.2 we use these explicit inequalities to provide the proofs of two propositions stated in Section 2.3.3 ; we also establish computationally accessible upper and lower bounds for the variance $V(q;a,b)$. Explicit estimates for the density $\delta(q;a,b)$ are proved in Section 2.5.3, including two theorems that give explicit numerical upper bounds for $\delta(q;a,b)$ for q above 1000. Finally, in Section 2.5.4 we describe the two methods we used to calculate numerical values for $\delta(q;a,b)$; we include some sample data from these calculations, including the 120 largest density values that ever occur.

2.5.1. Bounds for classical functions

The main goals of this section are to bound the number of zeros of $L(s,\chi)$ near a particular height and to estimate the size of $\frac{L'}{L}(s,\chi)$ inside the critical strip, both with explicit constants. To achieve this, we first establish some explicit inequalities for the Euler Gamma-function.

Proposition 2.5.1. *If* $\operatorname{Re} z \geq \frac{1}{8}$, *then*

$$\left| \log \Gamma(z) - \left(z - \frac{1}{2}\right) \log z + z - \frac{1}{2} \log 2\pi \right| \leq \frac{1}{4|z|}$$

and

$$\left| \frac{\Gamma'(z)}{\Gamma(z)} - \log(z+1) + \frac{1}{2z+2} + \frac{1}{z} \right| < 0.2.$$

PROOF. The first inequality follows from [**58**, equations (1) and (9) of Section 1.3], both taken with $n = 1$. As for the second inequality, we begin with the identity [**74**, equation (21)], taken with $a = 1$:

$$\Psi(z+1) = \log(z+1) - \frac{1}{2(z+1)} + f_1'(z).$$

Here $\Psi(z) = \frac{\Gamma'}{\Gamma}(z)$ has its usual meaning ; we use the identity $\frac{\Gamma'(z)}{\Gamma(z)} + \frac{1}{z} = \frac{\Gamma'(z+1)}{\Gamma(z+1)}$ to obtain

$$\frac{\Gamma'(z)}{\Gamma(z)} + \frac{1}{z} - \log(z+1) + \frac{1}{2(z+1)} = f_1'(z),$$

and therefore it suffices to show that $|f_1'(z)| \leq 0.2$ when $\operatorname{Re}(z) \geq \frac{1}{8}$. The notation $f_1(z) = \log F_{1,1/2}(z)$ is defined in [**74**, equation (9)], and therefore $f_1'(z) = F_{1,1/2}'(z)/F_{1,1/2}(z)$. By [**74**, Lemma 1.1.1], the denominator $F_{1,1/2}(z)$ is bounded below in modulus by $\sqrt{e/\pi}$; by [**74**, Lemma 2.2.1] taken with $a = n = 1$, the numerator is bounded above in modulus by

$$\left|F_{1,1/2}'(z)\right| < \log \frac{x+1}{x+1/2} - \frac{1}{2x+2},$$

where $x = \operatorname{Re} z$ (unfortunately [**74**, equation (27)] contains the misprint $f_{a,1/2}^{(n)}$ where $F_{a,1/2}^{(n)}$ is intended). The right-hand side of this inequality is a decreasing function of x, and its value at $x = \frac{1}{8}$ is $\log \frac{9}{5} - \frac{4}{9}$. We conclude that for $\operatorname{Re} z \geq \frac{1}{8}$, we have $|f_1'(z)| \leq \left(\log \frac{9}{5} - \frac{4}{9}\right)/\sqrt{e/\pi} < 0.2$, as needed. □

Lemma 2.5.1. *Let* $a = 0$ *or* $a = 1$*. For any real numbers* $\frac{1}{4} \le \sigma \le 1$ *and* T*, we have*

$$\left|\frac{\Gamma'(\frac{1}{2}(\sigma + iT + a))}{\Gamma(\frac{1}{2}(\sigma + iT + a))} - \frac{\Gamma'(\frac{1}{2}(2 + iT + a))}{\Gamma(\frac{1}{2}(2 + iT + a))}\right| < 7.812. \tag{2.5.1}$$

Proof. By symmetry we may assume that $T \ge 0$. We first dispose of the case $T \le 3$. When $a = 0$, a computer calculation shows that the maximum value of the left-hand side of equation (2.5.1) in the rectangle $\{\sigma + iT\colon \frac{1}{4} \le \sigma \le 1,\ 0 \le T \le 3\}$ occurs at $\sigma = \frac{1}{4}$ and $T = 0$: the value of the left-hand side at that point is a bit less than 7.812. When $a = 1$, a similar calculation shows that the left-hand side of equation (2.5.1) is always strictly less than 7.812.

For the rest of the proof, we may therefore assume that $T \ge 3$. By Proposition 2.5.1,

$$\begin{aligned}\frac{\Gamma'(\frac{1}{2}(\sigma + iT + a))}{\Gamma(\frac{1}{2}(\sigma + iT + a))} - \frac{\Gamma'(\frac{1}{2}(2 + iT + a))}{\Gamma(\frac{1}{2}(2 + iT + a))} &= \log\frac{\sigma + iT + a + 2}{2} - \frac{1}{\sigma + iT + a + 2} \\ &- \frac{2}{\sigma + iT + a} - \log\frac{4 + iT + a}{2} + \frac{1}{4 + iT + a} + \frac{2}{2 + iT + a} + \overline{O}(0.4),\end{aligned}$$

and therefore

$$\begin{aligned}\left|\frac{\Gamma'(\frac{1}{2}(\sigma + iT + a))}{\Gamma(\frac{1}{2}(\sigma + iT + a))} - \frac{\Gamma'(\frac{1}{2}(2 + iT + a))}{\Gamma(\frac{1}{2}(2 + iT + a))}\right| &\le \left|\log\left(1 - \frac{2 - \sigma}{4 + iT + a}\right)\right| \\ &+ \left|\frac{2 - \sigma}{(\sigma + iT + a + 2)(4 + iT + a)}\right| + 2\left|\frac{2 - \sigma}{(\sigma + iT + a)(2 + iT + a)}\right| + 0.4.\end{aligned}$$

Under the assumptions on σ, a, and T, we always have the inequality $\big|2 - \sigma/(4 + iT + a)\big| \le \frac{1}{2}$. The maximum modulus principle implies the inequality $\left|\frac{1}{z}\log(1 - z)\right| \le \log 4$ for $|z| \le \frac{1}{2}$, and so

$$\begin{aligned}\left|\frac{\Gamma'(\frac{1}{2}(\sigma + iT + a))}{\Gamma(\frac{1}{2}(\sigma + iT + a))} - \frac{\Gamma'(\frac{1}{2}(2 + iT + a))}{\Gamma(\frac{1}{2}(2 + iT + a))}\right| &\le \left|\frac{2 - \sigma}{4 + iT + a}\right| \log 4 \\ &+ \left|\frac{2 - \sigma}{(\sigma + iT + a + 3)(5 + iT + a)}\right| + 2\left|\frac{2 - \sigma}{(\sigma + iT + a)(2 + iT + a)}\right| + 0.4\end{aligned}$$

Finally we use the inequalities on σ, a, and T to conclude that

$$\left|\frac{\Gamma'(\frac{1}{2}(\sigma + iT + a))}{\Gamma(\frac{1}{2}(\sigma + iT + a))} - \frac{\Gamma'(\frac{1}{2}(2 + iT + a))}{\Gamma(\frac{1}{2}(2 + iT + a))}\right| \le \frac{2}{5}\log 4 + \frac{2}{5\sqrt{13}} + \frac{4}{3\sqrt{13}} + 0.4 < 1.4353,$$

which amply suffices to finish the proof. □

We turn now to estimates for quantities associated with Dirichlet L-functions. The next few results do not require GRH to be true, and in fact their proofs cite identities from the literature that hold more generally no matter where the zeros of $L(s,\chi)$ might lie. Accordingly, we use the usual notation $\rho = \beta + i\gamma$ to denote a nontrivial zero of $L(s,\chi)$, and all sums in this section of the form $\sum_\rho$ denote sums over all such nontrivial zeros of the Dirichlet L-function.

Lemma 2.5.2. *Let* $q \ge 2$*, and let* χ *be a nonprincipal character* (mod q). *For any real number* T*,*

$$\sum_\rho \frac{1}{|2+iT-\rho|^2} < \frac{1}{2}\log\big(0.609q(|T|+5)\big).$$

PROOF. It suffices to prove the lemma for primitive characters. For χ primitive, it is known [**60**, equation (10.37)] that as meromorphic functions on the complex plane,

$$\frac{L'(s,\chi)}{L(s,\chi)} = -\frac{1}{2}\log\frac{q}{\pi} - \frac{1}{2}\frac{\Gamma'(\frac{1}{2}(s+a))}{\Gamma(\frac{1}{2}(s+a))} + B(\chi) + \sum_\rho\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right), \tag{2.5.2}$$

where the constant $B(\chi)$ was described earlier in the proof of Lemma 2.3.2, and where $a = 0$ if $\chi(-1) = 1$ and $a = 1$ if $\chi(-1) = -1$. Taking real parts of both sides and using the identity (2.3.3), we obtain after rearrangement

$$\operatorname{Re}\sum_\rho \frac{1}{s-\rho} = \operatorname{Re}\frac{L'(s,\chi)}{L(s,\chi)} + \frac{1}{2}\log\frac{q}{\pi} + \frac{1}{2}\operatorname{Re}\frac{\Gamma'(\frac{1}{2}(s+a))}{\Gamma(\frac{1}{2}(s+a))}. \tag{2.5.3}$$

If we put $z = \frac{1}{2}(s+a)$ in Proposition 2.5.1, we see that for $\operatorname{Re} s \ge \frac{1}{8}$,

$$\begin{aligned}\operatorname{Re}\frac{\Gamma'(\frac{1}{2}(s+a))}{\Gamma(\frac{1}{2}(s+a))} &= \operatorname{Re}\log\frac{s+a+2}{2} - \operatorname{Re}\frac{1}{s+a+2} - \operatorname{Re}\frac{2}{s+a} + 0.2 \\ &\le \log|s+a+1| - \log 2 + 0 + 0.2 \le \log|s+3| - 0.493.\end{aligned}$$

Inserting this bound into equation (2.5.3) and putting $s = 2+iT$,

$$\operatorname{Re}\sum_\rho \frac{1}{2+iT-\rho} \le \operatorname{Re}\frac{L'(2+iT,\chi)}{L(2+iT,\chi)} + \frac{1}{2}\log\frac{q}{\pi} + \frac{1}{2}\log|5+iT| - 0.246.$$

Now notice that

$$\left|\frac{L'(2+iT,\chi)}{L(2+iT,\chi)}\right| = \left|-\sum_{n=1}^\infty \frac{\chi(n)\Lambda(n)}{n^{2+iT}}\right| \le \sum_{n=1}^\infty \frac{\Lambda(n)}{n^2} = -\frac{\zeta'(2)}{\zeta(2)} < 0.57, \tag{2.5.4}$$

and therefore

$$\begin{aligned}\operatorname{Re}\sum_{\rho}\frac{1}{2+iT-\rho} &\le 0.57+\frac{1}{2}\log\frac{q}{\pi}+\frac{1}{2}\log|5+iT|-0.246\\ &\le \frac{1}{2}\log q+\frac{1}{2}\log(|T|+5)+0.57-\frac{1}{2}\log\pi-0.246\\ &\le \frac{1}{2}\log\big(q(|T|+5)\big)-0.248\le\frac{1}{2}\log\big(0.609q(|T|+5)\big).\end{aligned}$$

We obtain finally

$$\begin{aligned}\sum_{\rho}\frac{1}{|2+iT-\rho|^2} &< \sum_{\rho}\frac{2-\beta}{|2+iT-\rho|^2}\\ &= \operatorname{Re}\sum_{\rho}\frac{1}{2+iT-\rho}\le\frac{1}{2}\log\big(0.609q(|T|+5)\big) \qquad (2.5.5)\end{aligned}$$

as claimed. □

Proposition 2.5.2. *For any nonprincipal character* χ *and any real number* T*, we have*

$$\#\{\rho\colon |T-\operatorname{Im}\rho|\le 2\}\le 4\log\big(0.609q(|T|+5)\big).$$

PROOF. This follows immediately from equation (2.5.5) and the inequalities

$$\sum_{\substack{\rho\\ |T-\gamma|\le 2}}1\le 8\sum_{\rho}\frac{1}{(2-\sigma)^2+(T-\gamma)^2}\le 8\sum_{\rho}\frac{2-\beta}{|2+iT-\rho|^2}.$$

□

Lemma 2.5.3. *Let* $s=\sigma+iT$ *with* $\frac{1}{4}\le\sigma\le 1$*. For any primitive character* χ (mod q) *with* $q\ge 2$*, if* $L(s,\chi)\ne 0$ *then*

$$\left|\frac{L'(s,\chi)}{L(s,\chi)}-\sum_{\substack{\rho\\ |T-\gamma|\le 2}}\frac{1}{s-\rho}\right|\le\sqrt{2}\log\big(0.609q(|T|+5)\big)+4.48.$$

PROOF. Applying equation (2.5.2) at $s=\sigma+iT$ and again at $2+iT$, we obtain

$$\begin{aligned}\frac{L'(s,\chi)}{L(s,\chi)}-\frac{L'(2+iT,\chi)}{L(2+iT,\chi)} &= \frac{1}{2}\frac{\Gamma'(\frac{1}{2}(2+iT+a))}{\Gamma(\frac{1}{2}(2+iT+a))}-\frac{1}{2}\frac{\Gamma'(\frac{1}{2}(s+a))}{\Gamma(\frac{1}{2}(s+a))}\\ &\qquad +\sum_{\rho}\left(\frac{1}{s-\rho}-\frac{1}{2+iT-\rho}\right),\end{aligned}$$

which implies

$$\left|\frac{L'(s,\chi)}{L(s,\chi)} - \sum_{\substack{\rho \\ |T-\gamma|\le 2}} \frac{1}{s-\rho}\right| \le \left|\frac{L'(2+iT,\chi)}{L(2+iT,\chi)}\right| + \frac{1}{2}\left|\frac{\Gamma'(\frac{1}{2}(2+iT+a))}{\Gamma(\frac{1}{2}(2+iT+a))} - \frac{\Gamma'(\frac{1}{2}(s+a))}{\Gamma(\frac{1}{2}(s+a))}\right|$$
$$+ \sum_{\substack{\rho \\ |T-\gamma|>2}} \left|\frac{1}{s-\rho} - \frac{1}{2+iT-\rho}\right| + \sum_{\substack{\rho \\ |T-\gamma|\le 2}} \frac{1}{|2+iT-\rho|}.$$

Using equation (2.5.4) and Lemma 2.5.1 to bound the first two terms on the right-hand side, we see that

$$\left|\frac{L'(s,\chi)}{L(s,\chi)} - \sum_{\substack{\rho \\ |T-\gamma|\le 2}} \frac{1}{s-\rho}\right| < 0.57 + 3.906$$
$$+ \sum_{\substack{\rho \\ |T-\gamma|>2}} \frac{2-\sigma}{|s-\rho||2+iT-\rho|} + \sum_{\substack{\rho \\ |T-\gamma|\le 2}} \frac{1}{|2+iT-\rho|}. \quad (2.5.6)$$

To prepare the last two sums for an application of Lemma 2.5.2, we note that when $|T-\gamma| > 2$,

$$\frac{2-\sigma}{|s-\rho||2+iT-\rho|} < 2\frac{|2+iT-\rho|}{|s-\rho|}\frac{1}{|2+iT-\rho|^2} < 2\sqrt{2}\frac{1}{|2+iT-\rho|^2};$$

on the other hand, when $|T-\gamma| \le 2$,

$$\frac{1}{|2+iT-\rho|} = \frac{|2+iT-\rho|}{|2+iT-\rho|^2} < \frac{2\sqrt{2}}{|2+iT-\rho|^2}.$$

Therefore equation (2.5.6) becomes, by Lemma 2.5.2,

$$\left|\frac{L'(s,\chi)}{L(s,\chi)} - \sum_{\substack{\rho \\ |T-\gamma|\le 2}} \frac{1}{s-\rho}\right| < 0.57 + 3.906 + 2\sqrt{2}\sum_{\rho} \frac{1}{|2+iT-\rho|^2}$$
$$< 4.48 + \sqrt{2}\log\big(0.609q(|T|+5)\big)$$

as claimed. □

We restore the assumption of GRH for the last proposition of this section, which is used in the proof of Lemma 2.5.7 below.

Proposition 2.5.3. *Assume GRH. Let* $s = \sigma + iT$ *with* $\frac{1}{4} \le \sigma \le 1$, $\sigma \ne \frac{1}{2}$. *If* χ *is any nonprincipal character* (mod q), *then*

$$\left|\frac{L'(s,\chi)}{L(s,\chi)}\right| \le \left(\frac{4}{|\sigma-\frac{1}{2}|} + \sqrt{2}\right)\log\big(0.609q(|T|+5)\big) + 4.48 + \frac{\log q}{2^\sigma - 1}.$$

Furthermore, if χ *is primitive and* $q \ge 2$*, then the summand* $(\log q)/(2^\sigma - 1)$ *can be omitted from the upper bound.*

Proof. Assume first that χ is primitive. Lemma 2.5.3 tells us that

$$\begin{aligned}\left|\frac{L'(s,\chi)}{L(s,\chi)}\right| &\le \sum_{\substack{\rho \\ |T-\gamma|\le 2}} \frac{1}{|s-\rho|} + \sqrt{2}\log\big(0.609q(|T|+5)\big) + 4.48 \\ &\le \frac{1}{|\sigma-\frac{1}{2}|}\#\{\rho\colon |T-\gamma|\le 2\} + \sqrt{2}\log\big(0.609q(|T|+5)\big) + 4.48\end{aligned}$$

under the assumption of GRH; the proposition for primitive χ now follows immediately from Proposition 2.5.2.

If χ is not primitive, then $L(s,\chi) = L(s,\chi^*)\prod_{p\mid q}\big(1 - \frac{\chi^*(p)}{p^s}\big)$; we then have the identity

$$\frac{L'(s,\chi)}{L(s,\chi)} = \frac{L'(s,\chi^*)}{L(s,\chi^*)} + \sum_{p\mid q}\frac{\chi^*(p)\log p}{p^s - \chi^*(p)}.$$

Therefore

$$\left|\frac{L'(s,\chi)}{L(s,\chi)} - \frac{L'(s,\chi^*)}{L(s,\chi^*)}\right| \le \sum_{p\mid q}\frac{\log p}{p^\sigma - 1} \le \frac{1}{2^\sigma - 1}\sum_{p\mid q}\log p \le \frac{\log q}{2^\sigma - 1},$$

which finishes the proof of the proposition in full. □

2.5.2. Bounds for the variance $V(q;a,b)$

This section has two main purposes. First, we provide the proofs of Propositions 2.3.5 and 2.3.6, two statements involving smoothed sums of the von Mangoldt function which were stated in Section 2.3.3. Second, we establish two sets of upper and lower bounds for the variance $V(q;a,b)$, one when q is prime and one valid for all q. All of these results are stated with explicit constants and are valid for explicit ranges of q.

Lemma 2.5.4. *For any real number* t*, we have* $\left|\frac{d}{dt}\big|\Gamma(-\frac{1}{2}+it)\big|\right| \le \big|\Gamma'(-\frac{1}{2}+it)\big|$.

Proof. We show more generally that if $f(t)$ is any differentiable complex-valued function that never takes the value 0, then $|f(t)|$ is also differentiable and $\left|\frac{d}{dt}|f(t)|\right| \le |f'(t)|$; the lemma then follows since Γ never takes the value 0. Write $f(t) = u(t) + iv(t)$ where u and v are real-valued; then

$$\tfrac{d}{dt}|f(t)| = \tfrac{d}{dt}\sqrt{u(t)^2 + v(t)^2} = \frac{u(t)u'(t) + v(t)v'(t)}{\sqrt{u(t)^2 + v(t)^2}}$$

while $|f'(t)| = |u'(t) + iv'(t)| = \sqrt{u'(t)^2 + v'(t)^2}$. The asserted inequality is therefore equivalent to $|u(t)u'(t)+v(t)v'(t)| \le \sqrt{u(t)^2 + v(t)^2}\sqrt{u'(t)^2 + v'(t)^2}$, which is a consequence of the Cauchy-Schwarz inequality. □

Lemma 2.5.5. *We have* $|\Gamma(s)| \le |\Gamma(\operatorname{Re} s)|$ *for all complex numbers* s.

Note that this assertion is trivially true if $\operatorname{Re} s$ is a nonpositive integer, under the convention $|\Gamma(-n)| = \infty$ for $n \ge 0$.

PROOF. We prove that the assertion holds whenever $\operatorname{Re} s > -n$, by induction on n. The base case $n = 0$ can be derived from the integral representation $\Gamma(s) = \int_0^\infty t^{s-1}e^{-t}dt$, which gives

$$|\Gamma(s)| \le \int_0^\infty |t^{s-1}|e^{-t}dt = \int_0^\infty t^{\operatorname{Re} s-1}e^{-t}dt = \Gamma(\operatorname{Re} s).$$

Now assume that the assertion holds whenever $\operatorname{Re} s > -n$. Given a complex number s for which $\operatorname{Re} s > -(n+1)$, we use the identity $\Gamma(s+1) = s\Gamma(s)$ and the induction hypothesis to write

$$|\Gamma(s)| = \frac{|\Gamma(s+1)|}{|s|} \le \frac{|\Gamma(\operatorname{Re} s+1)|}{|s|} = \frac{|\operatorname{Re} s|}{|s|}|\Gamma(\operatorname{Re} s)| \le |\Gamma(\operatorname{Re} s)|,$$

as desired. □

Lemma 2.5.6. *For any nonprincipal character* χ,

$$\sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \left|\Gamma\left(-\tfrac{1}{2} + i\gamma\right)\right| \le 14.27\log q + 16.25.$$

We remark that this lemma does not assume GRH, since the sum on the left-hand side only decreases if some of the zeros of $L(s,\chi)$ lie off the critical line.

PROOF. First, by Proposition 2.5.2 applied with $T = 0$, the number of zeros of $L(s, \chi)$ with $|\gamma| \le 2$ is at most $4 \log(3.045q)$; thus by Lemma 2.5.5,

$$\sum_{\substack{|\gamma|\le 2 \\ L(1/2+i\gamma,\chi)=0}} \big|\Gamma(-\tfrac12 + i\gamma)\big| \le \big|\Gamma(-\tfrac12)\big| \sum_{\substack{|\gamma|\le 2 \\ L(1/2+i\gamma,\chi)=0}} 1$$

$$\le 8\sqrt{\pi} \log(3.045q) \le 14.18 \log q + 15.79. \qquad (2.5.7)$$

We can write the remainder of the sum using Riemann-Stieltjes integration as

$$\sum_{\substack{|\gamma|>2 \\ L(1/2+i\gamma,\chi)=0}} \big|\Gamma(-\tfrac12 + i\gamma)\big| = \int_2^\infty \big|\Gamma(-\tfrac12 + it)\big| \, d\big(N(t,\chi) - N(2,\chi)\big)$$

$$= -\int_2^\infty \big(N(t,\chi) - N(2,\chi)\big) \frac{d}{dt} \big|\Gamma(-\tfrac12 + it)\big| \, dt;$$

the vanishing of the boundary terms is justified by the upper bound $N(t, \chi) \ll_q t \log t$ (see Proposition 2.2.8 for example) and the exponential decay of $\Gamma(s)$ on vertical lines. We conclude from Lemma 2.5.4 that

$$\sum_{\substack{|\gamma|\le 2 \\ L(1/2+i\gamma,\chi)=0}} \big|\Gamma(-\tfrac12 + i\gamma)\big| \le \int_2^\infty N(t,\chi) \big|\Gamma'(-\tfrac12 + it)\big| \, dt$$

$$\le \int_2^\infty \bigg(\Big(\frac{t}{\pi} + 0.68884 \Big) \log \frac{qt}{2\pi e} + 10.6035 \bigg) \big|\Gamma'(-\tfrac12 + it)\big| \, dt$$

by Proposition 2.2.8. Since $\log(qt/2\pi e) = \log q + \log(t/2\pi e)$, the right-hand side is simply a linear function of $\log q$; using numerical integration we see that

$$\sum_{\substack{|\gamma|\le 2 \\ L(1/2+i\gamma,\chi)=0}} \big|\Gamma(-\tfrac12 + i\gamma)\big| \le 0.09 \log q + 0.46.$$

Combining this upper bound with the bound in equation (2.5.7) establishes the lemma. □

Lemma 2.5.7. *Assume GRH. For any nonprincipal character χ,*

$$\int_{-3/4-i\infty}^{-3/4+i\infty} \left| \frac{L'(s+1,\chi)}{L(s+1,\chi)} \Gamma(s) \right| ds \le 101 \log q + 112.$$

PROOF. Proposition 2.5.3 with $\sigma = \frac{1}{4}$ tells us that for any real number t,

$$\begin{aligned}\left|\frac{L'(\frac{1}{4}+it,\chi)}{L(\frac{1}{4}+it,\chi)}\right| &\le 17.42\log\big(0.609q(|t|+5)\big) + 4.48 + \frac{\log q}{0.1892} \\ &\le 22.71\log q + 17.42\log(|t|+5) - 4.159,\end{aligned}$$

and therefore

$$\begin{aligned}&\int_{-3/4-i\infty}^{-3/4+i\infty} \left|\frac{L'(s+1,\chi)}{L(s+1,\chi)}\Gamma(s)\right| ds \\ &\qquad\le \int_{-\infty}^{\infty} \big(22.71\log q + 17.42\log(|t|+5) - 4.159\big)\big|\Gamma\big(-\tfrac{3}{4}+it\big)\big|\, dt.\end{aligned}$$

Again this integral is a linear function of $\log q$, and a numerical calculation establishes the particular constants used in the statement of the lemma. □

With these lemmas in hand, we are now able to provide the two proofs deferred until now from Section 2.3.3.

PROOF OF PROPOSITION 2.3.5. We begin with the Mellin transform formula, valid for any real number $c > 0$,

$$-\sum_{n=1}^{\infty} \frac{\chi(n)\Lambda(n)}{n} e^{-n/y} = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{L'(s+1,\chi)}{L(s+1,\chi)}\Gamma(s)y^s\, ds$$

(see [**60**, equations (5.24) and (5.25)]). We move the contour to the left, from the vertical line $\operatorname{Re} s = c$ to the vertical line $\operatorname{Re} s = -\frac{3}{4}$, picking up contributions from the pole of Γ at $s = 0$ as well as from each nontrivial zero of $L(s,\chi)$. The result is

$$\begin{aligned}-\sum_{n=1}^{\infty} \frac{\chi(n)\Lambda(n)}{n} e^{-n/y} = \frac{L'(1,\chi)}{L(1,\chi)} &+ \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \Gamma\big(-\tfrac{1}{2}+i\gamma\big)y^{-1/2+i\gamma} \\ &+ \frac{1}{2\pi i}\int_{-3/4-i\infty}^{-3/4+i\infty} \frac{L'(s+1,\chi)}{L(s+1,\chi)}\Gamma(s)y^s\, ds \quad (2.5.8)\end{aligned}$$

since we are assuming GRH. (Strictly speaking, we should consider truncations of these infinite integrals ; however, the exponential decay of $\Gamma(s)$ in vertical strips implies that the contributions at large height do vanish in the limit.)

The sum on the right-hand side can be bounded by

$$\Bigg| \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \Gamma\big(-\tfrac12+i\gamma\big)y^{-1/2+i\gamma}\Bigg| \le y^{-1/2} \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \big|\Gamma\big(-\tfrac12+i\gamma\big)\big| \\ \le \frac{14.27\log q + 16.25}{y^{1/2}}$$

by Lemma 2.5.6, while the integral can be bounded by

$$\left|\frac{1}{2\pi i}\int_{-3/4-i\infty}^{-3/4+i\infty} \frac{L'(s+1,\chi)}{L(s+1,\chi)}\Gamma(s)y^s\,ds\right| \le \frac{1}{2\pi y^{3/4}}\int_{-3/4-i\infty}^{-3/4+i\infty}\left|\frac{L'(s+1,\chi)}{L(s+1,\chi)}\Gamma(s)\right|ds \\ \le \frac{101\log q + 112}{2\pi y^{3/4}}$$

by Lemma 2.5.7. Using these two inequalities in equation (2.5.8) establishes the proposition. □

Proof of Proposition 2.3.6. Since $1 \le a < q$, we may write

$$\sum_{n\equiv a \,(\mathrm{mod}\ q)} \frac{\Lambda(n)}{n}e^{-n/q^2} = \frac{\Lambda(a)}{a}e^{-a/q^2} \\ + \overline{O}\Bigg(\sum_{\substack{q\le n\le q^2 \\ n\equiv a\,(\mathrm{mod}\ q)}}\frac{\Lambda(n)}{n} + \sum_{\substack{n>q^2 \\ n\equiv a\,(\mathrm{mod}\ q)}}\frac{\Lambda(n)}{n}e^{-n/q^2}\Bigg). \quad (2.5.9)$$

Since $\Lambda(n)/n \le (\log n)/n$, which is a decreasing function of n for $n \ge 3$, we have

$$\sum_{\substack{n>q^2 \\ n\equiv a\,(\mathrm{mod}\ q)}}\frac{\Lambda(n)}{n}e^{-n/q^2} \le \frac{\log q^2}{q^2}\sum_{j=q}^{\infty} e^{-(qj+a)/q^2} \le \frac{2\log q}{q^2}e^{-1}\sum_{k=0}^{\infty} e^{-j/q} \\ = \frac{2\log q}{q^2}e^{-1}\frac{1}{1-e^{-1/q}};$$

note here that $1 \le a < q$ so $q \ge 2$. As the function $t/(1-e^{-t})$ is bounded by $1/2(1-e^{-1/2})$ for $0 < t \le \frac12$, we conclude that

$$\sum_{\substack{n>q^2 \\ n\equiv a\,(\mathrm{mod}\ q)}}\frac{\Lambda(n)}{n}e^{-n/q^2} \le \frac{2\log q}{q}e^{-1}\frac{1}{2(1-e^{-1/2})} < 0.935\frac{\log q}{q}.$$

We bound the second term of equation (2.5.9) crudely :

$$\sum_{\substack{q\le n\le q^2\\ n\equiv a\ (\mathrm{mod}\ q)}} \frac{\Lambda(n)}{n} \le (\log q^2)\sum_{j=1}^{q-1}\frac{1}{qj+a} \le \frac{2\log q}{q}\sum_{j=1}^{q-1}\frac{1}{j} \le \frac{2\log q}{q}(\log q+1).$$

Finally, for the first term of equation (2.5.9), the estimate $e^{-t} = 1 + \overline{O}(t)$ for $t \ge 0$ allows us to write

$$\frac{\Lambda(a)}{a}e^{-a/q} = \frac{\Lambda(a)}{a}\left(1+\overline{O}\left(\frac{a}{q}\right)\right) = \frac{\Lambda(a)}{a} + \overline{O}\left(\frac{\log q}{q}\right).$$

Using these three deductions transforms equation (2.5.9) into the statement of the proposition. □

We now turn to the matter of giving explicit upper and lower bounds for $V(q;a,b)$. In the case where q is prime, we are already able to establish such estimates.

Proposition 2.5.4. *If* $q \ge 150$ *is prime, then*

$$2(q-1)(\log q - 2.42) - 47.238\log^2 q \le V(q;a,b) \le 2(q-1)(\log q - 0.99) \\ + 47.238\log^2 q.$$

PROOF. Combining Theorem 2.1.2 with Proposition 2.3.7, we see that

$$\begin{aligned} V(q;a,b) &= 2\phi(q)\big(\mathcal{L}(q) + K_q(a-b) + \iota_q(-ab^{-1})\log 2\big) + 2M^*(q;a,b) \\ &= 2\phi(q)\bigg(\mathcal{L}(q) + K_q(a-b) + \iota_q(-ab^{-1})\log 2 + \frac{\Lambda(r_1)}{r_1} + \frac{\Lambda(r_2)}{r_2} \\ &\qquad + H_0(q;a,b)\bigg) + \overline{O}\left(\frac{47.238\phi(q)\log^2 q}{q}\right) \end{aligned} \tag{2.5.10}$$

for any $q \ge 150$, where r_1 and r_2 denote the least positive residues of ab^{-1} and ba^{-1} (mod q). Since we are assuming q is prime, both $K_q(a-b)$ and $H_0(q;a,b)$ vanish, and we have

$$V(q;a,b) = 2(q-1)\left(\log\frac{q}{2\pi e^{\gamma_0}} + \iota_q(-ab^{-1})\log 2 + \frac{\Lambda(r_1)}{r_1} + \frac{\Lambda(r_2)}{r_2}\right) \\ + \overline{O}(47.238\log^2 q).$$

The function $\Lambda(n)/n$ is nonnegative and bounded above by $(\log 3)/3$, and the function ι_q takes only the values 0 and 1; therefore the quantity in large parentheses satisfies the bounds

$$\log q - 2.42 \le \log\frac{q}{2\pi e^{\gamma_0}} + \iota_q(-ab^{-1})\log 2 + \frac{\Lambda(r_1)}{r_1} + \frac{\Lambda(r_2)}{r_2} \le \log q - 0.99,$$

which establishes the proposition. □

We require two additional lemmas before we can treat the case of general (possibly composite) q.

Lemma 2.5.8. *With* H_0 *defined in Definition 2.3.1, we have*

$$-(4\log q)/q \le H_0(q;a,b) \le 4.56$$

for any reduced residues a *and* b (mod q).

Proof. Since $e(q;p,r) \ge 1$ always, we have

$$h_0(q;p,r) = \frac{1}{\phi(p^\nu)}\frac{\log p}{p^{e(q;p,r)}(1-p^{-e(q;p,1)})} \le \frac{1}{p-1}\frac{\log p}{p-1} \le 4\frac{\log p}{p^2}.$$

Therefore

$$\begin{aligned} H_0(q;a,b) &\le \sum_{p\mid q}\big(h_0(q;p,ab^{-1}) + h_0(q;p,ba^{-1})\big) \le 8\sum_{p\mid q}\frac{\log p}{p^2} \\ &< 8\sum_{n=1}^{\infty}\frac{\Lambda(n)}{n^2} = 8\left|\frac{\zeta'(2)}{\zeta(2)}\right| \le 4.56, \end{aligned}$$

which establishes the upper bound. On the other hand, note that $p^{e(q;p,1)}$ is an integer larger than 1 that is congruent to 1 (mod q/p^ν). Therefore $p^{e(q;p,1)} \ge q/p^\nu + 1$, and so

$$\begin{aligned} H_0(q;a,b) &\ge -2\sum_{p\mid q} h_0(q;p,1) = -2\sum_{p\mid q}\frac{1}{\phi(p^\nu)}\frac{\log p}{p^{e(q;p,1)}-1} \\ &\ge -2\sum_{p\mid q}\frac{1}{p^\nu(1-1/p)}\frac{\log p}{q/p^\nu} \ge -\frac{4}{q}\sum_{p\mid q}\log p \ge -\frac{4\log q}{q}, \end{aligned}$$

which establishes the lower bound. □

Lemma 2.5.9. *If* $q \geq 2$ *is any integer, then*

$$\sum_{p|q} \frac{\log p}{p-1} \leq 1.02 \log\log q + 3.04.$$

PROOF. We separate the sum into two intervals at the point $1 + \log q$. The contribution from the larger primes is at worst

$$\sum_{\substack{p|q \\ p \geq 1+\log q}} \frac{\log p}{p-1} \leq \frac{1}{\log q} \sum_{p|q} \log p \leq \frac{\log q}{\log q} = 1.$$

For the smaller primes, recall the usual notation $\theta(t) = \sum_{p \leq t} \log p$. We will use the explicit bound $\theta(t) \leq 1.01624t$ for $t > 0$ from Theorem 9 of **[67]** , and so the contribution from the smaller primes is bounded by

$$\begin{aligned} \sum_{\substack{p|q \\ p<1+\log q}} \frac{\log p}{p-1} \leq \sum_{p<1+\log q} \frac{\log p}{p-1} &= \int_{2^-}^{1+\log q} \frac{d\theta(t)}{t-1} \\ &= \frac{\theta(1+\log q)}{\log q} + \int_{2}^{1+\log q} \frac{\theta(t)}{(t-1)^2}\, dt \\ &\leq 1.01624 \left(\frac{1+\log q}{\log q} + \int_{2}^{1+\log q} \frac{t\, dt}{(t-1)^2} \right) \\ &= 1.01624 \left(1 + \frac{1}{\log q} + \log\log q - \frac{1}{\log q} + 1 \right) \\ &= 1.01624 \log\log q + 2.03248, \end{aligned}$$

which finishes the proof of the lemma. □

Proposition 2.5.5. *If* $q \geq 500$, *then*

$$2\phi(q)(\log q - 1.02 \log\log q - 7.34) \leq V(q; a, b) \leq 2\phi(q)(\log q + 6.1).$$

PROOF. We begin with equation (2.5.10), expanding the functions $\mathcal{L}$ and K_q according to Definition 2.1.3 :

$$V(q;a,b) = 2\phi(q)\bigg(\log\frac{q}{2\pi e^{\gamma_0}} - \sum_{p\mid q} \frac{\log p}{p-1} + \frac{\Lambda(q/(q,a-b))}{\phi(q/(q,a-b))} \\ + \iota_q(-ab^{-1})\log 2 + \frac{\Lambda(r_1)}{r_1} + \frac{\Lambda(r_2)}{r_2} + H_0(q;a,b) + \overline{O}\bigg(\frac{23.62\log^2 q}{q}\bigg)\bigg). \tag{2.5.11}$$

The last term on the first line is nonnegative and bounded above by $\log 2$, while the first three terms on the second line are nonnegative and bounded together by $\log 2 + \frac{2}{3}\log 3$ as in the proof of Proposition 2.5.4. The term $H_0(q;a,b)$ is bounded above by 4.56 and below by $(-4\log q)/q$ by Lemma 2.5.8. Therefore

$$2\phi(q)\bigg(\log q - \log 2\pi e^{\gamma_0} - \sum_{p\mid q}\frac{\log p}{p-1} - \frac{4\log q}{q} + \overline{O}\bigg(\frac{23.62\log^2 q}{q}\bigg)\bigg) \le V(q;a,b) \\ \le 2\phi(q)\bigg(\log q - \log 2\pi e^{\gamma_0} + \log 2 + \log 2 + \frac{2}{3}\log 3 + 4.56 + \overline{O}\bigg(\frac{23.62\log^2 q}{q}\bigg)\bigg). \tag{2.5.12}$$

The sum being subtracted on the top line is bounded above by $1.02\log\log q + 3.04$ by Lemma 2.5.9. Lastly, a calculation shows that the $\overline{O}$ error term is at most 1.83 for $q \ge 500$, and therefore

$$2\phi(q)\bigg(\log q - \log 2\pi e^{\gamma_0} - (1.02\log\log q + 3.04) - \frac{4\log q}{q} - 1.83\bigg) \le V(q;a,b) \\ \le 2\phi(q)\big(\log q - \log 2\pi e^{\gamma_0} + \log 2 + \log 2 + \tfrac{2}{3}\log 3 + 4.56 + 1.83\big),$$

which implies the assertion of the proposition. □

2.5.3. Bounds for the density $\delta(q;a,b)$

We use the results of the previous section to obtain explicit upper and lower bounds on $\delta(q;a,b)$; from these bounds, we can prove in particular that all of the largest values of these densities occur when the modulus q is less than an explicit bound. In the proof of Theorem 2.1.1, we expanded several functions, including an instance of sin, into their power series at the origin. While this yielded an excellent theoretical formula, for numerical purposes we will take a slightly different approach involving the error function $\operatorname{Erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}\,dt$. The following two lemmas allow us to write the density $\delta(q;a,b)$ in terms of the error function.

Lemma 2.5.10. *For any constants* $\nu > 0$ *and* ρ,

$$\int_{-\infty}^{\infty} t^4 e^{-\nu t^2/2}\,dt = \frac{3\sqrt{2\pi}}{\nu^{5/2}} \qquad \textit{and} \qquad \int_{-\infty}^{\infty} \frac{\sin\rho t}{t} e^{-\nu t^2/2}\,dt = \pi \operatorname{Erf}\left(\frac{\rho}{\sqrt{2\nu}}\right).$$

PROOF. For the first identity, a change of variables gives

$$\int_{\infty}^{\infty} t^4 e^{-\nu t^2/2}\,dt = \nu^{-5/2}\int_{\infty}^{\infty} w^4 e^{-w^2/2}\,dw = \nu^{-5/2} M_2(\infty) = \frac{3\sqrt{2\pi}}{\nu^{5/2}}$$

by Lemma 2.2.4. Our starting point for the second identity is [**1**, equation (7.4.6)] : for any constants $a > 0$ and x,

$$\int_0^{\infty} e^{-at^2}\cos 2xt\,dt = \frac{1}{2}\sqrt{\frac{\pi}{a}}e^{-x^2/a},$$

which can be rewritten as

$$\sqrt{\frac{\pi}{a}}e^{-x^2} = \int_{-\infty}^{\infty} e^{-at^2}\cos(2xt\sqrt{a})\,dt.$$

Integrating both sides from $x = 0$ to $x = w$ yields

$$\frac{\pi}{2\sqrt{a}}\operatorname{Erf}(w) = \int_{-\infty}^{\infty} e^{-at^2}\left(\int_0^w \cos(2xt\sqrt{a})\,dx\right)dt = \int_{-\infty}^{\infty} e^{-at^2}\frac{\sin(2wt\sqrt{a})}{2t\sqrt{a}}\,dt$$

(the interchanging of the integrals in the middle expression is justified by the absolute convergence of the integral). Setting $a = \frac{\nu}{2}$ and $w = \frac{\rho}{\sqrt{2\nu}}$, we obtain

$$\frac{\pi}{\sqrt{2\nu}}\operatorname{Erf}\left(\frac{\rho}{\sqrt{2\nu}}\right) = \int_{-\infty}^{\infty} e^{-\nu t^2/2}\frac{\sin\rho t}{t\sqrt{2\nu}}\,dt,$$

which establishes the lemma. □

Lemma 2.5.11. *Assume GRH and LI. Let* a *be a nonsquare* (mod q) *and* b *a square* (mod q). *If* $V(q;a,b) \ge 531$, *then*

$$\delta(q;a,b) = \frac{1}{2} + \frac{1}{2}\operatorname{Erf}\left(\frac{\rho(q)}{\sqrt{2V(q;a,b)}}\right) + \overline{O}\left(\frac{47.65\rho(q)}{V(q;a,b)^{3/2}} + 0.03506\frac{e^{-9.08\phi(q)}}{\phi(q)} + 63.68\rho(q)e^{-V(q;a,b)^{1/2}/2}\right).$$

PROOF. From Definition 2.2.6, we know that

$$\begin{aligned}\log \Phi_{q;a,b}(x) &= \sum_{\chi \pmod q} \sum_{\substack{\gamma>0 \\ L(1/2+i\gamma,\chi)=0}} \log J_0\bigg(\frac{2|\chi(a)-\chi(b)|x}{\sqrt{\frac14+\gamma^2}} \bigg) \\ &= \tfrac12 \sum_{\chi \pmod q} \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} \log J_0\bigg(\frac{2|\chi(a)-\chi(b)|x}{\sqrt{\frac14+\gamma^2}} \bigg)\end{aligned}$$

by the functional equation for Dirichlet L-functions. If $|x| \le \frac14$, then the argument of J_0 is at most $2\cdot 2\cdot \frac14/\frac12 = 2$ in absolute value. Since the Taylor expansion $\log J_0(x) = -x^2/4 + \overline{O}(.0311x^4)$ is valid for $|x| \le 2$, we see that

$$\begin{aligned}\log \Phi_{q;a,b}(x) = \tfrac12 \sum_{\chi \pmod q} \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} &\bigg(-\frac{|\chi(a)-\chi(b)|^2x^2}{\frac14+\gamma^2} \\ &\qquad + \overline{O}\bigg(.0311\frac{16|\chi(a)-\chi(b)|^4x^4}{(\frac14+\gamma^2)^2} \bigg) \bigg) \\ = -\tfrac12 x^2 \sum_{\chi \pmod q} \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} &\frac{|\chi(a)-\chi(b)|^2}{\frac14+\gamma^2} \\ + \overline{O}\bigg(\tfrac12 x^4 \sum_{\chi \pmod q} \sum_{\substack{\gamma\in\mathbb{R} \\ L(1/2+i\gamma,\chi)=0}} &.0311\frac{16\cdot 4|\chi(a)-\chi(b)|^2}{\frac14(\frac14+\gamma^2)} \bigg) \\ = -\tfrac12 V(q;a,b)x^2 + &\overline{O}(39.81V(q;a,b)x^4) \qquad (2.5.13)\end{aligned}$$

when $|x| \le \frac14$. Moreover, the error term in the expansion $\log J_0(x) = -x^2/4 + \overline{O}(.0311x^4)$ is always nonpositive as a consequence of Lemma 2.2.1(c), and hence the same is true for the recently obtained error term $\overline{O}(39.81V(q;a,b)x^4)$. This knowledge allows us to use the expansion $e^t = 1 + \overline{O}(t)$ for $t \le 0$, which yields

$$\Phi_{q;a,b}(x) = e^{-V(q;a,b)x^2/2}\big(1 + \overline{O}(39.81V(q;a,b)x^4)\big)$$

when $|x| \le \frac14$.

Proposition 2.2.10 says that when $V(q;a,b) \ge 531$,

$$\begin{aligned}\delta(q;a,b) = \frac12 + \frac{1}{2\pi}\int_{-V(q;a,b)^{-1/4}}^{V(q;a,b)^{-1/4}} \frac{\sin\rho(q)x}{x}\Phi_{q;a,b}(x)\,dx \\ + \overline{O}\bigg(0.03506\frac{e^{-9.08\phi(q)}}{\phi(q)} + 63.67\rho(q)e^{-V(q;a,b)^{1/2}/2} \bigg).\end{aligned}$$

Notice that $V(q;a,b)^{-1/4} \le 531^{-1/4} < \frac{1}{4}$, and so we may use our approximation for $\Phi_{q;a,b}(x)$ to deduce that

$$\begin{aligned}\delta(q;a,b) = \frac{1}{2}& \\ &+ \frac{1}{2\pi}\int_{-V(q;a,b)^{-1/4}}^{V(q;a,b)^{-1/4}} \frac{\sin\rho(q)x}{x} e^{-V(q;a,b)x^2/2}\big(1+\overline{O}(39.81V(q;a,b)x^4)\big)\,dx \\ &\qquad + \overline{O}\bigg(0.03506\frac{e^{-9.08\phi(q)}}{\phi(q)} + 63.67\rho(q)e^{-V(q;a,b)^{1/2}/2}\bigg). \qquad (2.5.14)\end{aligned}$$

The main term can be evaluated by the second identity of Lemma 2.5.10 :

$$\begin{aligned}&\frac{1}{2\pi}\int_{-V(q;a,b)^{-1/4}}^{V(q;a,b)^{-1/4}} \frac{\sin\rho(q)x}{x} e^{-V(q;a,b)x^2/2}\,dx \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\sin\rho(q)x}{x} e^{-V(q;a,b)x^2/2}\,dx \\ &\qquad + \overline{O}\bigg(\frac{1}{\pi}\int_{V(q;a,b)^{-1/4}}^{\infty} \bigg|\frac{\sin\rho(q)x}{x}\bigg| e^{-V(q;a,b)x^2/2}\,dx\bigg) \\ &= \frac{1}{2}\operatorname{Erf}\bigg(\frac{\rho(q)}{\sqrt{2V(q;a,b)}}\bigg) + \overline{O}\bigg(\frac{1}{\pi}\int_{V(q;a,b)^{-1/4}}^{\infty} \rho(q)V(q;a,b)^{1/4}xe^{-V(q;a,b)x^2/2}\,dx\bigg) \\ &= \frac{1}{2}\operatorname{Erf}\bigg(\frac{\rho(q)}{\sqrt{2V(q;a,b)}}\bigg) + \overline{O}\bigg(\frac{\rho(q)}{\pi V(q;a,b)^{3/4}} e^{-V(q;a,b)^{1/2}/2}\bigg).\end{aligned}$$

The error term in the integral in equation (2.5.14) can be estimated by the first identity of Lemma 2.5.10 :

$$\begin{aligned}&\frac{1}{2\pi}\int_{-V(q;a,b)^{-1/4}}^{V(q;a,b)^{-1/4}} \bigg|\frac{\sin\rho(q)x}{x}\bigg| e^{-V(q;a,b)x^2/2}39.81V(q;a,b)x^4\,dx \\ &\le 6.336\rho(q)V(q;a,b)\int_{-\infty}^{\infty} x^4e^{-V(q;a,b)x^2/2}\,dx \le 19.008\sqrt{2\pi}\cdot\rho(q)V(q;a,b)^{-3/2}.\end{aligned}$$

Therefore equation (2.5.14) becomes

$$\begin{aligned}\delta(q;a,b) &= \frac{1}{2} + \frac{1}{2}\operatorname{Erf}\bigg(\frac{\rho(q)}{\sqrt{2V(q;a,b)}}\bigg) + \overline{O}\bigg(\frac{\rho(q)}{\pi V(q;a,b)^{3/4}} e^{-V(q;a,b)^{1/2}/2}\bigg) \\ &\qquad + \overline{O}\bigg(\frac{47.65\rho(q)}{V(q;a,b)^{3/2}} + 0.03506\frac{e^{-9.08\phi(q)}}{\phi(q)} + 63.67\rho(q)e^{-V(q;a,b)^{1/2}/2}\bigg).\end{aligned}$$

Since $1/\pi V(q;a,b)^{3/4} \le 1/\pi(531)^{3/4} < 0.01$, this last estimate implies the statement of the lemma. □

We are now ready to bound $\delta(q;a,b)$ for all large prime moduli q.

Theorem 2.5.6. *Assume GRH and LI. If* $q \geq 400$ *is prime, then* $\delta(q; a, b) < 0.5262$ *for all reduced residues* a *and* $b \pmod q$*. If* $q \geq 1000$ *is prime, then* $\delta(q; a, b) < 0.51$.

PROOF. We may assume that a is a nonsquare $\pmod q$ and b is a square $\pmod q$, for otherwise $\delta(q; a, b) \leq \frac{1}{2}$. When $q \geq 331$ is prime, Proposition 2.5.4 and a quick calculation yield

$$\begin{aligned} V(q; a, b) &\geq 2(q-1)(\log q - 2.42) - 47.238 \log^2 q \\ &\geq 2q(\log q - 2.42) - 48 \log^2 q \geq 531. \end{aligned} \tag{2.5.15}$$

Therefore Lemma 2.5.11 applies, yielding (since $\rho(q) = 2$ and $\phi(q) = q - 1$)

$$\begin{aligned} \delta(q; a, b) &= \frac{1}{2} + \frac{1}{2} \operatorname{Erf}\left(\sqrt{\frac{2}{V(q; a, b)}}\right) \\ &\qquad + \overline{O}\left(\frac{95.3}{V(q; a, b)^{3/2}} + 0.03506 \frac{e^{-9.08q}}{q-1} + 127.36 e^{-V(q;a,b)^{1/2}/2}\right) \\ &\leq \frac{1}{2} + \frac{1}{2} \operatorname{Erf}\left(\sqrt{\frac{2}{2q(\log q - 2.42) - 48 \log^2 q}}\right) \\ &\qquad + \frac{95.3}{(2q(\log q - 2.42) - 48 \log^2 q)^{3/2}} + 0.03506 \frac{e^{-9.08q}}{q-1} \\ &\qquad + 127.36 e^{-\sqrt{q(\log q - 2.42)/2 - 12 \log^2 q}}, \end{aligned}$$

using the second inequality in equation (2.5.15). This upper bound is decreasing for $q \geq 331$, and so calculating it at $q = 400$ and $q = 1000$ establishes the inequalities given in the theorem. □

A similar bound for composite moduli q requires one last estimate.

Lemma 2.5.12. *For all* $q \geq 3$*, we have* $\rho(q) \leq 2q^{1.04/\log\log q}$.

PROOF. We first record some explicit estimates on the prime counting functions $\pi(y) = \sum_{p \leq y} 1$ and $\theta(y) = \sum_{p \leq y} \log p$. Rosser and Shoenfeld [**67**, Corollary 1 and Theorems 9 and 10] give, for $y \geq 101$, the bounds $0.84y \leq \theta(y) \leq$

$1.01624y$ and $\pi(y) \le 1.25506y/\log y$. Therefore

$$\pi(y) \le \frac{1.25506y}{\log y} \le \frac{1.25506\theta(y)/0.84}{\log\theta(y) - \log 1.01624} \le \frac{1.5\theta(y)}{\log\theta(y)} \tag{2.5.16}$$

(a calculation shows that the last inequality holds for $\theta(y) \ge 61$, which is valid in the range $y \ge 101$).

Now consider integers of the form $q(y) = \prod_{p \le y} p$, so that $\omega(q(y)) = \pi(y)$ and $\log q(y) = \theta(y)$. Equation (2.5.16) becomes

$$\omega(q(y)) \le 1.5(\log q(y))/\log\log q(y);$$

while the derivation was valid for $y \ge 101$, one can calculate that the inequality holds for $3 \le y \le 101$ as well. The following standard argument then shows that

$$\omega(q) \le \frac{1.5\log q}{\log\log q} \tag{2.5.17}$$

holds for all integers $q \ge 3$: if q has k distinct prime factors, then choose y to be the kth prime. Then the inequality (2.5.17) has been shown to hold for $q(y)$, and therefore it holds for q as well, since the left-hand side is k in both cases while the right-hand side is at least as large for q as it is for $q(y)$.

(This argument uses the fact that the right-hand side is an increasing function, which holds only for $q \ge e^e$; therefore technically we have proved (2.5.17) only for numbers with at least three distinct prime factors, since only then does the corresponding $q(y)$ exceed e^e. However, the right-hand side of (2.5.17) is always at least 4 in the range $q \ge 3$, and so numbers with one or two distinct prime factors easily satisfy the inequality.)

Finally, the inequality $\rho(q) \le 2^{\omega(q)+1}$ that was noted in Definition 2.1.1 allows us to conclude that $\rho(q) \le 2^{1+1.5(\log q)/\log\log q} < 2q^{1.04/\log\log q}$ for all $q \ge 3$, as desired. □

Theorem 2.5.7. *Assume GRH and LI. If* $q > 480$ *and* $q \notin \{840, 1320\}$, *then* $\delta(q; a, b) < 0.75$ *for all reduced residues* a *and* b (mod q).

PROOF. Again we may assume that a is a nonsquare (mod q) and b is a square (mod q). First we restrict to the range $q \ge 260000$; by Proposition 2.5.5 we have

$V(q;a,b) > 531$. Using Lemma 2.5.11, together with the upper bound for $\rho(q)$ from Lemma 2.5.12 and the lower bound for $V(q;a,b)$ from Proposition 2.5.5, we have

$$\begin{aligned}\delta(q;a,b) \le{}& \frac{1}{2} + \frac{1}{2}\operatorname{Erf}\left(\frac{2q^{1.04/\log\log q}}{2\sqrt{\phi(q)\big(\log q - 1.02\log\log q - 7.34\big)}}\right) \\ &+ \frac{33.7q^{1.04/\log\log q}}{\phi(q)^{3/2}\big(\log q - 1.02\log\log q - 7.34\big)^{3/2}} + 0.03506\frac{e^{-9.08\phi(q)}}{\phi(q)} \\ &+ 127.36q^{1.04/\log\log q}\exp\left(-\sqrt{\frac{\phi(q)}{2}\left(\log q - 1.02\log\log q - 7.34\right)}\right).\end{aligned} \tag{2.5.18}$$

Rosser and Schoenfeld [67, Theorem 15] have given the bound

$$\phi(q) > \frac{q}{e^{\gamma_0}\log\log q + 2.50637/\log\log q} \tag{2.5.19}$$

for $q \ge 3$. When this lower bound is substituted for $\phi(q)$ in the upper bound (2.5.18), the result is a smooth function of q that is well-defined and decreasing for $q \ge 260000$, and its value at $q = 260000$ is less than 0.75.

We now turn to the range $1000 \le q \le 260000$. We first compute explicitly, for each such modulus q, the lower bound for $V(q;a,b)$ in equation (2.5.12); the value of this sharper lower bound turns out always to exceed 531 in this range. Consequently, we may use Lemma 2.5.11 together with the lower bound for $V(q;a,b)$ from equation (2.5.12), obtaining

$$\begin{aligned}\delta(q;a,b) \le{}& \frac{1}{2} \\ &+ \frac{1}{2}\operatorname{Erf}\left(\frac{\rho(q)}{2\sqrt{\phi(q)\big(\log q - \log 2\pi e^{\gamma_0} - \sum_{p\mid q}\frac{\log p}{p-1} - \frac{4\log q}{q} - \frac{23.62\log^2 q}{q}\big)}}\right) \\ &+ \frac{17.85\rho(q)}{\phi(q)^{3/2}\big(\log q - \log 2\pi e^{\gamma_0} - \sum_{p\mid q}\frac{\log p}{p-1} - \frac{4\log q}{q} - \frac{23.62\log^2 q}{q}\big)^{3/2}} \\ &+ 0.03506\frac{e^{-9.08\phi(q)}}{\phi(q)} + 63.68\rho(q)\times \\ &\exp\left(-\sqrt{\frac{\phi(q)}{2}\left(\log q - \log 2\pi e^{\gamma_0} - \sum_{p\mid q}\frac{\log p}{p-1} - \frac{4\log q}{q} - \frac{23.62\log^2 q}{q}\right)}\right).\end{aligned}$$

This upper bound can be computed exactly for each q in the range $1000 \le q \le 260000$; the only five moduli for which the upper bound exceeds 0.75 are 1020, 1320, 1560, 1680, and 1848.

Finally, we use the methods described in Section 2.5.4, computing directly every value of $\delta(q;a,b)$ for the moduli $480 < q \le 1000$ and $q \in \{1020, 1320, 1560, 1680, 1848\}$ and verifying the inequality $\delta(q;a,b) < 0.75$ holds except for $q = 840$ and $q = 1320$, to complete the proof of the theorem. $\square$

2.5.4. Explicit computation of the densities

Throughout this section, we assume GRH and LI, and we let a denote a nonsquare (mod q) and b a square (mod q). In this section we describe the process by which we computed actual values of the densities $\delta(q;a,b)$, resulting for example in the data given in the tables and figures of this paper. In fact, we used two different methods for these computations, one that works for "small q" and one that works for "large q". For ease of discussion, we define the sets

$$\begin{aligned}
S_1 &= \{3 \le q \le 1000\colon q \not\equiv 2 \text{ (mod 4) and } \phi(q) < 80\} \\
S_2 &= \{101, 103, 107, 109, 113, 115, 119, 121, 123, 125, 129, 133, 141, 143, 145, 147, \\
&\qquad 153, 155, 159, 164, 165, 171, 172, 175, 176, 177, 183, 184, 188, 189, 195, 196, \\
&\qquad 200, 208, 212, 220, 224, 225, 231, 232, 236, 255, 260, 264, 276, 280, 288, 300, \\
&\qquad 308, 312, 324, 336, 348, 360, 372, 396, 420\} \\
S_3 &= \{3 \le q \le 1000\colon q \not\equiv 2 \text{ (mod 4) and } \phi(q) \ge 80\} \setminus S_2 \\
S_4 &= \{1020, 1320, 1560, 1680, 1848\}.
\end{aligned}$$

We omit integers congruent to 2 (mod 4) from these sets, since for odd q the prime number race (mod $2q$) is identical to the prime number race (mod q).

For the moduli q in the set $S_1 \cup S_2$, we numerically evaluated the integral in equation (2.2.10) directly; this method was used by Feuerverger and Martin **[22]** and is analogous to, and indeed based upon, the method used by Rubinstein and Sarnak **[68]**. We first used Rubinstein's computational package `lcalc` to calculate, for each character $\chi \pmod q$, the first $N(q)$ nontrivial zeros of $L(s,\chi)$ lying above the real axis. The term $\Phi_{q;a,b}$ in the integrand is a product of functions of the form $F(z,\chi)$, which is indexed by infinitely many zeros of $L(s,\chi)$; we approximated $F(z,\chi)$ by its truncation at $N(q)$ zeros, multiplied by a compensating quadratic polynomial as in [**68**, Section 4.3]. With this approximation to the integral (2.2.10), we truncated the range of integration to an interval $[-C(q), C(q)]$ and then discretized the truncated integral, replacing it by a sum over points spaced by $\varepsilon(q)$ as in [**68**, Section 4.1]. The result is an approximation to $\delta(q;a,b)$ that is valid up to at least 8 decimal places, provided we choose $N(q)$, $C(q)$, and $\varepsilon(q)$ carefully to get small errors. (All of these computations were performed using the computational software `Mathematica`.) Explicitly bounding the error in this process is not the goal of the present paper ; we refer the interested reader to [**68**] for rigorous error bounds of this kind, corresponding to their calculation of $\delta(q;N,R)$ for $q \in \{3,4,5,7,11,13\}$.

For the moduli q in the set $S_3 \cup S_4$ (and for any other moduli larger than 1000 we wished to address), we used an approach based on our asymptotic formulas for $\delta(q;a,b)$. We now outline a variant of the asymptotic formulas described earlier in this paper, one that was optimized somewhat for the the actual computations rather than streamlined for theoretical purposes.

We first note that a slight modification of the proof of Proposition 2.2.10 yields the estimate, for any $0 \le \kappa \le \frac{5}{24}$,

$$\begin{aligned}\delta(q;a,b) &= \frac{1}{2} + \frac{1}{2\pi}\int_{-\kappa}^{\kappa} \frac{\sin\rho(q)x}{x}\Phi_{q;a,b}(x)\,dx \\ &+ \overline{O}\biggl(\frac{1}{\pi}\int_{\kappa}^{5/24} \rho(q)|\Phi_{q;a,b}(x)|\,dx + 0.03506\frac{e^{-9.08\phi(q)}}{\phi(q)} + 63.67\rho(q)\bigl|\Phi_{q;a,b}\bigl(\tfrac{5}{24}\bigr)\bigr|\biggr)\end{aligned} \tag{2.5.20}$$

as long as $V(q;a,b) \ge 531$. In addition we have, for $|x| < \frac{3}{10}$, the inequalities

$$\begin{aligned}&-\tfrac{1}{2}V(q;a,b)x^2 - U(q;a,b)x^4 - 15.816U(q;a,b)x^6 \\ &\qquad\qquad \le \log\Phi_{q;a,b}(x) \le -\tfrac{1}{2}V(q;a,b)x^2 - U(q;a,b)x^4,\end{aligned}$$

where for convenience we have defined $U(q;a,b) = W_2(q;a,b)V(q;a,b)$; these inequalities can be proved using an argument similar to the calculation in equation (2.5.13), but employing the more precise estimate $\log J_0(z) = -z^2/4 - z^4/64 + \overline{O}(0.00386z^6)$ for $|z| \le 2$. Using the methods of Section 2.3.4, we also obtain the formula

$$\begin{aligned} U(q;a,b) = \frac{\phi(q)}{2}(3+\iota_q(a^2b^{-2}))\bigg(\log\frac{q}{2\pi e^{-\gamma_0}} - \sum_{p\mid q}\frac{\log p}{p-1} - \frac{\zeta(2)}{2}\bigg) \\ + \frac{\phi(q)}{2}\bigg(4\frac{\Lambda(q/(q,a-b))}{\phi(q/(q,a-b))} - \frac{\Lambda(q/(q,a^2-b^2))}{\phi(q/(q,a^2-b^2)} \\ - \big(\iota_q(-a^2b^{-2}) - 4\iota_q(-ab^{-1})\big)\Big(\log 2 + \frac{\zeta(2)}{4}\Big)\bigg) \\ + \frac{1}{4}\sum_{\chi \,(\mathrm{mod}\ q)} |\chi(a)-\chi(b)|^4\bigg(2\frac{L'(1,\chi)}{L(1,\chi)} - \frac{L''(1,\chi)}{L(1,\chi)} + \Big(\frac{L'(1,\chi)}{L(1,\chi)}\Big)^2\bigg). \quad (2.5.21) \end{aligned}$$

If we define $\kappa(q;a,b) = \min(\frac{\pi}{\rho(q)}, V(q;a,b)^{-1/4})$, then we know that $\kappa(q;a,b) \le \frac{5}{24}$ because of the lower bound $V(q;a,b) \ge 531$, and also that $(\sin\rho(q)x)/x$ is nonnegative for $|x| \le \kappa(q;a,b)$. Hence, equation (2.5.20) and the subsequent discussion establishes the following proposition :

Proposition 2.5.8. *Assume GRH and LI, and let* a *be a nonsquare* (mod q) *and* b *a square* (mod q). *If* $V(q;a,b) \ge 531$, *then*

$$\begin{aligned} \frac{1}{2} + \frac{1}{2\pi}\int_{-\kappa(q;a,b)}^{\kappa(q;a,b)} \frac{\sin\rho(q)x}{x} e^{-V(q;a,b)x^2/2 - U(q;a,b)x^4 - 15.816U(q;a,b)x^6} \, dx - Y(q;a,b) \\ \le \delta(q;a,b) \le \frac{1}{2} + \frac{1}{2\pi}\int_{-\kappa(q;a,b)}^{\kappa(q;a,b)} \frac{\sin\rho(q)x}{x} e^{-V(q;a,b)x^2/2 - U(q;a,b)x^4} \, dx + Y(q;a,b), \end{aligned}$$

where

$$\begin{aligned} Y(q;a,b) = \frac{\rho(q)}{\pi}\int_{\kappa(q;a,b)}^{5/24} e^{-V(q;a,b)x^2/2 - U(q;a,b)x^4} \, dx \\ + 0.03506\frac{e^{-9.08\phi(q)}}{\phi(q)} + 63.67\rho(q)e^{-25V(q;a,b)/1152 - (5/24)^4 U(q;a,b)} \end{aligned}$$

and formulas for $V(q;a,b)$ *and* $U(q;a,b)$ *are given in Theorem 2.1.2 and equation* (2.5.21), *respectively.*

The inequalities in Proposition 2.5.8 give accurate evaluations of $\delta(q;a,b)$ when $\phi(q)$ is large; we chose the inequality $\phi(q) \ge 80$ to be our working definition of "large". For each of the moduli q in the set $S_3 \cup S_4$, we computed

TAB. 2.7. The 20 smallest values of $\delta(244; a, 1)$ and of $\delta(997; a, 1)$, calculated using Proposition 2.5.8

q	a	a^{-1}	$\delta(q; a, 1)$	Error bound
244	243	243	0.558910	0.000022
244	123	123	0.559000	0.000018
244	3	163	0.562304	0.000020
244	7	35	0.563216	0.000022
244	31	63	0.563543	0.000022
244	153	185	0.563804	0.000021
244	11	111	0.564069	0.000024
244	29	101	0.564124	0.000024
244	17	201	0.564321	0.000023
244	33	37	0.564436	0.000024
244	19	167	0.564741	0.000024
244	23	191	0.564786	0.000023
244	107	187	0.565310	0.000024
244	69	145	0.565319	0.000022
244	53	221	0.565376	0.000022
244	85	89	0.565606	0.000022
244	129	157	0.565683	0.000021
244	173	189	0.565707	0.000023
244	177	193	0.565859	0.000023
244	103	199	0.565861	0.000024

q	a	a^{-1}	$\delta(q; a, 1)$	Error bound
997	2	499	0.508116457	0.000000014
997	5	399	0.508142372	0.000000015
997	7	285	0.508184978	0.000000015
997	11	272	0.508238549	0.000000016
997	17	176	0.508279881	0.000000016
997	29	722	0.508329803	0.000000016
997	37	512	0.508345726	0.000000016
997	41	535	0.508351018	0.000000016
997	8	374	0.508353451	0.000000016
997	43	371	0.508355411	0.000000016
997	47	297	0.508358709	0.000000016
997	61	474	0.508368790	0.000000016
997	163	367	0.508392448	0.000000016
997	103	242	0.508392587	0.000000016
997	113	150	0.508395577	0.000000016
997	181	661	0.508397690	0.000000016
997	127	840	0.508402416	0.000000016
997	157	870	0.508404812	0.000000016
997	283	613	0.508406794	0.000000016
997	179	518	0.508406994	0.000000016

every possible value of $V(q; a, b)$ and verified that they all exceed 531, so that Proposition 2.5.8 can be used. (The reason that the moduli in S_2 were calculated using the first method, rather than this one, is because at least one variance $V(q; a, b)$ was less than 531 for each of the moduli in S_2.) We then calculated the upper and lower bounds of Proposition 2.5.8, using numerical integration in `pari/gp`, to obtain all values of $\delta(q; a, b)$. The calculation of $V(q; a, b)$ and $U(q; a, b)$ involve the analytic terms $L(1, \chi)$, $L'(1, \chi)$, and $L''(1, \chi)$; we used the `pari/gp` package `computeL` (see **[19]**) to obtain these values accurate to 16 decimal places.

Table 2.7 gives a sample of the data we calculated with this second method, including the error bounds obtained. The error bounds are stronger for when q and $\phi(q)$ are large, explaining why the error bounds for the large prime $q = 997$ are so much better than for the smaller composite number $q = 244$.

We also take this opportunity to reinforce the patterns described in Section 2.4.1. For $q = 244$, the residue class $a = 123$ has the property that $q/(q, a-1) = 2$; thus the contribution of $K_{244}(122)$ to $\Delta(244; 123, 1)$ reduces the density $\delta(244; 123, 1)$. We see also the familiar small densities corresponding to $a = 243 \equiv -1 \pmod{244}$ and to small prime values of a. For $q = 997$, the small prime values of a (among those that are nonsquares modulo 997) appear in perfect order. We point out that the residue class $a = 8$ is almost in its correct limiting position, since the contribution to $\Delta(997; a, 1)$ is inversely correlated to $\frac{\Lambda(a)}{a}$, and $\frac{\Lambda(41)}{41} > \frac{\Lambda(43)}{43} > \frac{\Lambda(8)}{8} > \frac{\Lambda(47)}{47}$.

We mention that we undertook the exercise of calculating values $\delta(q; a, b)$ by both methods, for several intermediate values of q, as a way to verify our computations. For example, the calculations of $\delta(163; a, b)$ (see Table 2.3) were done using the integral formula (2.2.10) as described above. We calculated these same densities using Proposition 2.5.8; the error bounds obtained were all at most 4.6×10^{-6}, and the results of the first calculation all lay comfortably within the intervals defined by the second calculation.

Finally, the upper bounds for $\delta(q; a, b)$ in Theorems 2.5.6 and 2.5.7, together with the explicit calculation of the densities $\delta(q; a, b)$ for $q \in S_1 \cup S_2 \cup S_3 \cup S_4$, allow us to determine the most biased possible two-way races, that is, the largest values of $\delta(q; a, b)$ among all possible choices of q, a, and b. In particular, we verified Theorem 2.1.6 in this way, and we list the 120 largest densities in Table 2.8; there are precisely 117 distinct densities above $\frac{9}{10}$. (It is helpful to recall here that $\delta(q; a, 1) = \delta(q; a^{-1}, 1)$ and that $\delta(q; a, 1) = \delta(q; ab, b)$ for any nonsquare a and square b modulo q.)

TAB. 2.8. The top 120 most unfair prime number races

q	a	a^{-1}	$\delta(q;a,1)$
24	5	5	0.999988
24	11	11	0.999983
12	11	11	0.999977
24	23	23	0.999889
24	7	7	0.999834
24	19	19	0.999719
8	3	3	0.999569
12	5	5	0.999206
24	17	17	0.999125
3	2	2	0.999063
8	7	7	0.998939
24	13	13	0.998722
12	7	7	0.998606
8	5	5	0.997395
4	3	3	0.995928
120	71	71	0.988747
120	59	59	0.988477
60	11	11	0.987917
60	29	29	0.986855
120	109	109	0.986835
60	19	19	0.986459
120	89	89	0.986364
120	79	79	0.986309
120	101	101	0.984792
15	2	8	0.983853
120	13	37	0.980673
40	19	19	0.980455
60	7	43	0.979323
120	23	47	0.979142
15	14	14	0.979043
120	17	113	0.978762
120	7	103	0.978247
48	23	23	0.978096
120	43	67	0.978013
60	17	53	0.977433
48	41	41	0.977183
40	29	29	0.977161
20	3	7	0.976713
120	53	77	0.976527
60	23	47	0.975216
120	91	91	0.975051
120	83	107	0.975001
120	29	29	0.974634
120	19	19	0.974408
120	11	11	0.971988
48	31	31	0.970470
40	7	23	0.969427
40	13	37	0.969114
120	73	97	0.967355
20	19	19	0.966662
15	7	13	0.964719
120	31	31	0.963190
60	13	37	0.963058
60	59	59	0.962016
40	31	31	0.960718
48	5	29	0.960195
40	3	27	0.960099
16	7	7	0.959790
48	11	35	0.959245
120	119	119	0.957182
15	11	11	0.955226
120	41	41	0.955189
48	19	43	0.952194
5	2	3	0.952175
20	13	17	0.948637
120	61	61	0.948586
60	41	41	0.947870
16	3	11	0.947721
48	13	37	0.946479
40	17	33	0.946002
40	11	11	0.945757
40	39	39	0.942554
60	31	31	0.941802
48	7	7	0.939000
16	5	13	0.938369
168	125	125	0.936773
168	155	155	0.935843
168	47	143	0.932099
168	61	157	0.931981
84	41	41	0.931702
20	11	11	0.931367
168	139	139	0.931362
168	55	55	0.931346
48	47	47	0.929478
168	67	163	0.928944
84	71	71	0.928657
168	41	41	0.927933
84	55	55	0.927755
168	71	71	0.927349
16	15	15	0.926101

TAB. 2.9. The top 120 most unfair prime number races (continued)

q	a	a^{-1}	$\delta(q; a, 1)$
168	65	137	0.923960
168	53	149	0.923937
168	83	83	0.923868
21	5	17	0.923779
168	79	151	0.922597
40	21	21	0.922567
168	37	109	0.922359
168	17	89	0.920542
48	17	17	0.918910
56	27	27	0.918015
168	59	131	0.917874
168	23	95	0.917718
168	31	103	0.917278
168	29	29	0.915514
72	53	53	0.913533
21	2	11	0.911872
168	19	115	0.911412
168	11	107	0.909850
168	73	145	0.908239
168	5	101	0.908206
56	31	47	0.906135
84	67	79	0.905578
168	13	13	0.904525
168	97	97	0.904162
72	35	35	0.903755
84	47	59	0.902413
56	37	53	0.900863
84	53	65	0.899063
28	11	23	0.898807
168	127	127	0.898647

Chapitre 3

RESIDUE CLASSES CONTAINING AN UNEXPECTED NUMBER OF PRIMES

Author : Daniel Fiorilli

Abstract : We fix a non-zero integer a and consider arithmetic progressions $a \bmod q$, with q varying over a given range. We show that for certain specific values of a, the arithmetic progressions $a \bmod q$ contain, on average, significantly fewer primes than expected.

3.1. INTRODUCTION

The prime number theorem for arithmetic progressions asserts that

$$\psi(x;q,a) \sim \psi(x)/\phi(q)$$

for any a and q such that $(a,q)=1$. Another way to say this is that the primes are equidistributed in the $\phi(q)$ arithmetic progressions $a \bmod q$ with $(a,q) = 1$.

Fix an integer $a \neq 0$. We will be interested in the number of primes in the arithmetic progressions $a \bmod q$ with q varying in certain ranges, and we will show that for specific values of a, there are significantly fewer primes in these arithmetic progressions than in typical arithmetic progressions. Consider the average value of $\psi(x;q,a) - \psi(x)/\phi(q)$ over q. One might expect that no matter what the value of a is, the cancellations in these oscillating terms will force the average to be very small. However it turns out that the average is highly dependent on the arithmetical properties of a.

Here is the main result of the paper.

Theorem 3.1.1. *Fix an integer* $a \neq 0$ *and let* $M = M(x) \leq \log^B x$ *where* $B > 0$ *is a fixed real number. The average error term in the usual approximation for the number of primes* $p \equiv a \bmod q$ *with* $p \leq x$*, where* $(q, a) = 1$ *and* $q \leq x/M$*, is*

$$\frac{1}{\frac{\phi(a)}{a}\frac{x}{M}} \sum_{\substack{q \leq \frac{x}{M} \\ (q,a)=1}} \left(\psi(x; q, a) - \Lambda(a) - \frac{\psi(x)}{\phi(q)} \right) = \mu(a, M) + O_{a,\epsilon,B}\left(\frac{1}{M^{\frac{205}{538}-\epsilon}} \right) \tag{3.1.1}$$

with

$$\mu(a, M) := \begin{cases} 0 & \textit{if } \omega(a) \geq 2 \\ -\frac{1}{2}\log p & \textit{if } a = \pm p^e \\ -\frac{1}{2}\log M - C_5 & \textit{if } a = \pm 1, \end{cases}$$

where

$$C_5 := \frac{1}{2}\left(\log 2\pi + \gamma + \sum_p \frac{\log p}{p(p-1)} + 1 \right).$$

Remark 3.1.1. *Assuming Lindelöf's hypothesis, we can replace the error term in (3.1.1) by* $O_{a,\epsilon,B}\left(\frac{1}{M^{1/2-\epsilon}}\right)$.

Remark 3.1.2. *We subtracted* $\Lambda(a)$ *from* $\psi(x; q, a)$ *in (3.1.1) because the arithmetic progression* $a \bmod q$ *contains the prime power* p^e *for all* q *if* $a = p^e$.

Remark 3.1.3. *It may be preferable to replace* $\psi(x)$ *by* $\psi(x, \chi_0)$ *in Theorem 3.1.1, since the quantity*

$$\psi(x; q, a) - \psi(x, \chi_0)/\phi(q)$$

is the discrepancy (with signs) of the sequence of primes in the residue classes mod q*. One can do this with a negligible error term.*

3.2. Past Results

The study of the discrepancy $\psi(x; q, a) - x/\phi(q)$ on average has been a fruitful subject over the past decades. For example, the celebrated theorem of Bombieri-Vinogradov gives a bound on the sum of the mean absolute value of the maximum of this discrepancy over all $1 \leq a < q$ with $(a, q) = 1$, summed over $q \leq x^{1/2-o(1)}$. The Hooley-Montgomery refinement of the

Barban-Davenport-Halberstam Theorem gives an estimation of the variance of $\psi(x;q,a)-x/\phi(q)$, again for all values of a in the range $1\le a<q$ with $(a,q)=1$ for $q<x/\log^A x$. The mean value of $\psi(x;q,a)-x/\phi(q)$ was studied for fixed values of a for $q\ge x^{1/2}$ (see **[10]**,**[24]** or **[26]**), and bounds on this mean value turned out to be applicable to Titchmarsh's divisor problem, first solved by Linnik. The best result so far for this problem was obtained by Friedlander and Granville.

Theorem 3.2.1 (Friedlander, Granville)**.** *Let* $0<\lambda<1/4$, $A>0$ *be given. Then uniformly for* $0<|a|<x^\lambda$, $2\le Q\le x/3$ *we have*

$$\sum_{\substack{Q<q\le 2Q\\(q,a)=1}}\left(\psi(x;q,a)-\frac{x}{\phi(q)}\right)\ll_{\lambda,A}2^{\omega(a)}Q\log(x/Q)+\frac{x}{\log^A x}+Q\log|a|. \quad (3.2.1)$$

Remark 3.2.1. *If* a *is not a prime power the term* $Q\log|a|$ *may be deleted.*

Theorem 3.2.1 is a refinement of the deep results of Bombieri-Friedlander-Iwaniec **[10]** and of Fouvry **[24]**, and makes use of the dispersion method combined with Fourier analysis and involved estimates on Kloosterman sums.

The main method used in our paper, which we will refer to as the "divisor switching" technique, stemmed from the work of Dirichlet on the divisor problem. Variants of his "hyperbola method" were subsequently used in many different contexts, and have become a very important tool in analytic number theory. The variant which will be used in this paper is very similar to that of Hooley **[39]**.

3.3. Main results

A number of constants will appear throughout the paper. We will denote by γ the Euler-Mascheroni constant.

Definition 3.3.1. *We define*

$$C_1(a):=\frac{\zeta(2)\zeta(3)}{\zeta(6)}\frac{\phi(a)}{a}\prod_{p|a}\left(1-\frac{1}{p^2-p+1}\right),\quad C_5:=\frac{1}{2}\left(\log 2\pi+\gamma+\sum_p\frac{\log p}{p(p-1)}+1\right),$$

$$C_3(a):=C_1(a)\left(\gamma-1-\sum_p\frac{\log p}{p^2-p+1}+\sum_{p|a}\frac{p^2\log p}{(p-1)(p^2-p+1)}\right).$$

We will denote by $\omega(a)$ the number of distincts prime factors of a. Note that there are $\sim \frac{\phi(a)}{a}\frac{x}{M}$ terms in the sum over $1 \leq q \leq x/M$ with the condition $(q,a)=1$.

We will see later on that Theorem 3.1.1 can be made uniform for a in some range, which is the object of the more technical Theorem 3.7.1. A consequence of this is the following.

Theorem 3.3.1. *Fix* B, $\lambda < \frac{1}{4}$ *and* $\eta \leq \frac{1}{5}$, *three positive real numbers. We have for* $M \leq (\log x)^B$ *that the proportion of integers* a *in the range* $0 < |a| \leq x^{\lambda}$ *for which*

$$\frac{1}{\frac{\phi(a)}{a}\frac{x}{M}} \sum_{\substack{q \leq \frac{x}{M} \\ (q,a)=1}} \left(\psi(x;q,a) - \frac{\psi(x)}{\phi(q)} \right) = O_{B,\lambda}\left(\frac{1}{M^{\eta}} \right) \tag{3.3.1}$$

is at least $\frac{1-\frac{21}{8}\eta}{1+\eta+\eta^2}e^{-\gamma}$, *where* γ *is the Euler-Mascheroni constant.*

One can consider a different range for q.

Proposition 3.3.1. *Fix* A *and* $\lambda < \frac{1}{4}$, *two positive reals numbers. For* a *in the range* $0 < |a| \leq x^{\lambda}$ *such that* $\omega(a) \leq 10 \log\log x$ *we have*

$$\frac{1}{\frac{\phi(a)}{a}x} \sum_{\substack{q \leq x \\ (q,a)=1}} \left(\psi(x;q,a) - \Lambda(a) - \frac{\psi(x)}{\phi(q)} \right) = \frac{a}{\phi(a)} C_3(a) + O_{A,\lambda}\left(\frac{1}{\log^A x} \right). \tag{3.3.2}$$

More generally, for $M \geq 1$ *a fixed integer we have*

$$\frac{1}{\frac{\phi(a)}{a}\frac{x}{M}} \sum_{\substack{q \leq \frac{x}{M} \\ (q,a)=1}} \left(\psi(x;q,a) - \Lambda(a) - \frac{\psi(x)}{\phi(q)} \right) = \mu'(a,M) + O_{A,\lambda}\left(\frac{1}{\log^A x} \right), \tag{3.3.3}$$

where

$$\mu'(a,M) := \frac{a}{\phi(a)} M \left(C_1(a)\log M + C_3(a) - \sum_{\substack{r \leq M \\ (r,a)=1}} \frac{1}{\phi(r)}\left(1 - \frac{r}{M}\right) \right).$$

Note that $\mu'(a,1) = \frac{a}{\phi(a)}C_3(a)$.

By inverting the order of summation in Proposition 3.3.1, one gets the following corollary, which is an example of application of the results of Bombieri, Fouvry, Friedlander, Granville and Iwaniec (see **[10]**, **[24]** and **[26]**).

Corollary 3.3.1 (Titchmarsh's divisor problem)**.** *Fix A and $\lambda < \frac{1}{4}$, two positive real numbers. For a in the range $0 < |a| \leq x^{\lambda}$ such that $\omega(a) \leq 10 \log\log x$ we have*

$$\sum_{|a|<n\leq x} \Lambda(n)\tau(n-a) = C_1(a)x\log x + (2C_3(a) + C_1(a))x + O_{A,\lambda}\left(\frac{x}{\log^A x}\right). \tag{3.3.4}$$

Note that the constant 10 in Proposition 3.3.1 and Corollary 3.3.1 can be replaced by an arbitrary large real number.

3.4. Acknowledgements

I would like to thank my supervisor Andrew Granville for his advice, as well as my colleagues Farzad Aryan, Mohammad Bardestani, Tristan Freiberg and Kevin Henriot for many fruitful conversations. I would also like to thank John Friedlander for suggesting Remark 3.1.3. L'auteur est titulaire d'une bourse doctorale du Conseil de recherches en sciences naturelles et en génie du Canada.

3.5. Notation

Definition 3.5.1. *For $n \neq 0$ an integer (possibly negative), we define*

$$\Lambda(n) := \begin{cases} \log p & \text{if } n = p^e \\ 0 & \text{otherwise,} \end{cases} \qquad \vartheta(n) := \begin{cases} \log p & \text{if } n = p \\ 0 & \text{otherwise.} \end{cases}$$

Definition 3.5.2.

$$\psi(x;q,a) := \sum_{\substack{n\leq x \\ n\equiv a \bmod q}} \Lambda(n), \qquad \theta(x;q,a) := \sum_{\substack{n\leq x \\ n\equiv a \bmod q}} \vartheta(n).$$

The following definition is non-standard but will be useful in the proofs.

Definition 3.5.3.

$$\psi^*(x;q,a) := \sum_{\substack{|a|<n\leq x \\ n\equiv a \bmod q}} \Lambda(n), \tag{3.5.1}$$

$$\theta^*(x;q,a) := \sum_{\substack{|a|<n\leq x \\ n\equiv a \bmod q}} \vartheta(n). \tag{3.5.2}$$

We will need to consider the prime divisors of a which are less than M.

Definition 3.5.4. *For a an integer and $M > 0$ a real number, we define*

$$a_M := \prod_{\substack{p \mid a \\ p \leq M}} p. \tag{3.5.3}$$

The error term $E(M, a)$ will be defined depending on the context, so one has to pay attention to its definition in every statement.

3.6. LEMMAS

We begin by recalling the Hooley-Montgomery "divisor switching" technique.

Lemma 3.6.1. *Let a be an integer such that $0 < |a| \leq x^{1/4}$ and let $M = M(x)$ such that $1 \leq M < x$. We have*

$$\sum_{\substack{\frac{x}{M} < q \leq x \\ (q,a)=1}} \sum_{\substack{|a| < p \leq x \\ p \equiv a \bmod q}} \log p = \sum_{\substack{1 \leq r < (x-a)\frac{M}{x} \\ (r,a)=1}} \sum_{\substack{r\frac{x}{M} + a < p \leq x \\ p \equiv a \bmod r \\ p > |a|}} \log p + O(|a| \log x). \tag{3.6.1}$$

PROOF. Clearly,

$$\sum_{\substack{\frac{x}{M} < q \leq x \\ (q,a)=1}} \sum_{\substack{|a| < p \leq x \\ p \equiv a \bmod q}} \log p = \sum_{\substack{\frac{x}{M} < q \leq x \\ (q,a)=1}} \sum_{\substack{|a| < p \leq x \\ p \equiv a \bmod q \\ p > a + \frac{x}{M}}} \log p. \tag{3.6.2}$$

Now we will apply the switching technique which is somewhat similar to Dirichlet's hyperbola method. Setting $p - a = rq$ in (3.6.2), one can sum over r instead of summing over q. Now $(r, a) = 1$, else $(p, a) > 1$ so $p \mid a$, but this is impossible since $p > |a|$. Taking $a > 0$ for now, we get that (3.6.2) is equal to

$$\sum_{\substack{\frac{x}{M} < q \leq x - a \\ (q,a)=1}} \sum_{\substack{|a| < p \leq x \\ p \equiv a \bmod q \\ p > a + \frac{x}{M}}} \log p = \sum_{\substack{1 \leq r < (x-a)\frac{M}{x} \\ (r,a)=1}} \sum_{\substack{r\frac{x}{M} + a < p \leq x \\ p \equiv a \bmod r \\ p > |a|}} \log p. \tag{3.6.3}$$

If we had $a < 0$, the minor difference in (3.6.3) would be the $r = 1$ term where the additional condition $p \leq x + a$ is needed. Moreover, additional terms would be needed in passing from the right hand side of (3.6.2) to the left hand side of (3.6.3). Both these modifications can be made at the cost of adding an error term of size $\ll |a| \log x$.

□

Lemma 3.6.2. *We have the following estimates :*

$$\sum_{\substack{n \le M \\ (n,a)=1}} \frac{1}{\phi(n)} = C_1(a) \log M + C_1(a) + C_3(a) + O\left(2^{\omega(a)} \frac{\log M}{M}\right), \tag{3.6.4}$$

$$\sum_{\substack{n \le M \\ (n,a)=1}} \frac{n}{\phi(n)} = C_1(a) M + O\left(2^{\omega(a)} \log M\right). \tag{3.6.5}$$

Note that without loss of generality, we can replace a by a_M on the right side of (3.6.4) and (3.6.5).

PROOF. The proof of (3.6.4) is very similar to the proof of Lemma 13.1 in **[27]**. One first has to prove the following estimate :

$$\sum_{\substack{n \le M \\ (n,a)=1}} \frac{1}{n} = \frac{\phi(a)}{a} \left(\log M + \gamma + \sum_{p | a} \frac{\log p}{p-1} \right) + O\left(\frac{2^{\omega(a)}}{M}\right). \tag{3.6.6}$$

One then writes

$$\sum_{\substack{n \le M \\ (n,a)=1}} \frac{1}{\phi(n)} = \sum_{\substack{n \le M \\ (n,a)=1}} \frac{1}{n} \sum_{d | n} \frac{\mu^2(d)}{\phi(d)} = \sum_{\substack{d \le M \\ (d,a)=1}} \frac{\mu^2(d)}{d\phi(d)} \sum_{\substack{r \le M/d \\ (r,a)=1}} \frac{1}{r} \tag{3.6.7}$$

and inserts the estimate (3.6.6) into (3.6.7). The final step is to bound the tail of the sums and to compute the following constants :

$$\sum_{(d,a)=1} \frac{\mu^2(d)}{d\phi(d)} = \frac{\zeta(2)\zeta(3)}{\zeta(6)} \prod_{p | a} \left(1 - \frac{1}{p^2 - p + 1}\right),$$

$$\sum_{(d,a)=1} \frac{\mu^2(d)}{d\phi(d)} \log d = \frac{\zeta(2)\zeta(3)}{\zeta(6)} \prod_{p | a} \left(1 - \frac{1}{p^2 - p + 1}\right) \sum_{p \nmid a} \frac{\log p}{p^2 - p + 1}.$$

The proof of (3.6.5) goes along the same lines. □

Our very delicate analysis forces us to give some details about the "trivial" estimates for the prime counting functions.

Lemma 3.6.3. *We have for any $\epsilon > 0$ that*

$$\sum_{\substack{q \le x \\ (q,a)=1}} (\psi^*(x;q,a) - \theta^*(x;q,a)) \ll_\epsilon x^{1/2+\epsilon}. \tag{3.6.8}$$

PROOF.

$$\begin{aligned} &\sum_{\substack{q \le x \\ (q,a)=1}} (\psi^*(x;q,a) - \theta^*(x;q,a)) \\ &\qquad \le \sum_{q \le x} \sum_{\substack{|a|<p^e \le x \\ p^e \equiv a \bmod q \\ e \ge 2}} \log p \le \sum_{2 \le e \le \frac{\log x}{\log 2}} \sum_{p \le x^{1/e}} \sum_{\substack{q \le x \\ q \mid p^e - a}} \log p \\ &\qquad \le \log x \sum_{2 \le e \le \frac{\log x}{\log 2}} \sum_{p \le x^{1/e}} \tau(p^e - a) \ll_\epsilon x^{\epsilon/2} \sum_{2 \le e \le \frac{\log x}{\log 2}} \pi(x^{1/e}) \\ &\qquad \ll_\epsilon x^{1/2+\epsilon}. \end{aligned}$$

□

Lemma 3.6.4. *Let $a \neq 0$ be an integer and $1 \le Q \le x$. We have*

$$\sum_{\substack{q \le Q \\ (q,a)=1}} (\psi(x;q,a) - \Lambda(a) - \psi^*(x;q,a)) = O(|a| \log^2 |a|), \tag{3.6.9}$$

$$\sum_{\substack{q \le Q \\ (q,a)=1}} (\theta(x;q,a) - \vartheta(a) - \theta^*(x;q,a)) = O(|a| \log^2 |a|). \tag{3.6.10}$$

PROOF. Note that as soon as $q > 2|a|$, there are no integers congruent to a mod q in the interval $[1, |a|)$. We then have

$$\begin{aligned} \sum_{\substack{q \le Q \\ (q,a)=1}} (\psi(x;q,a) - \Lambda(a) - \psi^*(x;q,a)) &= \sum_{\substack{q \le Q \\ (q,a)=1}} \sum_{\substack{1 \le n < |a| \\ n \equiv a \bmod q}} \Lambda(n) \\ \le \sum_{\substack{q \le 2|a| \\ (q,a)=1}} \log |a| \sum_{\substack{1 \le n < |a| \\ n \equiv a \bmod q}} 1 &\ll \sum_{\substack{q \le 2|a| \\ (q,a)=1}} \log |a| \left(\frac{|a|}{q} \right) \ll |a| \log^2 |a|. \end{aligned}$$

The proof for θ and θ^* is similar. □

Lemma 3.6.5. *Let* $I \subset [1,x] \cap \mathbb{N}$. *We have*

$$\sum_{q \in I} \left(\frac{x}{\phi(q)} - \frac{\psi(x)}{\phi(q)} \right) \ll x e^{-C\sqrt{\log x}}, \tag{3.6.11}$$

where C *is an absolute positive constant.*

PROOF. This follows from Lemma 3.6.2 and the prime number theorem. □

To prove Lemma 3.6.8 we will need bounds on $\zeta(s)$.

Lemma 3.6.6. *Define* $\theta := \frac{32}{205}$ *and take any* $\epsilon > 0$. *In the region* $|\sigma + it - 1| > \frac{1}{10}$, *we have*

$$\zeta(\sigma + it) \ll_\epsilon (|t| + 1)^{\mu(\sigma)+\epsilon},$$

where

$$\mu(\sigma) = \begin{cases} 1/2 - \sigma & \textit{if } \sigma \leq 0 \\ 1/2 + (2\theta - 1)\sigma & \textit{if } 0 \leq \sigma \leq 1/2 \\ 2\theta(1 - \sigma) & \textit{if } 1/2 \leq \sigma \leq 1 \\ 0 & \textit{if } \sigma \geq 1. \end{cases}$$

PROOF. For the values outside the critical strip, see for example section II.3.4 of **[76]**. In the critical strip, we use an estimate due to Huxley **[41]**, which showed that $\zeta(1/2 + it) \ll_\epsilon (|t| + 1)^{\frac{32}{205}+\epsilon}$. The lemma then follows by convexity of μ. □

Remark 3.6.1. *Under Lindelöf's hypothesis, the conclusion of Lemma 3.6.6 holds with* $\theta = 0$.

Lemma 3.6.7 (Perron's formula)**.** *Let* $0 < \kappa < 1$, $y > 0$ *and define*

$$h(y) := \begin{cases} 0 & \textit{if } 0 < y < 1 \\ 1 - \frac{1}{y} & \textit{if } y \geq 1. \end{cases}$$

We have

$$h(y) = \frac{1}{2\pi i} \int_{(\kappa)} \frac{y^s}{s(s+1)} ds.$$

Moreover, for $T \geq 1$ *and positive* $y \neq 1$*, we have the estimate*

$$h(y) = \frac{1}{2\pi i}\int_{\kappa - iT}^{\kappa + iT} \frac{y^s}{s(s+1)}ds + O\left(\frac{y^\kappa}{T^2|\log y|}\right).$$

Finally, for $y = 1$*,*

$$0 = h(1) = \frac{1}{2\pi i}\int_{\kappa - iT}^{\kappa + iT} \frac{ds}{s(s+1)} + O\left(\frac{1}{T}\right).$$

PROOF. The first assertion is an easy application of the residue theorem.

Now take $y > 1$. We have again by the residue theorem that for any large integer $K \geq 3$ and for $T \geq 1$,

$$\begin{aligned}\frac{1}{2\pi i}\int_{\kappa - iT}^{\kappa + iT} \frac{y^s}{s(s+1)}ds - h(y) &= \frac{1}{2\pi i}\left(\int_{\kappa - iT}^{\kappa - K - iT} + \int_{\kappa - K - iT}^{\kappa - K + iT} + \int_{\kappa - K + iT}^{\kappa + iT}\right)\frac{y^s}{s(s+1)}ds \\ &\ll \frac{1}{T^2}\int_{\kappa - K}^{\kappa} y^\sigma d\sigma + \frac{y^{\kappa - K}}{|\kappa - K|^2}\int_{\kappa - K - iT}^{\kappa - K + iT}|ds| \ll \frac{y^\kappa}{T^2|\log y|} + T\frac{y^{\kappa - K}}{|\kappa - K|^2}.\end{aligned}$$

We deduce the second assertion of the lemma by letting K tend to infinity. The proof is similar in the case $0 < y < 1$.

The last case remaining is for $y = 1$. We have

$$\begin{aligned}\frac{1}{2\pi i}\int_{\kappa - iT}^{\kappa + iT} \frac{ds}{s(s+1)} &= \frac{1}{2\pi i}\log\left(\frac{1 + \frac{1}{\kappa - iT}}{1 + \frac{1}{\kappa + iT}}\right) \\ &= \frac{1}{2\pi i}\log\left(1 + O\left(\frac{1}{T}\right)\right) \\ &= O\left(\frac{1}{T}\right),\end{aligned}$$

which concludes the proof.

□

The following is a crucial lemma estimating a weighted sum of the reciprocal of the totient function.

Lemma 3.6.8. *Let* $a \neq 0$ *be an integer and* $M \geq 1$ *be a real number.*

If $\omega(a_M) \geq 1$,

$$\sum_{\substack{n\leq M\\(n,a)=1}} \frac{1}{\phi(n)}\left(1-\frac{n}{M}\right) = C_1(a_M)\log M + C_3(a_M) + \frac{\phi(a_M)}{a_M}\frac{\Lambda(a_M)}{2M} + E(M,a). \tag{3.6.12}$$

If $a_M = 1$,

$$\sum_{\substack{n\leq M\\(n,a)=1}} \frac{1}{\phi(n)}\left(1-\frac{n}{M}\right) = C_1(1)\log M + C_3(1) + \frac{1}{2}\frac{\log M}{M} + \frac{C_5}{M} + E(M,a). \tag{3.6.13}$$

There exists $\delta > 0$ *such that the error term* $E(M,a)$ *satisfies*

$$E(M,a) \ll_\epsilon \frac{\prod_{p|a_M}\left(1+\frac{1}{p^\delta}\right)}{M}\left(\frac{a_M}{M}\right)^{\frac{205}{538}-\epsilon}. \tag{3.6.14}$$

Remark 3.6.2. *Under Lindelöf's hypothesis,*

$$E(M,a) \ll_\epsilon \frac{\prod_{p|a_M}\left(1+\frac{1}{p^\delta}\right)}{M}\left(\frac{a_M}{M}\right)^{1/2-\epsilon}. \tag{3.6.15}$$

PROOF. Note first that we need only to consider the prime factors of a less or equal to M, since for $1 \leq n \leq M$, $(n,a) = 1 \Leftrightarrow (n,a_M) = 1$.

To calculate our sum we will write it as a contour integral and shift contours, showing that the contribution of the shifted contours is negligible and obtaining the main terms from the residues at the poles.

Setting $\kappa = \frac{1}{\log M}$ in Lemma 3.6.7,

$$\begin{aligned}\sum_{\substack{n\leq M\\(n,a)=1}} \frac{1}{\phi(n)}\left(1-\frac{n}{M}\right) &= \sum_{(n,a_M)=1} \frac{1}{\phi(n)} h\left(\frac{M}{n}\right)\\ &= \sum_{(n,a_M)=1} \frac{1}{\phi(n)}\frac{1}{2\pi i}\int_{\kappa-iT}^{\kappa+iT}\left(\frac{M}{n}\right)^s \frac{ds}{s(s+1)}\\ &\qquad + O\left(\frac{1}{T^2}\sum_{n\neq M}\frac{1}{\phi(n)|\log M/n|}\left(\frac{M}{n}\right)^\kappa + \frac{\log M}{TM}\right)\\ &= \frac{1}{2\pi i}\int_{\kappa-iT}^{\kappa+iT}\left(\sum_{(n,a_M)=1}\frac{1}{n^s\phi(n)}\right)\frac{M^s}{s(s+1)}ds + O_M\left(\frac{1}{T}\right).\end{aligned} \tag{3.6.16}$$

In the last step we used the elementary estimates

$$\sum_{n\leq M}\frac{1}{\phi(n)}\ll\log M\quad\text{and}\quad\sum_{n>M}\frac{1}{n^{\kappa}\phi(n)}\ll\log M.$$

Now taking Euler products we compute that

$$\sum_{(n,a_M)=1}\frac{1}{n^s\phi(n)}=\mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1)\tag{3.6.17}$$

where

$$\mathfrak{S}_{a_M}(s+1):=\prod_{\substack{p|a\\ p\leq M}}\left(1-\frac{1}{p^{s+1}}\right)\left(1+\frac{1}{(p-1)p^{s+1}}\right)^{-1}\tag{3.6.18}$$

and

$$Z_2(s+1):=\prod_p\left(1+\frac{1}{p(p-1)}\left(\frac{1}{p^{s+1}}-\frac{1}{p^{2s+2}}\right)\right),\tag{3.6.19}$$

which converges for $\operatorname{Re}s>-3/2$. Therefore, (3.6.16) becomes

$$\sum_{\substack{n\leq M\\(n,a)=1}}\frac{1}{\phi(n)}\left(1-\frac{n}{M}\right)=\frac{1}{2\pi i}\int_{\kappa-iT}^{\kappa+iT}\mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1)\frac{M^s}{s(s+1)}ds$$
$$+O_M\left(\frac{1}{T}\right).\tag{3.6.20}$$

The different results for different values of $\omega(a_M)$ come from the pole at $s=-1$. We see that $\mathfrak{S}_{a_M}(s+1)$ has a zero of order $\omega(a_M)$ at $s=-1$ whereas

$$\frac{\zeta(s+1)\zeta(s+2)}{s(s+1)}$$

has a pole of order two at $s=-1$. Hence the product has no pole if $\omega(a_M)\geq 2$, a pole of order one if $\omega(a_M)=1$, and a pole of order two if $\omega(a_M)=0$. We now shift the contour of integration to the left until the line $\operatorname{Re}(s)=\sigma$, where $-1-\frac{1}{2+4\theta}<\sigma<-1$ and $\theta:=\frac{32}{205}$. The right hand side of (3.6.20) becomes

$$=P_T+\frac{1}{2\pi i}\int_{\sigma-iT}^{\sigma+iT}\mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1)\frac{M^s}{s(s+1)}ds$$
$$+O_M\left(\frac{1}{T}\right)+O_a\left(\frac{\log^2 T}{T^2}\left(T^{1/6}+\frac{T^{1/2}}{M^{1/2}}+\frac{T^{7/6}}{M}\right)\right).\tag{3.6.21}$$

Here, P_T denotes the sum of all residues in the box $\sigma \leq \operatorname{Re} s \leq \frac{1}{\log M}$ and $|\operatorname{Im} s| \leq T$. The second error term in (3.6.21) comes from the horizontal integrals which we have bounded using Lemma 3.6.6 (note that $\theta < 1/6$). Taking $T \to \infty$ yields

$$\sum_{\substack{n \leq M \\ (n,a)=1}} \frac{1}{\phi(n)} \left(1 - \frac{n}{M}\right) = P_\infty + E(M, a), \tag{3.6.22}$$

where

$$E(M, a) := \frac{1}{2\pi i} \int_{(\sigma)} \mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1)\frac{M^s}{s(s+1)} ds.$$

Now on the line $\operatorname{Re} s = \sigma$ we have the bound (note that $0 < -1 - \sigma < \frac{1}{2+4\theta}$)

$$\mathfrak{S}_{a_M}(s+1) \ll a_M^{-1-\sigma} \prod_{p | a_M} \left(1 + \frac{1}{p^\delta}\right),$$

for some $\delta > 0$. Combining this with Lemma 3.6.6 yields

$$\begin{aligned} E(M, a) &\ll_\sigma \prod_{p|a_M} \left(1 + \frac{1}{p^\delta}\right) a_M^{-1-\sigma} \int_{-\infty}^{\infty} |\zeta(\sigma + 1 + it)||\zeta(\sigma + 2 + it)| \frac{M^\sigma}{(|t| + 1)^2} dt \\ &\ll \frac{\prod_{p|a_M} \left(1 + \frac{1}{p^\delta}\right)}{M} \left(\frac{a_M}{M}\right)^{-1-\sigma} \int_{-\infty}^{\infty} \frac{(|t| + 1)^{1/2-(\sigma+1)}(|t| + 1)^{2\theta(1-(\sigma+2))}}{(|t| + 1)^2} dt \\ &\ll_\sigma \frac{\prod_{p|a_M} \left(1 + \frac{1}{p^\delta}\right)}{M} \left(\frac{a_M}{M}\right)^{-1-\sigma}, \end{aligned} \tag{3.6.23}$$

since $1/2 - (\sigma + 1) + 2\theta(1 - (\sigma + 2)) < 1$ by our choice of σ. The claimed bound on $E(M, a)$ then follows by taking $\sigma := -1 - \frac{1}{2+4\theta} + \epsilon$ in (3.6.23).

It remains to compute P_∞ which is the sum of the residues of $\mathfrak{S}_{a_M}(s + 1)\zeta(s + 1)\zeta(s + 2)Z_2(s + 1)\frac{M^s}{s(s+1)}$ in the region $\sigma \leq \operatorname{Re} s \leq \frac{1}{\log M}$. Note that $\mathfrak{S}_{a_M}(s + 1)$ has poles on the lines $\operatorname{Re} s = -1 - \frac{\log(p-1)}{\log p}$, however these poles are cancelled by the zeros of $Z_2(s + 1)$. Thus the only possible singularities of $\mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1)\frac{M^s}{s(s+1)}$ in the region in question are at the points $s = 0$ and $s = -1$. Now a lengthy but straightfoward computation shows that we have a double pole at $s = 0$ with residue equal to $C_1(a_M) \log M + C_3(a_M)$. As for $s = -1$, we have to consider three cases.

If $\omega(a_M) \geq 2$, then $\mathfrak{S}_{a_M}(s+1) = O((s+1)^2)$ around $s=-1$, so $\mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1)\frac{M^s}{s(s+1)}$ is holomorphic and we don't have any residue.

If $\omega(a_M) = 1$, then $\mathfrak{S}_{a_M}(s+1)$ has a simple zero at $s=-1$ and thus $\mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1)\frac{M^s}{s(s+1)}$ has a simple pole with residue equal to $\frac{\phi(a_M)}{a_M}\frac{\Lambda(a_M)}{2M}$.

Finally, if $a_M = 1$, then $\mathfrak{S}_{a_M}(s) \equiv 1$ and thus $\mathfrak{S}_{a_M}(s+1)\zeta(s+1)\zeta(s+2)Z_2(s+1)\frac{M^s}{s(s+1)}$ has a double pole at $s=-1$ with residue equal to $\frac{1}{2}\frac{\log M}{M} + \frac{C_5}{M}$.

□

3.7. Further results and proofs

We will start by giving the fundamental result of this paper which works for M fixed as well as for M varying with x under the condition $M \leq (\log x)^{O(1)}$.

Proposition 3.7.1. *Fix $A > B > 0$ and $\lambda < \frac{1}{4}$, three positive real numbers. Let $M = M(x)$ be an integer such that $1 \leq M(x) \leq \log^B x$. For a in the range $0 < |a| \leq x^\lambda$ we have that*

$$\sum_{\substack{q \leq \frac{x}{M} \\ (q,a)=1}} \left(\psi(x;q,a) - \Lambda(a) - \frac{\psi(x)}{\phi(q)}\right) = x\Bigg(C_1(a)\log M + C_3(a) - \sum_{\substack{r \leq M \\ (r,a)=1}} \frac{1}{\phi(r)}\left(1 - \frac{r}{M}\right)\Bigg) + O_{A,B,\lambda}\left(\frac{\phi(a)}{a}2^{\omega(a)}\frac{x}{\log^A x}\right). \quad (3.7.1)$$

We can remove the condition of M being an integer at the cost of adding the error term $O(x\log\log M/M^2)$.

Proof. We will prove that

$$\sum_{\substack{q \leq \frac{x}{M} \\ (q,a)=1}} \left(\psi(x;q,a) - \Lambda(a) - \frac{x}{\phi(q)}\right) = x\Bigg(C_1(a)\log M + C_3(a) - \sum_{\substack{r \leq M \\ (r,a)=1}} \frac{1}{\phi(r)}\left(1 - \frac{r}{M}\right)\Bigg) + O_{A,B,\lambda}\left(\frac{\phi(a)}{a}2^{\omega(a)}\frac{x}{\log^A x}\right). \quad (3.7.2)$$

From this we can deduce the proposition since by Lemma 3.6.5 the difference between the left hand side of (3.7.1) and that of (3.7.2) is negligible.

Define $L := \log^{A+3} x$. Partitioning the sum into dyadic intervals and applying Theorem 3.2.1 gives

$$\sum_{\substack{q \le \frac{x}{L} \\ (q,a)=1}} \left(\psi(x; q, a) - \Lambda(a) - \frac{x}{\phi(q)} \right) = O_{A,\lambda} \left(2^{\omega(a)} \frac{x}{\log^{A+1} x} \right). \tag{3.7.3}$$

Therefore, we need to compute

$$\sum_{\substack{\frac{x}{L} < q \le \frac{x}{M} \\ (q,a)=1}} \left(\psi(x; q, a) - \Lambda(a) - \frac{x}{\phi(q)} \right),$$

which by lemmas 3.6.3 and 3.6.4 is equal to

$$\sum_{\substack{\frac{x}{L} < q \le \frac{x}{M} \\ (q,a)=1}} \left(\theta^*(x; q, a) - \frac{x}{\phi(q)} \right) + O(x^{2/3} + |a| \log^2 |a|). \tag{3.7.4}$$

We split the sum in (3.7.4) in three distinct sums as following :

$$\sum_{\substack{\frac{x}{L} < q \le x \\ (q,a)=1}} \theta^*(x; q, a) - \sum_{\substack{\frac{x}{M} < q \le x \\ (q,a)=1}} \theta^*(x; q, a) - x \sum_{\substack{\frac{x}{L} < q \le \frac{x}{M} \\ (q,a)=1}} \frac{1}{\phi(q)} = \mathrm{I} - \mathrm{II} - \mathrm{III}.$$

The third sum is easily treated using Lemma 3.6.2 :

$$\begin{aligned} \mathrm{III} = x \sum_{\substack{\frac{x}{L} < q \le \frac{x}{M} \\ (q,a)=1}} \frac{1}{\phi(q)} = x \Bigg(& C_1(a) \log(x/M) + C_1(a) + C_3(a) \\ & + O\left(2^{\omega(a)} \frac{\log(x/M)}{x/M} \right) \\ - \Bigg(& C_1(a) \log(x/L) + C_1(a) + C_3(a) + O\left(2^{\omega(a)} \frac{\log(x/L)}{x/L} \right) \Bigg) \Bigg) \\ & = C_1(a) x \log(L/M) + O\left(2^{\omega(a)} L \log x \right). \end{aligned} \tag{3.7.5}$$

For the first sum, we have

$$\mathrm{I} = \sum_{\substack{\frac{x}{L} < q \le x \\ (q,a)=1}} \theta^*(x; q, a) = \sum_{\substack{\frac{x}{L} < q \le x \\ (q,a)=1}} \sum_{\substack{|a| < p \le x \\ p \equiv a \bmod q}} \log p.$$

Using Lemma 3.6.1,

$$\begin{aligned} I &= \sum_{\substack{1\le r<(x-a)\frac{L}{x}\\ (r,a)=1}} \sum_{\substack{r\frac{x}{L}+a<p\le x\\ p\equiv a \bmod r\\ p>|a|}} \log p + O(|a|\log x) \\ &= \sum_{\substack{1\le r<(x-a)\frac{L}{x}\\ (r,a)=1}} \left(\theta^*(x;r,a) - \theta^*\left(r\frac{x}{L}+a;r,a\right)\right) + O(x^{1/3}) \\ &= \sum_{\substack{1\le r<L-a\frac{L}{x}\\ (r,a)=1}} \left(\frac{x}{\phi(r)} - \frac{rx}{\phi(r)L}\right) + O(|a|L\log^2 x)) + O_A\left(L\frac{x}{L\log^{A+1}x}\right). \end{aligned}$$

In the last step we used Lemma 3.6.4 (with $|a| < x^{\lambda}$) combined with the Siegel-Walfisz theorem in the form $\theta(x;r,a) - \frac{x}{\phi(r)} \ll_A \frac{x}{L\log^{A+1}x}$ for $r \le 2L$, as well as the estimate $\sum_{r\le R}\frac{|a|}{\phi(r)} \ll |a|\log R$. As $\left|a\frac{L}{x}\right| \le 1$ for x large enough and $\phi(r) \gg r/\log\log r$, this gives

$$\begin{aligned} I &= x\sum_{\substack{r<L\\ (r,a)=1}} \frac{1}{\phi(r)}\left(1-\frac{r}{L}\right) + O_A\left(\frac{x}{\log^{A+1}x}\right) + O\left(\frac{x\log\log L}{L}\right) \\ &= x\sum_{\substack{r<L\\ (r,a)=1}} \frac{1}{\phi(r)}\left(1-\frac{r}{L}\right) + O_A\left(\frac{x}{\log^{A+1}x}\right). \end{aligned}$$

We conclude the evaluation of I by applying Lemma 3.6.2 :

$$I = x\left(C_1(a)\log L + C_3(a) + O_A\left(\frac{2^{\omega(a)}}{\log^{A+1}x}\right)\right). \qquad (3.7.6)$$

Now with a similar computation using lemmas 3.6.1 and 3.6.4 as well as the Siegel-Walfisz theorem in the form $\theta(x;r,a) - \frac{x}{\phi(r)} \ll_{A,B} \frac{x}{M\log^{A+1}x}$ for $r \le 2M$, we show that

$$\begin{aligned}
\mathrm{II} &= \sum_{\substack{1\le r<(x-a)\frac{M}{x}\\ (r,a)=1}} \left(\theta^*(x;r,a)-\theta^*\left(r\frac{x}{M}+a,r,a\right)\right)+O(|a|\log x)\\
&= \sum_{\substack{1\le r<M-a\frac{M}{x}\\ (r,a)=1}} \left(\frac{x}{\phi(r)}-\frac{rx}{\phi(r)M}\right)+O(|a|M\log^2 x)+O_{A,B}\left(M\frac{x}{M\log^{A+1}x}\right)\\
&= x\sum_{\substack{1\le r\le M\\ (r,a)=1}} \frac{1}{\phi(r)}\left(1-\frac{r}{M}\right)+O_{A,B}\left(\frac{x}{\log^{A+1}x}\right).
\end{aligned}$$

In the last step we have used that M is an integer so the term for $r = M$ in the sum is given by $\frac{1}{\phi(M)}\left(1-\frac{M}{M}\right) = 0$. If $M \notin \mathbb{N}$, we have to add the error term $\frac{x}{\phi(\lfloor M\rfloor)}\left(1-\frac{\lfloor M\rfloor}{M}\right) = O(x\log\log M/M^2)$.

We conclude the proof by combining our estimates for I, II and III with the bound $\frac{a}{\phi(a)} \ll \log\log x$. □

PROOF OF PROPOSITION 3.3.1. Follows from Proposition 3.7.1. □

For M not necessarily fixed, we will need to use Lemma 3.6.8.

Theorem 3.7.1. *Let $M = M(x) \le \log^B x$ (not necessarily an integer) where $B > 0$ is a fixed real number. Fix $\lambda < \frac{1}{4}$ a positive real number, and let $a \neq 0$ be an integer such that $|a| \le x^\lambda$.*

If $\omega(a_M) \ge 1$,

$$\frac{1}{\frac{\phi(a)}{a}\frac{x}{M}} \sum_{\substack{q\le \frac{x}{M}\\ (q,a)=1}} \left(\psi(x;q,a)-\Lambda(a)-\frac{\psi(x)}{\phi(q)}\right) = -\frac{1}{2}\Lambda(a_R)+E(M,a). \qquad (3.7.7)$$

If $a_M = 1$,

$$\frac{1}{\frac{\phi(a)}{a}\frac{x}{M}} \sum_{\substack{q\le \frac{x}{M}\\ (q,a)=1}} \left(\psi(x;q,a)-\Lambda(a)-\frac{\psi(x)}{\phi(q)}\right) = -\frac{1}{2}\log M - C_5 + E(M,a). \quad (3.7.8)$$

Under the assumption that $\sum_{\substack{p|a\\ p>M}} \frac{1}{p} \leq 1$*, the error term* $E(M,a)$ *satisfies, for some* $\delta > 0$*,*

$$E(M,a) \ll_{a,\epsilon,B} \frac{a}{\phi(a)} \prod_{p|a_M} \left(1+\frac{1}{p^{\delta}}\right)\left(\frac{a_M}{M}\right)^{\frac{205}{538}-\epsilon} + M\log M \sum_{\substack{p|a\\ p>M}} \frac{\log p}{p} + \frac{2^{\omega(a)}}{\log^{B+1} x} + \frac{a}{\phi(a)}\frac{\log\log M}{M}. \quad (3.7.9)$$

Remark 3.7.1. *If we assume Lindelöf's hypothesis, we can replace the first term of the right hand side of* (3.7.9) *by* $\frac{a}{\phi(a)}\prod_{p|a_M}\left(1+\frac{1}{p^{\delta}}\right)\left(\frac{a_M}{M}\right)^{1/2-\epsilon}$.

PROOF OF THEOREM 3.7.1. Our starting point will be to set $A := 2B+1$ in Proposition 3.7.1 :

$$\sum_{\substack{q\leq \frac{x}{M}\\ (q,a)=1}} \left(\psi(x;q,a) - \Lambda(a) - \frac{\psi(x)}{\phi(q)}\right) = x\Bigg(C_1(a)\log M + C_3(a) - \sum_{\substack{r\leq M\\ (r,a)=1}} \frac{1}{\phi(r)}\left(1-\frac{r}{M}\right)\Bigg) + O_B\left(\frac{\phi(a)}{a}\frac{x}{M}\frac{2^{\omega(a)}}{\log^{B+1}x} + x\frac{\log\log M}{M^2}\right). \quad (3.7.10)$$

We now consider three cases depending on the number of prime factors of a_M.

Case 1 : $\omega(a_M) \geq 2$, which implies $\omega(a) \geq 2$. Applying Lemma 3.6.8 gives

$$\sum_{\substack{q\leq \frac{x}{M}\\ (q,a)=1}} \left(\psi(x;q,a) - \Lambda(a) - \frac{\psi(x)}{\phi(q)}\right) \ll_B x\Bigg(|C_1(a_M) - C_1(a)|\log M + |C_3(a_M) - C_3(a)| + \frac{\prod_{p|a_M}\left(1+\frac{1}{p^{\delta}}\right)}{M}\left(\frac{a_M}{M}\right)^{\frac{205}{538}-\epsilon}\Bigg) + \frac{\phi(a)}{a}\frac{x}{M}\frac{2^{\omega(a)}}{\log^{B+1}x} + x\frac{\log\log M}{M^2}. \quad (3.7.11)$$

If all the prime factors of a are less or equal to M, then $|C_i(a_M) - C_i(a)| = 0$. If not, we need upper bounds. By the definition of $C_1(a)$,

$$|C_1(a_M) - C_1(a)| = C_1(a)\left(\prod_{\substack{p|a\\p>M}}\left(1-\frac{1}{p}\right)^{-1}\left(1-\frac{1}{p^2-p+1}\right)^{-1} - 1\right)$$
$$\ll \frac{\phi(a)}{a}\sum_{\substack{p|a\\p>M}}\frac{1}{p},$$

as long as $\sum_{\substack{p|a\\p>M}} \frac{1}{p} \leq 1$. Moreover,

$$C_3(a_M) - C_3(a) = \left(C_1(a) + O\Big(\frac{\phi(a)}{a}\sum_{\substack{p|a\\p>M}}\frac{1}{p}\Big)\right)\frac{C_3(a_M)}{C_1(a_M)} - C_3(a)$$
$$= C_1(a)\left(\frac{C_3(a_M)}{C_1(a_M)} - \frac{C_3(a)}{C_1(a)}\right) + O\left(\frac{\phi(a)}{a}\sum_{\substack{p|a\\p>M}}\frac{\log M}{p}\right),$$

since $\sum_{p|a_M}\frac{p^2\log p}{(p-1)(p^2-p+1)} \ll \sum_{p\leq M}\frac{\log p}{p} \ll \log M$. Thus,

$$|C_3(a_M) - C_3(a)| \ll \frac{\phi(a)}{a}\sum_{\substack{p|a\\p>M}}\frac{p^2\log p}{(p-1)(p^2-p+1)} + \frac{\phi(a)}{a}\sum_{\substack{p|a\\p>M}}\frac{\log M}{p}$$
$$\ll \frac{\phi(a)}{a}\sum_{\substack{p|a\\p>M}}\frac{\log p}{p}.$$

Putting all this together and dividing by $\frac{x}{M}\frac{\phi(a)}{a}$ gives the claimed estimate.

Case 2 : $\omega(a_M) = 1$. The calculation is similar, however Lemma 3.6.8 gives a contribution of

$$-\frac{1}{2}\frac{a}{\phi(a)}\frac{\phi(a_M)}{a_M}\Lambda(a_M) = -\frac{1}{2}\Lambda(a_M) + O\left(\log M\sum_{\substack{p|a\\p>M}}\frac{1}{p}\right),$$

since $\Lambda(a_M) \leq \log M$.

Case 3 : $a_M = 1$. The contribution of

$$\frac{a}{\phi(a)}\left(-\frac{1}{2}\log M - C_5\right) = -\frac{1}{2}\log M - C_5 + O\left(\log M\sum_{\substack{p|a\\p>M}}\frac{1}{p}\right)$$

comes from Lemma 3.6.8.

□

PROOF OF THEOREM 3.1.1. It is a particular case of Theorem 3.7.1. The theorem is trivial if M is bounded. For M tending to infinity with x, we have that $a_M = a$ for x large enough, since a is fixed. Hence, the error term in Theorem 3.7.1 satisfies $E(M, a) \ll_{a,\epsilon,B} M^{-\frac{205}{538}+\epsilon}$. □

The final step is to count for how many integers a the error term in Theorem 3.7.1 is small.

Lemma 3.7.1. *Fix* B, $\lambda < \frac{1}{4}$, $\eta \leq \frac{1}{5}$ *and* δ, *four positive real numbers. Set* $X := x^\lambda$ *and let*

$$S(X) := \left\{ 1 \leq a \leq X : \omega(a_M) \geq 2, E(M, a) \ll_{B,\lambda} \frac{1}{M^\eta} \right\},$$

where $E(M, a)$ *is defined as in Theorem 3.7.1 and* $M \leq (\log x)^B$. *We have that*

$$|S(X)| \geq X \left(\frac{1 - \frac{21}{8}\eta}{1 + \eta + \eta^2} e^{-\gamma} + o(1) \right),$$

where γ *is the Euler-Mascheroni constant.*

PROOF. We first define the following sets :

$$S_1 := \left\{ 1 \leq a \leq X : \frac{a}{\phi(a)} \leq \log M \right\},$$

$$S_2 := \left\{ 1 \leq a \leq X : \prod_{p \mid a_M} \left(1 + \frac{1}{p^\delta} \right) \leq \log M \right\},$$

$$S_3 := \{1 \leq a \leq X : 2^{\omega(a)} \leq \log x\},$$

$$S_4 := \left\{ 1 \leq a \leq X : \sum_{\substack{p \mid a \\ p > M}} \frac{\log p}{p} \leq \frac{1}{M^{1+\eta} \log M}, \quad \prod_{\substack{p \mid a \\ p \leq M}} p \leq M^{1-\frac{21}{8}\eta}, \quad \omega(a_M) \geq 2 \right\}.$$

Note that if $a \in \cap_{i=1}^4 S_i$, then $E(M, a) \ll M^{-\eta}$, so

$$\bigcap_{i=1}^{4} S_i \subset S(X).$$

We now give a lower bound on the size of each of the S_i. We have that

$$\begin{aligned}
|S_1| &\geq X + O(1) - \left\{ 1 \leq a \leq X : \sum_{p|a} \frac{1}{p} \gg \log\log M \right\} \\
&\geq X + O(1) - \sum_{a \leq X} \frac{C}{\log\log M} \sum_{p|a} \frac{1}{p} \\
&= X + O(1) - \frac{C}{\log\log M} \sum_{p \leq x} \frac{1}{p} \left\lfloor \frac{X}{p} \right\rfloor \\
&= X\left(1 - O\left(\frac{1}{\log\log M}\right)\right).
\end{aligned}$$

Similarly, we get that

$$|S_2| \geq X\left(1 - O\left(\frac{1}{\delta \log\log M}\right)\right).$$

By the Hardy-Ramanujan Theorem (see [**37**]), we have that

$$|S_3| = X(1 - o(1)).$$

To estimate the size of S_4, we will need to use the theory of y-rough numbers, that is numbers whose prime factors are all greater or equal to y. Define the set T_1 consisting of the numbers of the form dr such that d is a an integer with at least two distinct prime factors in the interval $1 \leq d \leq M^{1-\frac{21}{8}\eta}$, and r is a $M^{1+\eta+\eta^2}$-rough integer in the interval $M^{1+\eta} \leq r \leq \frac{X}{d}$. Define also

$$\begin{aligned}
T_2 := &\left\{ a \in T_1 : \sum_{\substack{p|a \\ p > M^{1+\eta+\eta^2}}} \frac{\log p}{p} \leq \frac{1}{M^{1+\eta}\log M} \right\} \\
&= \left\{ a \in T_1 : \sum_{\substack{p|a \\ p > M}} \frac{\log p}{p} \leq \frac{1}{M^{1+\eta}\log M} \right\},
\end{aligned}$$

so $T_2 \subset S_4$. We have that

$$\begin{aligned}
|T_2| \geq |T_1| - &\sum_{\substack{d \leq M^{1-\frac{21}{8}\eta} \\ \omega(d) \geq 2}} \sum_{\substack{r \leq \frac{X}{d} \\ p'|r \Rightarrow p' \geq M^{1+\eta+\eta^2}}} M^{1+\eta+\frac{\eta^2}{2}} \sum_{p|r} \frac{\log p}{p} \\
&= |T_1| - M^{1+\eta+\frac{\eta^2}{2}} \sum_{\substack{d \leq M^{1-\frac{21}{8}\eta} \\ \omega(d) \geq 2}} \sum_{M^{1+\eta+\eta^2} \leq p \leq \frac{X}{d}} \frac{\log p}{p} \Phi\left(\frac{X}{dp}, M^{1+\eta+\eta^2}\right) \\
&= |T_1| - M^{1+\eta+\frac{\eta^2}{2}} \sum_{\substack{d \leq M^{1-\frac{21}{8}\eta} \\ \omega(d) \geq 2}} \sum_{M^{1+\eta+\eta^2} \leq p \leq X^{\frac{1}{2}}} \frac{\log p}{p} \Phi\left(\frac{X}{dp}, M^{1+\eta+\eta^2}\right) + O(X^{\frac{2}{3}}),
\end{aligned}$$

where $\Phi(x, y)$ is the number of y-rough integers up to x. Equation (1.13) of **[18]** shows that in our range of d and for x large enough we have

$$\Phi\left(\frac{X}{dp}, M^{1+\eta+\eta^2}\right) = \frac{X}{dp} \prod_{p \leq M^{1+\eta+\eta^2}} \left(1 - \frac{1}{p}\right)\left(1 + O_B\left(\exp\left(-\frac{\lambda}{3} M^{\frac{1}{2B}}\right)\right)\right),$$

so we get that

$$|T_1| - |T_2| \ll XM^{1+\eta+\frac{\eta^2}{2}} \log M \sum_{d \leq M^{1-\frac{21}{8}\eta}} \frac{1}{d} \sum_{M^{1+\eta+\eta^2} \leq p \leq X^{\frac{1}{2}}} \frac{\log p}{p^2} + O(X^{\frac{2}{3}}) \ll XM^{-\frac{\eta^2}{3}}.$$

We use the same ideas to estimate the number of elements of T_1 :

$$\begin{aligned}
|T_1| &= \sum_{\substack{d \leq M^{1-\frac{21}{8}\eta} \\ \omega(d) \geq 2}} \Phi\left(\frac{X}{d}, M^{1+\eta+\eta^2}\right) \\
&= X \sum_{\substack{d \leq M^{1-\frac{21}{8}\eta} \\ \omega(d) \geq 2}} \frac{1}{d} \prod_{p \leq M^{1+\eta+\eta^2}} \left(1 - \frac{1}{p}\right)\left(1 + O_B\left(\exp\left(-\frac{\lambda}{3} M^{\frac{1}{2B}}\right)\right)\right) \\
&= X\left(\frac{1 - \frac{21}{8}\eta}{1 + \eta + \eta^2} e^{-\gamma} + o(1)\right), \qquad (3.7.12)
\end{aligned}$$

by Merten's Theorem. Therefore, the estimate (3.7.12) also holds for $|T_2|$, and hence

$$|S_4| \geq X\left(\frac{1 - \frac{21}{8}\eta}{1 + \eta + \eta^2} e^{-\gamma} + o(1)\right).$$

We finish our proof by applying repeatedly, for $U, V \subset \mathbb{N} \cap [1, X]$, the following inequality :

$$\begin{aligned} |U(X) \cap V(X)| &= |U(X)| + |V(X)| - |U(X) \cup V(X)| \\ &\geq |U(X)| + |V(X)| - X. \end{aligned}$$

□

PROOF OF THEOREM 3.3.1. This is a consequence of Theorem 3.7.1 and Lemma 3.7.1. □

3.8. CONCLUDING REMARKS

Remark 3.8.1. *One could ask if the results of Theorem 3.1.1 are intrinsic to the sequence of prime numbers or if they are just a result of the weight $\Lambda(n)$ in the prime counting functions. However one can see that if we replace $\psi(x; q, a)$ by $\pi(x; q, a)$ and $\psi(x)$ by $\pi(x)$, the proof of Proposition 3.7.1 will go through with x replaced by $\mathrm{Li}(x)$ and an additional error term of*

$$O\left(x \frac{\log\log x}{\log^2 x}\right).$$

(This is because in the evaluation of II in the proof of Proposition 3.7.1, we have for $\frac{1}{\log x} \leq y \leq 1 + o(1)$ that $\mathrm{Li}(yx) - y\mathrm{Li}(x) \ll \frac{|\log y|}{(\log x)^2}$, a bound which cannot be improved for small values of y.) One has to prove the analogue of Theorem 3.2.1 which can be done using the Bombieri-Vinogradov theorem and a very delicate summation by parts. We conclude that an analogue of Theorem 3.1.1 holds with the natural prime counting functions in the range $M \leq \sqrt{\log x}$.

Theorem 3.8.1. *Fix an integer $a \neq 0$ and let $M = M(x) \leq \sqrt{\log x}$. We have*

$$\frac{1}{\frac{\phi(a)}{a}\frac{\mathrm{Li}(x)}{M}} \sum_{\substack{q \leq \frac{x}{M} \\ (q,a)=1}} \left(\pi(x; q, a) - \frac{\vartheta(a)}{\log a} - \frac{\pi(x)}{\phi(q)}\right) = \mu(a, M) + O_{a,\epsilon}\left(\frac{1}{M^{\frac{205}{538}-\epsilon}}\right), \tag{3.8.1}$$

where $\mu(a, M)$ is defined as in Theorem 3.1.1.

Remark 3.8.2. *Below we sketch an argument showing that the proportion of integers* $a \leq x^{\lambda}$ *for which the first term of the right hand side of* (3.7.9) *is* ≤ 1 *is not more than* $e^{-\gamma}$. *Setting* $X := x^{\lambda}$ *and using that* $M \ll \log X$, *we compute*

$$\#\Big\{a \leq X : \prod_{\substack{p|a \\ p\leq M}} p \leq M\Big\} = \sum_{\substack{n\leq M \\ \mu^2(n)=1}} \#\Big\{a \leq X : n \mid a, \Big(\frac{a}{n}, \prod_{\substack{p\leq M \\ p\nmid n}} p\Big) = 1\Big\},$$

which by the fundamental lemma of the combinatorial sieve is

$$\sim \sum_{\substack{n\leq M \\ \mu^2(n)=1}} \frac{X}{n} \prod_{\substack{p\leq M \\ p\nmid n}} \left(1 - \frac{1}{p}\right) \sim \frac{Xe^{-\gamma}}{\log M} \sum_{n\leq M} \frac{\mu^2(n)}{\phi(n)} \sim Xe^{-\gamma}.$$

Therefore, to extend the proportion of integers a *for which we get an admissible error term in Theorem 3.3.1, one would need to improve the bound* (3.7.9) *for* $E(M, a)$.

Chapitre 4

THE INFLUENCE OF THE FIRST TERM OF AN ARITHMETIC PROGRESSION

Author : Daniel Fiorilli

Abstract : The goal of this article is to study the discrepancy of the distribution of arithmetic sequences in arithmetic progressions. We will fix a sequence $\mathcal{A} = \{\mathbf{a}(n)\}_{n\geq 1}$ of non-negative real numbers in a certain class of arithmetic sequences. For a fixed integer $a \neq 0$, we will be interested in the behaviour of $\mathcal{A}$ over the arithmetic progressions $a \bmod q$, on average over q. Our main result is that for certain sequences of arithmetic interest, the value of a has a significant influence on this distribution, even after removing the first term of the progressions.

4.1. Introduction

The study of arithmetic sequences is a central problem in number theory. Undoubtedly, it is the sequence of prime numbers which has attracted the most attention amongst number theorists, leading to many theorems and conjectures. Other important sequences include sums of two squares, twin primes, divisor sequences and so on. In general, number theorists are interested in sequences with arithmetical content, and one can formally define wide classes of such sequences. Some phenomena occurring in the theory of prime numbers happen to be true for much wider classes of arithmetic sequences, such as the

Bombieri-Vinogradov theorem for example (see [63]). Another example is the Granville-Soundararajan uncertainty principle (see [33]).

We will fix an integer $a \neq 0$ and study the distribution of an arithmetic sequence $\mathcal{A} = \{\mathbf{a}(n)\}_{n\geq 1}$ in the progressions $a \bmod q$, on average over q. Under certain hypotheses, we will show how certain sequences remember the first term, that is how the value of a can influence the distribution of $\mathcal{A}$ in the progressions $a \bmod q$. Examples of such sequences include the sequences of primes, sums of two squares (or more generally values of positive definite binary quadratic forms), prime k-tuples (conditionally) and integers without small prime factors. We will see that in each of these examples, values of a which have the property that $\mathbf{a}(a) > 0$ have a negative influence. More mysteriously, there are other values of a having a negative influence, and it is not clear to me why these come up.

The structure of the paper is as follows. We begin in Section 4.2 by stating our concrete results for each of the arithmetic sequences mentioned earlier, to highlight the phenomena we will describe later on in more generality. In section 4.3, we give a framework to study general arithmetic sequences and state the hypotheses on which our main theorems will depend. These hypotheses will be crucial in the proofs of Section 4.5. Our general results are stated in Section 4.4, and proved in Section 4.5. As we will see in Section 4.6, most of the concrete examples we give satisfy the hypotheses of Section 4.3, but in some cases we need to slightly modify the analysis. We also see in this section exactly which hypotheses are needed for each result.

4.1.1. Acknowledgements

I would like to thank my supervisor Andrew Granville for suggesting this generalization and for his help and advice in general. I would also like to thank my colleagues Farzad Aryan, Mohammad Bardestani, Dimitri Dias, Tristan Freiberg and Kevin Henriot for many fruitful conversations. Ce travail a été rendu possible grâce à des bourses doctorales du Conseil de Recherche en

Sciences Naturelles et en Génie du Canada et de la Faculté des Études Supérieures et Postdoctorales de l'Université de Montréal.

4.2. EXAMPLES

Before we state the general result, let us look at concrete examples. Throughout, $\mathcal{A} = \{\mathbf{a}(n)\}_{n\geq 1}$ will be a fixed sequence of non-negative real numbers and $a \neq 0$ will be a fixed integer, on which every error term can possibly depend. We will adopt the convention that for negative values of a, $\mathbf{a}(a) := 0$ (and similarly for $\Lambda(a)$). Moreover, $M = M(x)$ will denote a function tending to infinity with x, and we will use $\sim$ as shorthand for $\sim_{M\to\infty}$ (similarly for $o(\cdot) = o_{M\to\infty}(\cdot)$). We define the following counting functions.

Definition 4.2.1.

$$\mathcal{A}(x) := \sum_{1\leq n\leq x} \mathbf{a}(n), \qquad \mathcal{A}_d(x) := \sum_{\substack{1\leq n\leq x: \\ d|n}} \mathbf{a}(n), \qquad \mathcal{A}(x; q, a) := \sum_{\substack{1\leq n\leq x \\ n\equiv a \bmod q}} \mathbf{a}(n).$$

4.2.1. Primes

The first example we give, which was studied more precisely in **[23]**, is the sequence of prime numbers.

Theorem 4.2.1. *Let $A > 0$ be a fixed real number. We have for $M = M(x) \leq (\log x)^A$ that*

$$\frac{1}{\frac{\phi(a)}{a}\frac{x}{M}} \sum_{\substack{q\leq \frac{x}{M} \\ (q,a)=1}} \left(\psi(x; q, a) - \Lambda(a) - \frac{\psi(x)}{\phi(q)}\right) \quad \textit{is} \quad \begin{cases} \sim -\frac{1}{2}\log M & \textit{if } a = \pm 1, \\ \sim -\frac{1}{2}\log p & \textit{if } a = \pm p^e, \\ = O\big(M^{-\frac{205}{538}+\epsilon}\big) & \textit{otherwise,} \end{cases}$$

where the constant implied in O depends on a, ϵ and A.

4.2.2. Integers represented by a fixed positive definite binary quadratic form, with multiplicity

The second example we consider is the sequence of integers which can be represented by a fixed positive definite binary quadratic form $Q(x,y)$ with integer coefficients, counted with multiplicity, that is

$$\mathbf{a}(n) := \#\{(x,y) \in \mathbb{Z}^2_{\geq 0} : Q(x,y) = n\}.$$

We will define $r_d(n)$ to be the total number of distinct representations of n by all of the inequivalent forms of discriminant d (which is not to be confused with $\mathbf{a}(n)$). By distinct representations, we mean that we count the representations up to automorphisms of the forms. We also define the function

$$\rho_a(q) := \frac{1}{q} \cdot \#\{1 \leq x, y \leq q : Q(x,y) \equiv a \bmod q\}.$$

Theorem 4.2.2. *Suppose that* $Q(x,y) = \alpha x^2 + \beta xy + \gamma y^2$ *is a fixed positive definite quadratic form (with integer coefficients) of discriminant* $d := \beta^2 - 4\alpha\gamma < 0$, *with* $d \equiv 1, 5, 9, 12, 13 \bmod 16$ *(for simplicity). Fix an integer* a *such that* $(a, 2d) = 1$. *We have for* $M = M(x) \leq x^{\lambda}$, *where* $\lambda < \frac{1}{12}$ *is a fixed real number, that*

$$\frac{1}{x/M} \sum_{q \leq \frac{x}{M}} \left(\mathcal{A}(x; q, a) - \mathbf{a}(a) - \frac{\rho_a(q)}{q} \mathcal{A}(x) \right)$$

$$= -C_Q \rho_a(4d) r_d(|a|) + O\left(\frac{1}{M^{\frac{1}{3}-\epsilon}} \right), \quad (4.2.1)$$

with

$$C_Q := \frac{A_Q}{2L(1, \chi_d)} \qquad \left(= \frac{w_d \sqrt{|d|}}{4\pi h_d} A_Q \right),$$

where A_Q *is the area of the region* $\{(x,y) \in \mathbb{R}^2_{\geq 0} : Q(x,y) \leq 1\}$, $\chi_d := \left(\frac{4d}{\cdot}\right)$, w_d *is the number of units of* $\mathbb{Q}(\sqrt{d})$ *and* h_d *is its class number. The constant implied in* O *depends on* a, ϵ, λ *and* Q.

Remark. *The number* $\rho_a(4d)$ *is either zero or equal to* $2^{\omega(2d)}$, $2^{\omega(2d)-2}$ *or* $3 \cdot 2^{\omega(2d)-2}$, *depending on* $Q(x,y)$ *(*$\omega(n)$ *denotes the number of distinct prime factors of* n*). For this reason, if* $\rho_a(4d) > 0$, *then it is independent of* a.

Therefore, there is no bias if $\rho_a(4d) = 0$ or if $|a|$ cannot be represented by a form of discriminant d. However, if this is not the case, then the bias is proportional to the number of such representations.

4.2.3. Sums of two squares, without multiplicity

The next example is the sequence of integers which can be written as the sum of two squares, without multiplicity. We define

$$\mathbf{a}(n) := \begin{cases} 1 & \text{if } n = \square + \square, \\ 0 & \text{else.} \end{cases}$$

For a fixed odd integer a, we define the multiplicative function $\mathbf{g}_a(q)$ on prime powers as follows. For $p \neq 2$ such that $p^f \parallel a$ with $f \geq 0$,

$$\mathbf{g}_a(p^e) := \frac{1}{p^e} \times \begin{cases} 1 & \text{if } p \equiv 1 \bmod 4 \\ 1 & \text{if } p \equiv 3 \bmod 4, e \leq f, 2 \mid e \\ \frac{1}{p} & \text{if } p \equiv 3 \bmod 4, e \leq f, 2 \nmid e \\ 1 + \frac{1}{p} & \text{if } p \equiv 3 \bmod 4, e > f, 2 \mid f \\ 0 & \text{if } p \equiv 3 \bmod 4, e > f, 2 \nmid f. \end{cases} \tag{4.2.2}$$

Moreover, $\mathbf{g}_a(2) := \frac{1}{2}$ and for $e \geq 2$, $\mathbf{g}_a(2^e) := \frac{1+(-1)^{\frac{a-1}{2}}}{2^{e+2}}$.

Theorem 4.2.3. *Fix an integer* $a \equiv 1 \bmod 4$. *We have for* $1 \leq M(x) \leq (\log x)^{\lambda}$, *where* $\lambda < 1/5$ *is a fixed real number, that*

$$\frac{1}{x/2M} \sum_{\frac{x}{2M} < q \leq \frac{x}{M}} \left(\mathcal{A}(x; q, a) - \mathbf{a}(a) - \mathbf{g}_a(q)\mathcal{A}(x)\right) \sim -\left(\frac{\log M}{\log x}\right)^{\frac{1}{2}} \frac{(-4)^{-l_a-1}(2l_a+2)!}{(4l_a^2-1)(l_a+1)!\pi} \prod_{\substack{p^f \parallel a: \\ p \equiv 3 \bmod 4, \\ f \text{ odd}}} \frac{\log(p^{\frac{f+1}{2}})}{\log M}, \tag{4.2.3}$$

where $l_a := \#\{p^f \parallel a : p \equiv 3 \bmod 4, 2 \nmid f\}$ *is the number primes dividing* a *to an odd power which are congruent to* 3 *modulo* 4.

Remark. *The right hand side of* (4.2.3) *is* $o((\log x)^{-\frac{1}{2}})$ *iff* $|a|$ *cannot be written as the sum of two squares. Also, if* $|a| = \square + \square$, *then it is equal to* $-\frac{1}{2\pi}\left(\frac{\log M}{\log x}\right)^{\frac{1}{2}}$.

Moreover, one can show that if $a \equiv 3 \bmod 4$, *then the left hand side of* (4.2.3) *is always* $o((\log x)^{-\frac{1}{2}})$.

4.2.4. Prime k-tuples

The next example concerns prime k-tuples. Let $\mathcal{H} = \{\mathcal{L}_1, ..., \mathcal{L}_k\}$ be a k-tuple of distinct linear forms $\mathcal{L}_i(n) = a_i n + b_i$, with $a_i, b_i \in \mathbb{Z}$, $a_i \geq 1$, and define

$$\mathcal{P}(n; \mathcal{H}) := \prod_{\mathcal{L} \in \mathcal{H}} \mathcal{L}(n).$$

We will suppose that $\mathcal{H}$ is admissible, that is for every prime p,

$$\nu_{\mathcal{H}}(p) := \#\{x \bmod p : \mathcal{P}(x; \mathcal{H}) \equiv 0 \bmod p\} < p.$$

Define

$$\mathbf{a}(n) := \prod_{\mathcal{L} \in \mathcal{H}} \Lambda(\mathcal{L}(n)) = \Lambda(a_1 n + b_1)\Lambda(a_2 n + b_2) \cdots \Lambda(a_k n + b_k).$$

The singular series associated to $\mathcal{H}$ is

$$\mathfrak{S}(\mathcal{H}) := \prod_p \left(1 - \frac{\nu_{\mathcal{H}}(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}.$$

Note that if $(\mathcal{P}(a; \mathcal{H}), q) > 1$, then $\mathcal{A}(x; q, a)$ is bounded. Fix $\delta > 0$. The Hardy-Littlewood conjecture stipulates that there exists a function $\mathbf{L}(x)$ tending to infinity with x such that if $(\mathcal{P}(a; \mathcal{H}), q) = 1$,

$$\mathcal{A}(x) = \mathfrak{S}(\mathcal{H})x + O\left(\frac{x}{\mathbf{L}(x)^{2+2\delta}}\right). \tag{4.2.4}$$

Define

$$\gamma(q) := \prod_{p \mid q} \left(1 - \frac{\nu_{\mathcal{H}}(p)}{p}\right).$$

Theorem 4.2.4. *Assume that* (4.2.4) *holds uniformly for all admissible* k*-tuples* $\tilde{\mathcal{H}}$ *such that* $|a_i| \leq \mathbf{L}(x)^{1+\delta}$ *and* $|b_i| = O(1)$. *Fix a* k*-tuple* $\mathcal{H} = \{\mathcal{L}_1, ..., \mathcal{L}_k\}$. *We have for*

$M = M(x) \leq \mathbf{L}(x)$ *that the average*

$$\frac{1}{\frac{\phi(\mathcal{P}(a;\mathcal{H}))}{\mathcal{P}(a;\mathcal{H})}\frac{x}{2M}} \sum_{\substack{\frac{x}{2M}<q\leq\frac{x}{M}: \\ (q,\mathcal{P}(a;\mathcal{H}))=1}} \left(\mathcal{A}(x;q,a) - \mathbf{a}(a) - \frac{\mathcal{A}(x)}{q\gamma(q)}\right) \textit{ is}$$

$$\begin{cases} \sim -\dfrac{(\log M)^{k-\omega(P(a;\mathcal{H}))}}{2(k-\omega(\mathcal{P}(a;\mathcal{H})))!} \displaystyle\prod_{p|\mathcal{P}(a;\mathcal{H})} \frac{p-\nu_{\mathcal{H}}(p)}{p-1}\log p & \textit{if } \omega(\mathcal{P}(a;\mathcal{H})) \leq k, \\ = O\left(\frac{1}{M^{\delta_k}}\right) & \textit{otherwise,} \end{cases}$$

where $\delta_k > 0$ *is a positive real number depending on* k*, and* $\omega(n)$ *denotes the number of distinct prime factors of* n*. The constant implied in* O *depends on* a*,* δ *and* $\mathcal{H}$*.*

In the case of twin primes, we have $\mathcal{H} = \{n, n+2\}$, so $\mathcal{P}(a, \mathcal{H}) = a(a+2)$, and the function $\nu_{\mathcal{H}}$ is given by $\nu_{\mathcal{H}}(2) = 1$ and $\nu_{\mathcal{H}}(p) = 2$ for odd p. We get that the average is

$$\begin{cases} \sim -\frac{(\log M)^2}{4} & \text{if } a = -1 \\ \sim -\frac{\log 3}{4}\log M & \text{if } a = 1, -3 \\ \sim -\frac{\log 2}{2}\log M & \text{if } a = 2, -4 \\ \sim -\frac{\log p \log q}{2}\frac{p-\nu_{\mathcal{H}}(p)}{p-1}\frac{q-\nu_{\mathcal{H}}(q)}{q-1} & \text{if } a(a+2) = \pm p^e q^f \\ O\left(\frac{1}{M^{\delta_2}}\right) & \text{if } \omega(a(a+2)) \geq 3. \end{cases}$$

4.2.5. Integers free of small prime factors

For $y = y(x)$ a function of x, define

$$\mathbf{a}_y(n) := \begin{cases} 1 & \text{if } p \mid n \Rightarrow p \geq y \\ 0 & \text{else,} \end{cases} \qquad \mathcal{A}(x,y) := \sum_{n\leq x} \mathbf{a}_y(n),$$

$$\gamma_y(q) := \prod_{\substack{p|q \\ p<y}} \left(1 - \frac{1}{p}\right), \qquad \mathcal{A}(x,y;q,a) := \sum_{\substack{n\leq x \\ n\equiv a \bmod q}} \mathbf{a}_y(n).$$

Theorem 4.2.5. *Fix* $a \neq 0$*,* $\delta > 0$ *and* $M = M(x) \leq (\log x)^{1-\delta}$*. If*

$$\nu_y(a, M) := \frac{1}{\frac{x}{2M}\frac{\phi(a)}{a}} \sum_{\substack{\frac{x}{2M}<q\leq\frac{x}{M} \\ (q,a)=1}} \left(\mathcal{A}(x,y;q,a) - \mathbf{a}_y(a) - \frac{\mathcal{A}(x,y)}{q\gamma_y(q)}\right),$$

then for $y \leq e^{(\log M)^{\frac{1}{2}-\delta}}$ *with* $y \to \infty$,

$$\nu_y(a, M) = \begin{cases} -\frac{1}{2} + o(1) & \text{if } a = \pm 1 \\ o(1) & \text{otherwise,} \end{cases}$$

and for $(\log x)^{\log\log\log x} \leq y \leq \sqrt{x}$,

$$\nu_y(a, M) = \frac{\mathcal{A}(x,y)}{x} \times \begin{cases} (-\frac{1}{2} + o(1)) \log M & \text{if } a = \pm 1 \\ -\frac{1}{2}\log p + o(1) & \text{if } a = \pm p^k \\ o(1) & \text{otherwise.} \end{cases}$$

(We have no result in the intermediate range.)

Remark. *For* x *large enough,* $\mathbf{a}_y(a) = 0$ *unless* $a = \pm 1$.

4.3. Definitions and Hypotheses

4.3.1. Arithmetic sequences

The goal of this section is to give a framework to study arithmetic sequences. This discussion is modeled on that in **[33]**.

We wish to study the sequence $\mathcal{A} = \{\mathbf{a}(n)\}_{n\geq 1}$ in arithmetic progressions, therefore one of our goals will be to prove the existence of a multiplicative function $\mathbf{g}_a(q)$ such that

$$\mathcal{A}(x; q, a) \sim \mathbf{g}_a(q)\mathcal{A}(x),$$

whenever $\mathbf{g}_a(q) \neq 0$. Let us give a heuristic way to do this with the help of an auxiliary multiplicative function $\mathbf{h}(d)$. First, denote by $\mathcal{S}$ a finite set of "bad primes", which are inherent to the sequence $\mathcal{A}$. We will assume that $\mathcal{A}$ is well distributed in the progressions $0 \bmod d$, that is there exists a multiplicative function $\mathbf{h}(d)$ such that for $(d, \mathcal{S}) = 1$,

$$\mathcal{A}_d(x) \approx \frac{\mathbf{h}(d)}{d}\mathcal{A}(x).$$

The fact that $\mathbf{h}(d)$ is multiplicative can be rephrased as "the events that $\mathbf{a}(n)$ is divisible by coprime integers are independent". Let us also assume that

$$\mathcal{A}(x; q, a) \approx \frac{1}{\phi(q/(q,a))} \sum_{\substack{n \le x: \\ (q,n)=(q,a)}} \mathbf{a}(n),$$

that is the sum is equally partitioned amongst the $\phi(q/(q,a))$ arithmetic progressions $b \bmod q$ with $(b,q) = (a,q)$. We then compute

$$\begin{aligned}\mathcal{A}(x; q, a) &\approx \frac{1}{\phi(q/(q,a))} \sum_{\substack{n \le x \\ (q,n)=(q,a)}} \mathbf{a}(n) = \frac{1}{\phi(q/(q,a))} \sum_{d | \frac{q}{(q,a)}} \mu(d) \mathcal{A}_{(q,a)d}(x) \\ &\approx \mathcal{A}(x) \frac{1}{\phi(q/(q,a))} \sum_{d | \frac{q}{(q,a)}} \mu(d) \frac{\mathbf{h}((q,a)d)}{(q,a)d} = \mathbf{g}_a(q)\mathcal{A}(x),\end{aligned}$$

where

$$\mathbf{g}_a(q) = \mathbf{g}_{(a,q)}(q) := \frac{1}{\phi(q/(q,a))} \sum_{d | \frac{q}{(q,a)}} \mu(d) \frac{\mathbf{h}((q,a)d)}{(q,a)d}$$

is a multiplicative function of q which depends on (q,a) (rather than depending on a). We have thus expressed the multiplicative function $\mathbf{g}_a(q)$ in terms of $\mathbf{h}(d)$. More explicitly, we have, when $p^f \parallel a$ (with $(pa, \mathcal{S}) = 1$), that

$$\mathbf{g}_a(p^e) = \begin{cases} \frac{\mathbf{h}(p^e)}{p^e} & \text{if } e \le f \\ \frac{1}{\phi(p^e)} \left(\mathbf{h}(p^f) - \frac{\mathbf{h}(p^{f+1})}{p} \right) & \text{if } e > f. \end{cases} \tag{4.3.1}$$

In particular, if $p \nmid a$,

$$\mathbf{g}_a(p^e) = \frac{1}{\phi(p^e)} \left(1 - \frac{\mathbf{h}(p)}{p} \right).$$

Another way to write this is

$$\mathcal{A}(x; q, a) \approx \frac{\mathbf{f}_a(q)}{q\gamma(q)} \mathcal{A}(x), \tag{4.3.2}$$

where

$$\gamma(q) := \frac{\phi(q)}{q} \prod_{p|q} \left(1 - \frac{\mathbf{h}(p)}{p} \right)^{-1} = \prod_{p|q} \frac{1 - 1/p}{1 - \mathbf{h}(p)/p},$$

and $\mathbf{f}_a(q)$ is a multiplicative function defined by $\mathbf{f}_a(q) := \mathbf{g}_a(q) q\gamma(q)$. Note that for $(a,q) = 1$, $\mathbf{f}_a(q) = 1$.

4.3.2. Hypotheses

In the following, $\delta > 0$ will denote a (small) fixed real number which will change from one statement to another. We will also fix an integer $a \neq 0$ with the property that $(a, \mathcal{S}) = 1$, where $\mathcal{S}$ is a finite set of bad primes. The function $\mathbf{L} : [0, \infty) \to [1, \infty)$ will be a given increasing smooth function such that $\mathbf{L}(x) \to \infty$ as $x \to \infty$ (think of $\mathbf{L}(x)$ as a power of $\log x$). We now assume the existence of a multiplicative function $\mathbf{f}_a(q) = \mathbf{f}_{(a,q)}(q)$, depending on (a, q), and of $\gamma(q) \neq 0$, which is independent of a (as in Section 4.3.3), such that for any fixed $a \neq 0$ and $q \geq 1$,

$$\mathcal{A}(x; q, a) \sim \frac{\mathbf{f}_a(q)}{q\gamma(q)} \mathcal{A}(x)$$

whenever $\mathbf{f}_a(q) \neq 0$. To simplify the notation, we will also assume the existence of a multiplicative function $\mathbf{h}(d)$ such that (4.3.1) holds (for $(qa, \mathcal{S}) = 1$).

Hypothesis 4.3.1. *There exists a positive increasing function* $\mathbf{R}(x)$ *(think of* $\mathbf{R}(x)$ *as a small power of* x*), with* $\mathbf{L}(x)^{1+\delta} \leq \mathbf{R}(x) \leq \sqrt{x}$*, such that*

$$\sum_{q \leq 2\mathbf{R}(x)} \max_{y \leq x} \left| \mathcal{A}(y; q, a) - \frac{\mathbf{f}_a(q)}{q\gamma(q)} \mathcal{A}(y) \right| \ll \frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}.$$

We will see later that if we use dyadic intervals, we can replace Hypothesis 4.3.1 by a weaker hypothesis.

Hypothesis 4.3.1*. *We have*

$$\sum_{q \leq 2\mathbf{L}(x)} \max_{y \leq x} \left| \mathcal{A}(y; q, a) - \frac{\mathbf{f}_a(q)}{q\gamma(q)} \mathcal{A}(y) \right| \ll \frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}.$$

Hypothesis 4.3.2. *For any* $z = z(x)$ *in the range* $\frac{1}{\mathbf{L}(x)} \leq z(x) \leq 1 + \frac{|a|}{x}$*, we have*

$$\frac{\mathcal{A}(zx)}{\mathcal{A}(x)} = z + O\left(\frac{1}{\mathbf{L}(x)^{1+\delta}} \right).$$

Moreover, for $n \leq x$*, we have the following bound :*

$$\mathbf{a}(n) \ll \frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}.$$

The next hypothesis is somewhat more specific to our analysis than the ones above, and it will allow us to use the analytic theory of zeta functions.

Hypothesis 4.3.3. *There exists a real number* $\mathbf{k} \geq 0$ *such that the sum*

$$\sum_{p \notin \mathcal{S}} \frac{\mathbf{h}(p) - \mathbf{k}}{p}$$

is convergent. More generally, for any real number t *and integer* $n \geq 1$, *we have*

$$\sum_{\substack{p \leq x \\ p \notin \mathcal{S}}} \frac{\mathbf{h}(p) - \mathbf{k}}{p^{1+it}} \leq (1/2 - \delta) \log(|t| + 2) + O(1),$$

$$\sum_{\substack{p \leq x \\ p \notin \mathcal{S}}} \frac{(\mathbf{h}(p) - \mathbf{k}) \log^n p}{p^{1+it}} \ll_{n,\epsilon} (|t| + 2)^{\epsilon}.$$

Finally, $\mathbf{h}(p) < p$ *and for any* $\epsilon > 0$,

$$\mathbf{h}(d) \ll_{\epsilon} d^{\epsilon}.$$

The final hypothesis will be useful when studying the full interval $1 \leq q \leq \frac{x}{M}$ rather than a dyadic one. It is not known for all the sequences we considered in Section 4.2 ; for this reason we used dyadic intervals in theorems 4.2.3, 4.2.4 and 4.2.5.

Hypothesis 4.3.4. *With the same* $\mathbf{R}(x)$ *as in Hypothesis 4.3.1, we have*

$$\sum_{q \leq \frac{x}{\mathbf{R}(x)}} \left(\mathcal{A}^*(x; q, a) - \frac{\mathbf{f}_a(q)}{q\gamma(q)} \mathcal{A}(x) \right) \ll \frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}},$$

where $\mathcal{A}^*(x; q, a)$ *is defined as in* (4.5.15).

4.3.3. The formula for the average

In this section we give a formula for the "average" $\mu_{\mathbf{k}}(a, M)$ which will appear in theorems 4.4.1 and 4.4.1*. The formula is rather complicated in its general form, however in concrete examples it can be seen that it reflects the nature of the sequence $\mathcal{A}$.

Definition 4.3.1.

$$\omega_{\mathbf{h}}(a) := \#\{p^f \parallel a \textit{ with } f \geq 1 : \mathbf{h}(p^f) = \mathbf{h}(p^{f+1})/p\}.$$

Definition 4.3.2. *Assume Hypothesis 4.3.3 and suppose that* $\mathcal{S} = \emptyset$. *For an integer* $a \neq 0$ *and a real number* $\mathbf{k} \geq 0$, *we define*

$$\mu_{\mathbf{k}}(a, M) := -\frac{1}{2}\frac{(\log M)^{1-\mathbf{k}-\omega_{\mathbf{h}}(a)}}{\Gamma(2-\mathbf{k}-\omega_{\mathbf{h}}(a))} \prod_{\substack{p^f \| a: \\ \mathbf{h}(p^f) = \frac{\mathbf{h}(p^{f+1})}{p}, \\ f \geq 0}} \frac{1 + \mathbf{h}(p) + ... + \mathbf{h}(p^f)}{(1-1/p)^{\mathbf{k}-1}} \log p$$
$$\times \prod_{\substack{p^f \| a: \\ \mathbf{h}(p^f) \neq \frac{\mathbf{h}(p^{f+1})}{p}, \\ f \geq 0}} \frac{\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p}{(1-1/p)^{\mathbf{k}}}. \quad (4.3.3)$$

Remark. *The first product on the right hand side of* (4.3.3) *is a finite product, since* a *is fixed and* $\mathbf{h}(p) < p$ *for all* p. *The second product is convergent, since for* $p \nmid a$ *we have* $\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p = 1 - \mathbf{h}(p)/p \approx 1 - \mathbf{k}/p$. *Of course both these statements rely on the assumption of Hypothesis 4.3.3.*

Remark. *One sees that for integer values of* $\mathbf{k}$, $\mu_{\mathbf{k}}(a, M) = 0$ *iff* $\omega_{\mathbf{h}}(a) \geq 2 - \mathbf{k}$, *by the location of the poles of* $\Gamma(s)$. *Moreover, since these are the only poles, we have* $\mu_{\mathbf{k}}(a, M) \neq 0$ *whenever* $\mathbf{k} \notin \mathbb{Z}$.

Remark. *If* $\mathcal{S} \neq \emptyset$, *we can still give a formula for* $\mu_{\mathbf{k}}(a, M)$, *assuming we understand well* $\mathbf{g}_a(p^e)$ *with* $p \in \mathcal{S}$. *However, this would complicate the already lengthy definition of* $\mu_{\mathbf{k}}(a, M)$, *so we only give individual descriptions in the examples.*

4.4. Main result

The main result of the paper is a formula for the average value of the discrepancy $\mathcal{A}(x; q, a) - \frac{\mathbf{f}_a(q)}{q\gamma(q)}\mathcal{A}(x)$, summed over $1 \leq q \leq Q$, with Q large enough in terms of x.

Theorem 4.4.1. *Assume that hypotheses 4.3.1, 4.3.2, 4.3.3 and 4.3.4 hold with* $\mathcal{S} = \emptyset$ *and the function* $\mathbf{L}(x)$. *Fix an integer* $a \neq 0$ *and let* $M = M(x)$ *be a function of* x *such that* $1 \leq M(x) \leq \mathbf{L}(x)$. *We have for any fixed real number* $A > 0$ *that*

$$\sum_{q \leq \frac{x}{M}} \left(\mathcal{A}(x; q, a) - \mathbf{a}(a) - \frac{\mathbf{f}_a(q)}{q\gamma(q)}\mathcal{A}(x) \right)$$
$$= \frac{\mathcal{A}(x)}{M} \left(\mu_{\mathbf{k}}(a, M)(1 + o(1)) + O\left(\frac{1}{\log^A M} \right) \right), \quad (4.4.1)$$

where $\mathbf{a}(a)$ *is the first term of* $\mathcal{A}(x; q, a)$ *for positive* a*, and whenever* a *is negative, we set* $\mathbf{a}(a) = 0$*. The constant implied in* O *depends on* a*,* A *and* $\mathcal{A}$*.*

We also give a dyadic version, which assumes a weaker form of Hypothesis 4.3.1, and does not assume Hypothesis 4.3.4 at all.

Theorem 4.4.1*. *Assume that hypotheses 4.3.1*, 4.3.2 and 4.3.3 hold with* $\mathcal{S} = \emptyset$ *and the function* $\mathbf{L}(x)$*. Fix an integer* $a \neq 0$ *and let* $M = M(x)$ *be a function of* x *such that* $1 \leq M(x) \leq \mathbf{L}(x)$*. We have for any fixed real number* $A > 0$ *that*

$$\sum_{\frac{x}{2M} < q \leq \frac{x}{M}} \left(\mathcal{A}(x; q, a) - \mathbf{a}(a) - \frac{\mathbf{f}_a(q)}{q\gamma(q)} \mathcal{A}(x) \right)$$

$$= \frac{\mathcal{A}(x)}{2M} \left(\mu_{\mathbf{k}}(a, M)(1 + o(1)) + O\left(\frac{1}{\log^A M} \right) \right). \quad (4.4.2)$$

The constant implied in O *depends on* a*,* A *and* $\mathcal{A}$*.*

Remark 4.4.1. *As we have seen in the examples of Section 4.2, theorems 4.4.1 and 4.4.1* easily generalize to arbitrary (given) sets* $\mathcal{S} \neq \emptyset$*, as long as we understand* $\mathbf{g}_a(p^e)$ *for each* $p \in \mathcal{S}$*.*

Remark 4.4.2. *If* $\mu_{\mathbf{k}}(a, M) \neq 0$*, then theorems 4.4.1 and 4.4.1* give asymptotics for the sum on the left hand side.*

Remark 4.4.3. *Suppose that* $\mathbf{k} = 0$ *(e.g. when* $\mathcal{A}$ *is the sequence of primes).*

If $\omega_{\mathbf{h}}(a) \geq 2$*, then* $\mu_0(a, M) = 0$*.*

If $\omega_{\mathbf{h}}(a) = 1$*, so there is a unique* $p_0^{f_0} \parallel a$*,* $f_0 \geq 1$*, such that* $\mathbf{h}(p_0^{f_0}) = \mathbf{h}(p_0^{f_0+1})/p_0$*, then*

$$\mu_0(a, M) = -\frac{1}{2}\left(1 - \frac{1}{p_0}\right)(1+\mathbf{h}(p_0)+...+\mathbf{h}(p_0^{f_0}))\log p_0 \prod_{\substack{p^f \parallel a \\ f \geq 0 \\ p \neq p_0}} \left(\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p\right).$$

If $\omega_{\mathbf{h}}(a) = 0$*, then*

$$\mu_0(a, M) = -\frac{\log M}{2} \prod_{\substack{p^f \parallel a \\ f \geq 0}} (\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p).$$

Remark 4.4.4. *Suppose that $\mathbf{k} = 1$ (e.g. when $\mathcal{A}$ is the sequence of integers which can be written as the sum of two squares, counted with multiplicity). Then*

$$\mu_1(a, M) = -\frac{1}{2} \prod_{\substack{p^f \| a \\ f \geq 0}} \frac{\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p}{1 - 1/p}.$$

Remark 4.4.5. *Suppose that $\mathbf{k}$ is an integer ≥ 2 (e.g. when $\mathcal{A}$ is the sequence of integers of the form $(m + c_1)(m + c_2) \cdots (m + c_{\mathbf{k}})$, where the c_i are distinct integers). Then $\mu_1(a, M) = 0$.*

4.5. Proof of the main result

The goal of this section is to prove theorems 4.4.1 and 4.4.1*.

4.5.1. An estimate for the main sum

In this section, we will assume that $\mathcal{S} = \emptyset$ for simplicity. Again, the results easily generalize to $\mathcal{S} \neq \emptyset$.

Proposition 4.5.1. *Assume Hypothesis 4.3.3. Let $M = M(x)$ and $\mathbf{R} = \mathbf{R}(x)$ be two positive functions of x such that $M(x)^{1+\delta} \leq \mathbf{R}(x) \leq \sqrt{x}$ for a fixed $\delta > 0$. We have*

$$\sum_{1 \leq r \leq \mathbf{R}} \frac{\mathbf{f}_a(r)}{r\gamma(r)} \left(1 - \frac{r}{\mathbf{R}}\right) - \sum_{1 \leq r \leq M} \frac{\mathbf{f}_a(r)}{r\gamma(r)} \left(1 - \frac{r}{M}\right) - \sum_{\frac{x}{\mathbf{R}} < q \leq \frac{x}{M}} \frac{\mathbf{f}_a(q)}{q\gamma(q)}$$
$$= \frac{\mu_k(a, M)}{M} \left(1 + O\left(\frac{\log\log M}{\log M}\right)\right) + O_{A,\delta}\left(\frac{1}{M \log^A M}\right).$$

The proof of Proposition 4.5.1 will require several lemmas.

Lemma 4.5.1. *With $\mathbf{f}_a(n)$ and $\gamma(n)$ defined as in Section 4.3.1, we have*

$$\frac{\mathbf{f}_a(n)}{n\gamma(n)} \ll \frac{1}{\phi(n)}.$$

Proof. By definition,

$$\begin{aligned}\frac{\mathbf{f}_a(n)}{n\gamma(n)} = \mathbf{g}_a(n) = \prod_{p^e \| n} \mathbf{g}_a(p^e) &\ll_{a,\mathcal{S}} \prod_{\substack{p^e \| n \\ p\nmid a, p\notin\mathcal{S}}} \mathbf{g}_a(p^e) \\ &= \prod_{\substack{p^e \| n \\ p\nmid a, p\notin\mathcal{S}}} \frac{1}{\phi(p^e)}\left(1 - \frac{\mathbf{h}(p)}{p}\right) \\ &\le \prod_{\substack{p^e \| n \\ p\nmid a, p\notin\mathcal{S}}} \frac{1}{\phi(p^e)} \ll_{a,\mathcal{S}} \frac{1}{\phi(n)}.\end{aligned}$$

□

Lemma 4.5.2. *Assume Hypothesis 4.3.3. Let* $h : [0,\infty) \to [0,\infty)$ *be a piecewise continuous function supported on* $[0,1]$, *taking a value halfway between the limit values at discontinuities, and suppose the integral*

$$\mathcal{M}h(s) := \int_0^1 h(x)x^{s-1}dx$$

converges absolutely for $\mathrm{Re}(s) > 0$. *Then,*

$$\sum_{n\le M} \frac{\mathbf{f}_a(n)}{n\gamma(n)} h\left(\frac{n}{M}\right) = \frac{1}{2\pi i}\int_{(1)} \mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}} Z_5(s)\mathcal{M}h(s)M^s ds, \tag{4.5.1}$$

where

$$\begin{aligned}\mathfrak{S}_2(s) := \prod_{\substack{p^f \| a \\ f\ge 1}} \Bigg[&\left(1 + \frac{\mathbf{h}(p)}{p^{s+1}} + \ldots + \frac{\mathbf{h}(p^f)}{p^{f(s+1)}}\right)\left(1 - \frac{1}{p^{s+1}}\right) \\ &+ \frac{\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p}{1 - 1/p}\frac{1}{p^{(f+1)(s+1)}}\Bigg]\left(1 - \frac{1}{p^{s+2}}\right)^{1-\mathbf{k}}, \\ Z_5(s) := \prod_{p\nmid a}&\left(1 + \frac{1}{p^{s+1}}\left(\frac{1}{\gamma(p)} - 1\right)\right)\left(1 - \frac{1}{p^{s+2}}\right)^{1-\mathbf{k}}. \end{aligned} \tag{4.5.2}$$

Moreover, $\mathfrak{S}_2(s)$ *is holomorphic in* $\mathbb{C}\setminus\{-2\}$ *and* $Z_5(s)$ *is holomorphic for* $\mathrm{Re}\, s > -1$.

PROOF. Define

$$Z_{\mathcal{A}}(s) := \sum_{n=1}^{\infty} \frac{\mathbf{g}_a(n)}{n^s} = \prod_p \left(1 + \frac{\mathbf{g}_a(p)}{p^s} + \frac{\mathbf{g}_a(p^2)}{p^{2s}} + \ldots\right).$$

A standard computation using the definition of $\mathbf{g}_a(n)$ (see (4.3.1)) yields that

$$Z_{\mathcal{A}}(s) = \mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}}Z_5(s).$$

The function $\mathfrak{S}_2(s)$ is clearly holomorphic in $\mathbb{C} \setminus \{-2\}$, and the fact that $Z_5(s)$ is holomorphic for $\operatorname{Re} s > -1$ follows from Hypothesis 4.3.3. Now, Mellin inversion gives that

$$h\left(\frac{n}{M}\right) = \frac{1}{2\pi i}\int_{(1)} \frac{M^s}{n^s}\mathcal{M}h(s)ds.$$

Multiplying by $\frac{\mathbf{f}_a(n)}{n\gamma(n)}$ and summing over n yields the result. □

4.5.1.1. *Properties of the Dirichlet series*

Lemma 4.5.3. *Assume Hypothesis 4.3.3. We have*

$$Z_5(s) = Z_5(-1) + O(|s+1|)$$

in the region $|s+1| \leq 3$, *with* $\operatorname{Re}(s) > -1$. *Note that by Hypothesis 4.3.3, the product defining* $Z_5(-1)$ *is convergent (see the proof of Proposition 4.5.2).*

PROOF. We will show that

$$\log \frac{Z_5(s)}{Z_5(-1)} \ll |s+1|,$$

from which the lemma clearly follows. Let s be a complex number with $\operatorname{Re} s > -1$. We compute

$$\log \frac{Z_5(s)}{Z_5(-1)} = \sum_{p\nmid a} \log\left(\frac{1+\frac{1}{p^{s+1}}(\frac{1}{\gamma(p)}-1)}{\frac{1}{\gamma(p)}} \cdot \frac{\left(1-\frac{1}{p^{s+2}}\right)^{1-\mathbf{k}}}{\left(1-\frac{1}{p}\right)^{1-\mathbf{k}}} \right)$$

$$= \sum_{p\nmid a} \left[\log\left(1-(1-\gamma(p))\left(1-\frac{1}{p^{s+1}}\right)\right) \right.$$

$$\left. + (1-\mathbf{k})\log\left(1+\frac{1}{p-1}\left(1-\frac{1}{p^{s+1}}\right)\right)\right]$$

$$= \sum_{p\nmid a}\left[\frac{\mathbf{h}(p)-1}{p-\mathbf{h}(p)}\left(1-\frac{1}{p^{s+1}}\right)+\frac{1-\mathbf{k}}{p-1}\left(1-\frac{1}{p^{s+1}}\right)\right]$$

$$+ O_\epsilon\left(|s+1|^2\sum_p \frac{\log^2 p}{p^{2-\epsilon}}\right)$$

$$= \sum_{p\nmid a}\left(\frac{\mathbf{h}(p)-1}{p-\mathbf{h}(p)}+\frac{1-\mathbf{k}}{p-1}\right)\left(1-\frac{1}{p^{s+1}}\right) + O\left(|s+1|^2\right)$$

$$= \sum_{p}\left(\frac{\mathbf{h}(p)-1}{p-\mathbf{h}(p)}+\frac{1-\mathbf{k}}{p-1}\right)\left(1-\frac{1}{p^{s+1}}\right) + O\left(|s+1|\right). \tag{4.5.3}$$

Note that by Hypothesis 4.3.3, the series

$$\sum_p \left(\frac{\mathbf{h}(p)-1}{p-\mathbf{h}(p)}+\frac{1-\mathbf{k}}{p-1}\right) = \sum_p \frac{\mathbf{h}(p)-\mathbf{k}}{p} + O(1)$$

converges. Moreover, summation by parts yields the following estimate :

$$S(t) := \sum_{p\leq t}\left(\frac{\mathbf{h}(p)-1}{p-\mathbf{h}(p)}+\frac{1-\mathbf{k}}{p-1}\right) = S(\infty) + O\left(\frac{1}{\log^2(t+2)}\right).$$

We then get that

$$\begin{aligned}
\sum_{p\le T}\left(\frac{\mathbf{h}(p)-1}{p-\mathbf{h}(p)}+\frac{1-\mathbf{k}}{p-1}\right)\left(1-\frac{1}{p^{s+1}}\right) &= \int_1^T\left(1-\frac{1}{t^{s+1}}\right)dS(t)\\
&= \left(1-\frac{1}{t^{s+1}}\right)S(t)\Big|_1^T-(s+1)\int_1^T\frac{S(t)}{t^{s+2}}dt\\
&= \left(1-\frac{1}{T^{s+1}}\right)\left(S(\infty)+O\left(\frac{1}{\log^2 T}\right)\right)-(s+1)\int_1^T\frac{S(\infty)}{t^{s+2}}dt\\
&\qquad\qquad + O\left(|s+1|\int_1^T\frac{dt}{t\log^2(t+2)}\right)\\
&= S(\infty)\left(1-\frac{1}{T^{s+1}}\right)+O\left(\frac{1}{\log^2 T}\right)+\frac{S(\infty)}{t^{s+1}}\Big|_1^T+O\left(|s+1|\right)\\
&= O\left(\frac{1}{\log^2 T}+|s+1|\right).
\end{aligned}$$

Taking $T\to\infty$ yields that (4.5.3) is $\ll |s+1|$. □

Lemma 4.5.4. *Let* $f(s)$ *be a holomorphic function over a domain* $\mathcal{D}$. *We have that* $\frac{f^{(n)}}{f}(s)$ *is a polynomial in the variables* $\left(\frac{f'(s)}{f(s)}\right)^{(0)}, \left(\frac{f'(s)}{f(s)}\right)^{(1)}, \ldots, \left(\frac{f'(s)}{f(s)}\right)^{(n-1)}$*, with integer coefficients.*

PROOF. The proof goes by induction, using the identity

$$\frac{f^{(n)}}{f}=\left(\frac{f^{(n-1)}}{f}\right)'+\frac{f^{(n-1)}}{f}\frac{f'}{f}.$$

□

Lemma 4.5.5. *Assume Hypothesis 4.3.3. Let* $Z_5(s)$ *be defined as in* (4.5.2) *and let* $n\ge 0$. *Then there exists* $\delta>0$ *such that, uniformly in the region* $-1<\sigma<-\frac{1}{2}$ *and* $t\in\mathbb{R}$, *we have*

$$Z_5^{(n)}(\sigma+it)\ll_n(|t|+2)^{1/2-\delta}. \tag{4.5.4}$$

In particular, if t *is fixed, then* $Z_5^{(n)}(\sigma+it)$ *is bounded near* $\sigma=-1$.

PROOF. First write $Z_5(s) = Z_3(s)Z_4(s)$, where

$$Z_3(s) := \prod_{p \nmid a} \left(1 + \frac{1}{p^{s+1}}\left(\frac{1}{\gamma(p)} - 1\right)\right)\left(1 - \frac{1-\mathbf{k}}{p^{s+2}}\right),$$

$$Z_4(s) := \prod_{p \nmid a} \left(1 - \frac{1}{p^{s+2}}\right)^{1-\mathbf{k}}\left(1 - \frac{1-\mathbf{k}}{p^{s+2}}\right)^{-1}.$$

The function $Z_4(s)$ is uniformly bounded in the region $\operatorname{Re} s \geq -1$, since the Eulerian product converges absolutely. As for $Z_3(s)$, we have for $-1 < \sigma < -\frac{1}{2}$ that

$$\log Z_3(\sigma + it) = \log \prod_{p \nmid a}\left(1 + \frac{1}{p^{\sigma+1+it}}\frac{\mathbf{k} - \mathbf{h}(p)}{p}\right) + O(1).$$

Hypothesis 4.3.3 gives

$$S(x,t) := \sum_{p \leq x} \frac{\mathbf{k} - \mathbf{h}(p)}{p^{1+it}} \leq (1/2 - \delta)\log(|t| + 2) + O(1).$$

Thus,

$$\begin{aligned}
\log \prod_{p \nmid a}\left(1 + \frac{1}{p^{\sigma+1+it}}\frac{\mathbf{k} - \mathbf{h}(p)}{p}\right) &= \sum_{p \nmid a} \frac{1}{p^{\sigma+1}}\frac{\mathbf{k} - \mathbf{h}(p)}{p^{1+it}} + O(1) \\
&= \int_1^\infty \frac{dS(x,t)}{x^{\sigma+1}} + O(1) \\
&= \frac{S(x,t)}{x^{\sigma+1}}\bigg|_1^\infty + (\sigma+1)\int_1^\infty \frac{S(x,t)}{x^{\sigma+2}}dx + O(1) \\
&\leq (1/2 - \delta)\log(|t| + 2)\int_1^\infty \frac{\sigma+1}{x^{\sigma+2}}dx + O(1) \\
&= (1/2 - \delta)\log(|t| + 2) + O(1),
\end{aligned}$$

which proves (4.5.4) for $n = 0$. The bound

$$\sum_{p \leq x} \frac{(\mathbf{k} - \mathbf{h}(p))\log^m p}{p^{1+it}} \ll_\epsilon (|t| + 2)^\epsilon$$

gives

$$\left(\frac{Z_5'(\sigma + it)}{Z_5(\sigma + it)}\right)^{(m)} \ll_\epsilon (|t| + 2)^\epsilon \tag{4.5.5}$$

for $m \geq 0$. We finish the proof of (4.5.4) for $n \geq 1$ by applying Lemma 4.5.4.

□

Lemma 4.5.6. *We have for $|\sigma + it - 1| > \frac{1}{10}$ that*

$$\zeta(\sigma + it) \ll_\epsilon (|t| + 2)^{\mu(\sigma)+\epsilon},$$

where

$$\mu(\sigma) = \begin{cases} 1/2 - \sigma & \text{if } \sigma \leq 0 \\ 1/2 - 2\sigma/3 & \text{if } 0 \leq \sigma \leq 1/2 \\ 1/3 - \sigma/3 & \text{if } 1/2 \leq \sigma \leq 1 \\ 0 & \text{if } \sigma \geq 1. \end{cases}$$

Moreover, these bounds are uniform for σ contained in any compact subset of $\mathbb{R}$.

PROOF. See Section II.3.4 of **[76]**, in particular (II.3.13) and Theorem 3.8. By studying the proof of the Phragment-Lindelöf principle (see Chapter 9 of **[20]** for instance), we see that the bounds we get are uniform in σ. □

Lemma 4.5.7. *Assume Hypothesis 4.3.3. Let*

$$Z(s) := \frac{\mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}}Z_5(s)}{s(s+1)},$$

with $\mathfrak{S}_2(s)$ and $Z_5(s)$ defined as in Lemma 4.5.2. There exists $\delta > 0$ such that uniformly for $|t| \geq 2$ and $-1 < \sigma < -\frac{1}{2}$,

$$Z^{(n)}(\sigma + it) \ll_n \frac{1}{|t|^{1+\delta}}.$$

PROOF. Define

$$Z_6(s) := \frac{\mathfrak{S}_2(s)\zeta(s+2)^{1-\mathbf{k}}Z_5(s)}{s(s+1)}.$$

Write $s = \sigma + it$, with $-1 < \sigma < -\frac{1}{2}$ and $|t| \geq 2$. We have for $m \geq 0$ that

$$\left(\frac{Z_6'(s)}{Z_6(s)}\right)^{(m)} = \left(\frac{\mathfrak{S}_2'(s)}{\mathfrak{S}_2(s)}\right)^{(m)} + (1-\mathbf{k})\left(\frac{\zeta'(s+2)}{\zeta(s+2)}\right)^{(m)} + \left(\frac{Z_5'(s)}{Z_5(s)}\right)^{(m)} - \left(\frac{2s+1}{s(s+1)}\right)^{(m)}.$$

We compute that

$$\left(\frac{\mathfrak{S}_2'(s)}{\mathfrak{S}_2(s)}\right)^{(m)} \ll_m 1, \qquad \left(\frac{2s+1}{s(s+1)}\right)^{(m)} \ll_m 1, \qquad \left(\frac{Z_5'(s)}{Z_5(s)}\right)^{(m)} \ll_{m,\epsilon} |t|^\epsilon.$$

(The first bound is clear, the second follows from the fact that $|t| \geq 2$ and the third comes from (4.5.5).) Applying Cauchy's formula for the derivatives as in Corollaire II.3.10 of **[76]** and then using the bound (II.3.55) of **[76]** yields

$$\left(\frac{\zeta'(s+2)}{\zeta(s+2)}\right)^{(m)} \ll_m \log^{m+1}(|t|).$$

Using Lemma 4.5.4,

$$Z_6^{(m)}(s) \ll_{\epsilon,m} |Z_6(s)||t|^{\epsilon}$$

for $m \geq 0$. We now use Lemma 4.5.5 to bound $|Z_5(s)|$, which gives

$$Z_6^{(m)}(s) \ll_m |\zeta(s+2)^{1-\mathbf{k}}||t|^{-3/2-2\delta}$$

for some $\delta > 0$. Now if $\mathbf{k} \leq 1$, we use Lemma 4.5.6 to bound $\zeta(s+2)^{1-\mathbf{k}}$. Otherwise, we use the bound $(\zeta(s+2))^{-1} \ll \log(|t|)$ (see (II.3.56) of **[76]**). In both cases we get

$$Z_6^{(m)}(s) \ll_m |t|^{-3/2-\delta}.$$

We now use Cauchy's formula for the derivatives, which states that

$$\zeta^{(k)}(s+1) = \frac{k!}{2\pi i}\oint_{|z|=r} \zeta(s+1+z)\,\frac{dz}{z^{k+1}}.$$

Selecting $r = \epsilon/2$ and applying Lemma 4.5.6, we get the bound[1]

$$\zeta^{(k)}(s+1) \ll_{k,\epsilon} |t|^{1/2+\epsilon}.$$

We conclude the existence of $\delta > 0$ such that

$$Z^{(n)}(s) = \sum_{i=0}^{n} \binom{n}{i} \zeta^{(i)}(s+1) Z_6^{(n-i)}(s) \ll_n \frac{1}{|t|^{1+\delta}}.$$

□

[1] This bound is still valid outside the zero-free region of $\zeta(s+1)$; this is why we considered the ordinary derivatives of $\zeta(s+1)$ instead of its logarithmic derivatives as with the other terms.

4.5.1.2. *The value of* $\mu_{\mathbf{k}}(a, M)$

Proposition 4.5.2. *Assume Hypothesis 4.3.3. If* $\mathbf{k} \in \mathbb{Z}$, *then*

$$\frac{1}{2\pi i}\int_{(-1/2)} \frac{\mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}}Z_5(s)}{s(s+1)}M^s ds$$
$$= -\frac{\mu_{\mathbf{k}}(a, M)}{M}\left(1 + O\left(\frac{\log\log M}{\log M}\right)\right) + O_A\left(\frac{1}{M\log^A M}\right)$$

where $\mu_{\mathbf{k}}(a, M)$ *is defined in Definition 4.3.2.*

Proof. We first need to understand the behaviour of

$$Z(s) := \frac{\mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}}Z_5(s)}{s(s+1)} \tag{4.5.6}$$
$$= (s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2}\frac{\mathfrak{S}_2(s)}{(s+1)^{\omega_{\mathbf{h}}(a)}}\zeta(s+1)((s+1)\zeta(s+2))^{1-\mathbf{k}}\frac{Z_5(s)}{s} \tag{4.5.7}$$

in the region $\mathcal{D} : -1 \leq \operatorname{Re} s \leq -1/2$. This function is holomorphic for $\operatorname{Re} s > -1$ by Lemma 4.5.2, and as we will see, the only point in $\mathcal{D}$ where $Z(s)$ is not necessarily locally bounded is $s = -1$. The functions

$$\zeta(s+1), \qquad ((s+1)\zeta(s+2))^{1-\mathbf{k}} \qquad \text{and} \qquad \frac{1}{s}$$

are holomorphic on $\mathcal{D}$ and do not vanish at $s = -1$. The function $Z_5(s)$ is holomorphic for $\operatorname{Re} s > -1$, and all its derivatives are locally bounded around any point of $\mathcal{D}$ by Lemma 4.5.5. We compute

$$Z_5(-1) = \prod_{p \nmid a} \frac{1 - \mathbf{h}(p)/p}{(1-1/p)^{\mathbf{k}}} \neq 0,$$

since $\mathbf{h}(p) < p$. As for the function $\mathfrak{S}_2(s)$, it is holomorphic on $\mathcal{D}$. However, this function can vanish at $s = -1$ if for a certain $p \mid a$ we have $\mathbf{h}(p^f) = \mathbf{h}(p^{f+1})/p$. In this case, we have for s close to -1 that

$$\mathfrak{S}_2(s)\prod_{p|a}\left(1-\frac{1}{p^{s+2}}\right)^{\mathbf{k}-1} = \prod_{\substack{p^f\|a:\\ \mathbf{h}(p^f)\neq\mathbf{h}(p^{f+1})/p,\\ f\geq 1}}\left[\frac{\mathbf{h}(p^f)-\mathbf{h}(p^{f+1})/p}{1-1/p}+O(|s+1|)\right]$$
$$\times\prod_{\substack{p^f\|a:\\ \mathbf{h}(p^f)=\mathbf{h}(p^{f+1})/p,\\ f\geq 1}}[(s+1)(1+\mathbf{h}(p)+...+\mathbf{h}(p^f))\log p+O(|s+1|^2)],$$

and since $\mathbf{h}(p^e)\geq 0$, this shows that every local factor has at most a simple zero at $s=-1$. We conclude that

$$\frac{\mathfrak{S}_2(s)}{(s+1)^{\omega_{\mathbf{h}}(a)}}$$

is holomorphic on $\mathcal{D}$ and does not vanish at $s=-1$. We now split in three distinct cases, depending on the analytic nature of $(s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2}$ near $s=-1$.

First case : $\mathbf{k}+\omega_{\mathbf{h}}(a)\geq 2$. In this case, $Z(s)$ and all of its derivatives are bounded near $s=-1$. To show this, note that it is true for the functions

$$(s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2},\qquad \frac{\mathfrak{S}_2(s)}{(s+1)^{\omega_{\mathbf{h}}(a)}},\qquad \zeta(s+1),$$
$$((s+1)\zeta(s+2))^{1-\mathbf{k}},\qquad \frac{1}{s}\qquad\text{and}\qquad Z_5(s),$$

so it is also true for $Z(s)$ by Leibniz's rule. We now shift the contour of integration to the left until the line $\operatorname{Re}s=-1+\frac{1}{\log M}$ to get

$$\begin{aligned}\frac{1}{2\pi i}\int_{(-1/2)}Z(s)ds &= \frac{i}{2\pi i}\int_{\mathbb{R}}Z\left(-1+\frac{1}{\log M}+it\right)M^{-1+\frac{1}{\log M}+it}dt\\ &=\frac{e}{M}\frac{1}{2\pi}\int_{\mathbb{R}}Z\left(-1+\frac{1}{\log M}+it\right)e^{it\log M}dt,\end{aligned}$$

which gives, after A integrations by parts,

$$\begin{aligned}\frac{1}{2\pi i}\int_{(-1/2)}Z(s)ds &\ll_A \frac{1}{M\log^A M}\int_{\mathbb{R}}\left|Z^{(A)}\left(-1+\frac{1}{\log M}+it\right)\right|\left|e^{it\log M}\right|dt\\ &\ll_A \frac{1}{M\log^A M}\left(O(1)+\int_{|t|\geq 2}\frac{1}{|t|^{1+\delta}}dt\right)\\ &\ll_A \frac{1}{M\log^A M}\end{aligned}$$

by Lemma 4.5.7. Note that the uniformity in σ was crucial. This shows that we can take $\mu_{\mathbf{k}}(a,M)=0$.

Second case : $\mathbf{k} + \omega_{\mathbf{h}}(a) = 1$. Let

$$c := \lim_{s\to 1^+} (s+1)Z(s) \neq 0$$

and define

$$Z_7(s) := Z(s) - \frac{c}{s+1}.$$

We can show using Lemma 4.5.3 that for s close to -1 with $\operatorname{Re} s > -1$, the following bound holds :

$$Z_7(s) \ll 1.$$

Lemma 4.5.3 implies that for s close to -1 with $\operatorname{Re} s > -1$, the function

$$Z_7'(s) = \frac{((s+1)Z(s))'}{s+1} - \frac{(s+1)Z(s)}{(s+1)^2} + \frac{c}{(s+1)^2}$$

satisfies

$$Z_7'(s) \ll \frac{1}{|s+1|}.$$

Using Lemma 4.5.7, we get that for $|t| \geq 2$,

$$Z_7'(s) \ll \frac{1}{|t|^{1+\delta}}.$$

Thus,

$$\begin{aligned}
\frac{1}{2\pi i}\int_{(-1+\frac{1}{\log M})} Z_7(s)M^s ds &= \frac{-1}{2\pi i \log M}\int_{(-1+\frac{1}{\log M})} Z_7'(s)M^s ds \\
&\ll \frac{1}{M\log M}\left|\int_{-\infty}^{\infty} Z_7'\left(-1+\frac{1}{\log M}+it\right)M^{it}dt\right| \\
&\ll \frac{1}{M\log M}\left(\left|\int_{-2}^{2} Z_7'\left(-1+\frac{1}{\log M}+it\right)M^{it}dt\right| + O(1)\right) \\
&\ll \frac{1}{M\log M}\left(\int_{-2}^{2} \frac{1}{\frac{1}{\log M}+|t|}dt + O(1)\right) \\
&\ll \frac{1}{M\log M}\left(\int_{0}^{\frac{1}{\log M}} \log M + \int_{\frac{1}{\log M}}^{2} \frac{1}{t}dt + O(1)\right) \\
&\ll \frac{\log\log M}{M\log M}.
\end{aligned}$$

Combining this bound with an easy residue computation yields

$$\begin{aligned}\frac{1}{2\pi i}\int_{(-1/2)} Z(s)M^s ds &= \frac{1}{2\pi i}\int_{(-1/2)} Z_7(s)M^s ds + \frac{1}{2\pi i}\int_{(-1/2)} \frac{c}{s+1}M^s ds \\ &= \frac{c}{M}\left(1 + O\left(\frac{\log\log M}{\log M}\right)\right).\end{aligned}$$

Now remarks 4.4.3 and 4.4.4 show that $c = -\mu_{\mathbf{k}}(a, M)$, which concludes this case.

Third case : $\mathbf{k} = \omega_{\mathbf{h}}(a) = 0$. Defining

$$c := \lim_{s\to -1^+}(s+1)^2 Z(s) \neq 0,$$

we get that the function $Z_8(s) := Z(s) - \frac{c}{(s+1)^2}$ satisfies the bound

$$Z_8(s) \ll \frac{1}{|s+1|}$$

by Lemma 4.5.3. An easy residue computation yields

$$\frac{1}{2\pi i}\int_{(-1/2)} Z(s)M^s ds = c\frac{\log M}{M} + \frac{1}{2\pi i}\int_{(-1+\frac{1}{\log M})} Z_8(s)M^s ds.$$

Proceeding in an analogous way to the previous case, we compute

$$\begin{aligned}\int_{(-1+\frac{1}{\log M})} Z_8(s)M^s ds &\ll \frac{1}{M}\left|\int_{-\infty}^{\infty} Z_8\left(-1 + \frac{1}{\log M} + it\right) M^{it} dt\right| \\ &\ll \frac{1}{M}\left(\left|\int_{-2}^{2} Z_8\left(-1 + \frac{1}{\log M} + it\right) M^{it} dt\right| + O(1)\right) \\ &\ll \frac{\log\log M}{M},\end{aligned}$$

from which we conclude

$$\begin{aligned}\frac{1}{2\pi i}\int_{(-1/2)} Z(s)M^s ds &= c\frac{\log M}{M}\left(1 + \frac{\log\log M}{\log M}\right) \\ &= -\frac{\mu_0(a, M)}{M}\left(1 + \frac{\log\log M}{\log M}\right)\end{aligned}$$

by Remark 4.4.3, since

$$c = \frac{1}{2}\prod_{\substack{p^f \| a \\ f\geq 0}} (\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p).$$

□

Lemma 4.5.8. *Let $z > 1$ be a real number. Then,*

$$\frac{1}{2\pi i}\int_{\mathrm{Re}\, s=-1/2}\frac{M^s}{(s+1)^z}ds = \frac{1}{M}\frac{(\log M)^{z-1}}{\Gamma(z)}.$$

PROOF. Let $R \geq 2$ be a large real number and consider $\mathcal{H}_R$ a Hankel contour centered at $s = -1$ and truncated at $-R \pm \epsilon i$ (see Théorème II.0.17 of [**76**]). Define C_R to be the union of two circle segments starting at the endpoints of $\mathcal{H}_R$ and ending at the points $\pm iR$. By Cauchy's formula,

$$\begin{aligned}\frac{1}{2\pi i}\int_{\mathrm{Re}\, s=-1/2}\frac{M^s}{(s+1)^z}ds &= \frac{1}{2\pi i}\int_{\mathrm{Re}\, s=0}\frac{M^s}{(s+1)^z}ds\\ &= \frac{1}{2\pi i}\int_{\mathcal{H}_R}\frac{M^s}{(s+1)^z}ds + \frac{1}{2\pi i}\int_{C_R}\frac{M^s}{(s+1)^z}ds\\ &= \frac{1}{2\pi i}\int_{\mathcal{H}_R}\frac{M^s}{(s+1)^z}ds + O\left(\frac{1}{R^{z-1}}\right),\end{aligned}$$

so by taking $R \to \infty$,

$$\begin{aligned}\frac{1}{2\pi i}\int_{\mathrm{Re}\, s=-1/2}\frac{M^s}{(s+1)^z}ds &= \frac{1}{2\pi i}\int_{\mathcal{H}_\infty}\frac{M^s}{(s+1)^z}ds\\ &= \frac{1}{M}\frac{1}{2\pi i}\int_{\mathcal{H}_\infty}\frac{e^{(s+1)\log M}}{(s+1)^z}ds\\ &= \frac{(\log M)^{z-1}}{M}\frac{1}{2\pi i}\int_{\mathcal{H}'_\infty}\frac{e^w}{w^z}dw\\ &= \frac{1}{M}\frac{(\log M)^{z-1}}{\Gamma(z)}\end{aligned}$$

by Hankel's formula. Here, $\mathcal{H}'_\infty$ denotes an infinite Hankel contour centered at $w = 0$. □

Proposition 4.5.3. *Assume Hypothesis 4.3.3. If $\mathbf{k} \notin \mathbb{Z}$, then*

$$\frac{1}{2\pi i}\int_{(-1/2)}\frac{\mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}}Z_5(s)}{s(s+1)}M^s ds$$
$$= -\frac{\mu_{\mathbf{k}}(a,M)}{M}\left(1+O\left(\frac{1}{\log M}\right)\right).$$

PROOF. As in Proposition 4.5.2, we need to study the function

$$Z(s) = (s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2} \frac{\mathfrak{S}_2(s)}{(s+1)^{\omega_{\mathbf{h}}(a)}} \zeta(s+1)((s+1)\zeta(s+2))^{1-\mathbf{k}} \frac{Z_5(s)}{s}$$

in the region $\mathcal{D} : -1 \leq \operatorname{Re} s \leq -1/2$. This function is holomorphic for $\operatorname{Re} s > -1$ by Lemma 4.5.5, and the only point in $\mathcal{D}$ where $Z(s)$ is not necessarily locally bounded is $s = -1$. However, the functions

$$\frac{\mathfrak{S}_2(s)}{(s+1)^{\omega_{\mathbf{h}}(a)}}, \qquad \zeta(s+1), \qquad ((s+1)\zeta(s+2))^{1-\mathbf{k}} \qquad \text{and} \qquad \frac{1}{s}$$

are holomorphic on $\mathcal{D}$ and do not vanish at $s = -1$. The function $Z_5(s)$ is holomorphic for $\operatorname{Re} s > -1$, all its derivatives are locally bounded around any point of $\mathcal{D}$, and $Z_5(-1) \neq 0$. Define

$$Z_9(s) := Z(s) - c(s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2},$$

where

$$c := \lim_{s \to -1^+} (s+1)^{2-\mathbf{k}-\omega_{\mathbf{h}}(a)} Z(s) \neq 0.$$

We have that

$$\begin{aligned} \frac{1}{2\pi i} \int_{(-1/2)} Z(s) M^s ds &= \frac{(-1)^{\lceil \mathbf{k} \rceil + \omega_{\mathbf{h}}(a)}}{2\pi i (\log M)^{\lceil \mathbf{k} \rceil + \omega_{\mathbf{h}}(a)}} \int_{(-1/2)} Z^{(\lceil \mathbf{k} \rceil + \omega_{\mathbf{h}}(a))}(s) M^s ds \\ &= \frac{(-1)^{\lceil \mathbf{k} \rceil + \omega_{\mathbf{h}}(a)}}{2\pi i (\log M)^{\lceil \mathbf{k} \rceil + \omega_{\mathbf{h}}(a)}} \left(\int_{(-1/2)} Z_9^{(\lceil \mathbf{k} \rceil + \omega_{\mathbf{h}}(a))}(s) M^s ds \right. \\ &\qquad \left. + c \frac{\Gamma(\mathbf{k} + \omega_{\mathbf{h}}(a) - 1)}{\Gamma(\mathbf{k} - \lceil \mathbf{k} \rceil - 1)} \int_{(-1/2)} (s+1)^{\mathbf{k} - \lceil \mathbf{k} \rceil - 2} M^s ds \right) \\ &= \frac{c}{M} \frac{(\log M)^{1-\mathbf{k}-\omega_{\mathbf{h}}(a)}}{\Gamma(2 - \mathbf{k} - \omega_{\mathbf{h}}(a))} \\ &\qquad + \frac{(-1)^{\lceil \mathbf{k} \rceil + \omega_{\mathbf{h}}(a)}}{2\pi i (\log M)^{\lceil \mathbf{k} \rceil + \omega_{\mathbf{h}}(a)}} \int_{(-1/2)} Z_9^{(\lceil \mathbf{k} \rceil + \omega_{\mathbf{h}}(a))}(s) M^s ds \end{aligned} \tag{4.5.8}$$

by Lemma 4.5.8. We will show the bound

$$Z_9^{(\lceil \mathbf{k} \rceil + \omega_{\mathbf{h}}(a))}(s) \ll |s+1|^{\mathbf{k} - \lceil \mathbf{k} \rceil - 1} \tag{4.5.9}$$

for s close to -1, which will yield (using Lemma 4.5.7)

$$\begin{aligned}\int_{(-1+\frac{1}{\log M})} & Z_9^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}(s)M^s ds \\ &\ll \frac{1}{M}\left|\int_{-\infty}^{\infty} Z_9^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}\left(-1+\frac{1}{\log M}+it\right)M^{it}dt\right| \\ &= \frac{1}{M}\left|\int_{-2}^{2} Z_9^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}\left(-1+\frac{1}{\log M}+it\right)M^{it}dt+O(1)\right| \\ &\ll \frac{1}{M}\left(\int_{-2}^{2}\left(\frac{1}{\log M}+|t|\right)^{\mathbf{k}-\lceil\mathbf{k}\rceil-1}dt+O(1)\right) \\ &\ll \frac{1}{M}\left(\int_{0}^{\frac{1}{\log M}}(\log M)^{1-\mathbf{k}+\lceil\mathbf{k}\rceil}+\int_{\frac{1}{\log M}}^{2} t^{\mathbf{k}-\lceil\mathbf{k}\rceil-1}dt+O(1)\right) \\ &\ll \frac{(\log M)^{\lceil\mathbf{k}\rceil-\mathbf{k}}+1}{M} \ll \frac{(\log M)^{\lceil\mathbf{k}\rceil-\mathbf{k}}}{M},\end{aligned}$$

from which we will conclude using (4.5.8) that

$$\begin{aligned}\frac{1}{2\pi i}\int_{(-1/2)} Z(s)M^s ds &= \frac{c}{M}\frac{(\log M)^{1-\mathbf{k}-\omega_{\mathbf{h}}(a)}}{\Gamma(2-\mathbf{k}-\omega_{\mathbf{h}}(a))}\left(1+O\left(\frac{1}{\log M}\right)\right) \\ &= -\mu_{\mathbf{k}}(a,M)\left(1+O\left(\frac{1}{\log M}\right)\right),\end{aligned}$$

achieving the proof. Let us now show that (4.5.9) holds. By Lemma 4.5.5, the function

$$Z_{10}(s) := (s+1)^{2-\mathbf{k}-\omega_{\mathbf{h}}(a)}Z(s)$$

as well as its derivatives are locally bounded around $s=-1$. Moreover, applying Lemma 4.5.3 gives the bound

$$Z_{10}(s) = Z_{10}(-1) + O(|s+1|). \tag{4.5.10}$$

Now we use Leibniz's formula :

$$\begin{aligned}Z^{(\lceil\mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}(s) &= \left((s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2}\,Z_{10}(s)\right)^{(\lceil\mathbf{k}\rceil+\omega_{\mathbf{h}}(a))} \\ &= \sum_{i=0}^{\lceil\mathbf{k}\rceil+\omega_{\mathbf{h}}(a)}\binom{\lceil\mathbf{k}\rceil+\omega_{\mathbf{h}}(a)}{i}\left((s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2}\right)^{(i)} Z_{10}^{(\lceil\mathbf{k}\rceil+\omega_{\mathbf{h}}(a)-i)}(s) \\ &= \left((s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2}\right)^{(\lceil\mathbf{k}\rceil+\omega_{\mathbf{h}}(a))} Z_{10}(s) + O(|s+1|^{\mathbf{k}-\lceil\mathbf{k}\rceil-1}) \\ &= \left((s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2}\right)^{(\lceil\mathbf{k}\rceil+\omega_{\mathbf{h}}(a))} Z_{10}(-1) + O(|s+1|^{\mathbf{k}-\lceil\mathbf{k}\rceil-1})\end{aligned}$$

by (4.5.10), so

$$\begin{aligned} Z_9^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}(s) &= Z^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))}(s) - c\left((s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2}\right)^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))} \\ &= (Z_{10}(-1)-c)\left((s+1)^{\mathbf{k}+\omega_{\mathbf{h}}(a)-2}\right)^{(\lceil \mathbf{k}\rceil+\omega_{\mathbf{h}}(a))} + O(|s+1|^{\mathbf{k}-\lceil \mathbf{k}\rceil-1}) \\ &= O(|s+1|^{\mathbf{k}-\lceil \mathbf{k}\rceil-1}) \end{aligned}$$

since $c = Z_{10}(-1)$.

□

Lemma 4.5.9. *Assume Hypothesis 4.3.3. Let $y \geq 1$ be a real number. Then,*

$$\frac{1}{2\pi i}\int_{(-1/2)} \mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}}Z_5(s)y^s\frac{ds}{s} \ll_\epsilon y^{-1+\epsilon}. \tag{4.5.11}$$

PROOF. Define

$$Z_{\mathcal{A}}(s) := \mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}}Z_5(s).$$

The goal is to bound the integral

$$\frac{1}{2\pi i}\int_{(-1/2)} Z_{\mathcal{A}}(s)y^s\frac{ds}{s} = \frac{1}{2\pi i}\int_{(-1+\epsilon)} Z_{\mathcal{A}}(s)y^s\frac{ds}{s}.$$

We will first show that this integral is $\ll_\epsilon y^{-1/2+\epsilon}$ using complex analysis, and then we will see how to improve this bound to $\ll_\epsilon y^{-1+\epsilon}$ by elementary means. In the region $-1+\epsilon < \sigma$, we have the bound

$$|Z(\sigma+it)| \ll_\epsilon |\zeta(\sigma+1+it)| \ll_\epsilon (|t|+2)^{\mu(\sigma+1)+\epsilon},$$

where $\mu(\sigma+1)$ is defined as in Lemma 4.5.6. Thus we get the bounds

$$\int_{-1+\epsilon-iT}^{-1+\epsilon+iT} Z(s)y^s\frac{ds}{s} \ll_\epsilon \frac{T^{1/2}}{y^{1-\epsilon}},$$

$$\int_{-1+\epsilon\pm iT}^{\epsilon\pm iT} Z(s)y^s\frac{ds}{s} \ll_\epsilon (Ty)^\epsilon\left(\frac{1}{T^{5/6}y^{1/2}} + \frac{1}{T^{1/2}y} + \frac{1}{T} + \frac{1}{T^{5/6}y^{1/2}}\right).$$

The last integral we need to bound is

$$\begin{aligned} \frac{1}{2\pi i}\int_{\mathrm{Re}\, s=\epsilon, |\mathrm{Im}\, s|>T} Z(s)y^s\frac{ds}{s} &= \sum_n \frac{\mathbf{f}_a(n)}{n\gamma(n)}\frac{1}{2\pi i}\int_{\mathrm{Re}\, s=\epsilon, |\mathrm{Im}\, s|>T}\left(\frac{y}{n}\right)^s\frac{ds}{s} \\ &\ll y^\epsilon \sum_n \frac{\mathbf{f}_a(n)}{n\gamma(n)}\frac{1}{n^\epsilon(1+T|\log(y/n)|)} \end{aligned}$$

by the effective version of Perron's formula (see Théorème II.2.3 of [**76**]). The last sum is

$$\ll \frac{y^{\epsilon}}{\sqrt{T}} \sum_{n\leq y\left(1-\frac{1}{\sqrt{T}}\right)} \frac{\mathbf{f}_a(n)}{n\gamma(n)} + \sum_{y\left(1-\frac{1}{\sqrt{T}}\right)\leq n\leq y\left(1+\frac{1}{\sqrt{T}}\right)} \frac{\mathbf{f}_a(n)}{n\gamma(n)} + \frac{y^{\epsilon}}{\sqrt{T}} \sum_{n\geq y\left(1+\frac{1}{\sqrt{T}}\right)} \frac{\mathbf{f}_a(n)}{n\gamma(n)} \frac{1}{n^{\epsilon}}$$

$$\ll \frac{y^{\epsilon}}{\sqrt{T}} \log y + \frac{1}{\sqrt{T}} \ll_{\epsilon} \frac{y^{\epsilon}}{\sqrt{T}} \log y$$

by Lemma 4.5.1. Taking $T = y$ yields that the left hand side of (4.5.11) is $\ll_{\epsilon} y^{-1/2+\epsilon}$. We now proceed to show this bound can be improved to $\ll_{\epsilon} y^{-1+\epsilon}$. The function $Z_{\mathcal{A}}(s)y^s/s$ has a double pole at $s = 0$ with residue equal to $C_1 \log y + C_2$, where C_1 and C_2 are real numbers independent of y. By the residue theorem and Mellin inversion,

$$\begin{aligned} \frac{1}{2\pi i} \int_{(-1/2)} Z_{\mathcal{A}}(s) y^s \frac{ds}{s} &= -C_1 \log y - C_2 + \frac{1}{2\pi i} \int_{(1)} Z_{\mathcal{A}}(s) y^s \frac{ds}{s} \\ &= \sum_{n\leq y} \frac{\mathbf{f}_a(n)}{n\gamma(n)} - C_1 \log y - C_2. \end{aligned} \tag{4.5.12}$$

Let us give an elementary estimate for the sum appearing on the right hand side of (4.5.12). Define

$$\nu(n) := \prod_{p|n} \frac{1-\mathbf{h}(p)}{p-1}.$$

Using the convolution identity

$$\frac{1}{\gamma(n)} = \sum_{rs=n} \mu^2(s)\nu(s),$$

we compute

$$\begin{aligned}
\sum_{n\le y}\frac{\mathbf{f}_a(n)}{n\gamma(n)} &= \sum_{s\le y}\frac{\mu^2(s)\nu(s)}{s}\sum_{r\le y/s}\frac{\mathbf{f}_a(rs)}{r} = \sum_{s\le y}\frac{\mu^2(s)\nu(s)}{s}\sum_{(a,s)|d|a}\mathbf{f}_a(d)\sum_{\substack{r\le y/s:\\ (a,rs)=d}}\frac{1}{r}\\
&= \sum_{s\le y}\frac{\mu^2(s)\nu(s)}{s}\sum_{(a,s)|d|a}\mathbf{f}_a(d)\sum_{\substack{r\le y/s:\\ \frac{d}{(d,s)}|r\\ (a,rs)=d}}\frac{1}{r}\\
&= \sum_{s\le y}\frac{\mu^2(s)\nu(s)}{s}\sum_{(a,s)|d|a}\mathbf{f}_a(d)\frac{(d,s)}{d}\sum_{\substack{l\le\frac{y(d,s)}{ds}:\\ (a/d,ls/(d,s))=1}}\frac{1}{l}\\
&= \sum_{s\le y}\frac{\mu^2(s)\nu(s)}{s}\sum_{\substack{(a,s)|d|a:\\ (a/d,s/(d,s))=1}}\mathbf{f}_a(d)\frac{(d,s)}{d}\sum_{\substack{l\le\frac{y(d,s)}{ds}:\\ (l,a/d)=1}}\frac{1}{l}\\
&= \sum_{s\le y}\frac{\mu^2(s)\nu(s)}{s}\sum_{\substack{(a,s)|d|a:\\ (a/d,s/(d,s))=1}}\mathbf{f}_a(d)\frac{(d,s)}{d}\frac{\phi(a/d)}{a/d}\Bigg(\log\left(\frac{y(d,s)}{ds}\right)+\gamma\\
&\qquad +\sum_{p|a/d}\frac{\log p}{p-1}+O\left(\frac{ds}{y(d,s)}\right)\Bigg).
\end{aligned} \tag{4.5.13}$$

Using the bound $\nu(n)\ll_\epsilon n^{-1+\epsilon}$, which is deduced from Hypothesis 4.3.3, we get that the error terms sum to $O_{a,\epsilon}(y^{-1+\epsilon})$. Moreover, we can extend the sum over $s\le y$ to all integers, at the cost of the error term $O_{a,\epsilon}(y^{-1+\epsilon})$. Having done this, (4.5.13) becomes

$$\sum_{n\le y}\frac{\mathbf{f}_a(n)}{n\gamma(n)} = \tilde{C}_1\log y+\tilde{C}_2+O_{a,\epsilon}(y^{-1+\epsilon}), \tag{4.5.14}$$

where $\tilde{C}_1$ and $\tilde{C}_2$ are real numbers which do not depend on y. Substituting (4.5.14) into (4.5.12) and using our previous bound, we get

$$(\tilde{C}_1-C_1)\log y+\tilde{C}_2-C_2+O_{a,\epsilon}(y^{-1+\epsilon}) = \frac{1}{2\pi i}\int_{(-1/2)}Z_{\mathcal{A}}(s)y^s\frac{ds}{s} = O_{a,\epsilon}(y^{-1/2+\epsilon}),$$

which of course implies that $\tilde{C}_1=C_1$ and $\tilde{C}_2=C_2$ since these numbers do not depend on y. We conclude from (4.5.12) and (4.5.14) that (4.5.11) holds.

□

4.5.1.3. *Proof of Proposition 4.5.1*

PROOF OF PROPOSITION 4.5.1. First we use Lemma 4.5.2 to write

$$\begin{aligned} S_5 := &\sum_{1\le r\le \mathbf{R}} \frac{\mathbf{f}_a(r)}{r\gamma(r)}\left(1-\frac{r}{\mathbf{R}}\right) - \sum_{1\le r\le M} \frac{\mathbf{f}_a(r)}{r\gamma(r)}\left(1-\frac{r}{M}\right) - \sum_{\frac{x}{R}<q\le\frac{x}{M}} \frac{\mathbf{f}_a(q)}{q\gamma(q)} \\ = &\frac{1}{2\pi i}\int_{(1)} \mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}} Z_5(s)\left(\frac{\mathbf{R}^s - M^s}{s+1} + \left(\frac{x}{\mathbf{R}}\right)^s - \left(\frac{x}{M}\right)^s\right)\frac{ds}{s}. \end{aligned}$$

Writing

$$\psi(s) := \frac{\mathbf{R}^s - M^s}{s+1} + \left(\frac{x}{\mathbf{R}}\right)^s - \left(\frac{x}{M}\right)^s,$$

it is trivial that $\psi(0) = 0$. Using Taylor series, we have for s close to 0 that

$$\psi(s) = (1+O(s))(s\log(\mathbf{R}/M) + O(s^2)) + s\log(x/\mathbf{R}) - s\log(x/M) + O(s^2),$$

which means that ψ has a double zero at $s = 0$. Thus,

$$\mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}} Z_5(s)\frac{\psi(s)}{s}$$

is holomorphic at $s = 0$. Using this fact,

$$\begin{aligned} S_5 &= \frac{1}{2\pi i}\int_{(-1/2)} \mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}} Z_5(s)\psi(s)\frac{ds}{s} \\ &= \frac{1}{2\pi i}\int_{(-1/2)} \mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^{1-\mathbf{k}} Z_5(s)(\mathbf{R}^s - M^s)\frac{ds}{s(s+1)} \\ &\qquad + O_\epsilon\left(\left(\frac{\mathbf{R}}{x}\right)^{1-\epsilon}\right) \end{aligned}$$

by Lemma 4.5.9. We conclude using propositions 4.5.2 and 4.5.3 that

$$\begin{aligned} S_5 &= \frac{\mu_{\mathbf{k}}(a,M)}{M}\left(1+O\left(\frac{\log\log M}{\log M}\right)\right) + O_A\left(\frac{1}{M\log^A M}\right) \\ &\qquad - \frac{\mu_{\mathbf{k}}(a,\mathbf{R})}{\mathbf{R}}\left(1+O\left(\frac{\log\log \mathbf{R}}{\log \mathbf{R}}\right)\right) \\ &\qquad\qquad + O_A\left(\frac{1}{\mathbf{R}\log^A \mathbf{R}}\right) + O_\epsilon\left(\left(\frac{\mathbf{R}}{x}\right)^{1-\epsilon}\right) \\ &= \frac{\mu_{\mathbf{k}}(a,M)}{M}\left(1+O\left(\frac{\log\log M}{\log M}\right)\right) + O_{A,\delta}\left(\frac{1}{M\log^A M}\right), \end{aligned}$$

since $M(x)^{1+\delta} \leq \mathbf{L}(x)^{1+\delta} \leq \mathbf{R}(x) \leq \sqrt{x}$. □

4.5.2. Proofs of theorems 4.4.1 and 4.4.1*

We first define the following counting function, which will come in handy for the proofs of this section :

$$\mathcal{A}^*(x;q,a) := \sum_{\substack{|a|<n\leq x\\ n\equiv a \bmod q}} \mathbf{a}(n). \tag{4.5.15}$$

Proof of Theorem 4.4.1. Let $1 \leq M(x) \leq \mathbf{L}(x)$ and let $\mathbf{R} = \mathbf{R}(x)$ be as in Hypothesis 4.3.1. We decompose the sum (4.4.1) as follows :

$$\begin{aligned}
&\sum_{q\leq \frac{x}{M}} \left(\mathcal{A}(x;q,a) - \mathbf{a}(a) - \frac{\mathbf{f}_a(q)}{q\gamma(q)}\mathcal{A}(x)\right)\\
&= \sum_{q\leq \frac{x}{M}} \left(\mathcal{A}^*(x;q,a) - \frac{\mathbf{f}_a(q)}{q\gamma(q)}\mathcal{A}(x)\right) + O(1)\\
&= \sum_{\frac{x}{\mathbf{R}}<q\leq x} \mathcal{A}^*(x;q,a) - \sum_{\frac{x}{M}<q\leq x} \mathcal{A}^*(x;q,a) - \mathcal{A}(x) \sum_{\frac{x}{\mathbf{R}}<q\leq \frac{x}{M}} \frac{\mathbf{f}_a(q)}{q\gamma(q)} \qquad (4.5.16)\\
&\qquad + \sum_{q\leq \frac{x}{\mathbf{R}}} \left(\mathcal{A}^*(x;q,a) - \frac{\mathbf{f}_a(q)}{q\gamma(q)}\mathcal{A}(x)\right) + O(1)\\
&= S_1 - S_2 - S_3 + S_4 + O(1).
\end{aligned}$$

Hypothesis 4.3.4 implies the bound

$$S_4 \ll \frac{\mathcal{A}(x)}{M(x)^{1+\delta}}.$$

To evaluate the sums S_1 and S_2 we use the Hooley-Montgomery divisor switching technique (see [**39**]). Setting $n = a + qr$, we have for positive a that

$$\begin{aligned}
S_2 = \sum_{\frac{x}{M}<q\leq x} \sum_{\substack{|a|<n\leq x\\ n\equiv a \bmod q}} \mathbf{a}(n) &= \sum_{1\leq r<(x-a)\frac{M}{x}} \sum_{\substack{a+r\frac{x}{M}<n\leq x\\ n\equiv a \bmod r}} \mathbf{a}(n)\\
&= \sum_{1\leq r<(x-a)\frac{M}{x}} \left(\mathcal{A}(x;r,a) - \mathcal{A}\left(a + r\frac{x}{M};r,a\right)\right). \qquad (4.5.17)
\end{aligned}$$

Using Hypothesis 4.3.1, we see that there exists $\delta > 0$ such that

$$\begin{aligned} S_2 &= \sum_{1 \le r < (x-a)\frac{M}{x}} \frac{\mathbf{f}_a(r)}{r\gamma(r)} \left(\mathcal{A}(x) - \mathcal{A}\left(a + r\frac{x}{M}\right)\right) + O\left(\frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+2\delta}}\right) \\ &= \sum_{1 \le r < (x-a)\frac{M}{x}} \frac{\mathbf{f}_a(r)}{r\gamma(r)} \left(\mathcal{A}(x) - \mathcal{A}\left(\frac{r}{M}x\right)\right) + O\left(\frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}\right) \qquad (4.5.18) \\ &= \mathcal{A}(x) \sum_{1 \le r < (x-a)\frac{M}{x}} \frac{\mathbf{f}_a(r)}{r\gamma(r)} \left(1 - \frac{\mathcal{A}\left(\frac{r}{M}x\right)}{\mathcal{A}(x)}\right) + O\left(\frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}\right) \end{aligned}$$

by hypotheses 4.3.2 and Lemma 4.5.1. Now, if a were negative, we would have to add an error term of size $\ll \frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}$ to (4.5.17) (by Hypothesis 4.3.2), which would yield the same error term in (4.5.18). Using Hypothesis 4.3.2 again, (4.5.18) becomes

$$= \mathcal{A}(x) \sum_{1 \le r < (x-a)\frac{M}{x}} \frac{\mathbf{f}_a(r)}{r\gamma(r)} \left(1 - \frac{r}{M}\right) + O\left(\frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}\right).$$

If M is an integer, then the M-th term of the sum is $\frac{\mathbf{f}_a(r)}{r\gamma(r)}\left(1 - \frac{M}{M}\right) = 0$. If not, the bound $\frac{\mathbf{f}_a(r)}{r\gamma(r)} \ll_\epsilon \frac{1}{\phi(r)}$ (see Lemma 4.5.1) implies that this last term is $\ll \mathcal{A}(x)\frac{\log\log M}{M^2}$. Thus,

$$S_2 = \mathcal{A}(x) \sum_{1 \le r \le M} \frac{\mathbf{f}_a(r)}{r\gamma(r)} \left(1 - \frac{r}{M}\right) + O\left(\frac{\mathcal{A}(x)}{M^{1+\delta}}\right)$$

since $M(x) \le \mathbf{L}(x)$. A similar calculation shows that

$$S_1 = \mathcal{A}(x) \sum_{1 \le r \le \mathbf{R}(x)} \frac{\mathbf{f}_a(r)}{r\gamma(r)} \left(1 - \frac{r}{\mathbf{R}(x)}\right) + O\left(\frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}\right).$$

Grouping terms, (4.5.16) becomes

$$\begin{aligned} &\sum_{q \le \frac{x}{M}} \left(\mathcal{A}(x; q, a) - \mathbf{a}(a) - \frac{\mathbf{f}_a(q)}{q\gamma(q)}\mathcal{A}(x)\right) = S_1 - S_2 - S_3 + S_4 + O(1) \\ &\quad = \mathcal{A}(x) \left(\sum_{1 \le r \le \mathbf{R}} \frac{\mathbf{f}_a(r)}{r\gamma(r)} \left(1 - \frac{r}{\mathbf{R}}\right) - \sum_{1 \le r \le M} \frac{\mathbf{f}_a(r)}{r\gamma(r)} \left(1 - \frac{r}{M}\right) - \sum_{\frac{x}{\mathbf{R}} < q \le \frac{x}{M}} \frac{\mathbf{f}_a(q)}{q\gamma(q)} \right) \\ &\qquad\qquad + O\left(\frac{\mathcal{A}(x)}{M^{1+\delta}}\right), \end{aligned}$$

which combined with Proposition 4.5.1 gives

$$= \frac{\mathcal{A}(x)}{M}\mu_{\mathbf{k}}(a,M)\left(1+O\left(\frac{\log\log M}{\log M}\right)\right)+O_A\left(\frac{\mathcal{A}(x)}{M\log^A M}\right),$$

that is

$$\sum_{q\leq\frac{x}{M}}\left(\mathcal{A}(x;q,a)-\mathbf{a}(a)-\frac{\mathbf{f}_a(q)}{q\gamma(q)}\mathcal{A}(x)\right)$$
$$=\frac{\mathcal{A}(x)}{M}\left(\mu_{\mathbf{k}}(a,M)\left(1+O\left(\frac{\log\log M}{\log M}\right)\right)+O_A\left(\frac{1}{\log^A M}\right)\right).$$

□

Proof of Theorem 4.4.1*. Let $1\leq M(x)\leq \mathbf{L}(x)$ and let $\mathbf{R}=\mathbf{R}(x)$ be as in Hypothesis 4.3.4. We decompose the sum (4.4.2) as follows :

$$\begin{aligned}
&\sum_{\frac{x}{2M}<q\leq\frac{x}{M}}\left(\mathcal{A}(x;q,a)-\mathbf{a}(a)-\frac{\mathbf{f}_a(q)}{q\gamma(q)}\mathcal{A}(x)\right)\\
&=\sum_{\frac{x}{2M}<q\leq\frac{x}{M}}\left(\mathcal{A}^*(x;q,a)-\frac{\mathbf{f}_a(q)}{q\gamma(q)}\mathcal{A}(x)\right)+O(1)\\
&=\sum_{\frac{x}{2M}<q\leq x}\mathcal{A}^*(x;q,a)-\sum_{\frac{x}{M}<q\leq x}\mathcal{A}^*(x;q,a)-\mathcal{A}(x)\sum_{\frac{x}{2M}<q\leq\frac{x}{M}}\frac{\mathbf{f}_a(q)}{q\gamma(q)}+O(1)\\
&=S_1-S_2-S_3+O(1).
\end{aligned} \tag{4.5.19}$$

Arguing as in the proof of Theorem 4.4.1, we set $n=a+qr$ to get that for positive a,

$$\begin{aligned}
S_2&=\sum_{1\leq r<(x-a)\frac{M}{x}}\left(\mathcal{A}(x;r,a)-\mathcal{A}\left(a+r\frac{x}{M};r,a\right)\right)\\
&=\mathcal{A}(x)\sum_{1\leq r<(x-a)\frac{M}{x}}\frac{\mathbf{f}_a(r)}{r\gamma(r)}\left(1-\frac{\mathcal{A}\left(\frac{r}{M}x\right)}{\mathcal{A}(x)}\right)+O\left(\frac{\mathcal{A}(x)}{\mathbf{L}(x)^{1+\delta}}\right)\\
&=\mathcal{A}(x)\sum_{1\leq r\leq M}\frac{\mathbf{f}_a(r)}{r\gamma(r)}\left(1-\frac{r}{M}\right)+O\left(\frac{\mathcal{A}(x)}{M^{1+\delta}}\right)
\end{aligned}$$

by Hypotheses 4.3.1*, 4.3.2 and Lemma 4.5.1. Now, if a were negative, we would have to add a negligible contribution. Thus, (4.5.19) becomes

$$\sum_{\frac{x}{2M}<q\leq\frac{x}{M}}\left(\mathcal{A}(x;q,a)-\mathbf{a}(a)-\frac{\mathbf{f}_a(q)}{q\gamma(q)}\mathcal{A}(x)\right)=\mathcal{A}(x)\Bigg(\sum_{1\leq r\leq 2M}\frac{\mathbf{f}_a(r)}{r\gamma(r)}\left(1-\frac{r}{2M}\right)$$

$$-\sum_{1\leq r\leq M}\frac{\mathbf{f}_a(r)}{r\gamma(r)}\left(1-\frac{r}{M}\right)-\sum_{\frac{x}{2M}<q\leq\frac{x}{M}}\frac{\mathbf{f}_a(q)}{q\gamma(q)}\Bigg)+O\left(\frac{\mathcal{A}(x)}{M^{1+\delta}}\right).$$

Going through the proof of Proposition 4.5.1, we see that this is

$$=\frac{\mathcal{A}(x)}{M}\mu_{\mathbf{k}}(a,M)\left(1+O\left(\frac{\log\log M}{\log M}\right)\right)$$

$$-\frac{\mathcal{A}(x)}{2M}\mu_{\mathbf{k}}(a,2M)\left(1+O\left(\frac{\log\log M}{\log M}\right)\right)+O_A\left(\frac{\mathcal{A}(x)}{M\log^A M}\right),$$

that is

$$\sum_{\frac{x}{2M}<q\leq\frac{x}{M}}\left(\mathcal{A}(x;q,a)-\mathbf{a}(a)-\frac{\mathbf{f}_a(q)}{q\gamma(q)}\mathcal{A}(x)\right)$$

$$=\frac{\mathcal{A}(x)}{2M}\left(\mu_{\mathbf{k}}(a,M)\left(1+O\left(\frac{\log\log M}{\log M}\right)\right)+O_A\left(\frac{1}{\log^A M}\right)\right),$$

since by the definition of $\mu_{\mathbf{k}}(a,M)$,

$$2\mu_{\mathbf{k}}(a,M)-\mu_{\mathbf{k}}(a,2M)=\mu_{\mathbf{k}}(a,M)\left(1+O\left(\frac{1}{\log M}\right)\right).$$

□

4.6. Further Proofs

In this section we prove the results of Section 4.2.

Proof of Theorem 4.2.1. Put

$$a(n):=\Lambda(n),$$

which gives $\mathcal{A}(x)=\psi(x)$ and $\mathcal{A}(x;q,a)=\psi(x;q,a)$. Define

$$\mathbf{f}_a(q):=\begin{cases}1 & \text{if } (a,q)=1\\ 0 & \text{otherwise,}\end{cases}$$

and $\gamma(q) := \frac{\phi(q)}{q}$. Define also the multiplicative function $\mathbf{h}(d)$ by $\mathbf{h}(1) = 1$, and $\mathbf{h}(d) = 0$ for $d > 1$. The prime number theorem in arithmetic progressions gives the asymptotic

$$\mathcal{A}(x; q, a) \sim \frac{\mathbf{f}_a(q)}{q\gamma(q)}\mathcal{A}(x),$$

for any fixed a and q such that $(a, q) = 1$. Now let us show that the hypotheses of Section 4.3.2 hold. Fix $A > 0$ and put $\mathbf{L}(x) := (\log x)^A$, $\mathbf{R}(x) := x^{1/2}(\log x)^{-B(A)}$, where $B(A) := A + 5$. Hypothesis 4.3.1 is the Bombieri-Vinogradov theorem. Hypothesis 4.3.2 follows from the prime number theorem. As $\mathbf{h}(p) = \mathbf{k} = 0$, Hypothesis 4.3.3 is trivial. Hypothesis 4.3.4 follows from Theorem 9 of **[10]**.

We now compute $\mu_{\mathbf{k}}(a, M)$. As $\mathbf{h}(p^e) = 0$, we have $\omega_{\mathbf{h}}(a) = \omega(a)$, the number of distinct prime factors of a. Thus, Remark 4.4.3 gives

$$\mu_0(a, M) = \begin{cases} -\frac{1}{2}\log M & \text{if } a = \pm 1, \\ -\frac{1}{2}\left(1 - \frac{1}{p}\right)\log p & \text{if } a = \pm p^e \\ 0 & \text{if } \omega(a) \geq 2 \end{cases}$$

so an application of Theorem 4.4.1 gives the result with a weaker error term. A better version of Proposition 4.5.1 follows from Huxley's subconvexity result **[41]**, yielding the stated error term (see **[23]** for a more precise proof). □

Proof of Theorem 4.2.2. Let $Q(x, y) := \alpha x^2 + \beta xy + \gamma y^2$ be a binary quadratic form, where α, β and γ are integers such that $\alpha > 0$, $(\alpha, \beta, \gamma) = 1$ and $d := \beta^2 - 4\alpha\gamma < 0$ (so $Q(x, y)$ is positive definite). Note that the set of d for which $d \equiv 1, 5, 9, 12, 13 \bmod 16$ includes a large subset of all fundamental discriminants. The set of bad primes is $\mathcal{S} := \{p : p \mid 2d\}$ in this case. Since $\mathcal{S} \neq \emptyset$, we will need to modify the proof of Theorem 4.4.1. We define

$$\chi_d := \left(\frac{4d}{\cdot}\right).$$

Note that for $(n, 2d) = 1$, we have the equalities

$$r_d(n) = \sum_{m|n} \chi_d(m) = \prod_{\substack{p^k \| n: \\ \chi_d(p)=1}} (k+1) \prod_{\substack{p^k \| n: \\ \chi_d(p)=-1, \\ k \text{ odd}}} 0. \tag{4.6.1}$$

An intuitive argument suggests that

$$\mathcal{A}(x; q, a) \sim \frac{R_a(q)}{q^2} \mathcal{A}(x),$$

where

$$R_a(q) := \#\{1 \leq x, y \leq q : Q(x, y) \equiv a \bmod q\}. \tag{4.6.2}$$

As this is a classical result, we leave its proof, as well as several other classical facts about binary quadratic forms, to Appendix A. The function

$$\mathbf{g}_a(q) := \frac{R_a(q)}{q^2}$$

is actually multiplicative (see Lemma A.0.2), and Lemma A.0.4 shows that for $p \nmid 2d$, $\mathbf{g}_a$ is given as in (4.3.1) with

$$\mathbf{h}(p^e) := \begin{cases} 1 + e\left(1 - \frac{1}{p}\right) & \text{if } \chi_d(p) = 1 \\ \frac{1}{p} & \text{if } \chi_d(p) = -1 \text{ and } 2 \nmid e \\ 1 & \text{if } \chi_d(p) = -1 \text{ and } 2 \mid e, \end{cases}$$

and for $p \mid 2d$, $R_a(p^e)$ is given as in (A.0.20) and (A.0.21). Since we are looking at large moduli, we need to use a result of Plaksin (Lemma 8 of **[65]**), which asserts that

$$\mathcal{A}(x; q, a) = \mathbf{g}_a(q) \mathcal{A}(x) + E(x, q), \tag{4.6.3}$$

where $E(x, q) \ll_{a,\epsilon} (x/q)^{\frac{3}{4}+\epsilon}$ if $q \leq x^{\frac{1}{3}}$, and $E(x, q) \ll_{a,\epsilon} x^{\frac{2}{3}+\epsilon} q^{-\frac{1}{2}}$ if $x^{\frac{1}{3}} < q \leq x^{\frac{2}{3}}$. Summing (4.6.3) over $q \leq x^{\frac{1}{2}}$, we get that the hypotheses 4.3.1 and 4.3.4 hold with $\mathbf{R}(x) := x^{\frac{1}{2}}$ and $\mathbf{L}(x) := x^{\lambda}$, provided $\lambda < \frac{1}{12}$. (Note that in the case $\beta = 0$, we can take the wider range $\lambda < \frac{1}{8}$, using Lemma 20 of **[64]**.) Hypothesis 4.3.2 follows from Gauss' estimate :

$$\mathcal{A}(x) = A_Q x + O(x^{\frac{1}{2}}),$$

where A_Q is the area of the region $\{(x,y) \in \mathbb{R}^2_{\geq 0} : Q(x,y) \leq 1\}$. Let us turn to Hypothesis 4.3.3. For $p \nmid 2d$,

$$\mathbf{h}(p) = \begin{cases} 2 - \frac{1}{p} & \text{if } \chi_d(p) = 1 \\ \frac{1}{p} & \text{if } \chi_d(p) = -1, \end{cases}$$

so we set $\mathbf{k} := 1$ and

$$\sum_{p \notin \mathcal{S}} \frac{\mathbf{h}(p) - \mathbf{k}}{p} = \sum_{p \nmid 2d} \frac{\chi_{-d}(p)}{p} + O(1) < \infty$$

by the prime number theorem for $\psi(x, \chi_{-d})$ (see [16]). Moreover,

$$\begin{aligned} \sum_{p \notin \mathcal{S}} \frac{(\mathbf{h}(p) - \mathbf{k})(\log p)^{n+1}}{p^{1+it}} &= O(1) + (-1)^{n+1} \left(\frac{L'}{L}\right)^{(n)} (1 + it, \chi_{-d}) \\ &\ll_{d,n} (\log(|t| + 2))^{n+2}, \end{aligned} \tag{4.6.4}$$

this last bound following from Cauchy's formula for the derivatives combined with the classical bound for $\frac{L'(s,\chi)}{L(s,\chi)}$ in a zero-free region (see Chapter 19 of [16]). As in the proof of Théorème II.3.22 of [76], we can deduce from (4.6.4) that (setting $\eta := 1/\log^2(|t| + 2)$)

$$\begin{aligned} \sum_{p \notin \mathcal{S}} \frac{\mathbf{h}(p) - \mathbf{k}}{p^{1+it}} + O(1) &= \log L(1 + it, \chi_{-d}) \\ &= \int_{1+it+\eta}^{1+it} \frac{L'(s, \chi_{-d})}{L(s, \chi_{-d})} ds + \log L(1 + it + \eta, \chi_{-d}) \\ &\ll \eta \log^2(|t| + 2) + \log \zeta(1 + \eta) = 2 \log\log(|t| + 2) + O(1). \end{aligned}$$

Having proven hypotheses 4.3.1, 4.3.2, 4.3.3 and 4.3.4, we now proceed to prove an analogue of Theorem 4.4.1 (since $\mathcal{S} = \{p : p \mid 2d\}$ is non-empty). In the proof of Lemma 4.5.2, we need to change the definition of $\mathfrak{S}_2(s) (= \mathfrak{S}_1(s))$ to (remember that $(a, 2d) = 1$)

$$\begin{aligned} \mathfrak{S}_2(s) = &\left(\left(1 - \frac{1}{2^{s+1}}\right)\left(1 + \frac{R_a(2)}{2^{s+2}}\right) + \frac{R_a(4)}{4} \frac{1}{2^{2s+2}}\right) \prod_{\substack{p \mid d \\ p \neq 2}} \left(1 - \frac{1}{p^{s+1}} + \frac{R_a(p)}{p^{s+2}}\right) \\ &\times \prod_{\substack{p^f \| a \\ f \geq 1 \\ p \notin \mathcal{S}}} \left[\left(1 + \frac{\mathbf{h}(p)}{p^{s+1}} + \ldots + \frac{\mathbf{h}(p^f)}{p^{f(s+1)}}\right)\left(1 - \frac{1}{p^{s+1}}\right) + \frac{\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p}{1 - 1/p} \frac{1}{p^{(f+1)(s+1)}}\right], \end{aligned}$$

(We also need to change the condition on the product defining $Z_5(s)$ to $p \nmid 2ad$) so

$$\begin{aligned}\mathfrak{S}_2(-1) &= \frac{R_a(4)}{4} \prod_{\substack{p \mid d \\ p \neq 2}} \frac{R_a(p)}{p} \prod_{\substack{p^f \| a \\ f \geq 1 \\ p \notin \mathcal{S}}} \frac{\mathbf{h}(p^f) - \mathbf{h}(p^{f+1})/p}{1 - 1/p} \\ &= \frac{R_a(4)}{4} \prod_{\substack{p^f \| d \\ p \neq 2}} \frac{R_a(p^f)}{p^f} \prod_{\substack{p^f \| a: \\ \chi_d(p) = 1}} \left(1 - \frac{1}{p}\right)(f+1) \prod_{\substack{p^f \| a: \\ \chi_d(p) = -1, \\ f \text{ even}}} \left(1 + \frac{1}{p}\right) \prod_{\substack{p^f \| a: \\ \chi_d(p) = -1, \\ f \text{ odd}}} 0 \\ &= \frac{R_a(4d)}{4d} \prod_{p \mid a} \left(1 - \frac{\chi_d(p)}{p}\right) r_d(|a|),\end{aligned}$$

by (4.6.1) and Lemma A.0.4. We conclude that Theorem 4.4.1 holds with

$$\mu_1(a, M) = -\frac{R_a(4d)}{4d} \cdot \frac{r_d(|a|)}{2L(1, \chi_d)},$$

which gives the result (with a weaker error term) by Dirichlet's class number formula. To get the better error term $O_\epsilon\left(\frac{1}{M^{1/3-\epsilon}}\right)$, one has to get a better estimate in Proposition 4.5.1. To do this, we go back to the proof of Proposition 4.5.2 and remark that (with the notation introduced there)

$$Z_5(s) = \prod_{p \nmid 2ad} \left(1 - \frac{\chi_d(p)}{p^{s+2}}\right),$$

so

$$Z(s) = \frac{\mathfrak{S}_3(s)\zeta(s+1)L(s+2, \chi_d)}{s(s+1)},$$

where

$$\mathfrak{S}_3(s) := \mathfrak{S}_2(s) \prod_{p \mid 2ad} \left(1 - \frac{\chi_d(p)}{p^{s+2}}\right)^{-1}.$$

Since $Z(s)$ is a meromorphic function on the whole complex plane, we can shift the contour of integration to the left until the line $\mathrm{Re}(s) = -\frac{4}{3} + \epsilon$. We have the following convexity bound on $L(s, \chi_d)$, for $0 \leq \sigma \leq 1$ and $t \in \mathbb{R}$:

$$L(\sigma + it, \chi_d) \ll_\epsilon (d(|t| + 3))^{\frac{1-\sigma}{2} + \epsilon}$$

(see (5.20) of [**44**]). Combining this bound with a standard residue calculation yields that

$$\frac{1}{2\pi i} \int_{(-1/2)} \frac{\mathfrak{S}_3(s)\zeta(s+1)L(s+2, \chi_d)}{s(s+1)} M^s ds = -\frac{\mu_1(a, M)}{M} + O_\epsilon\left(\frac{1}{M^{4/3-\epsilon}}\right),$$

from which we conclude the result. □

PROOF OF THEOREM 4.2.3. Set $\mathcal{S} := \{2\}$, $\mathbf{k} := \frac{1}{2}$ and $\mathbf{L}(x) := (\log x)^{\lambda}$ with $\lambda < 1/5$. We first prove Hypothesis 4.3.2 using a refinement of a theorem of Landau. We have

$$\mathcal{A}(x) = C\frac{x}{\sqrt{\log x}}\left(1 + O\left(\frac{x}{\log x}\right)\right), \tag{4.6.5}$$

with

$$C := \frac{1}{\sqrt{2}}\prod_{p\equiv 3 \bmod 4}\left(1-\frac{1}{p^2}\right)^{-\frac{1}{2}}.$$

(See for instance Exercice 240 of [**76**]). The distribution of $\mathcal{A}$ in the arithmetic progressions a mod q with $(a,q) = 1$ is uniform, however a result of the strength of Plaksin's (4.6.3) is far from being known. The best result so far for individual values of q (in terms of uniformity in q) is due to Iwaniec [**43**], which proved using the semi-linear sieve that if $(a,q) = 1$ and $a \equiv 1$ mod $(q,4)$, then

$$\mathcal{A}(x;q,a) = \frac{(2,q)}{(4,q)q\gamma(q)}\mathcal{A}(x)\left(1 + O\left(\frac{\log q}{\log x}\right)^{1/5}\right), \tag{4.6.6}$$

where

$$\gamma(q) := \prod_{\substack{p|q\\ p\equiv 3 \bmod 4}}\left(1+\frac{1}{p}\right)^{-1}.$$

An easy computation using the arithmetic properties of $\mathbf{a}(n)$ shows that

$$\mathcal{A}_{p^e}(x) = \begin{cases} \mathcal{A}\left(\frac{x}{p^{e+1}}\right) & \text{if } p \equiv 3 \mod 4 \text{ and } 2 \nmid e \\ \mathcal{A}\left(\frac{x}{p^{e}}\right) & \text{otherwise,} \end{cases}$$

and more generally,

$$\mathcal{A}_d(x) = \mathcal{A}\left(\frac{\mathbf{h}(d)}{d}x\right), \tag{4.6.7}$$

with

$$\mathbf{h}(p^e) := \begin{cases} \frac{1}{p} & \text{if } p \equiv 3 \mod 4 \text{ and } 2 \nmid e \\ 1 & \text{otherwise.} \end{cases}$$

This confirms that our choice of $\mathbf{k} = \frac{1}{2}$ was good, and Hypothesis 4.3.3 follows as in the proof of Theorem 4.2.2. Moreover, (4.6.6) can be extended to

$(a, q) = d$ for any fixed odd integer $d > 1$, by using the identity $\mathcal{A}(x; q, a) = \mathcal{A}\left(\frac{\mathbf{h}(d)}{d}x; \frac{q}{d}, \frac{a}{d}\right)$, hence Hypothesis 4.3.1* holds. As we have shown every hypothesis, we turn to the calculation of the average $\mu_{\frac{1}{2}}(a, M)$ (which is never zero since $\mathbf{k} \notin \mathbb{Z}$). We need to modify the definition of $\mathfrak{S}_2(s)$, changing the local factor at $p = 2$ to

$$\left(1 - \frac{1}{2^{s+2}}\right)^{1/2}\left(1 - \frac{1}{2^{2s+2}} + \frac{1}{2^{2s+3}}\right).$$

Doing so and proceeding as in the proof of Theorem 4.4.1*, we get the result.

□

Lemma 4.6.1. *Suppose that* $\mathcal{H} = \{a_1 n + b_1, \dots a_k n + b_k\}$ *is an admissible* k*-tuple of linear forms and* q, a *are two integers such that* $(q, a_i a + b_i) = 1$ *for* $1 \le i \le k$*. Then the modified* k*-tuple* $\tilde{\mathcal{H}} := \{a_1(qm + a) + b_1, \dots a_k(qm + a) + b_k\}$ *is also admissible. Moreover,*

$$\mathfrak{S}(\tilde{\mathcal{H}}) = \prod_{p \mid q}\left(1 - \frac{\nu_{\mathcal{H}}(p)}{p}\right)^{-1}\mathfrak{S}(\mathcal{H}).$$

Proof. First, since $\mathcal{H}$ is admissible, we have $(a_i, b_i) = 1$ for $1 \le i \le k$. Fix a prime p. We need to show that $\nu_{\tilde{\mathcal{H}}}(p) < p$. For a fixed i we have either $p \mid a_i$, in which case $p \nmid b_i$ so $a_i n + b_i \not\equiv 0 \bmod p$, or $p \nmid a_i$, in which case the only solution to $a_i n + b_i \equiv 0 \mod p$ is $n \equiv -a_i^{-1} b_i$. Hence, if $p \nmid a_i$, then there are only $\nu_{\mathcal{H}}(p) < p$ distinct possible values for $-a_i^{-1} b_i \bmod p$, thus regrouping these we can write

$$\prod_{i=1}^{k}(a_i n + b_i) \equiv C \prod_{i: p \mid a_i} b_i \prod_{j=1}^{\nu_{\mathcal{H}}(p)} (n + k_j)^{e_j} \bmod p,$$

where the k_j are distinct integers, $e_j \ge 1$ and $p \nmid C$. Using this and the fact that $(a_i, b_i) = 1$, we get

$$\prod_{i=1}^{k}(a_i(qm + a) + b_i) \equiv D \prod_{j=1}^{\nu_{\mathcal{H}}(p)} (qm + a + k_j)^{e_j} \bmod p,$$

with $p \nmid D$. If $p \nmid q$, then this has exactly $\nu_{\mathcal{H}}(p) < p$ solutions, therefore $\nu_{\tilde{\mathcal{H}}}(p) < p$. Otherwise, this becomes

$$\prod_{i=1}^{k}(a_i(qm+a)+b_i) \equiv \prod_{i=1}^{k}(a_i a+b_i) \not\equiv 0 \bmod p$$

since $(q, a_i a + b_i) = 1$ for $1 \leq i \leq k$. We conclude that $\tilde{\mathcal{H}}$ is admissible. The calculation of $\mathfrak{S}(\tilde{\mathcal{H}})$ follows easily. □

PROOF. [Proof of Theorem 4.2.4] Define $\mathcal{S} := \emptyset$ and

$$\mathbf{a}(n) := \prod_{\mathcal{L}\in\mathcal{H}} \Lambda(\mathcal{L}(n)) = \Lambda(a_1 n + b_1)\Lambda(a_2 n + b_2)\cdots\Lambda(a_k n + b_k).$$

In our context, some assumptions of Section 4.3.1 do not hold. The reason is that the asymptotic for $\mathcal{A}(x; q, a)$ depends on $(q, \mathcal{P}(a; \mathcal{H}))$ rather than depending only on (q, a). The correct conjecture in this case is that for integers a and q such that $(q, \mathcal{P}(a; \mathcal{H})) = 1$, (see **[53]**[2])

$$\mathcal{A}(x; q, a) \sim \frac{\mathcal{A}(x)}{q\gamma(q)},$$

with

$$\gamma(q) := \prod_{p|q}\left(1 - \frac{\nu_{\mathcal{H}}(p)}{p}\right).$$

This actually follows from the Hardy-Littlewood conjecture, by taking the modified k-tuple of linear forms $\tilde{\mathcal{L}}_i(m) := a_i(qm+a)+b_i = qa_i m + aa_i + b_i$, which is admissible if $\mathcal{H}$ is and $(q, \mathcal{P}(a; \mathcal{H})) = 1$ (see Lemma 4.6.1). Using this idea, we get that the assumption of (4.2.4) holding uniformly for $|a_i| \leq \mathbf{L}(x)^{1+\delta}$ implies Hypothesis 4.3.1*. We now prove an analogue of Proposition 4.5.1. Defining

$$Z_{\mathcal{H}}(s) := \sum_{n} \frac{\mathbf{f}_a(n)}{n^{s+1}\gamma(n)},$$

[2] Kawada imposes the additional condition that $R(\mathbf{b}) := \prod_{j=1}^{k} |a_j| \prod_{1\leq i,j\leq k} |a_i b_j - a_j b_i|$ is non-zero. However, we assume that our linear forms are admissible, distinct and $a_i \geq 1$; one can show that this implies $R(\mathbf{b}) \neq 0$.

where

$$\mathbf{f}_a(q) := \begin{cases} 1 & \text{if } (\mathcal{P}(a;\mathcal{H}), q) = 1 \\ 0 & \text{otherwise,} \end{cases}$$

one can compute that

$$Z_{\mathcal{H}}(s) = \mathfrak{S}_2(s)\zeta(s+1)\zeta(s+2)^k Z_0(s)$$

with

$$\mathfrak{S}_2(s) := \prod_{p \mid \mathcal{P}(a;\mathcal{H})} \left(1 - \frac{1}{p^{s+1}}\right)\left(1 + \frac{\nu_{\mathcal{H}}(p)}{p - \nu_{\mathcal{H}}(p)}\frac{1}{p^{s+1}}\right)^{-1}$$

and

$$Z_0(s) := \prod_p \left(1 + \frac{\nu_{\mathcal{H}}(p)}{p - \nu_{\mathcal{H}}(p)}\frac{1}{p^{s+1}}\right)\left(1 - \frac{1}{p^{s+2}}\right)^k$$

which converges for $\operatorname{Re} s > -3/2$. Note that $Z_{\mathcal{H}}(s)$ has a simple pole at $s = 0$. Also, $\mathfrak{S}_2(s)$ has a zero of order $\omega(\mathcal{P}(a, \mathcal{H}))$ at the point $s = -1$, and $Z_0(-1) = \mathfrak{S}(\mathcal{H})^{-1}$, so $Z_{\mathcal{H}}(s)$ is of order $\omega(\mathcal{P}(a, \mathcal{H})) - k$ at this point. The function

$$\psi(s) := \frac{(2M)^s - M^s}{s+1} + \left(\frac{x}{2M}\right)^s - \left(\frac{x}{M}\right)^s$$

vanishes to the second order at $s = 0$. Combining all this information, we obtain by shifting the contour of integration to the left that

$$\frac{1}{2\pi i}\int_{(1)} Z_{\mathcal{H}}(s)\psi(s)\frac{ds}{s} = \frac{1}{2M}\left(\mu_{1-k}(a, M)(1 + o(1)) + O\left(\frac{1}{M^{\delta_k}}\right)\right),$$

where

$$\mu_{1-k}(a, M) := \\ \begin{cases} -\dfrac{1}{2\mathfrak{S}(\mathcal{H})}\dfrac{(\log M)^{k-\omega(\mathcal{P}(a;\mathcal{H}))}}{(k - \omega(\mathcal{P}(a;\mathcal{H})))!}\displaystyle\prod_{p \mid \mathcal{P}(a;\mathcal{H})} \frac{p - \nu_{\mathcal{H}}(p)}{p}\log p & \text{if } \omega(\mathcal{P}(a;\mathcal{H})) \le k \\ 0 & \text{otherwise,} \end{cases}$$

and $\delta_k > 0$ is a small real number (one can take $\delta_k = \frac{1}{2+k}$). We conclude by proceeding as in the proof of Theorem 4.4.1*.

□

In the case of twin primes (that is $\mathbf{a}(n) := \Lambda(n)\Lambda(n+2)$), we give an explicit description of all integers $a \geq -1$ (without loss of generality since $-a(-a+2) = a(a-2)$) for which $\mu_{-1}(a, M) \neq 0$ (note the occurrence of Mersenne and Fermat primes) :

a	$a(a+2)$	$\omega(a(a+2))$
-1	-1	0
1	3	1
2	8	1
$p^e, p \neq 2 : p^e + 2 = q^f$	$p^e q^f$	2
$2^e : 2^{e-1} + 1 = q^f$	$2^{e+1}(2^{e-1}+1)$	2
$2^e - 2 : 2^{e-1} - 1 = q^f$	$2^{e+1}(2^{e-1}-1)$	2

PROOF. [Proof of Theorem 4.2.5] Define $\mathcal{S} := \emptyset$ and $\mathbf{L}(x) := (\log x)^{1-\delta}$. We split the proof in two cases, depending on the size of y.

Case 1 : $\log y \leq (\log M)^{\frac{1}{2}-\delta}$. The fundamental lemma of combinatorial sieve (see [**18**]) gives the following estimate, in the range $2 \leq y \leq x^{o(1)}$:

$$\mathcal{A}(x, y) = x \prod_{p \leq y} \left(1 - \frac{1}{p}\right) (1 + E(x, y)), \tag{4.6.8}$$

where $E(x, y) \ll x^{-\frac{1}{3}}$ for $2 \leq y < \frac{(\log x)^2}{16}$, and $E(x, y) \ll u^{-u}(\log y)^3$ for $\frac{(\log x)^2}{16} \leq y \leq x$, with the usual notation $u := \frac{\log x}{\log y}$ (so $y^u = x$). This shows that Hypothesis 4.3.2 holds. One shows that

$$\mathcal{A}_d(x, y) := \sum_{\substack{n \leq x \\ d \mid n}} \mathbf{a}_y(n) = \begin{cases} \mathcal{A}\left(\frac{x}{d}, y\right) & \text{if } p \mid d \Rightarrow p \geq y \\ 0 & \text{else,} \end{cases}$$

so we have $\mathcal{A}_d(x, y) = \mathcal{A}\left(\frac{\mathbf{h}_y(d)}{d} x, y\right)$, where

$$\mathbf{h}_y(d) := \begin{cases} 1 & \text{if } p \mid d \Rightarrow p \geq y \\ 0 & \text{else.} \end{cases}$$

Wolke [**78**] as shown a Bombieri-Vinogradov theorem for this sequence, which states that for any $A > 0$, there exists $B = B(A)$ such that for any $Q \leq$

$x^{\frac{1}{2}}/\log^B x$, we have, uniformly in the range $y \le \sqrt{x}$,

$$\sum_{q \le Q} \max_{(a,q)=1} \max_{z \le x} \left| \sum_{\substack{n \le z \\ n \equiv a \bmod q}} \mathbf{a}_y(n) - \frac{1}{\phi(q)} \sum_{\substack{n \le z \\ (n,q)=1}} \mathbf{a}_y(n) \right| \ll \frac{x}{\log^A x}. \tag{4.6.9}$$

(Notice that if $(a,q) > 1$, then $\mathcal{A}(x,y;q,a)$ is bounded.) We will only use this for $Q = 2\mathbf{L}(x)$, so from now on we suppose that $q \le 2(\log x)^{1-\delta}$. Arguing as in Section 4.3.1, we have for $\frac{x}{2\mathbf{L}(x)} \le z \le x$ that

$$\begin{aligned} \frac{1}{\phi(q)} \sum_{\substack{n \le z \\ (n,q)=1}} \mathbf{a}_y(n) &= \frac{1}{\phi(q)} \sum_{d|q} \mu(d) \mathcal{A}_d(z,y) = \frac{1}{\phi(q)} \sum_{d|q} \mu(d) \mathcal{A}\left(\frac{\mathbf{h}_y(d)}{d} z, y\right) \\ &= \frac{\mathcal{A}(z,y)}{\phi(q)} \sum_{d|q} \frac{\mathbf{h}_y(d)\mu(d)}{d} \left(1 + E_{d;q}(z,y)\right) \\ &= \frac{\mathcal{A}(z,y)}{q\gamma_y(q)} \left(1 + O(x^{-\frac{1}{3}+o(1)})\right), \end{aligned}$$

since by (4.6.8), in the range $d \le q \le (\log x)^{1-\delta}$ we have

$$E_{d;q}(z,y) \ll \left(\frac{d}{z}\right)^{\frac{1}{3}} \ll x^{-\frac{1}{3}+o(1)}.$$

Summing this over $q \le 2\mathbf{L}(x)$ and using (4.6.9), we get that

$$\sum_{q \le 2\mathbf{L}(x)} \max_{(a,q)=1} \max_{\frac{x}{2\mathbf{L}(x)} \le z \le x} \left| \mathcal{A}(z,y;q,a) - \frac{\mathcal{A}(z,y)}{q\gamma_y(q)} \right| \ll \frac{\mathcal{A}(x,y)}{\mathbf{L}(x)^{1+\delta}}. \tag{4.6.10}$$

Having a Bombieri-Vinogradov theorem in hand, we now prove an analogue of Proposition 4.5.1. A straightforward computation shows that

$$\begin{aligned} Z_{\mathcal{A}}(s) &:= \sum_{\substack{n \ge 1 \\ (n,a)=1}} \frac{1}{n^{s+1}\gamma_y(n)} \\ &= \zeta(s+1) \prod_{p|a} \left(1 - \frac{1}{p^{s+1}}\right) \prod_{\substack{p \nmid a \\ p < y}} \left(1 + \frac{1}{(p-1)p^{s+1}}\right) \qquad (4.6.11) \\ &= \mathfrak{S}_a(s)\zeta(s+1)\zeta(s+2)Z_{11}(s) \prod_{p \ge y} \left(1 + \frac{1}{(p-1)p^{s+1}}\right)^{-1}, \qquad (4.6.12) \end{aligned}$$

where

$$\mathfrak{S}_a(s) := \prod_{p|a} \left(1 - \frac{1}{p^{s+1}}\right) \left(1 + \frac{1}{(p-1)p^{s+1}}\right)^{-1},$$

$$Z_{11}(s) := \prod_p \left(1 + \frac{1}{(p-1)p^{s+2}} - \frac{1}{(p-1)p^{2s+3}}\right).$$

We will now use representation (4.6.11). Representation (4.6.12) will be useful for larger values of y, since then $\prod_{p<y}\left(1-\frac{1}{p^{s+2}}\right)^{-1}$ behaves like $\zeta(s+2)$ on the line $\mathrm{Re}(s) = -1 + \frac{1}{\log M}$. Note that by (4.6.11), $Z_{\mathcal{A}}(s)$ is defined on the whole complex plane, except at $s = 0$. As before, we need to compute the integral

$$I := \frac{1}{2\pi i}\int_{(2)} Z_{\mathcal{A}}(s)\psi(s)\frac{ds}{s},$$

where

$$\psi(s) := \frac{(2M)^s - M^s}{s+1} + \left(\frac{x}{2M}\right)^s - \left(\frac{x}{M}\right)^s,$$

which has a double zero at $s = 0$, so

$$\begin{aligned} I &= \frac{1}{2\pi i}\int_{(-1/2)} Z_{\mathcal{A}}(s)\psi(s)\frac{ds}{s} \\ &= \frac{1}{2\pi i}\int_{(-1/2)} Z_{\mathcal{A}}(s)((2M)^s - M^s)\frac{ds}{s(s+1)} + O_{a,\epsilon}\left(\left(\frac{M}{x}\right)^{\frac{1}{2}-\epsilon}\log y\right), \end{aligned}$$

by the same arguments as in Lemma 4.5.9 (and Merten's theorem). We now proceed as in the proof of Proposition 4.5.2. Moving the contour of integration to $\mathrm{Re}(s) = \sigma = -1 + \frac{1}{\log M}$ and using the bounds $Z_{\mathcal{A}}(\sigma + it) \ll_\epsilon (|t|+1)^{\frac{1}{2}+\epsilon}\log y$ and $Z'_{\mathcal{A}}(\sigma + it) \ll_\epsilon (|t|+1)^{\frac{1}{2}+\epsilon}(\log y)^2$ for $|t| \geq 2$ (by Cauchy's theorem for the derivatives), we can deduce that

$$I = \frac{2\mu_y(a,M) - \mu_y(a,2M)}{2M}\left(1 + O_a\left(\frac{(\log y)^2 \log\log M}{\log M}\right)\right) + o(1),$$

where

$$\mu_y(a,M) := \begin{cases} -\frac{1}{2}\prod_{p<y}\left(1-\frac{1}{p}\right)^{-1} & \text{if } a = \pm 1 \\ 0 & \text{else.} \end{cases}$$

We conclude the proof in the same lines as that of Theorem 4.4.1*.

Case 2 : $\mathbf{L}^{(1+\delta)\log\log \mathbf{L}} \leq y \leq \sqrt{x}$. Note that it is sufficient to consider this range, since $\mathbf{L}^{(1+\delta)\log\log \mathbf{L}} < (\log x)^{\log\log\log x}$. We have

$$\mathcal{A}(x,y) = \frac{x\omega(u)}{\log y}\left(1 + O\left(\frac{1}{\log y}\right)\right),$$

where $u := \frac{\log x}{\log y}$ and $\omega(u)$ is Buchstab's function (see Théorème III.6.4 of [**76**]). Therefore, we can use the properties of $\omega(u)$ to show that in the range $\frac{1}{\mathbf{L}(x)} \leq z \leq 1+\delta$,

$$\frac{\mathcal{A}(zx,y)}{\mathcal{A}(x,y)} = z\frac{\omega\left(u - O\left(\frac{\log \mathbf{L}}{\log y}\right)\right)}{\omega(u)}\left(1 + O\left(\frac{1}{\log y}\right)\right) = z\left(1 + O\left(\frac{\log \mathbf{L}}{\log y}\right)\right),$$

hence Hypothesis 4.3.2 holds if $y > x^{\frac{1}{\log\log x}}$. If $\mathbf{L}^{(1+\delta)\log\log \mathbf{L}} \leq y \leq x^{\frac{1}{\log\log x}}$, Hypothesis 4.3.2 follows from (4.6.8).

Now, since $q \leq 2\mathbf{L}(x) < y$, we have the equality

$$\frac{1}{\phi(q)}\sum_{\substack{n \leq x \\ (n,q)=1}} \mathbf{a}_y(n) = \frac{1}{\phi(q)}\sum_{n\leq x} \mathbf{a}_y(n) = \frac{\mathcal{A}(x,y)}{q\gamma_y(q)},$$

thus using (4.6.9) we conclude that Hypothesis 4.3.1* holds. We now turn to an analogue of Proposition 4.5.1, which we prove using (4.6.12). We need an estimate for

$$\begin{aligned} I &= \frac{1}{2\pi i}\int_{(-1/2)} Z_{\mathcal{A}}(s)\psi(s)\frac{ds}{s} \\ &= \frac{1}{2\pi i}\int_{(-1/2)} Z_{\mathcal{A}}(s)((2M)^s - M^s)\frac{ds}{s(s+1)} + O_{a,\epsilon}\left(\left(\frac{M}{x}\right)^{\frac{1}{2}-\epsilon}\right), \end{aligned}$$

since on the line $\sigma = -1 + \frac{1}{\log M}$, we have the bound

$$\prod_{p\geq y}\left(1 + \frac{1}{(p-1)p^{s+1}}\right)^{-1} \ll \prod_{p \geq \mathbf{L}^{(1+\delta)\log\log \mathbf{L}}}\left(1 + \frac{C_1}{p(\log p)^{1+\delta}}\right) \ll 1,$$

and similarly for the derivative of this product. We now study the function $Z(s) := \frac{Z_{\mathcal{A}}(s)}{s(s+1)}$. Using the bounds we just proved, we get that for $s = -1 + \frac{1}{\log M} + it$ with $|t| \geq 2$,

$$|Z(s)|, |Z'(s)| \ll_{\epsilon} (|t|+1)^{-\frac{3}{2}+\epsilon}.$$

If $\omega(a) \geq 2$, then $Z(s)$ and $Z'(s)$ are bounded near $s = -1$ and we conclude that $I = o(1)$. If $\omega(a) = 1$, then we define

$$Z_{12}(s) := Z(s) - \frac{c(M,y)}{s+1},$$

where

$$c(M, y) := -\frac{1}{2}\frac{\phi(a)}{a}\prod_{p|a}\log p\prod_{p\geq y}\left(1+\frac{1}{(p-1)p^{\frac{1}{\log M}}}\right)^{-1}.$$

One sees that for s close to -1 with $\mathrm{Re}(s) = \frac{1}{\log M}$,

$$\prod_{p\geq y}\left(1+\frac{1}{(p-1)p^{s+1}}\right)^{-1} = (1+O(|s+1|))\prod_{p\geq y}\left(1+\frac{1}{(p-1)p^{\frac{1}{\log M}}}\right)^{-1},$$

hence $|Z_{12}'(s)| \ll \frac{1}{|s+1|}$, and thus

$$I = -\frac{1}{2}\frac{\phi(a)}{a}\prod_{p|a}\log p\prod_{p\geq y}\left(1+\frac{1}{(p-1)p^{\frac{1}{\log M}}}\right)^{-1}(1+o(1)). \tag{4.6.13}$$

If $a = \pm 1$, then we take $Z_{13}(s) := Z(s) - \frac{c(M,y)}{(s+1)^2}$, and since $Z_{13}(s) \ll \frac{1}{|s+1|}$, we get that (4.6.13) holds. Finally, in our range of y,

$$\prod_{p\geq y}\left(1+\frac{1}{(p-1)p^{\frac{1}{\log M}}}\right)^{-1} = 1+O\left(\frac{1}{\log y}\right).$$

□

Annexe A

GENERALITIES ON BINARY QUADRATIC FORMS

In this section we review several classical facts about the distribution of positive definite binary quadratic forms $Q(x,y) = \alpha x^2 + \beta xy + \gamma y^2$ in arithmetic progressions. We recall the notations $d = \beta^2 - 4\alpha\gamma$, $\mathcal{S} = \{p \mid 2d\}$, $\chi_d = \left(\frac{4d}{\cdot}\right)$ and

$$R_a(q) = \#\{1 \leq x, y \leq q : Q(x,y) \equiv a \bmod q\}.$$

Lemma A.0.2. *The function* $R_a(q)$ *is multiplicative as a function of* q.

Proof. Define $S_a(q) := \{(x,y) \in (\mathbb{Z} \cap [1,q])^2 : Q(x,y) \equiv a \bmod q\}$ and let q_1, q_2 be two coprime integers. The "reduction mapping"

$$S_a(q_1q_2) \to S_a(q_1) \times S_a(q_2)$$

$$(x,y) \bmod q_1q_2 \mapsto ((x,y) \bmod q_1, (x,y) \bmod q_2)$$

is a bijection by the Chinese remainder theorem. □

Lemma A.0.3. *Take* $Q(x,y) := x^2 - dy^2$ *with* $d \equiv -1 \bmod 4$*, and let* $a \neq 0$ *be a fixed integer such that* $(a, 2d) = 1$*. We have that*

$$\frac{R_a(q)}{q^2} = \frac{\mathbf{f}_a(q)}{q\gamma(q)},$$

where

$$\gamma(q) := \prod_{p \mid q} \left(1 - \frac{\chi_d(p)}{p}\right)^{-1}$$

and $\mathbf{f}_a(q)$ *is a multiplicative function defined on primes as follows.*

For $p \nmid 2ad$, $\mathbf{f}_a(p^e) := 1$. *For* $p^f \parallel a$ *with* $f \geq 1$ *(so* $p \nmid 2d$*),*

$$\mathbf{f}_a(p^e) := \begin{cases} e+1+\frac{1}{p-1} & \textit{if } \chi_d(p)=1, e \leq f \\ f+1 & \textit{if } \chi_d(p)=1, e > f \\ \frac{1}{p+1} & \textit{if } \chi_d(p)=-1, e \leq f, 2 \nmid e \\ 1-\frac{1}{p+1} & \textit{if } \chi_d(p)=-1, e \leq f, 2 \mid e \\ 0 & \textit{if } \chi_d(p)=-1, e > f, 2 \nmid f \\ 1 & \textit{if } \chi_d(p)=-1, e > f, 2 \mid f. \end{cases} \tag{A.0.14}$$

For $p \mid 2d$ *(so* $p \nmid a$*),*

$$\mathbf{f}_a(p^e) = \begin{cases} 1+\left(\frac{a}{p}\right) & \textit{if } p \neq 2 \\ 1+\left(\frac{-4}{a}\right) & \textit{if } p=2, e \geq 2 \\ 1 & \textit{if } p=2, e=1. \end{cases} \tag{A.0.15}$$

PROOF. By Lemma A.0.2, it is enough to show that for any prime p and integer $e \geq 1$,

$$\frac{R_a(p^e)}{p^e} = \frac{\mathbf{f}_a(p^e)}{\gamma(p)}. \tag{A.0.16}$$

First case : $p \nmid 2d$.

We will proceed as in section 2.3 of **[9]**, by using Gauss sums. Writing $e(n) := e^{2\pi i n}$,

$$\begin{aligned} R_a(p^e) &= \frac{1}{p^e} \sum_{1 \leq m \leq p^e} e\left(-m\frac{a}{p^e}\right) \left(\sum_{1 \leq x \leq p^e} e\left(m\frac{x^2}{p^e}\right)\right) \left(\sum_{1 \leq y \leq p^e} e\left(-md\frac{y^2}{p^e}\right)\right) \\ &= p^e + \frac{1}{p^e} \sum_{1 \leq m \leq p^e-1} e\left(-m\frac{a}{p^e}\right) g(m;p^e)g(-md;p^e) \end{aligned}$$

where $g(m;q) := \sum_{n=1}^{q} e(mn^2/q)$ is a Gauss sum. We have the following properties (see [**8**]) :

$$\text{If } q \text{ is odd, then} \qquad g(1;q)^2 = \left(\frac{-1}{q}\right) q. \tag{A.0.17}$$

$$\text{If } (q,m)=1, \text{ then} \qquad g(m;q) = \left(\frac{m}{q}\right) g(1;q). \tag{A.0.18}$$

As for Ramanujan sums, (see for example (3.3) of [**44**])

$$\sum_{\substack{m=1\\(m,q)=1}}^{q} e(ma/q) = \phi(q)\frac{\mu(q/(q,a))}{\phi(q/(q,a))}. \tag{A.0.19}$$

Using these properties, we compute

$$\begin{aligned}
R_a(p^e) &= p^e + \frac{1}{p^e}\sum_{g=1}^{e}\sum_{\substack{1\le m\le p^e-1\\ p^{e-g}\| m}} e\left(-m\frac{a}{p^e}\right) g(m;p^e)g(-md;p^e)\\
&= p^e + \frac{1}{p^e}\sum_{g=1}^{e}\sum_{\substack{1\le m'\le p^g-1\\ p\nmid m'}} e\left(-m'\frac{a}{p^g}\right) p^{2e-2g} g(m';p^g)g(-m'd;p^g)\\
&= p^e + p^e\sum_{g=1}^{e}\left(\frac{d}{p^g}\right)p^{-g}\sum_{\substack{1\le m'\le p^g-1\\ p\nmid m'}} e\left(-m'\frac{a}{p^g}\right) \quad \text{by (A.0.17) and (A.0.18)}\\
&= p^e + p^e\sum_{g=1}^{e}\left(\frac{d}{p}\right)^g\left(1-\frac{1}{p}\right)\frac{\mu(p^g/(p^g,a))}{\phi(p^g/(p^g,a))} \qquad \text{by (A.0.19),}
\end{aligned}$$

which shows (after a straightforward computation) that (A.0.16) holds for $p \nmid 2d$.

Second case : $p \mid 2d$, $p \neq 2$**.**

In this case we have that $p \nmid a$, since $(a,\mathcal{S}) = 1$. The number of solutions of $x^2 - dy^2 \equiv a \bmod p$ is exactly $p\left(1+\left(\frac{a}{p}\right)\right)$. Moreover, such a solution must satisfy $x \not\equiv 0 \bmod p$, thus by Hensel's lemma we obtain that

$$\frac{R_a(p^e)}{p^e} = 1 + \left(\frac{a}{p}\right).$$

Third case : $p = 2$**.**

In this case, $2 \nmid a$. We have that $R_a(2) = 2$. Reducing the equation $x^2 - dy^2 \equiv a \bmod 2^e$ (using that $d \equiv -1 \bmod 4$), we get

$$x \not\equiv y \bmod 2, \qquad\qquad x^2 + y^2 \equiv a \bmod 4,$$

which shows that there are no solutions if $a \equiv 3 \bmod 4$. Suppose now that $a \equiv 1 \bmod 4$. For $e \geq 3$, an odd integer is a square mod 2^e if and only if it is congruent to 1 mod 8 ; in fact we have the following isomorphism :

$$(\mathbb{Z}/2^e\mathbb{Z})^\times \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{e-2}\mathbb{Z}.$$

Using these well-known facts, we find the number of solutions to $x^2 - dy^2 \equiv a \bmod 2^e$ such that x is odd is

$$\begin{aligned} &= 4\#\{y \bmod 2^e : dy^2 + a \equiv 1 \bmod 8\} \\ &= 2^{e-1}\#\{y \bmod 8 : y^2 \equiv d^{-1}(1-a) \bmod 8\} = 2^e \end{aligned}$$

since $d^{-1}(1-a) \equiv 0, 4 \bmod 8$. Now the number of solutions of $x^2 - dy^2 \equiv a \bmod 2^e$ such that x is even is just the number of solutions of $y^2 - d^{-1}x^2 \equiv -d^{-1}a \bmod 2^e$ such that y is odd, which as we have shown (and using that $-d^{-1} \equiv 1 \bmod 4$) is equal to 2^e. We conclude that

$$\frac{R_a(2^e)}{2^e} = \begin{cases} 2 & \text{if } a \equiv 1 \bmod 4 \\ 0 & \text{if } a \equiv 3 \bmod 4. \end{cases}$$

□

Lemma A.0.4. *Take* $Q(x,y) := \alpha x^2 + \beta xy + \gamma y^2$ *with* $(\alpha, \beta, \gamma) = 1$ *and* $d = \beta^2 - 4\alpha\gamma \equiv 1, 5, 9, 12, 13 \bmod 16$. *Let* $a \neq 0$ *be a fixed integer with* $(a, 2d) = 1$. *We have for* $(q, 2d) = 1$ *that*

$$\frac{R_a(q)}{q^2} = \frac{\mathbf{f}_a(q)}{q\gamma(q)},$$

where

$$\gamma(q) := \prod_{p \mid q} \left(1 - \frac{\chi_d(p)}{p}\right)^{-1}$$

and $\mathbf{f}_a(q)$ *is defined as in Lemma A.0.3. Moreover, for* $p \mid 2d$, $p \neq 2$ *(so* $p \nmid a$*),*

$$\frac{R_a(p^e)}{p^e} = \begin{cases} 1 + \left(\frac{\alpha a}{p}\right) & \text{if } p \neq 2, p \nmid \alpha \\ 1 + \left(\frac{\gamma a}{p}\right) & \text{if } p \neq 2, p \nmid \gamma \end{cases} \tag{A.0.20}$$

and

$$\frac{R_a(2^e)}{2^e} = \begin{cases} 1 & \textit{if } 2 \mid \beta, e = 1 \\ 1 + \left(\frac{-4}{\alpha a}\right) & \textit{if } 2 \mid \beta, 2 \nmid \alpha, e \geq 2 \\ 1 + \left(\frac{-4}{\gamma a}\right) & \textit{if } 2 \mid \beta, 2 \nmid \gamma, e \geq 2 \\ \frac{1}{2} & \textit{if } 2 \nmid \beta, 2 \mid \alpha\gamma \\ \frac{3}{2} & \textit{if } 2 \nmid \alpha\beta\gamma. \end{cases} \tag{A.0.21}$$

PROOF. First write $Q(x, y)$ in four different ways :

$$Q(x,y) = \frac{1}{4\alpha}((2\alpha x + \beta y)^2 - dy^2) \tag{A.0.22}$$

$$= \frac{1}{\alpha}\left(\left(\alpha x + \frac{\beta}{2}y\right)^2 - \frac{d}{4}y^2\right) \tag{A.0.23}$$

$$= \frac{1}{4\gamma}((\beta x + 2\gamma y)^2 - dx^2) \tag{A.0.24}$$

$$= \frac{1}{\gamma}\left(\left(\gamma y + \frac{\beta}{2}x\right)^2 - \frac{d}{4}x^2\right). \tag{A.0.25}$$

We will split in five distinct cases.

Case 1 : $p \nmid 2\alpha$**.** In this case, we use the representation (A.0.22). Note that the mapping $\phi_y : x \mapsto 2\alpha x + \beta y$ is an automorphism of $\mathbb{Z}/p^e\mathbb{Z}$, so

$$R_a(p^e) = \#\{1 \leq x, y \leq p^e : x^2 - dy^2 \equiv 4\alpha a \bmod p^e\}.$$

Going through the proof of Lemma A.0.3, we see that

$$\frac{R_a(p^e)}{p^e} = \frac{\mathbf{f}_{4\alpha a}(p^e)}{\gamma(p)} = \frac{\mathbf{f}_{\alpha a}(p^e)}{\gamma(p)} \qquad \left(= \frac{\mathbf{f}_a(p^e)}{\gamma(p)} \text{ if } p \nmid d\right).$$

Case 2 : $p \nmid 2\gamma$**.** In this case, we proceed in an analogous way to the first case, using the representation (A.0.24) to get that

$$\frac{R_a(p^e)}{p^e} = \frac{\mathbf{f}_{4\gamma a}(p^e)}{\gamma(p)} = \frac{\mathbf{f}_{\gamma a}(p^e)}{\gamma(p)} \qquad \left(= \frac{\mathbf{f}_a(p^e)}{\gamma(p)} \text{ if } p \nmid d\right).$$

Case 3 : $p \mid \alpha$**,** $p \mid \gamma$**,** $p \neq 2$**.** In this case $p \nmid \beta$, so $p \nmid d$. Writing $X := x + y$ and $Y := y$, we compute that

$$\alpha X^2 + \beta XY + \gamma Y^2 = \alpha x^2 + (2\alpha + \beta)xy + (\alpha + \beta + \gamma)y^2 =: \alpha' x^2 + \beta' xy + \gamma' y^2.$$

We have $p \mid \alpha'$, $p \nmid \beta'$ and $p \nmid \gamma'$, which reduces the problem to Case 2, and so

$$\frac{R_a(p^e)}{p^e} = \frac{\mathbf{f}_{(\alpha+\beta+\gamma)a}(p^e)}{\gamma(p)} = \frac{\mathbf{f}_a(p^e)}{\gamma(p)}.$$

Case 4.1 : $p = 2, 2 \mid \beta$**.** In this case, $d \equiv 0 \bmod 4$. We have that either $2 \nmid \alpha$, or $2 \nmid \gamma$. In the first event we use representation (A.0.23), which gives

$$R_a(2^e) = \#\{1 \le x, y \le 2^e : x^2 - d'y^2 \equiv \alpha a \bmod 2^e\}$$

with $d' := \frac{d}{4} \equiv -1 \bmod 4$. Going back to the proof of Lemma A.0.3, we get that

$$\frac{R_a(2^e)}{2^e} = \mathbf{f}_{\alpha a}(2^e).$$

In the event that $2 \nmid \gamma$, the result is

$$\frac{R_a(2^e)}{2^e} = \mathbf{f}_{\gamma a}(2^e).$$

Note that if $2 \nmid \alpha\gamma$, then since $\frac{d}{4} \equiv -1 \bmod 4$, we have $\alpha \equiv \gamma \bmod 4$, so

$$\mathbf{f}_{\alpha a}(2^e) = \mathbf{f}_{\gamma a}(2^e).$$

Case 4.2 : $p = 2, 2 \nmid \beta$**.** In this case, $2 \nmid d$ and $2 \nmid a$. An easy application of Hensel's lemma in either of the variables x or y (since one of them has to be odd) yields

$$\frac{R_a(2^e)}{2^e} = \frac{R_a(2)}{2},$$

and all the possibilities are contained in the following table.

α mod 2	β mod 2	γ mod 2	$R_a(2)$
0	1	0	1
0	1	1	1
1	1	0	1
1	1	1	3

□

BIBLIOGRAPHIE

[1] Milton Abramowitz, Irene A. Stegun, eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, New York : Dover, ISBN 0-486-61272-4 (1965).

[2] R.J. Backlund, *Über die Nullstellen der Riemannschen Zetafunktion.* (German) Acta Math. **41** (1916), no. 1, 345–375.

[3] Mark Borisovich Barban, *New applications of the "great sieve" of Ju. V. Linnik.* (Russian) Akad. Nauk Uzbek. SSR Trudy Inst. Mat. No. 22 (1961) 1–20.

[4] Mark Borisovich Barban, *On the average error term in the generalized prime number theorem.* (Russian) Dokl. Uzbek. SSR No.5 (1964) 5–7.

[5] Carter Bays, Richard H. Hudson, *The cyclic behavior of primes in the arithmetic progressions modulo* 11, J. Reine Angew. Math. **339** (1983), 215–220.

[6] Carter Bays, Kevin Ford, Richard H. Hudson, Michael Rubinstein, *Zeros of Dirichlet* L*-functions near the real axis and Chebyshev's bias*, J. Number Theory **87** (2001), no. 1, 54–76.

[7] Carter Bays, Richard H. Hudson, *A new bound for the smallest* x *with* $\pi(x) > \mathrm{li}(x)$. Math. Comp. **69** (2000), no. 231, 1285–1296 (electronic).

[8] Bruce C. Berndt, Ronald J. Evans, Kenneth S. Williams, *Gauss and Jacobi sums.* Canadian Mathematical Society Series of Monographs and Advanced Texts. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1998. xii+583 pp. ISBN : 0-471-12807-4

[9] Valentin Blomer, Jörg Brüdern, Rainer Dietmann, *Sums of smooth squares.* Compos. Math. **145** (2009), no. 6, 1401–1441.

[10] Enrico Bombieri, John B. Friedlander, Henryk Iwaniec, *Primes in arithmetic progressions to large moduli.* Acta Math. **156** (1986), no. 3-4, 203–251.

[11] Enrico Bombieri, John B. Friedlander, Henryk Iwaniec, *Primes in arithmetic progressions to large moduli. II.* Math. Ann. **277** (1987), no. 3, 361–393.

[12] Enrico Bombieri, John B. Friedlander, Henryk Iwaniec, *Primes in arithmetic progressions to large moduli. III.* J. Amer. Math. Soc. **2** (1989), no. 2, 215–224.

[13] Alexander Adolfovich Buchstab, *Asymptoticheskaya otsenka odnoy obshey teoritikocislovoy funktii* Mat. Sbornik, N.S., **2(44)** (1937), 1239-1246.

[14] Chen Jing Run, *On the representation of a larger even integer as the sum of a prime and the product of at most two primes.* Sci. Sinica **16** (1973), 157–176.

[15] Keresztély Corrádi, Imre Kátai, *A comment on K. S. Gangadharan's paper entitled "Two classical lattice point problems".* Magyar Tud. Akad. Mat. Fiz. Oszt. Közl. **17** 1967 89–97.

[16] Harold Davenport, *Multiplicative number theory.* Third edition. Revised and with a preface by Hugh L. Montgomery. Graduate Texts in Mathematics, 74. Springer-Verlag, New York, 2000. xiv+177 pp. ISBN : 0-387-95097-4.

[17] Harold Davenport, Heini Halberstam, *Primes in arithmetic progressions.* Michigan Math. J. **13** (1966) 485–489.

[18] Nicolaas Govert De Bruijn, *On the number of uncancelled elements in the sieve of Eratosthenes.* Nederl. Akad. Wetensch., Proc. 53, (1950) 803–812

[19] Tim Dokchitser, *Computing special values of motivic* L*-functions* Experiment. Math. **13** (2004), no. 2, 137–149.

[20] Harold M. Edwards, *Riemann's zeta function.* Pure and Applied Mathematics, Vol. 58. Academic Press, New York-London, 1974. xiii+315 pp.

[21] Leonhard Euler, *Remarques sur un beau rapport entre les séries des puissances tant directes que réciproques.* Mémoires de l'académie des sciences de Berlin **17** (1768) pp. 83–106.

[22] Andrey Feuerverger, Greg Martin, *Biases in the Shanks-Rényi prime number race*, Experiment. Math. **9** (2000), no. 4, 535–570.

[23] Daniel Fiorilli, *Residue classes containing an unexpected number of primes.* preprint : arXiv :1009.2699v1 [math.NT].

[24] Étienne Fouvry, *Sur le problème des diviseurs de Titchmarsh.* J. Reine Angew. Math. **357** (1985), 51–76.

[25] John B. Friedlander, Andrew Granville, *Limitations to the equi-distribution of primes. I.* Ann. of Math. (2) **129** (1989), no. 2, 363–382.

[26] John B. Friedlander, Andrew Granville, *Relevance of the residue class to the abundance of primes.* Proceedings of the Amalfi Conference on Analytic Number Theory (Maiori, 1989), 95–103, Univ. Salerno, Salerno, 1992.

[27] John B. Friedlander, Andrew Granville, Adolf Hildebrand, Helmut Maier, *Oscillation theorems for primes in arithmetic progressions and for sifting functions.* J. Amer. Math. Soc. **4** (1991), no. 1, 25–86.

[28] Patrick X. Gallagher, *The large sieve.* Mathematika **14** (1967) 14–20.

[29] Daniel A. Goldston, János Pintz, Cem Y. Yildirim, *Primes in tuples. I.* Ann. of Math. (2) **170** (2009), no. 2, 819–862.

[30] Ronald L. Graham, Donald E. Knuth, Oren Patashnik, *Concrete mathematics. A foundation for computer science* (2nd ed.), Addison-Wesley Publishing Company, Reading, MA, 1994.

[31] Andrew Granville, *Smooth numbers : computational number theory and beyond.* Algorithmic number theory : lattices, number fields, curves and cryptography, 267-323, Math. Sci. Res. Inst. Publ., **44**, Cambridge Univ. Press, Cambridge, 2008.

[32] Andrew Granville, Greg Martin, *Prime number races.* Amer. Math. Monthly **113** (2006), no. 1, 1–33.

[33] Andrew Granville, Kannan Soundararajan, *An uncertainty principle for arithmetic sequences.* Ann. of Math. (2) **165** (2007), no. 2, 593–635.

[34] Heini Halberstam, *Footnote to the Titchmarsh-Linnik divisor problem.* Proc. Amer. Math. Soc. **18** (1967) 187–188.

[35] Godfrey Harold Hardy, *On the Expression of a Number as the Sum of Two Squares.* Quart. J. Math. **46** (1915), 263–283.

[36] Godfrey Harold Hardy, John Edensor Littlewood, *Some problems of 'Partitio numerorum' ; III : On the expression of a number as a sum of primes.* Acta Math. **44** (1923), no. 1, 1–70.

[37] Godfrey Harold Hardy, Srinivasa Ramanujan, *The normal number of prime factors of a number* n. Quart. J. Math. **48** (1917), 76–92.

[38] James Lee Hafner, *New omega theorems for two classical lattice point problems.* Invent. Math. **63** (1981), no. 2, 181–186.

[39] Christopher Hooley, *On the Barban-Davenport-Halberstam theorem. I.* Collection of articles dedicated to Helmut Hasse on his seventy-fifth birthday, III. J. Reine Angew. Math. **274**/**275** (1975), 206–223.

[40] C. Hooley, *On the Barban-Davenport-Halberstam theorem. VII.*, J. London Math. Soc. (2) **16** (1977), no. 1, 1–8.

[41] Martin N. Huxley, *Exponential sums and the Riemann zeta function. V.* Proc. London Math. Soc. (3) **90** (2005), no. 1, 1–41.

[42] Martin N. Huxley, *Integer points, exponential sums and the Riemann zeta function.* Number theory for the millennium, II (Urbana, IL, 2000), 275-290, A K Peters, Natick, MA, 2002.

[43] Henryk Iwaniec, *The half dimensional sieve.* Acta Arith. 29 (1976), no. 1, 69–95.

[44] Henryk Iwaniec, Emmanuel Kowalski, *Analytic number theory.* American Mathematical Society Colloquium Publications, 53. American Mathematical Society, Providence, RI, 2004. xii+615 pp.

[45] Jerzy Kaczorowski, *On sign-changes in the remainder-term of the prime-number formula. I.* Acta Arith. **44** (1984), no. 4, 365–377.

[46] Jerzy Kaczorowski, *On sign-changes in the remainder-term of the prime-number formula. II.* Acta Arith. **45** (1985), no. 1, 65–74.

[47] Jerzy Kaczorowski, *On sign-changes in the remainder-term of the prime-number formula. III.* Acta Arith. **48** (1987), no. 4, 347–371.

[48] Jerzy Kaczorowski, *On sign-changes in the remainder-term of the prime-number formula. IV.* Acta Arith. **50** (1988), no. 1, 15–21.

[49] Jerzy Kaczorowski, *The* k*-functions in multiplicative number theory. V. Changes of sign of some arithmetical error terms.* Acta Arith. **59** (1991), no. 1, 37–58.

[50] Jerzy Kaczorowski, Kazimierz Wiertelak, *Oscillations of a given size of some arithmetic error terms.* Trans. Amer. Math. Soc. **361** (2009), no. 9, 5023–5039.

[51] Jerzy Kaczorowski, *Results on the distribution of primes.* J. Reine Angew. Math. **446** (1994), 89-113.

[52] Jerzy Kaczorowski, *On the distribution of primes (mod* 4*).* Analysis **15** (1995), no. 2, 159-171.

[53] Koichi Kawada, *The prime* k*-tuplets in arithmetic progressions.* Tsukuba J. Math. **17** (1993), no. 1, 43–57.

[54] Stanislaw Knapowski, Pál Turán, *Comparative prime-number theory. I. Introduction.* Acta Math. Acad. Sci. Hungar. **13** (1962) 299–314.

[55] John Edensor Littlewood, *On the class-number of the corpus* $P(\sqrt{-k})$, Proc. London Math. Soc. (2) **27** (1928), 358–372.

[56] John Edensor Littlewood, *Sur la distribution des nombres premiers.* Comptes Rendus de l'Académie des Sciences. Paris **158** (1914), 1869–1872.

[57] Ming-Chit Liu, Tianze Wang, *On the Vinogradov bound in the three primes Goldbach conjecture.* Acta Arith. **105** (2002), no. 2, 133–175.

[58] Yudell L. Luke, *Mathematical functions and their approximations*, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York/London, 1975. xvii+568 pp.

[59] Kevin S. McCurley, *Explicit estimates for the error term in the prime number theorem for arithmetic progressions*, Math. Comp. **42** (1984), no. 165, 265–285.

[60] Hugh L. Montgomery, Robert C. Vaughan, *Multiplicative number theory. I. Classical theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2007. MR MR2378655 (2009b :11001)

[61] Hugh L. Montgomery, *Primes in arithmetic progressions.* Michigan Math. J. **17** (1970) 33–39.

[62] Hugh L. Montgomery, Ulrike M.A. Vorhauer, *Changes of sign of the error term in the prime number theorem.* Funct. Approx. Comment. Math. **35** (2006), 235–247.

[63] Yoichi Motohashi, *An induction principle for the generalization of Bombieri's prime number theorem.* Proc. Japan Acad. **52** (1976), no. 6, 273–275.

[64] Vladimir Aleksandrovich Plaksin, *Asymptotic formula for the number of solutions of an equation with primes.* Izv. Akad. Nauk SSSR Ser. Mat. **45** (1981), no. 2, 321–397

[65] Vladimir Aleksandrovich Plaksin, *Asymptotic formula for the number of representations of a natural number by a pair of quadratic forms, the arguments of one of which are primes.* Izv. Akad. Nauk SSSR Ser. Mat. **48** (1984), no. 6, 1245–1265.

[66] M. M. Rao, *Probability theory with applications*, Probability and Mathematical Statistics, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1984.

[67] J. Barkley Rosser, Lowell Schoenfeld, *Approximate formulas for some functions of prime numbers* Illinois J. Math. **6** (1962), 64–94.

[68] Michael Rubinstein, Peter Sarnak, *Chebyshev's bias*, Experiment. Math. **3** (1994), no. 3, 173–197.

[69] L.G. Sathe, *On a problem of Hardy on the distribution of integers having a given number of prime factors. I.* J. Indian Math. Soc. (N.S.) **17** (1953), 63–82.

[70] L. G. Sathe, *On a problem of Hardy on the distribution of integers having a given number of prime factors. II.* J. Indian Math. Soc. (N.S.) **17** (1953), 83–141.

[71] Atle Selberg, *Note on a paper by L. G. Sathe.* J. Indian Math. Soc. (N.S.) **18** (1954), 83–87.

[72] Daniel Shanks, *Quadratic residues and the distribution of primes*, Math. Tables Aids Comput. **13** (1959), 272–284.

[73] Kannan Soundararajan, *Omega results for the divisor and circle problems.* Int. Math. Res. Not. 2003, no. 36, 1987–1998.

[74] John L. Spouge, *Computation of the gamma, digamma, and trigamma functions* SIAM J. Numer. Anal. **31** (1994), no. 3, 931–944.

[75] Pafnouti Tchebychev, *Lettre de M. le professeur Tchébychev à M.Fuss, sur un noveau théorème relatif aux nombres premiers contenus dans les formes $4n+1$ et $4n+3$.* Bull. Cl. phys.-math. Acad. Imp. Sci. St.Pétersbourg, **11** (1853), p.208.

[76] Gérald Tenenbaum, *Introduction à la théorie analytique et probabiliste des nombres.* Deuxième édition, Cours Spécialisés, 1. Société Mathématique de France, Paris, 1995. xv+457 pp.

[77] G. N. Watson, *A treatise on the theory of Bessel functions*, 2nd ed., Cambridge Mathematical Library, Cambridge University Press, 1995.

[78] Dieter Wolke, *Über die mittlere Verteilung der Werte zahlentheoretischer Funktionen auf Restklassen. II.* (German) Math. Ann. **204** (1973), 145–153.

[79] Xuan, Ti Zuo *Integers free of small prime factors in arithmetic progressions.* Nagoya Math. J. **157** (2000), 103-127.