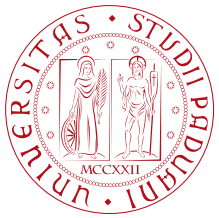

UNIVERSITÀ DEGLI STUDI DI PADOVA
Dipartimento di Matematica Pura ed Applicata

SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE MATEMATICHE
Indirizzo di Matematica
XXIII ciclo

TESI DI DOTTORATO DI RICERCA

ON SOME ADDITIVE PROBLEMS WITH PRIMES AND POWERS OF A FIXED INTEGER

Direttore della Scuola: Ch.mo Prof. Paolo Dai Pra

Coordinatore d'indirizzo: Ch.mo Prof. Franco Cardin

Supervisori: Ch.mo Prof. Bruno Chiarellotto
Ch.mo Prof. Alessandro Languasco

Dottorando: Valentina Settimi

Anno Accademico 2010/2011

I want to start by warmly thanking my supervisor Prof. Alessandro Languasco, because he has always encouraged and supported me during my research. He also spent a lot of time reading several drafts of this thesis with great attention. His comments and his help have been fundamental.

Next I would like to thank my supervisor Prof. Bruno Chiarellotto for having inspired my research. He has always been very supportive and patient, even when we had encountered some mathematical difficulties. I would like to thank him also for having supported my work financially and because he gave me the opportunity to move to Oxford.

I would also like to thank Prof. Alan Lauder for all the help he gave me in Oxford.

Finally, I would like to thank all of my colleagues at the Department of Pure and Applied Mathematics of the University of Padova.

Contents

Introduction

Ogni numero pari maggiore di 2 *è somma di due primi.*
Every even integer greater than 2 *is sum of two primes.*

Introduzione (italiano)

La congettura di Goldbach

Il 7 giugno del 1742, in una lettera indirizzata a Eulero, Goldbach formulò la seguente congettura

$$\begin{gathered} \textit{se } N \textit{ è un intero tale che } N = p_1 + p_2, \textit{ con } p_1 \textit{ e } p_2 \textit{ primi,} \\ \textit{allora, per ogni } 2 \leq k \leq N,\ N = p_1 + \ldots + p_k, \textit{ con } p_1, \ldots, p_k \textit{ primi.} \end{gathered}$$

Va sottolineato che al tempo di Goldbach il numero 1 era considerato un numero primo, contrariamente ad oggi. A margine della stessa lettera, Goldbach annotò un'altra congettura

$$\begin{gathered} \textit{se } N \textit{ è un intero maggiore di } 2, \\ \textit{allora } N = p_1 + p_2 + p_3, \textit{ con } p_1, \ p_2 \textit{ e } p_3 \textit{ primi.} \end{gathered}$$

Nella lettera di risposta, datata 30 giugno dello stesso anno, Eulero scrisse una terza congettura che egli stesso attribuì a Goldbach

$$\textit{se } N \textit{ è un intero positivo e pari, allora } N = p_1 + p_2, \textit{ con } p_1 \textit{ e } p_2 \textit{ primi.}$$

Oggi è noto che queste tre congetture sono equivalenti (si veda, *e.g.*, Pintz [Pin06b]).

Dopo l'esclusione di 1 dall'insieme dei numeri primi, queste congetture sono state tradotte nel moderno linguaggio dei primi. La prima, essendo strettamente legata alla primalità di 1, non ha un'interessante traduzione, al contrario la versione moderna della seconda congettura è la cosiddetta *congettura di Goldbach ternaria*

$$\begin{gathered} \textit{se } N \textit{ è un intero dispari maggiore di } 5, \\ \textit{allora } N = p_1 + p_2 + p_3, \textit{ con } p_1, \ p_2 \textit{ e } p_3 \textit{ primi,} \end{gathered} \tag{TGC}$$

mentre la versione moderna della terza congettura è la celebre *congettura di Goldbach*

$$\begin{gathered} \textit{se } N \textit{ è un intero pari maggiore di } 2, \\ \textit{allora } N = p_1 + p_2, \textit{ con } p_1 \textit{ e } p_2 \textit{ primi.} \end{gathered} \tag{GC}$$

Avendo ipotesi più restrittive, le versioni moderne sono più forti delle originali; in particolare non si ha l'equivalenza tra congettura di Goldbach e congettura di Goldbach ternaria, ma soltanto

$$\text{GC} \Rightarrow \text{TGC}.$$

Nonostante la formulazione molto semplice, la congettura di Goldbach è estremamente difficile da dimostrare, al punto che oggi, dopo più di 250 anni, è ancora un problema aperto. Per maggiori informazioni su questo argomento, si veda il bel lavoro monografico [Pin06b] di Pintz.

Risultati legati alla congettura di Goldbach

Benché la congettura di Goldbach sia ancora oggi un problema aperto, esistono numerosi risultati legati ad essa.

Nel 1923 Hardy e Littlewood, assumendo l'*ipotesi di Riemann generalizzata*[1] (GRH nel seguito), dimostrarono la TGC per N sufficientemente grande e la GC per quasi tutti gli interi pari (si vedano [HL23a] e [HL23b] rispettivamente).

Nel 1937 Vinogradov, in [Vin37], riuscì a dimostrare la TGC, per N sufficientemente grande, *incondizionatamente* e cioè senza assumere GRH.

Per quanto riguarda l'insieme eccezionale per la congettura di Goldbach, il principale risultato incondizionale si deve a Montgomery e Vaughan. Essi dimostrarono in [MV75] che esiste $\delta > 0$ per cui il numero di interi positivi pari e minori di X che non si possono scrivere come somma di due primi è $\ll X^{1-\delta}$. Recentemente in [Pin09], Pintz ha annunciato che la stima precedente è valida per $\delta = 1/3$.

Un importante risultato legato alla congettura di Goldbach si deve a Linnik: egli dimostrò, nel 1951 sotto GRH (in [Lin51]) e due anni dopo incondizionatamente (in [Lin53]) l'esistenza di una costante k tale che ogni intero pari sufficientemente grande è somma di due primi e al più k potenze di 2. A tale lavoro seguirono numerosi tentativi di trovare un upper bound per k e questo problema è noto come *problema di Goldbach-Linnik*.

Ad oggi, le migliori stime sono $k = 7$ sotto GRH e $k = 13$ incondizionatamente, ottenute da Heath–Brown e Putcha nel 2002 in [HBP02]. Va ricordato che Pintz e Ruzsa nel 2003, in [PR03], dimostrarono in maniera indipendente $k = 7$ sotto GRH (nello stesso lavoro, inoltre, gli autori annunciarono $k = 8$ incondizionatamente).

Nel 2007 Languasco, Pintz e Zaccagnini, in [LPZ07], hanno risolto una variante del problema di Goldbach-Linnik: fissato $k \geq 1$, trovare una formula asintotica per il numero di rappresentazioni di un intero minore di X come somma di due primi e k potenze di 2, valida per quasi tutti gli interi positivi e pari. Il punto importante nel loro lavoro è che, per ogni $k \geq 1$, il numero di eccezioni alla formula asintotica è $\ll_k X^{3/5}(\log X)^{10}$. Infatti

[1] *i.e.* per ogni carattere di Dirichlet χ e per ogni numero complesso s tali che $L(\chi, s) = 0$, se $0 < \Re(s) < 1$, allora $\Re(s) = 1/2$.

- la stima dell'insieme eccezionale per la congettura di Goldbach in [Pin09] è $\ll X^{2/3}$ e chiaramente $3/5 < 2/3$. Dunque, anche aggiungendo una sola potenza di 2, la stima migliora: ciò dipende dal riuscire a fare l'aritmetica in media della serie singolare;
- l'esponente $3/5$ è, allo stato dell'arte, il livello migliore che si possa ottenere: per abbassarlo si deve migliorare, negli esponenti, la stima di Vaughan (si veda [Vau97], Teorema 3.1).

Nel Capitolo 3 generalizziamo il risultato di Languasco, Pintz e Zaccagnini, trovando, fissati $g \geq 3$ e $k \geq 1$, una formula asintotica per il numero di rappresentazioni di un intero (minore di X) come somma di due primi e k potenze di g, valida per quasi tutti gli interi positivi verificanti opportune (e standard) condizioni aritmetiche. Il punto importante nel nostro lavoro è che, pur lavorando con potenze più sparse delle potenze di 2, riusciamo lo stesso ad ottenere $3/5$ come esponente nella stima dell'insieme eccezionale. Ciò dipende dal riuscire a gestire l'aritmetica, più complicata, facendo la media della serie singolare sulle potenze di g.

Un altro problema legato alla congettura di Goldbach è il *problema di Waring-Goldbach*, che riguarda la rappresentabilità un intero come somma di potenze prime. Il primo lavoro in questo campo si deve a Hua che, nel 1938 in [Hua38b], dimostrò che quasi tutti i naturali $n \equiv 3 \pmod{24}$ e $n \not\equiv 0 \pmod 5$ si possono scrivere come somma di tre quadrati di primi. Nello stesso lavoro dimostrò anche che ogni intero $n \equiv 5 \pmod{24}$ sufficientemente grande si può scrivere come somma di cinque quadrati di primi.

Combinando il problema di Goldbach-Linnik con il problema di Waring-Goldbach si hanno i cosiddetti *problemi misti* con potenze di primi e potenze di 2. Riguardo a questo argomento, ricordiamo il lavoro [LLZ99] di J. Liu, M.C. Liu e T. Zhan, in cui gli autori dimostrarono che, prendendo k opportunamente grande, si ha: sia che ogni intero pari sufficientemente grande è somma di quattro quadrati di primi e k potenze di 2, sia che ogni intero dispari sufficientemente grande è somma di un primo, due quadrati di primi e k potenze di 2. A tale articolo seguirono numerosi lavori per stimare k e, ad oggi, i risultati migliori sono: $k = 151$, relativamente allo studio di quattro quadrati di primi e k potenze di 2 (dimostrato da H. Li in [Li06]) e $k = 83$, relativamente allo studio di un primo, due quadrati di primi e k potenze di 2 (dimostrato da G. Lü e Sun in [LS09])

Un altro tipo di risultati collegati alla congettura di Goldbach sono quelli che riguardano i *problemi diofantei con numeri primi* che, in un certo senso, possono essere considerati come l'analogo reale dalla GC e della TGC. Riguardo a questo argomento, ricordiamo i lavori di Brüden, Cook e Perelli [BCP97] per la forma binaria e di Vaughan [Vau74] per la forma ternaria. Per esempio Vaughan dimostrò che, se λ_1, λ_2 e λ_3 sono numeri reali non nulli e non tutti dello stesso segno, η è un numero reale e λ_1/λ_2 è irrazionale, allora esistono infinite triplette ordinate (p_1, p_2, p_3) di numeri primi, tali che

$$|\, \eta + \lambda_1 p_1 + \lambda_2 p_2 + \lambda_3 p_3 \,| < (\max_{1\leq i\leq 3} p_i)^{-1/10} (\log(\max_{1\leq i\leq 3} p_i))^{20}.$$

Un problema diofanteo con due primi e potenze di 2, che può essere quindi considerato un analogo reale del problema di Goldbach-Linnik, è stato studiato da Parsell nel 2003 in

[Par03]. Qui egli ha dimostrato che i valori assunti dalle combinazioni lineari reali di due primi e k potenze di 2 possono approssimare arbitrariamente bene qualsiasi numero reale, prendendo k sufficientemente grande. Recentemente Languasco e Zaccagnini in [LZ10] hanno migliorato il lavoro di Parsell, abbassando il numero di potenze di 2 necessarie.

Ricordiamo infine che W.P. Li e Wang nel 2005, in [LW05], hanno studiato un'approssimazione diofantea ai problemi misti con quadrati di primi e potenze di 2: in particolare hanno dimostrato che i valori assunti delle combinazioni lineari reali di un primo, due quadrati di primi e k potenze di 2 possono approssimare arbitrariamente bene qualsiasi numero reale, prendendo k sufficientemente grande.

Nel Capitolo 2 miglioriamo il risultato W.P. Li e Wang, abbassando il loro lower bound per k di circa il 90%.

Struttura della tesi

Il Capitolo 1 è di carattere preliminare e raccoglie le notazioni e i risultati noti usati nella tesi; in particolare, nella Sezione 1.3 descriviamo alcune tecniche classiche che saranno fondamentali nei capitoli successivi: il metodo del cerchio di Hardy-Littlewood (utilizzato nel Capitolo 3 per studiare una variante del problema di Goldbach-Linnik), il metodo di Davenport-Heilbronn e la formula esplicita di Pintz (utilizzati nel Capitolo 2 per studiare un problema diofanteo).

Nel Capitolo 2 presentiamo il primo risultato di questa tesi, che riguarda un problema diofanteo con un primo, due quadrati di primi e k potenze di 2. Questo problema è stato già studiato da W.P. Li e Wang in [LW05]: nella Sezione 2.1 descriviamo il nostro risultato, confrontandolo dettagliatamente con tale precedente lavoro. In particolare spieghiamo quali novità abbiamo introdotto e come queste ci abbiano permesso di migliorare il risultato di Li-Wang, abbassando il lower bound di k di circa il 90%.

Nella Sezione 2.2 impostiamo il nostro problema e nella Sezione 2.3 elenchiamo i lemmi utilizzati per la sua risoluzione, dimostrando nel dettaglio i risultati nuovi (che sono varianti per quadrati di primi di risultati noti):

- Il Lemma 2.8 è una variante del Lemma 4 in Languasco-Zaccagnini [LZ10].
- Il Lemma 2.9 è una variante del Lemma 4 in Parsell [Par03].
- Il Lemma 2.12 è una variante del Lemma 1 in Brüdern-Cook-Perelli [BCP97].
- Il Lemma 2.13 è una variante della stima dell'integrale di Selberg dimostrata da Saffari-Vaughan in [SV77], §6 (in particolare i Claim 2.A e Claim 2.B sono l'analogo dei Lemmi 5 e 6 rispettivamente).

Nelle successive Sezioni 2.4, 2.5 e 2.6 dimostriamo opportune stime sugli archi principali, secondari e banali. Il capitolo si conclude con la Sezione 2.7 in cui, raccogliendo tutti i risultati parziali ottenuti, dimostriamo il nostro teorema.

Nel Capitolo 3 presentiamo il secondo risultato di questa tesi, che riguarda un problema con due primi e k potenze di $g \geq 3$. Più precisamente, il problema da noi studiato è

una variante del lavoro [LPZ07] di Languasco, Pintz e Zaccagnini riguardo al problema di Goldbach-Linnik. Nella Sezione 3.1 descriviamo il nostro risultato, confrontandolo dettagliatamente con quello in [LPZ07]: in particolare spieghiamo quali sono le novità da noi introdotte per lavorare con le potenze di g al posto delle potenze di 2.

Nella Sezioni 3.2 e 3.3 definiamo la notazione necessaria e impostiamo il nostro problema, mentre nelle successive Sezioni 3.4 e 3.5 dimostriamo opportune stime sugli archi principali e secondari. In particolare, nella Sezione 3.4, in analogia con [LPZ07], riusciamo ad ottenere una stima dell'insieme eccezionale che è ottimale, se si usa la stima di Vaughan. Nella Sezione 3.5, la parte interessante riguarda il trattamento delle potenze di g (si veda lo studio di $R^{(5)}_{\mathfrak{M}}(N)$).

La Sezione 3.6 è infine dedicata alla dimostrazione del nostro teorema. Nella successiva Sezione 3.7 dimostriamo i risultati nuovi utilizzati in tale Capitolo 3:

- Il Lemma 3.6 è una variante di un risultato di Romanov in [Rom34].
- Il Lemma 3.8 è un raffinamento del Lemma 1.2 in Murty-Rosen-Silverman [MRS96].
- Il Corollario 3.9 è un'applicazione del Lemma 3.8 alle potenze di g.
- Il Lemma 3.10 è una variante, per le potenze di g, del Lemma 6.2 in Languasco-Pintz-Zaccagnini [LPZ07].
- Il Lemma 3.14 riguarda la stima di un prodotto convergente.

Il capitolo si conclude con la Sezione 3.8, contenente il programma PARI-GP utilizzato nella dimostrazione del Lemma 3.14.

Appendice

L'Appendice A riguarda un problema coomologico da noi studiato durante il dottorato: contare i punti razionali di una curva algebrica definita sopra un campo finito. Tale problema è di grande interesse perché ha importanti applicazioni in crittografia (per esempio si applica al *problema del logaritmo discreto*). La nostra idea è di affrontarlo "alla maniera di Kedlaya" (si veda [Ked01]) ossia studiando l'azione del morfismo di Frobenius su particolari spazi di coomologia p-adica: la *coomologia di Monsky-Washnitzer*.

La tecnica di Kedlaya è stata già generalizzata, ad esempio, da Lauder [Lau04] usando la teoria di Dwork e da Chatel [CL09] usando la coomologia di Monsky-Washnitzer a supposto compatto. In questi lavori, le curve in esame sono sempre curve iperellittiche (*i.e.* rivestimenti doppi di $\mathbb{P}^1$): la nostra idea è generalizzare tali tecniche a curve più generali, come, ad esempio, le curve trigonali (*i.e.* rivestimenti tripli di $\mathbb{P}^1$).

Introduction (english)

The Goldbach conjecture

In a letter to Euler dated 7 June of 1742, Goldbach stated the following conjecture

if N is an integer such that $N = p_1 + p_2$, with p_1 and p_2 primes,
then, for every $2 \leq k \leq N$, $N = p_1 + \ldots + p_k$, with $p_1, \ldots, p_k$ prime.

We have to keep in mind that in Goldbach's time the number 1 was considered to be a prime, in contrast with the modern definition. In the margin of the same letter, Goldbach stated another conjecture

if N is a integer greater than 2,
then $N = p_1 + p_2 + p_3$, with p_1, p_2 and p_3 primes.

In his reply letter, dated 30 June of the same year, Euler wrote a third conjecture which is now ascribed to Goldbach

if N is a positive even integer, then $N = p_1 + p_2$, with p_1 and p_2 primes.

Today, these three conjectures are known to be equivalent (see, *e.g.*, Pintz [Pin06b]).

We can rewrite the conjectures above using the modern language of primes, that is without considering 1 to be a prime number. The first conjecture is strictly connected to the primality of 1 and therefore its modern version has no interest. On the contrary, the modern version of the second conjecture is the so called *ternary Goldbach conjecture*

if N is an odd integer greater than 5,
then $N = p_1 + p_2 + p_3$, with p_1, p_2 and p_3 primes, (TGC)

while the modern version of the third conjecture is the famous *Goldbach conjecture*

if N is an even integer greater than 2,
then $N = p_1 + p_2$, with p_1 and p_2 primes. (GC)

Since the modern conjectures have more restrictive hypothesis, they are stronger than original ones; in particular the Goldbach conjecture and the ternary Goldbach conjecture are not equivalent, but only

$$\text{GC} \Rightarrow \text{TGC}.$$

Despite its very simple statement, the Goldbach conjecture is extremely hard to prove and nowadays, after more than 250 years, it is still an open problem. For a beautiful survey on this subject, we refer to Pintz [Pin06b].

Results related to the Goldbach conjecture

Even if the Goldbach conjecture is still an open problem, nevertheless there exists a large number of results related to it.

In 1923 Hardy and Littlewood proved, under the *generalized Riemann hypothesis*[2] (GRH in the following) both TGC for sufficiently large N and GC for almost all the even integers (see [HL23a] and [HL23b] respectively).

[2] *i.e.* for every Dirichlet character χ and every complex number s such that $L(\chi, s) = 0$, if $0 < \Re(s) < 1$, then $\Re(s) = 1/2$.

In 1937 Vinogradov (see [Vin37]) was able to prove TGC for sufficiently large N, *unconditionally* that is without assuming GRH.

Concerning the exceptional set for the Goldbach conjecture, the most important unconditional result is by Montgomery and Vaughan. In [MV75] they proved that there exists $\delta > 0$ such that the number of positive even integers which are less than X and that cannot be written as sum of two primes is $\ll X^{1-\delta}$. Recently, Pintz in [Pin09] announced that the Montgomery-Vaughan estimate holds for $\delta = 1/3$.

A very important result related to the Goldbach conjecture is due to Linnik: he proved, in 1951 under GRH and two years later unconditionally (see [Lin51] and [Lin53] resp.) that there exists a constant k such that every sufficiently large even integer is a sum of two primes and at most k powers of 2. The problem of finding an upper bound for such k is known as the *Goldbach-Linnik problem* and it has been extensively studied.

To this day, the best upper bounds for k are $k = 7$ under GRH and $k = 13$ unconditionally, obtained by Heath–Brown and Putcha in 2002 in [HBP02]. We remark as well that Pintz and Ruzsa in 2003 independently proved $k = 7$ under GRH (see [PR03], in the same paper the authors also announced $k = 8$ unconditionally).

In 2007, Languasco, Pintz and Zaccagnini (see [LPZ07]) studied a variation of the Goldbach-Linnik problem: given $k \geq 1$, finding an asymptotic formula for the number of representations of a positive even integer less than X as sum of two primes and k powers of 2, which holds for almost all positive even integers. The important point in their work is that, for every $k \geq 1$, the number of exceptional values for the asymptotic formula is $\ll_k X^{3/5}(\log X)^{10}$. In fact

- by Pintz [Pin09], the size of the exceptional set for the Goldbach conjecture is $\ll X^{2/3}$, and $3/5 < 2/3$. Therefore, just adding a single power of 2, a better estimation can be found;
- the exponent $3/5$ is the best possible level, according to the state of the art: to lower it, we have to refine, in the exponents, the Vaughan estimation (see [Vau97], Theorem 3.1).

In Chapter 3 we generalize Languasco-Pintz-Zaccagnini's result: given $g \geq 3$ and $k \geq 1$, we find an asymptotic formula for the number of representations of an integer less than X as sum of two primes and k powers of g, which holds for almost all positive integers satisfying suitable (and standard) arithmetic conditions. The important point in our work is that, even using sparser powers, we still obtain $3/5$ as exponent in the estimation of the exceptional set.

Another problem connected to the Goldbach conjecture is the *Waring-Goldbach problem* which is about the representability of a positive integer as sum of prime powers. The first result in this field is dated 1938 and due to Hua who proved in [Hua38b] that almost all positive integers n, such that $n \equiv 3 \pmod{24}$ and $n \not\equiv 0 \pmod{5}$, can be written as sum of three prime squares. In the same work he also proved that every positive integer n sufficiently large and such that $n \equiv 5 \pmod{24}$ can be written as sum of five prime squares.

The combination of the Goldbach-Linnik problem with the Waring-Goldbach problem gives rise to the so called *mixed problem* with prime powers and powers of 2. In this field we recall the paper [LLZ99] by J. Liu, M.C. Liu e T. Zhan, where the authors proved that, given k suitably large, then both every even integer sufficiently large is sum of four prime squares and k powers of 2, and every odd integer sufficiently large is sum of one prime, two prime squares and k powers of 2. After Liu-Liu-Zhan's work, several estimates for k have been proved and, to this day, the best results are the following: $k = 151$, for the problem with four prime squares and k powers of 2 (proved by H. Li in [Li06]) and $k = 83$, for the problem with one prime, two prime squares and k powers of 2 (proved by G. Lü and Sun in [LS09]).

Other results related to the Goldbach conjecture are those about *diophantine problems with prime numbers* that can be considered as the real analogous of GC and TGC. In this field, we recall the works of Brüden, Cook and Perelli [BCP97] for the binary form and of Vaughan [Vau74] for the ternary one. For example, Vaughan proved that, if λ_1, λ_2 and λ_3 are non-zero real numbers not all of the same sign, η is real and λ_1/λ_2 is irrational, then there are infinitely many ordered triples (p_1, p_2, p_3) of primes for which

$$|\,\eta + \lambda_1 p_1 + \lambda_2 p_2 + \lambda_3 p_3\,| < (\max_{1\le i\le 3} p_i)^{-1/10}(\log(\max_{1\le i\le 3} p_i))^{20}.$$

A diophantine problem with two primes and powers of 2, that can be considered as a real analogous of the Goldbach-Linnik problem, was studied by Parsell in 2003 in [Par03]: he proved that the values obtained by linear real combinations of two primes and k powers of 2 can approximate every real number with an arbitrarily small error, for k sufficiently large. Recently in [LZ10], Languasco and Zaccagnini improved Parsell's result, lowering the lower bound for k.

We finally recall that W.P. Li and Wang in 2005, in [LW05], studied a diophantine approximation to a mixed problem with prime squares and powers of 2: more precisely, they proved that the values obtained by linear real combinations of one prime, two prime squares and k powers of 2 can approximate every real number with an arbitrarily small error, for k sufficiently large.

In Chapter 2 we improve W.P. Li and Wang's result, lowering their lower bound for k by about 90%.

Structure of the thesis

Chapter 1 is of preliminary character and it collects the notation and some well-known results used in this work. In particular, in Section 1.3, we describe some classical techniques, crucial in the following chapters: the Hardy-Littlewood circle method (used in Chapter 3 to study a variation of the Goldbach-Linnik problem), the Davenport-Heilbronn method and Pintz's explicit formula (used in Chapter 2 to study a diophantine problem).

In Chapter 2 we introduce our first theorem, which is about a diophantine problem with one prime, two prime squares and k powers of 2. This problem was already studied by W.P. Li e Wang in [LW05]: in Section 2.1 we describe our result, carefully comparing it

with this previous work. In particular, we analyse our techniques and we explain how they allowed us to improve Li-Wang's result, lowering their lower bound for k by about 90%.

In Section 2.2 we set up our problem and in Section 2.3 we collect the lemmas used to solve it, proving in details the new results (which are variations, for prime squares, of some well-known results):

- Lemma 2.8 is a variation of Lemma 4 in Languasco-Zaccagnini [LZ10].
- Lemma 2.9 is a variation of Lemma 4 in Parsell [Par03].
- Lemma 2.12 is a variation of Lemma 1 in Brüdern-Cook-Perelli [BCP97].
- Lemma 2.13 is a variation of Saffari-Vaughan estimation of Selberg integral in [SV77], §6 (in particular Claims 2.A and 2.B are variations of Lemma 5 and 6 respectively).

In Sections 2.4, 2.5 and 2.6 we prove some suitable estimates for the major, the minor and the trivial arcs.

The chapter ends with Section 2.7 where, gathering together the partial results, we prove our theorem.

In Chapter 3 we introduce our second theorem, which is about a problem with two primes and k powers of $g \geq 3$. More precisely, our work can be considered as a variation of [LPZ07], by Languasco, Pintz e Zaccagnini concerning the Goldbach-Linnik problem. In Section 3.1 we describe our result, carefully comparing it with the one in [LPZ07] and analysing the new techniques used to deal with powers of g, instead of with powers of 2.

In Sections 3.2 and 3.3 we fix the notation and set up our problem, while in Sections 3.4 and 3.5 we prove some suitable estimates for the major and the minor arcs. In particular, in Section 3.4, by analogy with [LPZ07], we obtain an estimate of the exceptional set which is optimal, using the Vaughan estimate (in Result (R.9)). The relevant part in Section 3.5 is the study of $R^{(5)}_{\mathfrak{M}}(N)$ which requires a careful treatment of the powers of g.

In Section 3.6 we finally prove our theorem. In the following Section 3.7 we prove the new results used in this chapter:

- Lemma 3.6 is a variation of a result of Romanov in [Rom34].
- Lemma 3.8 is a refinement of Lemma 1.2 in Murty-Rosen-Silverman [MRS96].
- Corollary 3.9 is an application of Lemma 3.8 to powers of g.
- Lemma 3.10 is a variation, for powers of g, of Lemma 6.2 in Languasco-Pintz-Zaccagnini [LPZ07].
- Lemma 3.14 is about the estimation of a convergent product.

The chapter ends with Section 3.8, containing the PARI-GP program used to prove Lemma 3.14.

Appendix

Appendix A is about point-counting cohomological problem we studied during the PhD: counting the rational points of an algebraic curve defined over a finite field. This is a very studied problem since it has significant applications in cryptography (*e.g.*, for the *discrete logarithm problem*). Our idea is to approach this problem "after Kedlaya" (see [Ked01]), that is studying the action of the Frobenius morphism over some suitable p-adic cohomological spaces: the *Monsky-Washnitzer cohomology*.

Kedlaya technique was already generalized, for example by Lauder [Lau04] using Dwork's theory and by Chatel [CL09] using the Monsky-Washnitzer cohomology with compact support. In all these works, the curves under consideration are hyperelliptic (*i.e.* double covers of $\mathbb{P}^1$): our idea is to generalize such techniques to more general curves, such as the trigonal curves (*i.e.* triple covers of $\mathbb{P}^1$).

Chapter 1

Preliminaries

In this chapter we gather the notation and results we will need in this work. We devote a separate chapter to collecting these notions, in order to simply the reading.

1.1 Notation

In this work, we use the following notation:

$\mathfrak{P}$	set of all prime numbers;
p_i	prime numbers;
g	≥ 3 fixed integer, base of powers;
k	≥ 1, number of powers (of 2 or g);
ν_i, m_i	$\in \mathbb{N}$ positive exponents;
λ_i, μ_i	$\in \mathbb{R}$, coefficients in diophantine problems;
ϖ	$\in \mathbb{R}$, number to be approximate in diophantine problems;
N	suitable integer, studied in Goldbach-Linnik problems;
X	$\in \mathbb{R}$ large parameter;
P, Q	$\in \mathbb{R}$, major and minor arcs levels;
η	$\in \mathbb{R}$, sufficiently small positive constant;
ϵ, ϵ_i	$\in \mathbb{R}$, arbitrarily small positive constants;
$\boldsymbol{C}, \boldsymbol{C_g}$	positive constants in the statement of our theorems;
C, C_i, D, c, c_i	positive constants;
α	$\in \mathbb{R}$ variable, used in exponential sums;
d	odd positive integer;
$\log a$	natural logarithm of a;
L, L'	$\in \mathbb{R}$ parameters suitably related to $\log X$;

$\mathcal{S}, \mathcal{S}_i$	$\subseteq \mathbb{R}$ suitable subsets of $\mathbb{R}$;
$I(\cdot,\cdot), I, J$	needed integrals;
$E(\cdot)$	exceptional set;
err_i	error terms;
$\imath$	imaginary unit;
$s = \sigma + \imath t$	complex number;
$\overline{s} = \sigma - \imath t$	complex conjugate;
$\Re(s)$	$= \sigma$, the real part of s;
$\rho_i = \beta_i + \imath\gamma_i$	generic zero of Dirichlet L-functions;
(a, b)	$\gcd(a, b)$;
$[a, b]$	$\mathrm{lcm}(a, b)$.

We remark that the values of the constants listed above are not necessarily the same at each occurrence.

We will also use the following classical definitions:

c_0 — *twin prime constant*, *i.e.* $c_0 = \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) \approx 0.660\,161$;

$\chi(n)$ — *Dirichlet character*, *i.e.* $\chi : \mathbb{Z} \to \mathbb{C}$ s.t. $\exists r \in \mathbb{N}_{>0}$ called *modulus* with

$$\begin{cases} \chi(n+r) = \chi(n) & \forall n \in \mathbb{Z}, \\ \chi(n) \neq 0 \Leftrightarrow (n, r) = 1, \\ \chi(mn) = \chi(m)\chi(n) & \forall m, n \in \mathbb{Z}; \end{cases}$$

$e(s)$ — $= e^{2\pi\imath s}$;

$\varphi(n)$ — *Euler φ-function*, *i.e.* $\varphi(n) = n \prod_{p|n} \frac{p-1}{p}$;

$\boldsymbol{\gamma}$ — *Euler constant*, *i.e.* $\boldsymbol{\gamma} \approx 0.577\,215$;

$\Gamma(s)$ — *Gamma function*, *i.e.* $\Gamma(s) = \int_0^{+\infty} x^{s-1} e^{-x} \mathrm{d}x$, if $\Re(s) > 0$, extended by analytic continuation to $\mathbb{C}$ except the non-positive integers (where the function has simple poles);

$J(X, h)$ — *Selberg integral*, *i.e.* $J(X, h) = \int_{\epsilon X}^{X} (\vartheta(x+h) - \vartheta(x) - h)^2 \mathrm{d}x$;

$J^*(X, h)$ — *"square-root" Selberg integral*, *i.e.* $J^*(X, h) = \int_{\epsilon X}^{X} \left(\vartheta(\sqrt{x+h}) - \vartheta(\sqrt{x}) - (\sqrt{x+h} - \sqrt{x})\right)^2 \mathrm{d}x$;

$K(\alpha, \eta)$ — kernel function for the Davenport-Heilbronn method, *i.e.*

$K(\alpha,\eta)=\left(\dfrac{\sin\pi\eta\alpha}{\pi\alpha}\right)^2$, for $\alpha\neq 0$;

$L(s,\chi)$ *Dirichlet L-function, i.e.* $L(s,\chi)=\displaystyle\sum_{n\leq 1}\frac{\chi(n)}{n^s}$, if $\Re(s)>1$,
given χ a Dirichlet character modulo $r>1$, and extended to $\mathbb{C}$ by analytic continuation;

$\Lambda(n)$ *von Mangoldt function, i.e.* $\Lambda(n)=\begin{cases}\log p & \text{if } n=p^k,\ \exists k>0\\ 0 & \text{otherwise;}\end{cases}$

$\mathfrak{M},\mathfrak{m},\mathfrak{t}$ *major*, *minor* and *trivial arcs*;

$\mu(n)$ *Möbius μ-function, i.e.* $\mu(n)=\begin{cases}(-1)^k & \text{if } n=p_1\dots p_k,\\ 1 & \text{if } n=1,\\ 0 & \text{otherwise;}\end{cases}$

$N(T)$ $=|\{\rho=\beta+\imath\gamma\,:\,\zeta(\rho)=0;\ 0<\beta<1;\ 0<\gamma\leq T\}|$;

$N(\sigma,T)$ $=|\{\rho=\beta+\imath\gamma\,:\,\zeta(\rho)=0;\ \beta\geq\sigma;\ |\gamma|\leq T\}|$;

$\pi(x)$ *prime-counting function, i.e.* $\pi(x)=|\{p:p\leq x\}|$;

$\psi(x)$ *Chebyshev ψ-function, i.e.* $\psi(x)=\displaystyle\sum_{n\leq x}\Lambda(n)$;

$\Psi(n)$ *Dedekind Ψ-function, i.e.* $\Psi(n)=n\displaystyle\prod_{p|n}\frac{p+1}{p}$;

$\mathfrak{S}(n)$ *singular series* for the Goldbach problem, *i.e.*
$\mathfrak{S}(n)=\begin{cases}0 & \text{if } n \text{ odd},\\ 2c_0\displaystyle\prod_{p|n;\ p>2}\frac{p-1}{p-2} & \text{if } n>0 \text{ even};\end{cases}$

$\vartheta(x)$ *Chebyshev ϑ-function, i.e.* $\vartheta(x)=\displaystyle\sum_{p\leq x}\log(p)$;

$\zeta(s)$ *Riemann ζ-function, i.e.* $\zeta(s)=\displaystyle\sum_{n\leq 1}\frac{1}{n^s}$, if $\Re(s)>1$,
extended to $\mathbb{C}\setminus\{1\}$ by analytic continuation ($s=1$ is a simple pole);

$f(x)\sim g(x)$ for $x\to x_0$, means $\displaystyle\lim_{x\to x_0}\frac{f(x)}{g(x)}=1$, with $x_0\in\mathbb{R}\cup\{\infty\}$;

$f(x)=o(g(x))$ for $x\to x_0$, means $\displaystyle\lim_{x\to x_0}\frac{f(x)}{g(x)}=0$, with $x_0\in\mathbb{R}\cup\{\infty\}$;

$f(x)=\mathcal{O}(g(x))$ for $x\to x_0$, means $|f(x)|\leq C\ g(x)|$ around x_0, with $x_0\in\mathbb{R}\cup\{\infty\}$ and C absolute constant;

$f(x)\ll g(x)$ for $x\to x_0$, means $f(x)=\mathcal{O}(g(x))$;

$f(x)\gg g(x)$ for $x\to x_0$, means $g(x)=\mathcal{O}(f(x))$.

Moreover we need to define the following functions:

$f(d)$ multiplicative function s.t. $f(d) = \begin{cases} \prod_{p|d} \frac{1}{p-2} & \text{if } d>2 \text{ odd}, \\ 0 & \text{otherwise;} \end{cases}$

$G(\alpha)$ exponential sum over powers of 2, *i.e.* $G(\alpha) = \sum_{1\le m\le L} e(2^m\alpha)$;

$G_g(\alpha)$ exponential sum over powers of g, *i.e.* $G_g(\alpha) = \sum_{1\le m\le L'} e(g^m\alpha)$;

$I(X;\mathcal{S})$ relevant integral in diophantine problem, *i.e.*
$I(X;\mathcal{S}) = \int_{\mathcal{S}} S_1(\lambda_1\alpha)S_2(\lambda_2\alpha)S_2(\lambda_3\alpha)G(\mu_1\alpha)\dots G(\mu_k\alpha)e(\varpi\alpha)K(\alpha,\eta)\mathrm{d}\alpha$;

$\mathcal{J}(X)$ $= [2X/3, X]$;

$\mathfrak{N}(X)$ number of solutions $(p_1,p_2,p_3,m_1,\dots,m_k)$ of the inequality
$|\,\lambda_1p_1+\lambda_2p_2^2+\lambda_3p_3^2+\mu_12^{m_1}+\dots+\mu_k2^{m_k}+\varpi\,| < \eta$
with $\epsilon X \le p_1, p_2^2, p_3^2 \le X$ and $1\le m_1,\dots,m_k\le L$;

$r''_{k,g}(N)$ counting function for our problem with two primes and k powers of g. *i.e.*
$r''_{k,g}(N) = |\,\{(p_1,p_2,\nu_1,\dots,\nu_k)\in\mathfrak{P}^2\times[1,L]^k \,:\, N=p_1+p_2+g^{\nu_1}+\dots+g^{\nu_k}\}\,|$
$= \sum_{1\le p_1,p_2\le X} \sum_{\substack{1\le\nu_1,\dots,\nu_k\le L \\ p_1+p_2+g^{\nu_1}+\dots+g^{\nu_k}=N}} 1$;

$R''_{k,g}(N)$ weighted counting function associated to $r''_{k,g}(N)$, *i.e.*
$R''_{k,g}(N) = \sum_{1\le m_1,m_2\le X} \sum_{\substack{1\le\nu_1,\dots,\nu_k\le L \\ m_1+m_2+g^{\nu_1}+\dots+g^{\nu_k}=N}} \Lambda(m_1)\Lambda(m_2)$;

$R''_{\mathcal{S}}(N)$ the restriction of $R''_{k,g}(N)$ to the set $\mathcal{S}$;

$r_{\mathrm{Gb}}(N)$ counting function for the Goldbach problem, *i.e.*
$r_{\mathrm{Gb}}(N) = |\,\{(p_1,p_2)\in\mathfrak{P}^2 \,:\, N=p_1+p_2\}\,| = \sum_{\substack{1\le p_1,p_2\le X \\ p_1+p_2=N}} 1$;

$R_{\mathrm{Gb}}(N)$ weighted counting function associated to $r_{\mathrm{Gb}}(N)$, *i.e.*
$R_{\mathrm{Gb}}(N) = \sum_{\substack{1\le m_1,m_2\le X \\ m_1+m_2=N}} \Lambda(m_1)\Lambda(m_2)$;

$R_{\mathcal{S}}(N)$ the restriction of $R_{\mathrm{Gb}}(N)$ to the set $\mathcal{S}$;

$S_1(\alpha)$ exponential sum over primes, *i.e.* $S_1(\alpha) = \sum_{\epsilon X\le p\le X} \log p\; e(p\alpha)$;

$S_2(\alpha)$ exponential sum over prime squares, *i.e.* $S_2(\alpha) = \sum_{\epsilon X\le p^2\le X} \log p\; e(p^2\alpha)$;

$S(\alpha)$ exponential sum over prime powers, *i.e.* $S(\alpha) = \sum_{1\le m\le X} \Lambda(m)e(m\alpha)$;

$\mathfrak{S}'(n) = \prod_{p|n;\ p>2} \frac{p-1}{p-2}$, so that $\mathfrak{S}(n) = 2c_0\mathfrak{S}'(n)$;

$\mathfrak{S}''(n) = \prod_{p|n;\ p>2} \frac{p+1}{p}$;

$\mathfrak{S}_-(n) = \Big(2 - \frac{1}{2^{m_0-1}} - \frac{1}{2^{m_0}}\Big) \prod_{m\geq 0} \prod_{p>2;\ p^m\|n} \Big(1 + \frac{1}{p} - \frac{1}{p^{m+1}} - \frac{1}{p^{m+2}}\Big)$,
with m_0 such that $2^{m_0}\|n$;

$t_{k,g}(N) = |\{(\nu_1,\dots,\nu_k) \in [1,L]^k \,:\, N = g^{\nu_1} + \dots + g^{\nu_k}\}| = \sum_{\substack{1\leq \nu_1,\dots,\nu_k \leq L \\ N = g^{\nu_1}+\dots+g^{\nu_k}}} 1$;

$t'_{k,g}(N) = |\{(\nu_1,\dots,\nu_k) \in [1,L']^k \,:\, N = g^{\nu_1} + \dots + g^{\nu_k}\}| = \sum_{\substack{1\leq \nu_1,\dots,\nu_k \leq L' \\ N = g^{\nu_1}+\dots+g^{\nu_k}}} 1$;

$T_1(\alpha)$ auxiliary exponential integral "over integers", *i.e.* $T_1(\alpha) = \int_{\epsilon X}^{X} e(\alpha t)\mathrm{d}t$;

$T_2(\alpha)$ auxiliary exponential integral "over squares", *i.e.* $T_2(\alpha) = \int_{\sqrt{\epsilon X}}^{\sqrt{X}} e(\alpha t^2)\mathrm{d}t$;

$U_1(\alpha)$ auxiliary exponential sum over integers, *i.e.* $U_1(\alpha) = \sum_{\epsilon X \leq n \leq X} e(\alpha n)$;

$U_2(\alpha)$ auxiliary exponential sum over squares, *i.e.* $U_2(\alpha) = \sum_{\epsilon X \leq n^2 \leq X} e(\alpha n^2)$;

$\xi(d) = \min\{\ell > 0 : 2^\ell \equiv 1 \pmod d\}$, with d odd;

$\xi_g(d) = \min\{\ell > 0 : g^\ell \equiv 1 \pmod d\}$, with $(d,g) = 1$;

$\overline{\xi_g}(d) = \min\{\ell > 0 : g^\ell \equiv 1 \pmod{d/(d,g)}\}$.

1.2 Useful results

In this section we collect some classical results that will be useful in the following.

R.1 (**Davenport-Heilbronn's kernel function**).
Given the kernel function for the Davenport-Heilbronn method $K(\alpha,\eta) = \big(\frac{\sin \pi\eta\alpha}{\pi\alpha}\big)^2$ *for* $\alpha \neq 0$, *then both*

$$K(\alpha,\eta) \ll \min(\eta^2; \alpha^{-2}),$$

$$\widehat{K}(t,\eta) = \int_{\mathbb{R}} K(\alpha,\eta)e(t\alpha)\mathrm{d}\alpha = \max(0; \eta - |t|),$$

where $\widehat{K}(t,\eta)$ *is the Fourier transform of* $K(\alpha,\eta)$.

Proof. This lemma is a well-known result by Davenport-Heilbronn (see [DH46], Lemma 4). See also Lemma 20.1 of Davenport [Dav05] for the proof. □

R.2 (**Prime Number Theorem (PNT)**).
Given the prime counting function $\pi(x)$, then

$$\pi(x) \sim \frac{x}{\log x}, \qquad x \to \infty.$$

Proof. It was conjectured by Gauss in 1792 and by Legendre in 1798. Almost a century later, in 1896, Hadamard [Had96] and de la Vallée Poussin [dlVP96] independently proved it by using analytic arguments (see Apostol [Apo76], Chapter 13, for an analytic proof). An elementary (but quite intricate) proof of the PNT was discovered in 1949 by Selberg [Sel49] and Erdős [Erd49]. □

R.3 (**Equivalent forms of the PNT**).
The PNT is logically equivalent to each of the following

i) $\vartheta(x) \sim x, \qquad x \to \infty,$

ii) $\psi(x) \sim x, \qquad x \to \infty,$

where $\vartheta(x)$ and $\psi(x)$ are the Chebyshev functions.

Proof. See Apostol [Apo76], Theorem 4.4. □

R.4 (**PNT with error term**).
The strongest known form for the PNT is:

$$\psi(x) = x + \mathcal{O}\left(x \exp\left(\frac{-c(\log x)^{3/5}}{(\log\log x)^{1/5}}\right)\right),$$

for an absolute constant $c > 0$, where $\psi(x)$ is the Chebyshev function[1].

Proof. See Ivić [Ivi85], Theorem 12.2. □

R.5 (**Vinogradov-Korobov's zero-free region**).
There exists an absolute constant $c > 0$ such that $\zeta(s) \neq 0$ for $s = \sigma + \imath t$ with

$$\sigma \geq 1 - \frac{c}{(\log(|t|+2))^{2/3}(\log\log(|t|+2))^{1/3}}.$$

Proof. See Montgomery [Mon71], Corollary 11.4. See also Ivić [Ivi85], Chapter 6, for an exhaustive analysis of the zero-free region. □

R.6 (**Dirichlet's theorem on diophantine approximations**).
Let $\alpha \in \mathbb{R}$, then, for every $X \in \mathbb{R}_{\geq 1}$, there exists $a/q \in \mathbb{Q}$ such that $(a,q) = 1$, $1 \leq q \leq X$ and

$$\left|\alpha - \frac{a}{q}\right| \leq \frac{1}{qX}.$$

[1] Similar statements hold true for $\vartheta(x)$ and $\pi(x)$ too.

Proof. See Vaughan [Vau97], Lemma 2.1. □

R.7 (**Law of best approximation for continued fraction**).
Let $x \in \mathbb{R}$ and let $a/q \in \mathbb{Q}$ be a convergent of the continued fraction for x. If $0 < q' < q$ and $a'/q' \neq a/q$, then

$$|a - qx| < |a' - q'x|.$$

Proof. See Hardy-Wright [HW10], Theorem 182, for the proof. See also Chapter 10 for a survey over continued fractions. □

R.8 (**Rieger's estimation**).
Let $T(X) = |\{(p_1, p_2, p_3, p_4) \in \mathfrak{P}^4 \,:\, p_1^2 + p_2^2 = p_3^2 + p_4^2;\ p_1p_2 \neq p_3p_4;\ p_i^2 \leq X\ \forall i\}|$, then

$$T(X) \ll X(\log X)^{-3}.$$

Proof. [Rie68], Satz 3. See also the estimate of H_{12} at page 106 of T. Liu [Liu04]. □

R.9 (**Vaughan's estimation of exponential sum over primes**).
If $(a, q) = 1$, $q \leq X$ and $|\alpha - a/q| < 1/q^2$, then

$$S(\alpha) \ll \left(\frac{X}{\sqrt{q}} + \sqrt{qX} + X^{4/5}\right)(\log X)^4,$$

where $S(\alpha) = \sum_{\epsilon X \leq m \leq X} \Lambda(m)\, e(m\alpha)$ and $\alpha \in \mathbb{R}$.

Proof. See Vaughan [Vau97], Theorem 3.1, for the "$\log p$-version", see Davenport [Dav00], §25, for the "$\Lambda(m)$-version". This result is a refinement of a famous estimate by Vinogradov (see, *e.g*, [Vin04], Theorem 1 of Chapter IX). □

R.10 (**Ghosh's estimation of exponential sum over prime squares**).
If $a, q \in \mathbb{N} \setminus \{0\}$ is such that $(a, q) = 1$ and $|\alpha - a/q| < q^{-2}$, then for any $\epsilon > 0$

$$S_2(\alpha) \ll_\epsilon X^{1/2+\epsilon}\Big(\frac{1}{q} + \frac{1}{X^{1/4}} + \frac{q}{X}\Big)^{1/4},$$

where $S_2(\alpha) = \sum_{\epsilon X \leq p^2 \leq X} \log p\, e(p^2\alpha)$ and $\alpha \in \mathbb{R}$.

Proof. [Gho81], Theorem 2. □

R.11 (**Gallagher's lemma on the truncated L^2-norm of exponentials sums**).
Let $\alpha \in \mathbb{R}$ and let

$$\mathscr{S}(\alpha) = \sum_{x \in \mathcal{S}} s(x) e(x\alpha)$$

be an absolutely convergent exponential sum, with $s(x) \in \mathbb{C}$ and $\mathcal{S} \subseteq \mathbb{R}$ an arbitrary sequence of real numbers. Let moreover $0 < \theta < 1$, $X \in \mathbb{R}$ and $\delta = \theta/X$, then

$$\int_{-X}^{X} \Big|\mathscr{S}(\alpha)\Big|^2 \mathrm{d}\alpha \ll_\theta \int_{-\infty}^{+\infty} \Big|\delta^{-1} \sum_{|x-y|<\frac{\delta}{2}} s(x)\Big|^2 \mathrm{d}y.$$

Proof. [Gal70], Lemma 1. See also Montgomery [Mon71], Lemma 1.9 for the proof. □

R.12 (**Saffari-Vaughan's estimation of the Selberg integral**).
Let $\epsilon > 0$ be an arbitrarily small constant. Then there exists a constant $c(\epsilon) > 0$ such that

$$J(X,h) \ll_\epsilon h^2 X \exp\Big(-c(\epsilon)\Big(\frac{\log X}{\log\log X}\Big)^{1/3}\Big)$$

uniformly for $X^{1/6+\epsilon} \le h \le X$, where $J(X,h)$ is the Selberg integral.

Proof. [SV77], §6. □

R.13 (**Truncated explicit formula for $\psi(x)$**).
Let $\rho = \beta + i\gamma$ run over the complex zeros of $\zeta(s)$. Then

$$\psi(x) = x - \sum_{|\gamma|\le T} \frac{x^\rho}{\rho} + \mathcal{O}\Big(\frac{x}{T}\log^2(xT) + \log x\Big)$$

uniformly in $T \ge 2$, where $\psi(x)$ is the Chebyshev function.

Proof. See Ivić [Ivi85], Theorem 12.1. See also Davenport [Dav00], Chapter 17 for an exhaustive analysis of the explicit formula for $\psi(x)$. □

R.14 (**Riemann-von Mangoldt formula**).
Let $N(T)$ be the number of the zeros of $\zeta(s)$ that lie in the region $0 < \beta < 1$ and $0 < \gamma \le T$, then

$$N(T) = \frac{T}{2\pi}\log\Big(\frac{T}{2\pi}\Big) - \frac{T}{2\pi} + \mathcal{O}(\log T).$$

Proof. See Ivić [Ivi85], Theorem 1.7. □

R.15 (**Zero-density estimates**).
Let $N(\sigma,T)$ be the number of the zeros of $\zeta(s)$ that lie in the region $\sigma \le \beta \le 1$ and $|\gamma| \le T$, then

$$N(\sigma,T) \ll \begin{cases} T\log T & \text{if } 0 \le \sigma \le \frac{1}{2}, \\ t^{D(\sigma)(1-\sigma)}(\log T)^B & \text{if } \frac{1}{2} \le \sigma \le 1, \end{cases}$$

where $D(\sigma)(1-\sigma) \le 1$ and $\frac{\partial}{\partial\sigma}(D(\sigma)(1-\sigma)) < 0$. Moreover, $D(\sigma)$ satisfies the following Ingham-Huxley's estimates

$$D(\sigma) \le \begin{cases} \dfrac{3}{2-\sigma} & \text{if } \frac{1}{2} \le \sigma \le \frac{3}{4}, \\ \dfrac{3}{3\sigma-1} & \text{if } \frac{3}{4} \le \sigma \le 1, \\ \dfrac{12}{5} & \text{if } \frac{1}{2} \le \sigma \le 1. \end{cases}$$

Proof. See Ivić [Ivi85], Chapter 11, for an exhaustive analysis about zero-density estimates. In particular see Theorem 11.1 for Ingham-Huxley's estimates. □

R.16 (**Euler's summation formula**).
Let $\lfloor\cdot\rfloor$ denote the floor function and let $x, y \in \mathbb{R}$, with $0 < y < x$. If $f(t) \in C^1([y, x])$ with (continue) derivative denoted by $f'(t)$, then

$$\sum_{y<n\leq x} f(n) = \int_y^x f(t)\mathrm{d}t + \int_y^x (t - \lfloor t\rfloor)f'(t)\mathrm{d}t + (y - \lfloor y\rfloor)f(y) - (x - \lfloor x\rfloor)f(x).$$

Proof. See Apostol [Apo76], Theorem 3.1. □

R.17 (**Abel's identity**).
For any arithmetic function $a(n)$, let

$$A(t) = \sum_{n\leq t} a(n),$$

where $A(t) = 0$, if $t < 1$. Let moreover $x, y \in \mathbb{R}$, with $0 < y < x$. If $f(t) \in C^1([y, x])$ with (continue) derivative denoted by $f'(t)$, then

$$\sum_{y<n\leq x} a(n)f(n) = A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)\mathrm{d}t.$$

Proof. See Apostol [Apo76], Theorem 4.2. □

R.18 (**Mertens' and Vasil'kovskaja's theorems**).
Let γ be the Euler constant.

1. Mertens' theorem *claims*

$$\prod_{p\leq x}\Big(1 - \frac{1}{p}\Big) = \frac{e^{-\gamma}}{\log x}\big(1 + \mathcal{O}\big((\log x)^{-1}\big)\big).$$

2. Vasil'kovskaja's theorem *claims that, given $L(x) = \exp((\log x)^{3/5}(\log\log x)^{-1/5})$, there exists an absolute constant $c > 0$ such that*

$$\prod_{p\leq x}\Big(1 - \frac{1}{p}\Big) = \frac{e^{-\gamma}}{\log x}\big(1 + \mathcal{O}\big(L(x)^{-c}\big)\big).$$

3. *Similarly, we have that, given the constant $b = \gamma - \sum_p \sum_{k\geq 2}(kp^k)^{-1}$, then*

$$\sum_{p\leq x}\frac{1}{p} = \log\log x + b + \mathcal{O}\big((\log x)^{-1}\big).$$

Proof. See Montgomery-Vaughan [MV07], Theorem 2.7(e) and page 55, for the proof of Mertens' theorem. See Theorem 2.7(d) for the proof of point 3.

For the proof of Vasil'kovskaja's theorem, see Languasco-Zaccagnini [LZ07], Lemma 5 (or see the proof on pages 80-81 of Prachar [Pra78] inserting the Vinogradov-Korobov [Vin58, Kor58a, Kor58b] zero-free region for the Riemann zeta function). □

1.3 Useful techniques

In this section we briefly describe some classical techniques that will be used throughout our work.

1.3.1 Hardy-Littlewood circle method

The circle method has its genesis in a paper of Hardy and Ramanujan [HR18] and it is used to approach many additive problems. The method can be roughly summarized as follows:

1. Turning an additive problem over integers (*e.g.*, the Goldbach problem) into an analytic problem, by means of Fourier analysis;
2. Dissecting the obtained integration interval into major and minor arcs which respectively give the expected main term and the expected error term of the additive problem.

To better explain how the method works, we describe here its application to the ternary Goldbach problem, as solved by Vinogradov [Vin37]. For this part we refer to Davenport [Dav00], §26.

Recalling that $\mathfrak{P}$ denotes the set of all prime numbers, the relevant counting function for the ternary Goldbach problem is

$$r_3(N) = |\{(p_1,p_2,p_3) \in \mathfrak{P}^3 \,:\, N = p_1+p_2+p_3\}|,$$

defined for any odd integer $N > 5$. Instead of considering $r_3(N)$, we deal with the associated weighted function

$$R_3(N) = \sum_{m_1+m_2+m_3=N} \Lambda(m_1)\Lambda(m_2)\Lambda(m_3),$$

where $\Lambda(m)$ is the von Mangoldt function. We remark that, by definition of $\Lambda(m)$, $R_3(N)$ is actually a weighed counting of the number of representations of N as sum of three prime powers. However it's easy to see that the contribution made to $R_3(N)$ by proper prime powers is $\ll N^{3/2}(\log N)^2$ and this error term doesn't affect the final result, as will be clear later.

Let us now consider the exponential sum

$$S(\alpha) = \sum_{1\le m\le N} \Lambda(m)e(m\alpha),$$

with $e(x) = e^{2\pi\imath x}$. It is straightforward that

$$S^3(\alpha) = \sum_{n\ge 1}\Big(\sum_{\substack{m_1+m_2+m_3=n\\ m_i\le N}} \Lambda(m_1)\Lambda(m_2)\Lambda(m_3)\Big)e(n\alpha).$$

Hence, by the Fourier coefficient formula (see, *e.g.*, Apostol [Apo76], §11.4), we obtain the following fundamental relation

$$R_3(N) = \int_0^1 S^3(\alpha)e(-N\alpha)\mathrm{d}\alpha. \tag{1.1}$$

We shall get that the $|S(\alpha)|$ is small, unless α is near a rational number with a small denominator and, in each such a case, we can find a suitable approximating function to $S(\alpha)$. This suggests to dissect the integration interval of (1.1) into two parts, $\mathfrak{M}$ and $\mathfrak{m}$ say, where $\mathfrak{m}$ is the subset of $[0,1]$ of those numbers which are not near rational numbers with small denominators and $\mathfrak{M}$ is made of intervals close to such rationals.

So we now split the interval $[0,1]$ using (a variation of) the so called *Farey dissection* (see, *e.g.*, Hardy-Wright [HW10], §3): let $P = (\log N)^B$ and $Q = N/P$, where $B > 0$ will be chosen later in term of another parameter $A > 0$. Thus, $P < Q$ holds for any sufficiently large N. Then, for every $1 \leq a \leq q \leq P$ such that $(a,q) = 1$, we define the interval

$$\mathfrak{M}(a,q) = \Big[\frac{a}{q} - \frac{1}{Q}, \frac{a}{q} + \frac{1}{Q}\Big].$$

Thanks to the restriction $q \leq P$, such intervals are not-overlapping: in fact, taken $\mathfrak{M}(a_1,q_1)$ and $\mathfrak{M}(a_2,q_2)$ such that $a_1/q_1 \neq a_2/q_2$, then

$$\Big|\frac{a_1}{q_1} - \frac{a_2}{q_2}\Big| = \Big|\frac{a_1q_2 - a_2q_1}{q_1q_2}\Big| \geq \frac{1}{q_1q_2} \geq \frac{1}{P^2} > \frac{2}{Q},$$

since $(\log N)^{3B} < N/2$ for sufficiently large N; while, if we assume that there exists $\alpha \neq 0$ such that $\alpha \in \mathfrak{M}(a_1,q_1) \cap \mathfrak{M}(a_2,q_2)$, we get the contradiction

$$\Big|\frac{a_1}{q_1} - \frac{a_2}{q_2}\Big| \leq \Big|\frac{a_1}{q_1} - \alpha\Big| + \Big|\alpha - \frac{a_2}{q_2}\Big| \leq \frac{2}{Q}.$$

We define the *major arcs* to be

$$\mathfrak{M} = \bigsqcup_{1\leq q\leq P} \bigsqcup_{\substack{1\leq a\leq q\\(a,q)=1}} \mathfrak{M}(a,q).$$

It's easy to see that $\mathfrak{M} \subset \big[\frac{1}{Q}, 1+\frac{1}{Q}\big]$ and so we define the *minor arcs* to be $\mathfrak{m} = \big[\frac{1}{Q}, 1+\frac{1}{Q}\big] \setminus \mathfrak{M}$.

We remark that, since $S(\alpha)$ and $e(n\alpha)$ have period 1 in α, we can shift the integration interval of (1.1) to $\big(\frac{1}{Q}, 1+\frac{1}{Q}\big]$ and so split it according to the Farey dissection described above. In this way we obtain

$$\begin{aligned} R_3(N) &= \int_{\frac{1}{Q}}^{1+\frac{1}{Q}} S^3(\alpha)e(-N\alpha)\mathrm{d}\alpha \\ &= \int_{\mathfrak{m}} S^3(\alpha)e(-N\alpha)\mathrm{d}\alpha + \sum_{1\leq q\leq P} \sum_{\substack{1\leq a\leq q\\(a,q)=1}} \int_{\mathfrak{M}(a,q)} S^3(\alpha)e(-N\alpha)\mathrm{d}\alpha. \end{aligned}$$

Now we deal with major and minor arcs separately.

Major arcs

Let us consider $\alpha \in \mathfrak{M}(a,q)$, for some fixed a and q. Therefore $\alpha = a/q + \beta$ where β is a real number such that $|\beta| \leq 1/Q$. Arguing as in pages 146-147 of Davenport [Dav00], we obtain

$$S(\alpha) = \frac{\mu(q)}{\varphi(q)} T(\beta) + \mathcal{O}\Big(N \exp(-c\sqrt{\log N})\Big)$$

where $\mu(q)$ is the Möbius function, $\varphi(q)$ is the Euler function, $T(\beta) = \sum_{m \leq N} e(m\beta)$ and $c > 0$ is a suitable constant (in the following it will not be necessarily the same at each occurrence). It's easy to see that

$$\begin{aligned}\int_{\mathfrak{M}(a,q)} S^3(\alpha) e(-N\alpha) \mathrm{d}\alpha = \frac{\mu(q)}{\varphi^3(q)} e\Big(\frac{-aN}{q}\Big) \int_{-1/Q}^{1/Q} T^3(\beta) e(-N\beta) \mathrm{d}\beta \\ + \mathcal{O}\bigg(\frac{N^3}{Q} \exp(-c\sqrt{\log N})\bigg),\end{aligned}$$

and so, recalling $Q = N/P$, $P = (\log N)^B$, we have

$$\begin{aligned}\int_{\mathfrak{M}} S^3(\alpha) e(-N\alpha) \mathrm{d}\alpha = \sum_{1 \leq q \leq P} \frac{\mu(q)}{\varphi^3(q)} c_q(N) \int_{-1/Q}^{1/Q} T^3(\beta) e(-N\beta) \mathrm{d}\beta \\ + \mathcal{O}\Big(N^2 \exp(-c\sqrt{\log N})\Big), \qquad (1.2)\end{aligned}$$

where $c_q(N) = \sum_{1 \leq a \leq q;\, (a,q)=1} e\big(\frac{aN}{q}\big)$ is the Ramanujan sum (see, *e.g.*, Montgomery-Vaughan [MV07], page 110).

We now estimate the integral at the right-hand side of (1.2) observing that $T(\beta)$ is such that $T(\beta) = \sum_{m \leq N} (e^{2\pi\imath\beta})^k \ll \min(N; |\beta|^{-1})$, which implies

$$\int_{1/Q}^{1-1/Q} |T^3(\beta)| \mathrm{d}\beta \ll Q^2.$$

Using the Fourier coefficient formula, we easily see that

$$\int_0^1 T^3(\beta) e(-N\beta) \mathrm{d}\beta = \frac{(N-1)(N-2)}{2} = \frac{N^2}{2} + \mathcal{O}(N).$$

Recalling that the integrand function above has period 1 and that $Q = N/(\log N)^B$, we get

$$\begin{aligned}\int_{-1/Q}^{1/Q} T^3(\beta) e(-N\beta) \mathrm{d}\beta &= \int_0^1 T^3(\beta) e(-N\beta) \mathrm{d}\beta - \int_{1/Q}^{1-1/Q} T^3(\beta) e(-N\beta) \mathrm{d}\beta \\ &= \frac{N^2}{2} + \mathcal{O}\big(N^2 (\log N)^{-2B}\big). \qquad (1.3)\end{aligned}$$

To complete the estimation of (1.2), there is left to study the following sum over q

$$\sum_{1 \leq q \leq P} \frac{\mu(q)}{\varphi^3(q)} c_q(N) = \sum_{q \geq 1} \frac{\mu(q)}{\varphi^3(q)} c_q(N) - \sum_{q > P} \frac{\mu(q)}{\varphi^3(q)} c_q(N).$$

Using the Möbius inversion formula, we easily obtain that $|c_q(N)| \leq \varphi(q)$ and so

$$\sum_{q>P} \frac{\mu(q)}{\varphi^3(q)} c_q(N) \ll \sum_{q>P} \frac{1}{\varphi^2(q)} \ll (\log N)^{-B+1},$$

where the last inequality follows from, *e.g.*, Hardy-Wright [HW10], Theorem 327. By the Euler product formula (see, *e.g.*, Apostol [Apo76], §11.5), we also get

$$\sum_{q\geq 1} \frac{\mu(q)}{\varphi^3(q)} c_q(N) = \prod_{p|N} \Big(1 - \frac{1}{(p-1)^2}\Big) \prod_{p\nmid N} \Big(1 + \frac{1}{(p-1)^3}\Big) = \mathfrak{S}_3(N),$$

where $\mathfrak{S}_3(N)$ is the singular series for the ternary Goldbach problem. Therefore the major arcs contribution is

$$\begin{aligned} \int_{\mathfrak{M}} S^3(\alpha) e(-N\alpha) \mathrm{d}\alpha &= \big(\mathfrak{S}_3(N) + \mathcal{O}\big((\log N)^{-B+1}\big)\big)\Big(\frac{N^2}{2} + \mathcal{O}\big(N^2 (\log N)^{-2B}\big)\Big) \\ &\quad + \mathcal{O}\Big(N^2 \exp(-c\sqrt{\log N})\Big) \\ &= \mathfrak{S}_3(N)\frac{N^2}{2} + \mathcal{O}\big(N^2(\log N)^{-B+1}\big). \end{aligned} \tag{1.4}$$

Minor arcs

We want to prove that the order of magnitude of the minor arcs contribution is smaller than the major arcs asymptotic behaviour proved in (1.4). The first step is observing that

$$\Big| \int_{\mathfrak{m}} S^3(\alpha) e(-N\alpha) \mathrm{d}\alpha \Big| \leq \max_{\alpha\in\mathfrak{m}} |S(\alpha)| \int_0^1 |S^2(\alpha)| \mathrm{d}\alpha.$$

By orthogonality, the integral in the right-hand side is $\sum_{m\leq N} \Lambda^2(m) \ll N \log N$, where the asymptotic inequality is a direct application of the Prime Number Theorem (see Result (R.4)). By Dirichlet's theorem on diophantine approximations (see Result (R.6)), we can apply Vaughan's estimation in Result (R.9) and therefore, for every $\alpha \in \mathfrak{m}$, we have

$$|S(\alpha)| \ll \left(\frac{N}{\sqrt{q}} + \sqrt{qN} + N^{4/5}\right)(\log N)^4 \ll N(\log N)^{-B/2+4}. \tag{1.5}$$

Hence, the minor arcs contribution is $\mathcal{O}\big(N^2(\log N)^{-B/2+5}\big)$.

Collecting (1.1), (1.4) and (1.5) and setting $B = 2(A+5)$, we finally get

Theorem 1.1 (Vinogradov's theorem [Vin37])**.** *Let N be a sufficiently large integer. Then, for any fixed $A > 0$, we have*

$$R_3(N) = \mathfrak{S}_3(N)\frac{N^2}{2} + \mathcal{O}\big(N^2(\log N)^{-A}\big).$$

1.3.2 Pintz's explicit formula

Pintz's explicit formula in the next Theorem 1.3 (that we cite from [Pin09], §18) is the main tool in estimating the contribution of the major arcs (obtained by the Farey dissection), to the weighted function associated to the Goldbach problem. To be more precise: in estimating the following function

$$R_{\mathfrak{M}}(n) = \int_{\mathfrak{M}} S^2(\alpha)\, e(-n\alpha)\mathrm{d}\alpha,$$

where $e(x) = e^{2\pi\imath x}$ and $S(\alpha) = \sum_{1\leq m\leq X} \Lambda(m)e(m\alpha)$, with $\Lambda(m)$ the von Mangoldt function and X a large parameter. Moreover, if $P \in [X^{2/5}, X^{41/100}]$ and $Q = X/P$, we have

$$\mathfrak{M} = \bigsqcup_{1\leq q\leq P} \bigsqcup_{\substack{1\leq a\leq q\\(a,q)=1}} \left[\frac{a}{q} - \frac{1}{qQ}, \frac{a}{q} + \frac{1}{qQ}\right],$$

as in Section 1.3.1. Pintz's formula shows that, beyond the eventual excepitonal zero, we can exactly evaluate the effect of all the so called "generalised exceptional zeros" for the L-functions. The technique used is a generalization of the "exceptional zero technique" introduced by Montgomery-Vaughan in [MV75], where they exactly computed the contribution made by the possible Siegel zero, to the asymptotic formula in the Goldbach problem.

We recall a classical and fundamental result about the zero-free region, defining the exceptional zeros.

Theorem 1.2 (Zero-free region)**.** *There exists an absolute constant $c > 0$ such that if χ is a Dirichlet character modulo r, then the region*

$$\left\{s = \beta + \imath\gamma : \beta > 1 - \frac{c}{\log(r(|\gamma| + 2))}\right\}$$

contains no zeros of $L(s,\chi)$ unless χ is a real character, in which case $L(s,\chi)$ has at most one, necessarily real, zero $\beta \in (1 - c/(\log 2r), 1]$. Such a zero β is called exceptional *(or* Siegel*)* zero*.*

Proof. See Montgomery-Vaughan [MV07], Theorem 11.3. □

Proceeding as in Jutila [Jut77], we consider the rectangle

$$R(\sigma, T) = \{\beta + \imath\gamma : \sigma \leq \beta \leq 1; |\gamma| \leq T\},$$

and we define $N(\sigma, T, \chi)$ to be the number of zeros of the L-function $L(s,\chi)$ which lie in $R(\sigma, T)$. Further let

$$N^*(\sigma, T, P) = \sum_{1\leq r\leq P} \sideset{}{^*}\sum_{\chi \bmod r} N(\sigma, T, \chi),$$

be the number of zeros of $L(s,\chi)$, for any primitive Dirichlet character χ modulo $r \leq P$ (*i.e.* the asterisk indicates that the sum is over primitive characters), which lie in $R(\sigma, T)$. Any such a zero (resp. character, modulus) is called *generalized exceptional zero* (resp. *character*, *modulus*).

To state Pintz's explicit formula, we need to consider the set $\mathcal{E}(H,T)$ of all generalized exceptional zeros which lie in $R(1-H/\log X,T)$, such that $|\mathcal{E}(H,T)| = N^*(1-H/\log X,T,P)$, with H and T sufficiently large constants and $P \leq X^{4/9-\epsilon}$. We also need to introduce the following notation:

$$\rho_i = (1-\delta_i) + \imath\gamma_i, \qquad \delta_i \leq \frac{H}{\log X}, \qquad |\gamma_i| \leq T,$$

$$\chi_i \pmod{r_i}, \tag{1.6}$$

$$r_i \leq P, \qquad P \leq X^{4/9-\epsilon},$$

Finally we include, among the generalized exceptional zeros, the pole $\rho_0 = 1$ of $\zeta(s) = L(s,\chi_0)$, where χ_0 can be considered as the primitive character modulo 1. So we can define $\mathcal{E}_0(H,T) = \mathcal{E}(H,T) \cup \{1\}$ and

$$\mathcal{A}(\rho_i) = \begin{cases} -1 & \text{if } \rho_i \in \mathcal{E}(H,T), \\ 1 & \text{if } i = 0. \end{cases} \tag{1.7}$$

We are now ready to state the needed theorem.

Theorem 1.3 (Pintz's explicit formula)**.** *Let $P \in [X^{\theta-\epsilon}, X^{\theta}]$, with $\theta < 4/9$ and $\epsilon > 0$. Then for any even $n \in [X/2, X]$*

$$R_{\mathfrak{M}}(n) = \sum_{\rho_i,\rho_j \in \mathcal{E}_0(H,T)} \mathfrak{S}(\chi_i,\chi_j,n)\mathcal{A}(\rho_i)\mathcal{A}(\rho_j)\frac{\Gamma(\rho_i)\Gamma(\rho_j)}{\Gamma(\rho_i+\rho_j)}n^{\rho_i+\rho_j-1} + \mathcal{O}\big(X\big(e^{-cH} + T^{-1/2}\big)\big),$$

where $c > 0$ is a fixed absolute constant, $\mathcal{A}(\rho)$ is defined above in (1.7)*, $\Gamma(\rho)$ is the Gamma function and $\mathfrak{S}(\chi_i,\chi_j,n)$ is the so-called* generalized singular series*.*

Proof. [Pin09], Theorem 25. □

Remark 1.4. *The exact definition of the generalized singular series is very elaborate (see Pintz [Pin06b], §15, pages 244-245), however here we just need the following properties of its:*

i) $|\mathfrak{S}(\chi_i,\chi_j,n)| \leq \mathfrak{S}(\chi_0,\chi_0,n) = \mathfrak{S}(n)$*;*

ii) for any sufficiently small $\epsilon > 0$ we have

$$|\mathfrak{S}(\chi_i,\chi_j,n)| \leq \epsilon,$$

except when

$$\begin{cases} r_i \mid C(\epsilon)n \\ r_j \mid C(\epsilon)n \\ \operatorname{cond}(\chi_i\overline{\chi_j}) \geq \epsilon^{-3}, \end{cases}$$

where $C(\epsilon)$ is a suitable constant depending only on ϵ and $\operatorname{cond}(\chi)$ *denotes the conductor*[2] *of χ.*

[2]The conductor of χ is its smallest induced modulus (see *e.g.* Apostol [Apo76], page 171).

The theorem yields the following corollary by Languasco, Pintz and Zaccagnini [LPZ07], which gives an estimation and an asymptotic estimation of $R_{\mathfrak{M}}(n)$.

Corollary 1.5 (Languasco-Pintz-Zaccagnini)**.** *Theorem 1.3 implies that for every even $n \in [X/2, X]$, we have:*

$$R_{\mathfrak{M}}(n) \ll n\mathfrak{S}(n).$$

In particular, if there exists a sufficiently small absolute constant $\eta > 0$ and a constant $C'(\eta)$ depending only on η such that

$$r_i \nmid C'(\eta)n \tag{1.8}$$

for every generalized (primitive) exceptional character χ_i, then

$$|R_{\mathfrak{M}}(n) - n\mathfrak{S}(n)| \leq \eta n\mathfrak{S}(n).$$

Moreover, denoting by a^* the odd square-free part of any $a \in \mathbb{N}$, we have $r_i^* \gg (\log X)^2$ for any generalized exceptional modulus.

Finally, if K denotes the number of generalized (primitive) exceptional character $\chi_i \bmod r_i$, then

$$0 \leq K \leq C_1(\eta),$$

with $C_1(\eta)$ a positive constant depending only on η.

Proof. [LPZ07], Corollary 4.2. □

We remark that $n\mathfrak{S}(n)$ is the diagonal term for $\rho_i = \rho_j = 1$ in Pintz's explicit formula above.

We conclude this section listing two useful results about generalized exceptional zeros and moduli.

Lemma 1.6 (Peneva)**.** *Suppose that $P \geq 1$ and $T \geq 2$. For $41/42 \leq \sigma \leq 1$ we have*

$$N^*(\sigma, T, P) \ll (P^4T)^{\frac{3}{2}(1-\sigma)}.$$

Proof. [Pen01, Pen04], Lemma 3′. □

In our situation, recalling that $P \in [X^{2/5}, X^{41/100}]$ and so $P \leq X^{1/2}$, and assuming $T \leq P$ and $H/\log X \leq 1/42$, we get

$$|\mathcal{E}(H, T)| \ll (X^{5/2})^{\frac{3}{2}(H/\log X)} = e^{15H/4}$$

that is, roughly speaking, we can bound the number of exceptional zeros only using the exceptional moduli.

Lemma 1.7 (Pintz)**.** *If $L(s, \chi) = 0$, with $s = 1 - \delta$ and $\chi \pmod r$ real character, then*

$$\delta \gg r^{-1/2}.$$

Proof. [Pin76], Theorem 3. □

As set in (1.6), our generalized exceptional zeros ρ_i are such that $\delta_i \leq H/\log X$. Thus Lemma 1.7 implies

$$\frac{H}{\log X} \geq \delta_i \gg \frac{1}{\sqrt{r_i}} \quad \Rightarrow \quad r_i \gg \Big(\frac{\log X}{H}\Big)^2 \gg (\log X)^2,$$

recalling that H is a sufficiently large constant.

1.3.3 Davenport-Heilbronn method

The notation used in this section is in accordance with the original works and not with the rest of the thesis.

The Hardy-Littlewood method described in Section 1.3.1, deals with the solution of equations in integers and it can be used, more generally, fixing k and ℓ positive integers, to study the solubility of an homogeneous equation of the form

$$\lambda_1 x_1^k + \ldots + \lambda_\ell x_\ell^k = 0,$$

where the λ_i's are integers (see Vaughan [Vau97], §9) or, equivalently, are all in rational ratio. To deal with the case when some of the λ_i's are not in rational ratio, Davenport and Heilbronn in [DH46] developed a new method, that can be therefore seen as a variation of the Hardy-Littlewood one.

To explain how it works, we briefly describe here the proof the main theorem in [DH46] (for more details see also Davenport [Dav05], §20, and Vaughan [Vau97], §11).

Theorem 1.8 (Davenport-Heilbronn)**.** *Let $\lambda_1, \ldots, \lambda_5$ be non-zero real numbers, not all of the same sign and such that there exist $1 \leq i, j \leq 5$ such that λ_i/λ_j is irrational. Then there exist arbitrarily large integers P such that*

$$|\lambda_1 x_1^2 + \ldots + \lambda_5 x_5^2| < 1 \tag{1.9}$$

has solutions in $1 \leq x_i \leq P$ for every $1 \leq i \leq 5$.

Remark 1.9. *1. We can easily deduce the corresponding result for the inequality*

$$|\lambda_1 x_1^2 + \ldots + \lambda_5 x_5^2| < \eta,$$

for any $\eta > 0$, just replacing λ_i with λ_i/η for every $1 \leq i \leq 5$, and then applying Theorem 1.8.

2. We can easily prove the same result for the inequality

$$|\lambda_1 x_1^2 + \ldots + \lambda_5 x_5^2 - \mu| < \eta,$$

for any real number μ, just replacing η with $\eta + |\mu|$.

3. Let k be a positive integer. Using Hua's inequality in [Hua38a], we can replace the squares by k^{th}-powers, so obtaining the corresponding result for

$$|\lambda_1 x_1^k + \ldots + \lambda_\ell x_\ell^k - \mu| < \eta,$$

provided that $\ell \geq 2^k + 1$.

Proof of Theorem 1.8. We start by observing that we can assume, without any loss of generality, that $\lambda_1/\lambda_2 \notin \mathbb{Q}$ and $\lambda_1/\lambda_2 < 0$. Let us set

$$Q = Q(x_1, \ldots, x_5) = \lambda_1 x_1^2 + \ldots + \lambda_5 x_5^2.$$

Given a large integer $P > 0$, our aim is to prove that $\sum_{1\le x_1,\dots,x_5\le P,\ |Q|<1} 1$ is "large". More precisely, we will prove that

$$\sum_{\substack{1\le x_1,\dots,x_5\le P\\ |Q|<1}} 1 \gg P^3. \tag{1.10}$$

The first step is to construct a function in Q, say $f(Q)$, which is positive for $|Q| < 1$ and zero for $|Q| \ge 1$. So let

$$f(Q) = \int_{-\infty}^{\infty} e(\alpha Q)\Big(\frac{\sin \pi\alpha}{\pi\alpha}\Big)^2 \mathrm{d}\alpha,$$

where $e(x) = e^{2\pi\imath x}$. It's easy to see that $f(Q) = \max(0; 1 - |Q|)$. Now let us define

$$S(\alpha) = \sum_{x=1}^{P} e(\alpha x^2) \qquad \text{and} \qquad I(\alpha) = \int_0^P e(\alpha x^2)\mathrm{d}x.$$

By easy computations, we obtain

$$\begin{aligned}\int_{-\infty}^{+\infty} S(\lambda_1\alpha)\dots S(\lambda_5\alpha)\Big(\frac{\sin \pi\alpha}{\pi\alpha}\Big)^2 \mathrm{d}\alpha &= \sum_{1\le x_1,\dots,x_5\le P} \int_{-\infty}^{+\infty} e(\alpha(\lambda_1 x_1^2 + \dots + \lambda_5 x_5^2))\Big(\frac{\sin \pi\alpha}{\pi\alpha}\Big)^2 \mathrm{d}\alpha \\ &= \sum_{\substack{1\le x_1,\dots,x_5\le P\\ |Q|<1}} (1 - |Q|).\end{aligned} \tag{1.11}$$

Since $1 \ge |1 - |Q||$, for $|Q| < 1$, then, to prove (1.10), it is sufficient to prove that

$$\sum_{\substack{1\le x_1,\dots,x_5\le P\\ |Q|<1}} (1 - |Q|) \gg P^3.$$

Arguing as in (1.11), but with $I(\alpha)$ replacing $S(\alpha)$, we obtain

$$\int_{-\infty}^{+\infty} I(\lambda_1\alpha)\dots I(\lambda_5\alpha)\Big(\frac{\sin \pi\alpha}{\pi\alpha}\Big)^2 \mathrm{d}\alpha = \int_{\substack{1\le x_1,\dots,x_5\le P\\ |Q|<1}} (1 - |Q|)\mathrm{d}x_1\dots\mathrm{d}x_5. \tag{1.12}$$

Using this setting, the idea of the proof is the following:

1. Proving that the right-hand side of (1.12) is $\gg P^3$, for P sufficiently large;
2. Proving that the difference between the left-hand sides of (1.11) and (1.12) is $o(P^3)$, for P sufficiently large;
3. Thus the right-hand side of (1.11) is $\gg P^3$, as wanted.

The key step is the second one and, in particular, the difficulty in proving it lies in estimating the contribution made by those α such that[3]

$$\frac{1}{4(\max_i |\lambda_i|)P} < |\alpha| < P^\delta,$$

for any fixed $\delta > 0$, to the integral on the left-hand side of (1.11). Here the hypothesis about the irrationality of λ_1/λ_2 is crucial (for example, we will need the continued fraction expansion for λ_1/λ_2). □

[3]For the lower bound, we follow Davenport [Dav05], §20, since it allows a sharper result. In the original paper it is just $\frac{1}{P} < |\alpha| < P^\delta$.

Chapter 2

On a Diophantine problem with one prime, two squares of primes and k powers of two

In this chapter we introduce our work about a diophantine problem with one prime, two prime squares and k powers of 2. This problem was already studied by W.P. Li and Wang [LW05]. Here we refine their result, improving their lower bound for k. More precisely, in the first section, we state our result and we compare it with Li-Wang's work. In Section 2.2 we set up our problem and we fix the notation, while in Section 2.3 we collect the lemmas needed to prove our theorem. In particular, Lemmas 2.8-2.9 and Lemmas 2.12-2.13 (with Claims 2.A-2.B) are our new results, used to deal with exponential sums over prime squares. In Sections 2.4 -2.6 we prove the partial results we need to complete the proof of our theorem, which is the subject of the last Section 2.7.

2.1 Introduction to our result

The very starting point for our work is the beautiful paper [Par03] by Parsell in which the author investigated the values taken by real linear combinations of two primes and a bounded number of powers of 2. Roughly speaking, Parsell proved that, under certain conditions, these values can be made arbitrarily close to any real number, by taking sufficiently many powers of 2.

More precisely he proved the following

Theorem (Parsell). *Suppose that λ_1 and λ_2 are real numbers such that λ_1/λ_2 is negative and irrational. Further suppose that $\mu_1, \dots, \mu_k$ are nonzero real numbers such that, for some i and j, the ratios λ_1/μ_i and λ_2/μ_j are rational. Finally, fix $\eta > 0$. Then there exists an integer k_0, depending at most on λ's, μ's and η, such that for every real number ϖ and every integer $k > k_0$, the inequality*

$$| \lambda_1 p_1 + \lambda_2 p_2 + \mu_1 2^{m_1} + \ldots + \mu_k 2^{m_k} + \varpi | < \eta \tag{2.1}$$

has infinitely many solutions in primes p_1 and p_2 and positive integers $m_1, \dots, m_k$.

After rearranging and multiplying through by a suitable constant, we can suppose that λ_1/μ_1 *and* λ_2/μ_2 *are rational,* $\lambda_1 > 1$, $\lambda_2 < -1$ *and* $|\lambda_1/\lambda_2| \geq 1$. *With these conditions, the constant* k_0 *is given explicitly as*

$$k_0 = 2 + \left\lceil \frac{\log((1 - 8\lambda_1\epsilon)\eta) - \log(2\boldsymbol{C}|\lambda_1\lambda_2|)}{\log(0.954)} \right\rceil$$

where $\epsilon > 0$ *is an arbitrarily small constant and*

$$\boldsymbol{C} = \boldsymbol{C}(q_1, q_2) = 25\sqrt{\log(2q_1)}\sqrt{\log(2q_2)},$$

with $a_i/q_i = \lambda_i/\mu_i$ *the reduced fraction having* $q_i > 0$.

Proof. [Par03], Theorem 1 (see pages 365 and 371 for the explicit formula for k_0). □

The main ingredient in Parsell's proof is the Davenport-Heilbronn version of the Hardy-Littlewood method. That is, Parsell first reduced his problem to studying the integral over $\mathbb{R}$ of some suitable exponential sums over primes and over powers of 2. Then he split the integration interval into major, minor and trivial arcs and, finally, he estimated the respective integrals one by one:

- On the major arc, Parsell used the truncated explicit formula for $\psi(x)$ (see Result (R.13)), the zero-density estimate by Ingham-Huxley (see Result (R.15)) and the zero-free region for the Riemann ζ-function (see Result (R.5)).

- On the minor arc, Parsell used a standard argument involving continued fractions for the irrational number λ_1/λ_2, to estimate the exponential sum over primes (see *e.g.* Vaughan [Vau74], Lemma 11). Then he used a result by Heath–Brown and Putcha (see [HBP02], Lemma 1 and some techniques in §5) to prove that the contribution of the exponential sum over powers of 2 is small, except on a set of very small measure.

- On the trivial arc, Parsell used trivial estimates.

We mention that recently, Languasco and Zaccagnini in [LZ10] improved Parsell's theorem, lowering the value of k_0.

Replacing one of the prime summands in (2.1), with the sum of two prime squares, we obtain the problem of studying the values taken by the form

$$\lambda_1 p_1 + \lambda_2 p_2^2 + \lambda_3 p_3^2 + \mu_1 2^{m_1} + \cdots + \mu_k 2^{m_k}.$$

This problem naturally descends from the combination of Parsell's theorem with the Waring-Goldbach problem with prime squares (see, *e.g.*, Hua [Hua38b]). The only result we know about it, is the following theorem by W.P. Li and Wang

Theorem (Li-Wang). *Suppose that* λ_1, λ_2 *and* λ_3 *are nonzero real numbers not all of the same sign and such that* λ_2/λ_3 *is irrational. Further suppose that* $\mu_1, \ldots, \mu_k$ *are nonzero real numbers such that, for some* i, j *and* ℓ, *the ratios* $\lambda_1/\mu_i, \lambda_2/\mu_j$ *and* λ_3/μ_ℓ *are rational. Finally, fix* $\eta > 0$. *Then there exists an integer* k_0, *depending at most on* λ*'s,* μ*'s,* η *and*

ϵ, where $\epsilon > 0$ is an arbitrarily small number, such that for every real number ϖ and every integer $k > k_0$, the inequality

$$|\,\lambda_1 p_1 + \lambda_2 p_2^2 + \lambda_3 p_3^2 + \mu_1 2^{m_1} + \ldots + \mu_k 2^{m_k} + \varpi\,| < \eta$$

has infinitely many solutions in primes p_1, p_2 and p_3 and positive integers $m_1, \ldots, m_k$. The constant k_0 is given explicitly as

$$k_0 = 3 + \Big\lceil \frac{\log(\lambda_1\eta) - \log(2^9 \boldsymbol{C}(|\lambda_1| + |\lambda_2| + |\lambda_3|)^2)}{\log(0.995)} \Big\rceil$$

where

$$\boldsymbol{C} = \boldsymbol{C}(q_1, q_2, q_3, \epsilon) = 5(1+\epsilon)^5 \sqrt{\Big(\frac{11^4 \cdot 43 \cdot \pi^{26}}{2^{27} \cdot 25} + \frac{\log 2}{2}\Big)} \sqrt{\log(2q_1)} \sqrt{\log(2q_2)} \sqrt{\log(2q_3)},$$

where $a_i/q_i = \lambda_i/\mu_i$ is the reduced fraction having $q_i > 0$.

Proof. [LW05], Theorem 1 (see page 21 for the explicit formula for k_0). □

Li-Wang's proof is along the lines of Parsell's one, but with in addition exponential sums over prime squares: the authors basically adapted the same techniques used by Parsell, to deal with prime squares when needed, and they used Ghosh's estimate in [Gho81] (see Result (R.10)) to deal with the exponential sum over prime squares on the minor arc.

Our idea is to improve Li-Wang's result (that is lowering their value of k_0) by using different and more fitting techniques. In particular, we prove the following result (that we write in a slightly different form, with respect to Parsell's and Li-Wang's theorems, splitting it into a theorem and a corollary).

Theorem A (Languasco-Settimi [LS11])**.** *Suppose that $\lambda_1 < 0$, $\lambda_2, \lambda_3 > 0$ are real numbers such that λ_2/λ_3 is irrational. Further suppose that $\mu_1, \ldots, \mu_k$ are nonzero real numbers such that $\lambda_i/\mu_i \in \mathbb{Q}$ for $i \in \{1, 2, 3\}$, and denote by a_i/q_i the reduced fraction of λ_i/μ_i having $q_i > 0$. Finally let $\eta > 0$ be a sufficiently small constant such that*

$$\eta < \min\Big(\Big|\frac{\lambda_1}{a_1}\Big|;\ \frac{\lambda_2}{a_2};\ \frac{\lambda_3}{a_3}\Big).$$

Then there exists an integer k_0, depending at most on $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$, $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)$, η and ϵ_1, where $\epsilon_1 > 0$ is an arbitrarily small constant, such that, for every real number ϖ and every integer $k \geq k_0$, the inequality

$$|\,\lambda_1 p_1 + \lambda_2 p_2^2 + \lambda_3 p_3^2 + \mu_1 2^{m_1} + \ldots + \mu_k 2^{m_k} + \varpi\,| < \eta \tag{2.2}$$

has infinitely many solutions in primes p_1, p_2 and p_3 and positive integers $m_1, \ldots, m_k$. The constant k_0 is given explicitly as

$$k_0 = 3 + \Big\lceil \frac{\log((3 - 2\sqrt{2} - \epsilon_4)\eta) - \log\left(4\boldsymbol{C}(|\lambda_1| + \lambda_2 + \lambda_3)\right)}{\log(0.9505087500)} \Big\rceil, \tag{2.3}$$

where $\epsilon_4 > 0$ is an arbitrarily small constant[1] *and*

$$\boldsymbol{C} = \boldsymbol{C}(q_1, q_2, q_3, \epsilon_1) = (1+\epsilon_1)\Big(\log 2 + C \cdot \mathfrak{S}'(q_1)\Big)^{1/2} \times \Big(\log^2 2 + D \cdot \mathfrak{S}''(q_2)\Big)^{1/4}\Big(\log^2 2 + D \cdot \mathfrak{S}''(q_3)\Big)^{1/4}, \tag{2.4}$$

with $C = 10.0219168340$, $D = 17,646,979.6536361512$. Moreover

$$\mathfrak{S}'(n) = \prod_{\substack{p|n \\ p>2}} \frac{p-1}{p-2} \qquad \textit{and} \qquad \mathfrak{S}''(n) = \prod_{\substack{p|n \\ p>2}} \frac{p+1}{p}. \tag{2.5}$$

Corollary A (Languasco-Settimi)**.** *Suppose that $\lambda_1, \lambda_2, \lambda_3$ are nonzero real numbers, not all of the same sign and such that λ_2/λ_3 is irrational. Further suppose $\mu_1, \dots, \mu_k$ are nonzero real numbers such that $\lambda_i/\mu_i \in \mathbb{Q}$ for $i \in \{1,2,3\}$, and denote by a_i/q_i the reduced fraction of λ_i/μ_i having $q_i > 0$. Finally let $\eta > 0$ be a sufficiently small constant such that*

$$\eta < \min\Big(\Big|\frac{\lambda_1}{a_1}\Big|;\ \Big|\frac{\lambda_2}{a_2}\Big|;\ \Big|\frac{\lambda_3}{a_3}\Big|\Big)$$

and $\tau \geq \eta > 0$. Then there exists an integer k_0, depending at most on $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$, η and ϵ_1, where $\epsilon_1 > 0$ is arbitrarily small constant, such that, for every real number ϖ and every integer $k \geq k_0$ the inequality

$$|\ \lambda_1 p_1 + \lambda_2 p_2^2 + \lambda_3 p_3^2 + \mu_1 2^{m_1} + \ldots + \mu_k 2^{m_k} + \varpi\ | < \tau$$

has infinitely many solutions in primes p_1, p_2 and p_3 and positive integers $m_1, \dots, m_k$.

Proof of the corollary. The corollary is a direct consequence of Theorem A since, reordering the λ's, we can get $\lambda_1 < 0$, $\lambda_2, \lambda_3 > 0$. Hence the theorem assures us that the inequality (2.2) has infinitely many solutions and the corollary immediately follows from the condition $\tau \geq \eta$. □

The rest of this chapter is devoted to a detailed proof of our Theorem A. We just want to remark here that, with respect to [LW05], our main gain comes from enlarging the size of the major arc, since this allows us to use sharper estimates on the minor arc. In particular

- On the major arc, we replaced Parsell's technique used in [LW05], with an argument involving a L^2-estimate of the exponential sum over prime squares (*i.e.* $S_2(\alpha)$). This is a standard tool when working on primes (see, *e.g.*, Languasco-Zaccagnini [LZ10] for an application to a similar problem), but it seems that it is the first time that this kind of technique is used for prime squares (so the relevant Lemmas 2.12 and 2.13 below could be of some independent interest).
  - Lemma 2.12 is a variation, for prime squares, of Lemma 2.11 (that we cite from Brüdern-Cook-Perelli [BCP97], Lemma 1).

[1] We call ϵ_4 in such a way since it will be the fourth arbitrarily small positive constant in the following.

  - Lemma 2.13 is a variation, for prime squares, of Saffari-Vaughan's estimation of the Selberg integral described in Result (R.12) (in particular, Claim 2.A and Claim 2.B are variations of [SV77], Lemma 5 and Lemma 6 respectively).

- On the minor arc, to treat the exponential sums over primes (*i.e.* $S_1(\alpha)$) and over prime squares (*i.e.* $S_2(\alpha)$), we follow the argument in Lemma 4 of Languasco-Zaccagnini [LZ10], instead of Parsell's mean value estimate used in [LW05]. Moreover, to deal with the exponential sum over powers of two (*i.e.* $G(\alpha)$), we insert Pintz-Ruzsa's algorithm (see [PR03]) to estimate the measure of the subset of the minor arc on which $|G(\alpha)|$ is "large", instead of Heath–Brown-Putcha's result used in [Par03] and [LW05]. These ingredients lead to a sharper estimate on the minor arc and let us reduce the lower bound k_0. More precisely

  - In Lemma 2.5, we use Lemma 4 of Languasco-Zaccagnini [LZ10] to treat $S_1(\alpha)$. and in Lemma 2.8 we developed a technique, analogous to Languasco-Zaccagnini's one, to treat $S_2(\alpha)$.

    Our Lemmas 2.5 and 2.8 (as we will analyze in details after their statements) improve the numerical constants in the definition (2.4) of $\boldsymbol{C}$, comparing with Li-Wang's ones.

    Moreover we remark that the works of Rosser-Schoenfeld [RS62] on $n/\varphi(n)$ and of Solé-Planat [SP11] on the Dedekind Ψ-function (see Lemmas 2.2 and 2.3 below) give for $\mathfrak{S}'(q)$ and $\mathfrak{S}''(q)$ sharper estimates, for large values of q, than $2\log(2q)$ used by Li-Wang in the definition of $\boldsymbol{C}$. This leads to another improvement in the value of $\boldsymbol{C}$.
  - To deal with the exponential sum over powers of two (*i.e.* $G(\alpha)$), we insert Pintz-Ruzsa's algorithm (defined in [PR03]). This ingredient allows us to improve the absolute value of the denominator of k_0 (compare our k_0 in (2.3) with Li-Wang's one). Comparing only such denominators, we see that our gain is about 90%.

In practice, the following example shows that the gain is actually slightly larger than 90%. For instance, taking $\lambda_1 = -\sqrt{5} = \mu_1^{-1}$, $\lambda_2 = \sqrt{3} = \mu_2^{-1}$, $\lambda_3 = \sqrt{2} = \mu_3^{-1}$, $\eta = 1$ and $\epsilon_1 = \epsilon_4 = 10^{-20}$, we get $k_0 = 286$, while W.P. Li-Wang's estimate gives $k_0 = 4120$.

2.2 Davenport-Heilbronn method

In this section we set the notation, in order to use (a variation of) the Davenport-Heilbronn method described in Section 1.3.3, to count the solutions of the inequality (2.2) in the statement of Theorem A.

Let ϵ, ϵ_1 be two sufficiently small positive constants, X be a large parameter, $M = |\mu_1| + \ldots + |\mu_k|$ and $L = \log_2(\epsilon X/(2M))$. Our aim is therefore to find the number $\mathfrak{N}(X)$ of solutions of the inequality (2.2), with $\epsilon X \le p_1, p_2^2, p_3^2 \le X$ and $1 \le m_1, \ldots, m_k \le L$.

Setting $e(u) = \exp(2\pi\imath u)$, the needed exponential sums are

$$S_1(\alpha) = \sum_{\epsilon X \le p \le X} \log p \; e(p\alpha) \qquad \text{exponential sum over primes,}$$

$$S_2(\alpha) = \sum_{\epsilon X \le p^2 \le X} \log p \; e(p^2\alpha) \qquad \text{exponential sum over prime squares,}$$

$$G(\alpha) = \sum_{1 \le m \le L} e(2^m\alpha) \qquad \text{exponential sum over powers of 2.}$$

The kernel function for the Davenport-Heilbronn method is $K(\alpha,\eta) = \left(\frac{\sin \pi\eta\alpha}{\pi\alpha}\right)^2$, for $\alpha \neq 0$. By Result (R.1) it verifies

K.i) $\widehat{K}(t,\eta) = \max(0\,;\,\eta - |t|)$, where $\widehat{K}(t,\eta)$ is the Fourier transform of $K(\alpha,\eta)$;

K.ii) $K(\alpha,\eta) \ll \min(\eta^2\,;\,\alpha^{-2})$.

If we define

$$I(X;\mathcal{S}) = \int_{\mathcal{S}} S_1(\lambda_1\alpha)S_2(\lambda_2\alpha)S_2(\lambda_3\alpha)G(\mu_1\alpha)\cdots G(\mu_k\alpha)e(\varpi\,\alpha)K(\alpha,\eta)\mathrm{d}\alpha,$$

for any $\mathcal{S} \subseteq \mathbb{R}$, then, by (K.i), it follows that

$$I(X;\mathbb{R}) \ll \eta(\log X)^3 \cdot \mathfrak{N}(X).$$

So we want to prove that, for $X \to +\infty$ running over a suitable integral sequence, we have

$$I(X;\mathbb{R}) \gg_{k,\boldsymbol{\lambda},\epsilon} \eta^2 X(\log X)^k, \tag{2.6}$$

where $\boldsymbol{\lambda} = (\lambda_1,\lambda_2,\lambda_3)$, thus obtaining

$$\mathfrak{N}(X) \gg_{k,\boldsymbol{\lambda},\epsilon} \eta X(\log X)^{k-3},$$

which implies Theorem A.

To prove the estimate (2.6), we dissect the real line into major $\mathfrak{M}$, minor $\mathfrak{m}$ and trivial $\mathfrak{t}$ arcs, by choosing $P = X^{2/5}/\log X$ and letting

$$\mathfrak{M} = \{\alpha \in \mathbb{R} \,:\, |\alpha| \le P/X\}, \quad \mathfrak{m} = \{\alpha \in \mathbb{R} \,:\, P/X < |\alpha| \le L^2\}, \quad \mathfrak{t} = \mathbb{R} \setminus (\mathfrak{M} \cup \mathfrak{m}). \tag{2.7}$$

Accordingly, we write

$$I(X;\mathbb{R}) = I(X;\mathfrak{M}) + I(X;\mathfrak{m}) + I(X;\mathfrak{t}). \tag{2.8}$$

Our aim is to prove that:

- On trivial arc,

$$|I(X;\mathfrak{t})| = o\big(XL^k\big) \tag{2.9}$$

  holds for all sufficiently large X.

- On major arc,

$$I(X;\mathfrak{M}) \ge c_1\eta^2 XL^k, \tag{2.10}$$

  with $c_1 = c_1(\epsilon_4,\boldsymbol{\lambda}) > 0$ constant.

- On minor arc,

$$|I(X;\mathfrak{m})| \leq c_2(k)\eta X L^k \tag{2.11}$$

holds for $X \to +\infty$ running over a suitable integral sequence, where $c_2(k) > 0$ depends only on k, $c_2(k) \to 0$ as $k \to +\infty$ and

$$c_1\eta - c_2(k) \geq c_3\eta, \tag{2.12}$$

for some absolute constant $c_3 > 0$ and for $k \geq k_0$.

Inserting (2.9)-(2.12) into (2.8), we finally obtain that the estimate in (2.6) holds, thus proving Theorem A.

2.3 Lemmas

In this section we collect the lemmas needed to prove the estimates (2.9)-(2.11). Some of these lemmas are known results; hence in many cases we will just give the relative references. Other ones are new and in this case we will prove them in details. In particular, we will prove Lemmas 2.8-2.9 and Lemmas 2.12-2.13 (with Claims 2.A-2.B) since they are needed to deal with exponential sums over prime squares.

Let $1 \leq n \leq (1-\epsilon)X/2$ be an integer and p_1, p_2 two prime numbers. We define the *twin-prime counting function* as follows

$$Z(X;2n) = \sum_{\epsilon X \leq p_1 \leq X} \sum_{\substack{p_2 \leq X \\ p_2 - p_1 = 2n}} \log p_1 \log p_2. \tag{2.13}$$

We recall that we denote by $\mathfrak{S}(n)$ the singular series for the Goldbach problem, and so $\mathfrak{S}(n) = 2c_0\mathfrak{S}'(n)$, where $\mathfrak{S}'(n)$ is defined in (2.5) and c_0 is the twin prime constant

$$c_0 = \prod_{p>2} \Big(1 - \frac{1}{(p-1)^2}\Big).$$

Notice that $\mathfrak{S}'(n)$ is a multiplicative function[2] and, according to Gourdon-Sebah [GS01], we have $0.66016181584 < c_0 < 0.66016181585$.

Further, given $\ell \in \mathbb{N} \setminus \{0\}$, let $r_{\ell,\ell}(m)$ be the number of representations of an integer m as $\sum_{i=1}^{\ell} 2^{u_i} - \sum_{i=1}^{\ell} 2^{v_i}$, with $1 \leq u_i, v_i \leq L$ integers, so that $r_{\ell,\ell}(m) = 0$ for sufficiently large $|m|$. Let us define

$$S(\ell, L) = \sum_{m \in \mathbb{Z}\setminus\{0\}} r_{\ell,\ell}(m)\mathfrak{S}(m).$$

The first lemma is about the behaviour of $S(\ell, L)$ for sufficiently large X.

[2] $\mathfrak{S}'(n)$ is the multiplicative part of $\mathfrak{S}(n)$

Lemma 2.1 (Khalfalah-Pintz). *For any given $\ell \geq 1$, there exists $A(\ell) \in \mathbb{R}$ such that*

$$\lim_{L\to+\infty}\Big(\frac{S(\ell,L)}{2L^{2\ell}}-1\Big)=A(\ell).$$

Proof. [KP06], Theorem 1. □

Moreover, Khalfalah-Pintz proved in Theorem 2 of [KP06] some numerical estimates for $A(\ell)$, when $1 \leq \ell \leq 7$. We just need here

$$A(1) < 0.2792521041. \tag{2.14}$$

The second lemma is an upper bound for $\mathfrak{S}'(n)$.

Lemma 2.2 (Languasco-Zaccagnini). *For $n \in \mathbb{N}$, $n \geq 3$, we have that*

$$\mathfrak{S}'(n) < \frac{n}{c_0\varphi(n)} < \frac{e^{\gamma}\log\log n}{c_0} + \frac{2.50637}{c_0\cdot\log\log n},$$

where $\gamma \approx 0.5772156$ is the Euler constant.

Proof. See [LZ10], Lemma 2. We just remark here that, for $n \geq 3$, the first estimate follows immediately observing that

$$\mathfrak{S}'(n)=\prod_{\substack{p|n\\ p>2}}\frac{(p-1)^2}{p(p-2)}\prod_{\substack{p|n\\ p>2}}\frac{p}{p-1}<\prod_{p>2}\frac{(p-1)^2}{p(p-2)}\prod_{p|n}\frac{p}{p-1}=\frac{1}{c_0}\frac{n}{\varphi(n)}.$$

The second estimate is a direct application of Theorem 15 of Rosser and Schoenfeld [RS62] which claims that, for every $n \geq 3$, we have

$$\frac{n}{\varphi(n)} < e^{\gamma}\log\log n + \frac{2.50637}{\log\log n}.$$

□

Let us now define

$$f(n)=\begin{cases} 1 & \text{for } n=1,2,\\ \frac{n}{c_0\varphi(n)} & \text{for } n\geq 3.\end{cases}$$

For every $n \geq 1$, the inequality $\mathfrak{S}'(n) \leq f(n)$ is sharper than Parsell's estimate $\mathfrak{S}'(n) \leq 2\log(2n)$ (see the estimate of $h(q)$ at page 369 of [Par03]). Since it is clear that computing the exact value of $f(n)$ for large values of n it is not easy (because it requires the knowledge of every prime factor of n), we also remark that, for every $n \geq 14$, the second estimate in Lemma 2.2 leads to a sharper bound than $\mathfrak{S}'(n) \leq 2\log(2n)$ used in [Par03] and [LW05].

The next lemma is an upper bound for $\mathfrak{S}''(n)$, which is defined in (2.5). We first remark that it is connected with the Dedekind Ψ-function $\Psi(n) = n\prod_{p|n}(p+1)/p$, since

$$\mathfrak{S}''(n)=\begin{cases}\frac{\Psi(n)}{n} & \text{for } n \text{ odd},\\ \frac{2\Psi(n)}{3n} & \text{for } n \text{ even}.\end{cases}$$

Lemma 2.3. *For $n \in \mathbb{N}$, $n \geq 31$, we have that*

$$\mathfrak{S}''(n) < e^{\gamma} \log\log n,$$

where $\gamma \approx 0.5772156$ is the Euler constant.

Proof. Clearly $\mathfrak{S}''(n) \leq \Psi(n)/n$, so the lemma immediately follows from Corollary 2 of Solé-Planat [SP11] which claims that $\Psi(n)/(n \log\log n) < e^{\gamma}$, for every $n > 30$. □

For every $n \geq 31$, the estimate in Lemma 2.3 is sharper than Li-Wang's one (that is $\mathfrak{S}''(n) \leq 2\log(2n)$, as in the last line of page 279 of [LW05]). We also remark that $\mathfrak{S}''(1) = \mathfrak{S}''(2) = 1$ and the computation of $\mathfrak{S}''(n)$ in the remaining interval $3 \leq n \leq 30$ is an easy task.

The next lemma is a famous result of Bombieri and Davenport that is needed in the proof of Lemma 2.5 below.

Lemma 2.4 (Bombieri-Davenport)**.** *There exists a positive constant B such that, for every positive integer n, we have*

$$Z(X; 2n) < B\mathfrak{S}(n)X,$$

where $Z(X; 2n)$ is defined in (2.13) *and $\mathfrak{S}(n)$ is the singular series for the Goldbach problem, provided that X is sufficiently large.*

Proof. [BD66], Theorem 2. □

Chen [Che78] proved that $B = 3.9171$ can be used in Lemma 2.4. The assumption of a suitable form of the twin prime conjecture (*i.e.* $Z(X; 2n) \sim \mathfrak{S}(n)X$, for $X \to +\infty$) implies that in this case we can take $B = 1 + \epsilon$, for every positive ϵ.

Now we state some lemmas we need to estimate $I(X; \mathfrak{m})$. The first one is an improvement of the mean-square estimate in Parsell [Par03], Lemma 3.

Lemma 2.5 (Languasco-Zaccagnini)**.** *Let X be a sufficiently large parameter and let $\lambda, \mu \neq 0$ be two real numbers such that $\lambda/\mu \in \mathbb{Q}$. Let $a, q \in \mathbb{Z} \setminus \{0\}$ be such that $q > 0$, $(a, q) = 1$ and $\lambda/\mu = a/q$. Let further $0 < \eta < |\lambda/a|$. We have*

$$\int_{\mathbb{R}} |S_1(\lambda\alpha)G(\mu\alpha)|^2 K(\alpha, \eta) \mathrm{d}\alpha < \eta X L^2 \Big((1-\epsilon)\log 2 + C \cdot \mathfrak{S}'(q) \Big) + \mathcal{O}_{M,\epsilon}(\eta X L),$$

where $C = 2B(1 + A(1)) = 10.0219168340$, $B = 3.9171$ is the constant in Lemma 2.4 and $A(1)$ is estimated in (2.14)*. Finally, let ϵ, L, M be as are defined at the beginning this section.*

Proof. [LZ10], Lemma 4. □

The constant $C = 10.0219168340$ should be compared with the value $C_1 = 11.4525218267$ obtained in [Par03] (we remark that in the proof of Lemma 3, Parsell also used the worst estimate $\mathfrak{S}'(n) \leq 2\log(2n)$). We also remark that, assuming the twin prime conjecture in Lemma 2.4 and taking $B = 1 + 10^{-20}$, we get $C = 2.5585042083$.

The next step is to find a mean value estimate for the exponential sum over prime squares. Instead of the estimate used by Li-Wang in [LW05], we prove a new result (see Lemma 2.8 below), following the argument used by Languasco and Zaccagnini in the proof of Lemma 2.5. In order to do this, we need the following Lemmas 2.6-2.7.

Lemma 2.6. *Let $\epsilon_1 > 0$ be an arbitrarily small constant. For $n \in \mathbb{Z}$, $n \neq 0$, $|n| \leq X$ and $n \equiv 0 \bmod 24$, let us define*

$$r(n) = \big|\big\{(p_1,p_2,p_3,p_4) \in \mathfrak{P}^4 : p_1^2+p_2^2-p_3^2-p_4^2 = n;\ p_j \leq X^{1/2},\ j=1,\dots,4\big\}\big|,$$

where $\mathfrak{P}$ is the set of all prime numbers. We have

$$r(n) \leq (1+\epsilon_1)c_4\frac{\pi^2}{16}\mathfrak{S}_-(n)\frac{X}{\log^4 X},$$

with $c_4 = (101)\cdot 2^{20}$ and

$$\mathfrak{S}_-(n) = \left(2-\frac{1}{2^{m_0-1}}-\frac{1}{2^{m_0}}\right)\prod_{\substack{p>2\\ p^m\|n\\ m\geq 0}}\left(1+\frac{1}{p}-\frac{1}{p^{m+1}}-\frac{1}{p^{m+2}}\right),$$

where $m_0 \in \mathbb{N}$ such that $2^{m_0} \parallel n$.

Proof. It follows by inserting the remark at page 385 of H. Li [Li06] in the proof of Lemma 2.2 of J. Liu-Lü [LL04]. □

We immediately remark that $\mathfrak{S}_-(n) \leq 2\mathfrak{S}''(n)$. In fact $\left(2-\frac{1}{2^{m_0-1}}-\frac{1}{2^{m_0}}\right) \leq 2$, for every $m_0 \in \mathbb{N}$, and

$$\prod_{\substack{p>2\\ p^m\|n\\ m\geq 0}}\left(1+\frac{1}{p}-\frac{1}{p^{m+1}}-\frac{1}{p^{m+2}}\right) \leq \prod_{\substack{p|n\\ p>2}}\left(1+\frac{1}{p}\right) = \mathfrak{S}''(n).$$

We also need the following result by H. Li

Lemma 2.7 (H. Li)**.** *Let d be a positive odd integer, let us define the quantity*

$$\xi(d) = \min\{\ell : 2^\ell \equiv 1 \pmod d\}.$$

Then the series

$$\sum_{\substack{d=1\\ 2\nmid d}}^{+\infty}\frac{\mu^2(d)}{d\xi(d)}$$

is convergent and its value is $c_5 < 1.620767$.

Proof. [Li06], Lemma 4. □

We are now ready to state our result for the exponential sum over prime squares, which is the analogue of Lemma 2.5, but with prime squares instead of primes.

Lemma 2.8. *Let X be a sufficiently large parameter and let $\lambda, \mu \neq 0$ be two real numbers such that $\lambda/\mu \in \mathbb{Q}$. Let $a, q \in \mathbb{Z}\setminus\{0\}$ be such that $q > 0$, $(a,q) = 1$ and $\lambda/\mu = a/q$. Let further $0 < \eta < |\lambda/a|$. We have*

$$\int_{\mathbb{R}} |S_2(\lambda\alpha)G(\mu\alpha)|^4 K(\alpha,\eta)\mathrm{d}\alpha < (1+\epsilon_1)\eta X L^4\Big(\log^2 2 + D\cdot\mathfrak{S}''(q)\Big),$$

where $\epsilon_1 > 0$ is an arbitrarily small constant and $D = c_4c_5\cdot 2^{-5}\cdot 3^{-1}\cdot\pi^2$, with c_4, c_5 defined as in Lemmas 2.6-2.7 respectively.

Lemma 2.8 should be compared with Lemma 4.2 of [LW05] (see also Lemma 4.3 of W.P. Li-Wang [LW07]) in which the value $D_1 = 2^{-27} \cdot 5^{-2} \cdot 11^4 \cdot 43 \cdot \pi^{26}$ plays the role of our D. Using the values $c_4 = 101 \cdot 2^{20}$ and $c_5 < 1.620767$ as in Lemmas 2.6-2.7, we see that $D < 17,646,979.6536361512$ while $D_1 = 1,581,925,383.0798448770$. We remark that $D < (0.0112) \cdot D_1$ and so the reduction factor here is close to the 98.8%. With an abuse of notation, in the statement of Theorem A we set $D = 17,646,979.6536361512$.

Proof. This proof is along the lines of Languasco-Zaccagnini's proof of Lemma 2.5. Letting

$$I = \int_{\mathbb{R}} |S_2(\lambda\alpha)G(\mu\alpha)|^4 K(\alpha,\eta)\mathrm{d}\alpha,$$

by definitions of $S_2(\alpha)$ and $G(\alpha)$ and by the property (K.i) of $K(\alpha,\eta)$, we immediately have

$$\begin{aligned} I = & \sum_{\epsilon X \le p_1^2,p_2^2,p_3^2,p_4^2 \le X} \log p_1 \log p_2 \log p_3 \log p_4 \\ & \times \sum_{1 \le m_1,m_2,m_3,m_4 \le L} \max\Big(0; \eta - |\lambda(p_1^2+p_2^2-p_3^2-p_4^2) + \mu(2^{m_1}+2^{m_2}-2^{m_3}-2^{m_4})|\Big). \quad (2.15) \end{aligned}$$

Let $\delta = \lambda(p_1^2+p_2^2-p_3^2-p_4^2) + \mu(2^{m_1}+2^{m_2}-2^{m_3}-2^{m_4})$. For a sufficiently small $\eta > 0$, we claim that

$$|\delta| < \eta \quad \Leftrightarrow \quad \delta = 0. \quad (2.16)$$

Recalling the hypothesis on a and q and assuming that $\delta \neq 0$ in (2.16), then, for $\eta < |\lambda/a|$ as in the hypothesis, the assumption $|\delta| < \eta$ leads to a contradiction. In fact we have

$$\begin{aligned} \frac{1}{|a|} > \frac{\eta}{|\lambda|} &> \Big| p_1^2+p_2^2-p_3^2-p_4^2 + \frac{q}{a}(2^{m_1}+2^{m_2}-2^{m_3}-2^{m_4}) \Big| \\ &= \Big| \frac{a(p_1^2+p_2^2-p_3^2-p_4^2) + q(2^{m_1}+2^{m_2}-2^{m_3}-2^{m_4})}{a} \Big| \ge \frac{1}{|a|}, \end{aligned}$$

since $a(p_1^2+p_2^2-p_3^2-p_4^2) + q(2^{m_1}+2^{m_2}-2^{m_3}-2^{m_4}) \neq 0$ is a linear integral combination. Inserting (2.16) in (2.15), we can write that

$$I = \eta \sum_{\substack{\epsilon X \le p_1^2,p_2^2,p_3^2,p_4^2 \le X \\ \lambda(p_1^2+p_2^2-p_3^2-p_4^2)+\mu(2^{m_1}+2^{m_2}-2^{m_3}-2^{m_4})=0}} \sum_{1 \le m_1,m_2,m_3,m_4 \le L} \log p_1 \log p_2 \log p_3 \log p_4. \quad (2.17)$$

Diagonal contribution. The diagonal contribution in (2.17) is equal to

$$\eta \sum_{\substack{\epsilon X \le p_1^2,p_2^2,p_3^2,p_4^2 \le X \\ p_1^2+p_2^2=p_3^2+p_4^2}} \log p_1 \log p_2 \log p_3 \log p_4 \sum_{\substack{1 \le m_1,m_2,m_3,m_4 \le L \\ 2^{m_1}+2^{m_2}=2^{m_3}+2^{m_4}}} 1. \quad (2.18)$$

The number of the solutions of $p_1^2+p_2^2 = p_3^2+p_4^2$, when $p_1p_2 \neq p_3p_4$, can be estimated using Rieger's Theorem (see Result (R.8)) and it is $\ll X(\log X)^{-3}$. This gives a contribution to

the first sum in (2.18) which is $\ll X\log X$. In the remaining case $p_1p_2=p_3p_4$, the first sum in (2.18) becomes

$$2\sum_{\epsilon X\le p_1^2,p_2^2\le X}\log^2 p_1\log^2 p_2 = 2\Big(\sum_{\sqrt{\epsilon X}\le p\le\sqrt{X}}\log^2 p\Big)^2 \le 2\left(\frac{\log X}{2}\right)^2\Big(\vartheta(\sqrt{X})-\vartheta(\sqrt{\epsilon X})\Big)^2$$
$$=\frac{(\log X)^2}{2}X(1-\sqrt{\epsilon})^2+o_\epsilon\big(X(\log X)^2\big)<(1-\epsilon)\frac{X}{2}(\log X)^2,$$

where we used the Prime Number Theorem for the ϑ-function (see Result (R.3)) and the fact that $\epsilon>0$ is a sufficiently small constant. The sum over powers of 2 in (2.18) can be evaluated by fixing first $m_1=m_3$ (thus getting exactly L^2 solutions) and then fixing $m_1\neq m_3$ (which gives other L^2-L solutions). Hence the contribution of the second sum in (2.18) is $2L^2-L$.

Combining these results we get that the total contribution of (2.18) is

$$<(1-\epsilon)\eta XL^2(\log X)^2<\eta XL^4(\log 2)^2.\blacksquare \tag{2.19}$$

Non-diagonal contribution. Now we have to estimate the contribution, say I', of the non-diagonal solutions of $\delta=0$ in (2.17). We achieve this by connecting I' with the singular series $\mathfrak{S}_-(n)$ of Lemma 2.6. First, we remark that

$$p_j>3,\ \forall j=1,\dots,4 \quad\Rightarrow\quad n=p_1^2+p_2^2-p_3^2-p_4^2\equiv 0 \bmod 24.$$

We now study the case $n\not\equiv 0 \bmod 24$ and $n\equiv 0 \bmod 24$ one by one.

- Let $n=p_1^2+p_2^2-p_3^2-p_4^2\not\equiv 0 \bmod 24$, then at least one of the p_j's must be equal to 2 or 3 and hence $r(n)$, defined as in the statement of Lemma 2.6, verifies[3] $r(n)\ll X^{1/2+\epsilon}$. Therefore, recalling that $\lambda/\mu=a/q\neq 0$ and $(a,q)=1$, if $2^{m_3}+2^{m_4}-2^{m_1}-2^{m_2}\neq 0$ and $(q/a)(2^{m_3}+2^{m_4}-2^{m_1}-2^{m_2})\not\equiv 0 \bmod 24$, then we have

$$\big|\big\{(p_1,\dots,p_4)\in\mathfrak{P}\,:\,p_1^2+p_2^2-p_3^2-p_4^2=\frac{q}{a}(2^{m_3}+2^{m_4}-2^{m_1}-2^{m_2})\big\}\big|\ll X^{1/2+\epsilon}.$$

- Let $n=p_1^2+p_2^2-p_3^2-p_4^2\equiv 0 \bmod 24$. By definition of L and M, we have

$$\Big|\Big(\frac{q}{a}\Big)(2^{m_3}+2^{m_4}-2^{m_1}-2^{m_2})\Big|\le\Big|\frac{q}{a}\Big|\frac{4\epsilon X}{2M}\le\frac{2\epsilon X}{|\lambda|}<X$$

  for ϵ sufficiently small. Therefore n satisfies the hypothesis of Lemma 2.6. Applying the lemma and recalling that $\log p_j\le(1/2)\log X$, $\mathfrak{S}_-(n)\le 2\mathfrak{S}''(n)$ and that $r((q/a)(2^{m_3}+2^{m_4}-2^{m_1}-2^{m_2}))\neq 0$ if and only if $a\mid(2^{m_3}+2^{m_4}-2^{m_1}-2^{m_2})$, we have

$$I'\le\frac{\eta}{16}\log^4 X\sum_{1\le m_1,m_2,m_3,m_4\le L}r\Big(\frac{q}{a}(2^{m_3}+2^{m_4}-2^{m_1}-2^{m_2})\Big)$$
$$<(1+\epsilon_1)c_4\frac{\pi^2}{128}\eta X\sum_{1\le m_1,m_2,m_3,m_4\le L}\mathfrak{S}''\Big(\frac{q}{a}(2^{m_3}+2^{m_4}-2^{m_1}-2^{m_2})\Big). \tag{2.20}$$

[3]Because in this case we have that $n+p_3^2+4=p_1^2+p_2^2$ or $n+p_3^2+9=p_1^2+p_2^2$ and so, for every fixed p_3 the number of solutions is bounded by the number of ways we can write an integer as a sum of two squares. It is $\ll(n+p_3^2+9)^\epsilon\ll X^\epsilon$, see, *e.g.*, Theorem 338 of Hardy-Wright [HW10].

Using the multiplicativity of $\mathfrak{S}''(n)$ (defined in (2.5)), we get

$$\mathfrak{S}''\Big(\frac{q}{a}(2^{m_3}+2^{m_4}-2^{m_1}-2^{m_2})\Big) \le \mathfrak{S}''(q)\mathfrak{S}''\Big(\frac{2^{m_3}+2^{m_4}-2^{m_1}-2^{m_2}}{a}\Big) \le \mathfrak{S}''(q)\mathfrak{S}''(2^{m_3}+2^{m_4}-2^{m_1}-2^{m_2}),$$

and so, by (2.20), for every sufficiently large X we can write

$$I' \le (1+\epsilon_1)c_4\frac{\pi^2}{128}\mathfrak{S}''(q)\eta X \sum_{1\le m_1,m_2,m_3,m_4\le L} \mathfrak{S}''(2^{m_3}+2^{m_4}-2^{m_1}-2^{m_2}).$$

Arguing now as in the estimation of Σ at pages 63-64 of Liu-Lü [LL04], we get

$$\sum_{1\le m_1,m_2,m_3,m_4\le L} \mathfrak{S}''(2^{m_3}+2^{m_4}-2^{m_1}-2^{m_2}) \le \frac{4}{3}c_5(1+\epsilon_1)L^4,$$

thus obtaining, for a sufficiently small ϵ_1, that

$$I' \le (1+\epsilon_1)c_4c_5\frac{\pi^2}{96}\mathfrak{S}''(q)\eta X L^4.\blacksquare \tag{2.21}$$

Hence, by (2.17), (2.19) and (2.21), we finally get

$$I < (1+\epsilon_1)\eta X L^4\Big(\log^2 2 + c_4c_5\frac{\pi^2}{96}\mathfrak{S}''(q)\Big),$$

this way proving Lemma 2.8. □

The next lemma is the analogous of Lemma 4 in [Par03], but for the exponential sum over prime squares and with a better choice of parameters with respect to Lemma 4.3 of [LW05]. We can use such parameters, since we have narrowed the size of minor arc.

Lemma 2.9. *Suppose that λ_2/λ_3 is irrational. Let $X = q^2$, where q is the denominator of a convergent to the continued fraction for λ_2/λ_3. Then, for arbitrarily small ϵ_2, we have*

$$\sup_{\alpha\in\mathfrak{m}} |S_2(\lambda_2\alpha)S_2(\lambda_3\alpha)| \ll_{\epsilon_2} X^{15/16+\epsilon_2}(\log X)^{1/2}.$$

Proof. Let $\alpha \in \mathfrak{m}$ and $Q = X^{1/4}/(\log X)^2 \le P$. Clearly $\lambda_2\alpha, \lambda_3\alpha \in \mathbb{R}$, so by Dirichlet's theorem in Result (R.6), for every $i = 2, 3$ there exist integers a_i, q_i such that $1 \le q_i \le X/Q$, $(a_i, q_i) = 1$ and $|\lambda_i\alpha q_i - a_i| \le Q/X$. We remark that $a_2a_3 \ne 0$ otherwise $|\lambda_i\alpha q_i| \le Q/X$, that is $|\alpha| \le Q/(q_i\lambda_i X) \le P/X$ and so we would have $\alpha \in \mathfrak{M}$.

Now suppose that $q_1, q_2 \le Q$. In this case we get

$$a_3q_2\frac{\lambda_2}{\lambda_3} - a_2q_3 = (\lambda_2\alpha q_2 - a_2)\frac{a_3}{\lambda_3\alpha} - (\lambda_3\alpha q_3 - a_3)\frac{a_2}{\lambda_3\alpha}$$

and hence

$$\left|a_3q_2\frac{\lambda_2}{\lambda_3} - a_2q_3\right| \le 2\left(1+\left|\frac{\lambda_2}{\lambda_3}\right|\right)\frac{Q^2}{X} < \frac{1}{2q},$$

for a sufficiently large X. Then, from the law of best approximation (see Result (R.7)) and the definition of $\mathfrak{m}$, we obtain the contradiction

$$X^{1/2} = q \leq |a_3 q_2| \ll q_2 q_3 \log^2 X \leq Q^2 \log^2 X \leq X^{1/2} \log^{-2} X.$$

Hence either $q_2 > Q$ or $q_3 > Q$. Assume, without loss of generality, that $q_2 > Q$. Using Ghosh's estimate in Result (R.10), we have

$$\sup_{\alpha \in \mathfrak{m}} |S_2(\lambda_2 \alpha) S_2(\lambda_3 \alpha)| \ll_{\epsilon_2} X^{1+\epsilon_2} \sup_{Q < q_2 \leq X/Q} \left(\frac{1}{q_2} + \frac{1}{X^{1/4}} + \frac{q_2}{X} \right)^{1/4} \ll_{\epsilon_2} X^{15/16+\epsilon_2} (\log X)^{1/2},$$

and Lemma 2.9 follows. □

To estimate the contribution of $G(\alpha)$ on the minor arc, we use Pintz-Ruzsa's method as developed in [PR03], §3-7.

Lemma 2.10 (Pintz-Ruzsa)**.** *Let $0 < c < 1$. Then there exists $\nu = \nu(c) \in (0,1)$ such that*

$$|E(\nu)| := |\{\alpha \in (0,1) : |G(\alpha)| > \nu L\}| \ll_{M,\epsilon} X^{-c}.$$

Proof. [PR03], §7. □

To obtain explicit values for ν, we used the version of Pintz-Ruzsa's algorithm already implemented to get the results in Languasco-Zaccagnini [LZ10]. We used the PARI/GP [The10] scripting language and the gp2c compiling tool to be able to compute fifty decimal digits (but we write here just ten) of the constant involved in the previous lemma. Running the program in our case, Lemma 2.10 gives the following result:

$$|G(\alpha)| \leq 0.9505087500 \cdot L\,, \tag{2.22}$$

for $\alpha \in (0,1) \setminus E$ and with $|E| \ll_{M,\epsilon} X^{-7/8-10^{-20}}$. The computing time to get (2.22) on a Apple MacBook Pro was equal to 29 minutes and 37 seconds (but to get 30 correct digits just 4 minutes and 6 seconds suffice). You can download the PARI/GP source code of our program together with the cited numerical values at the following link: `www.math.unipd.it/~languasc/PintzRuzsaMethod.html`.

Now we state some lemmas we need to work on the major arc. Let

$$J(X,h) = \int_{\epsilon X}^{X} (\vartheta(x+h) - \vartheta(x) - h)^2 \mathrm{d}x \tag{2.23}$$

and

$$J^*(X,h) = \int_{\epsilon X}^{X} \left(\vartheta(\sqrt{x+h}) - \vartheta(\sqrt{x}) - (\sqrt{x+h} - \sqrt{x}) \right)^2 \mathrm{d}x \tag{2.24}$$

be two different versions of the Selberg integral, with $\vartheta(x)$ the Chebyshev function. We also define the following exponential sums

$$U_1(\alpha) = \sum_{\epsilon X \leq n \leq X} e(\alpha n) \qquad \text{auxiliary exponential sum over integers,}$$

$$U_2(\alpha) = \sum_{\epsilon X \le n^2 \le X} e(\alpha n^2) \qquad \text{auxiliary exponential sum over squares.}$$

The famous Gallagher's lemma on the truncated L^2-norm of exponential sums (see Result (R.11)) applied to $(S_1 - U_1)(\alpha)$ gives the following well-known result.

Lemma 2.11. *For $1/X \le Y \le 1/2$, we have*

$$\int_{-Y}^{Y} |S_1(\alpha) - U_1(\alpha)|^2 \mathrm{d}\alpha \ll_\epsilon \frac{\log^2 X}{Y} + Y^2 X + Y^2 J\Big(X, \frac{1}{2Y}\Big),$$

where $J(X,h)$ is defined in (2.23).

Proof. See Brüdern-Cook-Perelli [BCP97], Lemma 1. □

Following Brüdern-Cook-Perelli, we prove the corresponding result for exponential sums over squares.

Lemma 2.12. *For $1/X \le Y \le 1/2$, we have*

$$\int_{-Y}^{Y} |S_2(\alpha) - U_2(\alpha)|^2 \mathrm{d}\alpha \ll_\epsilon \frac{\log^2 X}{YX} + Y^2 X + Y^2 J^*\Big(X, \frac{1}{2Y}\Big),$$

where $J^*(X,h)$ is defined in (2.24).

Proof. Letting now the needed integral to be

$$I = \int_{-Y}^{Y} |S_2(\alpha) - U_2(\alpha)|^2 \mathrm{d}\alpha.$$

By definitions of $S_2(\alpha)$ and $U_2(\alpha)$, we can write

$$I = \int_{-Y}^{Y} \Big| \sum_{\epsilon X \le p^2 \le X} \log p \; e(p^2\alpha) - \sum_{\epsilon X \le n^2 \le X} e(\alpha n^2) \Big|^2 \mathrm{d}\alpha = \int_{-Y}^{Y} \Big| \sum_{\epsilon X \le n^2 \le X} (\kappa(n) - 1) e(n^2\alpha) \Big|^2 \mathrm{d}\alpha,$$

where $\kappa(n) = \log p$, if $n = p$ prime, and $\kappa(n) = 0$, otherwise. Since the sum inside the absolute value in the right-hand side above is finite, we can apply Gallagher's lemma in Result (R.11), thus obtaining

$$I \ll Y^2 \int_{-\infty}^{\infty} \Big(\sum_{\substack{|y - n^2| \le 1/(4Y) \\ \epsilon X \le n^2 \le X}} (\kappa(n) - 1) \Big)^2 \mathrm{d}y = Y^2 \int_{-\infty}^{\infty} \Big(\sum_{\substack{x \le n^2 \le x + H \\ \epsilon X \le n^2 \le X}} (\kappa(n) - 1) \Big)^2 \mathrm{d}x,$$

where the last step follows from taking $x = y - 1/(4Y)$ and defining, for simplicity, $H = 1/(2Y)$.

It is easy to see that the inner sum in the last integral above is empty whenever $X < x$ or $\epsilon X > x + H$, thus we can restrict the integration range to $\mathcal{S} = [\epsilon X - H, X]$. Then we split $\mathcal{S}$ as $\mathcal{S} = \mathcal{S}_1 \sqcup \mathcal{S}_2 \sqcup \mathcal{S}_3$, where the symbol $\sqcup$ represents a disjoint union and

$$\mathcal{S}_1 = [\epsilon X - H, e_1], \quad \mathcal{S}_2 = [e_1, e_2], \quad \mathcal{S}_3 = [e_2, X],$$

with e_1, e_2 such that for any $x \in \mathcal{S}_2$ the condition $x \le n^2 \le x+H$ implies $\epsilon X \le n^2 \le X$. So, for every $x \in \mathcal{S}_2$ we must have $x \ge \epsilon X$ and $x+H \le X$, and hence $e_1 = \epsilon X$, $e_2 = X - H$. According to these definitions we can write

$$I \ll Y^2(I_1+I_2+I_3) = Y^2\left(\int_{\mathcal{S}_1}+\int_{\mathcal{S}_2}+\int_{\mathcal{S}_3}\right)\Big(\sum_{\substack{x\le n^2\le x+H\\ \epsilon X\le n^2\le X}}(\kappa(n)-1)\Big)^2\mathrm{d}x, \tag{2.25}$$

say. We now estimate I_i, for $i = 1, 2$ and 3, one by one.

Estimation of I_1. For every $x \in \mathcal{S}_1 = [\epsilon X - H, \epsilon X]$, we have

i) $n^2 \le x+H \Rightarrow n^2 \le \epsilon X + H = X\,(\epsilon + H/X) \le X$. The last inequality follows from

$$\epsilon \le 1 - \frac{H}{X}, \tag{2.26}$$

which is true since ϵ can be chosen small enough and $H/X \le 1/2$.

ii) $n^2 \ge \epsilon X \Rightarrow n^2 \ge x$.

So we can unify the conditions in the inner sum of I_1, obtaining

$$I_1 = \int_{\mathcal{S}_1}\Big(\sum_{\epsilon X\le n^2\le x+H}(\kappa(n)-1)\Big)^2\mathrm{d}x = \int_{\mathcal{S}_1}\Big(\sum_{\epsilon X\le n^2\le x+H}\kappa(n) - \sum_{\epsilon X\le n^2\le x+H}1\Big)^2\mathrm{d}x.$$

Recalling the definitions of $\kappa(x)$ and $\vartheta(x)$, we get

$$\sum_{\sqrt{\epsilon X}\le n\le\sqrt{x+H}}\kappa(n) = \vartheta(\sqrt{x+H}) - \vartheta(\sqrt{\epsilon X}).$$

Moreover, we trivially have

$$\sum_{\sqrt{\epsilon X}\le n\le\sqrt{x+H}}1 = \sqrt{x+H} - \sqrt{\epsilon X} + \mathcal{O}(1),$$

where $\mathcal{O}(1)$ is the contribution of the fractional part. Thus, using $|a+b|^2 \le 2|a|^2 + 2|b|^2$, we get

$$\begin{aligned} I_1 &= \int_{\mathcal{S}_1}\Big(\vartheta(\sqrt{x+H}) - \vartheta(\sqrt{\epsilon X}) - (\sqrt{x+H}-\sqrt{\epsilon X}) + \mathcal{O}(1)\Big)^2\mathrm{d}x \\ &\ll \int_{\epsilon X - H}^{\epsilon X}\Big(\vartheta(\sqrt{x+H}) - \vartheta(\sqrt{\epsilon X}) - (\sqrt{x+H}-\sqrt{\epsilon X})\Big)^2\mathrm{d}x + H. \end{aligned} \tag{2.27}$$

To estimate the previous integral, we use the trivial relation

$$\vartheta(y+h) - \vartheta(y) = \sum_{y\le p\le y+h}\log p \le \log(y+h)\,(\pi(y+h)-\pi(y)) \le \log(y+h)h, \tag{2.28}$$

for $h \geq 1$, which directly follows from the definition of $\pi(x)$.

In our situation $y = \sqrt{\epsilon X} \geq 0$ and $h = \sqrt{x+H} - \sqrt{\epsilon X}$. Thus, for $x \in [\epsilon X - H, \epsilon X + 2\sqrt{\epsilon X} - H + 1)$, we have $h < 1$, which implies that the values of the ϑ-function involved in (2.27) are equal and so the integrand there is reduced to the difference of the radicals (which is $h \ll 1$). In the remaining range we have $h \geq 1$, so we can use (2.28) thus obtaining

$$\begin{aligned}\Big(\vartheta(\sqrt{x+H}) - \vartheta(\sqrt{\epsilon X}) - (\sqrt{x+H} - \sqrt{\epsilon X})\Big)^2 &\ll \Big(\sqrt{x+H} - \sqrt{\epsilon X}\Big)^2 \Big(\log\sqrt{x+H} + 1\Big)^2 \\ &\ll_\epsilon \Big(\sqrt{x+H} - \sqrt{\epsilon X}\Big)^2 \Big(\log\sqrt{X}\Big)^2 \ll (\log X)^2 \Big(\sqrt{x+H} - \sqrt{\epsilon X}\Big)^2, \qquad (2.29)\end{aligned}$$

where we used $x \leq \epsilon X$ and $H \leq X/2$ by the definitions of H and Y.

Therefore, by (2.27) and (2.29), we have

$$\begin{aligned} I_1 &\ll (\log X)^2 \int_{\epsilon X - H}^{\epsilon X} \Big(\sqrt{x+H} - \sqrt{\epsilon X}\Big)^2 \,\mathrm{d}x + H \\ &\leq H(\log X)^2 \max_{\epsilon X - H \leq x \leq \epsilon X} \Big(\sqrt{x+H} - \sqrt{\epsilon X}\Big)^2 + H \\ &= H(\log X)^2 \Big(\sqrt{\epsilon X + H} - \sqrt{\epsilon X}\Big)^2 + H \\ &= \epsilon H X (\log X)^2 \Big(\sqrt{1 + \frac{H}{\epsilon X}} - 1\Big)^2 + H. \end{aligned}$$

By the trivial relation $1+y \leq (1+y/2)^2$, we get $\sqrt{|1+y|} - 1 \leq y/2$ and so $\sqrt{1+\frac{H}{\epsilon X}} - 1 \leq \frac{H}{2\epsilon X}$. Thus

$$I_1 \ll \epsilon H X \left(\frac{H}{\epsilon X}\right)^2 (\log X)^2 + H \ll_\epsilon \frac{H^3 (\log X)^2}{X} + H. \blacksquare \qquad (2.30)$$

Estimation of I_3. The estimation of I_3 is similar to the one of I_1. For every $x \in \mathcal{S}_3 = [X - H, X]$, we have

i) $n^2 \geq x \Rightarrow n^2 \geq X - H \geq \epsilon X$, where the last inequality follows from (2.26).

ii) $n^2 \leq X \Rightarrow n^2 \leq X - H + H \leq x + H$.

We can therefore unify the conditions in the inner sum of I_3, obtaining

$$\begin{aligned} I_3 &= \int_{\mathcal{S}_3} \Big(\sum_{x \leq n^2 \leq X} (\kappa(n) - 1)\Big)^2 \mathrm{d}x = \int_{X-H}^{X} \Big(\sum_{\sqrt{x} \leq p \leq \sqrt{X}} \log p - \sum_{\sqrt{x} \leq n \leq \sqrt{X}} 1\Big)^2 \mathrm{d}x \\ &\ll \int_{X-H}^{X} \Big(\vartheta(\sqrt{X}) - \vartheta\left(\sqrt{x}\right) - (\sqrt{X} - \sqrt{x})\Big)^2 \mathrm{d}x + H. \end{aligned}$$

We now work as for I_1, with $y = \sqrt{x}$ and $h = \sqrt{X} - \sqrt{x}$. For $x \in (X + 1 - 2\sqrt{X}, X]$, we have $h < 1$, so the values of the ϑ-function are the same and the integrand is equal to the

difference of the radicals (which is $h \ll 1$). In the remaining range we have $h \geq 1$ and so we can use (2.28), obtaining

$$I_3 \ll \int_{X-H}^{X} (\sqrt{X}-\sqrt{x})^2(\log\sqrt{X}+1)^2\mathrm{d}x + H \ll H(\log X)^2\Big(\max_{X-H\leq x\leq X}\big(\sqrt{X}-\sqrt{x}\big)\Big)^2 + H$$
$$= HX(\log X)^2\left(1-\sqrt{1-\frac{H}{X}}\right)^2 + H.$$

We now use the trivial relation $1-y \leq (1-y/2)^2$, that is $\sqrt{|1-y|}-1+y \leq y/2$. If $y \leq 1$, it implies $|1-\sqrt{1-y}| \leq y/(2\sqrt{1-y})$ and so, since $H \leq X/2$, we have

$$\Big|1-\sqrt{1-\frac{H}{X}}\Big| \leq \frac{H}{2X\sqrt{1-\frac{H}{X}}} \leq \frac{H}{2\sqrt{\epsilon}X},$$

where the last inequality is due to (2.26). Hence

$$I_3 \ll HX(\log X)^2\Big(\frac{H}{\sqrt{\epsilon}X}\Big)^2 + H \ll_\epsilon \frac{H^3(\log X)^2}{X} + H.\blacksquare \tag{2.31}$$

Estimation of I_2. First of all, we recall that for every $x \in \mathcal{S}_2 = [\epsilon X, X-H]$, the condition $x \leq n^2 \leq x+H$ implies $\epsilon X \leq n^2 \leq X$. Therefore

$$I_2 = \int_{\mathcal{S}_2}\Big(\sum_{x\leq n^2\leq x+H}(k(n)-1)\Big)^2\mathrm{d}x = \int_{\epsilon X}^{X-H}\Big(\sum_{\sqrt{x}\leq p\leq\sqrt{x+H}}\log p - \sum_{\sqrt{x}\leq n\leq\sqrt{x+H}}1\Big)^2\mathrm{d}x$$
$$= \int_{\epsilon X}^{X-H}\Big(\vartheta(\sqrt{x+H})-\vartheta\left(\sqrt{x}\right)-(\sqrt{x+H}-\sqrt{x})+\mathcal{O}(1)\Big)^2\mathrm{d}x$$
$$\ll \int_{\epsilon X}^{X}\Big(\vartheta(\sqrt{x+H})-\vartheta(\sqrt{x})-(\sqrt{x+H}-\sqrt{x})\Big)^2\mathrm{d}x + X$$
$$= J^*(X,H)+X, \tag{2.32}$$

where we used the definition of $J^*(X,H)$ in (2.24) and the inequality in (2.26) which implies $X-H-\epsilon X = X(1-\epsilon-H/X) \leq X.\blacksquare$

Therefore, by (2.25), (2.30)-(2.32) and recalling $H = 1/(2Y)$ and $Y \geq 1/X$, we have

$$\mathcal{I} \ll_\epsilon Y^2\left(\frac{(\log X)^2}{XY^3}+X+J^*\left(X,\frac{1}{2Y}\right)\right) = \frac{(\log X)^2}{XY}+XY^2+Y^2J^*\left(X,\frac{1}{2Y}\right),$$

and this proves Lemma 2.12. □

To estimate the "square-root" form $J^*(X,h)$ of the Selberg integral, we use the next result, which is analogous to Saffari-Vaughan's estimation of the Selberg integral given in Result (R.12).

Lemma 2.13. *Let ϵ_3 be an arbitrarily small positive constant. There exists a positive constant $c_6 = c_6(\epsilon_3)$ such that*

$$J^*(X,h) \ll_\epsilon h^2 \exp\Big(-c_6\Big(\frac{\log X}{\log\log X}\Big)^{1/3}\Big),$$

uniformly for $X^{7/12+\epsilon_3} \le h \le X$, where $J^*(X,h)$ is defined in (2.24).

Proof. We reduce our problem to estimate the integral

$$J^*_\psi(X,h) := \int_{\epsilon X}^{X} \Big(\psi(\sqrt{x+h}) - \psi(\sqrt{x}) - (\sqrt{x+h}-\sqrt{x})\Big)^2 \,\mathrm{d}x, \tag{2.33}$$

since, using $|a+b|^2 \le 2|a|^2 + 2|b|^2$, it is easy to see that

$$J^*(X,h) \ll J^*_\psi(X,h) + \int_{\epsilon X}^{X} \Big(\psi(\sqrt{x+h}) - \psi(\sqrt{x}) - \vartheta(\sqrt{x+h}) + \vartheta(\sqrt{x})\Big)^2 \,\mathrm{d}x.$$

The difference between $\vartheta(x)$ and $\psi(x)$ in $[\sqrt{x}, \sqrt{x+h}]$ is

$$\sum_{\substack{\sqrt{x}\le p^a \le \sqrt{x+h}\\ a\ge 2}} \log p \ll \log X \sum_{\substack{\sqrt{x}\le p^a \le \sqrt{x+h}\\ a\ge 2}} 1 \ll (\log X)^2 \sum_{\sqrt{x}\le p^2 \le \sqrt{x+h}} 1,$$

since for $p^a \le \sqrt{x+h}$ with $a \ge 2$, we have $\log p \le (1/4)\log(x+h)$ and $x, h \le X$ by assumption; moreover there are at most $\ll \log X$ prime powers in $[\sqrt{x}, \sqrt{x+h}]$. By definition of the prime-counting function $\pi(x)$, we have

$$\sum_{\sqrt{x}\le p^2 \le \sqrt{x+h}} 1 = \pi((x+h)^{1/4}) - \pi(x^{1/4}) \le (x+h)^{1/4} - x^{1/4} = x^{1/4}\Big(\Big(1+\frac{h}{x}\Big)^{1/4} - 1\Big).$$

Now, using the trivial relation $1 + y \le (1+y/2)^2$ twice, we obtain

$$\Big(1+\frac{h}{x}\Big)^{1/4} - 1 \le \Big(1+\frac{h}{2x}\Big)^{1/2} - 1 \le \Big(1+\frac{h}{4x}\Big) - 1,$$

and hence

$$\sum_{\sqrt{x}\le p^2 \le \sqrt{x+h}} 1 \le x^{1/4}\frac{h}{4x} \ll_\epsilon \frac{h}{X^{3/4}},$$

since $\epsilon X \le x \le X$. Therefore

$$J^*(X,h) \ll_\epsilon J^*_\psi(X,h) + \int_{\epsilon X}^{X} \frac{h^2}{X^{3/2}}(\log X)^4 \,\mathrm{d}x \ll_\epsilon J^*_\psi(X,h) + h^2\frac{(\log X)^4}{X^{1/2}}. \tag{2.34}$$

To estimate the right hand side of (2.34), we use the following results we will prove later.

Claim 2.A. *Let ϵ_3 be an arbitrarily small positive constant. There exists a positive constant $c_6 = c_6(\epsilon_3)$ such that*

$$\widetilde{J}^*_\psi(X,\delta) := \int_{\epsilon X}^{X} \Big(\psi(\sqrt{x+\delta x}) - \psi(\sqrt{x}) - (\sqrt{x+\delta x} - \sqrt{x})\Big)^2 \,\mathrm{d}x \ll_\epsilon \delta^2 X^2 \exp\Big(-c_6\Big(\frac{\log X}{\log\log X}\Big)^{1/3}\Big),$$

uniformly for $X^{-5/12+\epsilon_3} \le \delta \le 1$[4].

As we will prove later, Claim 2.A implies the following.

Claim 2.B. *Let ϵ_3 be an arbitrarily small positive constant. There exists a positive constant $c_6 = c_6(\epsilon_3)$ such that*

$$J^*_\psi(X,h) \ll_\epsilon h^2 \exp\Big(-c_6\Big(\frac{\log X}{\log\log X}\Big)^{1/3}\Big),$$

uniformly for $X^{7/12+\epsilon_3} \le h \le X$, where $J^*_\psi(X,h)$ is defined in (2.33).

Therefore, by (2.34) and Claim 2.B, we obtain

$$J^*(X,h) \ll_\epsilon h^2 \exp\Big(-c_6\Big(\frac{\log X}{\log\log X}\Big)^{1/3}\Big)$$

thus proving Lemma 2.13. □

We now prove the two claims used above.

Proof of Claim 2.A. We follow the line of Lemma 5 in Saffari-Vaughan [SV77]. To estimate $\widetilde{J}^*_\psi(X,\delta)$, we use the truncated explicit formula for $\psi(x)$ in Result (R.13), that is:

$$\psi(x) = x - \sum_{|\gamma|\le T} \frac{x^\rho}{\rho} + \mathcal{O}\Big(\frac{x}{T}\log^2(xT) + \log x\Big)$$

uniformly in $T \ge 2$ and for $\rho = \beta + i\gamma$ along the zeros of $\zeta(s)$. So

$$\widetilde{J}^*_\psi(X,\delta) = \int_{\epsilon X}^{X} \left| - \sum_{|\gamma|\le T} x^{\rho/2} \frac{((1+\delta)^{\rho/2}-1)}{\rho} + \mathcal{O}\left(\frac{\sqrt{X}}{T}\log^2(XT) + \log X\right)\right|^2 \mathrm{d}x,$$

where we used that $x \le X$, $\sqrt{x+\delta x} \ll \sqrt{X}$ and $\log(\sqrt{X}T) = (\log X)/2 + \log T \ll \log(XT)$. By the symmetry of the non-trivial zeros of $\zeta(s)$ with respect to the line $\Re(s) = 1/2$ and using $|a+b|^2 \le 2|a|^2 + 2|b|^2$, we get

$$\widetilde{J}^*_\psi(X,\delta) \ll_\epsilon \int_{\epsilon X}^{X} \Big| \sum_{\substack{|\gamma|\le T \\ \beta \ge 1/2}} x^{\rho/2} \frac{((1+\delta)^{\rho/2}-1)}{\rho}\Big|^2 \mathrm{d}x + \frac{X^2}{T^2}\log^4(XT) + X\log^2 X. \tag{2.35}$$

[4]The lower bound for δ is due to Ingham-Huxley's zero-density estimate given in Result (R.15).

As in Ivić [Ivi85], page 316, we define

$$c(\delta,\rho)=\frac{(1+\delta)^{\rho}-1}{\rho},$$

and so we can write

$$\left|c\left(\delta,\frac{\rho}{2}\right)\right|\ll\min\left(\frac{1}{|\gamma|};\delta\right), \tag{2.36}$$

since

$$\left|c\left(\delta,\frac{\rho}{2}\right)\right|\leq\frac{|1+\delta|^{\rho/2}+1}{|\rho|/2}\leq\frac{2(2^{\beta/2}+1)}{\sqrt{\beta^2+\gamma^2}}\ll\frac{1}{|\gamma|},$$

where we used the hypothesis $\delta\leq 1$ and the fact that $0<\beta<1$ for the non-trivial zeros of $\zeta(s)$. Moreover also

$$\left|c\left(\delta,\frac{\rho}{2}\right)\right|=\left|\int_1^{1+\delta}t^{\rho/2-1}\mathrm{d}t\right|\ll\int_1^{1+\delta}t^{\beta/2-1}\mathrm{d}t\leq\delta,$$

where we used $\beta/2-1<0$ and $t\geq 1$. Assuming $T\geq 1/\delta$, we can split the summation in (2.35) into two cases, defined accordingly to (2.36). That is

$$\widetilde{J}^{*}_{\psi}(X,\delta)\ll_{\epsilon} I_{[0,1/\delta)}+I_{[1/\delta,T]}+\frac{X^2}{T^2}\log^4(XT)+X(\log X)^2, \tag{2.37}$$

with

$$\begin{aligned}
I_{\mathcal{S}}&=\int_{\epsilon X}^{X}\Big|\sum_{\substack{|\gamma|\in\mathcal{S}\\ \beta\geq 1/2}}x^{\rho/2}c\left(\delta,\frac{\rho}{2}\right)\Big|^2\mathrm{d}x=\sum_{\substack{|\gamma_1|\in\mathcal{S}\\ \beta_1\geq 1/2}}\sum_{\substack{|\gamma_2|\in\mathcal{S}\\ \beta_2\geq 1/2}}c\Big(\delta,\frac{\rho_1}{2}\Big)c\Big(\delta,\frac{\overline{\rho}_2}{2}\Big)\int_{\epsilon X}^{X}x^{(\rho_1+\overline{\rho}_2)/2}\mathrm{d}x\\
&=\sum_{\substack{|\gamma_1|\in\mathcal{S}\\ \beta_1\geq 1/2}}\sum_{\substack{|\gamma_2|\in\mathcal{S}\\ \beta_2\geq 1/2}}c\Big(\delta,\frac{\rho_1}{2}\Big)c\Big(\delta,\frac{\overline{\rho}_2}{2}\Big)\frac{2X^{(\rho_1+\overline{\rho}_2)/2+1}(1-\epsilon^{(\rho_1+\overline{\rho}_2)/2+1})}{\rho_1+\overline{\rho}_2+2}\\
&\ll\sum_{\substack{|\gamma_1|\in\mathcal{S}\\ \beta_1\geq 1/2}}\sum_{\substack{|\gamma_2|\in\mathcal{S}\\ \beta_2\geq 1/2}}\Big|c\Big(\delta,\frac{\rho_1}{2}\Big)\Big|\Big|c\Big(\delta,\frac{\overline{\rho}_2}{2}\Big)\Big|\frac{X^{(\beta_1+\beta_2)/2+1}}{1+|\gamma_1-\gamma_2|}\\
&\ll\sum_{\substack{|\gamma_1|\in\mathcal{S}\\ \beta_1\geq 1/2}}\sum_{\substack{|\gamma_2|\in\mathcal{S}\\ 1/2\leq\beta_2\leq\beta_1}}\Big|c\Big(\delta,\frac{\rho_1}{2}\Big)\Big|\Big|c\Big(\delta,\frac{\overline{\rho}_2}{2}\Big)\Big|\frac{X^{\beta_1+1}}{1+|\gamma_1-\gamma_2|}.
\end{aligned} \tag{2.38}$$

In the previous chain of inequalities we used $|1-\epsilon^{(\rho_1+\overline{\rho}_2)/2+1}|\ll 1$ and $|\rho_1+\overline{\rho}_2+2|=\sqrt{(\beta_1+\beta_2+2)^2+(\gamma_1-\gamma_2)^2}\geq\sqrt{1+(\gamma_1-\gamma_2)^2}\gg 1+|\gamma_1-\gamma_2|$. We also ordered the zeros ρ with respect to their real part. Now we deal with $I_{[0,1/\delta)}$ and $I_{[1/\delta,T]}$ separately.

Estimation of $I_{[0,1/\delta)}$. For $|\gamma_1|, |\gamma_2| \in [0, 1/\delta)$, we have $|c(\delta, \rho_1/2)|, |c(\delta, \overline{\rho}_2/2)| \ll \delta$ by (2.36). From (2.38) we can write

$$I_{[0,1/\delta)} \ll \delta^2 X \sum_{\substack{|\gamma_1|<\delta^{-1}\\ \beta_1\geq 1/2}} X^{\beta_1} \sum_{\substack{|\gamma_2|<\delta^{-1}\\ 1/2\leq\beta_2\leq\beta_1}} \frac{1}{1+|\gamma_1-\gamma_2|} \ll \delta^2 X(\log X)^2 \sum_{\substack{|\gamma_1|<\delta^{-1}\\ \beta_1\geq 1/2}} X^{\beta_1}, \tag{2.39}$$

where the last inequality follows from

$$\begin{aligned}
\sum_{\substack{|\gamma_2|<\delta^{-1}\\ 1/2\leq\beta_2\leq\beta_1}} \frac{1}{1+|\gamma_1-\gamma_2|} &\ll \sum_{n=0}^{2/\delta} \sum_{\substack{|\gamma_2|<\delta^{-1}\\ 1/2\leq\beta_2\leq\beta_1\\ n\leq|\gamma_1-\gamma_2|\leq n+1}} \frac{1}{1+n} \ll \sum_{n=0}^{2/\delta} \frac{1}{1+n} \sum_{\substack{|\gamma_2|<\delta^{-1}\\ 1/2\leq\beta_2\leq\beta_1\\ n\leq|\gamma_1-\gamma_2|\leq n+1\\ \gamma_2\geq\gamma_1}} 1 \\
&= \sum_{n=0}^{2/\delta} \frac{1}{1+n} \sum_{\substack{|\gamma_2|<\delta^{-1}\\ 1/2\leq\beta_2\leq\beta_1\\ \gamma_1+n\leq\gamma_2\leq\gamma_1+n+1}} 1 \ll \sum_{n=0}^{2/\delta} \frac{1}{1+n} \log(\gamma_1+n) \\
&\ll \log\left(\gamma_1+\frac{2}{\delta}\right) \sum_{n=0}^{2/\delta} \frac{1}{1+n} \ll \log\left(\frac{3}{\delta}\right)\log\left(\frac{2}{\delta}\right) \\
&\ll (\log X)^2.
\end{aligned} \tag{2.40}$$

Here we used the relations $|\gamma_1-\gamma_2| \leq |\gamma_1|+|\gamma_2| < 2/\delta$, $\sum_{n=1}^{k} 1/n \ll \log k$ and $\log(\delta^{-1}) \ll \log X$, together with the Riemann-von Mangoldt formula (see Result (R.14)) which implies

$$N(0, t+1) - N(0, t) \ll \log t.$$

Denoting by $S_{[0,1/\delta)}$ the sum in the right hand side of (2.39), we get

$$\begin{aligned}
S_{[0,1/\delta)} &= \sum_{\substack{|\gamma|<1/\delta\\ \beta\geq 1/2}} X^{\beta} = \sum_{\substack{|\gamma|<1/\delta\\ \beta\geq 1/2}} (X^{\beta} - X^{1/2}) + \sum_{\substack{|\gamma|<1/\delta\\ \beta\geq 1/2}} X^{1/2} \\
&= \log X \int_{1/2}^{1} X^u N\Big(u, \frac{1}{\delta}\Big)\mathrm{d}u + X^{1/2} N\Big(\frac{1}{2}, \frac{1}{\delta}\Big) \ll \log X \max_{1/2\leq u\leq 1}\Big(X^u N\Big(u, \frac{1}{\delta}\Big)\Big).
\end{aligned}$$

We now recall Ingham-Huxley's zero-density estimates (see Result (R.15)) given by

$$N(\sigma, t) \ll \begin{cases} t\log t & \text{for } 0 \leq \sigma \leq \frac{1}{2}, \\ t^{\frac{12}{5}(1-\sigma)}(\log t)^B & \text{for } \frac{1}{2} \leq \sigma \leq 1. \end{cases} \tag{2.41}$$

And we also recall that, by the Vinogradov-Korobov zero-free region (see Result (R.5)), there are no zeros $\rho = \beta + i\gamma$ of the Riemann ζ-function having

$$\beta \geq 1 - \frac{c_7}{(\log(|\gamma|+2))^{2/3}(\log\log(|\gamma|+2))^{1/3}},$$

where $c_7 > 0$ is an absolute constant. In the following c_7 will not necessarily be the same at each occurrence.

In our situation, $|\gamma| \le T$, so $N(u,t) = 0$ for every $t \le T$ and $u \ge 1-K$ with

$$K = \frac{c_7}{(\log T)^{2/3}(\log\log T)^{1/3}}.$$

From the previous remarks, we obtain

$$\begin{aligned} S_{[0,1/\delta)} &\ll \log X \max_{1/2\le u\le 1-K} \left((\delta^{-1})^{(12/5)(1-u)}(\log(\delta^{-1}))^B X^u\right) \\ &\ll (\log X)^{B+1}\delta^{-\frac{12}{5}} \max_{1/2\le u\le 1-K} \left((\delta^{12/5}X)^u\right), \end{aligned}$$

since $\log(\delta^{-1}) \ll \log X$. We now observe that

$$\frac{\partial}{\partial u}((\delta^{12/5}X)^u) = (\delta^{12/5}X)^u \log(\delta^{12/5}X) > 0, \tag{2.42}$$

because $\delta \ge X^{-\frac{5}{12}+\epsilon_3}$ and so[5] $\delta^{12/5}X \ge (X^{-\frac{5}{12}+\epsilon_3})^{\frac{12}{5}}X > 1$. Therefore, the maximum above is attained at $u = 1-K$ and so

$$S_{[0,1/\delta)} \ll (\log X)^{B+1}\delta^{-\frac{12}{5}}\delta^{\frac{12}{5}(1-K)}X^{1-K} = X(\log X)^{B+1}(\delta^{12/5}X)^{-K}.$$

Inserting the last estimate into (2.39), we can write

$$A_{[0,1/\delta)} \ll \delta^2 X^2 (\log X)^{B+3}(\delta^{12/5}X)^{-K}.\blacksquare \tag{2.43}$$

Estimation of $I_{[1/\delta,T]}$. For $|\gamma_1|, |\gamma_2| \in [1/\delta, T]$, we have $|c(\delta,\rho_1/2)| \ll |\gamma_1|^{-1}$ and $|c(\delta,\overline{\rho}_2/2)| \ll |\gamma_2|^{-1}$ by (2.36). Thus, from (2.38) we get

$$\begin{aligned} I_{[1/\delta,T]} &\ll X \sum_{\substack{1/\delta\le|\gamma_1|\le T\\ \beta_1\ge 1/2}} \frac{X^{\beta_1}}{|\gamma_1|} \sum_{\substack{1/\delta\le|\gamma_2|\le T\\ 1/2\le\beta_2\le\beta_1}} \frac{1}{|\gamma_2|(1+|\gamma_1-\gamma_2|)} \\ &\ll X \sum_{\substack{1/\delta\le|\gamma_1|\le T\\ \beta_1\ge 1/2}} \frac{X^{\beta_1}}{|\gamma_1|^2} \sum_{\substack{|\gamma_1|\le|\gamma_2|\le T\\ 1/2\le\beta_2\le\beta_1}} \frac{1}{1+|\gamma_1-\gamma_2|} \ll X(\log T)^2 \sum_{\substack{1/\delta\le|\gamma_1|\le T\\ \beta_1\ge 1/2}} \frac{X^{\beta_1}}{|\gamma_1|^2}, \end{aligned}$$

where the last step follows from (2.40) with T instead of $1/\delta$. By a simple trick, we can rewrite the previous inequality as

$$I_{[1/\delta,T]} \ll X(\log T)^2(S'_{[1/\delta,T]} + S''_{[1/\delta,T]}), \tag{2.44}$$

with

$$S'_{[1/\delta,T]} = \sum_{\substack{1/\delta\le|\gamma|\le T\\ \beta\ge 1/2}} X^\beta\Big(\frac{1}{|\gamma|^2} - \frac{1}{T^2}\Big) \qquad \text{and} \qquad S''_{[1/\delta,T]} = \frac{1}{T^2}\sum_{\substack{1/\delta\le|\gamma|\le T\\ \beta\ge 1/2}} X^\beta.$$

[5] This explains how the lower bound of δ is linked to Ingham-Huxley's estimate, as said before in the statement of Claim 2.A.

For $S''_{[1/\delta,T]}$ we can argue as we did for $S_{[0,1/\delta)}$, just keeping in mind that this time $1/\delta \le |\gamma| \le T$. Hence

$$S''_{[1/\delta,T]} \ll \frac{\log X}{T^2} \max_{1/2\le u\le 1-K}\Big(X^u\Big[N(u,T)-N\Big(u,\frac{1}{\delta}\Big)\Big]\Big).$$

As for $S'_{[1/\delta,T]}$, we immediately obtain

$$S'_{[1/\delta,T]} = \sum_{\substack{1/\delta\le|\gamma|\le T\\ \beta\ge 1/2}} X^\beta \int_{|\gamma|}^T \frac{2}{t^3}\mathrm{d}t = 2\int_{1/\delta}^T \Big(\sum_{\substack{1/\delta\le|\gamma|\le t\\ \beta\ge 1/2}} X^\beta\Big)\frac{\mathrm{d}t}{t^3}.$$

The summation inside the previous integral is exactly the one in $S''_{[1/\delta,T]}$, but with t instead of T. Using $t\le T$, we can write

$$S'_{[1/\delta,T]} \ll \int_{1/\delta}^T \log X \max_{1/2\le u\le 1-K}\Big(X^u\Big[N(u,t)-N\Big(u,\frac{1}{\delta}\Big)\Big]\Big)\frac{\mathrm{d}t}{t^3}.$$

Therefore

$$\begin{aligned}
S'_{[1/\delta,T]}+S''_{[1/\delta,T]} &\ll \log X\Bigg(\int_{1/\delta}^T \max_{1/2\le u\le 1-K}\Big(X^u\Big[N(u,t)-N\Big(u,\frac{1}{\delta}\Big)\Big]\Big)\frac{\mathrm{d}t}{t^3}\\
&\qquad + \frac{1}{T^2}\max_{1/2\le u\le 1-K}\Big(X^u\Big[N(u,T)-N\Big(u,\frac{1}{\delta}\Big)\Big]\Big)\Bigg)\\
&\ll \log X\Bigg(\max_{1/\delta\le t\le T}\Big(\frac{1}{t^2}\max_{1/2\le u\le 1-K}\Big(X^u\Big[N(u,t)-N\Big(u,\frac{1}{\delta}\Big)\Big]\Big)\Big)\int_{1/\delta}^T\frac{\mathrm{d}t}{t}\\
&\qquad + \frac{1}{T^2}\max_{1/2\le u\le 1-K}\Big(X^u\Big[N(u,T)-N\Big(u,\frac{1}{\delta}\Big)\Big]\Big)\Bigg)\\
&\ll \log X\log(T\delta)\max_{1/\delta\le t\le T}\Big(\frac{1}{t^2}\max_{1/2\le u\le 1-K}\Big(X^u\,t^{(12/5)(1-u)}(\log t)^B\Big)\Big),
\end{aligned}$$

by the zero-density estimate in (2.41). Inserting this in (2.44) and recalling $t\le T$, we get

$$I_{[1/\delta,T]} \ll X(\log T)^{B+2}\log X\log(T\delta)\max_{1/2\le u\le 1-K}\Big(X^u\max_{1/\delta\le t\le T}\big(t^{(12/5)(1-u)-2}\big)\Big).$$

To compute the inner maximum above, we just remark that $(12/5)(1-u)-2<0$ (since $u>1/2$). Hence the maximum is attained at $t=1/\delta$, so

$$\begin{aligned}
I_{[1/\delta,T]} &\ll X(\log T)^{B+2}\log X\log(T\delta)\max_{1/2\le u\le 1-K}\big(X^u(\delta^{-1})^{(12/5)(1-u)-2}\big)\\
&= \delta^{-\frac{2}{5}}X(\log T)^{B+2}\log X\log(T\delta)\max_{1/2\le u\le 1-K}\big((X\delta^{12/5})^u\big).
\end{aligned}$$

By (2.42), the maximum above is attained at $u=1-K$, thus

$$\begin{aligned}
I_{[1/\delta,T]} &\ll \delta^{-\frac{2}{5}}X(\log T)^{B+2}\log X\log(T\delta)(X\delta^{12/5})^{1-K}\\
&= \delta^2X^2(\log T)^{B+2}\log X\log(T\delta)(X\delta^{12/5})^{-K}.\blacksquare
\end{aligned}\tag{2.45}$$

Conclusion of the proof. Inserting (2.43) and (2.45) into (2.37), we get

$$\widetilde{J}^*_\psi(X,\delta) \ll_\epsilon \delta^2 X^2 (X\delta^{12/5})^{-K} \log X \Big((\log X)^{B+2} + (\log T)^{B+2} \log(T\delta) \Big) + \frac{X^2}{T^2} (\log(XT))^4 + X(\log X)^2. \tag{2.46}$$

We now have to choose the optimal T. Remarking that:

i) We want $X \le X^2 T^{-2}$, that is $T \le \sqrt{X}$. This choice of T also implies that $\log T \ll \log X$ and hence (2.46) becomes

$$\widetilde{J}^*_\psi(X,\delta) \ll_\epsilon \delta^2 X^2 (X\delta^{12/5})^{-K} (\log X)^{B+4} + \frac{X^2}{T^2} (\log X)^4. \tag{2.47}$$

ii) We want $X^2 T^{-2} (\log X)^4 \le \delta^2 X^2 (X\delta^{12/5})^{-K} (\log X)^{B+4}$. This implies the inequality $T \ge \delta^{-1} (X\delta^{12/5})^{K/2} (\log X)^{-B/2}$. Recalling $1/\delta \le X^{5/12-\epsilon_3}$, we can take

$$T \ge X^{5/12-\epsilon_3} (X\delta^{12/5})^{K/2} (\log X)^{-B/2}.$$

Since $X^{5/12-\epsilon_3} (X\delta^{12/5})^{K/2} (\log X)^{-B/2} < \sqrt{X}$, the second condition on T is compatible with the first one. After this choice of T, (2.47) becomes

$$\widetilde{J}^*_\psi(X,\delta) \ll_\epsilon \delta^2 X^2 (X\delta^{12/5})^{-K} (\log X)^{B+4}. \tag{2.48}$$

By $T \le \sqrt{X}$, we also have

$$K = \frac{c_7}{(\log T)^{2/3} (\log\log T)^{1/3}} \ge \frac{c_8}{(\log X)^{2/3} (\log\log X)^{1/3}} =: K',$$

for a suitable positive constant c_8. From this and (2.48) we immediately obtain

$$\begin{aligned} \widetilde{J}^*_\psi(X,\delta) &\ll_\epsilon \delta^2 X^2 (X\delta^{12/5})^{-K'} (\log X)^{B+4} \\ &= \delta^2 X^2 (\log X)^{B+4} \exp\Big(- \frac{c_8 (\log X + (12/5)\log\delta)}{(\log X)^{2/3} (\log\log X)^{1/3}} \Big) \\ &\ll \delta^2 X^2 \exp\Big(- c_9 \Big(\frac{\log X}{\log\log X} \Big)^{1/3} \Big), \end{aligned}$$

for a sufficiently large X and for a suitable positive constant $c_9 = c_9(\epsilon_3)$. So Claim 2.A follows. □

Proof of Claim 2.B. We follow the line of Lemma 6 in Saffari-Vaughan [SV77]. We recall that ϵ_3 is an arbitrarily small positive constant and that $X^{7/12+\epsilon_3} \le h \le X$. Let now $2h \le v \le 3h$. To estimate $J^*_\psi(X,h)$, defined in (2.33), the first step is observing that, by $|a+b|^2 \le 2|a|^2 + 2|b|^2$, we get

$$\big(\psi(\sqrt{x+h}) - \psi(\sqrt{x}) - (\sqrt{x+h} - \sqrt{x}) \big)^2$$

$$\ll \big(\psi(\sqrt{x+v})-\psi(\sqrt{x})-(\sqrt{x+v}-\sqrt{x})\big)^2 + \big(\psi(\sqrt{x+v})-\psi(\sqrt{x+h})-(\sqrt{x+v}-\sqrt{x+h})\big)^2.$$

Thus, replacing h by $\int_{2h}^{3h}\mathrm{d}v$, we have

$$hJ^*_\psi(X,h) \ll \int_{\epsilon X}^{X}\int_{2h}^{3h}\Big(\big(\psi(\sqrt{x+v})-\psi(\sqrt{x})-(\sqrt{x+v}-\sqrt{x})\big)^2 + \big(\psi(\sqrt{x+v})-\psi(\sqrt{x+h})-(\sqrt{x+v}-\sqrt{x+h})\big)^2\Big)\mathrm{d}v\ \mathrm{d}x. \tag{2.49}$$

Setting $z=v-h$, $y=x+h$ and changing variables in the second integration, the right hand side of (2.49) becomes

$$\ll \int_{\epsilon X}^{X}\Big(\int_{2h}^{3h}\big(\psi(\sqrt{x+v})-\psi(\sqrt{x})-(\sqrt{x+v}-\sqrt{x})\big)^2\mathrm{d}v\Big)\mathrm{d}x + \int_{\epsilon X+h}^{X+h}\Big(\int_{h}^{2h}\big(\psi(\sqrt{y+z})-\psi(\sqrt{y})-(\sqrt{y+z}-\sqrt{y})\big)^2\mathrm{d}z\Big)\mathrm{d}y.$$

Since both the integrand functions are non-negative, we can extend the integration ranges merging x with y and v with z. Hence

$$hJ^*_\psi(X,h) \ll \int_{\epsilon X}^{X+h}\Big(\int_{h}^{3h}\big(\psi(\sqrt{x+v})-\psi(\sqrt{x})-(\sqrt{x+v}-\sqrt{x})\big)^2\mathrm{d}v\Big)\mathrm{d}x = \int_{\epsilon X}^{X+h}\Big(\int_{h/x}^{3h/x}\big(\psi(\sqrt{x+x\delta})-\psi(\sqrt{x})-(\sqrt{x+x\delta}-\sqrt{x})\big)^2\mathrm{d}\delta\Big)\,x\ \mathrm{d}x,$$

where, in the last step, we made the change of variable $\delta=v/x$, thus getting $\delta\geq h/x\geq X^{-5/12+\epsilon_3}$ as in the hypothesis of Claim 2.A[6].

We now use the relation $\epsilon X\leq x\leq X+h$ to remove the dependence on x of the integration interval in δ and to bound x itself. So we get

$$hJ^*_\psi(X,h) \ll (X+h)\int_{\epsilon X}^{X+h}\Big(\int_{h/(X+h)}^{3h/(\epsilon X)}\big(\psi(\sqrt{x+x\delta})-\psi(\sqrt{x})-(\sqrt{x+x\delta}-\sqrt{x})\big)^2\mathrm{d}\delta\Big)\mathrm{d}x = (X+h)\int_{h/(X+h)}^{3h/(\epsilon X)}\Big(\int_{\epsilon X}^{X+h}\big(\psi(\sqrt{x+x\delta})-\psi(\sqrt{x})-(\sqrt{x+x\delta}-\sqrt{x})\big)^2\mathrm{d}x\Big)\mathrm{d}\delta.$$

Finally, using Claim 2.A, we get

$$J^*_\psi(X,h) \ll_\epsilon \frac{X+h}{h}\int_{h/(X+h)}^{3h/(\epsilon X)}\Big(\delta^2X^2\exp\Big(-c_6\Big(\frac{\log X}{\log\log X}\Big)^{1/3}\Big)\Big)\mathrm{d}\delta \ll \frac{X^3}{h}\exp\Big(-c_6\Big(\frac{\log X}{\log\log X}\Big)^{1/3}\Big)\int_0^{3h/(\epsilon X)}\delta^2\mathrm{d}\delta$$

[6]Actually in Claim 2.A, we also have $\delta\leq 1$, but we can drop this condition since, for $\delta\geq 1$, the needed estimate follows from the Prime Number Theorem with error term in Result (R.4).

$$= \frac{X^3}{h}\frac{9h^3}{\epsilon^3 X^3}\exp\Big(-c_6\Big(\frac{\log X}{\log\log X}\Big)^{1/3}\Big) \ll_\epsilon h^2\exp\Big(-c_6\Big(\frac{\log X}{\log\log X}\Big)^{1/3}\Big).$$

This concludes the proof of Claim 2.B. □

Now we are ready to estimate $I(X;\mathfrak{M})$, $I(X;\mathfrak{m})$ and $I(X;\mathfrak{t})$.

2.4 The major arc

In this section we prove the inequality $I(X;\mathfrak{M}) \geq c_1\eta^2 XL$ in (2.10). We follow the argument in Languasco-Zaccagnini [LZ10], §4, suitably modified to deal also with prime squares. Letting

$$T_1(\alpha) = \int_{\epsilon X}^{X} e(t\alpha)\mathrm{d}t \ll_\epsilon \min\Big(X;\frac{1}{|\alpha|}\Big), \tag{2.50}$$

$$T_2(\alpha) = \int_{(\epsilon X)^{1/2}}^{X^{1/2}} e(t^2\alpha)\mathrm{d}t = \frac{1}{2}\int_{\epsilon X}^{X} v^{-1/2}e(v\alpha)\mathrm{d}v \ll_\epsilon X^{-1/2}\min\Big(X;\frac{1}{|\alpha|}\Big), \tag{2.51}$$

we can write

$$\begin{aligned} I(X;\mathfrak{M}) &= \int_{\mathfrak{M}} T_1(\lambda_1\alpha)T_2(\lambda_2\alpha)T_2(\lambda_3\alpha)G(\mu_1\alpha)\cdots G(\mu_k\alpha)e(\varpi\,\alpha)K(\alpha,\eta)\mathrm{d}\alpha \\ &+ \int_{\mathfrak{M}} \Big(S_1(\lambda_1\alpha)-T_1(\lambda_1\alpha)\Big)T_2(\lambda_2\alpha)T_2(\lambda_3\alpha)G(\mu_1\alpha)\cdots G(\mu_k\alpha)e(\varpi\,\alpha)K(\alpha,\eta)\mathrm{d}\alpha \\ &+ \int_{\mathfrak{M}} S_1(\lambda_1\alpha)\Big(S_2(\lambda_2\alpha)-T_2(\lambda_2\alpha)\Big)T_2(\lambda_3\alpha)G(\mu_1\alpha)\cdots G(\mu_k\alpha)e(\varpi\,\alpha)K(\alpha,\eta)\mathrm{d}\alpha \\ &+ \int_{\mathfrak{M}} S_1(\lambda_1\alpha)S_2(\lambda_2\alpha)\Big(S_2(\lambda_3\alpha)-T_2(\lambda_3\alpha)\Big)G(\mu_1\alpha)\cdots G(\mu_k\alpha)e(\varpi\,\alpha)K(\alpha,\eta)\mathrm{d}\alpha \\ &= J_1+J_2+J_3+J_4, \end{aligned} \tag{2.52}$$

say. In what follows we will prove that

$$J_1 \geq \frac{(3-2\sqrt{2})\eta^2 XL^k}{4\left(|\lambda_1|+|\lambda_2|+|\lambda_3|\right)} + \mathcal{O}_\epsilon\big(\eta^2 X^{1/5}L^{k+2}\big) \tag{2.53}$$

and

$$J_2+J_3+J_4 = o\big(\eta^2 XL^k\big), \tag{2.54}$$

thus obtaining, by (2.52)-(2.54), that

$$I(X;\mathfrak{M}) \geq \frac{(3-2\sqrt{2})-\epsilon_4}{4\left(|\lambda_1|+|\lambda_2|+|\lambda_3|\right)}\eta^2 XL^k.$$

This proves that the inequality (2.10) holds with $c_1 = 2^{-2}(3-2\sqrt{2}-\epsilon_4)\left(|\lambda_1|+|\lambda_2|+|\lambda_3|\right)^{-1}$ and $\epsilon_4 > 0$ an arbitrarily small constant.

In order to do this, we need these two estimates:

- The first one is

$$\int_0^1 |S_1(\alpha)|^2 \mathrm{d}\alpha \ll_\epsilon X \log X, \tag{2.55}$$

which is a consequence of the Prime Number Theorem (see Result (R.2)), together with Abel's identity (see Result (R.17)) and Parseval's formula (see, *e.g.*, Apostol [Apo74], Theorem 11.4). In fact

$$\int_0^1 |S_1(\alpha)|^2 \mathrm{d}\alpha = \sum_{\epsilon X \le p \le X} (\log p)^2 = X \log X + \mathcal{O}(X) - \epsilon X \log(\epsilon X) + \mathcal{O}(\epsilon X) \ll_\epsilon X \log X.$$

- The second one is

$$\int_0^1 |S_2(\alpha)|^4 \mathrm{d}\alpha \ll_\epsilon X \log^2 X, \tag{2.56}$$

which is based on Rieger's estimation in Result (R.8).

Estimation of J_2, J_3 and J_4.

We first estimate J_4. We remark that, by Euler's summation formula (see Result (R.16)), we have

$$T_i(\alpha) - U_i(\alpha) \ll 1 + X|\alpha| \quad \text{for every} \quad i = 1, 2. \tag{2.57}$$

So by definition (2.7) of $\mathfrak{M}$, the Cauchy-Schwarz inequality, and (2.55)-(2.57) we get

$$\begin{aligned}
&\int_{\mathfrak{M}} |S_1(\lambda_1\alpha)||S_2(\lambda_2\alpha)||T_2(\lambda_3\alpha) - U_2(\lambda_3\alpha)|\mathrm{d}\alpha \\
&\qquad \ll_{\boldsymbol{\lambda}} \int_{-1/X}^{1/X} |S_1(\lambda_1\alpha)||S_2(\lambda_2\alpha)|\mathrm{d}\alpha + X \int_{1/X}^{P/X} |\alpha||S_1(\lambda_1\alpha)||S_2(\lambda_2\alpha)|\mathrm{d}\alpha \\
&\qquad \ll_{\boldsymbol{\lambda}} X^{-1/4}\Big(\int_0^1 |S_1(\alpha)|^2\mathrm{d}\alpha\Big)^{1/2}\Big(\int_0^1 |S_2(\alpha)|^4\mathrm{d}\alpha\Big)^{1/4} \\
&\qquad\quad + X\Big(\int_{1/X}^{P/X} \alpha^4\mathrm{d}\alpha\Big)^{1/4}\Big(\int_0^1 |S_2(\alpha)|^4\mathrm{d}\alpha\Big)^{1/4}\Big(\int_0^1 |S_1(\alpha)|^2\mathrm{d}\alpha\Big)^{1/2} \\
&\qquad \ll_{\boldsymbol{\lambda},\epsilon} X^{1/2}\log X + P^{5/4}X^{1/2}\log X = o(X),
\end{aligned}$$

since $P = X^{2/5}/\log X$. Hence, using the trivial estimates $|G(\mu_i\alpha)| \le L$ and $K(\alpha,\eta) \ll \eta^2$, we can write

$$\begin{aligned}
J_4 = &\int_{\mathfrak{M}} S_1(\lambda_1\alpha)S_2(\lambda_2\alpha)\Big(S_2(\lambda_3\alpha) - U_2(\lambda_3\alpha)\Big)G(\mu_1\alpha)\cdots G(\mu_k\alpha)e(\varpi\,\alpha)K(\alpha,\eta)\mathrm{d}\alpha \\
&+ o_{\boldsymbol{\lambda},M,\epsilon}\big(\eta^2 X L^k\big).
\end{aligned}$$

Now using again the definition (2.7) of $\mathfrak{M}$, the Cauchy-Schwarz inequality and the trivial estimates $|G(\mu_i\alpha)| \le L$ and $K(\alpha,\eta) \ll \eta^2$, and applying the estimate in (2.55) and Lemmas 2.12-2.13 with $Y = P/X$, and finally observing that $|S_2(\lambda_2\alpha)| \ll_\epsilon X^{1/2}$ (by a direct application of the Prime Number Theorem in Result (R.3)), we have that

$$\begin{aligned} J_4 &\ll \eta^2 L^k X^{1/2}\Big(\int_{\mathfrak{M}} |S_2(\lambda_3\alpha) - U_2(\lambda_3\alpha)|^2 \mathrm{d}\alpha\Big)^{1/2}\Big(\int_{\mathfrak{M}} |S_1(\lambda_1\alpha)|^2\mathrm{d}\alpha\Big)^{1/2} + o_{\boldsymbol{\lambda},M,\epsilon}\big(\eta^2 X L^k\big) \\ &\ll_{\boldsymbol{\lambda},M,\epsilon} \eta^2 L^k X^{1/2}\Big(\int_0^1 |S_1(\alpha)|^2\mathrm{d}\alpha\Big)^{1/2} \exp\Big(-\frac{c_6(\epsilon_3)}{2}\Big(\frac{\log X}{\log\log X}\Big)^{1/3}\Big) + o_{\boldsymbol{\lambda},M}\big(\eta^2 X L^k\big) \\ &\ll_{\boldsymbol{\lambda},M,\epsilon} \eta^2 X L^{k+1/2} \exp\Big(-\frac{c_6(\epsilon_3)}{2}\Big(\frac{\log X}{\log\log X}\Big)^{1/3}\Big) = o\big(\eta^2 X L^k\big). \end{aligned}$$

The integral J_3 can be estimated analogously using the estimate in (2.51) for $T_2(\lambda_3\alpha)$, instead of $|S_2(\lambda_3\alpha)| \ll X^{1/2}$.

For J_2 we argue as follows. First of all, using again (2.57) and (2.51), we get

$$\begin{aligned} \int_{\mathfrak{M}} |T_1(\lambda_1\alpha) - U_1(\lambda_1\alpha)||T_2(\lambda_2\alpha)||T_2(\lambda_3\alpha)|\mathrm{d}\alpha &\ll_{\boldsymbol{\lambda}} X\int_{-1/X}^{1/X}\mathrm{d}\alpha + \int_{1/X}^{P/X}\frac{X|\alpha|}{X\alpha^2}\mathrm{d}\alpha \\ &\ll_{\boldsymbol{\lambda}} 1 + \log P = o(X), \end{aligned}$$

since $P = X^{2/5}/\log X$. Hence, using the trivial estimates $|G(\mu_i\alpha)| \le L$ and $K(\alpha,\eta) \ll \eta^2$, we can write

$$\begin{aligned} J_2 = \int_{\mathfrak{M}} \Big(S_1(\lambda_1\alpha) - U_1(\lambda_1\alpha)\Big) T_2(\lambda_2\alpha)T_2(\lambda_3\alpha)G(\mu_1\alpha)\cdots G(\mu_k\alpha)e(\varpi\,\alpha)K(\alpha,\eta)\mathrm{d}\alpha \\ + o_{\boldsymbol{\lambda},M}\big(\eta^2 X L^k\big). \end{aligned}$$

Using again the definition (2.7) of $\mathfrak{M}$, the Cauchy-Schwarz inequality and the trivial estimates $|G(\mu_i\alpha)| \le L$ and $K(\alpha,\eta) \ll \eta^2$, and applying Lemma 2.11 and Saffari-Vaughan's estimation of the Selberg integral in Result (R.12) with $Y = P/X$, we have

$$\begin{aligned} J_2 &\ll \eta^2 L^k\Big(\int_{\mathfrak{M}} |S_1(\lambda_1\alpha) - U_1(\lambda_1\alpha)|^2\mathrm{d}\alpha\Big)^{1/2}\Big(\int_{\mathfrak{M}} |T_2(\lambda_2\alpha)T_2(\lambda_3\alpha)|^2\mathrm{d}\alpha\Big)^{1/2} \\ &\qquad + o_{\boldsymbol{\lambda},M}\big(\eta^2 X L^k\big) \\ &\ll_{\boldsymbol{\lambda},M,\epsilon} \eta^2 L^k X^{1/2}\Big(\int_{\mathfrak{M}} |T_2(\lambda_2\alpha)T_2(\lambda_3\alpha)|^2\mathrm{d}\alpha\Big)^{1/2} \exp\Big(-\frac{c_6(\epsilon_3)}{2}\Big(\frac{\log X}{\log\log X}\Big)^{1/3}\Big) \\ &\qquad + o_{\boldsymbol{\lambda},M}\big(\eta^2 X L^k\big) \\ &\ll_{\boldsymbol{\lambda},M,\epsilon} \eta^2 X L^k \exp\Big(-\frac{c_6(\epsilon_3)}{2}\Big(\frac{\log X}{\log\log X}\Big)^{1/3}\Big) + o_{\boldsymbol{\lambda},M}\big(\eta^2 X L^k\big) = o\big(\eta^2 X L^k\big), \end{aligned}$$

since $\int_{\mathfrak{M}} |T_2(\lambda_2\alpha)T_2(\lambda_3\alpha)|^2\mathrm{d}\alpha \ll_{\boldsymbol{\lambda}} X$, by the estimate of $T_2(\alpha)$ in (2.51). Hence (2.54) holds. ∎

Estimation of J_1. Recalling that $P = X^{2/5}/\log X$, using the definition (2.7) of $\mathfrak{M}$, the estimates (2.50)-(2.51) for $T_1(\alpha)$ and $T_2(\alpha)$ and the definition (2.52) of J_1, we obtain

$$J_1 = \sum_{1\le m_1\le L} \cdots \sum_{1\le m_k\le L} \mathcal{J}\Big(\mu_1 2^{m_1} + \cdots + \mu_k 2^{m_k} + \varpi, \eta\Big) + \mathcal{O}_\epsilon\big(\eta^2 X^{1/5} L^{k+2}\big), \tag{2.58}$$

where $\mathcal{J}(u,\eta)$ is defined as

$$\begin{aligned}\mathcal{J}(u,\eta) &= \int_{\mathbb{R}} T_1(\lambda_1\alpha)T_2(\lambda_2\alpha)T_2(\lambda_3\alpha)e(u\alpha)K(\alpha,\eta)\mathrm{d}\alpha \\ &= \frac{1}{4}\int_{\epsilon X}^{X}\int_{\epsilon X}^{X}\int_{\epsilon X}^{X} \widehat{K}(\lambda_1u_1+\lambda_2u_2+\lambda_3u_3+u,\eta)u_2^{-1/2}u_3^{-1/2}\mathrm{d}u_1\mathrm{d}u_2\mathrm{d}u_3.\end{aligned}$$

Here the second relation follows by definitions (2.50)-(2.51) of $T_1(\alpha)$ and $T_2(\alpha)$ resp. and interchanging the integration order.

We recall that $\lambda_1<0$ and $\lambda_2,\lambda_3>0$, so if $|u|\le\epsilon X$, for

$$\frac{X|\lambda_1|}{2\left(|\lambda_1|+\lambda_2+\lambda_3\right)} \le u_2,u_3 \le \frac{X|\lambda_1|}{|\lambda_1|+\lambda_2+\lambda_3},$$

and for X sufficiently large and ϵ sufficiently small, we get

$$-\frac{\eta}{2}-(\lambda_2u_2+\lambda_3u_3+u) \le \lambda_1u_1 \le \frac{\eta}{2}-(\lambda_2u_2+\lambda_3u_3+u).$$

Hence there exists an interval for u_1, of length $\eta|\lambda_1|^{-1}$ and entirely contained in $[\epsilon X,X]$, such that $\widehat{K}(\lambda_1u_1+\lambda_2u_2+\lambda_3u_3+u,\eta)\ge\eta/2$ there. So, letting $b=(X|\lambda_1|)/(|\lambda_1|+\lambda_2+\lambda_3)$, we can write that

$$\mathcal{J}(u,\eta) \ge \frac{\eta^2}{8|\lambda_1|}\left(\int_{b/2}^{b} v^{-1/2}\mathrm{d}v\right)^2 = \frac{(3-2\sqrt{2})\eta^2X}{4\left(|\lambda_1|+\lambda_2+\lambda_3\right)}.$$

By definition of L, we have that $|\mu_12^{m_1}+\cdots+\mu_k2^{m_k}+\varpi|\le\epsilon X$, for X sufficiently large and hence, by (2.58), we obtain

$$J_1 \ge \frac{(3-2\sqrt{2})\eta^2XL^k}{4\left(|\lambda_1|+\lambda_2+\lambda_3\right)} + \mathcal{O}_\epsilon\big(\eta^2X^{1/5}L^{k+2}\big),$$

thus proving the bound (2.53).■

2.5 The trivial arc

In this section we prove the estimate $|I(X;\mathfrak{t})| = o\big(XL^k\big)$ in (2.9), following the argument in Parsell [Par03]. Recalling the definition (2.7) of $\mathfrak{t}$, the trivial estimate $|G(\mu_i\alpha)|\le L$ and using twice the Cauchy-Schwarz inequality, we get

$$\begin{aligned}|I(X;\mathfrak{t})| \ll{}& L^k\Big(\int_{L^2}^{+\infty}|S_1(\lambda_1\alpha)|^2K(\alpha,\eta)\mathrm{d}\alpha\Big)^{1/2} \\ &\times\Big(\int_{L^2}^{+\infty}|S_2(\lambda_2\alpha)|^4K(\alpha,\eta)\mathrm{d}\alpha\Big)^{1/4}\Big(\int_{L^2}^{+\infty}|S_2(\lambda_3\alpha)|^4K(\alpha,\eta)\mathrm{d}\alpha\Big)^{1/4}.\end{aligned}$$

By the estimation of $K(\alpha,\eta)$ given in (K.ii) and making a suitable change of variable, we have that for $i=2,3$

$$\int_{L^2}^{+\infty}|S_2(\lambda_i\alpha)|^4K(\alpha,\eta)\mathrm{d}\alpha \ll_{\boldsymbol{\lambda}} \int_{\lambda_iL^2}^{+\infty}\frac{|S_2(\alpha)|^4}{\alpha^2}\mathrm{d}\alpha \ll \sum_{n\geq\lambda_iL^2}\frac{1}{(n-1)^2}\int_{n-1}^{n}|S_2(\alpha)|^4\mathrm{d}\alpha$$
$$\ll_{\boldsymbol{\lambda}} L^{-2}\int_0^1|S_2(\alpha)|^4\mathrm{d}\alpha \ll_{\boldsymbol{\lambda},M,\epsilon} X,$$

where, in the last step, we used the estimation in (2.56). For the "S_1-integral" above, we argue in a similar way, obtaining

$$\int_{L^2}^{+\infty}|S_1(\lambda_1\alpha)|^2K(\alpha,\eta)\mathrm{d}\alpha \ll_{\boldsymbol{\lambda}} \int_{\lambda_1L^2}^{+\infty}\frac{|S_1(\alpha)|^2}{\alpha^2}\mathrm{d}\alpha \ll \sum_{n\geq\lambda_1L^2}\frac{1}{(n-1)^2}\int_{n-1}^{n}|S_1(\alpha)|^2\mathrm{d}\alpha$$
$$\ll_{\boldsymbol{\lambda}} L^{-2}\int_0^1|S_1(\alpha)|^2\mathrm{d}\alpha \ll_{\boldsymbol{\lambda},M,\epsilon}\frac{X}{\log X},$$

where in the last step we used the estimation in (2.55). Hence the asymptotic formula (2.9) holds.

2.6 The minor arc

In this section we prove the inequality $|I(X;\mathfrak{m})|\leq c_2(k)\eta XL^k$ in (2.11), using Pintz-Ruzsa's work in [PR03]. By definition we have

$$I(X;\mathfrak{m})=\int_{\mathfrak{m}}S_1(\lambda_1\alpha)S_2(\lambda_2\alpha)S_2(\lambda_3\alpha)G(\mu_1\alpha)\cdots G(\mu_k\alpha)e(\varpi\,\alpha)K(\alpha,\eta)\mathrm{d}\alpha.$$

Letting $c\in(0,1)$ to be chosen later, we first split $\mathfrak{m}$ as $\mathfrak{m}_1\sqcup\mathfrak{m}_2$, where $\sqcup$ denotes a disjoint union and $\mathfrak{m}_2$ is the set of those $\alpha\in\mathfrak{m}$ such that $|G(\mu_i\alpha)|>\nu(c)L$, for some $i\in\{1,\dots,k\}$ and with $\nu(c)$ defined in Lemma 2.10. We want to choose c in order to get $|I(X;\mathfrak{m}_2)|=o(\eta X)$, since, by Lemma 2.10, we know that $|\mathfrak{m}_2|\ll_{M,\epsilon}kL^2X^{-c}$.

To this end, using the trivial estimates $|G(\mu_i\alpha)|\leq L$ and $K(\alpha,\eta)\ll\eta^2$, the Cauchy-Schwarz inequality and Lemma 2.9, we obtain

$$\begin{aligned}|I(X;\mathfrak{m}_2)|&\leq\eta^2L^k\left(\sup_{\alpha\in\mathfrak{m}_2}|S_2(\lambda_2\alpha)S_2(\lambda_3\alpha)|\right)\left(\int_{\mathfrak{m}_2}|S_1(\lambda_1\alpha)|\mathrm{d}\alpha\right)\\&\ll_{\boldsymbol{\lambda}}\eta^2L^k|\mathfrak{m}_2|^{1/2}\left(\sup_{\alpha\in\mathfrak{m}}|S_2(\lambda_2\alpha)S_2(\lambda_3\alpha)|\right)\left(L^2\int_0^1|S_1(\alpha)|^2\mathrm{d}\alpha\right)^{1/2}\\&\ll_{\boldsymbol{\lambda},M,\epsilon,\epsilon_2}k^{1/2}\eta^2L^{k+3}X^{23/16+\epsilon_2-c/2},\end{aligned}$$

where $\epsilon_2>0$ is a sufficiently small constant and $X=q^2$, with q the denominator of a convergent to the continued fraction for λ_2/λ_3. Taking $c=7/8+10^{-20}$ as in (2.22) and applying Pintz-Ruzsa's lemma, we obtain that

$$|I(X;\mathfrak{m}_2)|=o(\eta X),\tag{2.59}$$

for $\nu = 0.9505087500$. We remark that neither the result of Kumchev [Kum06] nor the approach of Cook, Fox and Harman (see [CF01], [Har04], [CH06]) seem to give any improvement of the previous estimates.

Now we evaluate the contribution of $\mathfrak{m}_1$, using the Cauchy-Schwarz inequality and Lemmas 2.5, 2.8. So

$$\begin{aligned}|I(X;\mathfrak{m}_1)| &\leq (\nu L)^{k-3}\Big(\int_{\mathfrak{m}} |S_1(\lambda_1\alpha)G(\mu_1\alpha)|^2 K(\alpha,\eta)\mathrm{d}\alpha\Big)^{1/2} \\ &\quad \times \Big(\int_{\mathfrak{m}} |S_2(\lambda_2\alpha)G(\mu_2\alpha)|^4 K(\alpha,\eta)\mathrm{d}\alpha\Big)^{1/4}\Big(\int_{\mathfrak{m}} |S_2(\lambda_3\alpha)G(\mu_3\alpha)|^4 K(\alpha,\eta)\mathrm{d}\alpha\Big)^{1/4} \\ &< \nu^{k-3}\boldsymbol{C}\eta X L^k, \end{aligned} \tag{2.60}$$

where $\boldsymbol{C} = \boldsymbol{C}(q_1, q_2, q_3, \epsilon_1)$ is defined as in (2.4) (we remark that ϵ_1 is not necessarily the same at each occurrence).

Hence, by (2.59)-(2.60), we finally get that

$$|I(X;\mathfrak{m})| < (0.9505087500)^{k-3}\boldsymbol{C}\eta X L^k$$

holds, for X sufficiently large. This means that the inequality (2.11) holds with

$$c_2(k) = (0.9505087500)^{k-3}\boldsymbol{C}. \tag{2.61}$$

2.7 Proof of the theorem

We have to verify if there exists a $k_0 \in \mathbb{N}$ such that the condition (2.12) holds for X sufficiently large and $X = q^2$, where q is the denominator of a convergent of the continued fraction for λ_2/λ_3. Combining the inequalities (2.10)-(2.11), with $c_2(k)$ as in (2.61), we obtain that (2.12) holds for $k \geq k_0$, with k_0 as defined in (2.3). This completes the proof of our Theorem A.

Chapter 3

On the sum of two primes and k powers of $g \geq 3$

In this chapter we introduce our work about a problem with two prime and k powers of $g \geq 3$. It is a variation of Languasco-Pintz-Zaccagnini's work [LPZ07] about the Goldbach-Linnik problem, since we generalize their result to powers of g, instead of powers of 2. More precisely, in the first section we state our result and we compare it with Languasco-Pintz-Zaccagnini's one. In Section 3.2 we fix the general setting, introducing the needed notation, while in Section 3.3 we set up our problem. In Sections 3.4-3.5 we prove the partial results necessary in the proof of our theorem. In particular, in Section 3.4 we estimate contribution of the minor arcs, studying the size of the exceptional set; to do this, we just replace the trivial estimation for the exponential sum over powers of 2 used in [LPZ07], with the trivial estimation for the exponential sum over powers of g: in this way we obtain the same bound for the exceptional set, which is the optimal one using Vaughan's estimation in Result (R.9). On the contrary, in Section 3.5 a more careful treatment of the powers of g is needed (see the study of $R^{(5)}_{\mathfrak{M}}(N)$). In Section 3.6 we gather such partial results, proving our theorem. Finally, in the following Section 3.7, we collect and prove the new results used in this chapter. In particular, Lemma 3.6 is a variation, for powers of g, of a result of Romanov in [Rom34]; Lemma 3.8 is a sharper version of Lemma 1.2 in Murty-Rosen-Silverman [MRS96] and it does not depend on g, while Corollary 3.9 can be consider as an application of Lemma 3.8 to powers of g. In Lemma 3.10 we study an arithmetic sum which involves the singular series over powers of g and therefore a careful analysis of this arithmetic part is needed; this lemma is a variation, for powers of g, of Lemma 6.2 in Languasco-Pintz-Zaccagnini [LPZ07]. The last Lemma 3.14 is about the estimation of an infinite convergent product, which is effectively computed (with a specified error term) in the PARI-GP program written in Section 3.8.

3.1 Introduction to our result

The starting point for our work is the paper [LPZ07] by Languasco, Pintz and Zaccagnini concerning the Goldbach-Linnik problem, where the authors proved the validity, for almost all even integers, of a suitable asymptotic formula for the number of representations of a large even integer as sum of two primes and a bounded number of powers of 2.

More precisely, they proved the following

Theorem (Languasco-Pintz-Zaccagnini). *Let $k \geq 1$ be a fixed integer. Let $\eta > 0$ be given. Let $X > X_0(k, \eta)$ be a sufficiently large parameter and $L_2 = \log_2 X$. Let moreover $N \leq X$ be an even integer. If we define*

$$R''_k(N) = \sum_{1 \leq m_1, m_2 \leq X} \sum_{\substack{1 \leq \nu_1, \ldots, \nu_k \leq L_2 \\ m_1 + m_2 + 2^{\nu_1} + \ldots + 2^{\nu_k} = N}} \Lambda(m_1)\Lambda(m_2),$$

where $\Lambda(n)$ is the von Mangoldt function, then there exists a constant $\boldsymbol{C} = \boldsymbol{C}(k, N) \in [1, 2]$ such that

$$|R''_k(N) - \boldsymbol{C} N L_2^k| \leq \eta N L_2^k,$$

for all even integers $N \in [1, X]$, apart from at most $\mathcal{O}_k\big(X^{3/5}(\log X)^{10}\big)$ exceptions.

Proof. [LPZ07], Theorem, at page 2. □

We remark that $R''_k(N)$ is the weighted function associate to the following counting function (defined for N even)

$$r''_k(N) = |\, \{(p_1, p_2, \nu_1, \ldots, \nu_k) \in \mathfrak{P} \times [1, L_2]^k \,:\, N = p_1 + p_2 + 2^{\nu_1} + \ldots + 2^{\nu_k}\} \,|,$$

which counts the number of representations of an even integer N as sum of two primes and k powers of 2, being $\mathfrak{P}$ is the set of all prime numbers.

The main ingredient in Languasco-Pintz-Zaccagnini's proof is the Hardy-Littlewood circle method described in Section 1.3.1: they first reduced their problem to studying an integral, defined over the unit interval $[0, 1]$, which has as integrand function some suitable exponential sums over primes and over powers of 2; then they split the integration interval into major and minor arcs. Finally, they estimated the respective integrals one by one. The key step is the estimation on the major arcs. More precisely:

- On the minor arcs, the estimate is straightforward and the main ingredient is Vaughan's Lemma in Result (R.9), concerning the exponential sum over primes. Using it (and trivially estimating the exponential sums over powers of 2) Languasco, Pintz and Zaccagnini proved that the contribution on the minor arcs is small for almost all even integers.

- On the major arcs, the estimation is much more elaborated and the key point is Pintz's explicit formula in Section 1.3.2, which describes the behaviour of the exponential sum over primes on the major arcs. Using Pintz's formula, Languasco, Pintz and Zaccagnini reduced their problem to the study of a convergent series, which is actually related to the Romanov constant (see Pintz [Pin06a], §5, and Khalfalah-Pintz [KP06], Corollary 1), and of a suitable average of the Goldbach singular series.

Replacing the powers of 2 with powers of $g \geq 3$, we obtain the problem of studying the following counting function (defined for suitable integers N):

$$r''_{k,g}(N) = |\, \{(p_1, p_2, \nu_1, \ldots, \nu_k) \in \mathfrak{P}^2 \times [1, L]^k \,:\, N = p_1 + p_2 + g^{\nu_1} + \ldots + g^{\nu_k}\} \,|$$

with $L = \log_g X$, which is a natural generalization of the Goldbach-Linnik problem (see, *e.g.*, page 2 of Gallagher [Gal75]). Our idea is to prove, for this generalization, an analogous of Languasco-Pintz-Zaccagnini's theorem. More precisely, we prove the following.

Theorem B (Settimi)**.** *Let $k \geq 1$, $g \geq 3$ be fixed integers. Let $\eta > 0$ be given. Let $X > X_0(k, \eta, g)$ be a sufficiently large parameter and $L = \log_g X$. Let moreover $N \leq X$ be an integer that satisfies the following arithmetic conditions:*

$$\begin{cases} N \text{ even} & \text{if } g \text{ even,} \\ N \equiv k \pmod 2 & \text{if } g \text{ odd.} \end{cases} \tag{A.C.}$$

If we define

$$R''_{k,g}(N) = \sum_{1 \leq m_1, m_2 \leq X} \sum_{\substack{1 \leq \nu_1, \dots, \nu_k \leq L \\ m_1 + m_2 + g^{\nu_1} + \dots + g^{\nu_k} = N}} \Lambda(m_1)\Lambda(m_2),$$

where $\Lambda(n)$ is the von Mangoldt function, then there exists a constant $\boldsymbol{C_g} = \boldsymbol{C_g}(k, N)$ such that

$$|R''_{k,g}(N) - \boldsymbol{C_g} N L^k| \leq \eta N L^k,$$

for all $N \in [1, X]$ satisfying (A.C.)*, apart from at most $\mathcal{O}_g\big(X^{3/5}(\log X)^{10}\big)$ exceptions.*

Moreover

$$\boldsymbol{C_g} \leq 2c_0 e^{\gamma} \cdot 0.7574 \Big(\log\log g + 3 + \log 2 + \frac{\pi}{2} \Big) + \mathcal{O}\Big(\frac{1}{\log\log g} \Big),$$

where c_0 is the twin-prime constant and $\boldsymbol{\gamma}$ is the Euler constant[1]*.*

We remark that $R''_{k,g}(N)$ is the weighted function associate to the relevant counting function $r''_{k,g}(N)$ defined above and that the arithmetic conditions (A.C.) are standard for this kind of problem (see again [Gal75]).

This chapter is devoted to a detailed proof of our Theorem B, which follows the lines of Languasco-Pintz-Zaccagnini's proof of their theorem, but adapted to powers of g: the key point involves the arithmetic part, since a careful analysis of the non-coprimality to g is needed. More precisely:

- On minor arcs, we use the same estimation as [LPZ07].
- On major arcs, as in [LPZ07], we use Pintz's explicit formula (see Theorem 1.3 of Section 1.3.2) to reduce our problem to the study of a convergent series and of an arithmetic sum. This reduction step is more complicated than in [LPZ07], because of the possible non-coprimality to g of some parameters (see the estimation of $R^{(5)}_{\mathfrak{M}}(N)$ below).
  - To study the convergent series, we replaced Khalfalah-Pintz's technique used in [LPZ07], with a new result (see Lemma 3.6 and the auxiliary results in Lemma 3.8 and Corollary 3.9). In order to do this, we follow the argument in Murty-Rosen-Silverman [MRS96], but using sharper estimations (compare our Lemma 3.8 with Lemma 1 in [MRS96]).

[1] We recall (see the preliminaries in Chapter 1) that $\gamma \approx 0.5772156$ and, according to Gourdon-Sebah [GS01], $0.66016181584 < c_0 < 0.66016181585$.

– To study the arithmetic sum, we carefully adapt the technique in [LPZ07] to powers of g.

We conclude the comparison to [LPZ07], with the following remark.

Remark 3.1. *When $g = 2$, the arithmetic conditions* (A.C.) *trivially imply that N is even, which is exactly the Goldbach-Linnik problem. For this case we refer to [LPZ07], since the argument in our proof leads to a constant $\boldsymbol{C_g}$ which is, for $g = 2$, worst than Languasco-Pintz-Zaccagnini's $\boldsymbol{C}$. This is due to the fact that, in [LPZ07], following the line of Corollary 1 in Khalfalah-Pintz [KP06], a more refined analysis of the convergent series is performed.*

We do not try to replicate Khalfalah-Pintz's technique in our situation, since, even finding a more accurate bound for $\boldsymbol{C_g}$, the size of the exceptional set does not significantly change.

3.2 Definitions and general setting

In this section we set the notation in order to use the Hardy-Littlewood circle method, described in Section 1.3.1, to count the number of representations of an integer N (satisfying some suitable arithmetic conditions) as

$$N = p_1 + p_2 + g^{\nu_1} + \ldots + g^{\nu_k},$$

as in the statement of Theorem B.

To this end, let $k \geq 1$ and $g \geq 3$ be two fixed positive integers (the constants implied by the "$\mathcal{O}(\cdot)$" and "$\ll$" notations will silently depend on g and k). Let X be a large parameter, say $X > X_0(k, g, \eta)$ with $\eta > 0$ a sufficiently small constant, not necessarily the same at each occurrence. Let also $\mathcal{J}(X) = [2X/3, X]$, $L = \log_g X$ and $L' = L - L^{1/2}$. Finally, let $N \in \mathcal{J}(X)$[2].

Besides the counting function $r''_{k,g}(N)$ already defined, we also need the following functions

$$\begin{aligned} r_{\mathrm{Gb}}(N) &= |\,\{(p_1, p_2) \in \mathfrak{P}^2 \,:\, N = p_1 + p_2\}\,|, \\ t_{k,g}(N) &= |\,\{(\nu_1, \ldots, \nu_k) \in [1, L]^k \,:\, N = g^{\nu_1} + \ldots + g^{\nu_k}\}\,|, \\ t'_{k,g}(N) &= |\,\{(\nu_1, \ldots, \nu_k) \in [1, L']^k \,:\, N = g^{\nu_1} + \ldots + g^{\nu_k}\}\,|, \end{aligned} \tag{3.1}$$

which count respectively the number of representations of an even integer as sum of two primes (*i.e.* the "Goldbach representations") and the number of representations of a suitable integer as sum of k powers of g, with exponents restricted to two different intervals.

It's easy to see that the requirement $\nu_i \in [1, L]$ in the definition of $t_{k,g}(N)$, as well as in the definition of $r''_{k,g}(N)$, is not restrictive since $g^L = X \geq N \geq g^{\nu_1} + \ldots + g^{\nu_k} > g^{\nu_i}$ for every $1 \leq i \leq k$. On the contrary, the hypothesis $\nu_i \in [1, L']$ in the definition of $t'_{k,g}(N)$ is restrictive.

We can also define the following weighted function associated to $r_{\mathrm{Gb}}(N)$

$$R_{\mathrm{Gb}}(N) = \sum_{\substack{1 \leq m_1, m_2 \leq X \\ m_1 + m_2 = N}} \Lambda(m_1)\Lambda(m_2).$$

[2]We set $N \in \mathcal{J}(X)$ (but it would actually be sufficient $\mathcal{J}(X) = [\delta X, X]$ for any $0 < \delta < 1$), instead of $N \in [1, X]$, but in Section 3.6 we extend our result to the whole $[1, X]$, by using a dyadic argument.

The relationship between the weighted functions $R''_{k,g}(N)$ and $R_{\mathrm{Gb}}(N)$ is given by the formula

$$R''_{k,g}(N) = \sum_{1\leq n\leq N} R_{\mathrm{Gb}}(n)t'_{k,g}(N-n) + \mathcal{O}\big(NL^{k-1/2}\log\log N\big). \tag{3.2}$$

In fact, rearranging the sums in the definition of $R''_{k,g}(N)$, we have:

$$\begin{aligned}
R''_{k,g}(N) &= \sum_{0\leq m\leq N} \sum_{\substack{1\leq \nu_1,\dots,\nu_k\leq L\\ g^{\nu_1}+\dots+g^{\nu_k}=m}} \sum_{\substack{1\leq m_1,m_2\leq X\\ m_1+m_2=N-m}} \Lambda(m_1)\Lambda(m_2)\\
&= \sum_{0\leq m\leq N} R_{\mathrm{Gb}}(N-m) \sum_{\substack{1\leq \nu_1,\dots,\nu_k\leq L\\ g^{\nu_1}+\dots+g^{\nu_k}=m}} 1\\
&= \sum_{0\leq m\leq N} R_{\mathrm{Gb}}(N-m)\Big(\sum_{\substack{1\leq \nu_1,\dots,\nu_k\leq L'\\ g^{\nu_1}+\dots+g^{\nu_k}=m}} 1 + \sum_{L'\leq \nu\leq L} \sum_{\substack{1\leq \nu_1,\dots,\nu_{k-1}\leq L\\ g^{\nu_1}+\dots+g^{\nu_{k-1}}+g^{\nu}=m}} 1\Big)\\
&= \sum_{0\leq m\leq N} R_{\mathrm{Gb}}(N-m)t'_{k,g}(m) + \sum_{0\leq m\leq N} R_{\mathrm{Gb}}(N-m)\mathcal{O}\big(L^{k-1}(L-L')\big)\\
&= \sum_{0\leq n\leq N} R_{\mathrm{Gb}}(n)t'_{k,g}(N-n) + \mathcal{O}\big(L^{k-1/2}\big)\sum_{0\leq n\leq N} R_{\mathrm{Gb}}(n)\\
&= \sum_{1\leq n\leq N} R_{\mathrm{Gb}}(n)t'_{k,g}(N-n) + \mathcal{O}\big(NL^{k-1/2}\log\log N\big).
\end{aligned}$$

The last step is due to the fact that $R_{\mathrm{Gb}}(0) = 0$ and $R_{\mathrm{Gb}}(N) \ll N\mathfrak{S}(N) \ll N\log\log N$ by Theorem 3.11 of Halberstam-Richert [HR74] (see also the subsequent remark at page 317). We recall that $\mathfrak{S}(n)$ is the singular series of the Goldbach problem already defined in the previous chapters (see, *e.g.*, the preliminaries at Chapter 1) and it will be crucial in evaluating the major arcs contribution (see Section 3.5).

3.3 Hardy-Littlewood circle method

We want to use the Hardy-Littlewood circle method to study our relevant weighted function $R''_{k,g}(N)$, therefore we need the following exponential sums

$$\begin{aligned}
G_g(\alpha) &= \sum_{1\leq m\leq L'} e(g^m\alpha) && \text{exponential sum over powers of } g,\\
S(\alpha) &= \sum_{1\leq m\leq X} \Lambda(m)e(m\alpha) && \text{weighted exponential sum over primes,}
\end{aligned}$$

where $e(x) = e^{2\pi\imath x}$ as usual. We start by dissecting the unit circle, using the parameters

$$P \in [X^{2/5}, X^{41/100}], \qquad Q = X/P,$$

(the fact that P cannot be directly fixed and its upper bound depend on Pintz's explicit formula stated in Theorem 1.3 of Section 1.3.2). Let now

$$\mathfrak{M} = \bigsqcup_{1\le q\le P} \bigsqcup_{\substack{1\le a\le q\\(a,q)=1}} \left[\frac{a}{q} - \frac{1}{qQ}, \frac{a}{q} + \frac{1}{qQ}\right], \qquad \mathfrak{m} = \left[\frac{1}{Q}, 1+\frac{1}{Q}\right] \setminus \mathfrak{M}, \tag{3.3}$$

where $\sqcup$ means disjoint union. This construction is called the Farey dissection of the unit interval, with $\mathfrak{M}$ major arcs and $\mathfrak{m}$ minor arcs (see, *e.g.*, the fundamental paper [MV75] by Montgomery-Vaughan).

With this notation, we can split $R_{\mathrm{Gb}}(N)$ and $R''_{k,g}(N)$ as follows: for $\mathcal{S} \in \{\mathfrak{M}, \mathfrak{m}\}$, let us define

$$R_{\mathcal{S}}(n) = \int_{\mathcal{S}} S^2(\alpha) e(-n\alpha) \mathrm{d}\alpha.$$

Then

$$R_{\mathrm{Gb}}(n) = R_{\mathfrak{M}}(n) + R_{\mathfrak{m}}(n), \tag{3.4}$$

as at page 773 of Languasco-Pintz-Zaccagnini [LPZ07]. With a similar argument we have that, if we define

$$R''_{\mathcal{S}}(N) = \int_{\mathcal{S}} S^2(\alpha) G_g^k(\alpha) e(-N\alpha) \mathrm{d}\alpha, \tag{3.5}$$

then

$$R''_{k,g}(N) = R''_{\mathfrak{M}}(N) + R''_{\mathfrak{m}}(N) + \mathcal{O}\big(NL^{k-1/2}\log\log N\big). \tag{3.6}$$

In fact, by definition of $S(\alpha)$ and $G_g(\alpha)$ we obtain

$$\begin{aligned}
S^2(\alpha)G_g^k(\alpha) &= \sum_{1\le m_1,m_2\le X} \Lambda(m_1)\Lambda(m_2) e((m_1+m_2)\alpha) \sum_{1\le \nu_1,\dots,\nu_k\le L'} e(\alpha(g^{\nu_1}+\dots+g^{\nu_k}))\\
&= \sum_{1\le m_1,m_2\le X} \sum_{1\le \nu_1,\dots,\nu_k\le L'} \Lambda(m_1)\Lambda(m_2) e(\alpha(m_1+m_2+g^{\nu_1}+\dots+g^{\nu_k}))\\
&= \sum_{n\ge 1} \Big(\sum_{1\le m_1,m_2\le X} \sum_{\substack{1\le \nu_1,\dots,\nu_k\le L'\\ m_1+m_2+g^{\nu_1}+\dots+g^{\nu_k}=n}} \Lambda(m_1)\Lambda(m_2)\Big) e(\alpha n)\\
&= \sum_{n\ge 1} \Big(\sum_{1\le m\le n} R_{\mathrm{Gb}}(m) t'_{k,g}(n-m)\Big) e(\alpha n).
\end{aligned}$$

By relation (3.2) and the Fourier coefficients formula (see, *e.g.*, Apostol [Apo74], §11.4), we have

$$\begin{aligned}
R''_{k,g}(N) + \mathcal{O}\big(NL^{k-1/2}\log\log N\big) &= \sum_{1\le m\le N} R_{\mathrm{Gb}}(m) t'_{k,g}(N-m)\\
&= \int_{1/Q}^{1/Q+1} S^2(\alpha) G_g^k(\alpha) e(-N\alpha)\mathrm{d}\alpha
\end{aligned}$$

$$= R''_{\mathfrak{M}}(N) + R''_{\mathfrak{m}}(N).$$

Using again the relation (3.2) and the splitting formula (3.4) for $R_{\mathrm{Gb}}(N)$, we immediately get

$$R''_{k,g}(N) = \sum_{1\leq n\leq N} R_{\mathfrak{M}}(n)t'_{k,g}(N-n) + \sum_{1\leq n\leq N} R_{\mathfrak{m}}(n)t'_{k,g}(N-n) + \mathcal{O}\big(NL^{k-1/2}\log\log N\big),$$

since

$$R''_{\mathcal{S}}(N) = \sum_{1\leq n\leq N} R_{\mathcal{S}}(n)t'_{k,g}(N-n). \tag{3.7}$$

Remark 3.2. *$R''_{\mathcal{S}}(N) \neq 0$ implies that N satisfies the arithmetic conditions* (A.C.)*: in fact it is easy to see that $R_{\mathcal{S}}(n) \neq 0$ in* (3.7) *implies that n is even, which is equivalent to the fact that N verifies* (A.C.)*.*

3.4 Minor arcs

In this section we prove that the contribution of the minor arcs is small, except for few exceptional values. Here we follow the lines of [LPZ07], §3, and we use the same trivial estimation (see inequality 3.8 below).

The first step is estimating the L^2-average of $R''_{\mathfrak{m}}(N)$, for N satisfying (A.C.).

Lemma 3.3. *Let $R''_{\mathfrak{m}}(N)$ be as defined in* (3.5)*, then*

$$\sum_{N\in\mathcal{J}(X)} |R''_{\mathfrak{m}}(N)|^2 \ll XL^{2k}\log X \max_{\alpha\in\mathfrak{m}} |S(\alpha)|^2.$$

Proof. The trivial estimate for $G_g(\alpha)$ is

$$|G_g(\alpha)| = \Big|\sum_{1\leq m\leq L'} e(g^m\alpha)\Big| \leq \sum_{1\leq m\leq L'} |e(g^m\alpha)| = L' \leq L. \tag{3.8}$$

Using it, the definition of $R''_{\mathfrak{m}}(N)$ and Bessel's inequality (see, *e.g.*, [Apo74], §11.5), we get

$$\begin{aligned}\sum_{N\in\mathcal{J}(X)} |R''_{\mathfrak{m}}(N)|^2 &= \sum_{N\in\mathcal{J}(X)} \left|\int_{\mathfrak{m}} S^2(\alpha)G_g^k(\alpha)e(-N\alpha)\mathrm{d}\alpha\right|^2 \\ &\leq \int_{\mathfrak{m}} |S^2(\alpha)G_g^k(\alpha)|^2\mathrm{d}\alpha \\ &\leq L^{2k}\int_{\mathfrak{m}} |S(\alpha)|^4\mathrm{d}\alpha \\ &\leq L^{2k}\max_{\alpha\in\mathfrak{m}}|S(\alpha)|^2\int_{\mathfrak{m}} |S(\alpha)|^2\mathrm{d}\alpha\end{aligned}$$

$$\leq L^{2k} \max_{\alpha\in\mathfrak{m}} |S(\alpha)|^2 \int_{1/Q}^{1/Q+1} |S(\alpha)|^2 \mathrm{d}\alpha.$$

By Parseval's theorem and the Prime Number Theorem with error term (see Result (R.4)), we can write

$$\int_{1/Q}^{1+1/Q} |S(\alpha)|^2 \mathrm{d}\alpha = \sum_{1\leq m\leq X} \Lambda^2(m) = X\log X + \mathcal{O}(X).$$

Therefore we conclude

$$\sum_{N\in\mathcal{J}(X)} |R''_{\mathfrak{m}}(N)|^2 \ll XL^{2k}\log X \max_{\alpha\in\mathfrak{m}} |S(\alpha)|^2. \qquad \square$$

To bound $S(\alpha)$ for $\alpha\in\mathfrak{m}$, we use Vaughan's estimate in Result (R.9), that is

$$S(\alpha) \ll \left(\frac{X}{\sqrt{q}} + \sqrt{qX} + X^{4/5}\right)(\log X)^4, \qquad \text{for } (a,q)=1,\ \left|\alpha - \frac{a}{q}\right| < \frac{1}{q^2},$$

which implies

$$S(\alpha) \ll X^{4/5}(\log X)^4, \tag{3.9}$$

for every $\alpha\in\mathfrak{m}$. In fact, recalling the definition (3.3) of $\mathfrak{M}$, we have that $P<q<Q$ for every $\alpha\in\mathfrak{m}$. Moreover by construction $P\geq X^{2/5}$. Thus $X/q^{1/2}$ and $(qX)^{1/2}$ are bounded above by $X^{4/5}$.

Combining Lemma 3.3 and (3.9), we get

$$\sum_{N\in\mathcal{J}(X)} |R''_{\mathfrak{m}}(N)|^2 \ll XL^{2k}\log X(X^{4/5}(\log X)^4)^2 = X^{13/5}L^{2k}(\log X)^9. \tag{3.10}$$

Using this, we prove that, except for "few" exceptions, $R''_{\mathfrak{m}}(N)$ is "sufficiently small" (that is, by the splitting formula (3.6), its contribution to $R''_{k,g}(N)$ is negligible). More precisely, we conclude with the following lemma.

Lemma 3.4. *Letting*

$$E(X) = \{N\in\mathcal{J}(X) \,:\, N \text{ verifies (A.C.) and } |R''_{\mathfrak{m}}(N)| \geq NL^{k-1/2}\},$$

we have

$$|E(X)| \ll X^{3/5}L(\log X)^9.$$

Proof. Splitting the sum in (3.10) into $|R''_{\mathfrak{m}}(N)|\geq NL^{k-1/2}$ and $|R''_{\mathfrak{m}}(N)| < NL^{k-1/2}$, we obtain

$$X^{13/5}L^{2k}(\log X)^9 \gg \sum_{\substack{N\in\mathcal{J}(X)\\ |R''_{\mathfrak{m}}(N)|\geq NL^{k-1/2}}} |R''_{\mathfrak{m}}(N)|^2 + \sum_{\substack{N\in\mathcal{J}(X)\\ |R''_{\mathfrak{m}}(N)|< NL^{k-1/2}}} |R''_{\mathfrak{m}}(N)|^2$$

$$\gg \sum_{\substack{N\in\mathcal{J}(X)\\ |R''_{\mathfrak{m}}(N)|\geq NL^{k-1/2}}} (NL^{k-1/2})^2 \geq N^2L^{2k-1}|E(X)|.$$

Thus, recalling that $N \geq 2X/3$, we get

$$|E(X)| \ll X^{13/5}L(\log X)^9N^{-2} \leq X^{13/5}L(\log X)^9(2X/3)^{-2} \ll X^{3/5}L(\log X)^9. \qquad \square$$

3.5 Major arcs

In this section we estimate the contribution of the major arcs to $R''_{k,g}(N)$. Since $R''_{\mathfrak{M}}(N) = \sum_{1\leq n\leq N} R_{\mathfrak{M}}(n)t'_{k,g}(N-n)$ by (3.7), we start by studying $R_{\mathfrak{M}}(N)$. To this end, our main tools are the explicit formula by Pintz in [Pin09] and, in particular, the related corollary by Languasco-Pintz-Zaccagnini in [LPZ07]. They are, respectively, Theorem 1.3 and Corollary 1.5 in Section 1.3.2. We refer to our Section 1.3.2 for the definitions of generalized exceptional zeros, moduli and characters and we just summarize here the statement of Corollary 1.5:

(C.1) $R_{\mathfrak{M}}(n) \ll n\mathfrak{S}(n)$, for every even $n \in [X/2, X]$ (and thus $n \in \mathcal{J}(X) = [2X/3, X]$).

(C.2) $|R_{\mathfrak{M}}(n) - n\mathfrak{S}(n)| \leq \eta n\mathfrak{S}(n)$, if there exists a sufficiently small absolute constant $\eta > 0$ and a constant $C'(\eta)$ such that
$$r_i \nmid C'(\eta)n,$$
for every generalized exceptional modulus r_i.

(C.3) $r_i^* \gg (\log X)^2$, for any generalized exceptional modulus r_i, where a^* denotes the odd square-free part of $a \in \mathbb{N}$.

(C.4) $0 \leq K \leq C_1(\eta)$, where K is the number of generalized exceptional character χ_i and $C_1(\eta)$ is a positive constant.

We now split our interval $\mathcal{J}(X)$ according to points (C.1)-(C.2). More precisely let us call $\mathcal{J}_e(X)$ the set of those $n \in \mathcal{J}(X)$ which are even and let us denote

$$\mathcal{J}_2(X) = \{n \in \mathcal{J}_e(X) : n \text{ verifies point (C.2)}\},$$
$$\mathcal{J}_1(X) = \mathcal{J}_e(X) \setminus \mathcal{J}_2(X).$$

Using this partition of $\mathcal{J}_e(X)$ and the relation (3.7), we can split the contribution of the major arcs as

$$R''_{\mathfrak{M}}(N) = R^{(1)}_{\mathfrak{M}}(N) + R^{(2)}_{\mathfrak{M}}(N), \tag{3.11}$$

with

$$R^{(i)}_{\mathfrak{M}}(N) = \sum_{\substack{1\leq n\leq N\\ n\in\mathcal{J}_i(X)}} R_{\mathfrak{M}}(n)t'_{k,g}(N-n). \tag{3.12}$$

We observe (as done in Remark 3.2) that if $t'_{k,g}(N-n) \neq 0$, then n even if and only if N satisfies (A.C.): our problem is therefore reduced to estimating $R^{(1)}_{\mathfrak{M}}(N)$ and $R^{(2)}_{\mathfrak{M}}(N)$, for N verifying the arithmetic conditions.

Estimation of $R^{(1)}_{\mathfrak{M}}(N)$. By construction, for $n \in \mathcal{J}_1(X)$ holds the estimate $R_{\mathfrak{M}}(n) \ll n\mathfrak{S}(n)$. Thus by (3.12) and the definition of $t'_{k,g}(N-n)$ in (3.1), we have

$$\begin{aligned} R^{(1)}_{\mathfrak{M}}(N) &\ll \sum_{\substack{1\le n\le N\\ n\in\mathcal{J}_1(X)}} n\mathfrak{S}(n)t'_{k,g}(N-n) = \sum_{\substack{1\le n\le N\\ n\in\mathcal{J}_1(X)}} n\mathfrak{S}(n) \sum_{\substack{1\le\nu_1,\dots,\nu_k\le L'\\ g^{\nu_1}+\dots+g^{\nu_k}=N-n}} 1 \\ &= \sum_{\substack{1\le\nu_1,\dots,\nu_k\le L'\\ N-g^{\nu_1}-\dots-g^{\nu_k}\in\mathcal{J}_1(X)}} (N-g^{\nu_1}-\dots-g^{\nu_k})\mathfrak{S}(N-g^{\nu_1}-\dots-g^{\nu_k}) \\ &\le N \sum_{\substack{1\le\nu_1,\dots,\nu_k\le L'\\ N-g^{\nu_1}-\dots-g^{\nu_k}\in\mathcal{J}_1(X)}} \mathfrak{S}(N-g^{\nu_1}-\dots-g^{\nu_k}) \\ &= N \sum_{1\le\nu_1,\dots,\nu_{k-1}\in L'} \sum_{\substack{1\le\nu\le L'\\ \overline{m}-g^{\nu}\in\mathcal{J}_1(X)}} \mathfrak{S}(m'-g^{\nu}), \end{aligned} \tag{3.13}$$

where $m' = m'(N,\nu_1,\dots,\nu_{k-1}) = N-g^{\nu_1}-\dots-g^{\nu_{k-1}}$. It makes sense, since $N-g^{\nu_1}-\dots-g^{\nu_k}$ is even by the conditions (A.C.) on N and $1 \le N-g^{\nu_1}-\dots-g^{\nu_k} \le N$ is true for X sufficiently large.

We now observe that, given $\ell \in \mathcal{J}_e(X)$, by construction $\ell \in \mathcal{J}_1(X)$ is equivalent to the existence of $0 \le i \le K$ (with K defined in point (C.4)) such that $r_i | C'(\eta)\ell$ ($\Leftrightarrow \frac{r_i}{(C'(\eta),r_i)} | \ell$). So (3.13) implies

$$R^{(1)}_{\mathfrak{M}}(N) \ll N \sum_{\substack{1\le\nu_1,\dots,\nu_{k-1}\le L'\\ m'\equiv g \bmod 2}} \sum_{1\le i\le K} \sum_{\substack{1\le\nu\le L'\\ \frac{r_i}{(C'(\eta),r_i)}|(m'-g^{\nu})}} \mathfrak{S}(m'-g^{\nu}).$$

In order to estimate it, we define the quantity

$$A_g(m,r) = \sum_{\substack{1\le\nu\le L'\\ g^{\nu}<m\\ r|(m-g^{\nu})}} \mathfrak{S}(m-g^{\nu}) \tag{3.14}$$

for every $m \in [X/2, X]$ and $r \in \mathbb{N}$. We remark that if $m \not\equiv g \pmod 2$, then $m-g^{\nu}$ odd and so $\mathfrak{S}(m-g^{\nu}) = 0$, for any $1 \le \nu \le L'$, and therefore $A_g(m,r) = 0$. Thus the definition of $A_g(m,r)$ naturally contains the condition $m \equiv g \pmod 2$. Since $m' = N-g^{\nu_1}-\dots-g^{\nu_k} \in [X/2, X]$ for X sufficiently large, we can write

$$\begin{aligned} R^{(1)}_{\mathfrak{M}}(N) &\ll N \sum_{1\le i\le K} \sum_{1\le\nu_1,\dots,\nu_{k-1}\le L'} A_g\left(m', \frac{r_i}{(C'(\eta),r_i)}\right) \\ &\le N \sum_{1\le i\le K} \sum_{1\le\nu_1,\dots,\nu_{k-1}\le L'} A_g\left(m', \frac{r_i^*}{(C'(\eta),r_i^*)}\right), \end{aligned} \tag{3.15}$$

where * denotes the odd square-free part as before. The last step is due to $\frac{r_i^*}{(C'(\eta),r_i^*)} | \frac{r_i}{(C'(\eta),r_i)}$, which implies $A_g\big(m', \frac{r_i}{(C'(\eta),r_i)}\big) \le A_g\big(m', \frac{r_i^*}{(C'(\eta),r_i^*)}\big)$.

To conclude we need an upper bound for $A_g(m', r)$: we will prove in the next Lemma 3.10 that, for an arbitrarily $\omega > 0$, we have $A_g(m, r) \ll \omega L$, for any $m \in [X/2, X]$ such that $m \equiv g \pmod 2$ and r odd positive integer such that $r \gg (\log X)^2$. In our situation, the hypothesis of the lemma are satisfied since

- $m' = N - g^{\nu_1} - \ldots - g^{\nu_{k-1}} \equiv g \pmod 2$ by the arithmetic conditions (A.C.) on N;
- $N - g^{\nu_1} - \ldots - g^{\nu_{k-1}} \leq X$ is trivially true;
- $N - g^{\nu_1} - \ldots - g^{\nu_{k-1}} \geq X/2$ holds since $N \geq 2X/3$ and X sufficiently large;
- $\frac{r_i^*}{(C'(\eta), r_i^*)}$ is an odd positive integer by construction;
- By point (C.3) of the corollary above, we have $r_i^* \gg (\log X)^2$. Since $C'(\eta)$ is a constant depending only on η and not on X, we can therefore suppose, without loss of generality, that $r_i' = \frac{r_i^*}{(C'(\eta), r_i^*)} \gg (\log X)^2$.

Inserting Lemma 3.10 in (3.15) and choosing $\omega = \eta/K$, we obtain

$$R_{\mathfrak{M}}^{(1)}(N) \ll NL'^{k-1}\eta L \ll \eta N L^k.$$

The splitting formula (3.11) therefore implies

$$R''_{\mathfrak{M}}(N) = R_{\mathfrak{M}}^{(2)}(N) + \mathcal{O}\big(\eta N L^k\big).\blacksquare \tag{3.16}$$

Estimation of $R_{\mathfrak{M}}^{(2)}(N)$. By construction, for $n \in \mathcal{J}_2(X)$ holds the asymptotic estimate $|R_{\mathfrak{M}}(n) - n\mathfrak{S}(n)| \leq \eta n \mathfrak{S}(n)$ (and so $R_{\mathfrak{M}}(n) = (1+\mathcal{O}(\eta))n\mathfrak{S}(n)$) with $\eta > 0$ sufficiently small, as at point (C.2). Recalling the definition (3.12) of $R_{\mathfrak{M}}^{(2)}(N)$ and that $\mathcal{J}_2(X) = \mathcal{J}_e(X) \setminus \mathcal{J}_1(X)$, and arguing exactly as at pages 776-777 of Languasco-Pintz-Zaccagnini [LPZ07], we get

$$R_{\mathfrak{M}}^{(2)}(N) = (1 + \mathcal{O}(\eta)) \sum_{\substack{1 \leq n \leq N \\ n \in \mathcal{J}_2(X)}} n\mathfrak{S}(n) t'_{k,g}(N-n) = (1 + \mathcal{O}(\eta))\big(R_{\mathfrak{M}}^{(3)}(N) - R_{\mathfrak{M}}^{(1)}(N)\big),$$

with

$$R_{\mathfrak{M}}^{(3)}(N) = \sum_{\substack{1 \leq n \leq N \\ n \in \mathcal{J}_e(X)}} n\mathfrak{S}(n) t'_{k,g}(N-n).$$

Using the estimation of $R_{\mathfrak{M}}^{(1)}(N)$ just obtained, we see that

$$\begin{aligned} R_{\mathfrak{M}}^{(2)}(N) &= (1 + \mathcal{O}(\eta))R_{\mathfrak{M}}^{(3)}(N) + (1 + \mathcal{O}(\eta))\mathcal{O}\big(\eta N L^k\big) \\ &= (1 + \mathcal{O}(\eta))R_{\mathfrak{M}}^{(3)}(N) + \mathcal{O}\big(\eta N L^k\big), \end{aligned}$$

being η sufficiently small. So (3.16) implies

$$R''_{\mathfrak{M}}(N) = (1 + \mathcal{O}(\eta))R_{\mathfrak{M}}^{(3)}(N) + \mathcal{O}\big(\eta N L^k\big).\blacksquare \tag{3.17}$$

Therefore there is left to study $R_{\mathfrak{M}}^{(3)}(N)$.

Estimation of $R^{(3)}_{\mathfrak{M}}(N)$. We recall that

$$\mathfrak{S}(n) = 2c_0 \prod_{\substack{p|n \\ p>2}} \frac{p-1}{p-2}, \tag{3.18}$$

wherever n is even, and otherwise $\mathfrak{S}(n) = 0$. So, using the definition (3.1) of $t'_{k,g}(n)$, we get

$$R^{(3)}_{\mathfrak{M}}(N) = 2c_0 \sum_{\substack{1\le n\le N \\ n\in\mathcal{J}_e(X)}} \Big(n \prod_{\substack{p|n \\ p>2}} \frac{p-1}{p-2} \sum_{\substack{1\le \nu_1,\dots,\nu_k\le L' \\ g^{\nu_1}+\dots+g^{\nu_k}=N-n}} 1 \Big). \tag{3.19}$$

It is easy to see that, for every $\ell \in \mathbb{N}$

$$\prod_{\substack{p|\ell \\ p>2}} \frac{p-1}{p-2} = \sum_{d|\ell} \mu^2(d) f(d) \tag{3.20}$$

where $\mu(d)$ is the Möbius function (see, *e.g.*, Apostol [Apo76], page 24) and $f(d)$ is defined as

$$f(d) = \prod_{p|d} \frac{1}{p-2}, \tag{3.21}$$

whenever $d > 2$ is odd and $f(d) = 0$ otherwise. In fact, let m be the largest odd divisor of ℓ, eventually ℓ itself, then

$$\prod_{\substack{p|\ell \\ p>2}} \frac{p-1}{p-2} = \prod_{p|m} \Big(1 + \frac{1}{p-2}\Big) = \prod_{p|m} \Big(1 - \mu(p)f(p)\Big).$$

Both $\mu(d)$ and $f(d)$ are multiplicative functions and thus so is $(\mu f)(d)$. Therefore, using a well known result for multiplicative arithmetic functions (that can be found again in [Apo76], page 37), we obtain

$$\prod_{p|m} \Big(1 - \mu(p)f(p)\Big) = \sum_{d|m} \mu^2(d)f(d) = \sum_{d|\ell} \mu^2(d)f(d),$$

since $f(d) = 0$ whenever d even. So

$$\prod_{\substack{p|\ell \\ p>2}} \frac{p-1}{p-2} = \sum_{d|\ell} \mu^2(d)f(d).$$

Inserting the relation (3.20) in (3.19), we obtain

$$R^{(3)}_{\mathfrak{M}}(N) = 2c_0 \sum_{\substack{1\le n\le N \\ n\in\mathcal{J}_e(X)}} \Big(\sum_{d|n} \mu^2(d)f(d) \sum_{\substack{1\le \nu_1,\dots,\nu_k\le L' \\ g^{\nu_1}+\dots+g^{\nu_k}=N-n}} (N - g^{\nu_1} - \dots - g^{\nu_k}) \Big)$$

$$= 2c_0 \sum_{d\leq N} \mu^2(d)f(d) \sum_{\substack{1\leq \nu_1,\dots,\nu_k\leq L' \\ d|(N-g^{\nu_1}-\dots-g^{\nu_k})>0}} (N - g^{\nu_1} - \dots - g^{\nu_k}).$$

$$= 2c_0 N \sum_{d\leq N} \mu^2(d)f(d) \sum_{\substack{1\leq \nu_1,\dots,\nu_k\leq L' \\ d|(N-g^{\nu_1}-\dots-g^{\nu_k})>0}} \big(1 - \frac{g^{\nu_1} + \dots + g^{\nu_k}}{N}\big)$$

$$= (1 + \mathcal{O}_k(\eta))2c_0 N \sum_{d\leq N} \mu^2(d)f(d) \sum_{\substack{1\leq \nu_1,\dots,\nu_k\leq L' \\ d|(N-g^{\nu_1}-\dots-g^{\nu_k})>0}} 1,$$

where we used $N \geq 2X/3$ with X sufficiently large, and $N - g^{\nu_1} - \dots - g^{\nu_k}$ even by the arithmetic conditions (A.C.) on N.

We now split the external sum over d above into $(d, g) = 1$ and $(d, g) > 1$, obtaining

$$R^{(3)}_{\mathfrak{M}}(N) = (1 + \mathcal{O}_k(\eta))2c_0 N \Big(\sum_{\substack{d\leq N \\ (d,g)=1}} \mu^2(d)f(d) + \sum_{\substack{d\leq N \\ (d,g)>1}} \mu^2(d)f(d)\Big) \sum_{\substack{1\leq \nu_1,\dots,\nu_k\leq L' \\ d|(N-g^{\nu_1}-\dots-g^{\nu_k})>0}} 1$$

$$= (1 + \mathcal{O}_k(\eta))[R^{(4)}_{\mathfrak{M}}(N) + R^{(5)}_{\mathfrak{M}}(N)], \tag{3.22}$$

with

$$R^{(4)}_{\mathfrak{M}}(N) = 2c_0 N \sum_{\substack{d\leq N \\ (d,g)=1}} \mu^2(d)f(d) \sum_{\substack{1\leq \nu_1,\dots,\nu_k\leq L' \\ d|(N-g^{\nu_1}-\dots-g^{\nu_k})>0}} 1, \tag{3.23}$$

$$R^{(5)}_{\mathfrak{M}}(N) = 2c_0 N \sum_{\substack{d\leq N \\ (d,g)>1}} \mu^2(d)f(d) \sum_{\substack{1\leq \nu_1,\dots,\nu_k\leq L' \\ d|(N-g^{\nu_1}-\dots-g^{\nu_k})>0}} 1.\blacksquare \tag{3.24}$$

Let us introduce the notation $f = (d, g)$, $\bar{d} = d/f$ and $\bar{g} = g/f$. We remark that the fact that d is odd and square-free implies:

i) f is odd, since d is odd.

ii) $(\bar{d}, f) = 1$, since d is square-free.

iii) $(\bar{g}, \bar{d}) = 1$ by definition of gcd, and so $(\bar{d}, g) = 1$.

Inserting (3.22) in (3.17), we obtain

$$R''_{\mathfrak{M}}(N) = (1 + \mathcal{O}_k(\eta))[R^{(4)}_{\mathfrak{M}}(N) + R^{(5)}_{\mathfrak{M}}(N)] + \mathcal{O}\big(\eta N L^k\big), \tag{3.25}$$

hence there is left to study $R^{(4)}_{\mathfrak{M}}(N)$ and $R^{(5)}_{\mathfrak{M}}(N)$.

Estimation of $R^{(4)}_{\mathfrak{M}}(N)$. Let us define the quantities

$$\xi_g(d) = \min\{\ell \geq 1 : g^\ell \equiv 1 \pmod{d}\}, \tag{3.26}$$

$$j_g(d,k,N) = \sum_{\substack{1\le \nu_1,\dots,\nu_k\le \xi_g(d)\\ d|(N-g^{\nu_1}-\dots-g^{\nu_k})>0}} 1.$$

We remark that $\xi_g(d)$ is well defined since $(d,g)=1$ in $R^{(4)}_{\mathfrak{M}}(N)$. Choosing $\nu_1,\dots,\nu_{k-1}$ arbitrarily, there is at most one $1\le\nu\le\xi_g(d)$ such that $d|(N-g^{\nu_1}-\dots-g^{\nu_{k-1}}-g^{\nu})$. In fact, let us suppose by contradiction that there exists also $\tau\in[1,\xi_g(d)]$ with $\tau\neq\nu$ and $d|(N-g^{\nu_1}-\dots-g^{\nu_{k-1}}-g^{\tau})$. It implies that $g^{\nu}\equiv N-g^{\nu_1}-\dots-g^{\nu_{k-1}}\equiv g^{\tau} \pmod d$ that is $g^{\nu-\tau}\equiv 1 \pmod d$ which is in contradiction with the definition of $\xi_g(d)$. Therefore

$$j_g(d,k,N)\le \xi_g(d)^{k-1}. \tag{3.27}$$

Let us define also the following functions

$$\begin{aligned} S_g(m,d) &= \begin{cases} 1 & \text{if } \exists\ 1\le\nu\le L' \text{ s.t. } d|(m-g^{\nu})\\ 0 & \text{otherwise;}\end{cases} \qquad (3.28)\\ S'_g(m,N) &= \sum_{\substack{d\le N\\(d,g)=1}} \mu^2(d)f(d)S_g(m,d);\\ S''_g(N) &= \max_{1\le m\le N} S'_g(m,N). \end{aligned}$$

We use these functions to split the inner sum in (3.23) and estimate the associated error term for $R^{(4)}_{\mathfrak{M}}(N)$. We remark that we have an error only when, taken $1\le\nu_1,\dots,\nu_{k-1}\le L'$ arbitrarily, there exists an exponent $1\le\nu\le L'$ such that $d|(N-g^{\nu_1}-\dots-g^{\nu_{k-1}}-g^{\nu})$, thus

$$\begin{aligned} \frac{R^{(4)}_{\mathfrak{M}}(N)}{2c_0N} &= \sum_{\substack{d\le N\\(d,g)=1}} \mu^2(d)f(d) \sum_{\substack{1\le\nu_1,\dots,\nu_k\le L'\\ d|(N-g^{\nu_1}-\dots-g^{\nu_k})>0}} 1\\ &= \sum_{\substack{d\le N\\(d,g)=1}} \mu^2(d)f(d)\Big[j_g(d,k,N)\Big(\frac{L'}{\xi_g(d)}\Big)^k\\ &\qquad + \mathcal{O}(1)\sum_{1\le\nu_1,\dots,\nu_{k-1}\le L'} S_g(N-g^{\nu_1}-\dots-g^{\nu_{k-1}},d)\Big]\\ &= \sum_{\substack{d\le N\\(d,g)=1}} \mu^2(d)f(d)j_g(d,k,N)\Big(\frac{L'}{\xi_g(d)}\Big)^k\\ &\qquad + \mathcal{O}(1)\sum_{\substack{d\le N\\(d,g)=1}}\ \sum_{1\le\nu_1,\dots,\nu_{k-1}\le L'} \mu^2(d)f(d)S_g(N-g^{\nu_1}-\dots-g^{\nu_{k-1}},d)\\ &= \sum_{\substack{d\le N\\(d,g)=1}} \mu^2(d)f(d)j_g(d,k,N)\Big(\frac{L'}{\xi_g(d)}\Big)^k\\ &\qquad + \mathcal{O}(1)\sum_{1\le\nu_1,\dots,\nu_{k-1}\le L'} S'_g(N-g^{\nu_1}-\dots-g^{\nu_{k-1}},N) \end{aligned}$$

$$= \sum_{\substack{d\leq N\\(d,g)=1}} \mu^2(d)f(d)j_g(d,k,N)\Big(\frac{L'}{\xi_g(d)}\Big)^k + \mathcal{O}\big(L'^{k-1}S_g''(N)\big)$$

$$= \sum_{\substack{d\leq N\\(d,g)=1}} \mu^2(d)f(d)j_g(d,k,N)\Big(\frac{L'}{\xi_g(d)}\Big)^k + \mathcal{O}\big(L^{k-1}S_g''(N)\big), \tag{3.29}$$

recalling $L' \ll L$ and $1 \leq N - g^{\nu_1} - \ldots - g^{\nu_{k-1}} \leq N$. To conclude, we split the summation in (3.29) into $\Sigma_g^{(1)}(N)$ and $\Sigma_g^{(2)}(N)$ according to whether $\xi_g(d) \leq L^{1/2}$ or $\xi_g(d) > L^{1/2}$. Using the trivial estimation in (3.27) we get

$$\Sigma_g^{(2)}(N) = \sum_{\substack{d\leq N\\(d,g)=1\\ \xi_g(d)>L^{1/2}}} \mu^2(d)f(d)j_g(d,k,N)\Big(\frac{L'}{\xi_g(d)}\Big)^k$$

$$\leq L'^k \sum_{\substack{d\leq N\\(d,g)=1\\ \xi_g(d)>L^{1/2}}} \mu^2(d)\frac{f(d)}{\xi_g(d)} < \eta L^k,$$

for $X > c_1(\eta, g)$, where $c_1(\eta, g)$ is a suitable positive constant. The last step follows from the fact that the series $\sum_{(d,g)=1}\mu^2(d)\frac{f(d)}{\xi_g(d)}$ converges (as we will prove later in Remark 3.7 after Lemma 3.6). Observing that if $\xi_g(d)$ goes to $+\infty$, then also d goes to $+\infty$, we therefore have that our sum is dominated by the tail of a convergent series.

On the other hand we have

$$\Sigma_g^{(1)}(N) = L'^k \sum_{\substack{d\leq N\\(d,g)=1\\ \xi_g(d)\leq L^{1/2}}} \mu^2(d)\frac{f(d)j_g(d,k,N)}{\xi_g(d)^k}$$

$$= L'^k\Big(\boldsymbol{D_g}(k,N) - \sum_{\substack{d>N\\(d,g)=1}} \mu^2(d)\frac{f(d)j_g(d,k,N)}{\xi_g(d)^k} - \sum_{\substack{d\leq N\\(d,g)=1\\ \xi_g(d)>L^{1/2}}} \mu^2(d)\frac{f(d)j_g(d,k,N)}{\xi_g(d)^k}\Big),$$

with

$$\boldsymbol{D_g}(k,N) = \sum_{\substack{d\geq 1\\(d,g)=1}} \mu^2(d)\frac{f(d)j_g(d,k,N)}{\xi_g(d)^k}.$$

By the trivial estimate (3.27), we get $\boldsymbol{D_g}(k,N) \leq \sum_{(d,g)=1}\mu^2(d)\frac{f(d)}{\xi_g(d)}$, which converges again by Lemma 3.6 and Remark 3.7. Moreover

$$\sum_{\substack{d>N\\(d,g)=1}} \mu^2(d)\frac{f(d)j_g(d,k,N)}{\xi_g(d)^k} \ll \eta,$$

for $X > c_2(\eta, g)$, where $c_2(\eta, g)$ is a suitable positive constant, since this summation is the tail of a convergent series. Finally, we can take $X > c_3(\eta, g)$, with $c_3(\eta, g)$ a positive constant, in such a way that $L^{1/2}$ large enough to have

$$\sum_{\substack{d\le N\\(d,g)=1\\ \xi_g(d)>L^{1/2}}} \mu^2(d)\frac{f(d)j_g(d,k,N)}{\xi_g(d)^k} \ll \eta.$$

Collecting all these remarks and since $L' \ll L$, we get

$$\Sigma_g^{(1)}(N) = \boldsymbol{D_g}(k,N)L'^k + \mathcal{O}\big(\eta L^k\big).$$

So, if we define

$$\boldsymbol{D'_g}(k,N) = 2c_0\boldsymbol{D_g}(k,N),$$

then, by (3.29) and the estimations of $\Sigma_g^{(1)}(N)$ and $\Sigma_g^{(2)}(N)$ just obtained, we have

$$\begin{aligned} R_{\mathfrak{M}}^{(4)}(N) &= 2c_0N\big(\boldsymbol{D_g}(k,N)L'^k + \mathcal{O}\big(\eta L^k\big) + \mathcal{O}\big(L^{k-1}S_g''(N)\big)\big)\\ &= \boldsymbol{D'_g}(k,N)NL'^k + \mathcal{O}\big(\eta NL^k\big), \end{aligned} \tag{3.30}$$

where the last step follows from $S_g'(m,N) \ll \log L$ (as we will see later in Remark 3.13 after Lemma 3.10) that trivially implies $S_g''(N) \ll \eta L$.■

Estimation of $R_{\mathfrak{M}}^{(5)}(N)$. By definition, in $R_{\mathfrak{M}}^{(5)}(N)$ we sum over d such that $(d,g) = f > 1$. Recalling the notation $\bar d = d/f$ and that $(\bar d, g) = 1$, let us define the quantities

$$\begin{aligned} &\overline{\xi_g}(d) = \min\{\ell \ge 1 : g^\ell \equiv 1 \pmod{d/(d,g)}\} = \xi_g(\bar d), \\ &\overline{j_g}(d,k,N) = \sum_{\substack{1\le\nu_1,\dots,\nu_k\le\overline{\xi_g}(d)\\ d|(N-g^{\nu_1}-\dots-g^{\nu_k})>0}} 1. \end{aligned} \tag{3.31}$$

We remark that $\overline{\xi_g}(d)$ is well defined and moreover

i) whenever $(d,g) = 1$, then $\bar d = d$ and so $\overline{\xi_g}(d) = \xi_g(d)$; thus $\overline{j_g}(d,k,N) = j_g(d,k,N)$;

ii) if $(d,g) = d$ ($\Leftrightarrow d|g$), then $\bar d = 1$; thus

$$\begin{aligned} &\overline{\xi_g}(d) = \xi_g(1) = \min\{\ell \ge 1 : g^\ell \equiv 1 \pmod 1\} = 1, \\ &\overline{j_g}(d,k,N) = \sum_{\substack{1\le\nu_1,\dots,\nu_k\le 1\\ d|(N-g^{\nu_1}-\dots-g^{\nu_k})>0}} 1 = \begin{cases} 1 & \text{if } d|(N-kg)\ (\Leftrightarrow d|N)\\ 0 & \text{otherwise.}\end{cases} \end{aligned} \tag{3.32}$$

We also have that $d|(N - g^{\nu_1} - \dots - g^{\nu_k})$ trivially implies $\bar d|(N - g^{\nu_1} - \dots - g^{\nu_k})$. Thus

$$\overline{j_g}(d,k,N) = \sum_{\substack{1\le\nu_1,\dots,\nu_k\le\overline{\xi_g}(d)\\ d|(N-g^{\nu_1}-\dots-g^{\nu_k})>0}} 1 \le \sum_{\substack{1\le\nu_1,\dots,\nu_k\le\overline{\xi_g}(d)\\ \bar d|(N-g^{\nu_1}-\dots-g^{\nu_k})>0}} 1 = j_g(\bar d,k,N).$$

So it's easy to see that, (3.27) implies

$$\overline{j_g}(d,k,N) \leq j_g(\bar{d},k,N) \leq \xi_g(\bar{d})^{k-1} = \overline{\xi_g}(d)^{k-1}. \tag{3.33}$$

This inequality can be proven also by a direct argument: choosing $1 \leq \nu_1, \dots, \nu_{k-1} \leq \overline{\xi_g}(d)$ arbitrarily, there is at most one $1 \leq \nu \leq \overline{\xi_g}(d)$ such that $d|(N - g^{\nu_1} - \dots - g^{\nu_{k-1}} - g^{\nu})$. In fact, let us suppose by contradiction that there exists also $\tau \in [1, \overline{\xi_g}(d)]$ with $\tau \neq \nu$ and $d|(N - g^{\nu_1} - \dots - g^{\nu_{k-1}} - g^{\tau})$. It implies that $g^{\nu} \equiv N - g^{\nu_1} - \dots - g^{\nu_{k-1}} \equiv g^{\tau} \pmod{d}$. Thus also $g^{\nu} \equiv g^{\tau} \pmod{\bar{d}}$, that is $g^{\nu-\tau} \equiv 1 \pmod{\bar{d}}$ which is in contradiction with the definition of $\overline{\xi_g}(d)$ (we remark that g is a unit modulo $\bar{d}$, since $(\bar{d}, g) = 1$, and so we can invert g^{τ}) . Therefore $\overline{j_g}(d,k,N) \leq \overline{\xi_g}(d)^{k-1}$.

Let us define also the following functions

$$\overline{S'_g}(m,N) = \sum_{\substack{d \leq N \\ (d,g)>1}} \mu^2(d) f(d) S_g(m,d); \tag{3.34}$$

$$\overline{S''_g}(N) = \max_{1 \leq m \leq N} \overline{S'_g}(m,N).$$

Remark 3.5. *By definition, $S_g(m,d) = 1$ if there exists an exponent $1 \leq \nu \leq L'$ such that $d|(m - g^{\nu})$ and $S_g(m,d) = 0$ otherwise. So it is easy to see that, if $(d,g) = d$ ($\Leftrightarrow d|g$), then $S_g(m,d) = 1$ whenever $d|m$ and $S_g(m,d) = 0$ otherwise.*

We now observe that, given $1 \leq \nu \leq \overline{\xi_g}(d)$ and using the notation $d = \bar{d}f$ with $f = (g,d)$, then

$$d|(m - g^{\nu}) \Leftrightarrow \begin{cases} \bar{d}|(m - g^{\nu}), \\ f|(m - g^{\nu}) (\Leftrightarrow f|m), \end{cases}$$

because of $(\bar{d}, f) = 1$ and $f|g$. So if there exists $1 \leq \nu \leq \overline{\xi_g}(d)$ such that $d|(m - g^{\nu})$, then $d|(m - g^{\nu + \overline{\xi_g}(d)})$ since both

$$\begin{cases} m - g^{\nu+\overline{\xi_g}(d)} \equiv m \pmod{f} \equiv 0 \pmod{f} \\ m - g^{\nu+\overline{\xi_g}(d)} \equiv m - g^{\nu} \pmod{\bar{d}} \equiv 0 \pmod{\bar{d}}. \end{cases}$$

We now split the external sum in the definition (3.24) of $R^{(5)}_{\mathfrak{M}}(N)$, according to $1 < (d,g) < d$ and $(d,g) = d$. We also split inner sum and estimate the associated error term as did for $R^{(4)}_{\mathfrak{M}}(N)$, observing that we have an error only when, taken $1 \leq \nu_1, \dots, \nu_{k-1} \leq L'$ arbitrarily, there exists an exponent $1 \leq \nu \leq L'$ such that $d|(N - g^{\nu_1} - \dots - g^{\nu_{k-1}} - g^{\nu})$. Thus

$$\begin{aligned} \frac{R^{(5)}_{\mathfrak{M}}(N)}{2c_0 N} &= \sum_{\substack{d \leq N \\ 1<(d,g)<d}} \mu^2(d) f(d) \sum_{\substack{1 \leq \nu_1, \dots, \nu_k \leq L' \\ d|(N - g^{\nu_1} - \dots - g^{\nu_k}) > 0}} 1 + \sum_{\substack{d \leq N \\ d|g}} \mu^2(d) f(d) \sum_{\substack{1 \leq \nu_1, \dots, \nu_k \leq L' \\ d|N}} 1 \\ &= \sum_{\substack{d \leq N \\ 1<(d,g)<d}} \mu^2(d) f(d) \Big[\overline{j_g}(d,k,N) \Big(\frac{L'}{\overline{\xi_g}(d)} \Big)^k \end{aligned}$$

$$+\mathcal{O}(1)\sum_{1\le\nu_1,\dots,\nu_{k-1}\le L'} S_g(N-g^{\nu_1}-\dots-g^{\nu_{k-1}},d)\Big]$$
$$+L'^k\sum_{\substack{d\le N\\ d|(g,N)}}\mu^2(d)f(d). \tag{3.35}$$

Recalling the relations in (3.32), we get

$$L'^k\sum_{\substack{d\le N\\ d|(g,N)}}\mu^2(d)f(d)=L'^k\sum_{\substack{d\le N\\ d|(g,N)}}\mu^2(d)f(d)\frac{\overline{j_g}(d,k,N)}{\overline{\xi_g}(d)}.$$

Thus, if we introduce the notation

$$\overline{T'_g}(m,N)=\overline{S'_g}(m,N)-\sum_{\substack{d\le N\\ d|(g,m)}}\mu^2(d)f(d)=\sum_{\substack{d\le N\\ 1<(d,g)<d}}\mu^2(d)f(d)S_g(m,d),$$
$$\overline{T''_g}(N)=\max_{1\le m\le N}\overline{T'_g}(m,N),$$

we can rearrange (3.35) as follows

$$\begin{aligned}\frac{R^{(5)}_{\mathfrak{M}}(N)}{2c_0N}&=\sum_{\substack{d\le N\\ (d,g)>1}}\mu^2(d)f(d)\overline{j_g}(d,k,N)\Big(\frac{L'}{\overline{\xi_g}(d)}\Big)^k\\&\quad+\mathcal{O}(1)\sum_{\substack{d\le N\\ (d,g)>1}}\sum_{1\le\nu_1,\dots,\nu_{k-1}\le L'}\mu^2(d)f(d)S_g(N-g^{\nu_1}-\dots-g^{\nu_{k-1}},d)\\&=\sum_{\substack{d\le N\\ (d,g)>1}}\mu^2(d)f(d)\overline{j_g}(d,k,N)\Big(\frac{L'}{\overline{\xi_g}(d)}\Big)^k\\&\quad+\mathcal{O}(1)\sum_{1\le\nu_1,\dots,\nu_{k-1}\le L'}\overline{T'_g}(N-g^{\nu_1}-\dots-g^{\nu_{k-1}},N)\\&=\sum_{\substack{d\le N\\ (d,g)>1}}\mu^2(d)f(d)\overline{j_g}(d,k,N)\Big(\frac{L'}{\overline{\xi_g}(d)}\Big)^k+\mathcal{O}\big((L')^{k-1}\overline{T''_g}(N)\big)\\&=\sum_{\substack{d\le N\\ (d,g)>1}}\mu^2(d)f(d)\overline{j_g}(d,k,N)\Big(\frac{L'}{\overline{\xi_g}(d)}\Big)^k+\mathcal{O}\big(L^{k-1}\overline{T''_g}(N)\big),\end{aligned}\tag{3.36}$$

recalling $L'\ll L$ and $1\le N-g^{\nu_1}-\dots-g^{\nu_{k-1}}\le N$. To conclude, we split the summation above into $\overline{\Sigma^{(1)}_g}(N)$ and $\overline{\Sigma^{(2)}_g}(N)$ according to whether $\overline{\xi_g}(d)\le L^{1/2}$ or $\overline{\xi_g}(d)>L^{1/2}$ and we argue exactly as did for $R^{(4)}_{\mathfrak{M}}(N)$. That is, using the trivial estimate in (3.33) we get

$$\overline{\Sigma^{(2)}_g}(N)=\sum_{\substack{d\le N\\ (d,g)>1\\ \overline{\xi_g}(d)>L^{1/2}}}\mu^2(d)f(d)\overline{j_g}(d,k,N)\Big(\frac{L'}{\overline{\xi_g}(d)}\Big)^k$$

$$\leq L'^k \sum_{\substack{d\leq N\\(d,g)>1\\ \overline{\xi_g}(d)>L^{1/2}}} \mu^2(d)\frac{f(d)}{\overline{\xi_g}(d)} < \eta L^k,$$

for $X > c_1(\eta, g)$, where $c_1(\eta, g)$ is a suitable positive constant. The last step follows from the fact that the series $\sum_{(d,g)>1}\mu^2(d)\frac{f(d)}{\overline{\xi_g}(d)}$ converges (as we will prove later in Remark 3.7 after Lemma 3.6). Observing that if $\xi_g(d)$ goes to $+\infty$, then also d goes to $+\infty$, we therefore have that our sum is dominated by the tail of a convergent series.

On the other hand

$$\begin{aligned}\overline{\Sigma_g^{(1)}}(N) &= L'^k \sum_{\substack{d\leq N\\(d,g)>1\\ \overline{\xi_g}(d)\leq L^{1/2}}} \mu^2(d)\frac{f(d)\overline{j_g}(d,k,N)}{\overline{\xi_g}(d)^k}\\ &= L'^k\Big(\overline{\boldsymbol{D_g}}(k,N) - \sum_{\substack{d>N\\(d,g)>1}} \mu^2(d)\frac{f(d)\overline{j_g}(d,k,N)}{\overline{\xi_g}(d)^k} - \sum_{\substack{d\leq N\\(d,g)>1\\ \overline{\xi_g}(d)>L^{1/2}}} \mu^2(d)\frac{f(d)\overline{j_g}(d,k,N)}{\overline{\xi_g}(d)^k}\Big),\end{aligned}$$

with

$$\overline{\boldsymbol{D_g}}(k,N) = \sum_{\substack{d\geq 1\\(d,g)>1}} \mu^2(d)\frac{f(d)\overline{j_g}(d,k,N)}{\overline{\xi_g}(d)^k}.$$

By the trivial estimate in (3.33), we get $\overline{\boldsymbol{D_g}}(k,N) \leq \sum_{(d,g)>1}\mu^2(d)\frac{f(d)}{\overline{\xi_g}(d)}$, which converges again by Remark 3.7. Moreover

$$\sum_{\substack{d>N\\(d,g)>1}} \mu^2(d)\frac{f(d)\overline{j_g}(d,k,N)}{\overline{\xi_g}(d)^k} \ll \eta,$$

for $X > c_2(\eta, g)$, where $c_2(\eta, g)$ is a suitable positive constant, since this summation is the tail of a convergent series. Finally, we can take $X > c_3(\eta, g)$, with $c_3(\eta, g)$ a positive constant, in such a way that $L^{1/2}$ is large enough to have

$$\sum_{\substack{d\leq N\\(d,g)>1\\ \overline{\xi_g}(d)>L^{1/2}}} \mu^2(d)\frac{f(d)\overline{j_g}(d,k,N)}{\overline{\xi_g}(d)^k} \ll \eta.$$

Collecting all these remarks and since $L' \ll L$, we get

$$\overline{\Sigma_g^{(1)}}(N) = \overline{\boldsymbol{D_g}}(k,N)L'^k + \mathcal{O}\big(\eta L^k\big).$$

So, if we define

$$\overline{\boldsymbol{D'_g}}(k,N) = 2c_0\overline{\boldsymbol{D_g}}(k,N),$$

then, by (3.36) and the estimates of $\overline{\Sigma_g^{(1)}}(N)$ and $\overline{\Sigma_g^{(2)}}(N)$ just obtained, we have

$$\begin{aligned} R_{\mathfrak{M}}^{(5)}(N) &= 2c_0 N\big(\overline{\boldsymbol{D_g}}(k,N)L'^k + \mathcal{O}\big(\eta L^k\big) + \mathcal{O}\big(L^{k-1}\overline{T_g''}(N)\big)\big) \\ &= \overline{\boldsymbol{D_g'}}(k,N)NL'^k + \mathcal{O}\big(\eta NL^k\big), \end{aligned} \tag{3.37}$$

where the last step follows from $\overline{T_g'}(m,N) \ll \log L$ (as we will see later in Lemma 3.10 and Remark 3.13) that trivially implies $\overline{T_g''}(N) \ll \eta L$.■

Therefore, by the estimates of $R_{\mathfrak{M}}^{(4)}(N)$ and $R_{\mathfrak{M}}^{(5)}(n)$ in (3.30) and (3.37) resp., and by (3.25) we get

$$\begin{aligned} R_{\mathfrak{M}}''(N) &= (1+\mathcal{O}(\eta))\big[\boldsymbol{D_g'}(k,N)NL^k + \overline{\boldsymbol{D_g'}}(k,N)NL^k + \mathcal{O}\big(\eta NL^k\big)\big] + \mathcal{O}\big(\eta NL^k\big) \\ &= (1+\mathcal{O}(\eta))[\boldsymbol{D_g'}(k,N) + \overline{\boldsymbol{D_g'}}(k,N)]NL^k, \end{aligned} \tag{3.38}$$

and so there is left to study, for any $N \in \mathcal{J}(X)$ satisfying (A.C.), the following constant

$$\begin{aligned} \boldsymbol{C_g} = \boldsymbol{C_g}(k,N) &= \boldsymbol{D_g'}(k,N) + \overline{\boldsymbol{D_g'}}(k,N) = 2c_0\big[\boldsymbol{D_g}(k,N) + \overline{\boldsymbol{D_g}}(k,N)\big] \\ &= 2c_0\Big[\sum_{\substack{d\geq 1\\(d,g)=1}} \mu^2(d)\frac{f(d)j_g(d,k,N)}{\xi_g(d)^k} + \sum_{\substack{d\geq 1\\(d,g)>1}} \mu^2(d)\frac{f(d)\overline{j_g}(d,k,N)}{\overline{\xi_g}(d)^k}\Big] \\ &= 2c_0\sum_{d\geq 1}\mu^2(d)\frac{f(d)\overline{j_g}(d,k,N)}{\overline{\xi_g}(d)^k}, \end{aligned}$$

recalling that for $(d,g)=1$, $\overline{\xi_g}(d) = \xi_g(d)$ and $\overline{j_g}(d,k,N) = j_g(d,k,N)$, as already observed.

3.6 Proof of the theorem

Collecting the splitting formula (3.6) for $R_{k,g}''(N)$, Lemma 3.4 on minor arcs and the relation (3.38) on major arcs, we obtain

$$\begin{aligned} R_{k,g}''(N) &= (1+\mathcal{O}_k(\eta))\boldsymbol{C_g}NL^k + \mathcal{O}\big(NL^{k-1/2}\big) + \mathcal{O}\big(NL^{k-1/2}\log\log N\big) \\ &= (1+\mathcal{O}_k(\eta))\boldsymbol{C_g}NL^k \end{aligned}$$

which implies

$$|R_{k,g}''(N) - \boldsymbol{C_g}NL^k| \leq \eta NL^k$$

for every $N \in \mathcal{J}(X)$ satisfying (A.C.) with at most $\mathcal{O}\big(X^{3/5}(\log X)^9 L\big) = \mathcal{O}_g\big(X^{3/5}(\log X)^{10}\big)$ exceptions.

Extension to $[1,X]$. We can easily extend this result to every $N \in [1,X]$ satisfying (A.C.), by using a dyadic argument. That is, if we denote by $E(x,y)$ the subset of $[x,y]$ made up of those integers N that satisfies (A.C.) and such that $|R_{k,g}''(N) - \boldsymbol{C_g}NL^k| > \eta NL^k$, then

$$|E(1,X)| = \Big|\bigsqcup_{n=0}^{\mathcal{O}(\log X)} E\Big(X\Big(\frac{2}{3}\Big)^{n+1}, X\Big(\frac{2}{3}\Big)^{n}\Big)\Big|$$

$$\begin{aligned}
&\ll_g \sum_{n=0}^{\mathcal{O}(\log X)} \left(X\Big(\frac{2}{3}\Big)^n\right)^{3/5} \left(\log\left(X\Big(\frac{2}{3}\Big)^n\right)\right)^{10} \\
&\ll X^{3/5}(\log X)^{10} \sum_{n=0}^{\mathcal{O}(\log X)} \Big(\frac{2}{3}\Big)^{3n/5} \\
&\ll X^{3/5}(\log X)^{10},
\end{aligned}$$

where we used the convergence of $\sum_{n=0}^{\mathcal{O}(\log X)}((2/3)^{3/5})^n$ and we applied our result to $N \in \mathcal{J}(X(2/3)^n)$ every time.

Thus Theorem B holds, once we prove that the series defining $\boldsymbol{C_g}$ converges (we will actually prove that $\boldsymbol{C_g} \leq 2c_0e^{\gamma} \cdot 0.7574\big(\log\log g + 3 + \log 2 + \pi/2) + \mathcal{O}((\log\log g)^{-1})$, where the implicit constant is absolute).

3.7 Lemmas

Our first result is a variation of a result of Romanov [Rom34] (see Pintz-Ruzsa [PR03], page 188, for an easier proof) and it is along the lines of Theorem 1.1 in Murty-Rosen-Silverman [MRS96].

Lemma 3.6. *Let $f(d)$ and $\overline{\xi_g}(d)$ be as defined in* (3.21) *and* (3.31) *resp., then the series*

$$\sum_{d\geq 1} \mu^2(d) \frac{f(d)}{\overline{\xi_g}(d)}$$

is convergent.

Proof. Trivially

$$\sum_{d\geq 1} \mu^2(d) \frac{f(d)}{\overline{\xi_g}(d)} = \sum_{n\geq 1} \frac{1}{n} \sum_{\overline{\xi_g}(d)=n} \mu^2(d) f(d).$$

Let us define the functions

$$\begin{aligned}
F_g(n) &= \sum_{\overline{\xi_g}(d)=n} \mu^2(d) f(d), \\
F_g'(x) &= \sum_{n\leq x} F_g(n),
\end{aligned}$$

therefore the series in the statement of the lemma is

$$F_g'' = \sum_{n\geq 1} \frac{F_g(n)}{n}.$$

Following [MRS96], our aim is to use Abel's identity (see Result (R.17)) to show that F_g'' converges. More precisely, Abel's identity implies

$$\begin{aligned} F_g'' &= \lim_{y\to\infty}\sum_{n=1}^{y}\frac{F_g(n)}{n} = \lim_{y\to\infty}\Big(\frac{\sum_{n=1}^{y}F_g(n)}{y} - \frac{\sum_{n=1}^{1}F_g(n)}{1} + \int_1^y \frac{\sum_{n=1}^{t}F_g(n)}{t^2}\mathrm{d}t\Big)\\ &= \lim_{y\to\infty}\frac{F_g'(y)}{y} + \lim_{y\to\infty}\int_1^y \frac{F_g'(t)}{t^2}\mathrm{d}t. \end{aligned}$$

By the next Corollary 3.9, $F_g'(x) < \mathcal{C}e^{\gamma}(\log(x^2+1)+\log\log g+3)+\mathcal{O}((\log\log g)^{-1})$ where $\mathcal{C} < 0.7574$ is an absolute constant and $\boldsymbol{\gamma}$ is the Euler constant. Thus

$$\begin{aligned} F_g'' &< \lim_{y\to\infty}\frac{\mathcal{C}e^{\gamma}(\log(y^2+1)+\log\log g+3)+\mathcal{O}((\log\log g)^{-1})}{y}\\ &\quad + \lim_{y\to\infty}\int_1^y \frac{\mathcal{C}e^{\gamma}(\log(t^2+1)+\log\log g+3)+\mathcal{O}((\log\log g)^{-1})}{t^2}\mathrm{d}t\\ &= \mathcal{C}e^{\gamma}\lim_{y\to\infty}\int_1^y\frac{\log(t^2+1)}{t^2}\mathrm{d}t + \Big(\mathcal{C}e^{\gamma}(\log\log g+3)+\mathcal{O}\Big(\frac{1}{\log\log g}\Big)\Big)\lim_{y\to\infty}\int_1^y\frac{\mathrm{d}t}{t^2}\\ &= \mathcal{C}e^{\gamma}\Big(\log 2+\frac{\pi}{2}+\log\log g+3\Big)+\mathcal{O}\Big(\frac{1}{\log\log g}\Big). \end{aligned}$$

The last equality follows by partial integration, since we have

$$\begin{aligned} \int_1^y\frac{\log(t^2+1)}{t^2}\mathrm{d}t &= \Big(-\frac{1}{t}\log(t^2+1)\Big)\Big|_1^y + 2\int_1^y\frac{\mathrm{d}t}{t^2+1}\\ &= -\frac{\log(y^2+1)}{y}+\log 2+2\arctan y-\frac{\pi}{2}. \end{aligned}$$

Thus the needed limit is

$$\lim_{y\to\infty}\int_1^y\frac{\log(t^2+1)}{t^2}\mathrm{d}t = \log 2+2\cdot\frac{\pi}{2}-\frac{\pi}{2} = \frac{\pi}{2}+\log 2,$$

whose value is ≈ 2.2639435; so, for every fixed g, $F_g'' = \sum_{d\geq 1}\mu^2(d)\frac{f(d)}{\overline{\xi_g}(d)}$ converges. This proves Lemma 3.6. □

Remark 3.7. *We recall that the sum in* $R_{\mathfrak{M}}^{(5)}(N)$ *is*

$$\sum_{(d,g)>1}\mu^2(d)\frac{f(d)}{\overline{\xi_g}(d)}.$$

We have already observed that whenever $(d,g)=1$, $\overline{\xi_g}(d)=\xi_g(d)$*; so the sum in* $R_{\mathfrak{M}}^{(4)}(N)$ *is*

$$\sum_{(d,g)=1}\mu^2(d)\frac{f(d)}{\xi_g(d)} = \sum_{(d,g)=1}\mu^2(d)\frac{f(d)}{\overline{\xi_g}(d)}.$$

Since we can write

$$\sum_{d\geq 1}\mu^2(d)\frac{f(d)}{\overline{\xi_g}(d)} = \sum_{\substack{d\geq 1\\(d,g)=1}}\mu^2(d)\frac{f(d)}{\overline{\xi_g}(d)} + \sum_{\substack{d\geq 1\\(d,g)>1}}\mu^2(d)\frac{f(d)}{\overline{\xi_g}(d)},$$

then, proving that the left-hand side converges, we actually prove that both the series at the right-hand side converge, since their summands are non-negative.

Lemma 3.8. *Let $f(d)$ be as in definition* (3.21) *and $\mu(d)$ the Möbius function, then for every positive integers m*

$$\sum_{d|m}\mu^2(d)f(d) < \mathcal{C}e^{\gamma}(\log\log m + 3) + \mathcal{O}\bigg(\frac{1}{\log\log m}\bigg),$$

where $\mathcal{C} < 0.7574$ is an absolute constant and $\boldsymbol{\gamma} \approx 0.577216$ is the Euler constant.

Proof. We follow the argument in Corollary 2.3 of Murty-Rosen-Silverman [MRS96] which is a refinement of Erdős-Turán [Erd35] technique. Using (3.20) and splitting the product into "large primes" and "small primes". So we obtain

$$\sum_{d|m}\mu^2(d)f(d) = \prod_{\substack{p|m\\p>2}}\frac{p-1}{p-2} = \mathcal{P}_{\mathcal{L}}(m)\cdot\mathcal{P}_{\mathcal{S}}(m) \tag{3.39}$$

with

$$\mathcal{P}_{\mathcal{L}}(m) = \prod_{\substack{p|m\\p>\log m}}\frac{p-1}{p-2} \qquad \text{and} \qquad \mathcal{P}_{\mathcal{S}}(m) = \prod_{\substack{p|m\\2<p\leq\log m}}\frac{p-1}{p-2}.$$

Estimation of $\mathcal{P}_{\mathcal{L}}(m)$. Let us define the quantity $\omega^+(m) = |\{p\in\mathfrak{P} : p|m;\ p>\log m\}|$, then

$$m \geq \prod_{p|m}p \geq \prod_{\substack{p|m\\p>\log m}}p > \prod_{\substack{p|m\\p>\log m}}\log m = (\log m)^{\omega^+(m)}$$

and taking the logarithms of both sides

$$\omega^+(m) < \frac{\log m}{\log\log m}.$$

Therefore

$$\begin{aligned}\mathcal{P}_{\mathcal{L}}(m) &= \prod_{\substack{p|m\\p>\log m}}\Big(1+\frac{1}{p-2}\Big) < \prod_{\substack{p|m\\p>\log m}}\Big(1+\frac{1}{\log m-2}\Big)\\ &= \Big(1+\frac{1}{\log m-2}\Big)^{\omega^+(m)} < \Big(1+\frac{1}{\log m-2}\Big)^{\frac{\log m}{\log\log m}}\end{aligned}$$

$$= e^{\frac{\log m}{\log\log m}\log(1+\frac{1}{\log m-2})} \le e^{\frac{\log m}{\log\log m}\frac{1}{\log m-2}} \le e^{\frac{3}{\log\log m}}$$
$$= 1 + \frac{3}{\log\log m} + \mathcal{O}\bigg(\frac{1}{(\log\log m)^2}\bigg).$$

Here we used $\log(1+x) \le x$ for every $x > -1$, the Taylor expansion of the exp function and $\frac{\log m}{\log m-2} \le 3$ (since we can assume $\log m \ge 3$, because otherwise the product in (3.39) is empty). We remark that the implicit constant in the error term above can be effectively estimate (*e.g.*, it is easy to prove that it is $\le e^{32}/2$).■

Estimation of $\mathcal{P}_{\mathcal{S}}(m)$. Recalling the definition of the Riemann ζ-function and the related Euler product (see, *e.g.*, Apostol [Apo76], §11.5), we get

$$\begin{aligned}\mathcal{P}_{\mathcal{S}}(m) &= \prod_{\substack{p\mid m\\ 2<p\le\log m}}\Big(1+\frac{1}{p-2}\Big) \le \prod_{2<p\le\log m}\Big(1+\frac{1}{p-2}\Big)\\ &= \frac{1}{\zeta(2)}\prod_{2<p\le\log m}\Big(1+\frac{1}{p-2}\Big)\prod_{p}\Big(1-\frac{1}{p^2}\Big)^{-1}\\ &= \frac{4}{3}\frac{1}{\zeta(2)}\prod_{2<p\le\log m}\Big[\Big(1+\frac{1}{p-2}\Big)\Big(1-\frac{1}{p^2}\Big)^{-1}\Big]\prod_{p>\log m}\Big(1-\frac{1}{p^2}\Big)^{-1}. \qquad (3.40)\end{aligned}$$

Since $\frac{1}{1-1/p^2}$ is the sum of the geometric series with common ratio $1/p^2$, we have

$$\begin{aligned}\prod_{p>\log m}\Big(1-\frac{1}{p^2}\Big)^{-1} &= \prod_{p>\log m}\sum_{n\ge0}\Big(\frac{1}{p^2}\Big)^n = \prod_{p>\log m}\Big(1+\sum_{n\ge1}\frac{1}{p^{2n}}\Big)\\ &= 1+\mathcal{O}\bigg(\sum_{n>\log m}\frac{1}{n^2}\bigg) = 1+\mathcal{O}\bigg(\int_{\log m}^{+\infty}\frac{\mathrm{d}t}{t^2}\bigg)\\ &= 1+\mathcal{O}\bigg(\frac{1}{\log m}\bigg). \qquad (3.41)\end{aligned}$$

We can deal with the first product in (3.40) by splitting the difference of two squares as follows

$$\begin{aligned}\prod_{2<p\le\log m}\Big[\Big(1+\frac{1}{p-2}\Big)\Big(1-\frac{1}{p^2}\Big)^{-1}\Big] &= \prod_{2<p\le\log m}\Big[\Big(1+\frac{1}{p-2}\Big)\Big(1-\frac{1}{p}\Big)^{-1}\Big(1+\frac{1}{p}\Big)^{-1}\Big]\\ &= \prod_{2<p\le\log m}\frac{(p-1)p}{(p-2)(p+1)}\prod_{2<p\le\log m}\Big(1-\frac{1}{p}\Big)^{-1}\\ &= \frac{1}{2}\prod_{2<p\le\log m}\frac{(p-1)p}{(p-2)(p+1)}\prod_{p\le\log m}\Big(1-\frac{1}{p}\Big)^{-1}. \qquad (3.42)\end{aligned}$$

We now define

$$V(m) = \prod_{2<p\le\log m}\frac{(p-1)p}{(p-2)(p+1)},$$

and we estimate $\prod_{p\leq \log m}\left(1-\frac{1}{p}\right)^{-1}$ using Vasil'kovskaja's theorem in Result (R.18) (which is a sharper version of Mertens' theorem), that is:

$$\prod_{p\leq x}\left(1-\frac{1}{p}\right)^{-1} = e^{\gamma}\log x\left(1+\mathcal{O}\big(L(x)^{-c}\big)\right), \tag{3.43}$$

with $L(x)=\exp((\log x)^{3/5}(\log\log x)^{-1/5})$ and $c>0$ absolute constant.

So, by (3.40)-(3.43), we finally have

$$\begin{aligned}\mathcal{P}_{\mathcal{S}}(m) &\leq \frac{2}{3}\frac{V(m)}{\zeta(2)}e^{\gamma}\log\log m\left(1+\mathcal{O}\big(L(\log m)^{-c}\big)\right)\left(1+\mathcal{O}\left(\frac{1}{\log m}\right)\right)\\ &\leq \frac{2}{3}\frac{V(m)}{\zeta(2)}e^{\gamma}\log\log m\left(1+\mathcal{O}\big(L(\log m)^{-c}\big)\right),\end{aligned}$$

since $V(m)$ is dominated by a convergent product (see the next Lemma 3.14) and

$$\frac{1}{\log m}=\exp(-\log\log m)\ll \exp\Big(-c\frac{(\log\log m)^{3/5}}{(\log\log\log m)^{1/5}}\Big)=L(\log m)^{-c}.\blacksquare$$

We now put the estimates of $\mathcal{P}_{\mathcal{L}}(m)$ and $\mathcal{P}_{\mathcal{S}}(m)$, into (3.39). Observing that

$$\frac{1}{(\log\log m)^2}=\exp(-2\log\log\log m)\gg \exp\Big(-c\frac{(\log\log m)^{3/5}}{(\log\log\log m)^{1/5}}\Big)=L(\log m)^{-c},$$

and recalling again that $V(m)$ is dominated by a convergent product, we obtain

$$\begin{aligned}\sum_{d|m}\mu^2(d)f(d) &< \frac{2}{3}\frac{V(m)}{\zeta(2)}e^{\gamma}\log\log m\left(1+\frac{3}{\log\log m}+\mathcal{O}\left(\frac{1}{(\log\log m)^2}\right)\right)\\ &= \frac{2}{3}\frac{V(m)}{\zeta(2)}e^{\gamma}(\log\log m+3)+\mathcal{O}\left(\frac{1}{\log\log m}\right).\end{aligned}$$

Since $V(m)\leq 1.8687$ by the next Lemma 3.14, then

$$\mathcal{C}=\frac{2}{3}\frac{V(m)}{\zeta(2)}<\frac{2\cdot 1.8687}{3\pi^2/6}=\frac{7.4748}{\pi^2}<0.7574$$

and this proves the lemma. □

We remark that a direct application of the Murty-Rosen-Silverman strategy (see [MRS96] Lemma 1.2 for general setting and Remark at page 378 for sums over odd square-free integers) gives $\sum_{d|m}\mu^2(d)f(d)\leq \frac{1}{\zeta(2)}e^{\gamma}\log\log m+\mathcal{O}(1)$ and hence our Lemma 3.8 can be considered a sharper version of this result. We also remark that our Lemma 3.6 can be extended to $\sum_{d\geq 1}\mu^2(d)\frac{f(d)}{\xi_g(d)^{\epsilon}}$ for any $\epsilon>0$, as in Theorem 1.1 of [MRS96].

From Lemma 3.8, we get the following corollary.

Corollary 3.9. *Let $f(d)$ and $\overline{\xi_g}(d)$ be as defined in* (3.21) *and* (3.31) *resp., and let $\mu(d)$ be the Möbius function, then*

$$\sum_{\overline{\xi_g}(d)\le x} \mu^2(d)f(d) < \mathcal{C}e^{\gamma}(\log(x^2+1)+\log\log g+3)+\mathcal{O}\bigg(\frac{1}{\log\log g}\bigg),$$

for every $x \ge 1$, where $\mathcal{C} < 0.7574$ is an absolute constant and $\boldsymbol{\gamma} \approx 0.577216$ is the Euler constant.

Proof. By definition, $\overline{\xi_g}(d) \le x$ implies that there exists $\ell \le x$ such that $\bar{d}|(g^\ell - 1)$, with $\bar{d} = d/(d,g)$, eventually d itself or 1. Thus $\bar{d}|p_g(x)$, with

$$p_g(x) = \prod_{\ell\le x}(g^\ell - 1).$$

In order to apply Lemma 3.8, we need a divisibility condition for d, instead of for $\bar{d}$. So we argue as follow: using the notation $f = (d,g)$, $g = \bar{g}f$ and $d = \bar{d}f$, we have that $\bar{d}|p_g(x)$ implies the existence of an integer $a \ge 1$ such that $p_g(x) = a\bar{d}$. But

$$p_g(x) = a\bar{d} \quad \Rightarrow \quad p_g(x)f = a\bar{d}f \quad \Rightarrow \quad p_g(x)f\bar{g} = a\bar{d}f\bar{g} \quad (\Leftrightarrow p_g(x)g = ad\bar{g}).$$

Therefore $\overline{\xi_g}(d) \le x$ implies $d|p_g(x)g$ and so, by Lemma 3.8

$$\sum_{\overline{\xi_g}(d)\le x} \mu^2(d)f(d) \le \sum_{d|p_g(x)g} \mu^2(d)f(d) < \mathcal{C}e^{\gamma}(\log\log(p_g(x)g)+3)+\mathcal{O}\bigg(\frac{1}{\log\log(p_g(x)g)}\bigg).$$

To conclude, we observe that $p_g(x)g \le (\prod_{\ell\le x} g^\ell)g = g^{x(x+1)/2}g \le g^{x^2+1}$ since $x \ge 1$. Thus, taking twice the logarithms of both sides, we get

$$\log\log(p_g(x)g) \le \log(x^2+1)+\log\log g.$$

We also have $p_g(x)g \ge g$ since $p_g(x) = \prod_{\le n\le x}(g^n-1) \ge g-1 \ge 1$ and hence

$$\log\log(p_g(x)g) \ge \log\log g,$$

thus Corollary 3.9 follows. □

In the next lemma, we study the function $A_g(m,r)$ defined in (3.14), following the argument in [LPZ07] adapted to powers of g.

Lemma 3.10. *Let $m \in [X/2, X]$ be such that $m \equiv g \pmod 2$ and r a positive odd integer such that $r \gg (\log X)^2$. Then, given $A_g(m,r)$ as in* (3.14)*, we have*

$$A_g(m,r) \ll \omega L$$

for $\omega > 0$ arbitrary.

Proof. We first remark that $A_g(m,r)$ is defined for $m > g^\nu$. This leads to the hypothesis $m \geq X/2$ in the statement of the lemma, because $m \geq X/2 > X/g^{\sqrt{L}} \geq g^\nu$, for any $1 \leq \nu \leq L'$.

Since $m - g^\nu$ is even by hypothesis, according to the property (3.18) of the singular series and by relation (3.20), we have

$$\mathfrak{S}(m-g^\nu) = 2c_0 \prod_{\substack{p|(m-g^\nu)\\ p>2}} \frac{p-1}{p-2} = 2c_0 \sum_{d|(m-g^\nu)} \mu^2(d) f(d).$$

Thus, by definition (3.14) of $A_g(m,r)$ we get

$$\begin{aligned} A_g(m,r) &= 2c_0 \sum_{\substack{1\leq\nu\leq L'\\ g^\nu<m\\ r|(m-g^\nu)}} \sum_{d|(m-g^\nu)} \mu^2(d) f(d) \\ &\leq 2c_0 \sum_{d\leq X} \mu^2(d) f(d) \sum_{\substack{1\leq\nu\leq L'\\ g^\nu<m\\ g^\nu\equiv m \bmod [r,d]}} 1, \end{aligned} \tag{3.44}$$

since $d \leq m - g^\nu \leq m \leq X$ and with $[r,d] = \operatorname{lcm}(r,d)$.

Recalling the definitions (3.28) of $S_g(m,[r,d])$, we can bound the inner sum in (3.44) as

$$\sum_{\substack{1\leq\nu\leq L'\\ g^\nu<m\\ g^\nu\equiv m \bmod [r,d]}} 1 \leq S_g(m,[r,d]) \Big(\frac{L'}{\overline{\xi_g}([r,d])} + 1 \Big). \tag{3.45}$$

In fact, if there's no $\nu \in [1,L']$ such that $g^\nu \equiv m \pmod{[r,d]}$, then both sides of the inequality are null: $S_g(m,[r,d]) = 0$ by definition and the sum in the left-hand side is over an empty set. Otherwise, $S_g(m,[r,d]) = 1$ and so the inequality trivially follows by definition (3.31) of $\overline{\xi_g}([r,d])$.

Remark 3.11. *By* (3.32) *and Remark 3.5, whenever $[r,d]|g$, we have*

$$\overline{[r,d]} = 1 = \overline{\xi_g}([r,d])$$
$$S_g(m,[r,d]) = \begin{cases} 1 & \text{if } [d,r]|m, \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, if $[r,d]|m$, then $g^\nu \equiv m \equiv 0 \pmod{[d,r]}$ for every $1 \leq \nu \leq L'$. So whenever $[r,d]|(g,m)$, we have

$$\sum_{\substack{1\leq\nu\leq L'\\ g^\nu<m\\ g^\nu\equiv m \bmod [r,d]}} 1 = L',$$

and therefore inequality (3.45) *holds.*

Putting (3.45) in (3.44), we get

$$\begin{aligned} A_g(m,r) &\le 2c_0 \sum_{d\le X} \mu^2(d) f(d) S_g(m,[r,d]) \Big(\frac{L'}{\overline{\xi_g}([r,d])} + 1 \Big) \\ &\le 2c_0 L' \sum_{d\le X} \mu^2(d) \frac{f(d)}{\overline{\xi_g}([r,d])} + 2c_0 \sum_{d\le X} \mu^2(d) f(d) S_g(m,d) \\ &= 2c_0 L' \sum_{d\le X} \mu^2(d) \frac{f(d)}{\overline{\xi_g}([r,d])} + 2c_0 S_g^{\prime(tot)}(m,X), \end{aligned} \tag{3.46}$$

with $S_g^{\prime(tot)}(m,X) = S_g'(m,X) + \overline{S_g'}(m,X)$. Here we used the definition of $S_g'(m,X)$ and $\overline{S_g'}(m,X)$ in (3.28) and (3.34) resp., and the inequality $S_g(m,[r,d]) \le S_g(m,d)$, which follows from $d|[r,d]$.

Similarly we also have $\overline{\xi_g}([r,d]) \ge \overline{\xi_g}(d), \overline{\xi_g}(r)$. In fact $a|b$ implies $\frac{a}{(g,a)} | \frac{b}{(g,b)}$ since

$$a|b \;\Rightarrow\; b = an \;\Rightarrow\; \frac{b}{(b,g)} = \frac{an}{(an,g)} = \frac{an}{(a,g)(n,\frac{g}{(g,a)})}.$$

So both $\bar d$ and $\bar r$ divide $\overline{[r,d]}$. By this and splitting the sum in (3.46) into $d \le D$ and $d > D$, with $D = D(\omega)$ such that $\sum_{d>D} \mu^2(d) \frac{f(d)}{\overline{\xi_g}(d)} \le \omega/4$ (such a D exists, since the series $\sum_d \mu^2(d) \frac{f(d)}{\overline{\xi_g}(d)}$ converges by Lemma 3.6 and $f(d)/\overline{\xi_g}(d) \ge 0$), we get

$$\begin{aligned} A_g(m,r) &\le 2c_0 L' \Big(\frac{1}{\overline{\xi_g}(r)} \sum_{d\le D} \mu^2(d) f(d) + \sum_{D<d\le X} \mu^2(d) \frac{f(d)}{\overline{\xi_g}(d)} \Big) + \mathcal{O}\big(S_g^{\prime(tot)}(m,X)\big) \\ &\le 2c_0 L' \Big(\frac{1}{\overline{\xi_g}(r)} \sum_{d\le D} \mu^2(d) f(d) + \frac{\omega}{4} \Big) + \mathcal{O}\big(S_g^{\prime(tot)}(m,X)\big). \end{aligned}$$

It is easy to see that $\overline{\xi_g}(r) \ge \log_g(\bar r + 1)$, with $\bar r = r/(r,g)$: in fact otherwise we would have $g^{\overline{\xi_g}(r)} - 1 < \bar r$, against the definition of $\overline{\xi_g}(r)$, which implies $\bar r | (g^{\overline{\xi_g}(r)} - 1)$. Moreover $\sum_{d\le D} \mu^2(d) f(d) \ll \log D$ by Friedlander-Goldston [FG95], Lemma 2.1. So

$$A_g(m,r) \ll L' \Big(\frac{\log D}{\log_g(\bar r + 1)} + \frac{\omega}{4} \Big) + S_g^{\prime(tot)}(m,X) \ll \omega L + S_g^{\prime(tot)}(m,X), \tag{3.47}$$

since $\log D / \log_g(\bar r + 1) \ll \omega/4$, being $r \gg (\log X)^2$. To conclude, we have to estimate $S_g^{\prime(tot)}(m,X)$.

Estimation of $S_g^{\prime(tot)}(m,X)$. Letting now

$$p_g(m) = \prod_{1\le \nu \le L'} (m - g^\nu),$$

by the definition of $S_g(m,d)$, $S'_g(m,X)$ and $\overline{S'_g}(m,X)$ in (3.28) and (3.34), and recalling that $S_g^{\prime(tot)}(m,X) = S'_g(m,X) + \overline{S'_g}(m,X)$, we get

$$\begin{aligned} S_g^{\prime(tot)}(m,X) &= \sum_{1\leq\nu\leq L'} \sum_{d|(m-g^\nu)} \mu^2(d)f(d) \leq \sum_{d|p_g(m)} \mu^2(d)f(d) = \prod_{\substack{p|p_g(m)\\ p>2}} \frac{p-1}{p-2} \\ &= \exp\Big[\log\Big(\prod_{\substack{p|p_g(m)\\ p>2}} \big(1+\frac{1}{p-2}\big)\Big)\Big] = \exp\Big(\sum_{\substack{p|p_g(m)\\ p>2}} \log\big(1+\frac{1}{p-2}\big)\Big) \\ &\ll \exp\Big(\sum_{\substack{p|p_g(m)\\ p>2}} \frac{1}{p-2}\Big) \ll \exp\Big(\sum_{p|p_g(m)} \frac{1}{p}\Big), \end{aligned} \tag{3.48}$$

where we used the Taylor series expansion of the log function and $1/(p-2) \leq 3/p$, for all $p \geq 3$. We now split the last sum in (3.48) into $p > (\log X)^3$ and $p \leq (\log X)^3$ and we observe that

$$\sum_{\substack{p|p_g(m)\\ p>(\log X)^3}} \frac{1}{p} \leq \sum_{\substack{p|p_g(m)\\ p>(\log X)^3}} \frac{\log p}{p} \leq \frac{1}{(\log X)^3} \sum_{\substack{p|p_g(m)\\ p>(\log X)^3}} \log p \leq \frac{1}{(\log X)^3} \sum_{p|p_g(m)} \log p \leq \frac{1}{\log X}, \tag{3.49}$$

because it is easy to see that

$$\exp\Big(\sum_{p|p_g(m)} \log p\Big) = \prod_{p|p_g(m)} p \leq p_g(m) \leq \prod_{1\leq\nu\leq L'} X = X^{L'} \leq \exp((\log X)^2). \tag{3.50}$$

On the other hand, we have

$$\sum_{\substack{p|p_g(m)\\ p\leq(\log X)^3}} \frac{1}{p} \leq \sum_{p\leq(\log X)^3} \frac{1}{p} \ll \log\log((\log X)^3) + 1 \ll \log\log\log X, \tag{3.51}$$

by Result (R.18). So by (3.48)-(3.51) we get

$$S_g^{\prime(tot)}(m,X) \ll \exp\Big(\frac{1}{\log X} + \log\log\log X + 1\Big) \ll \log\log X. \blacksquare$$

Then by (3.47)

$$A_g(m,r) \ll \omega L + \log\log X \ll \omega L$$

which proves the lemma. □

Remark 3.12. *We can prove the same estimation for $S_g^{\prime(tot)}(m,X)$ in a different way: by Lemma 3.8 we have*

$$S_g^{\prime(tot)}(m,X) = \sum_{1\leq\nu\leq L'} \sum_{d|(m-g^\nu)} \mu^2(d)f(d) \leq \sum_{d|p_g(m)} \mu^2(d)f(d)$$

$$< \mathcal{C}e^{\gamma}(\log\log(p_g(m)) + 3) + \mathcal{O}\bigg(\frac{1}{\log\log(p_g(m))}\bigg),$$

where $\mathcal{C} < 0.7574$ *is an absolute constant and* $\boldsymbol{\gamma}$ *is the Euler constant. Arguing as in* (3.50) *and observing that* $\prod_{p|p_g(m)} p \geq 2$ *since* $m - g^{\nu}$ *is even, we get*

$$2 \leq p_g(m) \leq \exp((\log X)^2).$$

So, taking twice the logarithm, we obtain again

$$S_g'^{(tot)}(m, X) < \mathcal{C}e^{\gamma}(2\log\log(X)) + 3) + \mathcal{O}(1) \ll \log\log X.$$

Remark 3.13. *By construction*

$$S_g'(m, X),\ \overline{S_g'}(m, X),\ \overline{T_g'}(m, X) \geq 0,$$

and

$$\begin{aligned} S_g'^{(tot)}(m, X) &= S_g'(m, X) + \overline{S_g'}(m, X), \\ \overline{T_g'}(m, N) &= \overline{S_g'}(m, N) - \sum_{\substack{d \leq N \\ d|(g,m)}} \mu^2(d) f(d). \end{aligned}$$

Then

$$S_g'^{(tot)}(m, X) \ll \log\log X \quad \Rightarrow \quad S_g'(m, X), \overline{S_g'}(m, X), \overline{T_g'}(m, X) \ll \log\log X.$$

Lemma 3.14. *Given* $m \in \mathbb{N}$,

$$\prod_{2 < p \leq \log m} \frac{p(p-1)}{(p-2)(p+1)} < 1.8687.$$

Proof. It is easy to see that $V(m) = \prod_{2<p\leq\log m} \frac{p(p-1)}{(p-2)(p+1)} < \prod_{p>2} \frac{p(p-1)}{(p-2)(p+1)} = \prod_{p>2} \big(1 + \frac{2}{p^2-p-2}\big) = \mathcal{V}$ which is convergent. We apply Cohen's method (see [Coh07], §10.3.6, we also refer to Cohen's unpublished preprint [Coh] for more details) to prove that $\mathcal{V} \approx 1.8687$ with at least 100 correct digits (see Section 3.8 for the PARI-GP program we used to compute $\mathcal{V}$ and to estimate the error). To this end we will introduce the parameters B, M and J and we will decompose $\mathcal{V}$ into two parts: one is effectively computable and give rise to the numerical value, the other one measures the error and it depends on the parameters. To conclude, we will optimize the parameters in order to make the error to be less than the fixed tolerance of 100 correct digits.

We start by fixing a parameter $B > 2$ and splitting the product in $\mathcal{V}$ into $p \leq B$ and $p > B$. Clearly, the product $\prod_{p>2}^{B} \frac{p(p-1)}{(p-2)(p+1)}$ is finite, so its value can be effectively computed and we call it $\mathcal{T}_{\mathcal{B}}$.

By a simple trick, we can write

$$\mathcal{V} = \mathcal{T}_{\mathcal{B}} \prod_{p>B} \Big[\Big(1 - \frac{2}{p}\Big)^{-1}\Big(1 - \frac{1}{p}\Big)\Big(1 + \frac{1}{p}\Big)^{-1}\Big],$$

and taking the logarithm of both sides, using the Taylor series expansion of the log function and recalling that $\mathcal{V}$ converges, we obtain

$$\begin{aligned}\log \mathcal{V} &= \log \mathcal{T}_{\mathcal{B}} + \sum_{p>B} \Big[-\log\Big(1-\frac{2}{p}\Big) + \log\Big(1-\frac{1}{p}\Big) - \log\Big(1+\frac{1}{p}\Big)\Big] \\ &= \log \mathcal{T}_{\mathcal{B}} + \sum_{p>B} \Big[\sum_{m\geq 1} \frac{1}{mp^m}\big(2^m - 1 + (-1)^m\big)\Big] \\ &= \log \mathcal{T}_{\mathcal{B}} + \sum_{m\geq 2} \frac{2^m - 1 + (-1)^m}{m} \sum_{p>B} \frac{1}{p^m}, \end{aligned} \tag{3.52}$$

since the contribution of $m = 1$ is null.

We now fix another parameter $M > 2$ and we split the external sum in (3.52) into $2 \leq m \leq M$ and $m > M$. Let us define

$$\mathrm{err}_1 = \mathrm{err}_1(M, B) = \sum_{m>M} \frac{2^m - 1 + (-1)^m}{m} \sum_{p>B} \frac{1}{p^m}. \tag{3.53}$$

As an application of Möbius' inversion formula, we have

$$\sum_{p>B} \frac{1}{p^m} = \sum_{j\geq 1} \frac{\mu(j)}{j} \log(\zeta_{p>B}(jm)),$$

where $\zeta_{p>B}(n) = \prod_{p>B}(1-p^{-n})^{-1}$ (see Cohen [Coh07], page 209). So, collecting this formula and (3.52)-(3.53), we get

$$\log \mathcal{V} = \log \mathcal{T}_{\mathcal{B}} + \mathrm{err}_1 + \sum_{2\leq m\leq M} \frac{2^m - 1 + (-1)^m}{m} \sum_{j\geq 1} \frac{\mu(j)}{j} \log(\zeta_{p>B}(jm)).$$

We now fix the last parameter $J > 1$ and we split the inner sum above into $j \leq J$ and $j > J$, so that the finite sum

$$\mathcal{T}_{\mathcal{J}} = \sum_{2\leq m\leq M} \frac{2^m - 1 + (-1)^m}{m} \sum_{1\leq j\leq J} \frac{\mu(j)}{j} \log(\zeta_{p>B}(jm))$$

can be effectively computed. Let us define

$$\mathrm{err}_2 = \mathrm{err}_2(M, B, J) = \sum_{2\leq m\leq M} \frac{2^m - 1 + (-1)^m}{m} \sum_{j>J} \frac{\mu(j)}{j} \log(\zeta_{p>B}(jm)). \tag{3.54}$$

Using this notation we get

$$\log \mathcal{V} = \log \mathcal{T}_{\mathcal{B}} + \mathcal{T}_{\mathcal{J}} + \mathrm{err}_1 + \mathrm{err}_2$$

and so, taking the exp of both sides and using the Taylor series expansion of the exp function we obtain

$$\mathcal{V} = \mathcal{T}_{\mathcal{B}} e^{\mathcal{T}_{\mathcal{J}}} e^{\mathrm{err}_1 + \mathrm{err}_2} = \mathcal{T}_{\mathcal{B}} e^{\mathcal{T}_{\mathcal{J}}} (1 + \mathcal{O}(|\,\mathrm{err}_1 + \mathrm{err}_2\,|)). \tag{3.55}$$

As said before, $\mathcal{T}_{\mathcal{B}}e^{\mathcal{T}_{\mathcal{J}}}$ can be effectively computed and its value is ≈ 1.8687 with error less than 10^{-100} (see Section 3.8). So it is left to estimate the error term $|\operatorname{err}_1+\operatorname{err}_2|$, in particular, our aim is to set the parameters M, B and J in order to have the error term less than a fixed tolerance.

Estimation of err_2**.** It is easy to see that

$$|\log(\zeta_{p>B}(jm))| = \sum_{p>B}\sum_{n\geq 1}\frac{1}{np^{njm}} \leq \sum_{n>B}\frac{1}{n^{jm}} \leq \int_B^{+\infty}\frac{\mathrm{d}t}{t^{jm}} = \frac{B^{1-jm}}{jm-1}.$$

It follows that, by the definition (3.54), the triangular inequality and trivial bounds, we have

$$\begin{aligned}
|\operatorname{err}_2| &\leq \sum_{2\leq m\leq M}\frac{2^m}{m}\sum_{j>J}\frac{B^{1-jm}}{j(jm-1)}\\
&\leq \frac{B}{(J+1)(2J+1)}\sum_{2\leq m\leq M}\frac{2^m}{m}\sum_{j>J}\frac{1}{(B^m)^j}\\
&= \frac{B}{(J+1)(2J+1)}\sum_{2\leq m\leq M}\frac{2^m}{m}\Big(\frac{1}{B^{mJ}(B^m-1)}\Big)\\
&\leq \frac{B}{(J+1)(2J+1)}\sum_{2\leq m\leq M}\frac{2^m}{m}\Big(\frac{2}{B^{m(J+1)}}\Big)\\
&\leq \frac{B}{(J+1)(2J+1)}\sum_{2\leq m\leq M}\Big(\frac{2}{B^{J+1}}\Big)^m\\
&= \frac{B^{J+2}}{(J+1)(2J+1)(B^{J+1}-2)}\Big(\Big(\frac{2}{B^{J+1}}\Big)^{M+1}+\Big(\frac{2}{B^{J+1}}\Big)^2\Big)\\
&\ll \frac{B^{J+2}}{(J+1)(2J+1)(B^{J+1}-2)}\Big(\frac{2}{B^{J+1}}\Big)^2\\
&= \frac{4}{(J+1)(2J+1)(B^{J+1}-2)B^J},
\end{aligned} \tag{3.56}$$

where we used $B^m-1\geq B^m/2$ being $B^m>2$.■

Estimation of err_1**.** Using definition (3.53), trivial bounds and the geometric series formula, we get

$$\begin{aligned}
|\operatorname{err}_1| &\leq \sum_{m>M}\frac{2^m}{m}\sum_{n>B}\frac{1}{n^m} \leq \sum_{m>M}\frac{2^m}{m}\int_B^{+\infty}\frac{dt}{t^m} = \sum_{m>M}\frac{2^m}{m}\cdot\frac{B^{1-m}}{m-1}\\
&\leq \frac{B}{(M+1)M}\sum_{m>M}\Big(\frac{2}{B}\Big)^m = \frac{2^{M+1}}{(M+1)MB^{M-1}(B-2)}\cdot\blacksquare
\end{aligned} \tag{3.57}$$

We now fix 100 to be the number of correct digits we want in the effective computation of $\mathcal{V}$. So by (3.55), we are asking

$$\mathcal{V} - \mathcal{T}_{\mathcal{B}} e^{\mathcal{T}_{\mathcal{J}}} = \mathcal{O}\big(\mathcal{T}_{\mathcal{B}} e^{\mathcal{T}_{\mathcal{J}}} |\operatorname{err}_1 + \operatorname{err}_2|\big) \leq 10^{-100}.$$

Since the effective computation (see Section 3.8) gives that $\mathcal{T}_{\mathcal{B}} e^{\mathcal{T}_{\mathcal{J}}} \approx 1.8687 \leq 10$, we want

$$|\operatorname{err}_1|, |\operatorname{err}_2| \leq 10^{-101} \qquad \textit{i.e.} \qquad \log_{10}(|\operatorname{err}_1|), \log_{10}(|\operatorname{err}_2|) \leq -101. \tag{3.58}$$

Bound for $\operatorname{err}_1 = \operatorname{err}_1(M, B)$**.** We want to fix the parameters M and B in such a way that (3.58) holds. From (3.57) we get

$$\begin{aligned}
\log_{10}(|\operatorname{err}_1(M,B)|) &\leq \log_{10}\Big(\frac{2^{M+1}}{(M+1)MB^{M-1}(B-2)}\Big) \\
&= (M+1)\log_{10} 2 - \log_{10}(M+1) - \log_{10} M - (M-1)\log_{10} B \\
&\quad - \log_{10}(B-2) \\
&< (M-1)\log_{10} 2 - M\log_{10}(B-2) \\
&\leq (M-1)\frac{31}{100} - M\log_{10}(B-2),
\end{aligned}$$

since $\log_{10} M > \log_{10} 2$ by construction of M. Therefore, by this inequality and (3.58), if we fix $B = 50$, we get $\log_{10}(B-2) \geq 1.68$ and so

$$M\Big(\frac{31}{100} - \frac{168}{100}\Big) \leq -101 + \frac{31}{100} \qquad \textit{i.e.} \qquad M \geq 74,$$

so let us set $B = 50$ and $M = 74$.■

Bound for $\operatorname{err}_2 = \operatorname{err}_2(M, B, J)$**.** Using the values of B and M set above, we want to fix J in such a way that (3.58) holds. From (3.56) with $B = 50$ and $M = 74$ we get

$$\begin{aligned}
\log_{10}(|\operatorname{err}_2(M,B,J)|) &\leq \log_{10}\Big(\frac{4}{(J+1)(2J+1)(50^{J+1}-2)50^J}\Big) \\
&= \log_{10} 4 - \log_{10}(J+1) - \log_{10}(2J+1) - \log_{10}(50^{J+1}-2) \\
&\quad - J\log_{10} 50 \\
&\leq \log_{10} 4 - 2\log_{10} J - \log_{10} 2 - 2J\log_{10} 50 \\
&< \log_{10} 2 - 2J\log_{10} 50 \\
&\leq \frac{31}{100} - \frac{17}{5}J
\end{aligned}$$

where we used $50^{J+1} - 2 \geq 50^J$ and $\log_{10} J > 0$. So (3.58) implies

$$\frac{31}{100} - \frac{17}{5}J \leq -101 \qquad \textit{i.e.} \qquad J \geq 30,$$

so let us set $J = 30$.■

We now compute $\mathcal{T}_{\mathcal{B}} e^{\mathcal{T}_{\mathcal{J}}}$ with the fixed values $M = 74$, $B = 50$ and $J = 30$, thus obtaining $\mathcal{V} \approx 1.8687$ with at least 100 correct digits (see Section 3.8 below); this concludes the proof of Lemma 3.14. □

3.8 PARI-GP program

To compute $\mathcal{V}$ with 100 correct digits, we used the following PARI-GP program:

```
/****************************** V. SETTIMI *********************************
********************** COMPUTATION OF CONSTANT V **************************
************ WITH A PRECISION OF AT LEAST 100 DECIMAL DIGITS ***************
*******************************************************************************/

\\ Global variables:
\\ zetavector: vector of the needed values of the Riemann zeta function
\\ B1: level B of the primes used for the finite products

global(zetavector, B1);
global(defaulterror=0);


{V()=local(
B,J,M, S, P, freq,  maxim1,col,
defaultprecision,v,err1,err2,totalerr,finalerr
);

print("****** V. SETTIMI *******");
print("****** COMPUTATION OF CONSTANT V *******");
print("****** WITH A PRECISION OF AT LEAST 100 DECIMAL DIGITS *******");

defaultprecision=120;

default(realprecision,defaultprecision);
defaulterror=10^(-defaultprecision);

B= 50;
M= 74;
J= 30;
\\ B1 is the parameter B, as global variable
B1=B;

print("B= ", B);
print("M= ", M);
print("J= ", J);
print(" ");

\\ Initialization of zetavector: zetavector[i]=zeta(i), starting from i=2
maxim1=max(M,J);
zetavector=vector(M*J,col,0);
```

```
freq=vector(M*J,col,0);
for(j=2,maxim1,zetavector[j]=zeta(j); freq[j]=freq[j]+1);
for(m=2, M,
for (j=2,J,zetavector[m*j]=zeta(m*j); freq[m*j]=freq[m*j]+1));

\\ Computation of the truncated product over primes TB
TB=1;
forprime(p=3, B, TB = TB * (1+2/((p-2)*(p+1))); );

\\ Computation of the truncated sum TJ
TJ=0;
for(m=2, M, TJ= TJ+ ((2^m-1+(-1)^m)/m) * Sm(m,B,J));

\\ Computation of V
v = TB*exp(TJ);
print("The constant V is: ");
print(v);
print(" ");

\\ Estimation of error terms
print("With error: ");

err1= 2^(M+1)/(M*(M+1)*B^(M-1)*(B-2));
print("err1= ", err1*1.0);

err2= 4/(B^J*(B^(J+1)-2)*(J+1)*(2*J+1));
print("err2= ", err2*1.0);
print(" ");

totalerr= err1+err2;
finalerr=abs(v)*abs(totalerr);
print("Total Error: ",finalerr*1.0);
print("Correct digits: ",floor(abs(log(finalerr)/log(10)))-4);
print(" ");

}

/********************** COMPUTATION OF TAIL OF ****************************
********* ZETA FUNCTION WITH s>1 AND B THE UPPER BOUND ON PRIMES ***********
*****************************************************************************/

{zetaprodtail(s,B)= local(P, u);
P=1;
u=-s;
```

```
forprime(p=2, B , P=P*(1-p^u));
P= zetavector[s] * P;
return(P)
}


/******************** COMPUTATION OF TRUNCATED SUM ***********************
************** sum_{p prime > B} p^{-m}, WITH J UPPER BOUND ***************
*************************************************************************/

{Sm(m,B,J)= local(Smtail, u);
u=-m;
Smtail=0;
for(j=1,J,Smtail=Smtail+ moebius(j)*log(zetaprodtail(j*m,B))/j);
return(Smtail)
}
```

Running this program we obtain:

```
****** V. SETTIMI *******
****** COMPUTATION OF CONSTANT V *******
****** WITH A PRECISION OF AT LEAST 100 DECIMAL DIGITS *******
B= 50
M= 74
J= 30

The constant V is:
1.868785077418560788885072717487369900905147801909466688848496582866860656613
40534000810864344232976825772284577375815298 9

With error:
err1= 1.3393840960078452337258689653242259162751000600600600600600600600600 6
00600600600600600600600600600600600600600600600600600 6 E-107
err2= 4.877510331493800004230565838180856689582231623479640611391228223861 28
39596614234161819143310417768376520449572281047313 2 E-107

Total Error: 1.161803953400168124963140048429018642230162247313725223763015 1
429459898509463318941977391080732571425805063013005502393 6 E-106
Correct digits: 101
```

Appendix A

Point-counting using cohomology

A.1 Introduction

A.1.1 Point-counting problems

The first clear presentation of modular arithmetic can be found in Carl Friedrich Gauss' "Disquisitiones Arithmeticae", dated 1801, where, among other things, the author studies the number of solutions of equations modulo prime numbers.

More generally, we can ask for the number of simultaneous solutions of a system of polynomial equations modulo q, where $q = p^a$ for p prime number and a positive integer. Recalling that the set of solutions of a system of polynomial equations defines an algebraic variety, we can rewrite such a question geometrically, asking for the number of points of an algebraic variety over $\mathbb{F}_q$. This is the so called *point-counting problem*.

The naive solution is checking all the q^s points in $\mathbb{F}_q^s$, where s is the number of indeterminates occurring in the polynomials defining the variety. However, when q or s are too large, this becomes an "impossible task", that means it is too costly, even for a computer. Therefore, for effective solutions, smarter methods are needed.

Given $q = p^a$, where p is a prime number and a is a positive integer, let X be an algebraic variety over the finite field $\mathbb{F}_q$. We define

$$N(X) = |X(\mathbb{F}_q)|$$

to be the number of rational points of X. The crucial observation is that these $\mathbb{F}_q$-rational points are precisely those points of X that are fixed by the *Frobenius endomorphism*

$$\mathrm{Frob} : X \to X,$$

which is the identity on the topological space of X and raises the elements of the structural sheaf to their q-power (see Hartshorne [Har77], page 301). In general, for every $n \in \mathbb{N} \setminus \{0\}$ we define

$$N_n(X) = |X(\mathbb{F}_{q^n})|$$

to be the number of $\mathbb{F}_{q^n}$-rational points of X. This sequence of integers is used to the define the *zeta function of* X as follows

$$\zeta(X,t) = \exp\Big(\sum_{n\geq 1} \frac{N_n(X)}{n} t^n\Big).$$

A priori $\zeta(X,t)$ is an element of the ring of formal power series over $\mathbb{Q}$, but it turns out to be a rational function. This fundamental property is one of the famous *Weil conjectures* which investigate the behavior of the zeta function (see Hartshorne [Har77], Appendix C). These conjecture were introduced by Weil in [Wei49] and proved by Dwork [Dwo60], Grothendieck [Gro65] and Deligne [Del74].

A.1.2 Application to cryptography

Although the point-counting problem is an interesting problem in itself, in recent years it has received a lot of attention because of a number of applications, especially in cryptography. As example, we describe here how the point counting problem can be applied to the Diffie-Hellman protocol.

Suppose that Alice and Bob want to communicate with each other over an unsafe channel in a safe way, which means without letting their messages open to outsiders. Therefore, they have to encrypt their messages and this is, usually, done by means of a secret key, shared by Alice and Bob. So the problem is how can they safely exchange this key.

A well known method of exchanging keys is the so called *Diffie-Hellman protocol*, introduced in [DH76], which works as follows:

a) Alice and Bob first choose a commutative group G and an element $g \in G$. This information is public.

b) Alice arbitrarily chooses an integer a, then she computes $a \cdot g \in G$ and sends it to Bob.

c) Similarly, Bob arbitrarily chooses an integer b, computes $b \cdot g \in G$ and sends it to Alice.

d) Both Alice and Bob can then compute $k = a \cdot (b \cdot g) = b \cdot (a \cdot g)$, which is their shared secret key.

Such a key k is truly secret provided that an eavesdropper cannot compute it, given the public key (G, g) and both $a \cdot g$ and $b \cdot g$. This is known as the *Diffie-Hellman problem* and it is closely related to the so called *discrete logarithm problem*, which is the problem of retrieving n, given (G, g) and $n \cdot g$. Clearly, if we can solve the discrete logarithm problem, we can also solve the Diffie-Hellman one, however, if G and g are well chosen, both the problems are considered to be very hard. We refer to [CFA+06], §1.5-1.6 for a survey on these subjects.

The discrete logarithm and Diffie-Hellman problems can be related to point-counting algorithms, since, given a curve C defined over a finite field $\mathbb{F}_q$, if we consider its *Jacobian* $J = \mathrm{Jac}(C)$, which is an abelian variety over $\mathbb{F}_q$, then the set of its rational points $J(\mathbb{F}_q)$ is a

commutative group and so it can be used to perform the Diffie-Hellman protocol. However, to guarantee the security of the protocol, we need $|J(\mathbb{F}_q)|$ to have a suitably large prime factor. Therefore, to determine whether a curve is good for cryptography, one has to be able to compute the number of rational points of its Jacobian.

A.1.3 Weil cohomology

The turning point in finding smart solutions to point-counting problems is due to Weil [Wei49]. He started from the fact that, as observed in A.1.1, the rational points of a $\mathbb{F}_q$-variety are those points that stay fixed under the action of Frobenius. For varieties over $\mathbb{C}$, the problem of counting the fixed points of a given endomorphism had been extensively studied before Weil and an important result in this field was the *Lefschetz fixed point theorem* for $\mathbb{C}$-variety (see Greenberg-Harper [GH81], Theorem 30.9). Weil's idea was to find an analogous theorem for varieties over finite fields. In order to do this, he developed a suitable cohomology theory, which is defined as follows (we refer to the classical survey by Kleiman in [GGK$^+$68], in particular to §1.2).

Definition A.1. *Let $\mathbb{F}_q$ be a finite field of characteristic p and let K be a field of characteristic zero, called* coefficient field. *A* Weil cohomology *is a contravariant functor $X \mapsto H^\star(X)$ from the category of smooth, proper, geometrically irreducible $\mathbb{F}_q$-varieties to the category of graded K-algebras, satisfying the following three properties:*

1. *(*Poincaré duality*) If d is the dimension of X, then*
   - *i) the $H^i(X)$ are finite dimensional K-vector spaces and $H^i(X) = 0$, whenever $i < 0$ or $i > 2d$;*
   - *ii) $H^{2d}(X) \simeq K$;*
   - *iii) there are perfect pairings $H^i(X) \times H^{2d-i}(X) \to H^{2d}(X)$, for every $0 \leq i \leq 2d$.*
2. *(*Künneth formula*) For every $\mathbb{F}_q$-varieties X and Y, the projections $X \times Y \to X$ and $X \times Y \to Y$ canonically induce an isomorphism $H^\star(X) \otimes_K H^\star(Y) \xrightarrow{\sim} H^\star(X \times Y)$.*
3. *(*Cycle map*) Let $C^i(X)$ be the group of all the formal sums of subvarieties of codimension i in X. Then there exist group homomorphisms $\gamma_X : C^i(X) \to H^{2i}(X)$ which are functorial, multiplicative (*i.e. *compatible with the Künneth formula) and non-trivial.*

By functoriality, if $H^\star$ is a Weil cohomology, then the Frobenius endomorphism Frob induces maps

$$\mathrm{Frob}^* : H^i(X) \to H^i(X)$$

on cohomology, for every $\mathbb{F}_q$-variety X. Moreover, by the properties defining a Weil cohomology, we can deduce (see Kleiman [GGK$^+$68], Proposition 1.3.6 and §4) that the number of $\mathbb{F}_q^n$-rational points of X is given by the formula

$$N_n(X) = \sum_{i=0}^{2\dim(X)} (-1)^i \operatorname{Tr}((\mathrm{Frob}^n)^* | H^i(X)), \tag{A.1}$$

where Frob^n is the n-th iterate of Frob. This is the key relation in point-counting algorithms.

From (A.1), arguing as in Hartshorne [Har77], Lemmas 4.1-4.2 of Appendix C, we obtain the following explicit formula for the zeta function of X as an element of $\mathbb{Q}(t)$:

$$\zeta(X,t) = \prod_{i=0}^{2\dim(X)} \det\left(1 - \mathrm{Frob}^* \, t | H^i(X)\right)^{(-1)^{i+1}}.$$

Up to now, there are two main Weil cohomologies in characteristic p: the ℓ-adic cohomology, with values in $\mathbb{Q}_\ell$, and the p-adic rigid cohomology:

- In the 1960's Grothendieck introduced the *étale cohomology*, which is a very general type of cohomology that specializes to the usual singular cohomology for varieties over $\mathbb{C}$. For varieties over a finite field of characteristic p, if we take as coefficient field $\mathbb{Q}_\ell$, with ℓ a prime different from p, than this *étale ℓ-adic cohomology* is a Weil cohomology. See, *e.g.*, Harshorne [Har77], §3 of Appendix C or Milne's book [Mil80].

- The *p-adic rigid cohomology* was first introduced by Dwork in [Dwo60] and settled by Berthelot in [Ber86]. This p-adic cohomology is more handy because it is associated to a de Rham complex and not to a site as the étale one. For a comprehensive survey on it, we refer to Le Stum's monographic book [LS07].

A.1.4 Point-counting algorithms

The first point-counting algorithm is due to Schoof [Sch85] and it uses the ℓ-adic cohomology to attack elliptic curves over finite fields. More precisely, Schoof's algorithm uses ℓ_i-torsion points to compute the number of rational points modulo ℓ_i, where ℓ_i are small prime numbers different from the characteristic of the base field. Then it recovers the number of rational points via the Chinese Remainder Theorem. Schoof's algorithm was generalized by Pila [Pil90] for abelian varieties over finite fields and improved several times (see, *e.g.*, Elkies [Elk98]). However, a more fruitful approach is to use p-adic methods.

The first algorithm making use of p-methods is due to Satoh [Sat00] and it uses the p-adic canonical lift to attack elliptic curves over a finite fields of small characteristic.

A year later, Kedlaya [Ked01] exhibited a sort of generalization of Satoh's algorithm to hyperelliptic curve over finite fields of odd characteristic: instead of the p-adic canonical lift, which only fits for elliptic curves, Kedlaya's algorithm uses a rigid analytic lift. This kind of lift was introduced by Monsky-Washnitzer [MW68] and it applies to arbitrary non-singular affine varities[1]. The idea which underlies Kedlaya's method is to lift the curve, defined over the finite field $\mathbb{F}_q$, to one defined over $\mathbb{Z}_q$. Since the Frobenius morphism does not lift algebraically, one has to lift it analytically, using the so called *dagger* (or *weakly complete*) *algebras* (see the next Definition A.2). Then, as proved by Monsky-Washnitzer, we can compute the wanted rigid cohomology, using the classical algebraic de Rham cohomology.

[1] We remark that, when applied to an affine variety, rigid cohomology is isomorphic to the so called *Monsky-Washnitzer cohomology*. See [MW68] and also [vdP86].

Soon afterwards similar algorithms were developed for superelliptic curves (see Gaudry-Gürel [GG01]), for hyperelliptic curves in characteristic 2 and for $C_{a,b}$-curves (see Denef-Vercauteren [DV06b, DV06a]) and for non-degenerate curves (see Castryck-Denef-Vercauteren [CDV06]).

Another algorithm that can be seen as a generalization of Kedlaya's method, is the Chatel-Lubicz's one, introduced in [CL09], which uses the Monsky-Washnitzer cohomology *with compact support* to count the rational points of an hyperelliptic curve on a finite field of odd characteristic.

A different kind of p-adic algorithms was introduced by Lauder in [Lau04], using Dwork's *deformation theory* [Dwo63]. Lauder's idea is that, in order to compute the zeta function of a smooth hypersurface, one embeds it in a one-parameter family, such that the fiber at the origin is smooth and it has an easily computable zeta function. Dwork's theory associates to the one-parameter family a p-adic differential equation and, by computing a basis of solutions to such an equation around the origin (which is the "easy fiber") one can then recover the zeta function of the original hypersurface.

More recently, Lauder [Lau06] introduced another p-adic method for point-counting which is based on computing the zeta function by induction on the dimension. This method is called *fibration method* and it reduces the problem of calculating the action of the Frobenius on the rigid cohomology of a smooth projective variety over a finite field, to that of performing the same calculation for a smooth hyperplane section. In the inductive step, the algorithm uses the main technique developed for the deformation method.

We conclude this section remarking that, using p-adic algorithm, the main computational problem lies in the estimation of the analytic precision (since dagger algebra are involved) and the p-adic precision (since the coefficients are p-adic numbers) necessary to recover the action of the Frobenius.

A.2 Generalization of Chatel-Lubicz's algorithm

The starting point for our work is Chatel-Lubicz's algorithm for hyperelliptic curves. These curves are double covers of the projective line and our idea was to generalize the algorithm to non-plane curves, in particular to trigonal curves which are triple covers of the projective line. Before introducing our work, we briefly describe Chatel-Lubicz's algorithm.

Throughout this section let p be a prime number, let k be finite field of characteristic p and cardinality $q = p^a$, for some a positive integer a. Moreover, let $\mathcal{V}$ be a complete DVR having k as residue field and K as fraction field, where K is a characteristic zero field[2]. Let $\mathfrak{M}$ be the maximal ideal of $\mathcal{V}$. Finally we will denote by $\hookrightarrow$ an open immersion.

A.2.1 Chatel-Lubicz's algorithm

Let C_k be a smooth affine plane curve defined by a polynomial $f_k(X, Y)$ in $k[X, Y]$. Our aim is to find $H^i_{MW,c}(C_k/K)$, called the Monsky-Washnitzer cohomology with compact support

[2]For example $\mathcal{V}$ might be thought as the ring $W(k)$ of *Witt vectors* of k. See Serre [Ser79], Chapter II, §6.

spaces of C_k.

Let $\pi_k : C_k \to \mathbb{A}^1_k$ be the projection on the Y-axis. Actually, instead of considering $H^i_{MW,c}(C_k/K)$, it will be useful to deal with $H^i_{MW,c}(U_k/K)$, where U_k is the open subset of C_k defined as the étale locus of π_k (see Hartshorne [Har77], page 299). We remark that, since $U_k \overset{o}{\hookrightarrow} C_k$, than $H^i_{MW,c}(C_k/K) \hookrightarrow H^i_{MW,c}(U_k/K)$.

The first goal in Chatel-Lubicz's algorithm is therefore finding such an étale locus. It is done in three steps: (1) finding the branch locus $\Lambda_k \subseteq \mathbb{A}^1_k$; (2) removing Λ_k from $\mathbb{A}^1_k$ and so defining $V_k = \mathbb{A}^1_k \setminus \Lambda_k$; (3) finding $U_k = \pi_k^{-1}(V_k) \subseteq C_k$.

1. By definition, $y \in \mathbb{A}^1_k$ is in the branch locus of π_k if and only if there exists $x \in \mathbb{A}^1_k$ such that the point $(x, y) \in \mathbb{A}^2_k$ is actually in C_k and it is a ramification point of π_k. It means that $y \in \mathbb{A}^1_k$ is in the branch locus of π_k if and only if $f_k(X, y)$ has multiple roots as function of X, which is equivalent to $f_k(X, y)$ and $\frac{\partial f_k(X,y)}{\partial X}$ having common zeros. We remark that in this case the characterization of the brach locus is straightforward, because π_k is simply the projection on the Y-axis.

   To investigate such common zeros, we consider the *Bézout relation*

$$f_k(X, Y)u_k(X, Y) + \frac{\partial f_k(X, Y)}{\partial X} v_k(X, Y) = \Delta_k(Y),$$

   where $\Delta_k(Y) \in k[Y]$ is the *resultant* of $f_k(X, Y)$ and $\frac{\partial f_k(X,Y)}{\partial X}$ in $(k[Y])[X]$ with respect to X and $u_k(X, Y), v_k(X, Y) \in k[X, Y]$ are coprime in Y. Using this notation, we can finally define the branch locus of π_k as

$$\Lambda_k = \{\ y \in \mathbb{A}^1_k\ :\ \Delta_k(y) = 0\ \}.$$

2. By definition, we have

$$V_k = \operatorname{Spec}(k[Y]) \setminus \{y \in \mathbb{A}^1_k : \Delta_k(y) = 0\} = \operatorname{Spec}\Big(k\Big[Y, \frac{1}{\Delta_k(Y)}\Big]\Big).$$

   Let $B_k = k[Y, \frac{1}{\Delta_k(Y)}]$, so that $V_k = \operatorname{Spec} B_k$.

3. Finally, the étale locus of π_k is

$$U_k = \pi_k^{-1}(\operatorname{Spec} B_k) = \operatorname{Spec}\Big(k\Big[X, Y, \frac{1}{\Delta_k(Y)}\Big]/(f_k(X, Y))\Big),$$

   where the last step follows from the fact that π_k is the projection on the Y-axis. Let $A_k = k[X, Y, \frac{1}{\Delta_k(Y)}]/(f_k(X, Y))$, so that $U_k = \operatorname{Spec} A_k$.

Using the notation just introduced, we have the following commutative diagram

$$\begin{array}{ccc} C_k & \overset{o}{\hookleftarrow} & U_k \\ {\scriptstyle \pi_k}\downarrow & & \downarrow{\scriptstyle \pi_{k|_{U_k}}} \\ \mathbb{A}^1_k & \overset{o}{\hookleftarrow} & V_k \end{array} \tag{A.2}$$

where $\pi_{k|_{U_k}} : U_k \to V_k$ is finite étale (see [Har77], page 299). This $\pi_{k|_{U_k}}$ induces a ring homomorphism $B_k \to A_k$, that we will denote again by π_k with abuse of notation. Since $f_k(X, Y) = 0$ in A_k, if we set $d = \deg_X(f_k(X, Y))$, then $\{1, \dots, X^{d-1}\}$ is a basis af A_k as B_k-module.

Lifting in characteristic 0

The next step is lifting the diagram (A.2) in characteristic 0. It is possible thanks to Elkik [Elk73], Theorem 6, and we refer to Chatel's PhD thesis [Cha07], Chapter 3, for the details.

We recall that $\mathcal{V}$ is a complete DVR of mixed characteristic, having k as residue field and K as fraction field. We define $f(X,Y) \in \mathcal{V}[X,Y]$ to be a lift of $f_k(X,Y)$ such that it defines a smooth variety, say C, over $\mathcal{V}$. Moreover let $\pi : C \to \mathbb{A}^1_{\mathcal{V}}$ be the projection on the Y-axis, lifting π_k. By analogy with characteristic p, let $\Delta(Y) \in \mathcal{V}[Y]$ be the resultant of $f(X,Y)$ and $\frac{\partial f(X,Y)}{\partial X}$ with respect to X and let

$$f(X,Y)u(X,Y) + \frac{\partial f(X,Y)}{\partial X} v(X,Y) = \Delta(Y)$$

be the associated Bézout relation in $(\mathcal{V}[Y])[X]$, with $u(X,Y), v(X,Y) \in \mathcal{V}[X,Y]$ coprime in Y. Therefore the branch locus of π is

$$\Lambda = \{\ y \in \mathbb{A}^1_{\mathcal{V}}\ :\ \Delta(y) = 0\ \}.$$

Finally let $V = \mathbb{A}^1_{\mathcal{V}} \setminus \Lambda$ and $U = \pi^{-1}(V)$. Using this notation, we obtain the following commutative diagram

$$\begin{array}{ccc} C & \xleftarrow{o} & U \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi_{|U}} \\ \mathbb{A}^1_{\mathcal{V}} & \xleftarrow{o} & V \end{array}$$

where $\pi_{|U} : U \to V$ is finite étale and it induces a ring homomorphism $\pi : B \to A$ which is again finite étale, with

$$B = \mathcal{V}\Big[Y, \frac{1}{\Delta(Y)}\Big], \qquad\qquad A = \mathcal{V}\Big[X, Y, \frac{1}{\Delta(Y)}\Big]/(f(X,Y)).$$

As before, $\{1, \ldots, X^{d-1}\}$ is a basis of A as a finitely generated B-module.

Monsky-Washnitzer with compact support cohomology spaces

We don't give here the general definition of rigid cohomology with compact support (see, *e.g*, Le Stum [LS07], §6.4). However, for non-singular affine varieties, as in our situation, the rigid cohomology reduces to the Monsky-Washnitzer cohomology, which is more explicit. We write

$$H^i_{MW,c}(U_k/K, \mathcal{O}_{U_k}),$$

for the needed compact support Monsky-Washnitzer cohomology. We will denote it simply by $H^i_{MW,c}(U_k/K)$, when there is no ambiguity. Following van der Put [vdP86], §2, we may define it using $A^\dagger_K = A^\dagger \otimes_{\mathcal{V}} K$, where $A^\dagger$ is defined as follows (see Le Stum [LS07], page 117).

Definition A.2. *Let* $\underline{t} = (t_1, \ldots, t_n)$, $\underline{t}^{\underline{s}} = t_1^{s_1} \cdot \ldots \cdot t_n^{s_n}$ *and* $|\underline{s}| = \max_{1 \le i \le n} |s_i|$. *We define the* weak completion of $\mathcal{V}[\underline{t}]$ *as*

$$\mathcal{V}[\underline{t}]^\dagger = \Big\{\ \sum_{\underline{s}} a_{\underline{s}} \underline{t}^{\underline{s}} \in \mathcal{V}[[\underline{t}]]\ :\ \exists \rho > 1\ s.t.\ \lim_{|\underline{s}| \to \infty} |a_{\underline{k}}| \rho^{|\underline{s}|} = 0\ \Big\}.$$

A weakly complete algebra *is any quotient of* $\mathcal{V}[\underline{t}]^\dagger$ *over an ideal of finite type. Moreover, if* $A = \mathcal{V}[\underline{t}]/I$ *is the* $\mathcal{V}$*-algebra defined by the ideal* I*, then the* weak completion *of* A *is*

$$A^\dagger = \mathcal{V}[\underline{t}]^\dagger/(I \cdot \mathcal{V}[\underline{t}]^\dagger).$$

Hence, in our situation

$$A_K^\dagger = K\Big[X, Y, \frac{1}{\Delta(Y)}\Big]^\dagger/(f(X,Y)).$$

By finite étale descent (see Tsuzuki [Tsu99], Corollary 2.6.6), since the induced morphism $\pi : B \to A$ is finite étale, we have that

$$H^i_{MW,c}(U_k/K, \mathcal{O}_{U_k}) = H^i_{MW,c}(V_k/K, \pi_{k*}\mathcal{O}_{U_k}),$$

which is easier to handle because $V_k \subseteq \mathbb{A}^1_k$. The space on the right-hand side is defined using $A_K^\dagger$, endowed with the structure of $B_K^\dagger$-module given by π and

$$B_K^\dagger = K\Big[Y, \frac{1}{\Delta(Y)}\Big]^\dagger.$$

It's easy to see that $\{1, \dots, X^{d-1}\}$ is again a basis of $A_K^\dagger$ as $B_K^\dagger$-module.

By definition (see [LS07], Proposition 6.4.15) and recalling that C has dimension 1, we get

$$H^i_{MW,c}(V_k/K, \pi_{k*}\mathcal{O}_{U_k}) = H^i_{dR}((A_K^\dagger)_c[-1]),$$

which is the i-th cohomology space of the complex

$$0 \longrightarrow (A_K^\dagger)_c \xrightarrow{\nabla_c} (A_K^\dagger)_c \otimes_{B_K^\dagger} \Omega^1_{B_K^\dagger} \xrightarrow{\nabla_c} (A_K^\dagger)_c \otimes_{B_K^\dagger} \Omega^2_{B_K^\dagger} \xrightarrow{\nabla_c} \dots,$$

with $(A_K^\dagger)_c$ at degree 1. Since $A_K^\dagger$ is a finitely generated $B_K^\dagger$-module, according to [LS07], Proposition 6.1.18, we have the isomorphism

$$(A_K^\dagger)_c \simeq (B_K^\dagger)_c \otimes_{B_K^\dagger} A_K^\dagger, \tag{A.3}$$

where, by the definition of $B_K^\dagger$ and by [LS07], page 150, we have

$$(B_K^\dagger)_c = \Big\{ \sum_{i,j<0} a_{i,j} \frac{Y^i}{\Delta(Y)^j} \ : \ a_{i,j} \in K;\ \forall \eta > 1, \lim_{|(i,j)|\to\infty} |a_{i,j}|\eta^{|(i,j)|} = 0 \Big\}.$$

Since $\dim(C) = 1$, then $(A_K^\dagger)_c \otimes_{B_K^\dagger} \Omega^2_{B_K^\dagger} = 0$ and so we get

$$H^0_{MW,c}(U_k/K) = 0;$$
$$H^1_{MW,c}(U_k/K) = \frac{\mathrm{Ker}\left[(A_K^\dagger)_c \to (A_K^\dagger)_c \otimes_{B_K^\dagger} \Omega^1_{B_K^\dagger}\right]}{\mathrm{Im}\left[0 \to (A_K^\dagger)_c\right]} = \mathrm{Ker}\left[(A_K^\dagger)_c \to (A_K^\dagger)_c \otimes_{B_K^\dagger} \Omega^1_{B_K^\dagger}\right];$$

$$H^2_{MW,c}(U_k/K) = \frac{\mathrm{Ker}\left[(A^\dagger_K)_c \otimes_{B^\dagger_K} \Omega^1_{B^\dagger_K} \to 0\right]}{\mathrm{Im}\left[(A^\dagger_K)_c \to (A^\dagger_K)_c \otimes_{B^\dagger_K} \Omega^1_{B^\dagger_K}\right]} \simeq K.$$

So the only non-trivial space is $H^1_{MW,c}(U_k/K)$ and, to get it, we have to construct $(A^\dagger_K)_c$, $(A^\dagger_K)_c \otimes_{B^\dagger_K} \Omega^1_{B^\dagger_K}$, the connection $\nabla_c : (A^\dagger_K)_c \to (A^\dagger_K)_c \otimes_{B^\dagger_K} \Omega^1_{B^\dagger_K}$ and finally find $\mathrm{Ker}(\nabla_c)$.

By (A.3), $(A^\dagger_K)_c$ is constructed via $(B^\dagger_K)_c$ which can be explicitly written as a direct sum over the branch locus Λ (see [Cha07], §3.2.1) In the case of hyperelliptic curve it's easy to impose that the points in branch locus are rationals, thus obtaining an explicit construction for $(B^\dagger_K)_c$ and consequently for $(A^\dagger_K)_c$. In fact, if C_k is the affine model of a smooth genus g hyperelliptic curve over k given by the equation $Y^2 = \prod_{i=1}^{2g+1}(X - \overline{\lambda}_i)$, where $\overline{\lambda}_i \in k$ are distinct roots, then the brach locus in characteristic 0 is just $\Lambda = \{\lambda_1, \ldots, \lambda_{2g+1}, \infty\}$, where $\lambda_i \in \mathcal{V}$ lifts $\overline{\lambda}_i$, for every $1 \leq i \leq 2g+1$. According to [CL09], §2.2-2.3, we have that an element of $H^1_{MW,c}(U_k/K)$ can be therefore written as a vector $(m_{\lambda_1}, \ldots, m_{\lambda_{2g+1}}, m_{\lambda_\infty})$ where

$$m_\lambda = \sum_{j=0,1} Y^j \otimes \sum_{i=0}^{\infty} b^\lambda_{j,i}(t-\lambda)^i,$$

with $b^\lambda_{j,i} \in K$ and $b^\infty_{j,i} = 0$. From this construction, the explicit computation of the trace of the Frobenius follows.

A.2.2 Generalization

Our idea is to generalize Chatel-Lubicz's construction to trigonal curves. More precisely, we take a quartic $C_k \subseteq \mathbb{P}^2_k$ and we consider the projection of C_k from one of its points, say P_k, to a line, say L_k. Since C_k is a quartic and we project from one of its points, the projection is, outside the ramification points, a 3:1-map from C_k to $L_k \simeq \mathbb{P}^1_k$, that shows the trigonality of the curve.

Let k be a finite field of characteristic $p \neq 2, 3$ and let $C_k \subseteq \mathbb{P}^2_k$ be a quartic defined by the equation

$$Y^4 = f_k(X, Z),$$

with $f_k(X, Z) \in k[X, Y]$ homogeneous of degree 4, that is

$$f_k(X, Z) = aX^4 + bX^3Z + cX^2Z^2 + dXZ^3 + eZ^4.$$

If $e = 0 = a$, than C_k would be defined by the equation $Y^4 = bX^3Z + cX^2Z^2 + dXZ^3 = XZ(bX^2 + cXZ + dZ^2)$ which is a degenerate case, so we must have $a \neq 0$ or $e \neq 0$. Since $f_k(X, Z)$ is symmetric with respect to X and Z, we can suppose $e \neq 0$, without any loss of generality.

The next step is fixing the projection point. Suppose that e is such that there exists $\xi \in k$ with $\xi^4 = e$, therefore $[0 : \xi : 1] \in C_k$. Let $P_k = [0 : \xi : 1]$ be the projection point and $L_k : Y = 0$ be the projection line.

Now we have to define the projection map π_k from C_k to L_k via P_k. For every $Q = [x : y : z] \in C_k \setminus \{P_k\}$, the line of $\mathbb{P}^2_k$ passing through P_k and Q is given by the equation $\ell : \alpha X + \beta Y + \gamma Z = 0$, with $\alpha, \beta, \gamma \in k$ not all zero and such that

$$\left\{ \begin{array}{l} (\alpha X + \beta Y + \gamma Z)(P_k) = 0 \\ (\alpha X + \beta Y + \gamma Z)(Q) = 0 \end{array} \right. ; \quad \left\{ \begin{array}{l} \beta\xi + \gamma = 0 \\ \alpha x + \beta y + \gamma z = 0 \end{array} \right. ; \quad \left\{ \begin{array}{l} \gamma = -\beta\xi \\ \alpha x + \beta(y - \xi z) = 0 \end{array} \right. . \quad \text{(A.4)}$$

- If $y - \xi z \neq 0$, form the second equation in (A.4) we obtain $\beta = \frac{x}{\xi z - y}\alpha$ and so

$$\ell : X + \frac{x}{\xi z - y}(Y - \xi Z) = 0,$$

where we used the fact that $\alpha \neq 0$, otherwise we would have $\alpha = \beta = \gamma = 0$. Thus, if $Q = [x : y : z] \in C_k \setminus \{P_k\}$ such that $y \neq \xi z$, then

$$\pi_k(Q) = \ell \cap L_k = \left\{ [a : 0 : c] \in \mathbb{P}^2_k \ : \ a - \frac{\xi x}{\xi z - y} c = 0 \right\} = [\xi x : 0 : \xi z - y].$$

- If $y - \xi z = 0$, then $Q = [x : \xi z : z]$. We can assume $x \neq 0$, in fact otherwise $Q = [0 : \xi z : z]$ with $z \neq 0$, thus $Q = P_k$ against the hypothesis $Q \in C_k \setminus \{P_k\}$. So, from the second equation in (A.4), we get $\alpha = 0$. Hence

$$\ell : Y - \xi Z = 0,$$

where we used the fact that $\beta \neq 0$, otherwise we would have $\alpha = \beta = \gamma = 0$. Thus, if $Q = [x : \xi z : z] \in C_k \setminus \{P_k\}$, then

$$\pi_k(Q) = \ell \cap L = \left\{ [a : 0 : c] \in \mathbb{P}^2_k \ : \ -\xi c = 0 \right\} = [1 : 0 : 0].$$

We remark that $Q = [x : \xi z : z] \in C_k \setminus \{P_k\}$ if and only if $x \neq 0$ and

$$(\xi z)^4 = e z^4 = f_k(x, z) = a x^4 + b x^3 z + c x^2 z^2 + d x z^3 + e z^4,$$

that is

$$a x^3 + b x^2 z + c x z^2 + d z^3 = 0,$$

which has three, possibly coincident, solutions in x as function of a, b, c, d and z. That means that $[1 : 0 : 0]$ is the image of at most three different points of $C_k \setminus \{P_k\}$.

Summarizing, for every $Q = [x : y : z] \in C_k \setminus \{P_k\}$, both if $y = \xi z$ or $y \neq \xi z$, we have

$$\pi_k(Q) = [\xi x : 0 : \xi z - y].$$

Therefore

$$\begin{array}{rccccc} \pi_k : & C_k \setminus \{P_k\} & \to & L_k & \xrightarrow{\sim} & \mathbb{P}^1_k \\ & [x : y : z] & \mapsto & [\xi x : 0 : \xi z - y] & \mapsto & [\xi x : \xi z - y] \end{array},$$

with $y^4 = f_k(x, z)$. We remark that π_k is well-defined since if both $\xi x = 0$ and $y = \xi z$, then $[0 : \xi z : z] = P_k$.

Let us define the following open subsets of $\mathbb{P}^2_k$

$$W_0 = \{ [x : y : z] \in \mathbb{P}^2_k \ : \ z \neq 0 \} \qquad \text{and} \qquad W_1 = \{ [x : y : z] \in \mathbb{P}^2_k \ : \ x \neq 0 \}.$$

It's straightforward that $W_0 \cup W_1$ is an open covering of C_k, since $\mathbb{P}^2_k \setminus (W_0 \cup W_1) = [0 : 1 : 0]$ which is not a point of C_k.

Reduction to the affine coordinates: $Z \neq 0$

We now reduce ourself to the affine case, by setting $Z \neq 0$. We will denote $C_k \cap W_0$ by $C_{k,0}$ and, since $P_k \in W_0$, then $P_k \in C_{k,0}$. In this situation, the needed projection is

$$\begin{array}{rccccc} \pi_{k,0}: & C_{k,0} \setminus \{P_k\} & \to & \mathbb{P}^1_k & \to & \mathbb{A}^1_k \\ & (x,y) & \mapsto & [\xi x : \xi - y] & \mapsto & \begin{cases} \frac{\xi x}{\xi - y} & \text{if } y \neq \xi \\ P_\infty & \text{othewise} \end{cases} \end{array},$$

where $y^4 = f_k(x,1) = \tilde{f}(x) = ax^4 + bx^3 + cx^2 + dx + e$ and P_∞ is the point at infinity in $\mathbb{A}^1_k$. In the reduction to the affine coordinates, we will miss out P_∞, but the only thing we have to know is whether it is a branch point or not.

It's easy to see that P_∞ is in the brach locus of $\pi_{k,0}$ if and only if there exists $x \in k$ such that (x,ξ) is a point of $C_{k,0} \setminus \{P_k\}$ and it is a ramification point of $\pi_{k,0}$, that is $(x,\xi) \neq {}^3(0,\xi)$ is a multiple root of $Y^4 = \tilde{f}(X)$. Recalling that $\xi^4 = e$ and $x \neq 0$, it is equivalent to asking that $aX^3 + bX^2 + cX + d = 0$ has a multiple non-zero root in X. Therefore, P_∞ is in the brach locus of $\pi_{k,0}$ if and only if the determinant of $aX^3 + bX^2 + cX + d$ vanishes (*i.e.* $b^2c^2 + 4ac^3 - 4b^3d + 18abcd - 27a^2d^2 = 0$) and $d \neq 0$ or $c \neq 0$.

Let us now consider the following change of coordinates:

$$s = \frac{\xi x}{\xi - y}, \qquad t = \frac{y}{\xi},$$

we remark that $t \neq 1$, since $y \neq \xi$ in the affine setting. The inverse transformation is

$$y = \xi t, \qquad x = \frac{s(\xi - y)}{\xi} = s(1 - \frac{y}{\xi}) = s(1-t),$$

where $s, t \in k$ and $t \neq 1$. Thus

$$\begin{array}{rccc} \pi^{-1}_{k,0}: & C_{k,0} \setminus \{P_k\} & \leftarrow & \mathbb{A}^1_k \\ & \big(s(1-t), \xi t\big) & \leftarrow\!\shortmid & s \end{array},$$

with $t \neq 1$ such that

$$(\xi t)^4 = et^4 = \tilde{f}(s(1-t)) = as^4(1-t)^4 + bs^3(1-t)^3 + cs^2(1-t)^2 + ds(1-t) + e,$$

that is

$$as^4(1-t)^4 + bs^3(1-t)^3 + cs^2(1-t)^2 + ds(1-t) + e(1-t^4) = 0. \tag{A.5}$$

Since $1 - t^4 = (1+t^2)(1+t)(1-t)$ and $t \neq 1$, we can divide (A.5) by $t-1$, obtaining

$$as^4(1-t)^3 + bs^3(1-t)^2 + cs^2(1-t) + ds + e(1+t^2)(1+t) = 0.$$

Therefore $C_{k,0}$ is defined by $g(S,T) \in k[S,T]$, where

$$g(S,T) = aS^4(1-T)^3 + bS^3(1-T)^2 + cS^2(1-T) + dS + e(1+T^2)(1+T)$$

[3] We remark that $x \neq 0 \Leftrightarrow (x,\xi) \neq P_k$

$$= T^3(e - aS^4) + T^2(e + bS^3 + 3aS^4) + T(e - cS^2 - 2bS^3 - 3aS^4) + (e + dS + cS^2 + bS^3 + aS^4).$$

The next step is finding the branch locus $\Lambda_{k,0} \subseteq \mathbb{A}^1_k$ of $\pi_{k,0}$: $s \in \mathbb{A}^1_k$ is in the branch locus of $\pi_{k,0}$ if and only if there exists $t \in k \setminus \{1\}$ such that $(s(1-t), \xi t)$ is a point of $C_{k,0} \setminus \{P_k\}$ and it is a ramification point for $\pi_{k,0}$. That is, if and only if $g(s,T)$ has multiple non-1 roots as function of T, which is equivalent to asking whether $g(s,T)$ and $\frac{\partial g(s,T)}{\partial T}$ have a common zero, different from 1. Let

$$g(S,T)u(S,T) + \frac{\partial g(S,T)}{\partial T} v(S,T) = \Delta(S)$$

be the Bézout relation between $g(S,T)$ and $\frac{\partial g(S,T)}{\partial T}$ as elements of $(k[S])[T]$, with $\Delta(S) \in k[S]$ their resultant and $u(S,T), v(S,T) \in k[S,T]$ coprime in S.

Recalling that the characteristic of k is different from 2 or 3, we get $\frac{\partial g(S,T)}{\partial T} = 3T^2(e - aS^4) + 2T(e + bS^3 + 3aS^4) + (e - cS^2 - 2bS^3 - 3aS^4)$ and using the Extended Euclidean Algorithm we get[4]:

$$\begin{aligned}
\Delta(S) = {} & (-27a^3d^2 + 18a^2bcd - 4a^2c^3 - 4ab^3d + ab^2c^2)S^{14} \\
& + (-216a^3de + 72a^2bce - 16ab^3e)S^{13} \\
& + (-432a^3e^2 - 108a^2bde + 72a^2c^2e - 12ab^2ce)S^{12} \\
& + (-432a^2be^2 + 72a^2cde - 48ab^2de + 8abc^2e)S^{11} \\
& + (-144a^2ce^2 + 81a^2d^2e - 156ab^2e^2 - 36abcde + 8ac^3e + 4b^3de - b^2c^2e)S^{10} \\
& + (216a^2de^2 - 240abce^2 + 16b^3e^2)S^9 \\
& + (432a^2e^3 + 24abde^2 - 128ac^2e^2 + 12b^2ce^2)S^8 \\
& + (384abe^3 - 144acde^2 + 48b^2de^2 - 8bc^2e^2)S^7 \\
& + (96ace^3 - 81ad^2e^2 + 156b^2e^3 + 18bcde^2 - 4c^3e^2)S^6 \\
& + (-40ade^3 + 168bce^3)S^5 + (-16ae^4 + 84bde^3 + 56c^2e^3)S^4 \\
& + (48be^4 + 72cde^3)S^3 + (48ce^4 + 27d^2e^3)S^2 + 40de^4S + 16e^5.
\end{aligned}$$

Thus

$$\Lambda_{k,0} = \{\ s \in \mathbb{A}^1_k\ :\ \Delta(s) = 0\ \}.$$

As seen in Section A.2.1, to explicitly construct $H^i_{MW,c}(U_{k,0}/K)$, where $U_{k,0}$ is the étale locus of $\pi_{k,0}$, we need that $\Lambda_{k,0}$ is made of distinct rational points, but we cannot impose this condition without loss of generality. The same problem occurs if we consider the reduction in W_1 and in the general setting (we skip the details, since they are almost standard[5]).

We tried to overcome this obstacle, in many ways, for example by using *nondegenerate curves* (see, *e.g.*, Castryck et al. [CV09, CDV06] for a survey on nondegenerate curves and on their zeta function), but without success. We believe that the generalization we had in mind is achievable, but much more work is needed.

[4] We implemented a MAGMA program to perform this computation.

[5] We remark that we also tested the curve $C : X^4 + Y^4 - Z^4 = 0$, suggested by Michele Bolognesi, http://blogperso.univ-rennes1.fr/michele.bolognesi/.

Bibliography

[Apo74] T. Apostol, *Mathematical Analysis*, second ed., Addison-Wesley, 1974.

[Apo76] ———, *Introduction to Analytic Number Theory*, Springer, 1976.

[BCP97] J. Brüdern, R.J. Cook, and A. Perelli, *The values of binary linear forms at prime arguments*, Proc. of Sieve Methods, Exponential sums and their Application in Number Theory, Cambridge U.P., 1997, pp. 87–100.

[BD66] E. Bombieri and H. Davenport, *Small differences between prime numbers*, Proc. R. Math. Soc. **293** (1966), 1–18.

[CF01] R. J. Cook and A. Fox, *The values of ternary quadratic forms at prime arguments*, Mathematika **48** (2001), 137–149.

[CH06] R. J. Cook and G. Harman, *The values of additive forms at prime arguments*, Rocky Mountain J. Math. **36** (2006), 1153–1164.

[Che78] J. R. Chen, *On the Goldbach's problem and the sieve methods*, Sci. Sinica **21** (1978), 701–739.

[Coh] H. Cohen, *High precision computation of Hardy-Littlewood constants*, unpublished, `math.u-bordeaux1.fr/~cohen/hardylw.dvi`.

[Coh07] ———, *Number theory. Vol. II. Analytic and modern tools*, Springer, 2007.

[Dav00] H. Davenport, *Multiplicative Number Theory*, third ed., Springer, 2000.

[Dav05] ———, *Analytic methods for Diophantine equations and Diophantine inequalities*, second ed., Cambridge U. P., 2005.

[DH46] H. Davenport and H. Heilbronn, *On indefinite quadratic forms in five variables*, J. London Math. Soc. **21** (1946), 185–193.

[dlVP96] Ch. J. de la Vallée Poussin, *Analytical research on the theory of prime numbers. (Recherches analytiques sur la théorie des nombres premiers.)*, Ann. Soc. Sci. Bruxelles **20** (1896), 183–256, 281–362, 363–397.

[Erd35] P. Erdős, P.; Turán, *Über die Vereinfachung eines Landauschen Satzes. (On a simplification of a theorem of Landau.)*, Mitteil. Forsch.-Inst. Math. Mech. Univ. Tomsk **1** (1935), 144–147.

[Erd49] P. Erdős, *On a new method in elementary number theory which leads to an elementary proof of the prime number theorem*, Proc. Nat. Acad. Sci. U. S. A. **35** (1949), 374–384.

[FG95] J. B. Friedlander and D. A. Goldston, *Some singular series averages and the distribution of Goldbach numbers in short intervals*, Illinois J. Math. **39** (1995), 158–180.

[Gal70] P. X. Gallagher, *A large sieve density estimate near* $\sigma = 1$, Invent. Math. **11** (1970), 329–339.

[Gal75] ———, *Primes and Powers of* 2, Invent. Math. **29** (1975), 125–142.

[Gho81] A. Ghosh, *The distribution of* αp^2 *modulo* 1, Proc. London Math. Soc. **42** (1981), no. 2, 252–269.

[GS01] X. Gourdon and P. Sebah, *Some constants from number theory*, webpage available from `http://numbers.computation.free.fr/Constants/Miscellaneous/constantsNumTheory.html` (2001).

[Had96] J. Hadamard, *Sur la distribution des zéros de la fonction* $\zeta(s)$ *et ses conséquences arithmétiques*, Bull. Soc. Math. France **24** (1896), 199–220.

[Har04] G. Harman, *The values of ternary quadratic forms at prime arguments*, Mathematika **51** (2004), 83–96.

[HBP02] D. R. Heath-Brown and J.-C. Puchta, *Integers represented as a sum of primes and powers of two*, Asian J. Math. **6** (2002), 535–565.

[HL23a] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio numerorum'; III: On the expression of a number as a sum of primes*, Acta Math. **44** (1923), 1–70.

[HL23b] ———, *Some problems of 'Partitio numerorum'; V: A further contribution to the study of Goldbach's problem*, Proc. London Math. Soc. **22** (1923), 46–56.

[HR18] G. H. Hardy and S. Ramanujan, *Asymptotic formulæ in combinatory analysis*, Proc. London Math. Soc. (2) **17** (1918), 75–115, see also: Collected papers of Srinivasa Ramanujan, Chelsea, 2000, pp. 276–309.

[HR74] H. Halberstam and H.-E. Richert, *Sieve methods*, Academic Press, 1974.

[Hua38a] L. K. Hua, *On Waring's problem*, Quart. J. Math. Oxford **9** (1938), 199–202.

[Hua38b] ———, *Some results in the additive prime number theory*, Quart. J. Math. Oxford **9** (1938), 68–80.

[HW10] G. H. Hardy and E.M. Wright, *An introduction to the Theory of Numbers*, sixth ed., Oxford Science Publications, 2010.

[Ivi85] A. Ivić, *The Riemann Zeta-Function*, John Wiley and Sons, 1985.

[Jut77] M. Jutila, *On linnik's constant*, Math. Scand. **41** (1977), 45–62.

[Kor58a] N. M. Korobov, *Estimates of trigonometric sums and their applications*, Usp. Mat. Nauk **13** (1958), 185–192 (Russian).

[Kor58b] ———, *Weyl's sums estimates and the distribution of primes*, Dokl. Akad. Nauk SSSR **123** (1958), 28–31 (Russian).

[KP06] A. Khalfalah and J. Pintz, *On the representation of Goldbach numbers by a bounded number of powers of two*, Elementare und analytische Zahlentheorie, Franz Steiner Verlag, Stuttgart, 2006, pp. 129–142.

[Kum06] A.V. Kumchev, *On Weyl sums over primes and almost primes*, Michigan Math. J. **54** (2006), 243–268.

[Li06] H. Li, *Four prime squares and powers of* 2, Acta Arith. **125** (2006), 383–391.

[Lin51] Yu. V. Linnik, *Prime numbers and powers of two*, Trudy Mat. Inst. Steklov **38** (1951), 151–169, (Russian).

[Lin53] ———, *Addition of prime numbers with powers of one and the same number*, Mat. Sbornik **32** (1953), 3–60, (Russian).

[Liu04] T. Liu, *Representation of odd integers as the sum of one prime, two squares of primes and powers of 2*, Acta Arith. **115** (2004), 97–118.

[LL04] J. Liu and G. Lü, *Four squares of primes and 165 powers of 2*, Acta Arith. **114** (2004), 55–70.

[LLZ99] J. Liu, M.-C. Liu, and T. Zhan, *Squares of primes and powers of* 2, Monatsh. Math. **128** (1999), 283–313.

[LPZ07] A. Languasco, J. Pintz, and A. Zaccagnini, *On the sum of two primes and k powers of two*, Bull. London Math. Soc. **39** (2007), 771–780, `http://dx.doi.org/10.1112/blms/bdm062`.

[LS09] G. Lü and H. Sun, *Integers represented as the sum of one prime, two squares of primes and powers of 2*, Proc. Amer. Math. Soc. **137** (2009), 1185–1191.

[LS11] A. Languasco and V. Settimi, *On a diophantine problem with one prime, two squares of primes and s powers of two*, preprint (2011), `http://arxiv.org/abs/1103.1985`.

[LW05] W. P. Li and T. Z. Wang, *Diophantine approximation by a prime, squares of two primes and powers of two*, Pure Appl. Math. (Xi'an) **21** (2005), 295–299, (Chinese).

[LW07] ———, *Diophantine approximation with four squares of primes and powers of two*, Chinese Quart. J. Math. **22** (2007), 166–174.

[LZ07] A. Languasco and A. Zaccagnini, *A note on Mertens' formula for arithmetic progressions*, J. Number Theory **127** (2007), 37–46.

[LZ10] ———, *On a Diophantine problem with two primes and s powers of two*, Acta Arith. **145** (2010), 193–208, http://journals.impan.gov.pl/cgi-bin/aa/pdf?aa145-2-07.

[Mon71] H. L. Montgomery, *Topics in multiplicative number theory*, Lecture Notes in Mathematics, vol. 227, Springer, 1971.

[MRS96] M. R. Murty, M. Rosen, and J. H. Silverman, *Variations on a theme of Romanoff*, Internat. J. Math. **7** (1996), 373–391.

[MV75] H. L. Montgomery and R. C. Vaughan, *The exceptional set in Goldbach's problem*, Acta Arith. **27** (1975), 353–370.

[MV07] ———, *Multiplicative number theory. I. Classical theory*, Cambridge U. P., 2007.

[Par03] S. T. Parsell, *Diophantine approximation with primes and powers of two*, New York J. Math. **9** (2003), 363–371.

[Pen01] T. P. Peneva, *On the exceptional set for Goldbach's problem in short intervals*, Monatsh. Math. **132** (2001), no. 1, 49–65.

[Pen04] ———, *Corrigendum: "On the exceptional set for Goldbach's problem in short intervals"*, Monatsh. Math. **141** (2004), no. 3, 209–217.

[Pin76] J. Pintz, *Elementary methods in the theory of L-functions. II. On the greatest real zero of a real L-function*, Acta Arith. **31** (1976), 273–289.

[Pin06a] ———, *A note on Romanov's constant*, Acta Math. Hungar. **112** (2006), 1–14.

[Pin06b] ———, *Recent results on the Goldbach conjecture*, Elementare und analytische Zahlentheorie, Schr. Wiss. Ges. Johann Wolfgang Goethe Univ. Frankfurt am Main, 20, Franz Steiner Verlag, Stuttgart, 2006, pp. 220–254.

[Pin09] ———, *Landau's problems on primes*, J. Théor. Nombres Bordeaux **21** (2009), 357–404.

[PR03] J. Pintz and I. Z. Ruzsa, *On Linnik's approximation to Goldbach's problem, I*, Acta Arith. **109** (2003), 169–194.

[Pra78] K. Prachar, *Primzahlverteilung*, second ed., Springer, 1978.

[Rie68] G. J. Rieger, *Uber die Summe aus einem Quadrat und einem Primzahlquadrat*, J. Reine Angew. Math. **231** (1968), 89–100.

[Rom34] N. P. Romanov, *Über einige Sätze der additiven Zahlentheorie*, Math. Ann. **109** (1934), 668–678.

[RS62] J. B. Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois J. Math. **6** (1962), 64–94.

[Sel49] A. Selberg, *An elementary proof of the prime-number theorem*, Ann. of Math. (2) **50** (1949), 305–313.

[SP11] P. Solé and M. Planat, *Extreme values of the Dedekind ψ function*, Journal of Combinatorics and Number Theory **3** (2011), no. 1, 1–6.

[SV77] B. Saffari and R. C. Vaughan, *On the fractional parts of x/n and related sequences. II*, Ann. Inst. Fourier **27** (1977), 1–30.

[The10] The PARI Group, Bordeaux, *PARI/GP, version* `2.3.5`, 2010, available from `http://pari.math.u-bordeaux.fr/`.

[Vau74] R. C. Vaughan, *Diophantine approximation by prime numbers. I*, Proc. London Math. Soc. **28** (1974), 373–384.

[Vau97] ———, *The Hardy-Littlewood method*, second ed., Cambridge U. P., 1997.

[Vin37] I. M. Vinogradov, *The representation of an odd number as a sum of three primes*, Dokl. Akad. Nauk SSSR **15** (1937), 169–172, (in Russian).

[Vin58] ———, *A new estimate of the function $\zeta(1+it)$*, Izv. Akad. Nauk SSSR. Ser. Mat. **22** (1958), 161–164.

[Vin04] ———, *The method of trigonometrical sums in the theory of numbers*, Dover, 2004.

Appendix Bibliography

[Ber86] P. Berthelot, *Géométrie rigide et cohomologie des variétés algébriques de caractéristique p*, Mém. Soc. Math. France (N.S.) **3** (1986), no. 23, 7–32.

[CDV06] W. Castryck, J. Denef, and F. Vercauteren, *Computing zeta functions of nondegenerate curves*, Int. Math. Res. Pap. **72017** (2006).

[CFA+06] H. Cohen, G Frey, R. Avanzi, C. Doche, T. Lange, K. Nguyen, and F. Vercauteren (eds.), *Handbook of elliptic and hyperelliptic curve cryptography*, Discrete Mathematics and its Applications, Chapman & Hall/CRC, Boca Raton, FL, 2006.

[Cha07] G. Chatel, *Comptage de points : application des méthodes cristallines*, PhD Thesis (2007).

[CL09] G. Chatel and D. Lubicz, *A point counting algorithm using cohomology with compact support*, LMS J. Comput. Math. **12** (2009), 295–325.

[CV09] W. Castryck and J. Voight, *On nondegeneracy of curves*, Algebra Number Theory **3** (2009), no. 3, 255–281.

[Del74] P. Deligne, *La conjecture de Weil. I*, Inst. Hautes Études Sci. Publ. Math. (1974), no. 43, 273–307.

[DH76] W. Diffie and M. E. Hellman, *New directions in cryptography*, IEEE Trans. Information Theory **22** (1976), no. 6, 644–654.

[DV06a] J. Denef and F. Vercauteren, *Counting points on C_{ab} curves using Monsky-Washnitzer cohomology*, Finite Fields Appl. **12** (2006), no. 1, 78–102.

[DV06b] ———, *An extension of Kedlaya's algorithm to hyperelliptic curves in characteristic 2*, J. Cryptology **19** (2006), no. 1, 1–25.

[Dwo60] B. Dwork, *On the rationality of the zeta function of an algebraic variety*, Amer. J. Math. **82** (1960), 631–648.

[Dwo63] ———, *A deformation theory for the zeta function of a hypersurface*, Proc. Internat. Congr. Mathematicians (Stockholm, 1962), Inst. Mittag-Leffler, Djursholm, 1963, pp. 247–259.

[Elk73] R. Elkik, *Solutions d'équations à coefficients dans un anneau hensélien*, Ann. Sci. École Norm. Sup. (4) **6** (1973), 553–603 (1974).

[Elk98] N. D. Elkies, *Elliptic and modular curves over finite fields and related computational issues*, Computational perspectives on number theory, AMS/IP Stud. Adv. Math., vol. 7, Amer. Math. Soc., Providence, RI, 1998, pp. 21–76.

[GG01] P. Gaudry and N. Gürel, *An extension of Kedlaya's point-counting algorithm to superelliptic curves*, Advances in cryptology, Lecture Notes in Comput. Sci., vol. 2248, Springer, Berlin, 2001, pp. 480–494.

[GGK+68] J. Giraud, A. Grothendieck, S. L. Kleiman, M. Raynaud, and J. Tate, *Dix exposés sur la cohomologie des schémas*, Advanced Studies in Pure Mathematics, Vol. 3, North-Holland Publishing Co., Amsterdam, 1968.

[GH81] M. J. Greenberg and J. R. Harper, *Algebraic topology*, Mathematics Lecture Note Series, vol. 58, Benjamin/Cummings Publishing Co., Reading, Mass., 1981.

[Gro65] A. Grothendieck, *Formule de Lefschetz et rationalité des fonctions L*, Séminaire Bourbaki, vol. 9, Soc. Math. France, Paris, 1965, pp. 41–55.

[Har77] R. Hartshorne, *Algebraic geometry*, Springer-Verlag, New York, 1977.

[Ked01] K. S. Kedlaya, *Counting points on hyperelliptic curves using Monsky-Washnitzer cohomology*, J. Ramanujan Math. Soc. **16** (2001), 323–338.

[Lau04] A. G. B. Lauder, *Deformation theory and the computation of zeta functions*, Proc. London Math. Soc. (3) **88** (2004), 565–602.

[Lau06] ———, *A recursive method for computing zeta functions of varieties*, J. Comput. Math. **9** (2006), 222–269.

[LS07] B. Le Stum, *Rigid cohomology*, Cambridge Tracts in Mathematics, vol. 172, Cambridge U. P., Cambridge, 2007.

[Mil80] J. S. Milne, *Étale cohomology*, Princeton Mathematical Series, vol. 33, Princeton U. P., 1980.

[MW68] P. Monsky and G. Washnitzer, *Formal cohomology. I*, Ann. of Math. (2) **88** (1968), 181–217.

[Pil90] J. Pila, *Frobenius maps of abelian varieties and finding roots of unity in finite fields*, Math. Comp. **55** (1990), no. 192, 745–763.

[Sat00] T. Satoh, *The canonical lift of an ordinary elliptic curve over a finite field and its point counting*, J. Ramanujan Math. Soc. **15** (2000), no. 4, 247–270.

[Sch85] R. Schoof, *Elliptic curves over finite fields and the computation of square roots mod p*, Math. Comp. **44** (1985), no. 170, 483–494.

[Ser79] J.-P. Serre, *Local fields*, Graduate Texts in Mathematics, vol. 67, Springer-Verlag, 1979, Translated from the French by M. J. Greenberg.

[Tsu99] N. Tsuzuki, *On the Gysin isomorphism of rigid cohomology*, Hiroshima Math. J. **29** (1999), no. 3, 479–527.

[vdP86] M. van der Put, *The cohomology of Monsky and Washnitzer*, Mém. Soc. Math. France (N.S.) **23** (1986), 33–59.

[Wei49] A. Weil, *Numbers of solutions of equations in finite fields*, Bull. Amer. Math. Soc. **55** (1949), 497–508.