The Institute of Mathematical Sciences

Long chains of integers with few prime factors

Olivier Ramaré

IMSc Lecture Notes Series, no 12

The *IMSc Lecture Notes Series* is a publication of the Institute of Mathematical Sciences of Chennai. The scope of these publication ....

For more informations, see
`http://IMSC...`

Preface

This book is an elaboration on lectures given at the Institute of Mathematical Science in 2011. There was at that time a special year in number theory at IMSc and the courses were designed for students with limited knowledge of the theory at hand, but who were ready to invest enough energy. In this precise volume, the trek is roughly as follows. The first landmark is Mertens Theorem; from there we show how to compute averages of non-negative multiplicative functions; the next landmark we reach is the Brun-Titchmarsh Theorem via the Selberg sieve; after exploring the area, we move forward and derive a family of envelopping sieves that we hitherto employ to obtain results on the prime κ-tuple conjecture. In eight lectures, the reader is thus supposed to go from sea level to proving that, for each admissible 8-tuple $(h_1, h_2, \cdots, h_8)$, there are infinitely many integers n such that the product $(n + h_1)(n + h_2)\cdots(n + h_8)$ has at most 24 prime factors. It is somewhat illusory to believe this journey to be a refreshing stroll! The reader may however quit at anytime and, hopefully, will have acquired some bases in this aspect of analytic number theory. We have further tried to manage some more less demanding steps.

Olivier Ramaré

A video of this course can be found at
`http://www.imsc.res.in/conference_videos`

Key words and phrases. Large sieve inequality, Selberg sieve

2000 *Mathematics Subject Classification.* 11N, 11M20, 11P32

CNRS, Lab. Paul Painlevé, Université Lille 1, France

Contents

Preface iii

Table of content viii

Introduction 7

Chapter 1. Setting the background 9
- 1.1. Smooth summands 9
- 1.2. Multiplicative summands 11
- 1.3. Summing over primes 13
- 1.4. Handling coprimality conditions 15
- 1.5. A Lemma of van Lint & Richert 15

Chapter 2. Introducing the Selberg sieve 17
- 2.1. An initial estimate 17
- 2.2. Inversion formulas 17
- 2.3. The Brun-Titchmarsh inequality 19

Chapter 3. The Levin-Fainleib Theorem et alia 23
- 3.1. The Levin-Fainleib Theorem 23
- 3.2. A simple general inequality 28
- 3.3. A further consequence 29

Chapter 4. Some more exercises 31

Chapter 5. Introducing the large sieve 35
- 5.1. A hermitian tool 35
- 5.2. A pinch of number theory 37
- 5.3. Proof of the Brun-Titchmarsh inequality 37
- 5.4. Some historical comments and two open problems 39
- 5.5. Admissible tuples 40

Chapter 6. Some geometric considerations 43
- 6.1. Compact sets 43
- 6.2. Examples of compact sets 44
- 6.3. A family of arithmetical functions 45
- 6.4. Bordering system associated to a compact set 46

6.5. The compact relative to a κ-tuple of admissible shifts 47

Chapter 7. The Selberg sieve 49
7.1. A definition of sieve problems 49
7.2. An extremal problem 50
7.3. A general expression 52
7.4. On the number of prime twins 53
7.5. On a subset of prime twins 54
7.6. Computing the G-functions in the case of the prime κ-tuples 55

Chapter 8. Introduction to the weighted sieve 57
8.1. A beginner's stroll with historical comments 57
8.2. Results to be proven 59
Terminology 61

Chapter 9. The approach in the large 63
9.1. A general framework 63
9.2. Generalisation of a formula of Bombieri 65

Chapter 10. Another family of Selberg coefficients 67
10.1. Investigating the Selberg coefficients 67
10.2. Modifying the Selberg coefficients 69
10.3. Resuming the main proof 70

Chapter 11. Two reduction steps 73

Chapter 12. Specialisation of the proof 77
12.1. Final estimate 79
12.2. Conclusion 79
12.3. A first choice. Proof of Theorem 8.1 80

Chapter 13. Special computations for bounded values 83
13.1. Decomposition in affine pieces 83
13.2. Formulae to evaluate of I_1 and K_1 84
13.3. Formulae to evaluate of I_2 84
13.4. Results 85

Chapter 14. The GP-script 89
14.1. Precomputing Σ 89
14.2. Computing I_1, I_2 and K_1 90
14.3. Optimizing 93
14.4. Plots and other extraction of datas 96

Appendix A. On the convolution method 99
A.1. Introduction 99
A.2. Arithmetic functions 100
A.3. Dirichlet series 103

A.4. Summation by parts 106
A.5. Proof of Theorem A.1 107
A.6. Some digressions without proof 110

Notation 113

Bibliography 115

Index 121

Introduction

We will cover rapidly the ground from elementary number theory to the theory of the weighted sieve based on the Selberg sieve.

Here is an idea of the order different subjects will be introduced. The first four lectures are to be acquainted with the sieve in general. The last four ones present things in a more involved setting and present a weighted sieve.

(1) General sums : smooth summands, multiplicative summands, prime summands.
(2) Inversion formulae, van Lint & Richert Theorem, the Brun-Tichmarsh via the Selberg sieve.
(3) The Levin-Fainleib Theorem and remarks on primes in arithmetic progressions. The Brun-Titchmarsh inequality via hermitian inequalities.
(4) Computing G-functions and exercise session.
(5) Compact sets, inversion formulae. A general Selberg sieve bound. λ_d, $\lambda_d^\sharp$, a general van Lint-Richert Lemma, λ_d, $\lambda_d^\sharp$. Some problems as examples.
(6) The prime κ-tuple problem and an identity of Bombieri.
(7) A smoothed version of the Selberg λ_d. A general van Lint-Richert Lemma. The prime κ-tuple problem: the first two reduction steps.
(8) Final proof.

Theorem 0.1 (Diamond & Halberstam, 2008) *There are infinitely many integers n for which the product $n(n+2)(n+6)$ has at most 8 prime factors.*

The reader will find this result in table 11.1 of [**19**]. [check [**18**]].

Theorem 0.2 *There are infinitely many integers n for which the product*

$$\prod_{\ell\in\{0,1,3,4,6,9,10,14\}} (n+2\ell)$$

has at most 24 prime factors.

This result is contained in the course the author gave at HRI in january 2009, and is not yet published. As it turns out, C. Franze has recently improved on this result and the bound 23 is now available.

The appendix should contain a translation of a paper exposing the convolution method as well as the use of (convergent) Dirichlet series as generating series for multiplicative functions.

CHAPTER 1

Setting the background

The main objects in analytic number theory often look like

$$\sum_{n\le X} a(n)$$

for some function of "arithmetical" nature $a(n)$, where the adjective "arithmetical" needs to be defined. We review in this first lecture what the reader ought to know, and some more!

1.1. Smooth summands

This is the case when a is C^1. A hidden hypothesis we will comment later on is that a' is bounded. A first example of this situation is

$$\sum_{n\le X}\frac{1}{n} = \operatorname{Log} X + \gamma + \mathcal{O}(1/X), \quad (X \ge 1). \tag{1.1}$$

An even simpler example is given by

$$\sum_{n\le X} 1 = X + \mathcal{O}(1), \quad (X \ge 1). \tag{1.2}$$

We leave (1.1) to the reader and prove here that

$$\sum_{n\le X}\frac{\operatorname{Log} n}{n} = \tfrac12 \operatorname{Log}^2 X + \gamma_1 + \mathcal{O}(\operatorname{Log}(2X)/X), \quad (X \ge 1) \tag{1.3}$$

for some constant γ_1 that is called the Laurent-Stieltjes constant of index 1.

A preliminary remark on uniformity: In all three previous estimates, we have written "$X \ge 1$" while the estimate is most interesting when X is large. However, we need an estimate that is uniform in some range, and, for instance here, there exists a constant C such that, for any $X \ge 1$, we have

$$\left|\sum_{n\le X}\frac{\operatorname{Log} n}{n} - \tfrac12 \operatorname{Log}^2 X - \gamma_1\right| \le C\ \operatorname{Log}(2X)/X.$$

This would be **false** if we had written $+\mathcal{O}((\operatorname{Log} X)/X)$ in (1.3), for it cannot hold when $X = 1$. Such problems are usually trivial to sort, but a slip at this level may lead to mighty mistakes later on. There is nothing magic in the "Log$(2X)$" and we may as well have written "$1+\operatorname{Log} X$" or "Log$(3X)$".

Proof We simply write

$$\frac{\operatorname{Log} n}{n}=\frac{\operatorname{Log} X}{X}+\int_n^X\frac{\operatorname{Log} t-1}{t^2}dt.$$

This gives us

$$\begin{aligned}\sum_{n\le X}\frac{\operatorname{Log} n}{n}&=[X]\frac{\operatorname{Log} X}{X}+\sum_{n\le X}\int_n^X\frac{\operatorname{Log} t-1}{t^2}dt\\&=[X]\frac{\operatorname{Log} X}{X}+\int_1^X\Big(\sum_{n\le t}1\Big)\frac{\operatorname{Log} t-1}{t^2}dt\end{aligned}$$

where $[X]$ denotes the integer part of X. We continue by using (1.2) in the form $[t]=t-\{t\}$ ($\{t\}$ being the fractionnal part of t):

$$\sum_{n\le X}\frac{\operatorname{Log} n}{n}=\int_1^X\frac{\operatorname{Log} t-1}{t}dt+\operatorname{Log} X-\int_1^\infty\{t\}\frac{\operatorname{Log} t-1}{t^2}dt+\mathcal{O}\Big(\frac{\operatorname{Log}(2X)}{X}\Big).$$

and (1.3) follows readily. □

The technique we have developped in the above proof is known as *summation by parts*. The reader will find different versions of this, usually more intricate than the one above, relying either on Abel summation process or on Stieltjes integration. We have relied on (1.2), but see exercise 1.6 for a more general usage. We recommend to the reader the following two exercises.

Exercise 1.1 *Show that*

$$\sum_{n\le X}\operatorname{Log} n=X\operatorname{Log} X-X+\mathcal{O}(\operatorname{Log}(3X)),\quad (X\ge 1).$$

Exercise 1.2 *Show that*

$$\sum_{n\le X}\Big(\operatorname{Log}\frac{X}{n}\Big)^2\ll X,\quad (X\to\infty).$$

This case is thus well-understood. If we want to gain precision in the error term, then we appeal to the Euler-MacLaurin summation formula, but as the reader will see by analysing the example we treated, there is no way one can avoid fractionnal parts in the development. Note however that

- We do not know how to evaluate $\sum_{n\le X}n^{it}$ with enough precision when t is large with respect to X.

- The error term in (1.2) (i.e. the fractionnal part) is much more important than it looks. In our proofs, we want very often to show that the resulting error term is very small but, if it simply did not exist, then we would have $\zeta(s) = 1/(s-1)$. This implies that this error term is responsible for the functionnal equation of the Riemann zeta function as well as for its Euler-product!

1.2. Multiplicative summands

Let us start with a definition.

Definition 1.3 *A function a on the positive integers is said to be* multiplicative *when $a(1) = 1$ and $a(mn) = a(m)a(n)$ whenever m and n are coprime positive integers.*

As a result, if we decompose the integer n in prime powers, $n = \prod_i p_i^{\ell_i}$ say, then $a(n) = \prod_i a\big(p_i^{\ell_i}\big)$. Reciproqually, given any (double) sequence of values $b(p,\ell)$, the function defined by

$$a(n) = \prod_i b(p_i, \ell_i)$$

is indeed multiplicative. For example the function $a(n) = \phi(n)/n$ is multiplicative and we have

$$\frac{\phi(n)}{n} = \prod_{p|n}\Big(1 - \frac{1}{p}\Big).$$

The function $a(n) = 1$ is also multiplicative! Slightly more difficult: the function $a(n) = \mu^2(n)$ that takes the value 1 when n is *squarefree* (i.e. the only positive integer square that divides n is 1) and 0 otherwise is multiplicative. It is denoted by μ^2 because it is indeed the square of the *Moebius function* defined as follows:

(1.4)
$$\mu(n) = \begin{cases} 1 & \text{when } n = 1, \\ (-1)^r & \text{when } n = p_1 \cdots p_r, \text{ all the } p_i \text{ being prime and distinct,} \\ 0 & \text{otherwise.} \end{cases}$$

This Moebius function is also multiplicative. The reader may want to show that the functions $a(n) = n/\phi(n)$ and $a(n) = 2^{\omega(n)}$ (where $\omega(n)$ denotes the number of prime factors of n counted without multiplicity) are also multiplicative.

When dealing with multiplicativity, the following Lemma is often useful.

Lemma 1.4 *For any positive integer d, we denote by $\mathscr{D}(d)$ the set of all its (positive) divisors. When d_1 and d_2 are positive integers that are* *coprime, the map*

$$\begin{aligned}\mathscr{D}(d_1)\times\mathscr{D}(d_2)&\to\mathscr{D}(d_1d_2),\\(q_1,q_2)&\mapsto q_1q_2\end{aligned}$$

is one-to-one and onto.

Proof We simply mention that the map

$$\begin{aligned}\mathscr{D}(d_1d_2)&\to\mathscr{D}(d_1)\times\mathscr{D}(d_2),\\q&\mapsto(\gcd(q,d_1),\gcd(q,d_2))\end{aligned}$$

is the inverse of the one given in the Lemma. □

Let us proceed to evaluate

$$\sum_{n\le X}\frac{\phi(n)}{n}$$

We write

$$\frac{\phi(n)}{n}=\sum_{d|n}\frac{\mu(d)}{d}.$$

Indeed the RHS is easily proved to be multiplicative on using Lemma 1.4 and equals the LHS on prime powers. Once one has obtained this expression, the computations are straightforward:

$$\sum_{n\le X}\frac{\phi(n)}{n}=\sum_{n\le X}\sum_{d|n}\frac{\mu(d)}{d}=\sum_{d\le X}\frac{\mu(d)}{d}\sum_{\substack{n\le X,\\ d|n}}1.$$

This is a step you will see very often in analytic number theory: *the exchange of summation.* The inner sum is most easily estimated as $(X/d)+\mathcal{O}(1)$, and since $|\mu(d)|\le 1$, this gives us

$$\sum_{n\le X}\frac{\phi(n)}{n}=X\sum_{d\le X}\frac{\mu(d)}{d^2}+\mathcal{O}(\mathrm{Log}(2X))=X\sum_{d\ge1}\frac{\mu(d)}{d^2}+\mathcal{O}(\mathrm{Log}(2X))$$

which amounts to

$$\sum_{n\le X}\frac{\phi(n)}{n}=\tfrac{6}{\pi^2}X+\mathcal{O}(\mathrm{Log}(2X)).\tag{1.5}$$

This result is most satisfying, since the error term is very small. Determining whether this remainder term can or not be improved upon is a difficult task that we leave to the reader, and to which we do not know the answer. Here are some references concerning this very precise problem: [**92**, chapter 4], [**67**],[**68**], [**58**]

The technique we have developped in the above proof is by no means accidental and is called the *convolution method.* We detail it in the appendix. We will present later a more general result to handle sums of positive multiplicative function, but that is far from the precision one can attain by using the convolution method.

Exercise 1.5 *Show that*

$$\mu^2(n) = \sum_{d^2|n} \mu(d)$$

and deduce an estimation of $\sum_{n\le X} \mu^2(n)$ *with an error term of order* $\sqrt{X}$.

Exercise 1.6 *On using the proof of* (1.3) *but replacing* (1.2) *by* (A.15)*, show that*

$$\sum_{n\le X} \frac{\phi(n)}{n}\frac{\operatorname{Log} n}{n} = C_1 \operatorname{Log}^2 X + C_2 \operatorname{Log} X + C_3 + \mathcal{O}\biggl(\frac{(\operatorname{Log}(2X))^2}{X}\biggr)$$

for some constants C_1*,* C_2 *and* C_3*.*

Let us end this section by commenting on why the study of averages of multiplicative functions is central in analytic number theory. Multiplicative functions are one of the main way to characterize the multiplicative structure. Our main problem is not the multiplicative structure but its link with the additive one and

1.3. Summing over primes

We consider here some simple sums over prime summands, in the spirit of Mertens in 1897. We will see in the next chapter that general Theorems on averages of multiplicative functions require such estimates. We prove first that

Theorem 1.7 (Mertens)

$$\sum_{p\le X} \frac{\operatorname{Log} p}{p} = \operatorname{Log} X + \mathcal{O}(1).$$

The proof uses two ingredient. We define first the van Mangoldt function by

$$\Lambda(n) = \begin{cases} \operatorname{Log} p & \text{when } n = p^\nu \text{ with } \nu \ge 1,\\ 0 & \text{otherwise.}\end{cases} \tag{1.6}$$

This function has in fact already been introduced by Riemann in his famous 1859 report [**78**], but has been named after van Mangoldt. This latter mathematician was working on trying to decipher Riemann's work and on providing proof to all his claims. The fundamental property we will use reads as follows:

$$\operatorname{Log} n = \sum_{d|n} \Lambda(d). \tag{1.7}$$

Proof We do not provide any proof by verify it on $n = p_1^2 p_2$. On the one hand, we have

$$\operatorname{Log} n = 2\operatorname{Log} p_1 + \operatorname{Log} p_2$$

while on the other hand:

$$\sum_{d|n} \Lambda(d) = \underset{d=p_1,}{\operatorname{Log} p_1} + \underset{d=p_1^2,}{\operatorname{Log} p_1} + \underset{d=p_2}{\operatorname{Log} p_2} \operatorname{Log} p_1$$

□

We first prove rapidly a Tchebyschef estimate.

Theorem 1.8 (Tchebyschef) *We have*

$$\sum_{n\le X} \Lambda(n) \ll X, \quad (X \ge 1).$$

Proof Indeed, on using (1.7), we find that (recall that $[t]$ is the fractionnal part of t)

$$\sum_{n\le X} \operatorname{Log} n = \sum_{m\le X} \Lambda(m)[X/m].$$

We use that for $X/2$ and find that

$$\sum_{n\le X} \operatorname{Log} n - 2\sum_{n\le X/2} \operatorname{Log} n = \sum_{m\le X} \Lambda(m)\big([X/m] - 2[X/(2m)]\big)$$

which we now analyse. The result of exercise 1.1 tells us that the LHS equals $X \operatorname{Log} 2 + \mathcal{O}(\operatorname{Log}(2X))$. On the RHS, the function $y \mapsto [y] - 2[y/2]$ is in fact periodical of period 2, and moreover non-negative. More precisely, it takes the value 0 on $[0, 1)$ and the value 1 on $[1, 2)$. As a consequence,

$$\sum_{X/2<m\le X} \Lambda(m) \le X \operatorname{Log} 2 + \mathcal{O}(\operatorname{Log}(2X)).$$

Thus there exists a constant C such that, for any $X \ge 1$ we have

$$\sum_{X/2<m\le X} \Lambda(m) \le C\,X.$$

We use that for X, then $X/2$, then $X/4$, etc, until $X/2^k$, where $X/2^{k+1} < 1$. This gives us

$$\sum_{m\le X} \Lambda(m) \le C\left(X + \frac{X}{2} + \frac{X}{4} + \cdots\right) \le 2C\,X.$$

The Lemma is proved. □

Exercise 1.9 *Prove that* $\sum_{p\le X} \operatorname{Log} p \ll X$ *for* $X \ge 1$.

Exercise 1.10 *Prove that* $\sum_{p\le X} 1 \ll X/\operatorname{Log}(2X)$ *for* $X \ge 1$.

Here is what we wanted to prove.

Theorem 1.11 (Mertens) *We have*

$$\sum_{p\le X}\frac{\operatorname{Log} p}{p}=\operatorname{Log} X+\mathcal{O}(1),\quad (X\ge 1).$$

Proof Let us use again directly the result of exercise 1.1, namely

$$\sum_{n\le X}\operatorname{Log} n=X\operatorname{Log} X-X+\mathcal{O}(\operatorname{Log}(3X)).$$

On the other side, on introducing (1.7), we find that

$$\begin{aligned}\sum_{n\le X}\operatorname{Log} n&=\sum_{\ell m\le X}\Lambda(m)=\sum_{m\le X}\Lambda(m)\biggl(\frac{X}{m}+\mathcal{O}(1)\biggr)\\&=X\sum_{m\le X}\frac{\Lambda(m)}{m}+\mathcal{O}(X)\end{aligned}$$

by Theorem 1.8. The Theorem follows, once the reader has checked that $\sum_{m\le X}\frac{\Lambda(m)}{m}$ and $\sum_{p\le X}\frac{\operatorname{Log} p}{p}$ differ by at most $\mathcal{O}(1)$. □

Exercise 1.12 *Deduce from the above Theorem that there exists two positive constants c_1 and c_2 such that $\sum_{c_1X<p\le X}1\ge c_2X$ for $X\ge 2$.*

This deduction is also due to Mertens in 1897 in [**?**].

1.4. Handling coprimality conditions

1.5. A Lemma of van Lint & Richert

We have seen in the previous section how to handle coprimality conditions. The condition $(n,d)=1$ may be somewhat difficult when d is large with respect to n and we prove here a general Lemma due to [**90**]. This Lemma will be the ground of several generalizations later on.

Let h be a multiplicative and non-negative function. We consider

$$H_d(D)=\sum_{\substack{n\le D,\\(n,d)=1}}\mu^2(n)h(n)\tag{1.8}$$

and we abreviate H_1 in H.

Lemma 1.13 *For every positive integer d and any positive real number D, we have*

$$H(dD)\ge\sum_{\delta|d}\mu^2(\delta)h(\delta)H_d(D)\ge H(D).$$

Proof We write

$$\begin{aligned} H(D) &= \sum_{\delta|d} \sum_{\substack{n\le D,\\ (n,d)=\delta}} \mu^2(n)h(n) = \sum_{\delta|d} \mu^2(\delta)h(\delta) \sum_{\substack{m\le D/\delta,\\ (m,d)=1}} \mu^2(m)h(m) \\ &= \sum_{\delta|d} \mu^2(\delta)h(\delta)H_d(D/\delta). \end{aligned}$$

We use $H_d(D/\delta) \ge H_d(D/d)$ to get the first inequality and $H_d(D/\delta) \le H_d(D)$ for the second one. □

CHAPTER 2

Introducing the Selberg sieve

2.1. An initial estimate

We prove here in an elementary manner that, for any $D \geq 1$, we have

$$\sum_{d \leq D} \frac{\mu^2(d)}{\phi(d)} \geq \operatorname{Log} D. \tag{2.1}$$

Proof Let us introduce the concept of squarefree kernel $k(n)$ of the integer n: it is simply the product of all the prime divisors of n.

When d is a squarefree integer, we find that

$$\begin{aligned}\frac{1}{\phi(d)} &= \frac{1}{d}\prod_{p|d}\frac{1}{1-\frac{1}{p}} = \frac{1}{d}\prod_{p|d}\Bigl(1+\frac{1}{p}+\frac{1}{p^2}+\cdots\Bigr)\\ &= \sum_{\substack{m\geq 1,\\ k(m)=d}} \frac{1}{m}.\end{aligned}$$

As a consequence, we find that

$$\sum_{d\leq D}\frac{\mu^2(d)}{\phi(d)} = \sum_{d\leq D}\sum_{\substack{m\geq 1,\\ k(m)=d}}\frac{1}{m} \geq \sum_{m\leq D}\frac{1}{m} \geq \operatorname{Log} D$$

and the Lemma readily follows. □

2.2. Inversion formulas

The Moebius function is very combinatorial in nature. It appears in several inversion formulae and we discuss here one that will be of use later. All these formulae stem from the identity:

$$\sum_{d|n}\mu(d) = \begin{cases} 1 & \text{when } n=1,\\ 0 & \text{when } n>1\end{cases} \tag{2.2}$$

whcih we prove in two steps: both sides define multiplicative functions, and these multiplicative functions are equal on any prime power.

Let us assume we are given a function f on the positive integers and a parameter X. We look at the function F defined over the positive integers by

$$F(\delta) = \sum_{\substack{n \le X, \\ \delta | n}} f(n) \tag{2.3}$$

with the idea of recovering f from F. It is possible because, when seen as a linear system, it is diagonal, with coefficients equal to 1 on the diagonal. This easy argument shows that F determines f uniquely. We want an explicit formula. We have

$$f(n) = \sum_{\substack{\ell \le X, \\ n | \ell}} \mu(\ell/n) F(\ell) \tag{2.4}$$

Proof Indeed, let us call $g(d)$ the RHS above. We have

$$\sum_{\substack{n \le X, \\ \delta | n}} g(n) = \sum_{\substack{n \le X, \\ \delta | n}} \sum_{\substack{\ell \le X, \\ n | \ell}} \mu(\ell/n) F(\ell) = \sum_{\ell \le X} F(\ell) \sum_{\substack{n | \ell, \\ \delta | n}} \mu(\ell/n).$$

The last sum equals also

$$\sum_{m | \ell/\delta} \mu((\ell/\delta)/m).$$

It vanishes when $\delta \neq \ell$ and takes value 1 otherwise. We thus have

$$\sum_{\substack{n \le X, \\ \delta | n}} g(n) = F(\delta).$$

We have however seen that this equation defines f uniquely, and thus $g = f$. □

Another form of this inversion comes when we want to determine the function f from

$$F(q) = \sum_{d | q} f(d). \tag{2.5}$$

Here also, solving for the values $(F(q))$ leads to a diagonal system with coefficients 1 on the diagonal, but we rather have an explicit expression for f. This is provided to us by the following formula:

$$f(d) = \sum_{q | d} \mu(q/d) F(q) \tag{2.6}$$

which can be established as above.

Exercise 2.1

1. *Show that, when d is squarefree, the number of couples (d_1, d_2) solution of $[d_1, d_2] = d$ is $3^{\omega(d)}$.*
2. *Let $f(q)$ be the number of couples (q_1, q_2) solution of $[q_1, q_2] = q$. Show that f is a multiplicative function.*
3. *Show that $\sum_{q|n} f(q)$ is the number of couples (q_1, q_2) such that $[q_1, q_2]|n$ and that this latter number is $\tau(n)^2$, the square of the number of divisors of n.*
4. *Deduce from the above a general expression for $f(q)$.*

2.3. The Brun-Titchmarsh inequality

This Theorem reads as follows:

Theorem 2.2 *Let $M \geq 0$ and $N > q \geq 1$ be given and let a be an invertible residue class modulo q. The number Z of primes in the interval $[M+1, M+N]$ lying in the residue class a modulo q verifies*

$$Z \leq \frac{2N}{\phi(q)\,\mathrm{Log}(N/q)}.$$

This neat and effective version is due to [**63**]. Earlier versions essentially had $2 + o(1)$ instead of simply 2. The name "Brun-Titchmarsh" Theorem stems from [**61**]. Indeed, Titchmarsh proved such a theorem for $q = 1$ with a $\mathrm{Log}\,\mathrm{Log}(N/q)$ term instead of the 2 to establish the asymptotic for the number of divisors of the $p+1$, p ranging through the primes, and he used the method of Brun. The constant 2 (with a $o(1)$) appeared for the first time in [**82**]. See Theorem 5.1 for some further comments.

Proof As already mentioned, we restruct our attention to the case $q = 1$. Let z be a parameter that we will choose later. Let us take an arbitrary sequence $(\lambda_d)_{d \leq z}$ where we however specify that $\lambda_1 = 1$. We start with the following inequality

$$\sum_{M+1 \leq p \leq M+N} 1 \leq \sum_{M+1 \leq n \leq M+N} \Bigl(\sum_{d|n} \lambda_d\Bigr)^2 + z.$$

Why is it so? If p is a prime number from the interval $[M+1, M+N]$ and $> z$, then the coefficient $\beta(n) = \Bigl(\sum_{d|n} \lambda_d\Bigr)^2$ is simply equal to 1. Otherwise it is non-negative. We take vare of the primes $\leq z$ by adding z, since this is surely an upper bound for their number. Let us define

$$Z = \sum_{M+1 \leq p \leq M+N} 1.$$

On expanding the square above and denoting the lcm of d_1 and d_2 by $[d_1, d_2]$, we find that

$$Z \leq \sum_{d_1, d_2 \leq z} \lambda_{d_1} \lambda_{d_2} \sum_{\substack{n \in [M+1, M+N], \\ [d_1, d_2] | n}} 1 + z.$$

The inner summation over n is readily evaluated:

$$\sum_{\substack{n \in [M+1, M+N], \\ [d_1, d_2] | n}} 1 = \frac{N}{[d_1, d_2]} + \mathcal{O}(1).$$

As a conclusion of our first step, we find that

$$Z \leq N Z_0 + \mathcal{O}\biggl(\Bigl(\sum_{d \leq z} |\lambda_d| \Bigr)^2 \biggr) + z$$

where

$$Z_0 = \sum_{d_1, d_2 \leq z} \frac{\lambda_{d_1} \lambda_{d_2}}{[d_1, d_2]} \tag{2.7}$$

is asking to be evaluated. We forget the term containing $\sum_{d \leq z} |\lambda_d|$ for the time being. Our choice of λ_d will ensure that $|\lambda_d| \leq 1$ and this will be enough to control our error term.

The quantity Z_0 is a bilinear form in the variables λ_d and we seek its minimum under the condition $\lambda_1 = 1$. This could be difficult but we have Selberg diagonalization at our disposal, which we explain now. First note that

$$\frac{1}{[d_1, d_2]} = \frac{(d_1, d_2)}{d_1 d_2}$$

and we now have to separate d_1 and d_2 in the main term. We do so via the identity

$$q = \sum_{\delta | q} \phi(\delta)$$

(to check this one, verify that both sides are multiplicative functions and compute that they take a same value on prime powers). We thus get

$$\frac{1}{[d_1, d_2]} = \frac{1}{d_1 d_2} \sum_{\substack{\delta | d_1, \\ \delta | d_2}} \phi(\delta)$$

where we have split the condition $\delta | (d_1, d_2)$ in $\delta | d_1$ and $\delta | d_2$. We have thus reach the expression

$$Z_0 = \sum_{\delta \leq z} \phi(\delta) y_\delta^2, \quad y_\delta = \sum_{\substack{d \leq z, \\ \delta | d}} \lambda_d / d. \tag{2.8}$$

We can use the previous section to recover λ_d from y_δ:

$$\frac{\lambda_d}{d} = \sum_{\substack{\delta \le z,\\ d|\delta}} \mu(\delta/d) y_\delta.$$

In particular, the variables y_δ are linked by the condition

$$1 = \sum_{\delta \le z} \mu(\delta) y_\delta.$$

Our aim is now to minimize the quadratic form given in (2.8). There are several ways to do it and we use Lagrange multipliers here. We consider the function

$$H(y_1, \ldots, y_z, u) = \sum_{\delta \le z} \phi(\delta) y_\delta^2 - u\Big(\sum_{\delta \le z} \mu(\delta) y_\delta - 1\Big)$$

and we say that the minimum will be reached when

$$\frac{\partial H}{\partial y_1} = \frac{\partial H}{\partial y_2} = \cdots = \frac{\partial H}{\partial u} = 0.$$

We readily discover that this yields

$$y_\delta = \frac{\mu(\delta)}{\phi(\delta) Y}, \quad Y = \sum_{d \le z} \frac{\mu^2(d)}{\phi(d)}.$$

Let us note that we have not assumed at the beginning that λ_d vanishes when d is not squarefree, but the proof tells us to take $y_\delta = 0$ when δ is not squarefree, and this implies that λ_d shares this same property. Let us compute λ_d. We find that

$$\lambda_d = d \sum_{\substack{\delta \le z,\\ d|\delta}} \mu(\delta/d) \frac{\mu(\delta)}{\phi(\delta) Y} = \mu(d) \frac{d}{\phi(d) Y} \sum_{\substack{q \le z/d,\\ (q,d)=1}} \frac{\mu^2(q)}{\phi(q)}.$$

Note that Lemma 1.13 implies the neat bound $|\lambda_d| \le 1$ that we have already announced. As a conclusion, we find that

$$Z \le \frac{N}{Y} + \mathcal{O}(z^2) \le \frac{N}{\operatorname{Log} z} + \mathcal{O}(z^2)$$

by (2.1). We finally select $z = \sqrt{N}/\operatorname{Log} N$. □

Exercise 2.3 *Show that*

$$\sum_{\substack{m,n \le N,\\ (m,n)=1}} 1 = \frac{6}{\pi^2} N^2 + \mathcal{O}(N \operatorname{Log} N).$$

Exercise 2.4 *Show that*

$$\sum_{m,n\leq N}(m,n) = CN^2 \operatorname{Log} N + \mathcal{O}(N^2).$$

for a constant C.

Exercise 2.5 *Show that*

$$\sum_{m,n\leq N}(m,n)^2 = \left(\frac{2\zeta(2)}{3\zeta(3)} - \frac{1}{3}\right) N^3 + \mathcal{O}(N^2).$$

This last exercise is more difficult.

CHAPTER 3

The Levin-Fainleib Theorem et alia

We essentially encounters two different cases when studying the average of non-negative multiplicative functions: either the function we sum is roughly speaking of size of order $1/p$ at prime p, or it is of constant size on primes. Let us say for short that the first is case I and the latter is case II.

We first prove a Theorem that enables one to compute the average in case I. Then we prove a very general Theorem that bound above the mean value of a case II function by a mean value of a case I function. We finally handle the average of cases II functions.

3.1. The Levin-Fainleib Theorem

Here is a theorem inspired by [**36**] but where we take care of the values of our multiplicative function on powers of primes as well. The reader will find in [**62**] an appendix with a similar result. Moreover, we present a completely explicit estimate, which complicates the proof somewhat. In [**11**], the reader will find, inter alia, a presentation of many results in the area, a somewhat different exposition as well as a modified proof: the authors achieve there a better treatment of the error term by appealing to a preliminary sieving.

Theorem 3.1 *Let g be a non-negative multiplicative function. Let κ, L and A be three non-negative real parameters such that*

$$\begin{cases} \displaystyle\sum_{\substack{p\ge 2,\nu\ge 1\\ p^\nu\le Q}} g\big(p^\nu\big)\operatorname{Log}\big(p^\nu\big) = \kappa\operatorname{Log}Q + \mathcal{O}^*(L) \qquad (Q\ge 1),\\ \displaystyle\sum_{p\ge 2}\sum_{\nu,k\ge 1} g\big(p^k\big)g\big(p^\nu\big)\operatorname{Log}\big(p^\nu\big)\le A. \end{cases}$$

Then, when $D\ge \exp(2(L+A))$, we have

$$\sum_{d\le D} g(d) = C\,(\operatorname{Log}D)^\kappa\,(1+\mathcal{O}^*(B/\operatorname{Log}D))$$

where C is a positive constant and

$$B = 2(L+A)\big(1+2(\kappa+1)e^{\kappa+1}\big).$$

Furthermore, if

$$\sum_{p\le Q}\ \sum_{k\ge (\operatorname{Log} Q)/\operatorname{Log} p} g(p^k) = o(1) \tag{3.1}$$

as Q goes to infinity, then C is given by

$$C = \frac{1}{\Gamma(\kappa+1)}\prod_{p\ge 2}\biggl\{\Bigl(1-\frac{1}{p}\Bigr)^{\kappa}\sum_{\nu\ge 0} g\bigl(p^{\nu}\bigr)\biggr\}.$$

If in many applications the dependence in L is important, the one in A is most often irrelevant. In the context of the sieve, κ is called the *dimension* of the sieve: it is the parameter that determines the size of the average we are to compute and is, of course, of foremost importance. Let us mention in this direction that [**75**] obtains a one-sided result from one-sided hypothesis, following a path already thread in [**46**]. We show below the first two hypothesis do not imply that the LHS of (3.1) is even bounded. The same Theorem can be found in [**72**, Theorem 21.1] but where the condition (3.1) has been wrongly forgotten.

Proof Let us start with the idea of [**60**]:

$$\begin{aligned} G(D)\operatorname{Log} D &= \sum_{d\le D} g(d)\operatorname{Log}\frac{D}{d} + \sum_{d\le D} g(d)\operatorname{Log} d \\ &= \sum_{d\le D} g(d)\operatorname{Log}\frac{D}{d} + \sum_{\substack{p\ge 2,\nu\ge 1\\ p^{\nu}\le D}} g\bigl(p^{\nu}\bigr)\operatorname{Log}\bigl(p^{\nu}\bigr)\sum_{\substack{\ell\le D/p^{\nu}\\ (\ell,p)=1}} g(\ell). \end{aligned}$$

It is useful to introduce two functions:

$$\begin{cases} G_p(X) = \displaystyle\sum_{\substack{\ell\le X\\ (\ell,p)=1}} g(\ell) \\ T(D) = \displaystyle\sum_{d\le D} g(d)\operatorname{Log}\frac{D}{d} = \int_1^D G(t)\frac{dt}{t}, \end{cases} \tag{3.2}$$

so that we can rewrite the above as

$$G(D)\operatorname{Log}(D) = T(D) + \sum_{\substack{p\ge 2,\nu\ge 1\\ p^{\nu}\le D}} g\bigl(p^{\nu}\bigr)\operatorname{Log}\bigl(p^{\nu}\bigr)G_p(D/p^{\nu}).$$

The functions G_p and G are related by the following identity:

$$G_p(X) = G(X) - \sum_{k\ge 1} g\bigl(p^k\bigr)G_p(X/p^k).$$

This identity, when combined with our hypothesis, enables us to eliminate G_p:

$$\begin{aligned} G(D)\operatorname{Log}(D) &= T(D) + \sum_{\substack{p\ge 2,\nu\ge 1\\ p^\nu\le D}} g\bigl(p^\nu\bigr)\operatorname{Log}\bigl(p^\nu\bigr)G(D/p^\nu) + \mathcal{O}^*(AG(D)) \\ &= T(D) + \sum_{d\le D} g(d) \sum_{\substack{p\ge 2,\nu\ge 1\\ p^\nu\le D/d}} g\bigl(p^\nu\bigr)\operatorname{Log}\bigl(p^\nu\bigr) + \mathcal{O}^*(AG(D)) \\ &= T(D)(\kappa+1) + \mathcal{O}^*((L+A)G(D)). \end{aligned}$$

We rewrite the conclusion as

$$(\kappa+1)T(D) = G(D)\operatorname{Log} D\ (1+r(D)) \quad \text{with } r(D) = \mathcal{O}^*\left(\frac{L+A}{\operatorname{Log} D}\right).$$

We see the previous equation as a differential equation. We set

$$\exp E(D) = \frac{(\kappa+1)T(D)}{(\operatorname{Log} D)^{\kappa+1}} = \frac{G(D)}{(\operatorname{Log} D)^{\kappa}}(1+r(D))$$

getting for $D \ge D_0 = \exp(2(L+A))$

$$\begin{aligned} E'(D) = \frac{T'(D)}{T(D)} - \frac{(\kappa+1)}{D\operatorname{Log} D} &= \frac{-r(D)(\kappa+1)}{(1+r(D))D\operatorname{Log} D} \\ &= \mathcal{O}^*\left(\frac{2(L+A)(\kappa+1)}{D(\operatorname{Log} D)^2}\right) \end{aligned}$$

since $|r(D)| \le 1/2$ when $D \ge D_0$ and on computing $T'(D)$ through (3.2). Now, still for $D \ge D_0$, we have

$$E(\infty) - E(D) = \int_D^\infty E'(t)dt = \mathcal{O}^*\left(\frac{2(L+A)(\kappa+1)}{\operatorname{Log} D}\right).$$

Gathering our results, and using $\exp(x) \le 1 + x\exp(x)$ valid for $x \ge 0$, we infer that

$$\frac{G(D)}{(\operatorname{Log} D)^\kappa} = \frac{\exp E(D)}{1+r(D)} = \frac{e^{E(\infty)}}{1+r(D)}\left(1 + \mathcal{O}^*\left(\frac{2(L+A)}{\operatorname{Log} D}(\kappa+1)e^{\kappa+1}\right)\right).$$

We next use $1/(1+x) \le 1+2x$ valid when $0 \le x \le \frac{1}{2}$ and $(1+x)(1+y) \le (1+2x+y)$ valid for $x, y \ge 0$ and $y \le 1$ to infer

$$\frac{G(D)}{(\operatorname{Log} D)^\kappa} = e^{E(\infty)}\left(1 + \mathcal{O}^*\left(\frac{2(L+A)}{\operatorname{Log} D}\bigl(1+2(\kappa+1)e^{\kappa+1}\bigr)\right)\right).$$

This ends the main part of the proof. We are to identify $e^{E(\infty)} = C$. Note that the above proof is *apriori* wrong since $T'(D) \ne G(D)/D$ at the discontinuity points of G, but we simply have to restrict our attention to non integer D's and then proceed by continuity.

An expression for C. We start by noticing that summation by parts immediately yields that there exists a constant c_0 such that:

$$\sum_{p^k\le P} g(p^k) = \kappa \operatorname{Log}\operatorname{Log} P + c_0 + \mathcal{O}\bigl(1/\operatorname{Log}(3P)\bigr). \tag{3.3}$$

We then note that our second condition implies the convergence, for each prime p, of the series $\sum_{k\ge 0} g(p^k)$; it even implies that $\sum_{k\ge(\operatorname{Log} Q)/\operatorname{Log} p} g(p^k) \le \sqrt{A/\operatorname{Log} Q}$ which has the consequence that these sums are uniformly bounded.

$$\sum_{p>Q}\left|\operatorname{Log}\sum_{k\ge 0} g(p^k) - \sum_{k\ge 1} g(p^k)\right| \ll 1/\operatorname{Log}(2Q).$$

This follows from the inequality

$$\operatorname{Log}\sum_{k\ge 0} g(p^k) - \sum_{k\ge 1} g(p^k) \ll \Bigl(\sum_{\nu\ge 1} g(p^\nu)\Bigr)^2.$$

On summing over $p > Q$, we get that the sum above is

$$\ll \sum_{p>Q}\sum_{\nu,k\ge 1} g(p^\nu)g(p^k)\frac{\operatorname{Log} p}{\operatorname{Log} Q} \ll A/\operatorname{Log}(2Q)$$

as announced. On using the identity

$$\operatorname{Log}\prod_{p\le Q}\left\{\left(1-\frac{1}{p}\right)^{\kappa}\sum_{\nu\ge 0} g\bigl(p^\nu\bigr)\right\} = -\kappa\sum_{p\le Q}\frac{1}{p} + \sum_{\substack{p\le Q,\\ k\ge 1}} g(p^k)$$
$$+\sum_{p\le Q}\left(\kappa\operatorname{Log}(1-p^{-1}) + \frac{\kappa}{p} + \operatorname{Log}\sum_{k\ge 0} g(p^k) - \sum_{k\ge 1} g(p^k)\right).$$

On gathering some of the estimates above, we see that the last summand converges to a constant (with error term $\mathcal{O}(1/\operatorname{Log}(2Q))$). This finally gives us that there exists a constant c_2 such that

$$\operatorname{Log}\prod_{p\le Q}\left\{\left(1-\frac{1}{p}\right)^{\kappa}\sum_{\nu\ge 0} g\bigl(p^\nu\bigr)\right\} = c_2 + \sum_{\substack{p\le Q, k\ge 1,\\ p^k>Q}} g(p^k) + \mathcal{O}(1/\operatorname{Log}(2Q)).$$

Condition (3.1) tells us that the LHS indeed converges. This preliminary discussion tells us that the Eulerian product

$$\prod_{p\ge 2}\left\{\left(1-\frac{1}{p}\right)^{\kappa}\sum_{\nu\ge 0} g\bigl(p^\nu\bigr)\right\} \tag{3.4}$$

is convergent.

We define, for s a positive real number,

$$\begin{aligned} D(g,s) &= \sum_{d\geq 1}\frac{g(d)}{d^s} = s\int_1^\infty G(D)\frac{dD}{D^{s+1}} \\ &= sC\int_1^\infty (\operatorname{Log} D)^\kappa \frac{dD}{D^{s+1}} + \mathcal{O}\Big(sC\int_1^\infty (\operatorname{Log} D)^{\kappa-1}\frac{dD}{D^{s+1}}\Big) \\ &= C\big(s^{-\kappa}\Gamma(\kappa+1) + \mathcal{O}(s^{1-\kappa}\Gamma(\kappa))\big) \end{aligned}$$

and consequently

$$\begin{aligned} C &= \lim_{s\to 0^+} D(g,s)s^\kappa \Gamma(\kappa+1)^{-1} \\ &= \lim_{s\to 0^+} D(g,s)\zeta(1+s)^{-\kappa}\Gamma(\kappa+1)^{-1}. \end{aligned}$$

The absolute convergence of the Euler product (3.4) enables us to conclude easily. □

We now study an example that shows that (3.1) cannot be infered from the other hypothesis. We consider the multiplicative function g defined, for each prime p by $g(p^k) = 0$ except when $k = k_p = [\operatorname{Log} p] + n_p$ where n_p is one of 0, 1 or 2 and is chosen so that

$$\sum_{p\leq Q} g(p^k)\operatorname{Log}(p^k) = \tfrac{1}{2}\operatorname{Log} Q + \mathcal{O}(1).$$

This is possible because $\sum_{p\leq P}(\operatorname{Log} p)^2/p = \frac{1}{2}\operatorname{Log}^2 P + \mathcal{O}(1)$. If we take $n_p = 0$, we essentially would have to evaluate $\sum_{p\leq P}[\operatorname{Log} p](\operatorname{Log} p)/p$; this sum is asymptotic to $\frac{1}{2}\operatorname{Log}^2 P$, but the error term may be of size $\operatorname{Log} P$. One can show that it is possible to correct that with some mild modification n_p. We however have that

$$\sum_{p\leq Q}\sum_{k/p^k>Q} g(p^k) \sim \sum_{\exp\sqrt{\operatorname{Log} Q}<p\leq Q} 1/p \sim \tfrac{1}{2}\operatorname{Log}\operatorname{Log} Q$$

violating strongly (3.1). The reader will get a milder violation on taking $k_p = 2$. In this latter case, the Euler product expression is still valid, while in the former one, the Euler product does not converge.

Exercise 3.2 *Show that*

$$\sum_{d\leq D}\frac{\mu^2(d)3^{\omega(d)}}{\phi(d)} = \frac{1+o(1)}{6}\prod_{p\geq 2}\Big(1-\frac{3}{p^2}+\frac{2}{p^4}\Big)\operatorname{Log}^3 D.$$

Exercise 3.3 *Show that*

$$\sum_{d\leq D}\frac{\mu^2(d)}{\phi(d)} = \operatorname{Log} D + \mathcal{O}(1).$$

Exercise 3.4 *Let w be a function that belongs to $C^1[0,1]$. Show that*

$$\sum_{d\le D}\frac{\mu^2(d)}{\phi(d)}w\Big(\frac{\operatorname{Log} d}{\operatorname{Log} D}\Big)=\int_0^1 w(t)dt\ \operatorname{Log} D+\mathcal{O}(1).$$

Exercise 3.5 *Show that*

$$\sum_{n\le N}\frac{d(n)}{n}=\tfrac12\operatorname{Log}^2 N+\mathcal{O}(\operatorname{Log} N)$$

where $d(n)$ denotes the number of divisors of n.

3.2. A simple general inequality

We prove the following theorem that relies on a theme initially developed in [**40**]. The best result in this direction is in [**38**]. Of course, we also extend it to encompass values at powers of primes. The starting idea is still taken from the celebrated [**60**] proved in the preceding section.

Theorem 3.6 *Let $D\ge 2$ be a real parameter. Assume g is a multiplicative non-negative function such that*

$$\sum_{\substack{p\ge 2,\nu\ge 1\\ p^\nu\le Q}} g\big(p^\nu\big)\operatorname{Log}\big(p^\nu\big)\le KQ+K'\qquad(\forall Q\in[1,D])$$

for some constants $K,K'\ge 0$. Then for $D>\exp(K'-1)$, we have

$$\sum_{d\le D} g(d)\le\frac{(K+1)\,D}{\operatorname{Log} D-K'+1}\sum_{d\le D} g(d)/d.$$

Proof Let us set $\tilde G(D)=\sum_{d\le D} g(d)/d$. Using $\operatorname{Log}\frac{D}{d}\le\frac{D}{d}-1$, we get

$$\begin{aligned}G(D)\operatorname{Log} D&=\sum_{d\le D} g(d)\operatorname{Log}\frac{D}{d}+\sum_{d\le D} g(d)\operatorname{Log} d\\&\le D\tilde G(D)-G(D)+\sum_{\substack{p\ge 2,\nu\ge 1\\ p^\nu\le D}} g\big(p^\nu\big)\operatorname{Log}\big(p^\nu\big)\sum_{\substack{\ell\le D/p^\nu\\ (\ell,p)=1}} g(\ell)\end{aligned}$$

where we get the second summand by writing $\operatorname{Log} d=\sum_{p^\nu\|d}\operatorname{Log}\big(p^\nu\big)$. Finally

$$\begin{aligned}\sum_{\substack{p\ge 2,\nu\ge 1\\ p^\nu\le D}} g\big(p^\nu\big)\operatorname{Log}\big(p^\nu\big)\sum_{\substack{\ell\le D/p^\nu\\ (\ell,p)=1}} g(\ell)&=\sum_{\ell\le D} g(\ell)\sum_{\substack{p\ge 2,\nu\ge 1\\ p^\nu\le D/\ell\\ (p,\ell)=1}} g\big(p^\nu\big)\operatorname{Log}\big(p^\nu\big)\\&\le\sum_{\ell\le D} g(\ell)\Big(\frac{KD}{\ell}+K'\Big)\end{aligned}$$

from which the theorem follows readily. □

3.3. A further consequence

It is not difficult by following [**93**] to derive a stronger mean value result from Theorem 3.1. Since it will be required in one of the applications below, and since all the necessary material has been already exposed, we include one such result.

Theorem 3.7 *Let f be a non-negative multiplicative function and κ be non-negative real parameter such that*

$$\begin{cases} \displaystyle\sum_{\substack{p\ge 2,\nu\ge 1\\ p^\nu\le Q}} f\bigl(p^\nu\bigr)\operatorname{Log}\bigl(p^\nu\bigr) = \kappa Q + \mathcal{O}(Q/\operatorname{Log}(2Q)) \qquad (Q\ge 1),\\ \displaystyle\sum_{p\ge 2}\sum_{\substack{\nu,k\ge 1,\\ p^{\nu+k}\le Q}} f\bigl(p^k\bigr)f\bigl(p^\nu\bigr)\operatorname{Log}\bigl(p^\nu\bigr) \ll \sqrt{Q}, \end{cases}$$

then we have

$$\sum_{d\le D} f(d) = \kappa\, C\cdot D\,(\operatorname{Log} D)^{\kappa-1}\,(1+o(1))$$

where C is as in Theorem 3.1.

Proof We proceed as in Theorem 3.1. Write

$$S(D) = \sum_{d\le D} f(d).$$

By using Theorem 3.6 followed by an application of Theorem 3.1, we readily obtain the following apriori bound

$$S(D) \ll D(\operatorname{Log}(2D))^{\kappa-1}. \tag{3.5}$$

Consider now $S^*(D) = \sum_{d\le D} f(d)\operatorname{Log} d$. Proceeding as in the proof of Theorem 3.1, we get

$$\begin{aligned} S^*(D) &= \sum_{\substack{p\ge 2,\nu\ge 1\\ p^\nu\le D}} f\bigl(p^\nu\bigr)\operatorname{Log}\bigl(p^\nu\bigr) \sum_{\substack{\ell\le D/p^\nu\\ (\ell,p)=1}} f(\ell)\\ &= \sum_{\ell\le D} f(\ell) \sum_{\substack{p\ge 2,\nu\ge 1\\ p^\nu\le D/\ell,\\ (p,\ell)=1}} f\bigl(p^\nu\bigr)\operatorname{Log}\bigl(p^\nu\bigr) \end{aligned}$$

so that $S^*(D)$ equals

$$\sum_{\ell\le D} f(\ell) \sum_{\substack{p\ge 2,\nu\ge 1\\ p^\nu\le D/\ell}} f\bigl(p^\nu\bigr)\operatorname{Log}\bigl(p^\nu\bigr) - \sum_{\ell\le D} f(\ell) \sum_{\substack{p\ge 2,\nu,k\ge 1\\ p^{\nu+k}\le D/\ell,\\ (p,\ell)=1}} f\bigl(p^\nu\bigr)f\bigl(p^k\bigr)\operatorname{Log}\bigl(p^\nu\bigr)$$

We use our hypothesis on this expression and conclude that

$$S^*(D) = \kappa D \sum_{\ell \le D} f(\ell)/\ell + \mathcal{O}\biggl(Q \sum_{\ell \le D} \frac{f(\ell)}{\ell \operatorname{Log}(2Q/\ell)}\biggr) + \mathcal{O}\biggl(\sqrt{Q} \sum_{\ell \le D} \frac{f(\ell)}{\sqrt{\ell}}\biggr).$$

Both error terms are shown to be $\mathcal{O}(Q \operatorname{Log}(2Q)^{\kappa-1})$ by appealing to (3.5) while the main term is evaluated via Theorem 3.1. We finally use an integration by parts:

$$S(D) = 1 + \int_2^D S^*(t) \frac{dt}{t \operatorname{Log}^2 t} + \frac{S^*(D)}{\operatorname{Log} D}$$

to get the claimed asymptotic. □

CHAPTER 4

Some more exercises

For exercises, the reader should consult [**1**]. More elaborate and often more difficult ones are to be found in [**64**].

Exercise 4.1

1. *Show that, when m and n are squarefree and distinct, then*
$$\frac{\phi(m)}{m} \neq \frac{\phi(n)}{n}.$$
2. *Show that the same holds for the function $\ell \mapsto \sigma(\ell)/\ell$.*
3. *Does the function $\ell \mapsto \sigma(\ell)/\phi(\ell)$ verify also this property?*
4. *Investigate similarly the function $\ell \mapsto \prod_{p|\ell}(p+2)/(p+1)$ with respect to this property.*

Exercise 4.2

1. *Show that, for every prime number p and any positive integer a, one has*
$$\sigma(p^a) = \frac{p^{a+1}-1}{p-1}$$
where $\sigma(d)$ is the sum of the divisors of d.
2. *Prove the following identity, valid for any positive integer n:*
$$\sigma(n)^2 = n \sum_{d|n} \sigma(d^2)/d.$$

Exercise 4.3 *Show that the function that associates to every integer $n > 1$ the double of the sum of the integers belonging to $[1, n]$ that are coprime to n, and that associates the value 1 at the integer 1, is multiplicative.*

Exercise 4.4 *We define*
$$T(N) = \sum_{y<(N/8)^{1/3}} \operatorname{Log}(N - 8y^3).$$

1. *Show that*
$$T(N) = (N/8)^{1/3} \operatorname{Log} N + c_0 N^{1/3} + \mathcal{O}(\operatorname{Log}(2N))$$
for some constant c_0.

2. Show that

$$T(N) = \tfrac{1}{3}(N/8)^{1/3}\operatorname{Log} N + \sum_{\substack{(N/2)^{1/3}<m\leq N,\\ \exists y\leq (N/2)^{1/3}, m|N-8y^3}} \Lambda(m) + \mathcal{O}(N^{1/3}).$$

3. Deduce from the above that there exists infinitely many integers n of the shape $n = N - 8y^3$, where y is a positive integer verifying $8y^3 \leq N/2$, that have a prime factor large than $n^{1/6}$.

Exercise 4.5

1. Find an asymptotical formula for

$$\sum_{n\leq N}\sum_{d|n}\mu^2(n/d)3^{-\omega(d)}/n.$$

2. Find an asymptotical formula for

$$\sum_{n\leq N}\sum_{d|n}\mu^2(n/d)3^{-\omega(d)}/(n+1).$$

Exercise 4.6 *Show that*

$$\sum_{d\leq D} 9^{\omega(d)}/\phi(d) \sim C(\operatorname{Log} D)^9$$

for a positive constant C that is to be made explicit.

Exercise 4.7

1. Show that

$$\frac{n}{\phi(n)} = \sum_{d|n}\frac{\mu^2(d)}{\phi(d)}.$$

2. Show that

$$\sum_{n\leq N}\frac{n}{\phi(n)} = \frac{\zeta(2)\zeta(3)}{\zeta(6)}N + \mathcal{O}(\operatorname{Log}(2N)).$$

3. Show that there exists a constant C such that

$$\sum_{n\leq N}\frac{1}{\phi(n)} = \frac{\zeta(2)\zeta(3)}{\zeta(6)}\operatorname{Log} N + C + \mathcal{O}(\operatorname{Log}(2N)/N). \tag{4.1}$$

4. Show that the constant C from the previous question is given by

$$C = \frac{\zeta(2)\zeta(3)}{\zeta(6)}\biggl(\gamma - \sum_{p\geq 2}\frac{\operatorname{Log} p}{p^2-p+1}\biggr).$$

5. Show that the error term in (4.1) *cannot be any better than* $\operatorname{Log}\operatorname{Log}(4N)/N$.

Exercise 4.8 *We denote by $\tau(m)$ the number of (positive) divisors of m. Let $\mathcal{D}$ be the set of integers having only prime factors $\leq D$ where D is a parameter ≥ 1.*

1. *Show that*
$$\sum_{d\in\mathcal{D}} 1/d \ll \mathrm{Log}(2D).$$
2. *Show that, for every $q \ge 1$, we have*
$$\tau(m)^q \le \sum_{dk=m} \tau(d)^{q-1}\tau(k)^{q-1}.$$
3. *Show that, for every integer positive parameter q, we have*
$$\sum_{m\in\mathcal{D}} \tau(m)^q/m \ll (\mathrm{Log}(2D))^{2^q}.$$
4. *Let $D \ge 1$ be a fixed parameter and let $\tau(n; D)$ be the number of divisors of n that are below D. Let q be a positive integer parameter. Show that*
$$\sum_{n\le X} \tau(n; D)^q \ll X(\mathrm{Log}(2D))^{2^q}.$$
The reader may want to consider the largest divisor $t(n)$ of n in $\mathcal{D}$.

Exercise 4.9 *We denote by $\tau_r(n)$ the number of r-tuples of (positive) integers $(n_1, n_2, \cdots, n_r)$ that are such that $n_1 n_2 \cdots n_r = n$.*

1. *Show that*
$$\tau_r(p^a) = \binom{r-1+a}{r-1}.$$
2. *Show that $\tau_r(n_1 n_2 \cdots n_r) \le \tau_r(n_1)\tau_r(n_2)\cdots\tau_r(n_r)$ and deduce from it that*
$$\sum_{n\le N} \tau_r(n)^2/n \le (\mathrm{Log}\, N + 1)^{r^2}.$$
3. *Show that*
$$\sum_{n\le N} \tau_r(n)^2 \ll_r N(\mathrm{Log}\, N + 1)^{r^2-1}$$
where the symbol $\ll_r$ means "less or equal to a constant that may depend on r times ...".

CHAPTER 5

Introducing the large sieve

This chapter is an introduction to the arithmetical aspects of the large inequality. We are going to reprove here a weak version of the Brun-Titchmarsh Theorem (stated as Theorem 2.2).

Theorem 5.1 *Let M and N be two positive real numbers, N being further assumed to be > 1. There are at most $2N/\operatorname{Log} N$ prime numbers in the interval.*

When $M = 1$, the prime number Theorem gives a better result, since it replaces this 2 by a $1 + o(1)$ (and the inequality sign by an equality). The strength of this bound lies in its abscence of conditions linking M and N. It for instance tells us that, between 10^{100} and $10^{100} + 10\,000$, there are not more than 2 200 prime numbers. See Theorem 2.2 for some further commemts

The factor 2 in the statement is crucial, and any improvement on it would have momentous consequences on our knowledge of prime numbers.

We simplify the exposition below and prove the Theorem with a factor $4 + o(1)$ instead of 2; We further assume that $M \geq N^{1/4}$. However some more care in the coming proof would yield the required $2 + o(1)$ with nonconditions on M. Removing this last $o(1)$ is a difficult task. The road we take is far from the concepts of divisibility and sieve theory. There exists a way from this approach to the Selberg sieve one, but it would too long for us to indicate it.

We set

$$\mathcal{Q} = \left\{ q \leq Q,\ q \text{ squarefree} \right\} \tag{5.1}$$

where Q is a parameter at our disposal.

Let us mention finally that, in case the reader has a grasp on the french language, [**71**] proposes a french version of this exposition, with some more motivating material aimed at early students.

5.1. A hermitian tool

Our first Lemma has actors in a Hilbert space $\mathcal{H}$.

Lemma 5.2 (Approximate Parseval) *Let $(\varphi_q)_{q\in\mathcal{Q}}$ be a family of points in $\mathcal{H}$. Let f be another of $\mathcal{H}$. Then we have*

$$\sum_{q\in\mathcal{Q}} |[f|\varphi_q]|^2/M_q \leq \|f\|^2$$

where M_q is an upper bound for $\sum_{q'\in\mathcal{Q}} |[\varphi_q|\varphi_{q'}]|$.

See [**72**, Chapter 1] for more details.

In order to understand this Lemma, the best thing is to consider the case when the φ_q are two by two orthogonal. In that case, we can simply take $M_q = \|\varphi_q\|^2$ and the proposed inequality is nothing more than the Parseval inequality. Our Lemma, which is due to Selberg in the early seventies, rids us of the hypothesis of strict orthogonality, but it is strong only when the system we apply it to is "nearly orthogonal", i.e. M_q can be taken to be "almost" $\|\varphi_q\|^2$.

The hungarian mathematician Hálasz had introduced earlier general hermitian inequalities in this area (see [**35**, Hilfsatz 1]) but the general paternity of things is unclear.

Proof Let us consider the inequality

$$\left\| f - \sum_{q\in\mathcal{Q}} \xi_q\varphi_q \right\|^2 \geq 0$$

where (ξ_q) is a family of parameters that we may choose to the best of our interest. We would like $\sum_{q\in\mathcal{Q}} \xi_q\varphi_q$ to be the orthogonal projection of f on the space generated by (φ_q), but we do not know how to compute these coefficients ξ_q. We thus restrain ourselves to what we believe is only a close approximation, but that we know how to express. On developping the norm, we get

$$\|f\|^2 - 2\Re\sum_{q\in\mathcal{Q}} \overline{\xi_q}[f|\varphi_q] + \sum_{q,q'\in\mathcal{Q}} \xi_q\overline{\xi_{q'}}[\varphi_q|\varphi_{q'}] \geq 0,$$

where we separate q from q' by appealing to $|\xi_q||\xi_{q'}| \leq \frac{1}{2}(|\xi_q|^2 + |\xi_{q'}|^2)$. We shuffle the terms and inject M_q in the resulting inequality. We get

$$\|f\|^2 - 2\Re\sum_{q\in\mathcal{Q}} \overline{\xi_q}[f|\varphi_q] + \sum_{q\in\mathcal{Q}} |\xi_q|^2 M_q \geq 0. \tag{5.2}$$

The quadratic form in (ξ_q) has now been replaced by a diagonal one, for which getting the coefficients ξ_q that minise it is an easy task:

$$\xi_q = [f|\varphi_q]/M_q.$$

The Lemma follows by plug these values in (5.2). □

5.2. A pinch of number theory

Our second Lemma has to do with the Ramanujan sums:

$$\mathrm{c}_q(n)=\sum_{\substack{1\le a\le q,\\ (a,q)=1}}\exp(2i\pi an/q). \tag{5.3}$$

It is obvious that $\mathrm{c}_q(n)$ depends only on the residue class of n modulo q. Let us notice already that the summation carries over $\phi(q)$ integers a, where $\phi(q)$ is the Euler function, since $\phi(q)$ is mprecisely the number of integers not more that q and coprime to it.

When $q=p$ is a prime number, the condition $(a,p)=1$ reduces to $a\neq p$ and

$$\mathrm{c}_p(n)=\sum_{1\le a\le p}\exp(2i\pi an/p)-1=\begin{cases}p-1 & \text{when } p|n,\\ -1 & \text{when } p\nmid n.\end{cases} \tag{5.4}$$

On using the chinese remainder Theorem, we show the following multiplicativity rule:

$$\mathrm{c}_{q_1q_2}(n)=\mathrm{c}_{q_1}(n)\,\mathrm{c}_{q_2}(n)\qquad\text{when } (q_1,q_2)=1. \tag{5.5}$$

In our proof, q will be squarefree, and thus, the combination of (5.4) and of (5.5) is enough to compute $\mathrm{c}_q(n)$. Here is the property we will require:

Lemma 5.3 *When q is squarefree and n is prime to q, we have* $\mathrm{c}_q(n)=\mu(q)$.

In particular, this value is independant of n amd its modulus is 1. The assumption that q be squarefree is in fact not required for this Lemma.

Our third Lemma has already been proved in (2.1). It reads:

Lemma 5.4 *We have that*

$$\sum_{q\le Q}\mu^2(q)/\phi(q)\ge \operatorname{Log} Q.$$

Exercise 5.5 *Show that*

$$c_q(n)=\sum_{\substack{d|q,\\ d|n}}\mu(q/d)d.$$

5.3. Proof of the Brun-Titchmarsh inequality

We select for φ_q the following function:

$$\varphi_q(n)=\begin{cases}\mathrm{c}_q(n) & \text{when } M+1\le n\le M+N\\ 0 & \text{else,}\end{cases} \tag{5.6}$$

for q squarefree and of size at most Q.

The Hilbert space we consider is simply the space of functions over the integers in the interval $[M+1, M+N]$ equipped with the standard hermitiam product:

$$[g|h] = \sum_{M+1\le n\le M+N} f(n)\overline{g(n)}. \tag{5.7}$$

Our last Lemma gives a measure of the “almost orthogonality” of the family $(\varphi_q)_{q\in\mathcal{Q}}$.

Lemma 5.6

$$\sum_{q'\in\mathcal{Q}} |[\varphi_q|\varphi_{q'}]| \le M_q = \phi(q)(N+Q^4).$$

Proof We have

$$[\varphi_q|\varphi_{q'}] = \sum_{M+1\le n\le M+N} \biggl(\sum_{\substack{1\le a\le q,\\ (a,q)=1}} \mathrm{e}(na/q)\biggr)\biggl(\sum_{\substack{1\le a'\le q',\\ (a',q')=1}} \mathrm{e}(-na'/q')\biggr).$$

By summing first over a and a', we get

$$[\varphi_q|\varphi_{q'}] = \sum_{\substack{1\le a\le q,\\ (a,q)=1}} \sum_{\substack{1\le a'\le q',\\ (a',q')=1}} \sum_{M+1\le n\le M+N} \mathrm{e}\biggl(n\biggl(\frac{a}{q}-\frac{a'}{q'}\biggr)\biggr).$$

The inner summation is in fact the sum of a geometric progression. When $a/q \neq a'/q'$, it is at most, in modulus,

$$1\bigg/\biggl|\sin\biggl(\pi\biggl(\frac{a}{q}-\frac{a'}{q'}\biggr)\biggr)\biggr| \le qq'/2$$

by using the classical inequality $\sin x \ge 2x/\pi$ when $0\le x\le \pi/2$. □

Exercise 5.7 *In the proof above, fix a and take advantage on the summation over a' to show that one may take $M_q = \phi(q)(N+\mathcal{O}(Q^3\operatorname{Log}(2Q))$, .*

Exercise 5.8 *On using the fact that any two distinct points of the set $\{a/q, (a,q)=1, q\le Q\}$ are distant by at least $1/Q^2$, show that one can take $M_q/\phi(q) = N+\mathcal{O}(Q^2\operatorname{Log} Q)$.*

Proof of Theorem 5.1 We simply look at the characteristic function f of those prime numbers that lie inside the interval $[M+1, M+N]$; we assume for simplicity that Q is not more than M. In this case, we get directly

$$[f|\varphi_q] = \sum_{M+1\le n\le M+N} f(n)\overline{\varphi_q(n)} = \mu(q)\sum_{M+1\le n\le M+N} f(n) \tag{5.8}$$

simply because $\varphi_q(n) = \mu(q)$ for every prime numbers in our interval. On calling Z the number of these primes, we can rewrite the above equality as

$$[f|\varphi_q] = \mu(q)Z. \tag{5.9}$$

Our hermitain Lemma gives us

$$\sum_{q\in\mathcal{Q}} \left|(-1)^{\omega(q)}Z\right|^2/\phi(q) \le Z(N+Q^4),$$

i.e.

$$Z\sum_{q\le Q}\mu^2(q)/\phi(q) \le N+Q^4. \tag{5.10}$$

Lemma 5.4 leads to the upper bound $Z \le (N+Q^4)/\operatorname{Log} Q$ where the only thing left is to optimize Q. We select

$$Q = N^{1/4}/\operatorname{Log} N \tag{5.11}$$

and the Theorem follows readily. □

We have used in the proof the condition $Q \le M$. It is easy to remove it by taking for f the characteristic function of those primes tat lie within $[M+1, M+N]$ and are moreover strictly larger than Q. As a consequence of this change, $\sum_n f(n)$ does not qnymore equals to Z but to $Z + \mathcal{O}(Q)$, which is more than enough.

5.4. Some historical comments and two open problems

This proof shows clearly the sieving effect (in a somewhat vague sense) of the factor $c_r(m)$. The reader will find in [**23**] a very same use of this factor. It is also a main feature of earlier unpublished work of Selberg on pseudo-characters, a trace of which the reader will find in [**6**], [**50**], [**65**], [**66**] as well as in [**47**, chapter 18]. See further [**55**] and [**72**, chapter 11].

This way of proving the Brun-Titchmarsh inequality is fairly recent and is the basis of [**74**] where we prove that

Theorem 5.9 *There exists an N_0 such that for all $N \ge N_0$ and all $M \ge 1$ we have*

$$\pi(M+N) - \pi(M) \le \frac{2N}{\operatorname{Log} N + 3.53}.$$

The possibility of such an inequality, though with an unspecified value instead of 3.53 is due to [**90**]. [**82**] also mentions such a result without presenting any proof. [**5**] has the first value with the upper bound $Z \le 2N/(\operatorname{Log} N - 3 + o(1))$. [**63**] refined this -3 in a 5/6 and, in [**85**, section 22], the reader will find a proof leading to 2.81.

Here are two open problems:

Open Problem no I. *Determine*

$$\limsup_{N\to\infty}\max_{M\geq 0}\frac{\pi(M+N)-\pi(M)}{N/\operatorname{Log} N}.$$

We have shown that this maximum is ≤ 2, while the prime number Theorem shows that it is ≥ 1. I tend to believe that this maximum is indeed equal to 2, but this is shear guess since I do not even have a heuristical argument at my disposal.

Open Problem no II. *Determine*

$$\limsup_{M\to\infty}\frac{\pi(M+N)-\pi(M)}{N/\operatorname{Log} N}.$$

We know that this limit lies between 0 and 2, but we do not even know whether it is > 0 or not. We explain roughly in next section why the Hardy & Littlewood conjecture concerning prime κ-tuples implies that this limit is, when N becomes large, $\geq 1 + o(1)$

It would already be extremely interesting to find intervals $[M+1, M+N]$ where the number of primes within divided by $N/\operatorname{Log} N$ is > 1 and where, say, $M \geq 2N$. Here are some examples :

$M+1$	$M+N$	N	Z	ratio
5 639	5 659	21	7	1.0148 . . .
113 143	113 177	35	10	1.0158 . . .
21 817 283 854 511 261	21 817 283 854 511 311	51	14	1.0793 . . .

See [**22**], [**44**], [**81**] and [**80**]. In order to explore this problem some more, we need to define admissible tuples.

5.5. Admissible tuples

A common problem is to look at pairs $(n, n+2)$ for which each component is prime. Extending the problem to κ-tuples means looking for infinitely many integers n for which all the components of $(n+h_1, \ldots, n+h_\kappa)$ are simultaneously prime. Determining which tuples $(h_1, \ldots, h_\kappa)$ should have this property is a non trivial problem; Notice first that $(0,1)$ is clearly not a good choice! Here the obstruction comes from what happens modulo 2. In general the conjecture known as *the prime κ-tuples conjecture*, first stated by [**42**] is that obtructions can only be local. This warrants a definition:

Definition 5.10 *A κ-tuple $\mathfrak{s} = (h_1, \ldots, h_\kappa)$ of increasing integers is said to be* a κ-tuple of admissible shifts *if the set $\{h_1, \ldots, h_\kappa\}$ does not cover all of $\mathbb{Z}/p\mathbb{Z}$ for any prime p. We further impose $h_1 = 0$.*

The *length* $L(\mathfrak{s})$ of such tuple of admissible shifts being $h_\kappa - h_1 + 1$, it is enough to restrict p to be not more than this length in the statement. For example $(0, 2, 6, 9, 12)$ is admissible of length 13.

An interesting problem is to find as dense κ-tuples as possible, where the density is best quantified in terms of the length $L = h_\kappa - h_1 + 1$ compared to the number of primes less than L, i.e. $\pi(L)$. [**44**] proved that there exist κ-tuples of admissible shifts of size

$$\kappa \geq \pi(L) + (\operatorname{Log} 2 - \varepsilon)\frac{L}{\operatorname{Log}^2 L}$$

for every $\varepsilon > 0$ and provided N is large enough in terms of ε. Note that the Brun-Titchmarsh Theorem says that κ is bounded by $2L/\operatorname{Log} L$. If one is ready to believe the prime κ-tuple conjecture, such extreme examples of admissible shifts thus provides us with a lower bound for the best possible upper bound in the Brun-Titchmarsh Theorem. In order to avoid to appeal to the prime κ-tuple conjecture, it would be necessary to indeed exhibit specific examples of such tuples, but this is still beyond the power of nowadays algorithms and computers.

[**26**] contains conjectures and numerical computations related to this problem. [**22**] has built a 1 415-uple of admissible shifts of length 11 763, while $\pi(11\,763) = 1\,409$, but no one has been yet able to produce a corresponding prime 1 415-uple.

[**25**] has pushed computations further and found a 224-uple of length 1 417, while $\pi(1\,417) = 223$ The reader will find on the site of [**27**] a list of long prime tuples, for instance:

$$1\,906\,230\,835\,046\,648\,293\,290\,043{+}0, 4, 6, 10, 16, 18, 24, 28, 30, 34, 40, 46, 48, 54, 58, 60, 66, 70$$

due to J. Waldvogel & P. Leikauf in 2001. It contains 18 primes for a length of 70, while $\pi(70) = 19$.

CHAPTER 6

Some geometric considerations

The reader will find in [**72**, chapter 2] a more complete presentation of the notions detailled in this section.

6.1. Compact sets

We introduce in this section some vocabulary that allows us handle modular arithmetic. All of it is trivial enough but will make life easier later on.

∘∘ By a *compact set* $\mathcal{K}$, we mean a sequence $\mathcal{K} = (\mathcal{K}_d)_{d\ge 1}$ satisfying

(1) $\mathcal{K}_d \subset \mathbb{Z}/d\mathbb{Z}$ for all $d \ge 1$.

(2) For any divisor d of q, we have $\sigma_{q\to d}(\mathcal{K}_q) = \mathcal{K}_d$ where $\sigma_{q\to d}$ is the canonical surjection (also called the restriction map) from $\mathbb{Z}/q\mathbb{Z}$ to $\mathbb{Z}/d\mathbb{Z}$:

$$\begin{aligned} \sigma_{q\to d}: \quad \mathbb{Z}/q\mathbb{Z} &\to \mathbb{Z}/d\mathbb{Z} \\ x \bmod q &\mapsto x \bmod d. \end{aligned} \tag{6.1}$$

When $\mathcal{K}$ is not empty, we have $\mathcal{K}_1 = \mathbb{Z}/\mathbb{Z}$. As examples, we can take $\mathcal{K}_d = \mathbb{Z}/d\mathbb{Z}$ for all d or $\mathcal{K}_d = \mathcal{U}_d$, where $\mathcal{U}_d$ is the set of invertible classes modulo d. The intersection and union of compact sets is again a compact set.

We can also consider $\mathcal{K}$ a subset of $\hat{\mathbb{Z}} = \varprojlim \mathbb{Z}/d\mathbb{Z}$, in which case it is indeed a compact set. Furthermore we shall sometimes consider $\mathcal{K}_d$ as a subset of $\mathbb{Z}$: the set of relative integers whose reduction modulo d falls inside $\mathcal{K}_d$.

∘∘ We say that the compact set $\mathcal{K}$ is *multiplicatively split* if for any d_1 and d_2 coprime positive integers, the Chinese remainder map

$$\mathbb{Z}/d_1d_2\mathbb{Z} \longrightarrow \mathbb{Z}/d_1\mathbb{Z} \times \mathbb{Z}/d_2\mathbb{Z} \tag{6.2}$$

sends $\mathcal{K}_{d_1d_2}$ onto $\mathcal{K}_{d_1} \times \mathcal{K}_{d_2}$. In this case, the sets $\mathcal{K}_{p^\nu}$ for prime p and $\nu \ge 1$ determine $\mathcal{K}$ completely. Notice that when $\mathcal{K}$ is multiplicatively split:

$$|\mathcal{K}_{[d,d']}||\mathcal{K}_{(d,d')}| = |\mathcal{K}_d||\mathcal{K}_{d'}| \tag{6.3}$$

for any d and d', where $[d, d']$ is the lcm and (d, d') the gcd of d and d'. Here $|\mathcal{A}|$ stands for the cardinality of a set $\mathcal{A}$.

∘∘ A compact set is said to be *squarefree* if

$$\mathcal{K}_q = \sigma_{q\to d}^{-1}(\mathcal{K}_d)$$

whenever d divides q and has the same prime factors. For instance, $\mathcal{U}$ is squarefree since being prime to q or to its *squarefree kernel* is the same.

∘∘ A particularly successful hypothesis on $\mathcal{K}$ was introduced by [**49**] in the context of polynomials over a finite field and used in the case of the integers by [**30**] (see also [**83**]). It reads

$$\forall d|q,\ \forall a \in \mathcal{K}_d \text{ the quantity } \sum_{\substack{n\equiv a[d]\\ n\in\mathcal{K}_q}} 1 \text{ is independent of } a.$$

Another way to present this quantity would be to say it is the cardinality of $\sigma_{p^\nu\to p^{\nu-1}}^{-1}(\{a\})$. Since the introduction of this condition in our context is due to [**30**], we shall refer to it as the Johnsen-Gallagher condition. Note that this condition does not require $\mathcal{K}$ to be multiplicatively split, although all our examples will also satisfy this additional hypothesis.

Any squarefree compact set automatically satisfies the Johnsen-Gallagher hypothesis. Since the sieve kept to such sets for a very long time (the reason being that most classical problems fall within this framework), and the combinatorial sieve still does, this condition does not show up in classical expositions.

6.2. Examples of compact sets

The sequence $(\mathbb{Z}/q\mathbb{Z})$ is a compact set; it is indeed the largest one. The reader will check with no difficulties that it is multiplicatively split and squarefree.

We denote by $\mathcal{U}_q$ the set of invertible elements modulo q. This is also the multiplicative group of $\mathbb{Z}/q\mathbb{Z}$ when both sets are endowed with the multiplication. The sequence $(\mathcal{U}_q)_q$ is a compact set. It is again multiplicatively split and squarefree.

The sequence $(\mathcal{U}_q \bigcap (\mathcal{U}_q - 2))_q$ again defines a multiplicatively split squarefree compact set. In general, an additive shift of a compact set is still a compact set, and the intersection of two compact set is still a compact set. Both properties "being squarefree" and "being multiplicatively split" are equally preserved by the above operations.

Let us denote, for every modulus q, by $\mathcal{K}_q$ the set of squares modulo q. The sequence $(\mathcal{K}_q)_q$ defines a multiplicatively split compact set, but this compact set is not squarefree (look at what happens modulo 2 and 4).

Let us end this enumeration with the following example. When x belongs to $\mathbb{Z}/q\mathbb{Z}$, we have access to $\gcd(x,q)$ which is simply the gcd of q and y for any y belonging to the class of x modulo q (this is a proper definition once one has shown that this gcd indeed does not depend on the choice of y. We leave this part to the reader). We take for $\mathcal{K}_q$ the set of residue classes x for which $\gcd(x,q)$ is squarefree. The sequence $(\mathcal{K}_q)_q$ is a compact set. It is remarquable that *every* squarefree integer reduced modulo *any* q falls inside $\mathcal{K}_q$.

6.3. A family of arithmetical functions

Let us start with a multiplicatively split compact set $\mathcal{K}$. We consider the non-negative multiplicative function h defined by

$$h(d) = \prod_{p^\nu \| d} \left(\frac{p^\nu}{|\mathcal{K}_{p^\nu}|} - \frac{p^{\nu-1}}{|\mathcal{K}_{p^{\nu-1}}|} \right) \geq 0, \quad h(1) = 1 \tag{6.4}$$

where $q\|d$ means that q divides d in such a way that q and d/q are coprime. We shall say that q *divides* d *exactly*. Note that

$$\frac{d}{|\mathcal{K}_d|} = \sum_{\delta | d} h(\delta). \tag{6.5}$$

We further define

$$G_d(Q) = \sum_{\substack{\delta \leq Q, \\ [d,\delta] \leq Q}} h(\delta) \tag{6.6}$$

which we also denote by $G_d(\mathcal{K}, Q)$ when mentioning the compact set $\mathcal{K}$ is of any help. Let us note that in the extremal case $\mathcal{K}_d = \mathbb{Z}/d\mathbb{Z}$, we have $h(d) = 0$ except when $d = 1$ in which case we have $h(1) = 1$. This implies that $G_d(Q) = 1$ for all d's. These fairly unusual functions appear in the following form:

Lemma 6.1 *We have*

$$G_d(Q) = \sum_{\substack{q \leq Q \\ d|q}} \biggl(\sum_{f/d|f|q} \mu(q/f) f / |\mathcal{K}_f| \biggr).$$

Proof We plug (6.5) in the above RHS to get

$$S = \sum_{\substack{q \leq Q \\ d|q}} \biggl(\sum_{f/d|f|q} \mu(q/f) \sum_{\delta | f} h(\delta) \biggr).$$

After some shuffling, we get

$$S = \sum_{\delta \leq Q} h(\delta) \sum_{\substack{q, f \leq Q \\ d|f|q, \\ \delta | f}} \mu(q/f).$$

It means that $[d,\delta]|q$ that has to be $\leq Q$. Each summation over f equals to 0 when $q/[d,\delta]$ is not equal to one. This ends the proof of $S = G_d(Q)$. □

Often, the set $\mathcal{K}$ is squarefree, in which case the above expression simplifies and we recognize, up to a factor, the usual functions from the Selberg sieve (see (6.7) below). In particular, we know how to evaluate them. The reader should consult [**60**], [**36**] and [**37**] for the general theory.

We conclude by a lemma that is in fact a generalization of Lemma 1.13 but which is trivial in our setting. It will be further generalized in Lemma 10.3.

Lemma 6.2 *We have $G_\ell(Q\ell/d) \leq G_d(Q) \leq G_\ell(Q)$ for $\ell|d$.*

Proof Both inequalities are trivial consequences of the expression (6.6). The condition $[d,\delta] \leq Q$ implies that $[\ell,\delta] \leq Q$, when $\ell|d$, hence the second inequality. Furthermore, we have $[d,\delta] \leq [\ell,\delta]d/\ell$. Hence, when $[\ell,\delta] \leq Q\ell/d$, then $[d,\delta] \leq Q$ and the first inequality follows readily. □

When the compact set is squarefree, the reader will check from (6.4) *that $h(d) = 0$ as soon as d is not squarefee. In that case, the summand appearing in Lemma 6.1 vanishes whenever q/d and d are coprime. We can thus write $q = d\ell$ with $(\ell,d) = 1$ in this Lemma, which leads to*

$$G_d(Q) = \sum_{\delta/\ [d,\delta]\leq Q} h(\delta) = \sum_{\substack{q,\ell \\ (q,d)=1,\ell|d, \\ q\leq Q/d}} h(\ell)h(q)$$

i.e.

$$G_d(Q) = \frac{d}{|\mathcal{K}_d|} \sum_{\substack{\ell\leq Q/d \\ (\ell,d)=1}} h(\ell). \tag{6.7}$$

Since in classical literature $\mathcal{K}$ is always squarefree, authors tend to call $G_d(Q)$ what is in fact $|\mathcal{K}_d|G_d(Q)/d$ in our notation. We had the option of introducing another name, but we prefered to retain the same name in these lectures, for the reason that the most important value $G_1(Q)$ is unchanged. Note that it is usual to simply denote this latter value by $G(Q)$.

6.4. Bordering system associated to a compact set

We define here another sequence of sets $(\mathcal{L}_d)_{d\geq 1}$ complementary to $(\mathcal{K}_d)$: we set $\mathcal{L}_1 = \{1\}$ and $\mathcal{L}_{p^\nu} = \mathcal{K}_{p^{\nu-1}} - \mathcal{K}_{p^\nu}$, i.e. the set of elements of $x \in \mathbb{Z}/p^\nu\mathbb{Z}$ such that $\sigma_{p^\nu\to p^{\nu-1}}(x) \in \mathcal{K}_{p^{\nu-1}}$ but that do *not* belong to $\mathcal{K}_{p^\nu}$. We further define $\mathcal{L}_d$ by "multiplicativity". It is important to note, and that is different from what happens to $\mathcal{K}$, that *we do not have* $\mathcal{L}_\ell = \mathcal{L}_d/\ell\mathbb{Z}$ if $\ell|d$. Using $\mathbb{1}_\mathcal{A}$ to denote the characteristic function of $\mathcal{A}$, our definitions imply that

$$(6.8)\quad \begin{cases} \mathbb{1}_{\mathcal{L}_d} = \prod\limits_{p^\nu\|d}\bigl(\mathbb{1}_{\mathcal{K}_{p^{\nu-1}}} - \mathbb{1}_{\mathcal{K}_{p^\nu}}\bigr) = (-1)^{\omega(d)}\sum\limits_{\delta|d}\mu(d/\delta)\mathbb{1}_{\mathcal{K}_\delta} \\ \mathbb{1}_{\mathcal{K}_d} = \prod\limits_{p^\nu\|d}\bigl(\mathbb{1} - \mathbb{1}_{\mathcal{L}_p} - \mathbb{1}_{\mathcal{L}_{p^2}} - \cdots - \mathbb{1}_{\mathcal{L}_{p^\nu}}\bigr) = \sum\limits_{\delta|d}(-1)^{\omega(\delta)}\mathbb{1}_{\mathcal{L}_\delta}. \end{cases}$$

A remark on why one should introduce $\mathcal{L}$: to start with, let us note that classical sieve expositions stress more on the classes that one *excludes* modulo p, than on the classes that are retained, which in our setting means that the sets $\mathcal{L}_p$ are defined first, and the sets $\mathcal{K}_p$ are usually not specified. This is so because we usually exclude few classes, i.e. $\mathcal{L}_p$ is small while $\mathcal{K}_p$ is large. This notion of *small* and *large* is in fact what led to the nomenclature “large sieve”.

Introducing $\mathcal{K}_p$ allows us to get a geometrical setting, i.e. leads to a natural definition of $\mathcal{K}_d$ – while that of $\mathcal{L}_d$ is much less natural – and, in general, to smoother formulae for the main terms. However, when it comes to computing error terms, the fact that $\mathcal{L}_d$ has small cardinality in usual problems turns out to be extremely effective.

6.5. The compact relative to a κ-tuple of admissible shifts

We defined in definition 5.10 what is an admissible shift. Here is the associated compact set:

$$\mathcal{K}(\mathfrak{s})_d = \bigcap_{1\le i\le|\mathfrak{s}|}\bigl(\mathcal{U}_d - \mathfrak{s}(i)\bigr) \quad (\text{with } \mathcal{U}_d = (\mathbb{Z}/d\mathbb{Z})^*)$$

for a k-tuple $\mathfrak{s}$ of admissible shifts.

This compact set $\mathcal{K}(\mathfrak{s})$ is multiplicatively split and squarefree. The cardinality of $\mathcal{K}(\mathfrak{s})_p$ is $p - L(\mathfrak{s})$ when $p \ge L(\mathfrak{s}) + 1$, since all members of $\mathfrak{s}$ fall in this case in different residue classes modulo p. The cardinality of $\mathcal{K}(\mathfrak{s})_p$ is often larger for smaller p, but this has no impact whatsoever on the dimension on the sieve which remains $k = |\mathfrak{s}|$. Furthermore its associated bordering system is given by:

$$(6.9)\qquad \mathcal{L}(\mathfrak{s})_p = \bigl\{\mathfrak{s}(i) \bmod p,\ 1 \le i \le |\mathfrak{s}|\bigr\}.$$

CHAPTER 7

The Selberg sieve

The notion of sieving process goes back to Erathostenos, but its formalisation started only with the work of Legendre, and it is really Viggo Brun who started the modern line in [**9**]. This line of approach is said to be comobinatorial. Atle Selberg has later introduced another sieving process which we present below. The third branch in sieves has been initiated by Yu Linnik ate about the same time (1940-1945), but it will emerge as a full-fledged sieve process only from the seventies onward. Chapter 5 somehow is linked wityh this later line of thoughts.

The reader should also have a look at the excellent presentation of this material in [**37**]. The more advanced reader has to read fully [**29**].

7.1. A definition of sieve problems

To properly set the sieve problem, one needs two objects:

(1) A finite host sequence $\mathcal{A}$; for instance, as was the case upto now in these lectures, $\mathcal{A} = [M+1, M+N]$.

(2) A compact set $\mathcal{K}$, i.e. a finite collection of well-behaved – see chapter 6 – subsets $\mathcal{K}_d$ of $\mathbb{Z}/d\mathbb{Z}$.

The question is then to understand

$$\mathcal{S} = \{n \in \mathcal{A} \quad / \quad \forall d \leq D, \quad n \in \mathcal{K}_d\} \tag{7.1}$$

and, in particular, to evaluate its cardinality. We met this question already with $\mathcal{K}_d = (\mathbb{Z}/d\mathbb{Z})^*$ the set of invertible elements modulo d to reach the prime numbers and prove the Brun-Titchmarsh Theorem.

Il s'agit là du problème classique du crible dans notre optique. Notre présentation diffère des présentations que l'on peut trouver dans [**7**] ou dans les lectures on sieves de [**85**] sur plusieurs points :

(1) L'approche usuelle consiste à regarder les classes que l'on ôte et non celles que l'on garde. Ceci induit un manque de régularité des expressions que l'on manipule. On trouve toutefois la trace de notre façon de faire dans [**8**] où ils démontrent le théorème de Brun-Titchmarsh de façon étonnante. Cet article est d'ailleurs l'ancêtre de ce travail.

(2) Nous regardons ce qui se passe modulo p, mais aussi modulo $p^2, p^3, \ldots$ ce qui induit des complications notoires. [**30**] donne une solution partielle à cette question dans le cas où la suite hôte est un intervalle et [**83**] une solution plus complète, mais la complexité des expressions l'oblige là encore à contrôler le terme d'erreur par le grand crible ce qui limite l'utilisation à la suite hôte des entiers dans un intervalle. Par ailleurs, l'exposition de cette solution demande une dizaine de pages sans que la longueur contribue à la clarté.

(3) On se ramène usuellement à la condition $P(n) \equiv 0[q]$ pour un certain polynôme P par différentes astuces, ce qui est inutile ici.

7.2. An extremal problem

In our presentation of the Selberg sieve, we consider the following extremal problems

$$(7.2)\qquad \begin{cases} \sum_d \lambda_d^\sharp = 1 \quad , \quad \lambda_d^\sharp = 0 \quad \text{if } d \ge D \\ \text{Main term of } \displaystyle\sum_{M<n\le M+N} \Bigl(\sum_{d/n\in\mathcal{K}_d} \lambda_d^\sharp\Bigr)^2 \quad \text{minimal} \end{cases}$$

and

$$(7.3)\qquad \begin{cases} \lambda_1 = 1 \quad , \quad \lambda_d = 0 \quad \text{if } d \ge D \\ \text{Main term of } \displaystyle\sum_{M<n\le M+N} \Bigl(\sum_{d/n\in\mathcal{L}_d} \lambda_d\Bigr)^2 \quad \text{minimal.} \end{cases}$$

We switch from one problem to the other using (6.8) :

$$(7.4)\qquad \begin{cases} (-1)^{\omega(d)}\lambda_d = \displaystyle\sum_{d|\ell} \lambda_\ell^\sharp \quad , \quad \lambda_\ell^\sharp = \sum_{\ell|d} \mu(d/\ell)(-1)^{\omega(d)}\lambda_d, \\ \displaystyle\sum_{d/n\in\mathcal{L}_d} \lambda_d = \sum_{d/n\in\mathcal{K}_d} \lambda_d^\sharp. \end{cases}$$

Solving the first problem is very easy because $\mathcal{K}$ is multiplicatively split, and is performed via the diagonalization process of Selberg. Indeed, we write

$$\begin{aligned} \sum_{M<n\le M+N} \Bigl(\sum_{d/n\in\mathcal{K}_d} \lambda_d^\sharp\Bigr)^2 &= \sum_{d_1,d_2\le D} \lambda_{d_1}^\sharp \lambda_{d_2}^\sharp \sum_{\substack{M<n\le M+N \\ n\in\mathcal{K}_{[d_1,d_2]}}} 1 \\ &= \sum_{d_1,d_2\le D} \lambda_{d_1}^\sharp \lambda_{d_2}^\sharp \frac{|\mathcal{K}_{[d_1,d_2]}|}{[d_1,d_2]} N + \text{ error term} \end{aligned}$$

Set $\rho(d) = |\mathcal{K}_d|/d$ and let h be the solution of $1/\rho = \mathbb{1} \star h$ as in (6.4). We then have

$$\sum_{d_1,d_2 \le D} \lambda^\sharp_{d_1} \lambda^\sharp_{d_2} \frac{|\mathcal{K}_{[d_1,d_2]}|}{[d_1,d_2]} = \sum_{d_1,d_2 \le D} \lambda^\sharp_{d_1} \rho(d_1) \lambda^\sharp_{d_2} \rho(d_2) (\mathbb{1} \star h)((d_1,d_2))$$
$$= \sum_{q \le D} h(q) \biggl(\sum_{q|d \le D} \lambda^\sharp_d \rho(d) \biggr)^2 .$$

We comment on the above relations: first we note that any two randomly chosen integers have a small gcd, so that we indeed reduce the difficulty by exchanging lcm with gcd; the next problem is still the fact that d_1 and d_2 are linked and the introduction of h is a key idea to separate them fully. Pursuing the proof, we define

$$y_q = \sum_{q|d \le D} \lambda^\sharp_d \rho(d) \tag{7.5}$$

and recover the $\lambda^\sharp_d$'s from the y_q's by*

$$\rho(d) \lambda^\sharp_d = \sum_{d|q \le D} \mu(q/d) y_q \tag{7.6}$$

which enables us to establish that

$$1 = \sum_d \lambda^\sharp_d = \sum_q h(q) y_q . \tag{7.7}$$

We minimize the quadratic form $\sum h(q) y_q^2$ subject to the condition (7.7). On using Lagrange multipliers, we see optimal† y_q's should all be equal to $1/\sum_d h(d)$ i.e. $1/G_1(D)$.

Gathering our results we infer

$$\lambda^\sharp_d = \frac{d}{|\mathcal{K}_d|} \sum_{q \le D/d} \mu(q)/G_1(D) \quad \text{and} \quad \lambda_d = (-1)^{\omega(d)} G_d(D)/G_1(D). \tag{7.8}$$

Proof Equation (7.6) together with the fact that y_q is constant with value $1/G_1(D)$ gives the value of $\lambda^\sharp_d$. We next use the first equation

*Equation (7.5) may be seen as a linear system expressing the y_q's in terms of the $(\lambda^\sharp_d \rho(d))$'s. This system being in triangular form, the $(\lambda^\sharp_d \rho(d))$'s are uniquely determined in terms of the y_q's. The reader will check that the RHS of (7.6) verifies this system, and hence, is equal to $\lambda^\sharp_d \rho(d)$.

†When $h(q)$ vanishes, the corresponding value of y_q has no influence whatsoever; the corresponding λ_q will always appear with coefficient $h(q)$, The solution y_q we choose is the one that yields uniform formulae.

of (7.4) to get

$$\begin{aligned}G_1(D)(-1)^{\omega(d)}\lambda_d &= \sum_{d|\ell}\lambda_\ell^\sharp = \sum_{d|\ell}\frac{\ell}{|\mathcal{K}_\ell|}\sum_{q\le D/\ell}\mu(q)\\ &= \sum_{k\le D}\Biggl(\sum_{\substack{k=q\ell,\\ d|\ell}}\frac{\ell}{|\mathcal{K}_\ell|}\mu(q)\Biggr) = G_d(D)\end{aligned}$$

by Lemma 6.1. □

From the information theory point of view, going from $(\lambda_d^\sharp)$ to (λ_d) may be explained by the following remark : when writing $n \in \mathcal{K}_{p^\nu}$, we forget we already know that $n \in \mathcal{K}_{p^{\nu-1}}$; Removing this redundancy leads to $(\mathcal{L}_d)$ and to (λ_d).

Note that Lemma 6.2 tells us simply that $|\lambda_d| \le 1$.

As for the cardinality of $\mathcal{S}$ (defined in (12.2)), we directly get

$$\begin{aligned}|\mathcal{S}| &\le \sum_{n\le N}\Biggl(\sum_{d/n\in\mathcal{K}_d}\lambda_d^\sharp\Biggr)^2 = \sum_{n\le N}\Biggl(\sum_{d/n\in\mathcal{L}_d}\lambda_d\Biggr)^2\\ &\le \frac{N}{G_1(D)} + \Biggl(\sum_d |\mathcal{L}_d||\lambda_d|\Biggr)^2 \qquad (7.9)\end{aligned}$$

Going from $(\lambda_d^\sharp)$ to (λ_d) is thus extremely important in reducing the error term, thanks to Lemma 6.2. The reader should notice that this switching of variables is fully mechanical and relies only the identities (7.4). This fact will be used in section 9.2.

In [**83**] and [**66**], the reader will find another exposition and in [**30**] closely related material.

7.3. A general expression

Let $(u_n)_{n\in\mathbb{Z}}$ be a weighted sequence, the weights u_n being non-negative and such that $\sum_n u_n < +\infty$. Let $\mathcal{K}$ be a multiplicatively split compact set. We assume that there exists a multiplicative function $\sigma^\sharp$, a parameter X and a function $R_d^\sharp$ such that

$$\sum_{n\in\mathcal{K}_d} u_n = \sigma^\sharp(d)X + R_d^\sharp. \qquad (7.10)$$

We assume further that $\sigma^\sharp$ is non-negative and decreases on powers of primes (a likely hypothesis if one conceives of $\sigma^\sharp(d)$ as being a density), which translates into $\sigma^\sharp(q) \ge \sigma^\sharp(d)$ whenever $q|d$. Equivalently, we assume the existence of σ and R_d such that

$$\sum_{n\in\mathcal{L}_d} u_n = \sigma(d)X + R_d \qquad (7.11)$$

but the non-increasing property on chains of multiples is way less obvious to state. Switching from (7.10) to (7.11) is readily done through (6.8). There comes

$$\begin{cases} (-1)^{\omega(d)}\sigma(d) = \sum_{\delta|d} \mu(d/\delta)\sigma^\sharp(\delta), \\ \sigma^\sharp(d) = \sum_{\delta|d} (-1)^{\omega(\delta)}\sigma(\delta). \end{cases} \tag{7.12}$$

All the analysis of section 7.2 applies, except we are to change the definition of our G-functions. First, h is the solution of

$$\frac{1}{\sigma^\sharp(d)} = \sum_{q|d} h(q) \tag{7.13}$$

(compare with (6.5)), that is to say

$$h(d) = \prod_{p^\nu \| \delta} \left(\frac{1}{\sigma^\sharp(p^\nu)} - \frac{1}{\sigma^\sharp(p^{\nu-1})} \right) \geq 0. \tag{7.14}$$

Proceeding as in section 7.2, but with $\rho = \sigma^\sharp$, we get

$$\sum_{n \in \mathcal{S}} u_n \leq \frac{X}{G_1(z)} + \sum_{d_1, d_2} \lambda_{d_1}\lambda_{d_2} R_{[d_1,d_2]}, \tag{7.15}$$

with $\mathcal{S}$ defined by (12.2). Notice that we still have $|\lambda_d| \leq 1$ as in the simpler case of intervals.

7.4. On the number of prime twins

We will give an upper bound for the number of prime twins up to N, as N goes to infinity, by applying Selberg sieve. The compact set we take is simply $\mathcal{K} = \mathcal{U} \cap (\mathcal{U} - 2)$ as was the case then. It is multiplicatively split as well as squarefree. For the associated function h, we readily find that

$$\begin{cases} h(2) = 1 \quad \text{and} \quad h(2^\nu) = 0 & \text{if } \nu \geq 2, \\ h(p) = 2/(p-2) \quad \text{and} \quad h(p^\nu) = 0 & \text{if } p \geq 3 \text{ and } \nu \geq 2. \end{cases}$$

This gives us

$$G_1(Q) = \sum_{q \leq Q} \mu^2(q) \prod_{\substack{p|q \\ p \neq 2}} \frac{2}{p-2} \tag{7.16}$$

to which we apply Theorem 3.1 with $\kappa = 2$ to get

$$G_1(Q) \sim \frac{1}{4} \prod_{p \geq 3} \frac{(p-1)^2}{p(p-2)} (\operatorname{Log} Q)^2. \tag{7.17}$$

We again choose $Q = \sqrt{N}/\operatorname{Log} N$ to find that

$$\left|\left\{p \le N \ /p+2 \text{ is prime}\right\}\right| \le 16(1+o(1)) \prod_{p\ge 3} \frac{p(p-2)}{(p-1)^2} N/(\operatorname{Log} N)^2$$

a bound that is 8 times larger than its conjectured value. [**86**] establishes the above inequality for all $N > 1$ with no $o(1)$ term. If we were to use the Bombieri-Vinogradov Theorem , we would get a bound only 4 times off the expected one. Note that [**95**] reduces this constant to 3.3996; that such an improvement holds only when we look at prime twins located on the initial segment $[1, N]$, contrarily to the above bound which remains valid for *any* interval of length N.

7.5. On a subset of prime twins

Our aim here is to give an upper bound for the number of primes p not more than N that are such that $p+2$ is a prime, while $p+1$ is squarefree. The compact set $\mathcal{K}$ we choose is defined by split multiplicativity: for prime p, $\mathcal{K}_p$ is $\mathcal{U}_p \cap (\mathcal{U}_p + 2)$ while $\mathcal{K}_{p^2}$ is the set of invertibles that are not congruent to -2 modulo p and not congruent to -1 modulo p^2. For higher powers of p, $\mathcal{K}_{p^\nu}$ is defined by trivially lifting $\mathcal{K}_{p^2}$, and so will be of no interest. This yields

$$\begin{cases} |\mathcal{K}_2| = 1, \ |\mathcal{K}_4| = 1, \\ |\mathcal{K}_p| = p-2, \ |\mathcal{K}_{p^2}| = p(p-2) - 1 = p^2 - 2p - 1 \text{ if } p \ge 3. \end{cases}$$

But now the host sequence is that of primes p weighted with a $\operatorname{Log} p$ each so that

$$\sigma(d) = |\mathcal{K}_d|/\phi(d) \tag{7.18}$$

Of course $\mathcal{L}_d \cap \mathcal{U}_d$ has at most one class (class -2 modulo p and class -1 modulo p^2), implying that the error term

$$R_d = \sum_{\substack{p\le N \\ p\in\mathcal{L}_d\cap\mathcal{U}_d}} \operatorname{Log} p - \frac{|\mathcal{L}_d \cap \mathcal{U}_d| N}{\phi(d)} \tag{7.19}$$

may be controlled by

Lemma 7.1 (Bombieri-Vinogradov) *For any $B \ge 0$, there exists an $A \ge 0$ such that*

$$\sum_{q\le Q} \max_{y\le N} \max_{a \ mod^* q} \left| \sum_{\substack{p\le N \\ p\equiv a[q]}} \operatorname{Log} p - \frac{N}{\phi(q)} \right| \ll N/(\operatorname{Log} N)^B$$

for $Q = \sqrt{N}/(\operatorname{Log} N)^A$.

By taking $B = 2$, this yields

$$\sum_{d_1,d_2\le D} |\lambda_{d_1}\lambda_{d_2}R_{[d_1,d_2]}| \ll N/(\operatorname{Log} N)^2$$

provided $D^2 = \sqrt{N}/(\operatorname{Log} N)^A$. As for the main term, we check that

$$\begin{cases} h(2) = 0,\ h(4) = 1, \\ h(p) = \dfrac{1}{p-2},\ h(p^2) = \dfrac{p-1}{p^3-4p^2+3p+2} \text{ if } p \ge 3. \end{cases}$$

Theorem 3.1 applies with $\kappa = 1$. We finally get

Theorem 7.2 *The number of primes $p \le N$ that are such that $p+1$ is squarefree and $p+2$ is prime does not exceed*

$$4(1+o(1)) \prod_{p\ge 3} \frac{p^2-2p-1}{(p-1)^2} \frac{N}{\operatorname{Log}^2 N}$$

as N goes to infinity.

This bound is 4 times larger than what is conjectured but the main point here is that this bound is indeed smaller than the one one gets for prime twins (see preceding section) by a large factor, namely

$$2\prod_{p\ge 3} \frac{p(p-2)}{p^2-2p-1} = 3.426\ldots$$

7.6. Computing the G-functions in the case of the prime κ-tuples

We want to compute the G-functions is the case of a κ-tuple $\mathfrak{s}$ of admissible shifts and when the host sequence is the one of primes.

Exercise 7.3 *We want to bound above the number of integers n between 1 and N that are such that n^2+1 is also a prime number. We set $T(X) = X^2 + 1$. We define $\mathcal{K}_q$ to be the set of classes x modulo q that are such that $T(x) = x^2+1$ is invertible modulo q, i.e.*

$$x \in \mathcal{K}_q \iff T(x) \in \mathcal{U}_q.$$

(1) Show that $|\mathcal{K}_2| = 1$, $\mathcal{K}_p = \mathbb{Z}/p\mathbb{Z}$ when $p \equiv 3[4]$, and $|\mathcal{K}_p| = p-2$ otherwise.

(2) Let p be a prime number and let $\alpha \ge 2$. We denote by σ the canonical surjection from $\mathbb{Z}/p^\alpha\mathbb{Z}$ to $\mathbb{Z}/p\mathbb{Z}$. Show that $x \in \mathbb{Z}/p^\alpha\mathbb{Z}$ lies inside $\mathcal{K}_{p^\alpha}$ if and only if $\sigma(x)$ is in $\mathcal{K}_p$. Show that $\mathcal{K} = (\mathcal{K}_q)_{q\ge 1}$ is multiplicatively split as well as squarefree.

(3) Show that $g(q) = \sum_{d|q} \mu(q/d)d/|\mathcal{K}_d|$ is a positive multiplicative function.

(4) *Show that there exists a positive constant C_1 such that the number Z of integers $n \le N$ such that $n^2 + 1$ is a prime number verifies that, for every $N \ge 2$:*

$$Z \le C_1 N / \operatorname{Log} N.$$

CHAPTER 8

Introduction to the weighted sieve

The theory of the weighted sieve can be developped in a general context, but we will follow here the problem that is its usual application, namely the one of twin primes. Primes p such that $p+2$ is also a prime have been termed *twins* by the german mathematician P. Stäckel, or so says the first chapter of [**89**]. The assertion claiming it to be a conjecture of Euclid seems to be frivolous and the first mention of it I found lies in [**15**]. The general problem of solving linear equations in prime variables had of course already been addressed much earlier (in 1742) by Goldbach and most of the early effort on primes concentrated on two problems: showing that every plausible arithmetic progressions contained infinitely many of them and proving the Goldbach conjecture. There is no record of any progress on the problem of twin primes until [**9**] proved that the sequence of such primes is much smaller than the sequence of primes. Its method developped into the theory of sieves, which has now several branches. The sieve is very efficient to determine upper bounds but is much less powerful to provide us with lower bounds, let them be extremely weak.

8.1. A beginner's stroll with historical comments

Let us examin the twin prime problem more closely. We would like to show that the sum

$$S = \sum_{n \le N} \Bigl(3 - \sum_{\substack{p \mid n(n+2),\\ p \le N+2}} 1 \Bigr) \tag{8.1}$$

goes to infinity (where p runs through primes). Indeed this would imply that there exist infinitely many integers n such that $n(n+2)$ has at most two prime factors p: these are bound to be n and $n+2$ who would thus be prime! Let us try to pursue this strategy. We develop (8.1) and get

$$\begin{aligned} S &= 3\sum_{n\le N+2} 1 - \sum_{p\le N+2}\ \sum_{\substack{n\le N,\\ n(n+2)\equiv 0[p]}} 1 \\ &= (3+o(1))N - \sum_{p\le N+2}\frac{2N}{p} - \sum_{p\le N+2}\Biggl(\sum_{\substack{n\le N,\\ n(n+2)\equiv 0[p]}} 1 - \frac{2N}{p}\Biggr) \end{aligned}$$

The last sum is at most $\mathcal{O}(N/\operatorname{Log} N)$. However the second sum is readily seen to be equivalent to $-N\operatorname{Log}\operatorname{Log} N$, since

$$\sum_{p\le N+2} 1/p = \operatorname{Log}\operatorname{Log} N + \mathcal{O}(1).$$

This ruins the whole argument since this implies that S tends to $-\infty$!

Let us try to see how to improve on the above trial. A first remark consists in noticing that the *host sequence* from which we are trying to detect prime twins is too large: indeed if n were restricted to range only prime numbers p' such that $p'+2$ is a prime, the approach would work! Except that we do not know enough on this sequence... So the problem becomes finding a larger sequence with which we can still work but for which we have enough information. This will be provided to us by a process akin to the Selberg sieve: we shall thus have a sequence of weights $\beta(n)$ at our disposal that takes value 1 on prime twins, is otherwise nonnegative and gives more weight to integers such that $n(n+2)$ has few prime factors. This latter process is what we call the *weighted sieve.*

The second remark, which is connected to the previous one, is that one could improve on the sieving coefficient $\Bigl(3-\sum_{\substack{p|n(n+2),\\ p\le N+2}} 1\Bigr)$: indeed, this coefficient becomes very negative when the integer $n(n+2)$ has many prime factors. We have tried in [**73**] to include divisors that have more prime factors, but the fact that we do not know anymore that $\beta(n)$ is indeed 0 or 1 renders the combinatoric dificult, and a lower bound sieve approach intractable (as far as I have been able to understand). We got however stronger results on including such divisors, results that we describe in [**73**].

The process consists thus then in adding a weight*, as in (8.1), to a sieve device, and this is how it appeared in history, somewhat contrarily to the way we have motivated this method.

We find in [**56**] (and later in [**57**]) the first combination of an upper bound sieve and a system of weights to prove the existence of integers having few prime factors. In the aforementioned paper, Kuhn proves

*The word "weight" is overloaded in this theory and we shall adopt a clearer terminology at the end of this introduction.

that the interval $[X, X+\sqrt{X}]$ contains numbers having at most 4 prime factors, provided X be large enough.

[**76**] (translated in [**77**]) proved that there exists a constant r such that every large even integer N can be written as a prime and an integer having at most r prime factors. We will say *a* P_r. This developped in the main Theorem of [**12**] where the author proves that $r = 2$ is possible. Concerning P_2 in short interval, [**13**] proves that every interval $[X, X + X^{1/2}]$ contains a P_2, when X is large enough. [**39**] reduced this exponent to 0.477, [**48**] to 0.45 while [**94**] reduced it further to 0.44.

8.2. Results to be proven

In these lectures, we will first prove the following basic result:

Theorem 8.1 *Let $\mathfrak{s}$ be a κ-tuple of admissible shifts. We can find infinitely many primes p such that $\prod_{2\le i\le|\mathfrak{s}|}(p+\mathfrak{s}(i))$ has at most $\kappa(\operatorname{Log}\kappa+4)$ prime factors*

Proving the κ-tuple conjecture amounts to replacing the $\kappa(\operatorname{Log}\kappa+4)$ by κ in the above. This is largely out of reach of today's techniques, but the bound 2κ is a plausible goal. However the bound $\kappa\operatorname{Log}\kappa$ has stayed for a very long time. The reader will find in [**73**] some more material on this question and we prove the existence of infinitely many κ-tuples with *exactly* $\kappa\operatorname{Log}\kappa+\mathcal{O}(\kappa)$ prime factors. This can be extended to prove the existence of infinitely many κ-tuples with *exactly* $\lambda\kappa\operatorname{Log}\kappa+\mathcal{O}(\kappa)$ prime factors for any chosen $\lambda \ge 1$. On a slightly different note, Heath-Brown in [**43**] developped an idea of Selberg on the twin prime conjecture, and investigated the problem of bounding the number of prime factors of each $n+h_i$. He has obtained that, given any admissible κ-tuple, there are infinitely many tuples $(n, n+h_2, \cdots, n+h_\kappa)$ such that each $n+h_i$ has at most $2(1+o(1))\operatorname{Log}\kappa/\operatorname{Log}2$ prime factors.

It thus came as a surprise when Craig Franze proved in june 2010 in [**28**] a bound of the shape $(1/2)\kappa\operatorname{Log}\kappa+\mathcal{O}(\kappa)$. This is the first time the barrier $\kappa\operatorname{Log}\kappa$ is broken. He proves also precise bounds when $\kappa \in \{5, 6, 7, 8, 9, 10\}$ that improve on the ones that were known before.

We are not going to detail here the main breakthrough that is Franze's result, but prove more modestly the following.

Theorem 8.2 *Let $\mathfrak{s}$ be a κ-tuple of admissible shifts. We can find infinitely many primes p such that $\prod_{1\le i\le|\mathfrak{s}|}(n+\mathfrak{s}(i))$ has at most $n_0(\kappa)$ prime factors, where $n_0(\kappa)$ is given in the table below.*

Prior to Franze's result, the method we proposed led to the best values known for every values of κ and improved on these values as soon as κ was somewhat large (starting from $\kappa = 8$). Both methods can most probably be mixed, and there are already quite a number of issues that are not optimaly resolved in the proof we present.

We included ine the table below results from [**69**], [**96**], [**79**], [**43**], [**18**], [**45**], table 11.1 of [**19**]. It emerges from this table that the method we propose equals the best of the others for small values of κ and start showing its teeth when $\kappa = 8$. Even as it stands Franze bounds are already better.

κ	Porter	Xie	Salerno	Heath-Brown	Diamond & Halberstam	Ho & Tsang	Franze	$n_0(\kappa)$
1	3							2
2	9		6	5	5	5		5
3	14		10	9	8	8		8
4	20	14	14	13	12	12		12
5	27	18	18	17	16	16	15	16
6	33	23	23	21	20	20	19	20
7	40	27		26	25	24	23	24
8	46	32		32	29	29	27	28
9	53	37		39	34	33	31	32
10	60	42		45	39	38	34	37
11						44		41
12						48		46
13						53		51
14						58		55
15						63		60
16						69		65
17						74		70
18						80		75
19						85		80
20						91		85
21						97		90
22						103		95

$n_0(\kappa)$

Maybe we should explicitly state that the values shown are upper bounds to what is accessible via the method developped here, though we believe our choice of parameters to be very close to the optimal one.

Special values: [**97**] shows that there are infinitely many primes such that $(p+2)(p+6)(p+8)$ has less than 12 prime factors. The reader may also consult [**33**], [**34**] and [**52**] with benefit. Finally, the papers [**32**] and

[**14**] (also presented in [**54**]) are just outside the scope of this presentation but use techniques and ideas very close to what is presented here.

Terminology

Our problem will consist in estimating a sum of the shape

$$\sum_n \Bigl(3 - \sum_{\substack{p|n(n+2),\\ p\le y}} 1\Bigr)\beta(n).$$

We shall call the sequence $\beta(n)$ in the above summation, which will come from a process close to the (upper) Selberg sieve, the *host sequence*. They are sometimes refered to in the litterature as the Selberg weights.

The coefficient $(3 - \sum_{\substack{p|n(n+2),\\ p\le y}} 1)$ or what will be placed there for a similar effect the *sieve coefficient*. These coefficients will be extremely linked with lower sieve procedures and in particular may well be negative.

There are of course some arbitrariness in how to split our coefficients into two parts, but all that will become clear.

Two issues will make matters somewhat more complicated: the construction of the Selberg coefficients depends on yet another sequence will be at this level be called (and be treated as) an host sequence. Secondly, we shall modify these Selberg coefficients, which will then be used as the current host sequence by employing... some weights! The word “weight” will be reserved for these, except in the expression “the weighted sieve”.

CHAPTER 9

The approach in the large

We restrict our attention in these lectures to the prime κ-tuple and to sieving coefficients taking care of prime divisors only. It is however not more difficult to accomodate a general setting. We follow the beginning of [**73**] closely.

9.1. A general framework

For an admissible shift $(h_1, \cdots, h_\kappa)$, we define

$$\mathcal{K}(h_1, \cdots, h_\kappa)_d = \bigcap_{1\le i\le\kappa} \bigl(\mathcal{U}_d - h_i\bigr) \quad (\text{with } \mathcal{U}_d = (\mathbb{Z}/d\mathbb{Z})^*). \tag{9.1}$$

When n lies in $\mathcal{K}(h_1, \cdots, h_\kappa)_d$, then $n + h_i$ falls in $\mathcal{U}_d$, i.e. is prime to d, for all i from 1 to κ.

The definition (9.1) shows clearly that this compact set is multiplicatively split and square-free. We need another one, $\mathcal{K}^*$, that will be successively $\mathcal{U}_d - h_1, \mathcal{U}_d - h_2$, ..., $\mathcal{U}_d - h_\kappa$ in [**73**] and which we will simply take to be $\mathcal{K}$ here. In general we simply assume that

(H_1) $\mathcal{K}$ is a multiplicatively split compact set;
(H_2) $\mathcal{K}^*$ is a square-free multiplicatively split compact set that contains $\mathcal{K}$.

We associate to $\mathcal{K}^*$ (and to any multiplicatively split compact set) its bordering system $(\mathcal{L}^*_d)_{d\ge1}$ as in section 6.4. Each $\mathcal{L}^*_d$ is the subset of $\mathbb{Z}/d\mathbb{Z}$ defined by

- $\mathcal{L}^*_1 = \{1\}$ and $\mathcal{L}^*_d = \emptyset$ when d is not square-free.
- When d_1 and d_2 are co-prime, $\mathcal{L}^*_{d_1d_2}$ is in bijection via the Chinese remainder map to $\mathcal{L}^*_{d_1} \times \mathcal{L}^*_{d_2}$.
- $\mathcal{L}^*_p = \mathbb{Z}/p\mathbb{Z} \setminus \mathcal{K}^*_p$.

This may look complicated, but the situation clears when one looks at characteristic functions, see (6.8). We will also use the bordering system $(\mathcal{L}_d)_{d\ge1}$ associated with $\mathcal{K}$. Note that the condition $\mathcal{K}^* \supset \mathcal{K}$ has the following consequence: when p is a prime number and a is a positive integer, we have

$$\mathbb{1}_{\mathcal{K}_{p^a}} \cdot \mathbb{1}_{\mathcal{L}^*_p} = \mathbb{1}_{\mathcal{K}_{p^a}} \cdot (\mathbb{1} - \mathbb{1}_{\mathcal{K}^*_p}) = 0. \tag{9.2}$$

Having this preparation at hand, we can present the main actor of this paper, namely the sum

$$S((a_{d^*})_{d^*};\mathcal{K},\mathcal{K}^*)=\sum_{n\le N}\biggl(\sum_{d^*/n\in\mathcal{L}^*_{d^*}}a_{d^*}\biggr)\biggl(\sum_{d/n\in\mathcal{K}_d}\tilde{\lambda}^\sharp_d\biggr)^2. \tag{9.3}$$

We assume that a_{d^*} vanishes when $d^*>D^*$. The coefficients $(\tilde{\lambda}^\sharp_d)_d$ are completely free for us to choose. We simply assume that they vanish when $d>Q$, for some parameter Q. We need some more material from sieve theory. We define the coefficients $(\tilde{\lambda}_d)_d$ by

$$\tilde{\lambda}_d=(-1)^{\omega(d)}\sum_{d|\ell}\tilde{\lambda}^\sharp_\ell. \tag{9.4}$$

We have (see (7.4) or [**72**, (11.5)])

$$\tilde{\lambda}^\sharp_d=\sum_{d|\ell}\mu(\ell/d)(-1)^{\omega(\ell)}\tilde{\lambda}_\ell\quad\text{and}\quad\sum_{d/n\in\mathcal{K}_d}\tilde{\lambda}^\sharp_d=\sum_{d/n\in\mathcal{L}_d}\tilde{\lambda}_d. \tag{9.5}$$

In practice, the condition $n\in\mathcal{K}_d$ leads to easier treatment of the main term, while the $\tilde{\lambda}_d$'s will be smaller, leading to a better treatment of the error term. We finally introduce the multiplicative function

$$h(d)=\prod_{p^\nu\|d}\Bigl(\frac{p^\nu}{|\mathcal{K}_{p^\nu}|}-\frac{p^{\nu-1}}{|\mathcal{K}_{p^{\nu-1}}|}\Bigr).$$

We need to handle averages of this function and we follow [**37**, Chapter 5] (see also [**36**]). Condition (Ω_1) therein is introduced in the fourth part of the first chapter, page 49, but it is expedient to assume a much stronger hypothesis, namely

$$h(p)\ll\kappa/p. \tag{H_4}$$

Our main hypothesis on h is a slight simplification of what is called $(\Omega_2(\kappa,L))$ at the beginning of [**37**, Chapter 5]. It reads

$$\sum_{p\le x}\frac{h(p)\operatorname{Log}p}{p}=\kappa\operatorname{Log}x+\mathcal{O}(1). \tag{H_3}$$

This implies classically, via Theorem 3.1, that

$$\sum_{\delta\le x}h(\delta)=A(\operatorname{Log}x)^\kappa+\mathcal{O}((\operatorname{Log}(2x))^{\kappa-1})$$

when $x\le Q$, by using [**37**, Lemma 5.3, 5.4]. We deduce from (H_3) and (H_4) the following weaker form which will be easier to use:

$$\sum_{\delta\le x}h(\delta)=A(\operatorname{Log}x)^\kappa+\mathcal{O}(\alpha Y).\quad(1\le x\le Q) \tag{9.6}$$

where we set

$$Y=A(\operatorname{Log}Q)^\kappa. \tag{9.7}$$

When $\mathcal{K} = \mathcal{K}(h_1, \cdots, h_\kappa)$, the constant A is equal to the constant $\mathscr{C}$ of Theorem **??**. It is > 0 when $(h_1, \cdots, h_\kappa)$ is admissible, or, and this is an equivalent statement, when $\mathcal{K}(h_1, \cdots, h_\kappa)$ is non-empty.

9.2. Generalisation of a formula of Bombieri

The quantity (9.3) has a summation over four variables (n, d^*, d and d'). We take care here of the summation over n. We set

$$\gamma(d^*) = |\mathcal{L}^*_{d^*}|/d^* \tag{9.8}$$

as well as

$$\boxed{S_0((a_{d^*})_{d^*};\mathcal{K},\mathcal{K}^*) = \sum_{d^*,\delta} \gamma(d^*) a_{d^*} h(\delta) \Biggl(\sum_{\substack{\delta|d,\\ (d,d^*)=1}} |\mathcal{K}_d| \tilde{\lambda}^\sharp_d / d \Biggr)^2 .} \tag{9.9}$$

Lemma 9.1 *We have*

$$S((a_{d^*})_{d^*};\mathcal{K},\mathcal{K}^*) = N S_0((a_{d^*})_{d^*};\mathcal{K},\mathcal{K}^*) + \mathcal{O}\Biggl(\sum_{d_1^*,d_2,d_3} |a_{d_1^*}||\tilde{\lambda}_{d_2}||\tilde{\lambda}_{d_3}||\mathcal{L}^*_{d_1}||\mathcal{L}_{d_2}||\mathcal{L}_{d_3}| \Biggr).$$

Proof We first revert to $(\mathcal{L}_d)$ on invoking (9.5) and get

$$S((a_{d^*})_{d^*};\mathcal{K},\mathcal{K}^*) = \sum_{d_1^*,d_2,d_3} a_{d_1^*} \tilde{\lambda}_{d_2} \tilde{\lambda}_{d_3} \sum_{\substack{n \le N,\\ n \in \mathcal{L}^*_{d_1^*} \cap \mathcal{L}_{d_2} \cap \mathcal{L}_{d_3}}} 1.$$

Note that $\mathcal{L}^*_{d_1^*} \cap \mathcal{L}_{d_2} \cap \mathcal{L}_{d_3}$ vanishes when d_1^* is not square-free, or when there is a prime p and two distinct powers $a \geq 1$ and $b \geq 1$ that divides respectively d_2 and d_3. The reader will conclude that this set defines modulo $[d_1^*, d_2, d_3]$ a subset of cardinality at most $|\mathcal{L}^*_{d_1}||\mathcal{L}_{d_2}||\mathcal{L}_{d_3}|$. Concerning the main term, we divide it by N and write it as

$$\mathrm{M} = \sum_{d_1^*,d_2,d_3} a_{d_1^*} \tilde{\lambda}_{d_2} \tilde{\lambda}_{d_3} \frac{|\mathcal{L}^*_{d_1^*} \cap \mathcal{L}_{d_2} \cap \mathcal{L}_{d_3}|}{[d_1^*, d_2, d_3]}.$$

It can be defined as the limit when N goes to infinity of $S((a_{d^*})_{d^*};\mathcal{K},\mathcal{K}^*)/N$. To compute this limit we use (9.5) and switch to the $\tilde{\lambda}^\sharp_d$. This gives us

$$\mathrm{M} = \sum_{d_1^*,d_2,d_3} a_{d_1^*} \tilde{\lambda}^\sharp_{d_2} \tilde{\lambda}^\sharp_{d_3} \frac{|\mathcal{L}^*_{d_1^*} \cap \mathcal{K}_{d_2} \cap \mathcal{K}_{d_3}|}{[d_1^*, d_2, d_3]}.$$

Note that $\mathcal{K}_{d_2} \cap \mathcal{K}_{d_3} = \mathcal{K}_{[d_2,d_3]}$. We use (9.2) to introduce the condition $(d_1^*, d_2 d_3) = 1$. This gives us

$$\frac{|\mathcal{L}^*_{d_1^*} \cap \mathcal{K}_{d_2} \cap \mathcal{K}_{d_3}|}{[d_1^*, d_2, d_3]} = \frac{|\mathcal{L}^*_{d_1^*}|}{d_1^*} \frac{|\mathcal{K}_{d_2} \cap \mathcal{K}_{d_3}|}{[d_2, d_3]}.$$

We complete the separation of d_2 and d_3 via the diagonalisation process of Selberg, i.e. we write

$$\frac{|\mathcal{K}_{d_2} \cap \mathcal{K}_{d_3}|}{[d_2, d_3]} = \frac{|\mathcal{K}_{d_2}||\mathcal{K}_{d_3}|}{d_2 d_3} \sum_{\substack{\delta | d_1, \\ \delta | d_2}} h(\delta).$$

The Lemma follows readily. □

This Lemma generalizes [**6**, Theorem 18]. This same formula occurs as [**34**, Section 7.3.1, Lemma 1]. This is also [**84**, (5.6')]. Our proof is much shorter than the initial one of Bombieri. Greaves's proof is also remarkably short and shares with the above one the fact of treating the variable d_1^* in a distinct manner. Our switching between $\mathcal{L}$ and $\mathcal{K}$ as usual enables us to extend the proof to the case when $\mathcal{K}$ is not assumed to be square-free.

CHAPTER 10

Another family of Selberg coefficients

It is time to narrow our family of host sequences. But to do so, we will first develop some material to motivate our choice.

10.1. Investigating the Selberg coefficients

Let $M = \mathrm{lcm}(d \le Q)$ and let us look at $\mathcal{K}_M \subset \mathbb{Z}/M\mathbb{Z}$. We assume momentarily that the compact set satisfies the Johnsen-Gallagher condition (6.4).

Let us expand the characteristic function of $\mathcal{K}_M$ in Fourier series:

$$\begin{aligned}
\mathbb{1}_{\mathcal{K}_M}(n) &= \sum_{b \bmod M} \biggl(\frac{1}{M} \sum_{c \in \mathcal{K}_M} e(-bc/M)\biggr) e(bn/M) \\
&= \sum_{d|M} \sum_{a \bmod {}^* d} \biggl(\frac{1}{M} \sum_{c \in \mathcal{K}_M} e(-ac/d)\biggr) e(an/d) \\
&= \sum_{d|M} \sum_{a \bmod {}^* d} \biggl(\frac{|\mathcal{K}_M|}{M|\mathcal{K}_d|} \sum_{c \in \mathcal{K}_d} e(-ac/d)\biggr) e(an/d)
\end{aligned}$$

where we have used the Johnsen-Gallagher condition (see (JG)). We define

$$\psi_d^*(n) = \sum_{a \bmod {}^* d} \biggl(\frac{1}{|\mathcal{K}_d|} \sum_{c \in \mathcal{K}_d} e(-ac/d)\biggr) e(an/d). \tag{10.1}$$

These functions are, up to a multiplicative factor, the pseudo-characters introduced by Selberg in 1973 (see [**6**], [**65**] as well as [**53**]). An L^2 approximation of $\mathbb{1}_{\mathcal{K}_M}$ is thus given by

$$\frac{|\mathcal{K}_M|}{M} \sum_{d \le Q} \psi_d^*(n). \tag{10.2}$$

It is also possible to define $\psi_d^*(n)$ by Moebius inversion since we readily verify that

$$\sum_{d|q} \psi_d^* = \frac{q}{|\mathcal{K}_q|} \mathbb{1}_{\mathcal{K}_q}. \tag{10.3}$$

The coefficient $q/|\mathcal{K}_q|$ is somewhat mysterious and explained in [**72**, Section 9.4]. Inverting the above equation leads to the definition

$$\psi_d^* = \sum_{q|d} \mu(d/q) \frac{q}{|\mathcal{K}_q|} \mathbb{1}_{\mathcal{K}_q}. \tag{10.4}$$

This definition is valid whether $\mathcal{K}$ verifies condition (JG) or not, and is thus the one we take in general. Note that $\psi_d^*(n) = h(d)$ as soon as n belongs to $\mathcal{K}_d$. As a consequence, the function

$$\sum_{d\le Q} \psi_d^* \tag{10.5}$$

is constant over $\mathcal{K}_M$. This is the usual Selberg coefficient up to a normalising multiplier. Indeed, we find that ....

10.1.1. An intermezzo: the sieve bound via local models. Equation (10.1) can be rewritten in the form

$$\psi_d^*(n) = \sum_{a \bmod {}^* d} \hat{\psi}^*(a/d) e(an/d). \tag{10.6}$$

with

$$\hat{\psi}^*(a/d) = \frac{1}{|\mathcal{K}_d|} \sum_{c\in\mathcal{K}_d} e(-ac/d) \tag{10.7}$$

In case $\mathcal{K}_d = \mathcal{U}_d$, we have $\psi_d^*(n) = \mu(d)\,\mathrm{c}_d(n)/\phi(d)$ where $\mathrm{c}_d(n)$ is the Ramanujan sum defined in (5.3). Here is a preliminary Lemma:

Lemma 10.1 *We have*

$$\sum_{a\ mod^*\ d} |\hat{\psi}^*(a/d)|^2 = h(d).$$

Proof We have

$$\sum_{a \bmod {}^* d} |\hat{\psi}^*(a/d)|^2 = \frac{1}{|\mathcal{K}_d|^2} \sum_{c_1,c_2\in\mathcal{K}_d} \mathrm{c}_d(c_1-c_2) = \frac{1}{|\mathcal{K}_d|^2} \sum_{\delta|d} \delta\mu(d/\delta) \sum_{\substack{c_1,c_2\in\mathcal{K}_d,\\ c_1\equiv c_2[\delta]}} 1$$

$$= \frac{1}{|\mathcal{K}_d|^2} \sum_{\delta|d} \delta\mu(d/\delta) \frac{|\mathcal{K}_d|^2}{|\mathcal{K}_\delta|}$$

by the Jonhsen-Gallagher condition. □

We consider the following function φ_q:

$$\varphi_q(n) = \begin{cases} \psi_q^*(n) & \text{when } M+1 \le n \le M+N \\ 0 & \text{else,} \end{cases} \tag{10.8}$$

for q squarefree and of size at most Q.

The Hilbert space we consider is simply the space of functions over the integers in the interval $[M+1, M+N]$ equipped with the standard hermitiam product:

$$[g|h] = \sum_{M+1 \leq n \leq M+N} f(n)\overline{g(n)}. \tag{10.9}$$

Our last Lemma gives a measure of the "almost orthogonality" of the family $(\varphi_q)_{q \in \mathcal{Q}}$, where $\mathcal{Q}$ is given by (5.1).

Lemma 10.2

$$\sum_{q' \in \mathcal{Q}} |[\varphi_q|\varphi_{q'}]| \leq M_q = \phi(q)(N + Q^4).$$

Proof We have

$$\begin{aligned}[\varphi_q|\varphi_{q'}] &= \sum_{M+1 \leq n \leq M+N} \psi_q^*(n)\overline{\psi_{q'}^*(n)} \\ &= \sum_{M+1 \leq n \leq M+N} \sum_{a \bmod {}^* q} \hat{\psi}^*(a/q)e(an/q) \sum_{a' \bmod {}^* q'} \overline{\hat{\psi}^*(a'/q')e(-a'n/q')}.\end{aligned}$$

By summing first over a and a', we get

$$[\varphi_q|\varphi_{q'}] = \sum_{\substack{1 \leq a \leq q, \\ (a,q)=1}} \sum_{\substack{1 \leq a' \leq q', \\ (a',q')=1}} \hat{\psi}^*(a/q)\overline{\hat{\psi}^*(a'/q')} \sum_{M+1 \leq n \leq M+N} \mathrm{e}\left(n\left(\frac{a}{q} - \frac{a'}{q'}\right)\right).$$

The inner summation is in fact the sum of a geometric progression. When $a/q \neq a'/q'$, it is at most, in modulus,

$$1\Big/\left|\sin\left(\pi\left(\frac{a}{q} - \frac{a'}{q'}\right)\right)\right| \leq qq'/2$$

by using the classical inequality $\sin x \geq 2x/\pi$ when $0 \leq x \leq \pi/2$. □

On taking

10.2. Modifying the Selberg coefficients

The expression (10.5) above calls immediately for a modification, namely

$$\sum_{d \leq Q} \zeta_d \psi_d^* \tag{10.10}$$

for some arbitrary coefficients $(\zeta_d)_{d \leq Q}$. We readily find that

$$\begin{aligned}\sum_{d \leq Q} \zeta_d \psi_d^*(n) &= \sum_{d \leq Q} \zeta_d \sum_{\substack{q|d, \\ n \in \mathcal{K}_q}} \mu(d/q)\frac{q}{|\mathcal{K}_q|} \\ &= \sum_{q/n \in \mathcal{K}_q} \frac{q}{|\mathcal{K}_q|} \sum_{q|d \leq Q} \zeta_d \mu(d/q)\end{aligned}$$

so that we take

$$\tilde{\lambda}_q^\sharp = \frac{q}{|\mathcal{K}_q|}\sum_{q|d\le Q}\zeta_d\mu(d/q)/Y \tag{10.11}$$

where Y is the size parameter defined in (9.7). This choice is not very relevant here, since our quantities will all be homogenous in Y. Note that, when d^* is square-free and co-prime with δ,

$$Y\sum_{\substack{\delta|d,\\(d,d^*)=1}}|\mathcal{K}_d|\tilde{\lambda}_d^\sharp/d = \sum_{\substack{\delta|d,\\(d,d^*)=1}}\sum_{d|q\le Q}\zeta_q\mu(q/d) = \sum_{\delta|q}\zeta_q\sum_{\substack{\delta|d|q,\\(d,d^*)=1}}\mu(q/d).$$

Let us write $q=\delta\ell$. It is obvious that every prime factor of ℓ divides d^*, for otherwise the relevant contribution vanishes. Furthermore

$$\sum_{\substack{\delta|d|\delta\ell,\\(d,d^*)=1}}\mu(\delta\ell/d) = \mu(\ell)$$

(since only $d=\delta$ appears in this sum) which does not vanish only when $\mu^2(\ell)\neq 0$, which here implies that $\ell|d^*$. We define

$$G_d(Q) = \sum_{\substack{f\le Q,\\ [f,d]\le Q}} h(f)\zeta_{[d,f]}. \tag{10.12}$$

This leads to

$$\tilde{\lambda}_d = (-1)^{\omega(d)}G_d(Q)/Y \tag{10.13}$$

As already noted, the usual normalisation $Y=G_1(Q)$ is not required here. The next Lemma in a generalisation of Lemma 1.13.

Lemma 10.3 *When $\zeta\ge 0$ decreases on chains of multiples, the inequalities $G_\ell(Q\ell/d)\le G_d(Q)\le G_\ell(Q)$ hold whenever $\ell|d$.*

This has the nice consequence that $|\lambda_d|\le G_1(Q)/Y\ll 1$, while $\lambda_d^\sharp$ can be much bigger.

Proof The proof is a copy of the one of Lemma 6.2. The condition $[f,d]\le Q$ implies that $[f,\ell]\le Q$, which proves the first claim (notice that $h\ge 0$). In the other direction, let f be such that $[f,\ell]\le Q$. We have $[f,d]\le[f,\ell](d/\ell)$ and the Lemma follows readily. □

10.3. Resuming the main proof

With such a choice of the $\tilde{\lambda}_d$'s, $S_0((a_{d^*})_{d^*};\mathcal{K},\mathcal{K}^*)$ becomes

$$\boxed{Y^2S_0((a_{d^*})_{d^*};\mathcal{K},\mathcal{K}^*) = \sum_{\substack{d^*,\delta,\\(d^*,\delta)=1}}\gamma(d^*)a_{d^*}h(\delta)\Biggl(\sum_{\substack{\ell\le Q/\delta,\\ \ell|d^*}}\mu(\ell)\zeta_{\delta\ell}\Biggr)^2.} \tag{10.14}$$

To compare with earlier work, our family of parameters (ζ_d) has this name in [**6**] and [**79**]. Salerno in [**79**] chooses for ζ_d a step function with only two steps. Greaves in [**34**, Section 7.3.2, (2.4)] calls this parameter $y(d)$ and chooses a logarithmic smoothing of $\mathbb{1}_{d\le Q}$. Selberg in [**84**, (7.6), (7.9), (7.11)] uses weights that are similar to (10.15). The paper [**31**, (3.14)] uses also a similar shape though the sum they study is somewhat different. Heath-Brown in [**43**, (4)] determines his $\tilde{\lambda}_d$ directly, but the sum he studies differs notably from ours.

We shall further restrict our attention to weights of shape

$$\zeta_d = w\Big(\frac{\operatorname{Log} d}{\operatorname{Log} Q}\Big) = w(\alpha \operatorname{Log} d) \tag{10.15}$$

for some non-negative non-increasing function w on $(0,1]$. We quote explicitly:

$$\alpha = 1/\operatorname{Log} Q. \tag{10.16}$$

We further assume that w is continuous with $w(1) = 0$, prolonged to $(0,\infty)$ by setting $w(t) = 0$ when $t \geq 1$, piecewise differentiable and such that w' is bounded. These hypothesis ensure that $w(t) = -\int_t^1 w'(u)du$ which is what we need. We thus have

$$\zeta_d = -\int_{\alpha \operatorname{Log} d}^{1} w'(u)du. \tag{10.17}$$

CHAPTER 11

Two reduction steps

We continue in this chapter to develop the proof in a general context. We first restrict d^* to integers without any small prime factors. This step may seem harmless and usual, but is in fact crucial; it will rid us of many constant terms in asymptotic expressions and will enable us to disregard most of the coprimality conditions. This introduces a parameter P_0 which should disappear from the main term. This will be most easily treated when we will restrict the sieve coefficient to prime divisors, but it is a difficult step in general.

To restrict d^*, we limit our investigation to integers with fairly few divisors, as quantified by (H_5) below. Our second step here will be to remove the coprimality condition $(d^*, \delta) = 1$ from (10.14).

Lemma 11.1 *Let $d^* > 1$ be an integer and p be its smallest prime factor. We have, with the choice given by* (10.17)*,*

$$\left| \sum_{\substack{\ell' \leq Q/\delta, \\ \ell' | d^*}} \mu(\ell') \zeta_{\delta \ell'} \right| \leq \tau(d^*) \|w'\|_\infty \alpha \operatorname{Log} p.$$

Proof Indeed, it is enough to consider the case when d^* is square-free. Let us set $d^* = p d_0^*$. We can dispense with the condition $\ell' \leq Q/\delta$ since it is included in w (for $w(1) = 0$). We thus find that

$$\sum_{\substack{\ell' \leq Q/\delta, \\ \ell' | d^*}} \mu(\ell') \zeta_{\delta \ell'} = \sum_{\ell' | d_0^*} \mu(\ell') \big(\zeta_{\delta \ell'} - \zeta_{\delta p \ell'} \big)$$

which gets majorized as announced. □

We assume that

$$(H_5) \qquad \sum_{d^*} \tau(d^*)^2 \gamma(d^*) |a_{d^*}| = o\big(\operatorname{Log} Q\big).$$

In case (a_{d^*}) is simply the characteristic function of the primes $\leq D^*$, the upper bound is $\mathcal{O}(\operatorname{Log} \operatorname{Log} D^*)$.

Removing the small prime factors of d^*. Getting rid of the small prime factors of d^* will simplify the computation of the main term, essentially by removing constant terms. Let P_0 be a parameter to be chosen later. We set

$$\mathfrak{f}_0 = \prod_{p \le P_0} p. \tag{11.1}$$

We find that, on invoking (H_5),

$$\begin{aligned}\Bigg| \sum_{\substack{\delta,\\ (d^*,\mathfrak{f}_0)\neq 1}} \gamma(d^*)h(\delta)a_{d^*}\Bigg(\sum_{\substack{\ell'\le Q/\delta,\\ \ell'|d^*}}\mu(\ell')\zeta_{\delta\ell'}\Bigg)^2\Bigg| \\ \le \alpha^2\|w'\|_\infty^2 Y(\operatorname{Log} P_0)^2\sum_{d^*}\tau(d^*)^2\gamma(d^*)|a_{d^*}| \ll \|w'\|_\infty^2\alpha Y(\operatorname{Log} P_0)^2.\end{aligned}$$

This is more than enough. It is also $\mathcal{O}(\alpha^2\|w'\|_\infty^2 Y(\operatorname{Log}\operatorname{Log} P)(\operatorname{Log} P_0)^2)$ when (a_{d^*}) is simply the characteristic function of the primes $\le D^*$.

Removing the coprimality condition. We now remove the condition $(d^*, \delta) = 1$ in (10.14). Indeed on using (H_4), we find that

$$\begin{aligned}\Bigg| \sum_{\substack{d^*,\delta,\\ (d^*,\mathfrak{f}_0)=1,\\ (d^*,\delta)\neq 1}} \gamma(d^*)h(\delta)a_{d^*}\Bigg(\sum_{\substack{\ell'\le Q/\delta,\\ (d^*,\mathfrak{f}_0)=1,\\ \ell'|d^*}}\mu(\ell')\zeta_{\delta\ell'}\Bigg)^2\Bigg| \\ \le \alpha^2\|w'\|_\infty^2 \sum_{P_0<p\le P}\operatorname{Log}^2 p \sum_{\substack{d^*,\delta,\\ (d^*,\mathfrak{f}_0)=1,\\ p|d^*,p|\delta}}\gamma(d^*)h(\delta)|a_{d^*}|\tau(d^*)^2 \\ \ll \alpha^2\|w'\|_\infty^2 \sum_{P_0<p\le P}\frac{\operatorname{Log}^2 p}{p} \sum_{\substack{d^*,\delta,\\ (d^*,\mathfrak{f}_0)=1,\\ p|d^*,p|\delta}}\gamma(d^*)h(\delta/p)|a_{d^*}|\tau(d^*)^2 \\ \ll \alpha^2\|w'\|_\infty^2 Y \sum_{\substack{d^*\\ (d^*,\mathfrak{f}_0)=1}}\gamma(d^*)|a_{d^*}|\tau(d^*)^2\sum_{p|d^*}\frac{\operatorname{Log}^2 p}{p}.\end{aligned}$$

Note that d^* has at most $(\operatorname{Log} D^*)/\operatorname{Log} P_0$ prime factors and that we have assumed that $\alpha\operatorname{Log} D^* \ll 1$. As a consequence, the bound above is

$$\ll \|w'\|_\infty^2 Y^2\alpha(\operatorname{Log} P_0)/P_0. \tag{11.2}$$

We set

(11.3)

$$\boxed{Y^2 S_0^{(1)}((a_{d^*})_{d^*\le D^*}) = \sum_{\substack{d^*,\delta,\\ (d^*,\mathfrak{f}_0)=1}} \gamma(d^*)h(\delta)a_{d^*}\biggl(\sum_{\substack{\ell'\le Q/\delta,\\ \ell'|d^*}} \mu(\ell')\zeta_{\delta\ell'}\biggr)^2}$$

and we can replace $S_0((a_{d^*})_{d^*\le D^*})$ by $S_0^{(1)}((a_{d^*})_{d^*\le D^*})$ up to an error term of size at most (up to a multiplicative constant)

$$\|w'\|_\infty^2 Y\alpha\bigl((\operatorname{Log} P_0)^2 + (\operatorname{Log} P_0)/P_0\bigr) \ll \|w'\|_\infty^2 Y\alpha(\operatorname{Log} P_0)^2. \tag{11.4}$$

CHAPTER 12

Specialisation of the proof

In these lectures, we restrict our attention to the case when $\mathcal{K}$ is defined in (9.1), $\mathcal{K}^* = \mathcal{K}$ and d^* is 1 or a prime number. In such a case, we have

$$\gamma(1) = 1, \quad \gamma(p) = \kappa/p \tag{12.1}$$

except for a finite number of primes p for which $\mathcal{L}_p$ collapses somewhat. But since P_0 goes to infinity, we can discard such a case. We see from (11.3) that

$$Y^2 S_0^{(1)}(\mathbb{1}_{d^*=1}) = G(Q) \tag{12.2}$$

and that

$$Y^2 S_0^{(1)}((a_p)_{p\le D^*}) = \kappa \sum_{\substack{P_0<p\le D^*,\\ \delta\le Q}} \frac{h(\delta)}{p}\Big(w\big(\alpha \operatorname{Log}(p\delta)\big) - w\big(\alpha \operatorname{Log}\delta\big)\Big)^2$$

$$= \kappa \sum_{\substack{P_0<p\le D^*,\\ \delta\le Q}} \frac{h(\delta)}{p}\Big(w\big(\alpha \operatorname{Log}(p\delta)\big)^2 - 2w\big(\alpha \operatorname{Log}\delta\big)w\big(\alpha \operatorname{Log}(p\delta)\big) + w\big(\alpha \operatorname{Log}\delta\big)^2\Big)$$

$$= M_1 - 2M_2 + M_3$$

say. We assume that

$$Q \le D^* \tag{12.3}$$

for we would otherwise reach trivial results. Estimation of M_1, M_2 and M_3 relies on the following simple Lemma:

Lemma 12.1 *Let W be a C^1-function over $(0,\infty)$ that vanishes when its variable is larger than 1. We assume that $\|W'\|_1$ is finite. We have*

$$\sum_{\delta\le Q} h(\delta) W(\alpha \operatorname{Log}\delta)/Y = \kappa \int_0^1 u^{\kappa-1} W(u)du + \mathcal{O}(\alpha\|W'\|_1).$$

with $\alpha = 1/\operatorname{Log} Q$.

Proof Indeed, we use summation by parts to write

$$\sum_{\delta\le Q} h(\delta)W(\alpha\operatorname{Log}\delta) = -\sum_{\delta\le Q} h(\delta)\int_{\alpha\operatorname{Log}\delta}^{1} W'(u)du = \int_0^1 \sum_{\delta\le Q^u} h(\delta)\ W'(u)du$$
$$= A\int_0^1 \operatorname{Log}^\kappa(Q^u)W'(u)du + \mathcal{O}(\alpha Y\|W'\|_1)$$

on appealing to (9.6) and the Lemma follows readily. □

We use Lemma 12.1 with $W(u) = w^2(u+\alpha\operatorname{Log}p)$ for M_1, with $W(u) = w(u)w(u+\alpha\operatorname{Log}p)$ for M_2 and with $W(u) = w^2(u)$ for M_3. Their derivatives are bounded in terms in terms of w. Note that, in M_1 and M_2, the variable p can be bounded above by Q. We thus get

$$YS_0^{(1)}((a_p)_{p\le D^*})/\kappa^2 = \int_0^1 \sum_{P_0<p\le Q} \frac{(w(u)-w(u+\alpha\operatorname{Log}p))^2}{p} u^{\kappa-1}du$$
$$+\int_0^1 w(u)^2u^{\kappa-1}du\operatorname{Log}\frac{\operatorname{Log}D^*}{\operatorname{Log}Q} + \mathcal{O}(\alpha\operatorname{Log}\operatorname{Log}Q).$$

We deal with the sum over primes in the following Lemma.

Lemma 12.2 *Let F be a C^1-function over $[0,1]$ that vanishes at 0. We assume that $\|F'\|_\infty$ is finite. We have*

$$\sum_{P_0<p\le Q} \frac{F(\alpha\operatorname{Log}p)}{p} = \int_0^1 \frac{F(v)}{v}dv + \mathcal{O}(\|F'\|_\infty(\alpha\operatorname{Log}P_0 + 1/\operatorname{Log}P_0)).$$

Proof Indeed, on calling S the sum to be studied, partial summation yields

$$S = \sum_{P_0<p\le Q} \frac{1}{p}\int_0^{\alpha\operatorname{Log}p} F'(v)dv$$
$$= \int_\xi^1 \Big(-\operatorname{Log}v + \mathcal{O}(1/\operatorname{Log}P_0)\Big)F'(v)dv + \int_0^\xi \Big(-\operatorname{Log}\xi + \mathcal{O}(1/\operatorname{Log}P_0)\Big)F'(v)dv$$
$$= \int_\xi^1 \frac{F(v)}{v}dv + \mathcal{O}(1/\operatorname{Log}P_0) = \int_0^1 \frac{F(v)}{v}dv + \mathcal{O}(|F'(0)|\alpha\operatorname{Log}P_0 + \|F'\|_1/\operatorname{Log}P_0)$$

with $\xi = \alpha\operatorname{Log}P_0$. □

On using this Lemma, we reach

$$(12.4)\quad YS_0^{(1)}((a_p)_{p\le D^*})/\kappa^2 = \int_0^1\int_0^1 \frac{(w(u)-w(u+v))^2}{v}u^{\kappa-1}dudv$$
$$+\int_0^1 w(u)^2u^{\kappa-1}du\operatorname{Log}\frac{\operatorname{Log}D^*}{\operatorname{Log}Q} + \mathcal{O}(\alpha\operatorname{Log}(P_0\operatorname{Log}Q) + 1/\operatorname{Log}P_0).$$

12.1. Final estimate

We find that, on collecting Lemma 9.1, (11.2), (11.4) and (12.4), we get

$$\sum_{n\le N}\Big(b-\sum_{p/n\in\mathcal{L}_p}1\Big)\Big(\sum_{d/n\in\mathcal{K}_d}\tilde{\lambda}^{\sharp}_d\Big)^2\big/\big(NG(Q)\kappa^2\big) \tag{12.5}$$
$$=\Big(\frac{b}{\kappa}-\mathrm{Log}\,\frac{\mathrm{Log}\,D^*}{\mathrm{Log}\,Q}\Big)\int_0^1 w(u)^2u^{\kappa-1}du-\int_0^1\int_0^1\frac{(w(u)-w(u+v))^2}{v}u^{\kappa-1}dudv$$
$$+\mathcal{O}\bigg(\alpha\frac{\mathrm{Log}\,P_0}{P_0}+D^*Q^2N^{-1}+\alpha\,\mathrm{Log}^2\,P_0+\alpha\,\mathrm{Log}(P_0\,\mathrm{Log}\,Q)+1/\,\mathrm{Log}\,P_0\bigg).$$

We select

$$P_0=\exp\sqrt{\mathrm{Log}\,Q} \tag{12.6}$$

so that the error term above reduces to $\mathcal{O}\big((\mathrm{Log}\,\mathrm{Log}\,Q)/\sqrt{\mathrm{Log}\,Q}+D^*Q^2N^{-1}\big)$. It is high time to take some notation:

$$I_1(w,\kappa)=\int_0^1 w(u)^2u^{\kappa-1}du. \tag{12.7}$$

We set

$$\int_0^1\int_0^1\frac{(w(u)-w(u+v))^2}{v}u^{\kappa-1}dudv=K_1(w,\kappa)+I_2(w,\kappa)$$

where

$$K_1(w,\kappa)=-\int_0^1\mathrm{Log}(1-u)\,w(u)^2u^{\kappa-1}du, \tag{12.8}$$

and

$$I_2(w,\kappa)=\int_0^1\int_0^{1-u}\frac{(w(u)-w(u+v))^2}{v}u^{\kappa-1}dudv. \tag{12.9}$$

These denominations may look somewhat obscure, but there are taken in accordance to another sets of notes I have distributed on the subject.

12.2. Conclusion

We select some $\theta\ge 1$ and some $\epsilon\in(0,1/2]$. We choose $D^*=Q^\theta$ and $Q=N^{\frac{1-\epsilon}{2+\theta}}$. When

$$\frac{b}{\kappa}>\mathrm{Log}\,\theta+\frac{K_1(w,\kappa)+I_2(w,\kappa)}{I_1(w,\kappa)} \tag{12.10}$$

then there are infinitely many integers n such that

$$b-\sum_{\substack{p|(n+h_1)(n+h_2)\dots(n+h_\kappa),\\ p\le D^*}}1>0. \tag{12.11}$$

This means that $(n+h_1)(n+h_2)\dots(n+h_\kappa)$ has at most $[b]$ prime factors less than D^*. Each $n+h_i$ can have at most $[(1+2\theta^{-1})/(1-\epsilon)]$ prime factors strictly greater than D^*, so the total number of prime factors is

$$\kappa[(1+2\theta^{-1})/(1-\epsilon)]+\biggl[\kappa\operatorname{Log}\theta+\kappa\frac{K_1(w,\kappa)+I_2(w,\kappa)}{I_1(w,\kappa)}\biggr]. \tag{12.12}$$

We want θ to be as close to 1 as possible.

Extending the class of smoothings w. Our main problem is thus to minimize of

$$\frac{K_1(w,\kappa)+I_2(w,\kappa)}{I_1(w,\kappa)}. \tag{12.13}$$

This minimisation should take place over functions w that are C^1 over $[0,\infty)$, and vanish from 1 onwards. An approximation argument readily extends the class in such a way that the continuity in 1 is not required anymore. In fact we can also dispense with the C^1 condition: a continuous function that is C^1 per interval is enough, since (10.17) is the writing we need. We have also used the fact that w was of bounded variations. We further see that, on using an approximation argument, we can also dispense with the condition that w be continuous in 1. We will not need it, but we can also show that it is enough to assume that w is C^1 per interval, without being continuous. A step function would perfectly do.

12.3. A first choice. Proof of Theorem 8.1

We select w to be function equal to 1 over $[0,1]$ and 0 otherwise. The quantity $I_2(w,\kappa)$ vanishes, and $K_1(w,\kappa)$ is computed in the following Lemma:

Lemma 12.3 *We have*

$$-\int_0^1 \operatorname{Log}(1-u)u^{\kappa-1}du=\frac{1}{\kappa}\Bigl(1+\frac12+\cdots+\frac{1}{\kappa}\Bigr).$$

Proof We use the development

$$-\operatorname{Log}(1-u)=\sum_{k\ge1}u^k/k$$

valid for $|u|<1$ and use the Lebesgue dominated convergence Theorem to infer that

$$\begin{aligned}-\int_0^1 \operatorname{Log}(1-u)u^{\kappa-1}du&=\sum_{k\ge1}\int_0^1 u^{k+\kappa-1}du/k\\&=\sum_{k\ge1}\frac{1}{k(\kappa+k)}=\frac{1}{\kappa}\sum_{k\ge1}\Bigl(\frac1k-\frac{1}{k+\kappa}\Bigr)\end{aligned}$$

and the last summation amounts to a finite sum, due to a telescoping effect. □

We select $\theta = 1$, $\epsilon = 0.1$ and use the classical bound

$$1 + \frac{1}{2} + \cdots + \frac{1}{\kappa} \leq \operatorname{Log} \kappa + 1.$$

CHAPTER 13

Special computations for bounded values

Since we are not able to get an explicit solution, if it exists, to the optimization problem arising from (12.12), we do some direct optimization for small values of κ. We look only at continuous and locally affine functions. Note that they are Lipschitz so the definition of I_2 does not present any problem. Note furthermore that the functions we shall choose at the end are non-increasing, indeed ensuring that $|\tilde{\lambda}_d| \le 1$.

These choices are the one that enabled us to build tablein the introduction. Note that in what follows we consider chunks of two functions, one with parameters a, α, b, β and a second one with parameters a', α', b', β'. We will shorten the first by calling it w and the latter by calling it w', *which does not have anything to do with the derivative!*

13.1. Decomposition in affine pieces

We select functions

$$w_{a,\alpha,b,\beta} = \begin{cases} 0 & \text{when } t \notin [a,b] \\ \frac{(\beta-\alpha)t+\alpha b-\beta a}{b-a} = \gamma t + \eta & \text{when } t \in [a,b]. \end{cases} \tag{13.1}$$

from which we build

$$w = \sum_{1\le i\le I} w_{a_i,\alpha_i,b_i,\beta_i} \tag{13.2}$$

with $0 = a_1 < b_1 = a_2 < b_2 = a_3 < \cdots < b_I = 1$ and $\beta_i = \alpha_{i+1}$. The three functionnals I_1, K_1 and I_2 are quadratic forms. The first two are fo diagonals forms, so we have

$$I_1(w,\kappa) = \sum_{1\le i\le I} I_1(w_{a_i,\alpha_i,b_i,\beta_i},\kappa)$$

and similarly for $K_1(w,\kappa)$. The problem is more complicated for I_2, and we get

$$I_2(w,\kappa) = \sum_{1\le i,j\le I} I_2(w_{a_i,\alpha_i,b_i,\beta_i}, w_{a_j,\alpha_j,b_j,\beta_j},\kappa)$$

where we have kept the notation I_2 for the induced hermitian product defined by

$$I_2(w,w',\kappa)=\int_0^1\int_0^{1-u}\frac{(w(u)-w(u+v))\overline{(w'(u)-w'(u+v))}}{v}u^{\kappa-1}dudv.$$

13.2. Formulae to evaluate of I_1 and K_1

We first see that

$$(13.3)\quad I_1(w_{a,\alpha,b,\beta},\kappa)=\gamma^2\frac{b^{\kappa+2}-a^{\kappa+2}}{\kappa+2}+2\gamma\eta\frac{b^{\kappa+1}-a^{\kappa+1}}{\kappa+1}+\eta^2\frac{b^\kappa-a^\kappa}{\kappa}.$$

Then we set

$$(13.4)\qquad \Sigma(x,\kappa)=\sum_{\ell\geq1}\frac{x^{\kappa+\ell}}{\ell(\ell+\kappa)}.$$

We readily find that

$$K_1(w_{a,\alpha,b,\beta},\kappa)=\gamma^2(\Sigma(b,\kappa{+}2)-\Sigma(a,\kappa{+}2))+2\gamma\eta(\Sigma(b,\kappa{+}1)-\Sigma(a,\kappa{+}1))+\eta^2(\Sigma(b,\kappa)-\Sigma(a,\kappa)).$$

13.3. Formulae to evaluate of I_2

Computing I_2 leads to much more bulky formulaes. We first assume that $0\leq a\leq b\leq a'\leq b'\leq1$. We have, by polarization,

$$\begin{aligned}I_2(w_{a,\alpha,b,\beta},w_{a',\alpha',b',\beta'},\kappa)&=\int_a^{b'}\int_t^{b'}\frac{(w'(u)-w'(t))(w(u)-w(t))}{u-t}dut^{\kappa-1}dt\\&=\int_a^b\int_{a'}^{b'}\frac{(w'(u)-w'(t))(w(u)-w(t))}{u-t}dut^{\kappa-1}dt\\&\quad+\int_b^{b'}\int_t^{b'}\frac{(w'(u)-w'(t))(w(u)-w(t))}{u-t}dut^{\kappa-1}dt.\end{aligned}$$

On examining the values of the functions w and w' in the intervals of integration, this expression simplifies into:

$$\begin{aligned}I_2(w_{a,\alpha,b,\beta},w_{a',\alpha',b',\beta'},\kappa)&=-\int_a^b\int_{a'}^{b'}\frac{w'(u)w(t)}{u-t}dut^{\kappa-1}dt\\&=-\int_a^b\left(\gamma'(b'-a')+(\gamma't+\eta')\operatorname{Log}\frac{b'-t}{a'-t}\right)w(t)t^{\kappa-1}dt\\&=-\gamma'(b'-a')\left(\frac{\gamma(b^{\kappa+1}-a^{\kappa+1})}{\kappa+1}+\frac{\eta(b^\kappa-a^\kappa)}{\kappa}\right)\\&-\gamma'\gamma\mathfrak{p}(\kappa+2,a,b,a'b')-(\gamma\eta'+\eta\gamma')\mathfrak{p}(\kappa,a,b,a'b')-\eta'\eta\mathfrak{p}(\kappa,a,b,a'b')\end{aligned}$$

with

$$(13.5)\quad \mathfrak{p}(\kappa, a, b, a'b') = \int_a^b t^{\kappa-1} \operatorname{Log} \frac{b'-t}{a'-t} dt$$

$$= \frac{b^\kappa - a^\kappa}{\kappa} \operatorname{Log} \frac{b'}{a'} - b'^\kappa \left(\Sigma\Big(\kappa, \frac{b}{b'}\Big) - \Sigma\Big(\kappa, \frac{a}{b'}\Big)\right) + a'^\kappa \left(\Sigma\Big(\kappa, \frac{b}{a'}\Big) - \Sigma\Big(\kappa, \frac{a}{a'}\Big)\right).$$

We finally have to handle the case when $w_{a,\alpha,b,\beta} = w_{a',\alpha',b',\beta'}$. In that case

$$I_2(w_{a,\alpha,b,\beta}, w_{a,\alpha,b,\beta}, \kappa) = \int_0^a \int_a^b \frac{w(u)^2}{u-t} du t^{\kappa-1} dt$$

$$+ \int_a^b \int_t^1 \frac{(w(u)-w(t))(w(u)-w(t))}{u-t} du t^{\kappa-1} dt + \int_b^1 \int_t^1 \frac{(w(u)-w(t))w(u)}{u-t} du t^{\kappa-1} dt$$

i.e. $I_2(w_{a,\alpha,b,\beta}, w_{a,\alpha,b,\beta}, \kappa)$ is equal to

$$W + \int_a^b \int_t^b \frac{(w(u)-w(t))(w(u)-w(t))}{u-t} du t^{\kappa-1} dt + \int_a^b \int_b^1 \frac{w(t)w(t)}{u-t} du t^{\kappa-1} dt$$

$$= W + \gamma^2 \int_a^b \frac{(b-t)^2}{2} t^{\kappa-1} dt + \int_a^b w(t)^2 t^{\kappa-1} \operatorname{Log} \frac{1-t}{b-t} dt$$

$$= W + \frac{\gamma^2}{2} \left(\frac{b^2(b^\kappa - a^\kappa)}{\kappa} - 2\frac{b(b^{\kappa+1} - a^{\kappa+1})}{\kappa+1} + \frac{b^{\kappa+2} - a^{\kappa+2}}{\kappa+2} \right)$$

$$+ \gamma^2 \mathfrak{p}(\kappa+2, a, b, b, 1) + 2\gamma\eta \mathfrak{p}(\kappa+1, a, b, b, 1) + \eta^2 \mathfrak{p}(\kappa, a, b, b, 1)$$

where W is defined by

$$W = \int_0^a \int_a^b \frac{w(u)^2}{u-t} du t^{\kappa-1} dt$$

$$= \int_0^a \left(\gamma^2 \frac{b^2-a^2}{2} + (\gamma^2 t + 2\gamma\eta)(b-a) + (\gamma t + eta)^2 \operatorname{Log} \frac{b-t}{a-t} \right) t^{\kappa-1} dt$$

$$= \left(\gamma^2 \frac{b^2-a^2}{2} \right) \frac{a^\kappa}{\kappa} + \gamma^2 (b-a) \frac{a^{\kappa+1}}{\kappa+1} + 2\gamma\eta(b-a)\frac{a^\kappa}{\kappa}$$

$$(13.6)\quad + \gamma^2 \mathfrak{p}(\kappa+2, 0, a, a, b) + 2\eta\gamma \mathfrak{p}(\kappa+1, 0, a, a, b) + \eta^2 \mathfrak{p}(\kappa, 0, a, a, b).$$

13.4. Results

We ran the Pari/GP-script described in next chapter to get the results below. The process has been to optimize the resulting quadratic form under the quadratic constraint $I_1(\kappa, w) = 1$; this is a classical problem which is solved by getting the eigenvectors associated with the some matrix. The script is detailled in next chapter, and the optimisation process in coded in section 14.3. We then found rational coefficients close enough to the optimal ones obtained and recomputed the resulting $n_0(\kappa)$.

Here are some more details for $\kappa = 2$ and $\kappa = 8$.

We reach $n_0(2) \leq 5$ by using $\theta^{-1} = 0.38$ and the simplest affine function w that takes the values $w(0) = 1$ and $w(1) = 61/500$.

We reach $n_0(8) \leq 28$ by selecting $\theta^{-1} = 0.49$

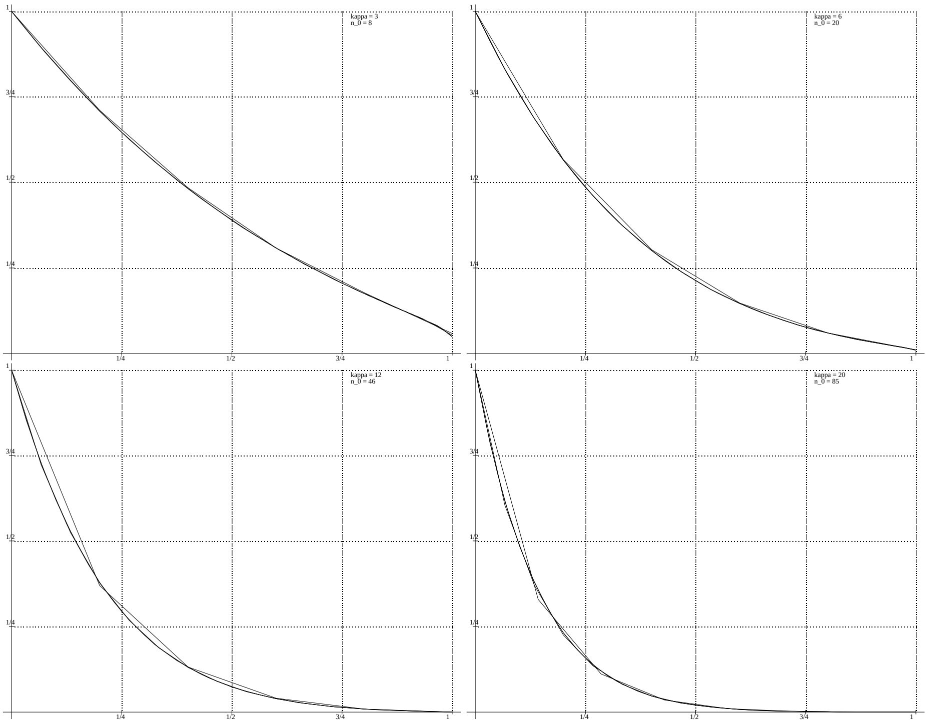
$$w(0) = 250, \quad w(1/2) = 40, \quad w(1) = 1.$$

We also plotted the different optimal solutions with a fixed κ and decreasing subdivisions. The optimal solution proposed by our script seems to converge to some convex decreasing function that nearly vanishes at $t = 1$. This last observation is not quite true when $\kappa = 3$.

κ	$1/\theta$	nb	values	(l)	$n_0(\kappa)$
1	0.28	1	[1000,272]	1	2
2	0.38	1	[1000,122]	1	5
3	0.48	1	[1000,68]	40	8
4	0.46	1	[1000,43]	30	12
5	0.46	1	[1000,30]	30	16
6	0.46	1	[1000,22]	40	20
7	0.48	1	[1000,16]	40	24
8	0.49	2	[250, 40, 1]	40	28
9	0.48	5	[100000, 45689, 19831, 7575, 2321, 226]	40	32
10	0.48	3	[10000, 2409, 437, 12]	40	37
11	0.48	5	[100000, 39706, 15070, 4978, 1275, 83]	40	41
12	0.48	5	[100000, 37037, 13157, 4051, 952, 51]	40	46
13	0.48	4	[100000, 26637, 6407, 1119, 28]	30	51
14	0.49	6	[1000000, 389243, 149816, 53062, 16236, 3824, 211]	30	55
15	0.48	8	[10000000, 4732626, 2252128, 1033725, 447270, 178197, 62553, 17622, 1525]	30	60
16	0.48	7	[10000000, 4034866, 1638226, 630749, 221735, 67486, 15914, 901]	30	65
17	0.49	5	[100000, 26256, 6739, 1475, 231, 4]	30	70
18	0.49	6	[1000000, 308342, 95709, 27254, 6586, 1169, 31]	30	75
19	0.49	6	[1000000, 290904, 85598, 23094, 5268, 873, 19]	30	80
20	0.49	7	[10000000, 3297711, 1114086, 357423, 103842, 25646, 4708, 139]	40	85
21	0.49	7	[10000000, 3134579, 1011485, 310121, 85960, 20173, 3486, 87]	30	90
22	0.49	8	[10000000, 3472269, 1247118, 434059, 141344, 41582, 10438, 1967, 64]	20	95

In this table, the column (l) indicates up to which level of precision in the decomposition in w we went: we tried to divide the unit interval in 1 piece, then in 2 equal pieces, and so on, till the number indicated. The column nb contains the number of equal subdivisions of the unit interval we have used, and for instance 5 tells us we have used the subdivision $[0, 1/5, 2/5, 3/5, 4/5, 1]$.

We reproduce below the plot of the optimal functions got by subdividing the unit interval in 5, then 15, then 30 and then 40 subintervals of equal length, when $\kappa = 3, 6, 12$ and $\kappa = 20$ (except that with started with 7 subdivisions in this last case and we increased the precision to 700 digits to compensate loss in precision):



These numerical datas introduce the following question.

Open Problem no III. *Show that the quantity $(K_1(w,3)+I_2(w,3))/I_1(w,3)$ reaches its minimum at some convex decreasing function w. Get an explicit expression for this minimum and, if possible, get an explicit expression of the optimal w_0, or at least, compute $w_0(1)$.*

Open Problem no IV.

Answer to the same questions for a general κ, or asymptotically in κ.

CHAPTER 14

The GP-script

We take this opportunity to present a full script written for the system Pari/GP, see [**88**]. This symbol denotes a collection of high level computation tools developped and maintained by mathematicians. The code is free to use and accessible to read, check and/or improve upon. It is furthermore well adapted to number theory. The script we present below is far from being optimized, our aim here being twofold: to introduce the reader to PARI/GP and to explain the algorithm we have employed. This script can be turned into a C-program and compiled by `gp2c`. The procedure to do so is described on a specific problem in [**3**].

This script is stored in file `GP/AffineHRIW.gp`.

14.1. Precomputing Σ

The parameter κ being fixed, we store values of Σ (given by (13.4)) that have been already computed. To do so, we handle two lists, `PrecomputedArgs` for the arguments and `PrecomputedVals` for the values taken. The integer `NbPrecomputedValUsed` is here for statistical reasons: we increase it by one each time we reuse an already computed value. The function that computes Σ is called ... `Sigmi` for the name `Sigma` is protected in Pari/GP.

```
global(PrecomputedArgs, PrecomputedVals, NbPrecomputedValUsed);
PrecomputedArgs = []; PrecomputedVals = [];
NbPrecomputedValUsed = 0;

{getvalue(uk)=
   local(ell = 1);
   while(ell <= length(PrecomputedArgs),
     if(uk == PrecomputedArgs[ell],
        NbPrecomputedValUsed++;
        return(PrecomputedVals[ell]),
        if(uk < PrecomputedArgs[ell], return(0), ell++)));
   /* No precomputed values, return 0: */
   return(0)}

{getindex(uk)= /* Sigmi(uk) has *not* been precomputed. */
```

```
   local(k = 1);
   while(k <= length(PrecomputedArgs),
     if(uk < PrecomputedArgs[k], return(k), k++));
   return(k)}

{vecinsert(vec, what, where)=
   forstep(k =  length(vec), where+1, -1, vec[k] = vec[k-1]);
   vec[where] = what; return(vec)}

{Sigmi(u, kappa)=
   local(res, where);

   if(u == 0, return(0),
      res = getvalue(u + kappa);
      if(res != 0, return(res),);
      if(u == 1,
        for(ell = 1, kappa, res += 1/ell);
        res = res/kappa,
        /* else: */
        res = sumpos(ell = 1, u^ell/ell/(ell+kappa))*u^kappa);
     where = getindex(u + kappa);
     if(where <= length(PrecomputedArgs),
        PrecomputedArgs = concat(PrecomputedArgs, [0]);
        PrecomputedArgs = vecinsert(PrecomputedArgs, u + kappa, where);
        PrecomputedVals = concat(PrecomputedVals, [0]);
        PrecomputedVals = vecinsert(PrecomputedVals, res, where),
        /* else: */
        PrecomputedArgs = concat(PrecomputedArgs, [u+kappa]);
        PrecomputedVals = concat(PrecomputedVals, [res]));
      return(res))}
```

14.2. Computing I_1, I_2 and K_1

We start the actual computations. The function `Partial` computes $\mathfrak{p}$ given by (13.5), while the function `getW` computes W given by (13.6). The final functions of this part are `getvectorI1`, `getvectorK1` and `getvectorI2` which take κ as argument as well as the subdivision of the unit interval we consider. Note that we use the indeterminates `X`, `Y`, `Xp` and `Yp` in the `getvectorI1|K1|I2` family of functions,

```
{getbaseI1(kappa, a, b, gammaa, etaa, ap, bp, gammaap, etaap)=
   if(b <= ap, return(0),
      return(gammaa^2*(b^(kappa+2)-a^(kappa+2))/(kappa+2)
         +2*gammaa*etaa*(b^(kappa+1)-a^(kappa+1))/(kappa+1)
         +etaa^2*(b^kappa-a^kappa)/kappa));}
```

```
{getI1(kappa, a, b, gammaa, etaa, ap, bp, gammaap, etaap)=
   if(a < ap,
      getbaseI1(kappa, a, b, gammaa, etaa, ap, bp, gammaap, etaap),
      getbaseI1(kappa, ap, bp, gammaap, etaap, a, b, gammaa, etaa))}

{Partial(kappa, a, b, ap, bp)=
   if((u == 0)||(a == b)||(ap == bp), return(0),
      return(log(bp/ap)*(b^kappa-a^kappa)/kappa
        -bp^kappa*(Sigmi(b/bp, kappa)-Sigmi(a/bp, kappa))
        +ap^kappa*(Sigmi(b/ap, kappa)-Sigmi(a/ap, kappa))))}

{getbaseK1(kappa, a, b, gammaa, etaa, ap, bp, gammaap, etaap)=
   if(b <= ap, return(0),
      return(gammaa^2*(Sigmi(b, kappa+2) - Sigmi(a, kappa+2))
        +2*gammaa*etaa*(Sigmi(b, kappa+1) - Sigmi(a, kappa+1))
        +etaa^2*(Sigmi(b, kappa) - Sigmi(a, kappa))));}

{getK1(kappa, a, b, gammaa, etaa, ap, bp, gammaap, etaap)=
   if(a < ap,
      getbaseK1(kappa, a, b, gammaa, etaa, ap, bp, gammaap, etaap),
      getbaseK1(kappa, ap, bp, gammaap, etaap, a, b, gammaa, etaa))}

{getW(kappa, a,b, gammaa, etaa)=
   return((gammaa^2*(b^2-a^2)/2+2*gammaa*etaa*(b-a))*a^kappa/kappa
     +gammaa^2*(b-a)*a^(kappa+1)/(kappa+1)
     +gammaa^2*Partial(kappa+2, 0, a, a, b)
     +2*gammaa*etaa*Partial(kappa+1, 0, a, a, b)
     +etaa^2*Partial(kappa, 0, a, a, b))}

{getbaseI2(kappa, a, b, gammaa, etaa, ap, bp, gammaap, etaap)=
   if(b <= ap,
      if(b == 0, return(0),
         return(-gammaap*(bp-ap)*(gammaa*(b^(kappa+1)-a^(kappa+1))/(kappa+1)
                 +etaa*(b^(kappa)-a^(kappa))/(kappa))
           -gammaap*gammaa*Partial(kappa+2, a, b, ap, bp)
           -(gammaap*etaa+etaap*gammaa)*Partial(kappa+1, a, b, ap, bp)
           -etaap*etaa*Partial(kappa, a, b, ap, bp))),
      return( gammaa^2/2*(b^2*(b^(kappa)-a^(kappa))/(kappa)
          -2*b*(b^(kappa+1)-a^(kappa+1))/(kappa+1)
          +(b^(kappa+2)-a^(kappa+2))/(kappa+2))
          +gammaa^2*Partial(kappa+2, a, b, b, 1)
          +2*gammaa*etaa*Partial(kappa+1, a, b, b, 1)
          +etaa^2*Partial(kappa, a, b, b, 1)
          +getW(kappa, a, b, gammaa, etaa)))}

{getI2(kappa, a, b, gammaa, etaa, ap, bp, gammaap, etaap)=
   if(a < ap,
      getbaseI2(kappa, a, b, gammaa, etaa, ap, bp, gammaap, etaap),
      getbaseI2(kappa, ap, bp, gammaap, etaap, a, b, gammaa, etaa))}

{getgammaaetaa(a, b, alpha, beta) =
   return([(beta-alpha)/(b-a), (b*alpha-a*beta)/(b-a)])}

{getvectorI1(kappa, endpoints)=
   local( taille, var, a, b, mymat , aux, adder, auxbis);
   taille = length(endpoints);
   mymat = matrix( taille, taille);
   for(k = 1, taille-1,
      a = endpoints[k]; b = endpoints[k+1];
      var = getgammaaetaa(a, b, X, Y);
      aux = getI1(kappa, a, b, var[1], var[2], a, b, var[1], var[2]);
      add = subst(subst(aux, X, 1), Y, 0);
      mymat[k, k] += add;
      auxbis = subst(subst(aux, X, 0), Y, 1);
      mymat[k+1, k+1] += auxbis;
      add += auxbis;
      auxbis = (subst(subst(aux, X, 1), Y, 1) - add)/2;
      mymat[k, k+1] += auxbis; mymat[k+1, k] += auxbis);
   print("I1(", kappa, ") is ready.");
   return(mymat)}

{getvectorK1(kappa, endpoints)=
   local( taille, var, a, b, mymat, aux, adder, auxbis);
   taille = length(endpoints);
   mymat = matrix( taille, taille);
   for(k = 1, taille-1,
      a = endpoints[k]; b = endpoints[k+1];
      var = getgammaaetaa(a, b, X, Y);
      aux = getK1(kappa, a, b, var[1], var[2], a, b, var[1], var[2]);
      add = subst(subst(aux, X, 1), Y, 0);
      mymat[k, k] += add;
      auxbis = subst(subst(aux, X, 0), Y, 1);
      mymat[k+1, k+1] += auxbis;
      add += auxbis;
      auxbis = (subst(subst(aux, X, 1), Y, 1) - add)/2;
      mymat[k, k+1] += auxbis; mymat[k+1, k] += auxbis);
   print("K1 is ready.");
   return(mymat)}

{getvectorI2(kappa, endpoints)=
   local( taille, var, a, b, ap, bp, varp, mymat, aux, adder, auxbis);
   taille = length(endpoints);
   mymat = matrix( taille, taille);
   for(k = 1, taille-1,
     a = endpoints[k]; b = endpoints[k+1];
     var = getgammaaetaa(a, b, X, Y);
     for(l = k, taille-1,
     if(l > k,
        ap = endpoints[l]; bp = endpoints[l+1];
        varp = getgammaaetaa(ap, bp, Xp, Yp);

        aux = getI2(kappa, a, b, var[1], var[2], ap, bp, varp[1], varp[2]);
        add = subst(subst(subst(subst(aux, X, 1), Y, 0), Xp, 1), Yp, 0);
        mymat[k, l] += add; mymat[l, k] += add;
        auxbis = subst(subst(subst(subst(aux, X, 0), Y, 1), Xp, 0), Yp, 1);
        add += auxbis;
        mymat[k+1, l+1] += auxbis; mymat[l+1, k+1] += auxbis;
        auxbis = subst(subst(subst(subst(aux, X, 1), Y, 0), Xp, 0), Yp, 1);
        add += auxbis;
        mymat[k, l+1] += auxbis; mymat[l+1, k] += auxbis;
        auxbis = subst(subst(subst(subst(aux, X, 0), Y, 1), Xp, 1), Yp, 0);
        add += auxbis;
        mymat[k+1, l] += auxbis; mymat[l, k+1] += auxbis,
        /* else: */
        aux = getI2(kappa, a, b, var[1], var[2], a, b, var[1], var[2]);
        add = subst(subst(aux, X, 1), Y, 0);
        mymat[k, k] += add;
        auxbis = subst(subst(aux, X, 0), Y, 1);
        add += auxbis;
        mymat[k+1, k+1] += auxbis;
        auxbis = (subst(subst(aux, X, 1), Y, 1) - add)/2;
        mymat[k, l+1] += auxbis; mymat[l+1, k] += auxbis;
      )));
   print("I2 is ready.");
   return(mymat)}
```

14.3. Optimizing

This part is dedicated to the optimization process. The optimal slopes for the step function are computed in `OptimizeK1andI2surI1`. We then range over a selected set of values of $\tau = 1/\theta$ and get the best one in

`Optimizedfunction`. The reader may then use `Optimizedfunction(4, regularpoints(10))` to get the result for a subdivision in 10 subintervals of equal length and $\kappa = 4$. In order to check the computations, we only take a rational approximation of the best slopes, and this is the work done in `abetterversionof`.

```
{getlowerbounddd0(kappa, val, tau)=
   kappa*floor(2*tau+1)+ floor(kappa*(-log(tau)+ val));}

{abetterversionof(thekernel)=
   local(thevector, maxelement = 0, minelement = 0, multiplier, finalden = 1);

   thevector = vector(length(thekernel~));
   multiplier = 10^(ceil(length(thekernel~)*2/3)+1);
   for(k = 1, length(thekernel~),
      thevector[k] = thekernel[k];
      maxelement = max(thevector[k], maxelement);
      minelement = min(thevector[k], minelement));
   if(maxelement > -minelement,
      coef = maxelement, coef = minelement);
   for(k = 1, length(thekernel~),
      thevector[k] = round(multiplier*thevector[k]/maxelement)/multiplier;
      finalden = lcm(finalden, denominator(thevector[k]));
   for(k = 1, length(thekernel~), thevector[k] *= finalden);
   return(thevector)}

{localmateigen(mymat) =
   local(res, realprecision_ini);
   realprecision_ini = default(realprecision,,1);
   default(realprecision, realprecision_ini-5);
   res = mateigen(mymat);
   default(realprecision, realprecision_ini);
   return(res)}

{OptimizeK1andI2surI1(kappa, tau, I1K1I2vectordata, silent = 1)=
   local(myeigenvectors, myvaleigenvectors, mymat1, mymat2, mymat,
         goodindex, value, bestvalue = 0, taille, thevector, res, aux);
   taille = length(I1K1I2vectordata[1]);

   mymat1 = I1K1I2vectordata[1];
   mymat2 = I1K1I2vectordata[2] + I1K1I2vectordata[3];
   mymat = mymat1^(-1)*mymat2;

   myeigenvectors = localmateigen(mymat);
   for(k = 1, taille,
      myvaleigenvectors = mymat*(myeigenvectors[,k]);
      value = 0;
      for(l = 1, taille,
         if((value == 0)&&(myeigenvectors[l,k] != 0),
            value = myvaleigenvectors[l]/myeigenvectors[l,k],
            if((value != 0)&&(myeigenvectors[l,k] != 0),
               aux = myvaleigenvectors[l]/myeigenvectors[l,k];
               if(abs(1-value/aux) > 0.00000001,
                  print("!!! Bad precision loss!!",)),)));
      if((bestvalue == 0)||(value < bestvalue),
         bestvalue = value; goodindex = k,););
   thevector = abetterversionof(myeigenvectors[,goodindex]);

   res = thevector*mymat2*(thevector~)/(thevector*mymat1*(thevector~));
   if(res<0, silent = 0;
      print("Value of tau = 1/theta = ", tau); print(thevector),);

   if(silent == 0,
      print("\nIn truth, minimum is ", bestvalue);
      print("   and our simplified version yields ", res+0.0);
      print("  Aux = ", (thevector)*mymat2*(thevector~));
      print("   I1 = ", (thevector)*mymat1*(thevector~)),);

   return([thevector, res, bestvalue])}

{Optimizedfunction(kappa, endpoints, lowertau = 0, uppertau = 1,
       steptau = 0.02, UsePrecomputedVals = 0, I1K1I2vectordata = [])=
   local(nb, nbbest = 10000, taubest, compteur = 0, nboptimal,
         resstuff, bestcoefs);

   if(UsePrecomputedVals == 0,
      PrecomputedVals = []; PrecomputedArgs = [],);
   NbPrecomputedValUsed = 0;
   if(I1K1I2vectordata == [],
      I1K1I2vectordata = [getvectorI1(kappa, endpoints),
                          getvectorK1(kappa, endpoints),
                          getvectorI2(kappa, endpoints)],);

   if(lowertau == 0, lowertau = 1/15,);
   resstuff = OptimizeK1andI2surI1(kappa, tau, I1K1I2vectordata, 0);

   forstep(tau = min(1, uppertau), lowertau, -steptau,
           compteur++ ;
```

```
            nboptimal = getlowerboundd0(kappa, resstuff[3], tau);
            nb = getlowerboundd0(kappa, resstuff[2], tau);
            if(nboptimal < nb,
               print("!!! Truncation has induced a severe loss!!");
               print("Increase multiplier in abetterversionof."),);
            if(nb <= nbbest,
               if(nb < nbbest,
                  print("\n[nb = ",nb,"] ",
                        "Best coefficients : ", resstuff[1]),);
               nbbest = nb; bestcoefs = resstuff[1]; taubest = tau,);
            if(compteur%20 == 1, print1("~[",nbbest,"]"),));

   print1("\nWe have precomputed ", length(PrecomputedArgs), " values of Sigmi");
   print(" and used them ", NbPrecomputedValUsed, " times.");
   print("\n >> n_0(", kappa, ")  = ", nbbest, " facteurs premiers. <<");
   print(" with Log Q / Log P = tau = ", taubest);
   print(" and the values: ", bestcoefs);
   return([nbbest, taubest, endpoints, bestcoefs])}

{regularpoints(nb)= vector(nb+1, k, (k-1)/nb)}
```

14.4. Plots and other extraction of datas

Plotting is not a main issue, but we record the code below so that it may be reused (almost) verbatim.

```
{rescaling(values)=
   for(k = 2, length(values), values[k] = values[k]/values[1] + 0.0);
   values[1] = 1; return(values)}

{InnerPlotFunction(kappa, n0, endpoints, values, WindowNumber = 1,
                   Color = 2, Comment = 0, Height = 0, OnFile = 0)=
   local(s, HPadding = 0.02, VPadding = 0.02, PSoffset = 0.18,
         NbVerticalPoints = 150, NbHorizontalPoints = 150);
   s = plothsizes(); s = [s[1],s[2]];
   plotinit(WindowNumber , s[1]-1, s[2]-1);
   plotcolor(WindowNumber , Color);
   values = rescaling(values);
   if(OnFile != 1, PSoffset = 0,);
   plotscale(WindowNumber, -HPadding, 1 + HPadding + PSoffset,
             -VPadding - PSoffset, 1 + VPadding);
   plotlines(WindowNumber, endpoints, values);
   if(Comment == 0,,
      /*--- Axis: ---*/
      plotlines(WindowNumber, [-HPadding, 1 + HPadding], [0, 0]);
      plotlines(WindowNumber, [0, 0], [-VPadding, 1 + VPadding]);
      /*--- Horizontal ticks: ---*/
      for(m = 1, 4,
         plotlines(WindowNumber,
                   [-HPadding/3, HPadding/3], [m/4, m/4]);
         plotmove(WindowNumber ,
                  -2*HPadding/3, m/4 + VPadding/3);
         plotstring(WindowNumber , Str(m/4));
         plotmove(WindowNumber , m/4, 0);
         for(n = 1, NbVerticalPoints,
            plotrpoint(WindowNumber, 0, 1/NbVerticalPoints)));
      /*--- Vertical ticks: ---*/
      for(m = 1, 4,
         plotlines(WindowNumber,
                   [m/4,m/4], [-VPadding/3, VPadding/3]);
         plotmove(WindowNumber , m/4 - 2*HPadding/3, -VPadding);
         plotstring(WindowNumber , Str(m/4));
         plotmove(WindowNumber , 0, m/4);
         for(n = 1, NbHorizontalPoints,
            plotrpoint(WindowNumber, 1/NbHorizontalPoints, 0)));
      /*------------------------*/
      plotmove(WindowNumber , 0.77, 1 - Height - VPadding);
      plotstring(WindowNumber , Str("kappa = ", kappa));
      plotmove(WindowNumber , 0.77, 1 - Height - 2*VPadding);
      plotstring(WindowNumber , Str("n_0 = ", n0)));}

{PlotFunction(kappa, plots, OnFile, Prefix)=
   local(listtodraw = [1, 1, 1], FileName = "Affine", olddefault);
   FileName = concat(Prefix, FileName);
   FileName = concat(FileName, Str(kappa));
   InnerPlotFunction(kappa, plots[1][1], plots[1][2], plots[1][3],
                           1, 1, 1, 0, OnFile);
   FileName = concat(concat(FileName, "-"), Str(length(plots[1][2])-1));
   for(p = 2, length(plots),
      InnerPlotFunction(kappa, plots[p][1], plots[p][2], plots[p][3],
                           p, p, 0, 0, OnFile);
      FileName = concat(concat(FileName, "-"), Str(length(plots[p][2])-1));
      listtodraw = concat(listtodraw, [p, 1 ,1]));
   if(OnFile == 1,
      olddefault = default(psfile);
      FileName = concat(FileName, ".ps");
      default(psfile, FileName);
      psdraw(listtodraw);
      default(psfile, olddefault), );
   plotdraw(listtodraw);}

{DoIt(kappa, listnbsub, OnFile = 0, Prefix = "")=
   local (plots = [], res, lastn0 = 0);
   res  = Optimizedfunction(kappa, regularpoints(listnbsub[1]), 0, 1, 0.01, 0);
   plots = [[res[1], res[3], res[4]]];
   lastn0 = res[1];
   for(n = 2, length(listnbsub),
      /* We can reuse precomputed values: */
      res  = Optimizedfunction(kappa, regularpoints(listnbsub[n]), 0, 1, 0.01, 1);
      if(res[1] != lastn0, print("!!!Warning!!! Varying n0!!!"));
      plots = concat(plots, [[res[1], res[3], res[4]]]));
   PlotAffineFunction(kappa, plots, OnFile, Prefix);}
```

The command `DoIt(20,[7,15,30,40],1)` produces the postcript file `Affine20-7-15-30-40.ps` that we convert to the eps format by the shell command `ps2eps`.

APPENDIX A

On the convolution method

Estimation of the average order of an arithmetic function by the method of convolution. Perrine Berment and Olivier Ramaré

A.1. Introduction

The arithmetic functions are very often poorly understood, and posses a behavior that appears irregular and inconsistent. More specifically, in this article we focus on the study of the function

$$f_0(n) = \prod_{p|n}(p-2). \tag{A.1}$$

The values of f_0 between 1 and 54 are as follows

$$1,0,1,0,3,0,5,0,1,0,9,0,11,0,3,0,15,0,17,0,5,0,21,0,3,0,1,0,27,\\ 0,29,0,9,0,15,0,35,0,11,0,39,0,41,0,3,0,45,0,5,0,15,0,51,0.$$

This does not give us much informations even if we consider only the values at odd integers of the same interval :

$$1,1,3,5,1,9,11,3,15,17,5,21,3,1,27,29,9,15,35,11,39,41,3,45,5,15,51.$$

For more informations, we can try to determine its average order, i.e. an approximation of $(1/X)\sum_{n\le X} f_0(n)$. Here we find regularity. We will in fact demonstrate that :

Theorem A.1 *Let X be a positive real number. For all real numbers σ in $]1/2,1]$, we have*

$$(1/X)\sum_{n\le X} f_0(n) = \mathcal{C}X + \mathcal{O}(X^\sigma)$$

where the implied constant in the $\mathcal{O}$-symbol depends on σ and where

$$\mathcal{C} = \tfrac{1}{2}\prod_{p\ge 2}\left(1-\frac{3}{p(p+1)}\right) = 0.14630\cdots$$

Note that in this statement and from now onwards, the letter p denotes a prime number.

The average order has the effect of concealing certain unusual values taken by the function considered. Note that as is the case for equivalents, we choose a well understood and quite simple function which allows one in addition to have small error term. Now what constitutes for a fairly simple function obviously depends on authors !

The method we explain here belongs to the folklore and did not appear in any systematic exposition as far as we know. It is very flexible in its use and give rise to very good error terms. It took 50 years to build a complete theory of average orders of multiplicative functions (see below for a definition), but this was mainly oriented towards a maximum extension of the class in question rather than to the quality of the error term. Note the theorem of Ikehara generalized by Delange [**16**], the theorem of Wirsing [**93**], another result of Delange [**17**], and the outstanding work of [**35**]. The reader will find exposition of these materials in the books by Apostol [**1**] and Tenenbaum [**?**].

Also note that we could very easily replace the $\mathcal{O}$-term in the theorem by an explicit inequality. It is little more difficult to write down the first few exact numeric digits of $\mathcal{C}$, but we leave aside this issue here and move forward.

Now we present an outline of the method of convolution. This is to determine the average order of an arithmetic function f. For this, we take a model function g which looks like f and of which we know the average order. The model for f_0 is the function that associates n to n, of which we obviously know the average order. How does one know that g is a model for f ? We will define convolution product $\star$ and show that there exists a function h satisfying $f = h \star g$, and where h is “smaller” than f. The average order of f will be obtained by determining one of the h, g, which will essentially be ruled by that of g.

Here is a word about the choice of the function f_0. This function has no prior geometric interpretation; it is not entirely true since its value at a square-free integer, say q, is the number of primitive Dirichlet characters modulo q. We have chosen f_0 precisely for its arbitrariness; its particular form allows us to simplify some parts of the exposition.

The proof itself is very short and is explained in section A.5, but we detail here the concepts used.

We would like to thank Hervé Queffelec heartily for his suggestions and remarks that have been essential to write this article.

A.2. Arithmetic functions

Arithmetic functions are frequently used in later parts, we begin by recalling some definitions and properties. First of all, an arithmetic function is defined on $\mathbb{N}^*$ with values in $\mathbb{C}$. We call it arithmetic if it has any arithmetic meaning ..., precisely !

Among these functions, the multiplicative functions play a special role.

Definition A.2 *We say that a function $f : \mathbb{N}^* \to \mathbb{C}$ is multiplicative if $f(1) = 1$, and also*

$$f(nm) = f(n)f(m), \quad \textit{when} \quad \gcd(n, m) = 1.$$

The multiplicative arithmetic functions are of major interest since their value at an integer is determined uniquely by the values at prime powers involved in the composition of this integer. Indeed,

$$f(n) = \prod_{p^\alpha \| n} f(p^\alpha), \tag{A.2}$$

where the notation $p^\alpha \| n$ means that $p^\alpha | n$ and $p^{\alpha+1} \nmid n$.

This in particular implies that a multiplicative function is completely determined by its values on the powers of prime numbers.

A.2.1. Bestiary.

Here are some well-known multiplicative functions :

- $\varphi(n)$: the Euler indicator. It is equal to the number of integers between 1 and n which are prime to n. The Chinese lemma allows us to show that it is multiplicative and $\varphi(p^\alpha) = p^{\alpha-1}(p-1)$.
- $d(n)$: the number of divisors on n.
- $\sigma(n)$: the sum of the divisors of n.
- $\delta_{n=k}$: the indicator function is 1 when $n = k$ and 0 otherwise.
- $\theta_\alpha(n)$: the function which sends n to n^α.
- 1 is a more common notation for θ_0.
- $\mu(n)$: the Mobius function. It is multiplicative and is -1 on each prime number and 0 on all of their higher powers.
- $\mu^2(n)$: the indicator function of the square-free integers. It is 0 if n is divisible by any square greater than 1 and 1 otherwise.
- $\lambda(n)$: the Liouville function is multiplicative and is equal to $(-1)^k$ on all p^k.

A.2.2. Convolution Product.

We define the convolution product of arithmetic function f and g by :

$$(f \star g)(n) = \sum_{d|n} f(n/d)g(d). \tag{A.3}$$

This product is associative and commutative. The function $\delta_{n=1}$ is the identity element, since for any arithmetic function g, we have

$$(\delta_1 \star g)(n) = \sum_{\ell m=n} \delta_1(\ell) g(m) = g(n).$$

This product is also distributive with respect to the addition of two arithmetic functions and both laws give the set of all arithmetic functions a commutative algebra structure with unity over $\mathbb{C}$. We could also enrich this structure by considering the derivation

$$\partial : (f(n))_{n\geq 1} \mapsto (f(n)\log n)_{n\geq 1}$$

which is linear and satisfy the addition $\partial(f \star g) = (\partial f) \star g + f \star (\partial g)$ but we go out of scope here. The reader will find a fairly detailed exposition of this structure in the book by Bateman & Diamond [**2**].

Exercise A.3 *Show that, if $D(f,s)$ converges absolutely, it is the same for $D(\partial f, r)$ for $r > s$. What about the converse ? Is it possible to weaken the condition by $r \geq s$?*

Here is the main theorem on multiplicative functions :

Theorem A.4 *If f and g are two multiplicative functions, then so is $f \star g$.*

Let us start with a lemma.

Lemma A.5 *Let m and n be two co-prime integers. For any function F, we have*

$$\sum_{d|mn} F(d) = \sum_{d_1|m} \sum_{d_2|n} F(d_1 d_2).$$

Proof Given an integer n, we denote by $\mathcal{D}(n)$ the set of all positive divisors of n. For example, we have $\mathcal{D}(12) = \{1,2,3,4,6,12\}$. Consider the following two functions :

$$\begin{aligned} \Phi : \mathcal{D}(m) \times \mathcal{D}(n) &\to \mathcal{D}(mn), & \Psi : \mathcal{D}(mn) &\to \mathcal{D}(m) \times \mathcal{D}(n), \\ (d_1, d_2) &\mapsto d_1 d_2 & d &\mapsto (\gcd(d,m), \gcd(d,n)). \end{aligned}$$

The proof is to show that $\Psi \circ \Phi = Id$ and $\Phi \circ \Psi = Id$ and the statement in the lemma is the functional translation of that.

Let $(d_1, d_2) \in \mathcal{D}(m) \times \mathcal{D}(n)$. We have $\gcd(d_1 d_2, m) = d_1$ and, similarly $\gcd(d_1 d_2, n) = d_2$, which guarantees that $\Psi \circ \Phi = Id$. Then, if d is a divisor of mn, we have

$$\gcd(d, mn) = \gcd(d,m)\gcd(d,n)$$

which allows us to conclude that $\Phi \circ \Psi = Id$. □

The proof of Theorem A.4 is a simple application of this lemma.

Proof Clearly, we have $(f \star g)(1) = f(1)g(1) = 1$. Let m and n be two co-prime integers. Then, by definition :

$$(f \star g)(mn) = \sum_{d|mn} f\left(\frac{mn}{d}\right) g(d).$$

Here we recall lemma A.5, which gives us

$$\begin{aligned}(f \star g)(mn) &= \sum_{d_1|m} \sum_{d_2|n} f\left(\frac{mn}{d_1 d_2}\right) g(d_1 d_2) \\ &= \sum_{d_1|m} \sum_{d_2|n} f(m/d_2) f(n/d_2) g(d_1) g(d_2) = (f \star g)(m)(f \star g)(n)\end{aligned}$$

as desired. □

This allows us to obtain certain multiplicative function identities, as for example, we have : $d(n) = (1 \star 1)(n)$. The reader will note that Lemma A.5 is in fact equivalent to this multiplicativity !

Exercise A.6 *Show that $1 \star \lambda$ is the characteristic function of squares.*

Exercise A.7 *Let f and g be the functions defined by $f(n) = d(n)^2$ and $g(n) = d(n^2)$. Show that $f = 1 \star g$.*

Exercise A.8 *Show that $\theta_1 = 1 \star \varphi$.*

A.3. Dirichlet series

We only talk about Dirichlet series with real arguments and restrict ourselves in the area of absolute convergence which will suffice here. For studying the complex case, see [**87**] and [**24**].

When we have an arithmetic function, say f, we can form its Dirichlet series which is, for any real argument s :

$$D(f, s) = \sum_{n \geq 1} f(n)/n^s. \tag{A.4}$$

A priori this definition is formal, since there might not even be an s for which the series converges (it is the case when $f(n) = e^n$).

Corollary A.9 *Let f be an arithmetic function such that $D(f, s)$ converges absolutely for some s. Then, for all $r > s$, the series $D(f, r)$ converges absolutely.*

Proof We have

$$D(f, r) = \sum_{n \geq 1} \frac{f(n)}{n^s} \frac{n^s}{n^r}.$$

If $r > s$ then $n^s/n^r < 1$, hence $D(|f|, r) < D(|f|, s)$, i.e. the Dirichlet series of f converge for all $r > s$. □

This property introduces us the notion of abscissa of convergence.

Definition A.10 *We call the* abscissa of convergence *of the function f, the smallest real number s such that the Dirichlet series $D(f,s)$ converges. If $D(f,s)$ converge for all s, we then say that the abscissa of convergence is $-\infty$.*

Note that it is not certain that the series in question converges on its abscissa of convergence. By a theorem of Landau, see [**21**], it can never happen if the abscissa is finite and f is non negative. We note that the abscissa of convergence can be $-\infty$ and the function f is non negative, without implying that f has bounded support i.e. it vanishes everywhere except on a finite set. The function $f(n) = e^{-n}$ is a counter-example.

Corollary A.11 *Suppose that the Dirichlet series of a multiplicative function f converge absolutely for some s. Then, $D(f,s)$ has the following Euler product expansion:*

$$D(f,s) = \prod_{p\geq 2} \sum_{k\geq 0} \frac{f(p^k)}{p^{ks}}.$$

Proof Write n as a product of prime powers and thanks to multiplicativity of f, we get the result. For a complete proof, the reader can consult [**87**]. □

We can associate to each arithmetic function f, a Dirichlet series and this series converges at least at one point if the function grows reasonably. Such Dirichlet series defines the function as shown in the following property. We will use it in future.

Corollary A.12 *Let f and g be two arithmetic functions such that their respective Dirichlet series converge absolutely for some s. Suppose further that $D(f,r) = D(g,r)$ pour tout $r > s$. Then $f = g$.*

Proof If $h_1 = f - g$, we have $D(h_1, r) = 0$ for all $r > s$. Since this series converges for $r = s+1$, we deduce that the absolute value of $h_2(n) = h_1(n)/n^{r+1}$ is bounded and $D(h_2, r) = 0$ for all $r > -1$ and we have to prove that $h_2 = 0$. Suppose this is not the case and call n_0 is the smallest integer n such that $h_2(n) \neq 0$. A direct comparison with an integral gives us for $r > 1$:

$$\begin{aligned}|n_0^r D(h_2,r) - h_2(n_0)| &\leq \max_n |h_2(n)| \sum_{n\geq n_0+1} \frac{n_0^r}{n^r} \\ &\leq \max_n |h_2(n)| n_0^r \int_{n_0}^{\infty} \frac{dt}{t^r} \leq \max_n |h_2(n)| n_0/(r-1),\end{aligned}$$

which tends to 0 as r tends to infinity. But $D(h_2, r) = 0$, which guarantees that $h_2(n_0) = 0$ contrary to our hypothesis. The reader can modify the proof in two ways : first replace the method of contradiction by a proof by induction. Then a small change gives us that

$$n_0^r D(h_2, r) - h_2(n_0) = \mathcal{O}((1 + n_0^{-1})^{-r})$$

whence we have $\mathcal{O}(1/r)$. □

A.3.1. The Riemann ζ function.

The function ζ is an important arithmetic function since it occurs in Euler's formula, it is therefore a link between natural numbers and prime numbers. Its Dirichlet series is as simple as the one associated with the constant function 1 (we denote this by θ_0 and 1 in our bestiary). It is defined for $s > 1$ by

$$\zeta(s) = \sum_{n \geq 1} n^{-s}.$$

This Dirichlet series is the simplest and well known, however we still do not know it enough. For further study of this function, the reader is refer to [**87**] and [**24**]. By Proposition 2, the function ζ has an Euler product

$$\zeta(s) = \prod_{p \geq 2} (1 - p^{-s})^{-1}, \tag{A.5}$$

which converges absolutely for $s > 1$.

A.3.2. Dirichlet series and convolution product.

The two operations of arithmetic functions translates very nicely in terms of Dirichlet series :

- Regarding addition (+) : given two functions f and g whose Dirichlet series converge absolutely for s, we have
$$D(f + g; s) = D(f; s) + D(g; s).$$
- Regarding multiplication ($\star$) : given two functions f and g whose Dirichlet series converge absolutely for s, then so is true for the Dirichlet series of $f \star g$ at s, and we have
$$D(f \star g; s) = D(f; s)D(g; s).$$

This equality of absolute convergent series is easily verified as we can interchange the terms of the series whenever required. It shows in particular that the operator which maps an arithmetic function to its Dirichlet series trivializes the convolution in the same way the Fourier transform trivializes the convolution of the functions of the real line.

The notion of abscissa of convergence leads us to define the size of an arithmetic functions : the function f_1 is bigger than the function f_2 if its abscissa of convergence is more ! The convolution method consists in writing $f = g \star h$ where g has the same abscissa convergence as f and h has smaller abscissa of convergence. We will simply say that h is a small perturbation and f is a perturbed version of g.

It is clear that the abscissa of convergence of the Dirichlet series of $f \star g$ is bounded by the maximum abscissa of absolute convergence Dirichlet series associated with f and g. This increase is often equal if the two abscissas are not equal ... and none of the factors is zero !

Exercise A.13 *Show that*

(1) $D(\varphi, s) = \zeta(s-1)/\zeta(s)$,
(2) if $f(n) = d(n)^2$, *then* $D(f,s) = \zeta(s)^4/\zeta(s)$,
(3) if $f(n) = d(n^2)$, *then* $D(f,s) = \zeta(s)^3/\zeta(s)$,
(4) $D(\lambda, s) = \zeta(2s)/\zeta(s)$,
(5) $D(\mu^2, s) = \zeta(s)/\zeta(2s)$.

Exercise A.14 *Show that the Dirichlet series associated to the Mobius function* μ *is* $1/\zeta(s)$ *and deduce an example where the abscissa of absolute convergence of a product can be strictly less than the greater of the two factors abscissae (consider* $1 \star \mu$ *).*

A.4. Summation by parts

We make a detour here to address a very useful technique for the kind of problem that concerns us here, but is surprisingly poorly understood. This is the summation by parts, in our context, of course. We describe this process by an example. Begin by recalling that, for all $t \in \mathbb{R}$, we have

$$\sum_{n \le t} 1 = t + \mathcal{O}(1).$$

Now suppose that we want to get an approximation of $\sum_{n \le X} 1/n$.

We need only to note that :

$$\frac{1}{n} = \frac{1}{X} + \int_n^X \frac{dt}{t^2}. \tag{A.6}$$

We then have :

$$\sum_{n \le X} \frac{1}{n} = \frac{\sum_{n \le X} 1}{X} + \int_1^X \sum_{n \le t} 1 \frac{dt}{t^2} = \log X + \mathcal{O}(1).$$

Similarly, we can, for example, evaluate the sum $\sum_{n \le X} \log n$.

To further illustrate this technique, we prove of the following corollary of Theorem A.1 :

Theorem A.15
We have $\sum_{n \le X} f_0(n)/n = 2\mathcal{C}X + \mathcal{O}(X^\sigma)$ *for all real* σ *in* $]1/2, 1]$.

Proof Indeed, we use (A.6) to obtain :

$$\sum_{n\le X} f_0(n)/n = \sum_{n\le X} f_0(n)\Big(\frac{1}{X}+\int_n^X \frac{dt}{t^2}\Big)= \frac{\sum_{n\le X} f_0(n)}{X} + \int_1^X \sum_{n\le t} f_0(n)\frac{dt}{t^2}$$

$$= \mathcal{C}X + \mathcal{O}(X^\sigma) + \mathcal{C}\int_1^X dt + \mathcal{O}\Big(\int_1^X t^{\sigma-1}dt\Big)$$

which gives us the required result. □

The reader may find similar theorems by the average order means of $f = g \star h$ knowing that of g, as in [**70**, lemma 3.2].

A.5. Proof of Theorem A.1

First Step. Let's start by considering the Dirichlet series of f_0 and its Euler product expansion. We have, by definition :

$$D(f_0, s) = \prod_{p\ge 2}\sum_{k\ge 0} \frac{\prod_{\ell | p^k}(\ell-2)}{p^{ks}}.$$

Take a closer look at each factor. In the sum over k, the contribution of $k = 0$ is non-trivial and is 1 ; the Euler factor in p becomes :

$$1 + \sum_{k\ge 1} \frac{\prod_{\ell | p^k}(\ell-2)}{p^{ks}}.$$

Since ℓ and p are prime numbers, this forces that $\ell = p$. We have

$$\sum_{k\ge 1}(p-2)/p^{ks} = \frac{p-2}{p^s-1}.$$

Here the Dirichlet series associated to f_0 is :

$$D(f_0, s) = \prod_{p\ge 2}\Big(1+\frac{p-2}{p^s-1}\Big). \tag{A.7}$$

We note that the product $\prod_{p\ge 2}(1+\frac{1}{p^{s-1}-1})$ correspond to $\zeta(s-1)$. Therefore keep this factor out of our product. We write

$$D(f_0, s) = \prod_{p\ge 2}\Big(1+\frac{p-2}{p^s-1}\Big) = \prod_{p\ge 2}\Big(1-\frac{2p^{s-1}+p-3}{(p^s-1)p^{s-1}}\Big)\Big(\frac{1}{1-1/p^{s-1}}\Big)$$

$$= H(s)\zeta(s-1). \tag{A.8}$$

The product defining $H(s)$ converge absolutely for those s for which the series $\sum \frac{2p^{s-1}+p-3}{(p^s-1)p^{s-1}}$ converge absolutely, which takes place for $s > 3/2$ by extending the sum to all integers. The abscissa of absolute convergence of $\zeta(s-1)$ is equal to 2, it is also possible for $D(f_0, s)$, so much so that that the series H converge in a wider domain. If we write the series H as a Dirichlet series of a function, then it will indeed be smaller than

f_0 in the sense defined in subsection A.3.2. We further comment on the possibility of the above situation. Note first that we use this notion "size" as a guide in our calculations and a rough heuristic will suffice. But we can also show a posterior that the abscissa of convergence (and absolute convergence here since f_0 is non negative) is actually equal to 2. Indeed, suppose the theorem A.1 is proved. A summation by parts gives us

$$\begin{aligned}\sum_{1\le n\le N} f_0(n)/n^s &= \sum_{1\le n\le N} f_0(n)\Big(\frac{1}{N^s}+s\int_n^N\frac{dt}{t^{s+1}}\Big)\\ \text{(A.9)}\qquad &= \frac{\sum_{1\le n\le N} f_0(n)}{N^s}+s\int_1^N\sum_{1\le n\le t} f_0(n)\frac{dt}{t^{s+1}}\\ &= \mathcal{C}N^{2-s}+s\,\mathcal{C}\int_1^N\frac{dt}{t^{s-1}}+\mathcal{O}(N^{1+\sigma-s})+\mathcal{O}\Big(\int_1^N dt/t^{s-\sigma}\Big).\end{aligned}$$

By taking $\sigma = 0.6$ for example, we see that the series defining $D(f_0, s)$ converge for $s > 2$ and diverge for $s < 2$. We continue this discussion in the last section of this article.

Second Step. We must now write the two functions $H(s)$ and $G(s) = \zeta(s-1)$ as Dirichlet series. The case of G is easy since $G(s) = D(\theta_1, s)$. Let us now turn to H. We look for a function h such that

$$H(s)=\sum_{n\ge1} h(n)/n^s \tag{A.10}$$

and we restrict ourselves on multiplicative functions. We therefore look for a function h such that :

$$\sum_{k\ge0}\frac{h(p^k)}{p^{ks}} = 1-\frac{2p^{s-1}+p-3}{(p^s-1)p^{s-1}}. \tag{A.11}$$

The condition $h(1) = 1$ is used to settle the case $k = 0$. We just need to take care of the sum corresponding to $k \ge 1$. We set $z = 1/p^s$, and get the right hand side of (A.11) as a rational function in z as follows :

$$-\frac{\frac{2}{pz}+p-3}{(\frac{1}{z}-1)\frac{1}{pz}} = \frac{2z+p^2z^2-3pz^2}{z-1} = -2\sum_{k\ge1} z^k-(p^2-3p)\sum_{k\ge2} z^k.$$

Identifying term by term, we see that the multiplicative function h is defined by

$$\begin{cases} h(p) = -2,\\ h(p^k) = -(p^2-3p+2) \quad \text{for } k\ge 2,\end{cases} \tag{A.12}$$

solves our problem.

Third Step. Since $D(f_0, s) = H(s)\zeta(s-1)$ and we know the Dirichlet series expansion of $H(s)$ and $\zeta(s-1)$, $D(f_0, s)$ is a product of two Dirichlet series and is therefore an arithmetic convolution product. Property 3 allows us to identify it term by term and conclude that $f_0 = \theta_1 \star h$.

Here we present another method which is more pedestrian. Consider the multiplicative function h defined by (A.12) and recall that θ_1 is a function which maps $n \mapsto n$. We note that : $\theta_1 \star h$ is still a multiplicative function which is therefore defined by its values on the prime powers. Now, for prime p and natural number $k \geq 1$:

$$\begin{aligned}(\theta_1 \star h)(p^k) &= \sum_{\ell/p^k} \theta_1(\frac{p^k}{\ell})h(\ell) = \sum_{\ell/p^k} \frac{p^k}{\ell} h(\ell)\\ &= p^k\left(1 - \frac{2}{p} - \sum_{2\leq t\leq k} \frac{p^2-3p+2}{p^t}\right) = f(p^k).\end{aligned}$$

By multiplicativity, this implies that $f_0(n) = (\theta_1 \star h)(n)$. We write this identity in an explicit form below for further use :

$$f_0(n) = \sum_{\ell m=n} h(\ell)\theta_1(m) = \sum_{\ell m=n} h(\ell)m.$$

We are now beginning to calculate the average order of f_0. The above equality gives us

$$\sum_{n\leq X} f_0(n) = \sum_{\ell m\leq X} h(\ell)m = \sum_{\ell\leq X} h(\ell) \sum_{m\leq X/\ell} m. \tag{A.13}$$

Now, we know that, for all natural numbers N :

$$\sum_{n\leq N} n = N(N+1)/2,$$

which implies that for real numbers $M \geq 1$, we have :

$$\sum_{m\leq M} m = \tfrac{1}{2}M(M+1) + \mathcal{O}(M) = \tfrac{1}{2}M^2 + \mathcal{O}(M). \tag{A.14}$$

First note that this estimate is valid as soon as M is non negative. It turns out that the proposed method is much simplified if we are satisfied with a weaker error term. The estimate (A.14) in fact implies also that

$$\sum_{m\leq M} m = \tfrac{1}{2}M^2 + \mathcal{O}(M^\sigma) \tag{A.15}$$

for all $\sigma \in [1, 2]$ and all $M \geq 0$. We now have all the tools to conclude. We take the proof of the equation (A.13) considered earlier and see that the condition $\ell \leq X$ is superfluous. It gives then directly

$$\sum_{n\leq X} f_0(n) = \frac{X^2}{2} \sum_{\ell\geq 1} \frac{h(\ell)}{\ell^2} + \mathcal{O}\Big(X^\sigma \sum_{\ell\geq 1} \frac{|h(\ell)|}{\ell^\sigma}\Big).$$

Since $H(s)$ converge absolutely for $s > 3/2$, the sum $\sum_{\ell\geq 1} |h(\ell)|/\ell^\sigma$ is finite for all $\sigma > 3/2$. The proof of the theorem A.1 is complete by renaming σ.

By the same method . . .

The reader, following a method similar to that proposed in the proof of theorem A.1, find the average order using the function φ.

A.6. Some digressions without proof

The Dirichlet series were introduced in [**20**] by P.G. Lejeune-Dirichlet in 1937 to show the existence of infinitely many prime numbers in an arithmetic progressions (of the form $a + nq$ where a and q are relatively prime). Dedekind, first a student and then a friend of Dirichlet has established several properties of these series enriching the book by [**59**]. The structure of the next step is due to the memoir of Cahen [**10**], which is famous for the inaccuracy of its worth! The development of the theory has gone well underway at this time and in 1915 appeared the splendid little monograph [**41**] of Hardy & Riesz which to date remains the basis work on the question. The reader can find a part of this material in [**87**].

Here we focus on two points :

(1) To what extent the average order and the abscissa of convergence are related?

(2) Does writing $D(f, s) = D(h, s)D(g, s)$ (established in (A.8)) allows one to conclude that the abscissa of absolute convergence of the series $D(f, s)$ is the same as that of $D(g, s)$?

Regarding the first point, we have seen in section A.5 that knowledge of the average order of the function f allowed us to deduce the abscissa of absolute convergence of $D(f, s)$. The converse is false simply because it is quite possible that f has no average order. These two notions are linked by the following theorem (due to Cahen [**10**]) :

Theorem A.16 *If the abscissa of absolute convergence σ_0 of $D(f, s)$ is strictly positive, it is given by*

$$\sigma_0 = \limsup_{n\to\infty} \frac{\log \sum_{1\leq n\leq N} |f(n)|}{\log N}.$$

There is an analogous theorem to determine the abscissa of convergence (we did not establish its existence !), and it is also possible to treat the case σ_0 is non negative (but the formula is different). The reader will note that this formula is the exact counterpart of the Hadamard formula giving the radius of convergence of power series, subject to recall the identity we have (almost !) demonstrated in (A.9) :

$$D(|f|,s) = s\int_1^\infty \Big(\sum_{n\le t} |f(n)|\Big) dt/t^{s+1}.$$

Let us now turn to the second question. We assume that here we have a decomposition of the form $D(f,s) = D(h,s)D(g,s)$, where we know the abscissa of absolute convergence, say σ_0 of $D(g,s)$ and where that of $D(h,s)$ is strictly smaller. Can we conclude that σ_0 is still the abscissa of absolute convergence σ_0' of $D(f,s)$? It is clear that $\sigma_0' \le \sigma_0$, but can it be smaller ? This is obviously true if $h = 0$, is it so if $h \ne 0$? The authors of this article do not know how to answer this general question, but it is permissible in this case of application to add a hypothesis : we assume that for all $\delta > 0$, the modulus of $D(h,s)$ is bounded below when s is in the complex half-plane $\Re s \ge \sigma_0 + \delta$. This hypothesis does not cost us anything in practice since we get $D(h,s)$ as an Euler product converge since it is neither infinite nor zero. But we must now consider the s of the complex field, we had manage to avoid till now ! Here is the theorem we are interested :

Theorem A.17 *Let $D(h,s)$ be a Dirichlet series absolutely convergent for $\Re s \ge \sigma$ and is bounded below by a constant > 0, then $1/D(h,s)$ is still a Dirichlet series absolutely convergent for $\Re s \ge \sigma$.*

The result allows us to write $D(g,s) = D(h,s)^{-1}D(f,s)$ and to conclude that $\sigma_0' \ge \sigma_0$, which gives us many $\sigma_0 = \sigma_0'$.

Many studies compare the abscissa of simple or uniform absolute convergence of the three equal components $D(f,s) = D(h,s)D(g,s)$; the reader will find a presentation and their extensions to the case of several factors and the largest improvement (optimal) in [**51**].

Notation

Notation used throughout these notes is standard ... in one way or the other! Here is a guideline:

— $e(y) = \exp(2i\pi y)$.

— The use of the letter p for a variable always implies this variable is a prime number.

— $[d, d']$ stands for the lcm and (d, d') for the gcd of d and d', while $[t]$ denotes the integer part of the real number t. In this context $\{t\}$ denotes the fractionnal part of t.

— $|\mathcal{A}|$ stands for the cardinality of the set $\mathcal{A}$ while $\mathbb{1}_{\mathcal{A}}$ stands for its characteristic function.

— $q\|d$ means that q divides d in such a way that q and d/q are coprime. In words we shall say that *q divides d exactly.*

— The squarefree kernel of the integer $d = \prod_i p_i^{\alpha_i}$ is $\prod_i p_i$, the product of all prime factors of d.

— $\omega(d)$ is the number of prime factors of d, counted without multiplicity.

— $\phi(d)$ is the Euler totient, i.e. the cardinality of the multiplicative group of $\mathbb{Z}/d\mathbb{Z}$.

— $\sigma(d)$ is the number of positive divisors of d, except in section 4 where it will denote a density.

— $\tau(d)$ is the number of positive divisors of d.

— $\tau_k(d)$ is the number of k-tuples of (positive) integers $(d_1, \cdots, d_k)$ such that $d_1 \cdots d_k = d$, so that $\tau_2 = \tau$.

— $\mu(d)$ is the Moebius function, that is 0 when d is divisible by a square > 1 and otherwise $(-1)^r$ otherwise, where r is the number of prime factors of d.

— $c_q(n)$ is the Ramanujan sum. It is the sum of $e(an/q)$ over all a modulo q that are prime to q.

— $\Lambda(n)$ is van Mangoldt function: which is $\operatorname{Log} p$ is n is a power of the prime p and 0 otherwise.

— The notation $f = \mathcal{O}_A(g)$ means that there exists a constant B such that $|f| \leq B\,g$ but that this constant may depend on A. When we put in several parameters as subscripts, it simply means the implied constant depends on all of them.

- — The notation $f = \mathcal{O}^*(g)$ means that $|f| \le g$, that is a $\mathcal{O}$-like notation, but with an implied constant equal to 1.
- — The notation $f \star g$ denotes the arithmetic convolution of f and g, that is to say the function h on positive integers such that $h(d) = \sum_{q|d} f(q)g(d/q)$. exists for every real number x.
- — $\mathcal{U}$ is the compact set $(\mathcal{U}_d)_d$ where, for each d, $\mathcal{U}_d$ is the set of invertible elements modulo d.
- — π is ... the usual real number about $3.141\,5\ldots$! But also identifies the counting function of the primes: $\pi(6) = 3$ for instance. We have tried to avoid this notation when not too awkward, just as we did not use the Chebyshev ϑ and ψ functions except in chapter 5. We also used the variations $\vartheta(x;\chi)$, $\vartheta(x;q,a)$, $\psi(x,\chi)$ and $\psi(x;q,a)$.
- — The letter ψ^* is for local model in chapter 10, see (10.1).
- — $\mathbb{1}$ denotes a characteristic function in one way or another. For instance, $\mathbb{1}_{\mathcal{K}_d}$ is 1 if $n \in \mathcal{K}_d$ and 0 otherwise, but we could also write it as $\mathbb{1}_{n\in\mathcal{K}_d}$, closer to what is often called the Dirac δ-symbol. We also use $\mathbb{1}_{(n,d)=1}$ and $\mathbb{1}_{q=q'}$.

Bibliography

[1] T.M. Apostol. *Introduction to analytic number theory.* Springer-Verlag, New York, 1976. Undergraduate Texts in Mathematics.

[2] P.T. Bateman and H.G. Diamond. *Analytic number theory.* World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2004. An introductory course.

[3] D. Berkane, O. Bordellès, and O. Ramaré. Explicit upper bounds for the remainder term in the divisor problem and two applications. *Math. of Comp.*, pages 1–23, 2011.

[4] E. Bombieri. On the large sieve method. *Mathematika*, 12:201–225, 1965.

[5] E. Bombieri. A note on the large sieve. *Acta Arith.*, 18:401–404, 1971.

[6] E. Bombieri. Le grand crible dans la théorie analytique des nombres. *Astérisque*, 18:103pp, 1987/1974.

[7] E. Bombieri. Le grand crible dans la théorie analytique des nombres. *Astérisque*, 18:103pp, 1987/1974.

[8] E. Bombieri and H. Davenport. On the large sieve method. *Abh. aus Zahlentheorie und Analysis zur Erinnerung an Edmund Landau*, Deut. Verlag Wiss., Berlin:11–22, 1968.

[9] V. Brun. La série $\frac{1}{5}+\frac{1}{7}+\frac{1}{11}+\frac{1}{13}+\frac{1}{17}+\frac{1}{19}+\frac{1}{29}+\frac{1}{31}+\frac{1}{41}+\frac{1}{43}+\frac{1}{59}+\frac{1}{61}+\cdots$ où les dénominateurs sont ”nombres premiers jumeaux” est convergente ou finie. *Darboux Bull.*, 43(2):100–104, 124–128, 1919.

[10] E. Cahen. Sur la fonction $\zeta(s)$ de Riemann et sur des fonctions analogues. 1894. `http://www.numdam.org/item?id=ASENS_1894_3_11__75_0`.

[11] J. Cazaran and P. Moree. On a claim of Ramanujan in his first letter to Hardy. *Expositiones Mathematicae*, 17:289–312, 1999. based on a lecture given 01-12-1997 by J. Cazaran at the Hardy symposium in Sydney.

[12] Jing-run Chen. On the representation of a larger even integer as the sum of a prime and the product of at most two primes. *Sci. Sin.*, 16:157–176, 1973.

[13] Jing-run Chen. On the distribution of almost primes in an interval. *Sci. Sin.*, 18:611–627, 1975.

[14] János D.A. Goldston amd J. Pintz and C.Y. Yıldırım. Primes in tuples. III. On the difference $p_{n+\nu}-p_n$. *Funct. Approx. Comment. Math.*, 35:79–89, 2006.

[15] A. de Polignac. ””. *Comptes Rendus Acad. Sciences Paris*, pages 400, 738–739, 1849.

[16] H. Delange. Généralisation du théorème de Ikehara. *Ann. Sci. Ecole Norm. Sup. (3)*, 71:213–242, 1954. `http://www.numdam.org/item?id=ASENS_1954_3_71_3_213_0`.

[17] H. Delange. Un théorème sur les fonctions arithmétiques multiplicatives et ses applications. *Ann. Sci. École Norm. Sup. (3)*, 78:1–29, 1961. `http://www.numdam.org/item?id=ASENS_1961_3_78_1_1_0`.

[18] H. Diamond and H. Halberstam. Some applications of sieves of dimension exceeding 1. In *Sieve methods, exponential sums, and their applications in number*

theory (Cardiff, 1995), volume 237 of *London Math. Soc. Lecture Note Ser.*, pages 101–107. Cambridge Univ. Press, Cambridge, 1997.

[19] H.G. Diamond, H. Halberstam, and W.F. Galway. *A higher-dimensional sieve method*, volume 177 of *Cambridge tracts in mathematics*. Cambridge University Press, 2008.

[20] P.G.L. Dirichlet. Beweis des Satzes, das jede unbegrenzte arithmetische Progression, deren erstes Glied und Differenz ganze Zahlen ohne gemeinschaftlichen Factor sind, unendlich viele Primzahlen enthält. *Abhandlungen der Königlichen Preussischen Akademie der Wissenschaften zu Berlin*, 1937. Scan de l'article original : `http://bibliothek.bbaw.de/bibliothek-digital/digitalequellen/schriften/anzeige?band=07-abh/1837&seite:int=00000286` et traduction : `http://arxiv.org/abs/0808.1408`.

[21] F. Dress. Thormes d'oscillations et fonction de Möbius. *Sémin. Théor. Nombres, Univ. Bordeaux I*, Exp. No 33:33pp, 1983/84. `http://resolver.sub.uni-goettingen.de/purl?GDZPPN002545454`.

[22] P. Dusart. *Autour de la fonction qui compte le nombre de nombres premiers*. PhD thesis, Limoges, `http\string://www.unilim.fr/laco/theses/1998/T1998_01.pdf`, 1998. 173 pp.

[23] P.D.T.A. Elliott. On maximal variants of the Large Sieve. II. *J. Fac. Sci. Univ. Tokyo, Sect. IA*, 39(2):379–383, 1992.

[24] W.J. Ellison. *Les nombres premiers*. Hermann, Paris, 1975. En collaboration avec Michel Mendès France, Publications de l'Institut de Mathématique de l'Université de Nancago, No. IX, Actualités Scientifiques et Industrielles, No. 1366.

[25] T.J. Engelsma. K-tuple, permissible patterns. *http://www.opertech.com/primes/k-tuples.html*, 2009.

[26] P. Erdös and H. Riesel. On admissible constellations of consecutive primes. *BIT*, 28(3):391–396, 1988.

[27] T. Forbes. Prime k-tuplets. *http://anthony.d.forbes.googlepages.com/ktuplets.htm*. From 1996.

[28] C. Franze. *A Lower Bound Sieve Method with Applications*. PhD thesis, Central Michigan University, 2010.

[29] J. Friedlander and H. Iwaniec. *Opera de cribro*, volume 57 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2010.

[30] P.X. Gallagher. Sieving by prime powers. *Acta Arith.*, 24:491–497, 1974.

[31] D.A. Goldston, S.W. Graham, J. Pintz, and C.Y. Yıldırım. Small gaps between products of two primes. *Proc. London Math. Soc.*, 98(3):741–774, 2009.

[32] D.A. Goldston, J. Pintz, and C.Y. Yıldırım. Primes in Tuples. I. *Ann. of Math.*, 170(2):819862, 2009. available at arxiv under reference math.NT/0508185.

[33] G. Greaves. The weighted linear sieve and Selberg's λ^2-method. *Acta Arith.*, 47(1):71–96, 1986.

[34] G. Greaves. *Sieves in number theory*, volume 43 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag, Berlin, 2001. xii+304 pp.

[35] G. Halász. Über die Mittelwerte multiplikativer zahlentheorischer funktionen. *Acta Math. Acad. Sci. Hungar.*, 19:365–403, 1968.

[36] H. Halberstam and H.E. Richert. Mean value theorems for a class of arithmetic functions. *Acta Arith.*, 43:243–256, 1971.

[37] H. Halberstam and H.E. Richert. Sieve methods. *Academic Press (London)*, page 364pp, 1974.

[38] H. Halberstam and H.E. Richert. On a result of R. R. Hall. *J. Number Theory*, 11:76–89, 1979.

[39] H. Halberstam and H.E. Richert. Almost-primes in short intervals. *[A] Recent progress in analytic number theory, Symp. Durham 1979*, 1:69–101, 1981.

[40] R.R. Hall. Halving an estimate obtained from Selberg's upper bound method. *Acta Arith.*, 25:347–351, 1974.

[41] G. H. Hardy and M. Riesz. *The general theory of Dirichlet's series.* Cambridge Tracts in Mathematics and Mathematical Physics, No. 18. Stechert-Hafner, Inc., New York, 1964. Première édition en 1915.

[42] 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.*, 44:1–70, 1922.

[43] D.R. Heath-Brown. Almost-prime k-tuples. *Mathematika*, 44(2):245–266, 1997.

[44] D. Hensley and I. Richards. Primes in intervals. *Acta Arith.*, 4(25):375–391, 1974.

[45] Kwan-Hung Ho and Kai-Man Tsang. On almost prime k-tuples. *J. Number Theory*, 120(1):33–46, 2006.

[46] H. Iwaniec. A new form of the error term in the linear sieve. *Acta Arith.*, 37:307–320, 1980.

[47] H. Iwaniec and E. Kowalski. *Analytic number theory.* American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, 2004. xii+615 pp.

[48] H. Iwaniec and M. Laborde. p_2 in short intervals. *Ann. Inst. Fourier*, 4(31):37–56, 1981.

[49] J. Johnsen. On the large sieve method in $gf[q,x]$. *Mathematika*, 18:172–184, 1971.

[50] M. Jutila. Zero-density estimates for L-functions. *Acta Arith.*, 32:55–62, 1977.

[51] J.-P. Kahane and H. Queffélec. Ordre, convergence et sommabilité de produits de séries de Dirichlet. *Ann. Inst. Fourier (Grenoble)*, 47(2):485–529, 1997. http://www.numdam.org/item?id=AIF_1997__47_2_485_0.

[52] Jia Hai Kan. On the number of solutions of $N-p=P_r$. *J. Reine Angew. Math.*, 414:117–130, 1991.

[53] I. Kobayashi. A note on the Selberg sieve and the large sieve. *Proc. Japan Acad.*, 49(1):1–5, 1973.

[54] E. Kowalski. Écarts entre nombres premiers successifs (d'après Goldston, Pintz, Yıldırım, . . .). *Astérisque*, (311):Exp. No. 959, viii, 177–210, 2007. Séminaire Bourbaki. Vol. 2005/2006.

[55] E. Kowalski and P. Michel. Zeros of families of automorphic l-functions close to 1. *Pacific J. Math.*, 207(2):411–431, 2002.

[56] P. Kuhn. Zur Viggo Brun'schen Siebmethode. I. *Norske Vid. Selsk. Forhdl., Trondheim*, 14:145–148, 1941.

[57] P. Kuhn. Neue Abschätzungen auf Grund der Viggo Brunschen Siebmethode. *[A] 12. Skand. Mat.-Kongr., Lund 1953*, pages 160–168, 1954.

[58] Y.-K. Lau and Y.-F.S. Pétermann. Frequency of oscillations of an error term related to the Euler function. *Mathematika*, 47(1-2):161–164, 2000.

[59] P.G. Lejeune-Dirichlet. *Lectures on Number Theory, edited by R. Dedekind. Second edition. (Vorlesungen über Zahlentheorie, herausgegeben von R. Dedekind. Zweite Auflage.).* Braunschweig. Vieweg , 1871. Première édition en 1863.

[60] B.V. Levin and A.S. Fainleib. Application of some integral equations to problems of number theory. *Russian Math. Surveys*, 22:119–204, 1967.

[61] Yu.V. Linnik. The dispersion method in binary additive problems. *Leningrad*, page 208pp, 1961.

[62] G. Martin. An asymptotic formula for the number of smooth values of a polynomial. *J. Number Theory*, 93(2):108–182, 2002.

[63] H.L. Montgomery and R.C. Vaughan. The large sieve. *Mathematika*, 20(2):119–133, 1973.

[64] H.L. Montgomery and R.C. Vaughan. *Multiplicative Number Theory: I. Classical Theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2006.

[65] Y. Motohashi. Primes in arithmetic progressions. *Invent. Math.*, 44(2):163–178, 1978.

[66] Y. Motohashi. Sieve Methods and Prime Number Theory. *Tata Lectures Notes*, page 205, 1983.

[67] Y.-F.S. Pétermann. On an estimate of Walfisz and Saltykov for an error term related to the Euler function. *J. Théor. Nombres Bordx.*, 10(1):203–236, 1998.

[68] S.S. Pillai and S.D. Chowla. On the error terms in some asymptotic formulae in the theory of numbers. I. *Journal L. M. S.*, 5:95–101, 1930.

[69] J. W. Porter. Some numerical results in the Selberg sieve method. *Acta Arith.*, 20:417–421, 1972.

[70] O. Ramaré. On Snirel'man's constant. *Ann. Scu. Norm. Pisa*, 21:645–706, 1995. `http://math.univ-lille1.fr/~ramare/Maths/Article.pdf`.

[71] O. Ramaré. Le théorème de Brun-Titchmarsh : une approche moderne. pages 1–10, 2005. `http://math.univ-lille1.fr/\~{}ramare/Maths/Nantes.pdf`.

[72] O. Ramaré. *Arithmetical aspects of the large sieve inequality*, volume 1 of *Harish-Chandra Research Institute Lecture Notes*. Hindustan Book Agency, New Delhi, 2009. With the collaboration of D. S. Ramana.

[73] O. Ramaré. On long κ-tuples with few prime factors. *Proc. of the London Math. Soc.*, page 39pp, 2011.

[74] O. Ramaré and J.-C. Schlage-Puchta. Improving on the Brun-Titchmarsh theorem. *Acta Arith.*, 131(4):351–366, 2008.

[75] D.A. Rawsthorne. Selberg's sieve estimate with a one-sided hypothesis. *Acta Arith.*, 49:281–289, 1982.

[76] A. Rényi. Über die Darstellung der geraden Zahlen als Summe einer Primzahl und einer Fast-Primzahl. *Izv. Akad. Nauk SSSR, Ser. Mat.*, 12:57–78, 1948.

[77] A. Rényi. On the representation of an even number as the sum of a prime and of an almost prime. *Transl., Ser. 2, Am. Math. Soc.*, 19:299–321, 1962.

[78] B. Riemann. Üeber die Anzahl der Primzahlen unter einer gegebenen Grösse. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, 1859.

[79] S. Salerno. A note on Selberg sieve. *Acta Arith.*, 45(4):279–288, 1986.

[80] A. Schinzel. Remarks on the paper: Sur certaines hypothèses concernant les nombres premiers. *Acta Arith.*, 7(1):1–8, 1961.

[81] A. Schinzel and W. Sierpiǹski. Sur certaines hypothèses concernant les nombres premiers. *Acta Arith.*, 4(3):185–208, 1958.

[82] A. Selberg. On elementary problems in prime number-theory and their limitations. *C.R. Onzième Congrès Math. Scandinaves, Trondheim, Johan Grundt Tanums Forlag*, pages 13–22, 1949.

[83] A. Selberg. Remarks on multiplicative functions. *Lectures Notes in Mathematics (Berlin)*, 626:232–241, 1976.

[84] A. Selberg. Sifting problems, sifting density, and sieves. In D. Goldfeld" "K.E. Aubert, E. Bombieri, editor, *Number Theory, Trace Formulas and Discrete Groups*, pages 467–484, Oslo, 1987. Academic Press, San Diego London.

[85] A. Selberg. Collected papers. *Springer-Verlag*, II:251pp, 1991.

[86] H. Siebert. Montgomery's weighted sieve for dimension two. *Monatsh. Math.*, 82(4):327–336, 1976.

[87] G. Tenenbaum. *Introduction à la théorie analytique et probabiliste des nombres*, volume 1 of *Cours Spécialisés*. Société Mathématique de France, Paris, second edition, 1995.

[88] *PARI/GP, version* `2.4.3`. Bordeaux, 2008. `http://pari.math.u-bordeaux.fr/`.

[89] H. Tietze. *Gelöste und ungelöste mathematische Probleme aus alter und neuer Zeit*, volume Bd. 1, xx+256 pp. (10 plates); Bd. 2, iv+298 pp. (8 plates). Verlag C. H. Beck, München, 1959.

[90] J.E. van Lint and H.E. Richert. On primes in arithmetic progressions. *Acta Arith.*, 11:209–216, 1965.

[91] A.I. Vinogradov. The density hypothesis for Dirichet l-series. *Izv. Akad. Nauk SSSR Ser. Mat.*, 29, 1965.

[92] A. Walfisz. *Weylsche Exponentialsummen in der neueren Zahlentheorie.* , 1963.

[93] E. Wirsing. Das asymptotische Verhalten von Summen über multiplikative Funktionen. *Math. Ann.*, 143:75–102, 1961.

[94] Jie Wu. p_2 dans les petits intervalles. *Théorie des nombres, Smin., Paris/Fr. 1989-90, Prog. Math.*, 102:233–267, 1992.

[95] Jie Wu. Chen's double sieve, Goldbach's conjecture and the twin prime problem. *Acta Arith.*, 114(3):215–273, 2008.

[96] Sheng Gang Xie. On the k-twin primes problem. *Acta Math. Sinica*, 26(3):378–384, 1983.

[97] Sheng Gang Xie. The prime 4-tuplet problem. *Sichuan Daxue Xuebao*, 26(Special Issue):168–171, 1989.

Index

$G_d(Q)$, 45
$G_d(\mathcal{K}, Q)$, 45
Y, 65
$\mathbb{1}_{\mathcal{K}_d}$, 46
$\mathbb{1}_{\mathcal{L}_d}$, 46
$\mathcal{U}_q$, 44
$\beta_{\mathcal{K}}$, 50
$\mathrm{c}_q(n)$, 36
ψ_d^*, 67
$\sigma_{q\to d}$, 43
ζ_d, 71
$h(d)$, 45

Admissible shift, 40, 41
Apostol, T., 31

Bombieri, E., 39, 49, 54, 67
Bombieri-Vinogradov Theorem, 53, 54
Bordering system, 46, 47, 63
Brun, V., 49, 57

Cazaran, J., 23
Cazaran, J. & Moree, P., 23
Chen, Jingrun, 58
Chowla, S.D., 12
Compact set, 43
  Johnsen-Gallagher condition, 44
  multiplicatively split, 43, 47
  squarefree, 43

defalpha, 71
Diamond, H.G., 7, 59
Diamond, H.G. & Halberstam, H., 59
Diamond, H.G., Halberstam, H. & Galway, W.F., 7, 59
Dimension of the sieve, 24
Dusart P., 41

Elliott, P.D.T.A., 39
Erdös P., 41
Erdös P. & Riesel H., 41
Euler-MacLaurin summation formula, 10
Exchange of summation, 12

Fainleib, A.S., 24, 27, 45
Forbes T., 41
Franze, C., 59
Friedlander, J., 49
Friedlander, J. & Iwaniec, H., 49

Gallagher, P.X., 44, 49, 52
Galway, W.F., 7, 59
Goldston, D.A., 60
Goldston, D.A., Pintz, J. & Yıldırım, C.Y., 60
Greaves, G., 60

Hálasz, G., 36
Halberstam, H., 7, 23, 27, 45, 49, 58, 59
Halberstam, H. & Richert, H.E., 23, 27, 45, 49, 58
Hall R.R., 27
Hardy, G.H., 40
Hardy, G.H. & Littlewood, J.E., 40
Heath-Brown, R., 59
Hensley D., 40
Hensley D. & Richards I., 40
Ho, Kwan-Hung, 59
Ho, Kwan-Hung & Tsang, Kai-Man, 59
Host sequence, 49, 58, 61

Iwaniec, H., 24, 39, 49, 58
Iwaniec, H. & Kowalski, E., 39
Iwaniec, H. & Laborde, M., 58

Johnsen-Gallagher condition, 44
Jutila, M., 39

Kan, Jia Hai, 60
Kobayashi, I., 67
Kowalski, E., 39, 60
Kuhn P., 58

Laborde, M., 58
Lau, Y.-K., 12
Leikauf P., 41
Length of a κ-tuple of admissible shifts, 40
Levin, B.V., 24, 27, 45
Levin, B.V. & Fainleib, A.S., 24, 27, 45
Linnik, Yu V., 19, 49

Littlewood, J.E., 40

Martin, G., 23
Montgomery, H.L., 19, 39
Montgomery, H.L. & Vaughan, R.C., 19, 39
Moree, P., 23
Motohashi, Y., 52, 67
Multiplicatively split, 43, 47

Open Problem no I, 39
Open Problem no II, 39
Open Problem no III, 85
Open Problem no IV, 85

Pétermann, Y.-F.S., 12
Pillai, S.S., 12
Pintz, J., 60
de Polignac A., 57
Porter, J.W., 59
Prime twin, 57
Prime κ-tuples conjecture, 40
Pseudo-character, 67

Ramanujan sum, 36
Rawsthorne, D.A., 24
Rényi, A., 58
Richards I., 40
Richert, H.E., 23, 27, 45, 49, 58
Richert, H.E., 15
Riemann B., 13
Riesel H., 41

Salerno, S., 59
Schlage-Puchta, J.-C., 39
Selberg, A., 19, 39, 44, 49, 52
Siebert, H., 53
Sieve coefficients, 61
Sieve dimension, 24
Squarefree kernel, 17, 44, 111
Summation by parts, 10, 77

Tiezte H., 57
Tsang, Kai-Man, 59

Uniformity in error terms, 9

van Lint, J.E., 15
van Lint J.E. & Richert H.E., 45
Vaughan, R.C., 19, 39
Vinogradov, A.I., 54

Waldvogel J., 41
Waldvogel J. & Leikauf P., 41
Walfisz, A., 12
Wirsing E., 28
Wu, Jie, 53, 58

Xie, Sheng Gang, 59

Yıldırım, C.Y., 60

IMSc Lecture Notes Series:

[1] *Giovanni Corvaja:* Integral points on varieties
[2] *Ram Murty:* Special Values of Zeta and L-functions
[3] *Rob Tubbs:* Hilbert's Seventh Problem: Solutions and Extensions
[4] *Carlo Gasbarri:* Arithmetic of number fields, the geometry of algebraic curves and Nevanlinna theory
[5] *Kannan Soundararajan:* A probabilistic model for L-functions
[6] *Winfried Kohnen:* L-functions and generalized modular forms
[7] *Patrice Philippon:* Some aspects of Mahler's method in transcendence number theory
[8] *Michel Waldschmidt:* Multiple zeta values
[9] *Yuri Nesterenko:* Transcendental Number theory: Elimination theory, linear independence and algebraic independence, irrationality of zeta values
[10] *Jean-Marc Deshouillers:* Introduction to automatic sequences

In preparation:

[11] *Antal Balog:* Additive combinatorics
[12] *Olivier Ramaré:* Long chains of integers with few prime factors
[13] *Sinnou David:* Lehmer's Problem and Baker's Theory
[14] *Kumar Murty:* Families of L-functions
[15] *Next Generation:* Modern Mathematics

More informations concerning ordering, publishing instructions as well as videos related to these lectures notes may be found at
`http://IMSC...`

This is what will appear at the back of the book.