# Université de Montréal


# La distribution des zéros des fonctions $L$

par


## Antoine Comeau-Lapointe

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


Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en mathématiques

24 août 2016

# Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé

## La distribution des zéros des fonctions $L$

présenté par

## Antoine Comeau-Lapointe

a été évalué par un jury composé des personnes suivantes :

*Matilde Lalin*

(président-rapporteur)

*Dimitris Koukoulopoulos*

(directeur de recherche)

*Abraham Broer*

(membre du jury)

Mémoire accepté le

*9 août 2016*

# SOMMAIRE

Selon la philosophie de Katz et Sarnak, la distribution des zéros des fonctions $L$ est prédite par le comportement des valeurs propres de matrices aléatoires. En particulier, le comportement des zéros près du point central révèle le type de symétrie de la famille de fonctions $L$. Une fois la symétrie identifiée, la philosophie de Katz et Sarnak conjecture que plusieurs statistiques associées aux zéros seront modélisées par les valeurs propres de matrices aléatoires du groupe correspondant. Ce mémoire étudiera la distribution des zéros près du point central de la famille des courbes elliptiques sur $\mathbb{Q}[i]$. Brumer a effectué ces calculs en 1992 sur la famille de courbes elliptiques sur $\mathbb{Q}$. Les nouvelles problématiques reliées à la généralisation de ses travaux vers un corps de nombres seront mises en évidence.

**Mots-clés : Fonctions $L$, Zéros près du point central, Philosophie de Katz et Sarnak, Théorie des matrices aléatoires, Courbes elliptiques**

# SUMMARY

The Katz and Sarnak philosophy states that the distribution laws of zeros of $L$-functions follow the distribution laws of eigenvalues of random matrices. The zeros near the central point would reveal the symmetry type of our family of $L$-functions. Once the symmetry has been identified, it is conjectured that many statistics associated to the zeros would be predicted by the eigenvalues of the corresponding group of random matrices. This thesis will study the low-lying zeros of the family of elliptic curves over $\mathbb{Q}[i]$. Brumer computed the symmetry type of the family of elliptic curves over $\mathbb{Q}$ in 1992. New challenges arising from this generalisation over number fields of his work will be revealed in this thesis.

**Keywords : $L$-functions, Low-lying zeros, Katz and Sarnak philosophy, Random matrix theory, Elliptic curves**

# La distribution des zéros des fonctions $L$

Antoine Comeau-Lapointe

## Table des matières

**1 Introduction** **4**
1.1 Résumé . . . 4
1.2 Structure du mémoire . . . 6
1.3 Remerciements . . . 7

**2 Prérequis** **8**
2.1 Notations conventionnelles . . . 8
2.2 Fonctions arithmétiques . . . 8
2.3 Fonctions sur les réels . . . 8
2.4 Notation asymptotique . . . 9
2.5 Caractères de Dirichlet . . . 10
2.6 Géométrie algébrique . . . 13
2.6.1 Le corps de fonctions d'une variété . . . 14
2.6.2 Points singuliers et réguliers . . . 15
2.6.3 Modèles affines et projectifs . . . 15
2.6.4 Morphismes de variétés . . . 16

**3 Fonction zêta de Riemann** **17**
3.1 Produit d'Euler . . . 17
3.2 Prolongement analytique . . . 17
3.2.1 Sommation de Poisson . . . 17
3.2.2 Équation fonctionnelle . . . 18
3.3 Région sans zéro . . . 20
3.4 Théorème des nombres premiers . . . 23

**4 Fonctions $L$** **28**
4.1 Définitions . . . 28
4.2 Les zéros des fonctions $L$ . . . 31
4.3 Région sans zéro . . . 36
4.4 Formule explicite . . . 38
4.5 Théorème des nombres premiers . . . 40

**5 Théorie des matrices aléatoires** **43**
5.1 Théorèmes de convergence et d'universalité . . . 43
5.2 Mesure de Haar . . . 44
5.3 Continuité des valeurs propres . . . 46
5.4 La distribution des zéros des fonctions $L$ . . . 47

**6 Courbes elliptiques** **52**
6.1 Équation de Weierstrass . . . . . 52
6.2 Le groupe des points rationnels . . . . . 53
6.3 La fonction $L$ de Hasse-Weil . . . . . 55

**7 La symétrie de la famille des courbes elliptiques sur $\mathbb{Q}$** **58**
7.1 Introduction . . . . . 58
7.2 Calcul du type de symétrie . . . . . 59
7.3 Restriction de minimalité . . . . . 70
7.4 Application . . . . . 71

**8 La symétrie de la famille des courbes elliptiques sur $\mathbb{Q}(i)$** **73**
8.1 Le corps $\mathbb{Q}(i)$ . . . . . 73
8.2 La fonction $L$ de Hasse-Weil pour $E/\mathbb{Q}(i)$ . . . . . 74
8.3 Sommation de Poisson . . . . . 74
8.4 Condition du conducteur . . . . . 76
8.5 L'estimation centrale . . . . . 78
8.6 Directions futures de recherche . . . . . 80

**Références** **82**

# 1 Introduction

## 1.1 Résumé

Malgré la simplicité des nombres naturels, les techniques utilisées afin d'obtenir des informations sur leur structure sont d'une grande complexité. Par exemple, comment devrions-nous nous y prendre si nous désirons avoir une idée de la quantité de nombres premiers plus petits que $x$? Nous savons maintenant que cette quantité est approximativement $x/\log x$. Ce résultat, appelé le *théorème des nombres premiers*, fut formellement prouvé en 1896 par de la Vallée-Poussin et Hadamard indépendamment. La preuve est fondamentalement basée sur l'analyse complexe, ce qui est surprenant lorsque l'on considère le caractère élémentaire des nombres premiers.

La fonction complexe nous révélant le comportement des nombres premiers est donnée par la formule suivante pour $s \in \mathbb{C}$

$$\zeta(s) := \sum_{n\geq 1} \frac{1}{n^s}$$

convergente absolument dans la région $\mathrm{Re}(s) > 1$. Il est possible d'obtenir des informations sur les nombres premiers par la forme spécifique de cette formule. Le théorème des nombres premiers découle d'une intégration complexe effectuée sur $\zeta'/\zeta$ en utilisant le théorème de Cauchy. Les zéros de $\zeta$ sont pour cette raison intimement liée à la distribution des nombres premiers. Il est par contre nécessaire que $\zeta$ soit définie sur tout le plan complexe. Par les travaux de Riemann en 1859, $\zeta$ possède un prolongement méromorphe sur tout le plan complexe, avec un pôle simple en $s = 1$. Il y a une infinité de zéros et l'hypothèse de Riemann dit qu'ils sont tous situés sur la ligne $\mathrm{Im}(s) = 1/2$.

En utilisant les mêmes idées, des informations sur divers autres objets mathématiques peuvent ainsi être obtenues. On considère alors les formules

$$L(s, f) := \sum_{n\geq 1} \frac{\lambda_f(n)}{n^s}$$

appelées *fonctions L* pour $\{\lambda_f(n)\}_{n\geq 1}$ une fonction complexe multiplicative ($\lambda_f(mn) = \lambda_f(m)\lambda_f(n)$ si $(m, n) = 1$) associée à un certain objet $f$ arbitraire. Les objets peuvent par exemple être une extension des rationnels, une variété algébrique ou une forme modulaire. Les informations sur le comportement de $\sum_{p\leq x} \lambda_f(p)$ seront encore une fois contenues dans la position de ses zéros. L'hypothèse de Riemann se généralise pour toutes les fonctions $L$, on l'appelle *la grande hypothèse de Riemann.*

En 1973, Montgomery émit une conjecture sur la corrélation entre les paires de zéros de $\zeta$. Cette corrélation correspond essentiellement à observer le comportement moyen de la distance entre deux zéros de $\zeta$. En discutant de sa recherche avec Freeman Dyson, ce dernier a remarqué que la fonction de densité trouvée par Montgomery était la même que celle pour la corrélation entre les paires de valeurs propres de matrices aléatoires. Cette relation conjecturale est maintenant généralisée à toutes les fonctions $L$.

Il existe plusieurs types de matrices aléatoires. En considérant plutôt une *famille de fonctions L*, c'est-à-dire un ensemble de fonctions $L$ similaires, plusieurs statistiques associées à leurs zéros semblent être modélisées par les statistiques des valeurs propres d'un certain type de matrices aléatoires. Plusieurs expériences numériques semblent confirmer cette conjecture.

On appelle ce phénomène, *la philosophie de Katz et Sarnak*. On appelle *le type de symétrie* le groupe de matrices aléatoires associé à la famille.

Les courbes elliptiques sont les courbes algébriques les plus simples encore aujourd'hui partiellement comprises. Une courbe elliptique $E/\mathbb{Q}$ est définie comme étant le sous-ensemble de $\mathbb{C}^2$ solution de l'équation $y^2 = x^3 + ax + b$, où $a, b \in \mathbb{Z}$. Il est possible de réduire ces coefficients modulo $p$, où $p$ est premier, et de compter le nombre de solutions $(x, y)$ à l'intérieur de $(\mathbb{Z}/p\mathbb{Z})^2$. Dénotons cette quantité par $N_p$ et posons $a_p = p + 1 - N_p$. Une fonction $L$, appelée *de Hasse-Weil*, est construite à partir de ces informations pour tout $p$ ainsi

$$L(s, E) := \prod_p L_p(s, E)^{-1}$$

où

$$L_p(s, E) := 1 - a_p p^{-s} + p^{1-2s}$$

pour les $p$ tels que la réduction de la courbe modulo $p$ se comporte bien (la courbe réduite n'est pas singulière). La définition de $L_p(s, E)$ pour les autres $p$ est similaire. Il y aura un nombre fini de tels $p$ problématiques pour chaque courbe. La conjecture de Birch et Swinnerton-Dyer établit un lien entre l'ordre du zéro à $s = 1$, appelé *rang analytique*, et le rang algébrique du groupe des points rationnels de la courbe. Ces deux quantités seraient égales par cette conjecture. Pour les courbes elliptiques ayant $a, b \in \mathcal{O}_k$, les entiers de $k$ une extension des rationnels, la fonction $L$ de Hasse-Weil et la conjecture de Birch et Swinnerton-Dyer se généralisent. Les travaux de ce mémoire s'effectueront sur $k = \mathbb{Q}(i)$.

Brumer a tenté de calculer en 1992 le type de symétrie de la famille des courbes elliptiques rationnelles par la distribution des zéros proches du point central sous la grande hypothèse de Riemann. Le résultat est conditionnel à l'étendue du support de la transformée de Fourier d'une fonction test. Sans cette condition, la symétrie de la famille est orthogonale. Cette fonction test sert à capturer les premiers zéros de la fonction $L$. Plus le support de la transformée de Fourier est large, plus la fonction test est localisée à l'origine. Le zéro central est d'intérêt par la conjecture de Birch et Swinnerton-Dyer, ce qui explique notre désir de localiser la fonction test à l'origine. En 2005, Young a augmenté le support permis, toujours sous la grande hypothèse de Riemann. Une application de pouvoir utiliser un support plus large est qu'il est possible de réduire la borne sur le rang analytique moyen de la famille. Les travaux de Young impliquent que cette borne serait $\leq 25/14$. Seul un support permis plus large que $(-1, 1)$ serait en mesure de distinguer la symétrie orthogonale de la famille. Le support permis par Brumer était $(-5/9, 5/9)$ et Young l'a augmenté à $(-7/9, 7/9)$ en analysant en détail une certaine somme au lieu de l'estimer trivialement.

Précisément, le résultat de Brumer est le suivant. Chaque courbe de notre famille est associée à une paire d'entiers $(a, b)$ qui sont les coefficients du polynôme de la courbe. On utilise une fonction lisse $\omega(x, y)$ à support compact afin de sélectionner les courbes de notre famille ayant $|a| \leq X^{1/3}$ et $|b| \leq X^{1/2}$. Lorsque $X$ tend vers l'infini, toutes les courbes seront alors considérées. Le sens de l'exposant est pour contrôler une certaine quantité reliée aux courbes elliptiques. Pour capturer ces courbes par le biais de cette fonction à support compact, on définit une fonction sur les éléments de notre famille

$$\omega_X(E_{a,b}) := \omega\left(\frac{a}{X^{1/3}}, \frac{b}{X^{1/2}}\right)$$

où $E_{a,b}$ est la courbe elliptique $y^2 = x^3 + ax + b$. On définit la quantité

$$W_X(\mathcal{F}) := \sum_{E \in \mathcal{F}} \omega_X(E)$$

qui compte essentiellement le nombre de courbes considérées par rapport à $X$. On définit ensuite pour une courbe fixée une quantité reliée aux premiers zéros de sa fonction $L$ :

$$D(E, \phi) := \sum_{\gamma_E} \phi\left(\frac{\gamma_E}{2\pi} \log X\right)$$

où $\gamma_E$ est la partie imaginaire des zéros de $L(s, E)$. Les facteurs à l'intérieur ne servent qu'à renormaliser leur position. La fonction $\phi$ doit être lisse et de décroissance super-exponentielle, ce qui explique pourquoi cette somme ne considère que les premiers zéros. On somme cette fonction sur notre famille tout en gardant un contrôle sur les courbes choisies par $\omega_X$.

$$\mathcal{D}(\omega_X, \mathcal{F}, \phi) := \sum_{E \in \mathcal{F}} D(E, \phi)\omega_X(E)$$

En divisant cette quantité par $W_X(\mathcal{F})$, on obtient une densité moyenne du positionnement des premiers zéros des fonctions $L$ de notre famille. Le résultat de Brumer est

$$\lim_{X \to \infty} \frac{\mathcal{D}(\omega_X, \mathcal{F}, \phi)}{W_X(\mathcal{F})} = \hat{\phi}(0) + \frac{1}{2}\phi(0) \quad \text{lorsque } X \to \infty$$

si le support de $\hat{\phi}$, la transformée de Fourier de $\phi$, est contenu dans $(-5/9, 5/9)$. La quantité vers laquelle cette limite converge correspond à la quantité que l'on retrouve en effectuant les mêmes calculs sur les positions des valeurs propres de matrices orthogonales.

Finalement, les travaux de ce mémoire prouvent que ce résultat est encore valide sur la famille de courbes elliptiques sur $\mathbb{Q}(i)$. Nous allons voir que les travaux de Brumer s'ajustent assez aisément sans compromettre l'étendue du support. Il s'agit d'un premier pas vers une généralisation pour les autres extensions des rationnels. Beaucoup est révélé sur les nouvelles problématiques introduites par l'utilisation de $\mathbb{Q}(i)$. Le support permis est encore $(-5/9, 5/9)$ et la borne sur le rang analytique moyen sera dans notre cas $\leq 23/10$.

## 1.2 Structure du mémoire

Après les prérequis, un traitement classique de la fonction zêta de Riemann sera fait afin de se préparer à la généralité du Chapitre 4. La preuve du prolongement analytique trouvée par Riemann et la preuve de théorème des nombres premiers seront données.

Le Chapitre 4 détaillera les théorèmes pouvant être obtenus sur les fonctions $L$ d'un point de vue général. La formule explicite reliant une somme sur les zéros à une somme sur les nombres premiers est d'ailleurs utilisée comme point de départ afin d'effectuer les calculs des Chapitres 7 et 8 sur les zéros proches du point central.

La théorie des matrices aléatoires sera traitée au Chapitre 5. Le sens précis de la similarité entre la distribution des valeurs propres et la distribution des zéros des fonctions $L$ sera éclairci. La dernière section de ce chapitre discutera de fonctions analogues aux fonctions $L$ où le phénomène est également présent. Par contre, la source de ce phénomène est comprise pour

ces fonctions.

Ensuite, le Chapitre 6 parlera des courbes elliptiques et de la construction de leur fonction $L$ associée.

Le Chapitre 7 rassemblera la matière de tous les chapitres précédents, qui sont pourtant de branches distinctes des mathématiques. Le type de symétrie de la famille des courbes elliptiques sur $\mathbb{Q}$ sera calculé à partir des zéros proches du point central.

Le Chapitre 8 présentera la contribution de ce mémoire sur la famille des courbes elliptiques sur $\mathbb{Q}(i)$ en appliquant les mêmes stratégies que Brumer.

## 1.3 Remerciements

Je remercie le Fonds de recherche du Québec - Nature et technologies (FRQNT) pour le financement de ce mémoire. Cette bourse m'a permis de me concentrer sur ce projet à plein temps et poursuivre mes passions.

Je remercie également mon directeur Dimitris Koukoulopoulos de sa grande aide, de m'avoir guidé à travers cette voie et de m'avoir proposé cet intéressant projet de recherche.

Finalement, je remercie ma famille, en particulier Hélène et François, et ma copine Geneviève pour leur soutien durant la rédaction.

# 2 Prérequis

## 2.1 Notations conventionnelles

Les notations suivantes seront librement utilisées à travers le mémoire. Nous dénotons la partie réelle et imaginaire d'un nombre complexe par $s = \sigma + it$. Les zéros d'une fonction $L$ s'écrivent $\rho = \beta + i\gamma$ spécialement. La lettre $p$ désigne toujours un nombre premier. Finalement, on définit la fonction $e(x) := e^{2\pi ix}$.

## 2.2 Fonctions arithmétiques

L'unicité de la factorisation en nombres premiers d'un nombre naturel entraîne plusieurs questions intéressantes. Par exemple, combien de facteurs premiers distincts est-ce que $n$ possède? Il est naturel de considérer la fonction qui associe chaque entier positif à son nombre de facteurs premiers distincts. Cette fonction est un exemple d'une fonction arithmétique, c'est-à-dire, une fonction définie sur les naturels vers les complexes. Voici une liste de fonctions arithmétiques usuelles, où $n = p_1^{a_1} p_2^{a_2} \dots p_k^{a_k}$ est la factorisation en nombres premiers de $n$:

- $\omega(n) = k$, la fonction qui compte le nombre de nombres premiers distincts qui divisent $n$.
- $\Omega(n) = a_1 + a_2 + \cdots + a_k$, la fonction qui compte le nombre total de nombres premiers qui composent $n$.
- $\mu(n) = \begin{cases} (-1)^k & \text{si } a_i = 1 \, \forall \, i \in \{1, \dots, k\}, \text{ c'est-à-dire, } n \text{ est sans carré} \\ 0 & \text{sinon,} \end{cases}$
  la fonction de Möbius.
- $\sigma_s(n) = \sum_{d|n} d^s$, la somme des diviseurs de $n$ où chaque diviseur est mis à la puissance $s$ où $s$ est un nombre complexe. On dénote $\sigma_1(n)$ par $\sigma(n)$.
- $\tau_l(n) = \#\{(d_1, \dots, d_l) \in \mathbb{N}^l \text{ tel que } n = d_1 \dots d_l\}$, le nombre de façons de séparer $n$ en $l$ facteurs. Remarquons que $\sigma_0 = \tau_2$.
- $d(n) = \sum_{d|n} 1$, le nombre de diviseurs de $n$. On remarque que $d(n) = \sigma_0(n) = \tau_1(n)$.
- $\varphi(n) = \#\{1 \leq a < n : (a, n) = 1\}$, la fonction indicatrice d'Euler qui compte combien de nombres plus petits que $n$ sont copremiers à $n$. La notation $(a, b)$ désigne le plus grand commun diviseur de $a$ et $b$.
- $\Lambda(n) = \begin{cases} \log(p_1) & \text{si } k = 1, \\ 0 & \text{sinon,} \end{cases}$
  la fonction von Mangoldt qui vaut $\log(p)$ aux puissances de $p$ où $p$ est premier.

## 2.3 Fonctions sur les réels

Il est possible de définir des fonctions ayant comme domaine les réels en sommant les valeurs d'une fonction arithmétique évaluées aux $n \leq x$. Bien sûr, ces fonctions seront constantes entre deux entiers et auront leurs sauts qu'aux entiers. Pour le reste de ce mémoire, $p$ signifie un nombre premier.

- $\pi(x) = \#\{p \leq x\}$, la fonction qui compte le nombre de nombres premiers $\leq x$.

- $\psi(x) = \sum_{n \leq x} \Lambda(n)$ la deuxième fonction de Chebyshev.
- $\theta(x) = \sum_{p \leq x} \log(p)$.

Dû au comportement irrégulier de la factorisation en nombres premiers des nombres naturels lorsqu'on les observe en ordre (1,2,3,4,...), ces fonctions sont inévitablement irrégulières. Afin de les étudier, une notation cruciale à maîtriser en théorie analytique des nombres est la notation asymptotique.

## 2.4 Notation asymptotique

Étant donnée la nature irrégulière de certaines fonctions, il est difficile de manipuler arithmétiquement celles-ci. On abandonne donc une certaine précision afin d'obtenir des formules possédant de meilleures qualités arithmétiques tout en gardant un certain contrôle de l'erreur. On définit les notations principalement utilisées ainsi:

- $f(x) = \mathcal{O}(g(x))$ si $\exists\, c > 0\, \exists\, X > 0$ tel que $|f(x)| \leq c|g(x)|\, \forall\, x > X$
  (notation équivalente: $f(x) \ll g(x)$)
- $f(x) \asymp g(x)$ si $f(x) = \mathcal{O}(g(x))$ et $g(x) = \mathcal{O}(f(x))$
- $f(x) \sim g(x)$ si $\lim_{x\to\infty} f(x)/g(x) = 1$

Pour ce qui est de la première notation, appelée «grand O», on remarque que le symbole d'égalité a une signification spéciale. On utilise surtout la notation $\mathcal{O}(f(x))$ pour représenter un terme d'erreur. La notation $\ll$ est utilisée si l'on désire simplement borner les quantités.

Essentiellement, $\mathcal{O}(g(x))$ représente un ensemble de fonctions. Cet ensemble est composé de toutes les fonctions ayant une croissance plus petite ou égale à $g(x)$. Une équation sous la forme suivante,

$$f(x) = h(x) + \mathcal{O}(g(x))$$

signifie qu'il existe une certaine fonction $E(x)$ élément de l'ensemble des fonctions bornées en croissance par $g(x)$ telle que l'égalité est vraie, c'est-à-dire, $f(x) = h(x) + E(x)$. En général, la fonction $E(x)$ est une fonction qui capture les irrégularités de $f$. Il est suffisant de n'avoir qu'une borne sur sa croissance puisque pour de très grand $x$, l'erreur devient négligeable par rapport au terme principal. Par exemple, on peut simplifier la fonction plancher $\lfloor x \rfloor = x + \mathcal{O}(1)$ qui devient beaucoup plus facile à manipuler.

Pour donner une meilleure idée de ce qu'on entend par croissance et du rôle de la constante $c$ dans la définition formelle, la fonction $f(x) = x^2$, par exemple, n'est pas en $\mathcal{O}(x)$ puisque peu importe par quelle constante $c$ on multiplie $x$, la fonction $x^2$ sera plus grande que $cx$ pour les $x > c$.

La deuxième notation signifie que $f$ est proportionnelle à $g$, ce qui veut dire que les deux fonctions ont la même croissance et aucune des deux ne domine l'autre. On aura donc, si la limite existe, $\lim_{x\to\infty} f(x)/g(x) = a$ où $0 < a < \infty$. La troisième notation est le cas précis où $a = 1$.

## 2.5 Caractères de Dirichlet

Pour $q \in \mathbb{N}\backslash\{0\}$, le groupe $\mathbb{Z}/q\mathbb{Z}$ est présent dans $\mathbb{C}^*$ sous l'isomorphisme suivant: $\psi(n \bmod q) = \mathrm{e}(n/q)$. En fait, il existe $q$ homomorphismes de $\mathbb{Z}/q\mathbb{Z}$ vers le cercle unité complexe donnés par

$$\psi_a(n) = \mathrm{e}\left(\frac{an}{q}\right)$$

pour $0 \leq a \leq q-1$. On appelle ces fonctions sur $\mathbb{Z}/q\mathbb{Z}$ des caractères additifs.

On obtient les caractères de Dirichlet modulo $q$ en considérant les homomorphismes de $(\mathbb{Z}/q\mathbb{Z})^*$ vers le cercle unité complexe. On sait par la théorie des représentations qu'il y aura $|(\mathbb{Z}/q\mathbb{Z})^*| = \varphi(q)$ homomorphismes possibles. Puisqu'on souhaite obtenir une fonction arithmétique $\chi(n)$, on étend le domaine d'un homomorphisme $\tilde{\chi}(n)$ de $(\mathbb{Z}/q\mathbb{Z})^*$ à $\mathbb{Z}$ ainsi:

- $\chi(n) = \tilde{\chi}(n \bmod q)$ pour tout $(n, q) = 1$,
- $\chi(n) = 0$ si $(n, q) \neq 1$

alors $\chi(n)$ est périodique mod $q$. On appelle la fonction arithmétique $\chi(n)$ ainsi obtenue un caractère de Dirichlet modulo $q$. Chaque modulo $q$ possède un caractère nommé principal provenant de l'homomorphisme trivial $\tilde{\chi}(n) = 1$. Les propriétés importantes des caractères de Dirichlet sont la complète multiplicativité $\chi(mn) = \chi(m)\chi(n)$ pour tout $m, n \in \mathbb{Z}$ et l'orthogonalité des caractères entre eux qui nous servira à sélectionner dans une somme seulement les termes où $n$ est congrus à un certain $a$ mod $q$ pour $(a, q) = 1$. Expliquons maintenant cette propriété dans un contexte plus général.

On dénote par $G$ un groupe abélien fini et $\hat{G}$ ses $|G|$ caractères. En définissant la multiplication sur les caractères composant $\hat{G}$ ainsi $(\chi_1\chi_2)(g) = \chi_1(g)\chi_2(g)$ pour tout $g \in G$, on obtient une structure de groupe sur $\hat{G}$ où le caractère trivial est l'identité. On exprime ainsi les relations d'orthogonalité pour tout $\chi \in \hat{G}$,

$$\frac{1}{|G|}\sum_{g\in G}\chi(g) = \begin{cases} 1 & \text{si } \chi = \chi_0 \text{ le caractère trivial,} \\ 0 & \text{sinon.} \end{cases}$$

Également, pour tout $g \in G$

$$\frac{1}{|G|}\sum_{\chi\in\hat{G}}\chi(g) = \begin{cases} 1 & \text{si } g = 1, \\ 0 & \text{sinon.} \end{cases}$$

On prouve la première égalité ainsi. Si $\chi = \chi_0$, l'égalité est évidente. Sinon, on se rappelle qu'un groupe agit sur lui-même, chaque élément $h \in G$ engendre une permutation des éléments de $G$ en multipliant chaque élément par $h$. C'est ce qui justifie la dernière égalité

$$\chi(h)\sum_{g\in G}\chi(g) = \sum_{g\in G}\chi(hg) = \sum_{g\in G}\chi(g).$$

Il existe un $h \in G$ tel que $\chi(h) \neq 1$ sans quoi on aurait $\chi = \chi_0$. La somme doit donc égaler 0 pour que l'égalité soit vraie.

Pour la deuxième relation d'orthogonalité, on a encore que si $g = 1$ l'égalité est triviale. Sinon, on utilise exactement le même argument que pour la première relation

$$\chi_1(g) \sum_{\chi \in \hat{G}} \chi(g) = \sum_{\chi \in \hat{G}} \chi_1 \chi(g) = \sum_{\chi \in \hat{G}} \chi(g).$$

Il suffit maintenant de montrer qu'il existe un $\chi_1 \in \hat{G}$ tel que $\chi_1(g) \neq 1$ pour $g \neq 1$. On utilise la décomposition en groupes cycliques d'un groupe abélien fini. Un caractère $\chi$ sur le produit direct de groupes $G_A \times G_B$ peut être représenté comme la multiplication d'un caractère $\chi_A$ sur $G_A$ et un $\chi_B$ sur $G_B$, c'est-à-dire, $\chi((g_A, g_B)) = \chi_A(g_A)\chi_B(g_B)$. Donc, comme la projection de notre $g$ ci-haut ne vaut pas 1 sur tous les groupes cycliques du produit, on peut choisir les caractères triviaux pour tous les groupes cycliques sauf un, où la projection de $g$ n'est pas 1, nous permettant d'y choisir un caractère ne valant pas 1. Le caractère résultant a donc notre propriété désirée.

La méthode pour capturer les $n$ congrus à $a$ mod $q$ pour $(a, q) = 1$ est de considérer la somme suivante:

$$\frac{1}{\varphi(q)} \sum_{\chi \pmod q} \chi(na^{-1})$$

où la somme est sur tous les caractères de Dirichlet modulo $q$. Cette somme vaut 1 si $n \equiv a \pmod q$ et 0 sinon par la deuxième relation d'orthogonalité. En insérant cette somme dans une somme, on sommera seulement sur les $n$ congrus à $a$ mod $q$. On remarque que $\chi(a^{-1}) = \chi(a)^{-1} = \overline{\chi(a)} =: \overline{\chi}(a)$. Alors

$$\sum_{\substack{n \leq x \\ n \equiv a \pmod q}} f(n) = \frac{1}{\varphi(q)} \sum_{\chi \pmod q} \overline{\chi}(a) \sum_{n \leq x} f(n)\chi(n). \tag{2.1}$$

Les caractères de Dirichlet ne sont pas tous authentiques. Certains proviennent d'un plus petit modulo $q'$ facteur de $q$. On dit qu'un caractère $\chi'$ modulo $q'$ induit un caractère $\chi$ modulo $q$ pour $q'|q$ si

$$\chi(a) = \begin{cases} \chi'(a) & \text{si } (a, q) = 1 \\ 0 & \text{sinon.} \end{cases} = \chi'(a) \cdot \chi_0(a)$$

où $\chi_0$ est le caractère trivial modulo $q$. Alors $\chi$ est $\chi'$, mais avec davantage de 0. On appelle *conducteur* du caractère le plus petit $q'|q$ tel qu'il existe un caractère $\chi'$ modulo $q'$ qui induit $\chi$. Si le conducteur est égal au modulo $q$, le caractère ne provient pas d'un plus petit modulo et on dit que le caractère est *primitif*.

On note ici que le symbole de Legendre est un caractère réel de Dirichlet modulo un nombre premier $p$ impair

$$\chi(n) = \left(\frac{n}{p}\right) = \begin{cases} 1 & \text{si } n \text{ est un carré mod } p, \\ -1 & \text{si } n \text{ n'est pas un carré mod } p, \\ 0 & \text{si } p|n. \end{cases}$$

Il est possible de généraliser afin d'avoir des caractères réels modulo tous les entiers en utilisant le symbole de Kronecker.

Pour terminer cette section, on étudiera les sommes de Gauss. Pour tout caractère de Dirichlet mod $q$, on définit sa somme de Gauss ainsi

$$\tau(\chi) = \sum_{a \pmod q} \chi(a) e\Big(\frac{a}{q}\Big).$$

La relation importante pour cette somme est la suivante

$$\chi(n) = \frac{1}{\tau(\overline{\chi})} \sum_{1 \le a \le q} \overline{\chi}(a) e\Big(\frac{an}{q}\Big) \qquad \text{pour } (n,q) = 1. \tag{2.2}$$

En fait, si $\chi$ est primitif, cette égalité est valide pour tout $n$ car les deux côtés vaudront zéro quand $(n,q) \neq 1$ comme on le montrera. Cette égalité peut être vue comme une relation entre $\chi$ et sa transformée de Fourier discrète. Démontrons maintenant cette relation où l'on dénote $\overline{n}$ pour l'inverse de $n$ mod $q$.

$$\sum_{1 \le a \le q} \overline{\chi}(a) e\Big(\frac{an}{q}\Big) = \sum_{1 \le a \le q} \overline{\chi}(a\overline{n}) e\Big(\frac{a\overline{n}n}{q}\Big) = \overline{\chi}(\overline{n})\tau(\overline{\chi}) = \chi(n)\tau(\overline{\chi}).$$

La première égalité est due à la permutation des éléments de $\mathbb{Z}/q\mathbb{Z}$ lorsqu'on les multiplie par un $n$ inversible, le changement $a \to a\overline{n}$ ne fait donc que permuter les termes de la somme. La dernière est due à $\chi(n)\chi(\overline{n}) = \chi(n\overline{n}) = \chi(1) = 1$. Dans le cas où $\chi$ est primitif, montrons que cette égalité est vraie aussi si $(n,q) > 1$.

Posons $1 < d = (n,q)$. Alors on peut écrire $n = n_1 d$ et $q = q_1 d$. En mettant en évidence la congruence de $a$ mod $q_1$ dans la somme, on obtient

$$\sum_{1 \le a \le q} \overline{\chi}(a) e\Big(\frac{an}{q}\Big) = \sum_{b=1}^{q_1} \sum_{j=1}^{d} \overline{\chi}(b + jq_1) e\Big(\frac{(b+jq_1)n_1}{q_1}\Big) = \sum_{b=1}^{q_1} e\Big(\frac{bn_1}{q_1}\Big) \sum_{j=1}^{d} \overline{\chi}(b+jq_1).$$

On va montrer que la somme sur $\overline{\chi}$ vaut zéro, ce qui termine la preuve car $\tau(\overline{\chi})\chi(n)$ vaut aussi zéro pour $(n,q) > 1$.

Afin de montrer que la somme sur $\chi$ vaut zéro, un fait est nécessaire à démontrer d'abord. Si un caractère de Dirichlet mod $q$ respecte, pour $q_1 | q$, $\chi(n) = 1 \, \forall \, n \equiv 1 \bmod q_1$ et $(n,q) = 1$, alors le caractère n'est pas primitif si $q_1 < q$.

Pour $n$ tel que $(n,q) = 1$, $l$ tel que $(n + lq_1, q) = 1$ et $\overline{n}$ l'inverse de $n$ mod $q$, on a

$$\chi(n) = \chi(n)\chi(1 + l\overline{n}q_1) = \chi(n + lq_1),$$

car $1 + l\overline{n}q_1 \equiv 1 \bmod q_1$ et $(1 + l\overline{n}q_1, q) = 1$ puisque $1 + l\overline{n}q_1 \equiv \overline{n}(n + lq_1) \bmod q$. Alors par hypothèse $\chi(1 + l\overline{n}q_1) = 1$. Le conducteur du caractère est donc $q_1$ ou un diviseur de $q_1$.

De retour à notre preuve, on sait maintenant qu'il existe $k \in \mathbb{Z}$ tel que $(1 + kq_1, q) = 1$ et $\chi(1 + kq_1) \neq 1$, car $\chi$ est primitif et $q_1 < q$. Considérons maintenant

$$\chi(1 + kq_1) \sum_{j=1}^{d} \chi(b + jq_1) = \sum_{j=1}^{d} \chi(b + \Big(bk + (1 + kq_1)j\Big)q_1) = \sum_{j=1}^{d} \chi(b + jq_1).$$

Puisque $(1+kq_1, q) = 1$, $(bk+(1+kq_1)j)$ prendra toutes les valeurs mod $d$, il s'agit donc d'une permutation des termes de la somme qui explique pourquoi on revient sur la même somme dans la dernière égalité. Comme $\chi(1+kq_1) \neq 1$, pour que l'égalité ci-haute soit vraie, on doit avoir

$$\sum_{j=1}^{d} \chi(b+jq_1) = 0$$

ce qui termine la preuve.

Finalement, il est relativement facile et utile de calculer la norme de $\tau(\chi)$ pour $\chi$ primitif. On se sert de la relation qu'on vient de prouver et on met le tout au carré

$$|\tau(\chi)|^2|\overline{\chi}(n)|^2 = \left|\sum_{a=1}^{q} \chi(a)e\Big(\frac{an}{q}\Big)\right|^2$$

valide pour $1 \leq n \leq q$. On somme sur tous les $n$ pour obtenir

$$\varphi(q)|\tau(\chi)|^2 = \sum_{n=1}^{q}\left|\sum_{a=1}^{q} \chi(a)e\Big(\frac{an}{q}\Big)\right|^2$$

Comme $|z|^2 = z\overline{z}$, on évalue directement le produit des deux sommes et on change l'ordre de sommation afin d'isoler l'exponentielle pour obtenir

$$\sum_{n=1}^{q}\left|\sum_{a=1}^{q} \chi(a)e\Big(\frac{an}{q}\Big)\right|^2 = \sum_{a=1}^{q}\sum_{b=1}^{q}\chi(a)\overline{\chi}(b)\sum_{n=1}^{q} e\Big(\frac{(a-b)n}{q}\Big).$$

La somme sur $n$ vaut zéro si $a \neq b$ car on somme des points équidistants sur le cercle unité complexe. La somme vaut $q$ si $a = b$ car l'exponentielle vaudra 1 pour tout $n$. Alors,

$$\sum_{a=1}^{q}\sum_{b=1}^{q}\chi(a)\overline{\chi}(b)\sum_{n=1}^{q} e\Big(\frac{(a-b)n}{q}\Big) = q\sum_{a=1}^{q}\chi(a)\overline{\chi}(a) = q\varphi(q)$$

$$\Rightarrow \varphi(q)|\tau(\chi)|^2 = \varphi(q)q.$$

Alors $|\tau(\chi)| = \sqrt{q}$.

## 2.6 Géométrie algébrique

La géométrie algébrique s'intéresse aux zéros des polynômes. Puisqu'un polynôme effectue seulement les opérations d'addition et de multiplication, cette section travaillera avec un corps arbitraire $k$ étant donnée sa structure additive et multiplicative. Il est nécessaire d'étudier les polynômes sur $\bar{k}$, la fermeture algébrique de $k$, par notre désir d'étudier leurs zéros. On s'intéresse à l'anneau $k[X] := k[X_1, X_2, \ldots, X_n]$ des polynômes en $n$ variables avec coefficients en $k$. Comme chaque $X_i$ prendra ses valeurs dans $\bar{k}$, on définit l'espace affine de dimension $n$ sur $\bar{k}$

$$\mathbb{A}^n(\bar{k}) = \{(x_1, \ldots, x_n) \,|\, x_i \in \bar{k}\},$$

dénoté simplement par $\mathbb{A}^n$ si le corps est évident par le contexte.

Un polynôme $f \in k[X]$ induit une fonction $\mathbb{A}^n \to \bar{k}$ et ses zéros forment un sous-ensemble

de $\mathbb{A}^n$. Le sous-ensemble $\mathbb{A}^n(k) \subset \mathbb{A}^n(\bar{k})$ est appelé l'ensemble des points $k$-rationnels de $\mathbb{A}^n(\bar{k})$.

On définit de manière similaire un nouvel espace, un peu moins intuitif, qui sera plus large que l'espace affine pour une même dimension. Il s'agit de l'espace projectif construit par l'espace $\mathbb{A}^{n+1}$ sans son point $(0, 0, \ldots, 0)$ modulo la relation d'équivalence

$$(x_0, \ldots, x_n) \sim (y_0, \ldots, y_n)$$

s'il existe $\lambda \in \bar{k}^*$ tel que $(y_0, \ldots, y_n) = (\lambda x_0, \ldots, \lambda x_n)$. On dénote les classes d'équivalence par $[x_0, \ldots, x_n]$. Les points $k$-rationnels de $\mathbb{P}^n$ sont le sous-ensemble

$$\mathbb{P}^n(k) := \{[x_0, \ldots, x_n] \in \mathbb{P}^n \,|\, x_i \in k\}.$$

Intuitivement, il s'agit de l'espace de toutes les droites de $\mathbb{A}^{n+1}$ passant par l'origine. Les polynômes étudiés sur cet espace devront être homogènes de degré $n + 1$ afin que leurs zéros soient bien définis, c'est-à-dire, ne dépendent pas du choix du représentant du point de $\mathbb{P}^n$.

Les idéaux de $\bar{k}[X] = \bar{k}[X_0, \ldots, X_n]$ sont tous générés par un nombre fini de polynômes par le théorème de la base de Hilbert. Si les générateurs sont homogènes, on dit que l'idéal est homogène. On associe à un idéal homogène $I$ de $\bar{k}[X]$ un sous-ensemble de $\mathbb{P}^n$

$$V_I = \{P \in \mathbb{P}^n \,|\, f(P) = 0 \text{ pour tous les polynômes homogènes } f \text{ de } I\}.$$

Il s'agit donc de l'intersection des zéros de chaque générateur. On appelle un tel $V_I$ un ensemble algébrique projectif. Inversement, un ensemble algébrique projectif $V$ est associé à un idéal homogène $I(V)$ de $\bar{k}[X]$ généré par

$$\{f \in \bar{k}[X] \,|\, f \text{ est homogène et } f(P) = 0 \text{ pour tout } P \in V\}.$$

Si $I(V)$ se génère par des polynômes homogènes avec coefficients en $k$, on dit que $V$ est défini sur $k$ et on le dénote par $V/k$. Les points $k$-rationnels de $V$ sont dénotés par

$$V(k) = V \cap \mathbb{P}^n(k).$$

Si l'idéal homogène $I(V)$ de $\bar{k}[X]$ est premier, on dit que $V$ est une variété projective. La notion de variété affine est également définie dans l'espace affine de la même manière, sans toutefois la restriction d'homogénéité sur les idéaux de $\bar{k}[X_1, \ldots, X_n]$.

### 2.6.1 Le corps de fonctions d'une variété

Un corps important pour l'étude d'une variété affine est son corps de fonctions. On débute par la définition de

$$\bar{k}[V] := \frac{\bar{k}[X]}{I(V)},$$

qui est un anneau intègre. Il s'agit de tous les polynômes définis sur la variété. Le quotient nous assure que deux polynômes différents de cet anneau sont différents sur la variété. Le corps de fractions de cet anneau est appelé le corps de fonctions de la variété affine $V$ et est dénoté par $\bar{k}(V)$. Ce corps est composé de toutes les fonctions rationnelles définies sur la variété.

La dimension $\dim(V)$ d'une variété $V$ est définie comme étant le degré de transcendance de

$\bar{k}(V)$ par rapport à $\bar{k}$. Pour une variété définie par un idéal généré par un seul polynôme non constant, son degré sera $n-1$ où $n$ est le nombre de variables.

Une définition alternative du corps des fonctions d'une variété projective est donnée ainsi: il s'agit des fonctions $f(X)/g(X)$ pour $f, g \in \bar{k}[X] = \bar{k}[X_0, \ldots, X_n]$ où $f, g$ sont homogènes du même degré afin d'être bien définies sur l'espace projectif. Il faut également que $g \notin I(V)$ afin d'empêcher que la fonction rationnelle ne soit pas définie sur tous les points de la variété. Le corps des fonctions de la variété projective est donc l'espace de ces fonctions rationnelles modulo la relation d'équivalence $f_1/g_1 \sim f_2/g_2$ si $f_1 g_2 - f_2 g_1 \in I(V)$ car ces deux fonctions seront les mêmes sur la variété.

### 2.6.2 Points singuliers et réguliers

Une notion de régularité sur les points d'une variété existe dans le sens où il existe un plan tangent au point. Un point qui se comporte mal, c'est-à-dire que le plan tangent n'existe pas, est appelé singulier. Pour $V$ une variété affine et $P \in V$, soient $f_1, \ldots, f_m$ des générateurs de $I(V)$. Le point $P$ est non singulier si la matrice $(\partial f_i/\partial X_j)_{1 \leq i,j \leq n}$ a rang $n - \dim(V)$. Si tous les points de la variété sont non singuliers, la variété est dite non singulière ou lisse.

### 2.6.3 Modèles affines et projectifs

Le plan $\mathbb{A}^n$ est inclus de plusieurs manières à l'intérieur de $\mathbb{P}^n$ par les injections

$$\phi_i : \mathbb{A}^n \to \mathbb{P}^n$$

$$(x_1, \ldots, x_n) \mapsto [x_1, \ldots, x_{i-1}, 1, x_i, \ldots, x_n].$$

Dénotons l'image de cette fonction $U_i := \mathrm{Im}(\phi_i)$. On a alors une bijection de $U_i$ à $\mathbb{A}^n$ et l'inverse est donné par

$$\phi_i^{-1} : U_i \to \mathbb{A}^n$$

$$[x_0, \ldots, x_n] \mapsto \left( \frac{x_0}{x_i}, \frac{x_1}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \ldots, \frac{x_n}{x_i} \right).$$

On note que $x_i \neq 0$ sur $U_i$ et que l'on divise par $x_i$ afin que le représentant du point de $U_i$ ait la valeur 1 à la coordonnée $i$. Un polynôme $f$ en $n$ variables est envoyé à un polynôme $f^*$ homogène en $n+1$ variables ainsi

$$f^*(X_0, \ldots, X_n) = X_i^{\deg(f)} f \left( \frac{X_0}{X_i}, \frac{X_1}{X_i}, \ldots, \frac{X_{i-1}}{X_i}, \frac{X_{i+1}}{X_i}, \ldots, \frac{X_n}{X_i} \right).$$

Ce processus est appelé homogénéisation par rapport à $X_i$. On renverse ce processus en posant $X_i = 1$:

$$f(X_1, \ldots, X_n) = f^*(X_1, \ldots, X_{i-1}, 1, X_{i+1}, \ldots, X_n).$$

Cette substitution est appelée déhomogénéisation par rapport à $X_i$.

Une variété projective $V$ est alors associée à $n+1$ variétés affines. Pour $U_i$ fixé, $\phi_i^{-1}(V \cap U_i)$ sera une variété affine où $I(\phi_i^{-1}(V \cap U_i))$ sera la déhomogénéisation des polynômes de $I(V)$ par rapport à $x_i$. Remarquons que $V \cap U_0, \ldots, V \cap U_n$ couvrent $V$ complètement. Inversement, pour $V$ une variété affine, en homogénéisant les générateurs de $I(V)$ par rapport à un certain $x_i$ on obtient les générateurs d'un idéal homogène associé à une variété projective $\overline{V}$ appelée

fermeture projective de $V$. Pour $i$ fixé, le passage du modèle affine à projectif est réversible, c'est-à-dire, pour $V$ une variété affine on a $V = \phi_i^{-1}(\overline{V} \cap U_i)$ et pour $V$ une variété projective on a soit $\phi_i^{-1}(V \cap U_i) = \varnothing$, soit $V = \overline{\phi_i^{-1}(V \cap U_i)}$. On appelle *points à l'infini* d'une variété affine les points de l'espace projectif $\overline{V} \setminus \phi_i(V)$.

Le corps de fonctions $\bar{k}(V)$ d'une variété projective $V$ est défini par le corps de fonctions de la variété affine associée $\phi^{-1}(V \cap U_i)$. Les corps sont tous isomorphes pour les différents choix de $i$.

Un point $P$ de la variété projective $V$ est non singulier si pour un $U_i$ tel que $P \in U_i$ on a que le point $\phi_i^{-1}(P)$ est non singulier sur la variété affine $\phi_i^{-1}(V \cap U_i)$.

### 2.6.4 Morphismes de variétés

Il existe un concept de morphisme sur les variétés projectives. Soient $V_1 \subset \mathbb{P}^m$ et $V_2 \subset \mathbb{P}^n$ deux variétés projectives. On appelle une fonction $\phi : V_1 \to V_2$ un *morphisme* si elle peut être construite à partir de $n+1$ fonctions de même degré $f_i \in \bar{k}[V_1]$ par

$$\phi(P) = [f_0(P), \ldots, f_n(P)]$$

pour tous les points $P \in V_1$.

Deux variétés $V_1$ et $V_2$ sont dites isomorphes s'il existe deux morphismes $\phi : V_1 \to V_2$ et $\psi : V_2 \to V_1$ tels que $\psi \circ \phi$ et $\phi \circ \psi$ sont les identités sur $V_1$ et $V_2$ respectivement.

# 3 Fonction zêta de Riemann

## 3.1 Produit d'Euler

La définition de la fonction $\zeta$ de Riemann est la formule suivante:

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

où $s \in \mathbb{C}$. Elle est l'exemple prototype d'une série de Dirichlet qui sont des séries de la forme $\sum_{n\leq 1} a_n/n^s$. Par contre, cette série ne converge que pour $\mathrm{Re}(s) > 1$ et est analytique dans ce domaine. La question est donc, existe-t-il un prolongement analytique de cette fonction sur tout le plan complexe? La réponse est oui, avec un pôle simple en $s = 1$ de résidu 1, et fut prouvée par Riemann en 1859.

Dans la région où notre formule converge, il est possible d'exprimer $\zeta(s)$ sous une autre forme, appelée produit d'Euler, qui montre explicitement un lien avec les nombres premiers. On a l'égalité suivante, toujours dans la région $\mathrm{Re}(s) > 1$ :

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \frac{1}{p^{3s}} + \dots\right),$$

où le produit est sur les nombres premiers. Cette égalité est simplement due à l'unicité de la factorisation en nombres premiers d'un nombre, qu'on remarque facilement lorsqu'on développe le produit. On est maintenant en présence d'une série géométrique. En appliquant la formule pour son évaluation on obtient finalement

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \left(\frac{1}{1 - p^{-s}}\right)$$

pour $\mathrm{Re}(s) > 1$. Le but de ce chapitre est de montrer comment nous pouvons utiliser nos outils d'analyse complexe, plus précisément le théorème de Cauchy, afin d'obtenir des résultats sur la distribution des nombres premiers.

## 3.2 Prolongement analytique

### 3.2.1 Sommation de Poisson

Un outil très important dans la preuve du prolongement analytique est la sommation de Poisson qui permet de remplacer une fonction $f(x)$ par sa transformée de Fourier $\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e^{-2\pi i x\xi}dx$ à l'intérieur d'une somme de valeurs discrètes de $f$. L'équation est:

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{m\in\mathbb{Z}} \hat{f}(m)$$

si $f \in \mathcal{S}(\mathbb{R})$, l'espace de Schwartz qui contient des fonctions infiniment différentiables de forte décroissance, c'est-à-dire, $f^{(j)}(x) \ll_{j,n} 1/(1 + |x|^n)$ pour $j \in \mathbb{N} \cup 0$ et $n \in \mathbb{N}$. La décroissance est donc plus rapide que polynomiale. La signification des lettres $j, n$ en indice du symbole $\ll$ est que la constante bornant la croissance dépend de $j$ et $n$. La notation $f^{(j)}$ signifie la $j^{\text{ième}}$ dérivée de $f$.

*Preuve.* On définit la fonction $F(x) = \sum_{n\in\mathbb{Z}} f(n+x)$ qui aura période 1 puisqu'elle est une translation de distance $x$ de tous les points de $\mathbb{Z}$ dans $\mathbb{R}$. Nous allons voir que $F$ sera exprimable en série de Fourier

$$F(x) = \sum_{m\in\mathbb{Z}} a_m e(mx).$$

On calcule les $a_m$ ainsi

$$a_m := \int_0^1 F(x)e(-mx)dx = \int_0^1 \sum_{n\in\mathbb{Z}} f(n+x)e(-mx)dx$$

La convergence de la somme est démontrée par le critère de Weierstrass qui est satisfait par le fait que $f(x+n) \ll \frac{1}{n^2}$ pour $x \in [0,1]$ et $n \neq 0$ par les propriétés de décroissance d'une fonction de l'espace de Schwartz. On veut maintenant changer l'ordre d'intégration et de sommation par le critère de Weierstrass. En exprimant la somme infinie ainsi

$$\sum_{n\in\mathbb{Z}} f(n+x) = \lim_{N\to\infty} \sum_{-N\leq n\leq N} f(n+x),$$

on a que

$$a_m = \lim_{N\to\infty} \sum_{-N\leq n\leq N} \int_0^1 f(n+x)e(-mx)dx.$$

La fonction $e(x)$ a période 1, on peut donc ajouter $-mn$ sans problème à l'intérieur,

$$\begin{aligned} a_m &= \lim_{N\to\infty} \sum_{-N\leq n\leq N} \int_0^1 f(n+x)e(-m(n+x)))dx \\ &= \lim_{N\to\infty} \int_{-N}^{N+1} f(x)e(-mx)dx \\ &= \hat{f}(m) \end{aligned}$$

en rassemblant les bouts d'intégrales consécutifs sommés sous une seule intégrale.

La transformée de Fourier d'une fonction de l'espace de Schwartz est aussi élément de l'espace de Schwartz. Ceci implique $a_m = \hat{f}(m) \ll 1/m^2$ et donc la série de Fourier de $F(x)$ converge par Weierstrass. Alors, en évaluant $F(0)$ par sa définition originale et par sa représentation en série de Fourier en utilisant $a_m = \hat{f}(m)$, on obtient le théorème.

□

### 3.2.2 Équation fonctionnelle

Le but sera de trouver une formule pour $\zeta$ analytique dans tout le plan complexe, excepté un pôle simple de résidu 1 en $s = 1$. On remarquera également une symétrie respectée par la formule qui donnera l'équation fonctionnelle de $\zeta$. Le point de départ de Riemann dans sa preuve est la fonction gamma

$$\Gamma(s) := \int_0^\infty y^{s-1}e^{-y}dy$$

qui converge pour $\mathrm{Re}(s) > 0$. En effectuant le changement de variable $y = n^2\pi x$ et en évaluant à $s/2$ on obtient,

$$\pi^{-\frac{s}{2}}\Gamma\Big(\frac{s}{2}\Big)n^{-s} = \int_0^\infty x^{\frac{s}{2}-1}e^{-n^2\pi x}dx.$$

On voit ici qu'on a obtenu le facteur $n^{-s}$ dans le terme de gauche de notre équation. Pour retrouver la fonction $\zeta$, on somme donc sur tous les $n$ de 1 à l'infini. La restriction du domaine est maintenant dans la région $\mathrm{Re}(s) > 1$. On peut encore changer l'ordre de sommation et d'intégration par le théorème de convergence dominée ici. En effet, la fonction $f(y) := e^{-y^2\pi x}$ est élément de $\mathcal{S}(\mathbb{R})$, alors la somme a une convergence super-exponentielle. On peut donc la borner par son premier terme

$$\sum_{1\leq n\leq N} f(n) \ll e^{-\pi x}$$

pour tout $N$. L'équation devient maintenant

$$\pi^{-\frac{s}{2}}\Gamma\Big(\frac{s}{2}\Big)\zeta(s) = \int_0^\infty x^{\frac{s}{2}-1}\sum_{n=1}^{\infty} e^{-n^2\pi x}dx. \tag{3.1}$$

Si l'on pose $\omega(x) = \sum_{n=1}^{\infty} e^{-n^2\pi x}$, la somme dans notre intégrale, on aura que

$$\theta(x) := \sum_{n\in\mathbb{Z}} e^{-n^2\pi x} = 2\omega(x) + 1,$$

une somme à laquelle la sommation de Poisson est applicable. La fonction $g(y) = e^{-y^2\pi}$ est sa propre transformée de Fourier. En traitant $x$ comme une constante ici et en posant $f(y) = g(y\sqrt{x})$, on obtient que $\hat{f}(\xi) = \hat{g}(\xi/\sqrt{x})/\sqrt{x} = g(\xi/\sqrt{x})/\sqrt{x}$ par les propriétés de la transformée de Fourier. Par Poisson

$$\theta(x) = \sum_{n\in\mathbb{Z}} f(n) = x^{-1/2}\sum_{n\in\mathbb{Z}} e^{\frac{-n^2\pi}{x}} = x^{-1/2}\theta(1/x).$$

Ceci implique par la relation entre $\omega$ et $\theta$

$$\omega(1/x) = -\frac{1}{2} + \frac{1}{2}x^{1/2} + x^{1/2}\omega(x).$$

On voudrait se débarrasser de la région proche de 0 dans l'intégrale de l'équation (3.1) puisqu'elle pourrait y diverger. Pour ce faire, on sépare l'intégrale en deux parties, de 0 à 1 et de 1 à l'infini. Après le changement de variable $x' = 1/x$ dans l'intégrale sur $[0, 1]$ on obtient

$$\pi^{-\frac{s}{2}}\Gamma\Big(\frac{s}{2}\Big)\zeta(s) = \int_1^\infty x^{\frac{s}{2}-1}\omega(x)dx + \int_1^\infty x^{-\frac{s}{2}-1}\omega(1/x)dx.$$

On utilise maintenant notre relation entre $\omega(x)$ et $\omega(1/x)$ pour rassembler les deux intégrales sous une seule et on évalue l'intégrale des deux autres termes qui ne contiennent pas $\omega$ pour obtenir:

$$\pi^{-\frac{s}{2}}\Gamma\Big(\frac{s}{2}\Big)\zeta(s) = \frac{1}{s(s-1)} + \int_1^\infty (x^{\frac{1}{2}s-1} + x^{-\frac{1}{2}s-\frac{1}{2}})\omega(x)dx. \tag{3.2}$$

Cette formule est valide pour $\mathrm{Re}(s) > 1$. Par contre, puisque $\omega(x) \ll e^{-\pi x}$ comme montré ci-haut, l'intégrale converge absolument pour n'importe quel $s \in \mathbb{C}$. On a également la convergence uniforme sur n'importe quel compact du plan. Comme la fonction $\Gamma$ est définie sur tout le plan complexe, cette équation nous donne le prolongement analytique de $\zeta$, c'est-à-dire, une fonction analytique pour tout $s \in \mathbb{C}$ et qui vaut $\sum_{n=1}^{\infty} n^{-s}$ dans la région $\mathrm{Re}(s) > 1$. On remarque que le côté droit de l'équation (3.2) demeure inchangé en remplaçant $s$ par $1 - s$. Ceci nous donne donc l'équation fonctionnelle pour $\zeta$:

$$\pi^{-\frac{1}{2}s}\Gamma\Big(\frac{1}{2}s\Big)\zeta(s) = \pi^{-\frac{1}{2}(1-s)}\Gamma\Big(\frac{1}{2}(1-s)\Big)\zeta(1-s). \tag{3.3}$$

### 3.3 Région sans zéro

On voit par (3.2) que la fonction

$$\xi(s) = \frac{1}{2}s(s-1)\pi^{-\frac{1}{2}s}\Gamma\Big(\frac{1}{2}s\Big)\zeta(s)$$

sera analytique sur tout le plan complexe. L'équation fonctionnelle devient plus élégante,

$$\xi(s) = \xi(1-s)$$

pour tout $s \in \mathbb{C}$. Comme $\Gamma$ n'a aucun zéro et que son pôle à $s = 0$ annule le zéro du facteur $s$ à ce point, le seul pôle possible pour $\zeta$ est à $s = 1$. On voit également que $\zeta$ aura des zéros à $-2, -4, -6, \ldots$ afin d'éliminer les pôles de $\Gamma$ à ces points. Ces zéros de $\zeta$ sont appelés les zéros triviaux. Également par (3.2) et du fait que $\Gamma(\overline{s}) = \overline{\Gamma(s)}$, nous avons la symétrie suivante: $\zeta(\overline{s}) = \overline{\zeta(s)}$.

La fonction $\zeta(s)$ n'a aucun zéro dans la région $\mathrm{Re}(s) > 1$ puisque le produit d'Euler nous donne

$$\frac{1}{\zeta(s)} = \prod_p (1 - p^{-s}) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$$

qui converge absolument pour $\mathrm{Re}(s) > 1$, tout comme $\Gamma(s/2)$. L'équation fonctionnelle nous assure donc que $\zeta(s)\Gamma(s/2)$ n'a aucun zéro dans la région $\mathrm{Re}(s) < 0$. La fonction $\zeta(s)$ n'a donc que les zéros triviaux discutés ci-haut dans cette région. Les autres zéros possibles doivent donc se situer dans la région $\{s \in \mathbb{C} : 0 \leq \mathrm{Re}(s) \leq 1\}$ appelée *la bande critique*. Par l'équation fonctionnelle et la symétrie $\zeta(\overline{s}) = \overline{\zeta(s)}$, un zéro $\rho$ implique que $\overline{\rho}, 1-\rho$ et $1-\overline{\rho}$ sont également tous des zéros de $\zeta(s)$. D'ailleurs, la fameuse hypothèse de Riemann prédit que tous les zéros non triviaux ont $\mathrm{Re}(\rho) = 1/2$, c'est-à-dire, sont situés directement sur la ligne de symétrie. Ces zéros sont appelés zéros non triviaux. Comme nous allons le voir, il est possible de prouver qu'il n'y a aucun zéro dans une certaine région à l'intérieur de la bande critique.

**Théorème 3.1.** *La fonction $\zeta(s)$ n'a aucun zéro dans la région*

$$\sigma \geq 1 - \frac{c}{\log(|t|+2)},$$

*pour un certain $c > 0$.*

*Preuve.* La preuve se servira de la dérivée logarithmique de $\zeta$ au lieu de $\zeta$ elle-même afin d'avoir des pôles à la position des zéros de $\zeta$. On calcule ainsi la dérivée du logarithme de zêta à partir du produit d'Euler,

$$(\log(\zeta(s))' = \left(-\sum_p \log(1-p^{-s})\right)' = -\sum_p \frac{p^{-s}\log(p)}{1-p^{-s}} = -\sum_p \log(p)\left(\frac{1}{1-p^{-s}} - 1\right)$$

$$= -\sum_p \sum_{n\geq 1} \log(p) p^{-ns}.$$

où la convergence justifie le changement d'ordre entre la sommation et la dérivée. On a donc

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n\geq 1} \Lambda(n) n^{-s}$$

dans la région $\mathrm{Re}(s) > 1$. Puisque $n^{-s} = e^{-s\log(n)}$, on a:

$$\mathrm{Re}\bigg( -\frac{\zeta'(s)}{\zeta(s)}\bigg) = \sum_{n\geq 1} \Lambda(n) n^{-\sigma}\cos(t\log n)$$

toujours dans la région $\mathrm{Re}(s) > 1$.

La preuve repose essentiellement sur l'inégalité trigonométrique suivante, qui nous donnera une inégalité reliant la partie imaginaire d'un zéro de $\zeta$ avec sa partie réelle,

$$3 + 4\cos(\theta) + cos(2\theta) \geq 0$$

valide pour tout $\theta$ puisqu'elle découle simplement de l'expansion de $2(1+\cos(\theta))^2$. En posant $\theta = t\log n$, en multipliant par $\Lambda(n)n^{-\sigma} \geq 0$ et en sommant, on obtient:

$$\sum_{n\geq 1} \Lambda(n) n^{-\sigma}\Big(3 + 4\cos(t\log n) + \cos(2t\log n)\Big) \geq 0$$

$$\Rightarrow 3\bigg( -\frac{\zeta'(\sigma)}{\zeta(\sigma)}\bigg) + 4\bigg( -\mathrm{Re}\frac{\zeta'(\sigma+it)}{\zeta(\sigma+it)}\bigg) + \bigg( -\mathrm{Re}\frac{\zeta'(\sigma+2it)}{\zeta(\sigma+2it)}\bigg) \geq 0. \tag{3.4}$$

On bornera chacun des trois termes et on obtiendra l'inégalité voulue nous donnant des informations sur la partie réelle d'un zéro de $\zeta$.

Afin de borner les deux derniers termes, on utilisera la formule suivante pour $\xi(s)$ (voir Davenport [1] Chapitre 12),

$$\xi(s) = e^{A+Bs}\prod_{\rho}\left(1 - \frac{s}{\rho}\right)e^{s/\rho} \tag{3.5}$$

valide dans tout le plan et où le produit sur $\rho$ signifie les zéros de la fonction $\xi(s)$ comptés avec multiplicité qui sont en fait les zéros non triviaux de $\zeta(s)$. Cette égalité découle de la théorie sur les fonctions entières (voir Davenport [1] Chapitre 11) puisqu'on peut montrer $\xi(s) = e^{\mathcal{O}(|s|\log(|s|+2))}$ en utilisant l'approximation de Stirling pour borner $\Gamma$ et une certaine expression pour $\zeta$ facile à borner.

On peut également montrer que la somme sur la norme des zéros diverge, donc que $\zeta$ a une infinité de zéros. Si la somme convergeait, c'est-à-dire

$$\sum_{\rho}\frac{1}{|\rho|} < \infty,$$

alors en utilisant l'inégalité $|(1-z)e^z| \leq e^{2|z|}$, $z \in \mathbb{C}$, qui découle de la série de puissances de $(1-z)e^z$, à l'intérieur du produit dans (3.5) on aurait

$$|\xi(s)| < e^{C|s|}$$

pour un certain $C < \infty$. Par contre, lorsque $s \to \infty$ par les réels, on a $\zeta(s) \to 1$ et $\log\Gamma(s) \sim s\log s$ par Stirling, donc $\xi$ ne satisfait pas cette borne.

Revenons maintenant à l'inégalité (3.4). On aimerait avoir une borne sur $-\zeta'(s)/\zeta(s)$ en utilisant le produit (3.5) pour $\xi$. Avec l'équation $\xi(s) = \frac{1}{2}s(s-1)\pi^{-\frac{1}{2}s}\Gamma\Big(\frac{1}{2}s\Big)\zeta(s)$, on applique la dérivée logarithmique pour obtenir:

$$-\frac{\zeta'(s)}{\zeta(s)} = \frac{1}{s-1} - B - \frac{1}{2}\log\pi + \frac{1}{2}\frac{\Gamma'(s/2+1)}{\Gamma(s/2+1)} - \sum_{\rho}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right)$$

Le pôle de $\zeta(s)$ à $s=1$ est explicite dans cette formule. Toujours par Stirling, le terme en $\Gamma$ est borné par $A\log t$ dans la région $t \geq 2$ et $1 \leq \sigma \leq 2$. Dans cette région, on a donc:

$$-\mathrm{Re}\frac{\zeta'(s)}{\zeta(s)} < A'\log t - \sum_{\rho}\mathrm{Re}\left(\frac{1}{s-\rho} + \frac{1}{\rho}\right),$$

avec $A' > A$ absorbant les autres termes. Chaque terme dans la somme est positif car, pour $\rho = \beta + i\gamma$,

$$\mathrm{Re}\frac{1}{s-\rho} = \frac{\sigma-\beta}{|s-\rho|^2},$$
$$\mathrm{Re}\frac{1}{\rho} = \frac{\beta}{|\rho|^2}.$$

Alors on peut borner supérieurement le deuxième terme de (3.4) en ne gardant qu'un seul terme $1/(s-\rho)$ dans la somme correspondant à un zéro quelconque. La relation entre la partie réelle et imaginaire émerge de ce terme. On pose également $t = \gamma$ pour être proche du zéro choisi puisque le pôle s'y trouvant influence son voisinage par analyticité. On obtient,

$$-\mathrm{Re}\frac{\zeta'(\sigma+it)}{\zeta(\sigma+it)} < A'\log t - \frac{1}{\sigma-\beta}.$$

On borne le troisième terme simplement par

$$-\mathrm{Re}\frac{\zeta'(\sigma+2it)}{\zeta(\sigma+2it)} < A'\log t$$

Finalement, on borne le premier terme par:

$$-\frac{\zeta'(\sigma)}{\zeta(\sigma)} < \frac{1}{\sigma-1} + A''$$

valide pour $1 < \sigma \leq 2$ pour un certain $A'' > 0$. Cette inégalité provient du pôle simple de $\zeta$ à $s=1$. Donc, on obtient l'inégalité suivante en utilisant les trois bornes dans (3.4)

$$\frac{4}{\sigma-\beta} < \frac{3}{\sigma-1} + A'''\log t.$$

Comme nous avons la liberté de choisir $\sigma > 1$ et que nous souhaitons mettre en relation $\beta$ et $t$, il est possible de poser $\sigma = 1 + \delta/\log t$, où $\delta > 0$ arbitraire, pour obtenir

$$\beta < 1 + \frac{\delta}{\log t} - \frac{4\delta}{(3 + A'''\delta)\log t}.$$

La constante $\delta$ peut être choisie selon $A'''$ afin d'obtenir

$$\beta < 1 - \frac{c}{\log t},$$

pour $t \geq 2$ où $c$ est une certaine constante positive.

Il est possible de montrer que $\zeta$ n'a aucun zéro sur $\sigma = 1$ pour tout $t$ en utilisant la même idée d'appliquer l'inégalité trigonométrique, mais cette fois sur $\log \zeta(s)$. L'argument est toutefois beaucoup plus court et simple, et fut amélioré trois ans plus tard par de la Vallée Poussin afin de donner la preuve de la région sans zéro présentée ci-haut. On peut donc affirmer que pour tout $t$, il n'y a aucun zéro dans cette région

$$\sigma \geq 1 - \frac{c}{\log(|t| + 2)},$$

pour un certain $c > 0$. La force des estimations sur les nombres premiers dépend grandement de la profondeur de la région sans zéro comme on le remarquera dans la prochaine section. L'hypothèse de Riemann engendre donc des estimations très précises. □

## 3.4 Théorème des nombres premiers

Le théorème principal de cette section est le suivant.

**Théorème 3.2.** *Pour une certaine constante $c > 0$*

$$\psi(x) = x + \mathcal{O}\Big(x \exp(-c\sqrt{\log(x)})\Big).$$

Ce théorème donne une formule asymptotique pour $\psi(x)$ par le biais de l'intégration complexe dans la bande critique et montre l'utilité de tout ce que nous avons fait jusqu'à présent. On remarque donc que même si les nombres premiers sont des objets assez simples à définir, obtenir des informations sur leur distribution demande de grands efforts. Afin d'utiliser le théorème de Cauchy pour l'intégration complexe, il est nécessaire de transformer les zéros de $\zeta$ en pôles. C'est la raison pourquoi on travaillera avec la dérivée logarithmique de $\zeta$. D'ailleurs, la formule que l'on obtient est pour $\psi(x)$ puisque les coefficients de $-\zeta'/\zeta$ dans la représentation en série de Dirichlet sont égaux à $\Lambda(n)$. Cette section rendra évidente l'importance du prolongement analytique de $\zeta$ étant donné que les zéros non triviaux se trouvent hors de la zone de convergence de la série de Dirichlet de $\zeta$.

On préférerait avoir une estimation sur $\pi(x)$ plutôt que $\psi(x)$ par simplicité.

**Théorème des nombres premiers.**

$$\pi(x) = \mathrm{li}(x) + \mathcal{O}\Big(x \exp(-c\sqrt{\log(x)})\Big)$$

*où* $\mathrm{li}(x) := \int_2^x dx/\log(x)$.

*Preuve.* Premièrement, on a

$$|\psi(x) - \theta(x)| \ll \sqrt{x}\log^2(x),$$

car

$$\psi(x) - \theta(x) = \sum_{\substack{p^k \leq x \\ k \geq 2}} \log p \leq \log(x) \sum_{\substack{p^k \leq x \\ k \geq 2}} 1.$$

Il y a $\log(x)/\log(p) \ll \log(x)$ choix de $k$, et comme $k \geq 2$, on a $p \leq \sqrt{x}$. La différence est donc absorbée par l'erreur. Alors, par le Théorème 3.2

$$\theta(x) = x + \mathcal{O}\Big(x \exp(-c\sqrt{\log(x)})\Big).$$

On utilise la sommation partielle pour relier $\pi(x)$ à $\theta(x)$

$$\pi(x) = \int_{2^-}^{x} \frac{1}{\log y} d\theta(y) = \frac{\theta(x)}{\log x} + \int_{2^-}^{x} \frac{\theta(y)}{y \log^2 y} dy.$$

Pour la partie significative de $\theta(y)$, l'intégrale vaut

$$\int_{2^-}^{x} \frac{1}{\log^2 y} dy = \mathrm{li}(x) - \frac{x}{\log x} + \frac{2}{\log 2}.$$

On traite trivialement les termes d'erreur pour obtenir

$$\pi(x) = \mathrm{li}(x) + \mathcal{O}\Big(x \exp(-c\sqrt{\log(x)})\Big).$$

Mentionnons ici le fait que

$$\mathrm{li}(x) \sim \frac{x}{\log x}.$$

□

L'outil principal utilisé mettant en évidence le lien entre les zéros et la somme des coefficients est l'inversion de Mellin. On définit la transformée de Mellin d'une fonction $\phi : (0, +\infty) \to \mathbb{C}$ à support compact ainsi

$$\hat{\phi}(s) = \int_0^\infty \phi(x) x^{s-1} dx$$

donnant une nouvelle fonction qui est maintenant de $\mathbb{C}$ à $\mathbb{C}$ en supposant qu'elle est définie partout. Il est possible de récupérer $\phi(x)$ à partir de $\hat{\phi}(s)$ ainsi

$$\phi(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} x^{-s} \hat{\phi}(s) ds$$

pour un $c \in \mathbb{R}$ tel que $\hat{\phi}(s)$ est analytique dans la bande $c - \delta < \mathrm{Re}(s) < c + \delta$ pour un $\delta > 0$. On remarque que la fonction $\Gamma(s)$ est la transformée de Mellin de $e^{-x}$ pour $\mathrm{Re}(s) > 0$.

On débute avec la somme suivante:

$$\sum_{n=1}^{\infty} \Lambda(n)\phi(n),$$

où la présence de la fonction à support compact $\phi$ dans la somme sert à ce qu'on la remplace par l'inversion de Mellin de sa transformée de Mellin, c'est-à-dire, on la remplace par elle-même. La raison étant que l'on obtiendra une intégrale complexe dans la formule et que l'on retrouvera notre fonction $\zeta'/\zeta$.

Si on choisit sagement notre fonction $\phi$, on retrouvera $\psi(x)$ ainsi qu'une petite erreur dans la somme ci-haute. Le choix judicieux est de choisir, pour $1 \leq x \leq y$ fixés

$$\phi(z) = \begin{cases} 0 & z < 0 \\ z & 0 \leq z < 1 \\ 1 & 1 \leq z \leq x \\ 1 + \frac{x-z}{y} & x < z \leq x + y \\ 0 & z > x + y. \end{cases}$$

Puisque cette fonction continue vaut 1 entre $z = 1$ et $z = x$, on obtient en l'utilisant dans la somme

$$\sum_{n=1}^{\infty} \Lambda(n)\phi(n) = \sum_{n \leq x} \Lambda(n) + \sum_{x < n \leq x+y} \Lambda(n)\phi(n) = \psi(x) + \sum_{x < n \leq x+y} \Lambda(n)\phi(n).$$

La somme sur les $n$ de $x$ à $x + y$ représente notre erreur qu'on souhaite contrôler. On borne trivialement $\phi(n) \leq 1$ et $\Lambda(n) \leq \log(x + y) \leq \log(2x) \ll \log(x)$. Comme on somme $y$ termes, on obtient la borne triviale suivante

$$\sum_{x < n \leq x+y} \Lambda(n)\phi(n) \ll y \log(x).$$

On obtient donc l'équation

$$\sum_{n=1}^{\infty} \Lambda(n)\phi(n) = \psi(x) + \mathcal{O}(y \log(x)).$$

Il est maintenant temps d'appliquer l'inversion de Mellin dans la série.

$$\sum_{n=1}^{\infty} \Lambda(n)\phi(n) = \frac{1}{2\pi i} \sum_{n=1}^{\infty} \Lambda(n) \int_{2-i\infty}^{2+i\infty} n^{-s} \hat{\phi}(s) ds$$

L'égalité est triviale, car on a remplacé $\phi(n)$ par l'inversion de sa transformée de Mellin qui revient donc sur $\phi(n)$. L'intégrale converge absolument par une borne sur $\hat{\phi}(s)$ qu'on donnera dans quelques instants, ce qui nous permet d'intervertir l'ordre de sommation et d'intégration. Comme on intègre sur $\sigma = 2$, la série converge et on retrouve la dérivée logarithmique de $\zeta$:

$$\frac{1}{2\pi i} \sum_{n=1}^{\infty} \Lambda(n) \int_{2-i\infty}^{2+i\infty} n^{-s} \hat{\phi}(s) ds = \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} \hat{\phi}(s) ds = \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} -\frac{\zeta'}{\zeta}(s)\hat{\phi}(s) ds.$$

On vient donc de mettre en relation une somme sur les puissances de nombres premiers avec une intégrale complexe d'une fonction méromorphe. L'inversion de Mellin est une stratégie très élégante fréquemment utilisée en théorie des nombres analytique. Elle permet entre autres de relier des sommes sur les nombres premiers avec des sommes sur les zéros de certaines fonctions méromorphes, car on peut utiliser le théorème de Cauchy en fermant la région d'intégration. Dans le cas présent, on cherchera à éviter les zéros, d'où l'importance d'avoir calculé une région sans zéro.

Grâce au prolongement analytique de $\zeta$, on peut changer la ligne d'intégration tant qu'on ne passe pas par un pôle. On dénote par $\mathcal{Z}$ la frontière de la région suivante, $\{s = \sigma + it | \sigma >$

$1 - c/3\log(|t|+2)\}$ où $c$ est la constante de notre région sans zéro. Pour appliquer le théorème des résidus de Cauchy sur $\mathcal{Z}$ et la ligne verticale $\sigma = 2$, il est nécessaire de fermer la région par deux lignes horizontales à la hauteur $-T$ et $T$. Pour tout $T$, seulement un pôle est capturé à $s = 1$ de résidu $\hat{\phi}(1)$. Par les estimations que l'on calculera bientôt sur $-\zeta'/\zeta$ et $\hat{\phi}(s)$, la fonction qu'on intègre tend vers 0 pour $t \to \infty$ et donc l'intégrale sur les lignes horizontales tend vers 0 lorsque $T \to \infty$. On a donc par Cauchy

$$\frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} -\frac{\zeta'}{\zeta}(s)\hat{\phi}(s)ds - \frac{1}{2\pi i}\int_{\mathcal{Z}} -\frac{\zeta'}{\zeta}(s)\hat{\phi}(s)ds = \hat{\phi}(1)$$

Le signe moins devant la deuxième intégrale est parce qu'on intègre de bas en haut. Donc

$$\sum_{n=1}^{\infty}\Lambda(n)\phi(n) = \hat{\phi}(1) + \frac{1}{2\pi i}\int_{\mathcal{Z}} -\frac{\zeta'}{\zeta}(s)\hat{\phi}(s)ds.$$

On calcule maintenant les bornes sur $-\zeta'/\zeta$ et $\hat{\phi}(s)$. Dans la définition de la région $\mathcal{Z}$, on a divisé $c$ par 3 car on veut que $\mathcal{Z}$ soit à distance d'au moins $\gg 1/\log(|t|+2)$ des zéros de $\zeta$, où la constante implicite est absolue. La raison étant qu'il existe une borne dans notre région (voir Koukoulopoulos [2] Lemme 3.2.3) pour $\zeta'/\zeta(s)$ qui dépend de la distance entre $s$ et les zéros de $\zeta$,

$$\frac{\zeta'}{\zeta}(s) = -\frac{1}{s-1} + \sum_{\rho:|\gamma-t|\leq 1}\frac{1}{s-\rho} + \mathcal{O}(\log|t|+2)).$$

Cette borne nous donne

$$-\frac{\zeta'}{\zeta}(s) \ll \log^2(|t|+2)$$

pour $s \in \mathcal{Z}$. Il reste maintenant à borner $\hat{\phi}(s)$ dans la région $1/2 \leq \sigma \leq 2$. Par définition

$$\hat{\phi}(s) = \int_0^{x+y}\phi(z)z^{s-1}dz.$$

On borne d'abord $\phi(z)$ par 1. On applique ensuite l'intégration par partie. On dérive $\phi(z)$ (la dérivée ne sera pas continue, il faut considérer l'intégrale en trois parties) et on intègre $z^{s-1}$. Alors

$$\hat{\phi}(s) \ll \frac{-1+(x+y)^s+(x/y)((x+y)^s-x^s)}{s(s+1)} \ll \frac{x^\sigma}{|s|}\min\left(1, \frac{x}{|s|y}\right).$$

En utilisant maintenant toutes nos bornes dans l'intégrale sur $\mathcal{Z}$, on obtient

$$\frac{1}{2\pi i}\int_{\mathcal{Z}} -\frac{\zeta'}{\zeta}(s)\hat{\phi}(s)ds \ll \int_{\mathcal{Z}}\frac{x^\sigma}{|s|}\min\left(1,\frac{x}{|s|y}\right)\log^2(|t|+2)ds \ll x^{\sigma(T)}\log^3(T)$$

où on définit $T = x/y$ et $\sigma(T) = 1 - c/3\log(T)$. Comme on est dans la région $1/2 \leq \sigma \leq 2$, on a que $|t| \asymp |s|$. Alors à hauteur $t = T$ sur $\mathcal{Z}$ dans l'intégrale, le min devient $x/|s|y$. Ce changement provoque la convergence de l'intégrale sur le reste de $\mathcal{Z}$ puisqu'au dénominateur, $|s|$ est à la puissance 2. L'intégrale sur les queues sera essentiellement $1/|s|$ évalué à hauteur $T$, ce qui nous donnera une quantité en accord avec notre borne. Pour la partie entre $-T$ et $T$, la valeur maximale que $x^\sigma$ prendra est $x^{\sigma(T)}$ car $\sigma(T)$ est la plus grande valeur de $\sigma$ dans cette zone. On peut donc le sortir de l'intégrale. De même pour le facteur log qui trouvera sa plus grande valeur à $\log^2(T)$. Il reste donc à intégrer $1/|s|$ sur $\mathcal{Z}$ de $-T$ à $T$ qui donne la troisième puissance du log dans l'estimation.

Il est facile de voir que $\hat{\phi}(1) = x + y/2 - 1/2 = x + \mathcal{O}(y)$. On remarquera que c'est ce résidu qui donne le terme principal de notre estimation sur $\psi(x)$. En rassemblant enfin toutes nos estimations, on obtient

$$\psi(x) = x + \mathcal{O}\Big((xT^{-1} + x^{\sigma(T)})\log^3(x)\Big).$$

Le choix de $T$ est libre tant que $1 \leq T \leq x$. Pour retirer la dépendance en $y$ de notre estimation, on pose $T = \exp(\sqrt{\log(x)}/3)$. Un simple calcul donne

$$\psi(x) = x + \mathcal{O}\Big(x\exp(-c\sqrt{\log(x)})\Big)$$

pour $x \leq 2$ où $c > 0$ est une nouvelle constante. Alors,

$$\psi(x) \sim x.$$

Sous l'hypothèse de Riemann, on obtient une estimation beaucoup plus forte. À la place d'intégrer sur $\mathcal{Z}$, on intègre sur la ligne verticale $\sigma = 1/2 + \epsilon$ et dans l'estimation, à la place d'avoir $x^{\sigma(T)}$ on aura $x^{1/2+\epsilon}$ pour $\epsilon > 0$. Avec $T = x$, on obtient

$$\psi(x) = x + \mathcal{O}_\epsilon(x^{1/2+\epsilon}).$$

Étant donnée la force de l'estimation par rapport à ce que nous sommes capables d'obtenir sans l'hypothèse, il ne fait aucun doute que l'hypothèse de Riemann est très difficile à prouver.

# 4 Fonctions $L$

## 4.1 Définitions

Le chapitre précédent nous a montré qu'il était possible d'obtenir des informations sur les nombres premiers à partir de la fonction zêta de Riemann. Les techniques engendraient des estimations sur la somme des coefficients de la dérivée logarithmique jusqu'à $x$. En utilisant ces mêmes techniques, il sera possible d'obtenir des informations sur divers autres objets mathématiques en définissant des fonctions $L$ à partir de ceux-ci. Ce chapitre étudiera les fonctions $L$ d'un point de vue général, c'est-à-dire, sans lien avec aucun objet concret. On pourra obtenir des résultats simplement à partir de la définition abstraite. Les résultats ainsi obtenus pourront donc être appliqués à un cas concret sans avoir à refaire les preuves. Bien sûr, en travaillant directement avec une fonction $L$ les estimations pourront s'améliorer dans quelques cas. Les estimations obtenues dans le cas général sont suffisantes pour plusieurs applications. On commence par la définition d'une fonction $L$. Cette définition provient d'un cadre respecté par plusieurs fonctions $L$ reliées à des objets de diverses branches des mathématiques.

La fonction $L(f,s) : \mathbb{C} \to \mathbb{C}$, où $s$ est la variable complexe et $f$ ne sert qu'à indiquer l'objet associé, est une *fonction* $L$ si elle respecte les conditions suivantes :

1. Il existe une série de Dirichlet représentant $L(f,s)$ dans la région de convergence, c'est-à-dire
$$L(f,s) = \sum_{n\geq 1} \frac{\lambda_f(n)}{n^s},$$
où $\lambda_f(1) = 1$ et $\lambda_f(n) \in \mathbb{C}$. La série doit converger absolument dans la région $\mathrm{Re}(s) > 1$. Cette série de Dirichlet doit pouvoir s'exprimer en produit d'Euler :
$$\sum_{n\geq 1} \frac{\lambda_f(n)}{n^s} = \prod_p (1-\alpha_1(p)p^{-s})^{-1} \dots (1-\alpha_d(p)p^{-s})^{-1},$$
où $\alpha_i(p) \in \mathbb{C}$ et $d$ est dit le *degré* de la fonction $L$. En particulier, $\lambda_f$ est une fonction multiplicative avec $\lambda_f(p) = \alpha_1(p) + \cdots + \alpha_d(p)$. Ce produit doit également converger absolument dans la région $\mathrm{Re}(s) > 1$. On appelle *paramètres locaux à* $p$ les quantités $\alpha_i(p)$ et ils doivent respecter $|\alpha_i(p)| < p$ pour tout $p$ et tout $1 \leq i \leq d$ afin que $L(f,s) \neq 0$ pour $\mathrm{Re}(s) > 1$.

2. Il existe un *facteur gamma* qui sera présent dans l'équation fonctionnelle
$$\gamma(f,s) := \pi^{-ds/2} \prod_{j=1}^{d} \Gamma\Big(\frac{s+\kappa_j}{2}\Big),$$
où les quantités $\kappa_j \in \mathbb{C}$ sont appelés les *paramètres locaux à l'infini* et doivent satisfaire $\mathrm{Re}(\kappa_j) > -1$. Soit ces nombres sont réels, soit ils sont en paire de conjugués complexes.

3. Il existe un entier $q(f) \geq 1$, appelé *conducteur* de $L(f,s)$, tel que $\alpha_i(p) \neq 0$ pour tout $p \nmid q(f)$ et $1 \leq i \leq d$. On qualifie de *ramifié* un nombre premier divisant $q(f)$.

4. La fonction $L$ possède un prolongement méromorphe dont les seuls pôles possibles sont en $s = 0$ et $s = 1$. L'équation fonctionnelle suivante doit être respectée
$$\Lambda(f,s) = \epsilon(f)\Lambda(\overline{f}, 1-s),$$

où

$$\Lambda(f,s) := q(f)^{s/2}\gamma(f,s)L(f,s)$$

est appelée la *fonction L complète*. L'objet $\overline{f}$ est appelé le *dual* de $f$ et doit respecter $\lambda_{\overline{f}}(n) = \overline{\lambda_f(n)}$, $\gamma(\overline{f},s) = \gamma(f,s)$ et $q(\overline{f}) = q(f)$. La quantité $\epsilon(f)$ est appelée *signe* de $L(f,s)$ et respecte $|\epsilon(f)| = 1$. De plus, la fonction $\Lambda(f,s)$ doit être holomorphe, sauf possiblement à des pôles à $s = 0$ et $s = 1$. Si ces pôles existent, ils auront le même ordre par l'équation fonctionnelle. On dénote par $r(f)$ l'ordre des pôles. S'il n'y a pas de pôles et si $L(f,0) \neq 0$ et $L(f,1) \neq 0$, on pose $r(f) = 0$. Si $L(f,0) = 0$ et $L(f,1)$=0, on pose $r(f) = -k$ où $k$ est l'ordre de ces zéros qui est égal par l'équation fonctionnelle.

Notre définition n'englobe pas toutes les fonctions intéressantes. Beaucoup de fonctions possédant une représentation en série de Dirichlet échouent à satisfaire quelques conditions. Par exemple, une fonction $L$ de Dirichlet d'un caractère non primitif définie ci-après ne satisfait pas notre définition. En effet, l'équation fonctionnelle aura des facteurs supplémentaires.

L'avantage de travailler d'un point de vue général est que les estimations dépendront non seulement de $s$, mais également des paramètres. Par exemple, les principales estimations seront sur le nombre de zéros $\rho = \beta + i\gamma$ tels que $|\gamma| \leq X$ ou sur la somme des coefficients jusqu'à $X$ de la série de Dirichlet de la dérivée logarithmique de $L(f,s)$, autrement appelé le théorème des nombres premiers. Le but sera de trouver des estimations uniformes en ces paramètres et en $s$, ce qui nous permettra de varier non seulement $s$, mais également la fonction $L$ à travers une famille par exemple. Il est pratique de définir les fonctions et quantités suivantes afin d'alléger les bornes :

$$\mathfrak{q}_\infty(s) := \prod_{j=1}^{d}(|s+\kappa_j|+3),$$

$$\mathfrak{q}(f,s) := q(f)\mathfrak{q}_\infty(s),$$

$$\mathfrak{q}(f) := \mathfrak{q}(f,0).$$

On appelle $\mathfrak{q}(f,s)$ le *conducteur analytique* de $L(f,s)$. On note que $\mathfrak{q}(f)$ rassemble le degré, le conducteur et les paramètres locaux à l'infini sous une seule quantité.

Donnons maintenant deux exemples. La fonction zêta de Riemann $L(f,s) = \zeta(s)$ est de degré $d = 1$ avec $\lambda_f(n) = 1$ pour tout $n$, $\alpha_1(p) = 1$ pour tout $p$, $\kappa_1 = 0$, $q(f) = 1$, $\epsilon(f) = 1$ et $r(f) = 1$.

Le deuxième exemple est les fonctions $L$ de Dirichlet. À partir d'un caractère $\chi$ modulo $q$, on définit

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

qui converge pour $\mathrm{Re}(s) > 0$ si $\chi$ est non principal et pour $\mathrm{Re}(s) > 1$ sinon. L'ensemble des fonctions $L$ de Dirichlet provenant de caractères primitifs est un exemple d'une famille de fonctions $L$. La raison pour laquelle on se concentre sur les caractères primitifs est que pour un caractère $\chi$ mod $q$ induit par $\chi_1$ mod $q_1$ on a la relation

$$L(s,\chi) = \prod_p\left(1-\frac{\chi(p)}{p^s}\right)^{-1} = \prod_{p\nmid q}\left(1-\frac{\chi_1(p)}{p^s}\right)^{-1} = \prod_{p\nmid q_1}\left(1-\frac{\chi_1(p)}{p^s}\right)^{-1}\prod_{p\mid q}\left(1-\frac{\chi_1(p)}{p^s}\right)$$

$$= L(s, \chi_1) \prod_{p|q} \left(1 - \frac{\chi_1(p)}{p^s}\right).$$

Il faut voir pour l'avant-dernière égalité que $\chi_1(p) = 0$ si $p|q_1$ alors le deuxième produit vient enlever seulement les termes où $p|q$ et $p \nmid q_1$ multipliés en trop. Les facteurs supplémentaires de l'équation fonctionnelle proviennent de ce fait pour les caractères non primitifs. Le degré de ces fonctions $L$ est 1, $\lambda_f(n) = \chi(n)$ pour tout $n$ et $\alpha_1(p) = \chi(p)$ pour tout $p$. Le conducteur vaut $q$ le modulo du caractère.

Une nouvelle fonction $L$ peut être construite à partir de deux fonctions $L$. Le procédé est appelé la convolution de Rankin-Selberg et est défini ainsi. Soient $L(f, s)$ et $L(g, s)$ deux fonctions $L$ de degré $d$ et $e$, avec paramètres locaux à l'infini $\kappa_i$ et $\nu_j$ et paramètres locaux $\alpha_i(p)$ et $\beta_j(p)$ respectivement. La convolution de Rankin-Selberg de $f$ et $g$, si elle existe, est une fonction $L$ de degré $d \cdot e$ telle que:

- Pour les $p \nmid q(f)q(g)$, les $d \cdot e$ paramètres locaux sont simplement les $d \cdot e$ combinaisons possibles entre les $d$ paramètres de $f$ et les $e$ paramètres de $g$ multipliés entre eux. Précisément, on dénote le facteur du produit d'Euler à un tel $p$

$$L_p(f \otimes g, s) = \prod_{\substack{1 \le i \le d \\ 1 \le j \le e}} (1 - \alpha_i(p)\beta_j(p)p^{-s})^{-1}.$$

- Pour les $p$ ramifiés par $f$ ou $g$, c'est-à-dire, $p|q(f)q(g)$, la restriction sur les paramètres locaux est seulement celle donnée par la définition: la norme est bornée par $p$. Autrement dit, le facteur à $p$ est donné par

$$H_p(s) = \prod_{j=1}^{de} (1 - \gamma_j(p)p^{-s})^{-1} \qquad \text{avec } |\gamma_j(p)| < p.$$

- Le facteur gamma doit respecter

$$\gamma(f \otimes g, s) = \pi^{-des/2} \prod_{i,j} \Gamma\left(\frac{s + \mu_{i,j}}{2}\right),$$

  où $\mathrm{Re}(\mu_{i,j}) \le \mathrm{Re}(\kappa_i + \nu_j)$ et $|\mu_{i,j}| \le |\kappa_i| + |\nu_j|$.

- Le conducteur de $L(f \otimes g, s)$ doit diviser $q(f)^e q(g)^d$ et il doit y avoir un pôle à $s = 1$ si $g = \overline{f}$.

Alors

$$L(f \otimes g, s) = \prod_{p \nmid q(f)q(g)} L_p(f \otimes g, s) \prod_{p|q(f)q(g)} H_p(s)$$

est la convolution de Rankin-Selberg de $f$ et $g$ et a comme dual $L(\overline{f} \otimes \overline{g}, s)$. Il est fort possible qu'il n'existe aucune fonction $L$ satisfaisant ces conditions. Dans ce cas, $f$ et $g$ n'admettent pas de convolution de Rankin-Selberg.

Il est également possible de construire de nouvelles fonctions $L$ en multipliant deux fonctions $L$, $L(f, s)$ et $L(g, s)$. On vérifie facilement que les critères sont respectés et que le degré est la somme des deux degrés, le conducteur $q(f)q(g)$, le facteur gamma $\gamma(f, s)\gamma(g, s)$, la racine

$\epsilon(f)\epsilon(g)$ et le conducteur analytique $\mathfrak{q}(f,s)\mathfrak{q}(g,s)$. Également, si $L(f,s)$ est une fonction $L$ entière, alors pour tout $t \in \mathbb{R}$ la fonction $L(h,s) = L(f,s+it)L(\overline{f},s-it)$ est une fonction $L$ avec conducteur $q(h) = q(f)^2$, facteur gamma $\gamma(h,s) = \gamma(f,s+it)\gamma(f,s-it)$, racine $\epsilon(h) = 1$ et conducteur analytique $\mathfrak{q}(h,s) = \mathfrak{q}(f,s+it)\mathfrak{q}(f,s-it)$.

## 4.2 Les zéros des fonctions $L$

Les fonctions $L$ ont des *zéros triviaux* annulant les pôles du facteur gamma puisque $\Lambda(f,s)$ est holomorphe, sauf possiblement aux points $s = 0, 1$. Les zéros triviaux sont donc situés aux points $-2m-\kappa_j \neq 0$ pour chaque $j$ et tous les entiers $m \geq 0$. Les autres zéros *non triviaux* sont situés dans la bande critique $0 \leq \mathrm{Re}(s) \leq 1$ puisque le produit d'Euler converge absolument pour $\mathrm{Re}(s) > 1$ et l'équation fonctionnelle nous assure qu'il n'y aura pas de zéros dans la région $\mathrm{Re}(s) < 0$ par ce fait. On remarque qu'un zéro $\rho \neq 0$ de $\Lambda(f,s)$ implique que $\overline{\rho}$ est un zéro de $\Lambda(\overline{f},s)$ et l'équation fonctionnelle implique que $1-\overline{\rho}$ est également un zéro de $\Lambda(f,s)$. Cette section donnera une estimation sur $N(T,f)$ la quantité de zéros non triviaux dans la région $0 \leq \mathrm{Im}(s) \leq T$. Le théorème principal de cette section est

$$N(T,f) = \frac{T}{\pi}\log\frac{qT^d}{(2\pi e)^d} + \mathcal{O}(\log \mathfrak{q}(f,iT)).$$

On commence par exprimer $\Lambda(f,s)$ sous une forme où les zéros seront explicites. Des estimations sur $L(f,s)$ et l'approximation de Stirling démontrent que la fonction entière $(s(1-s))^r\Lambda(f,s)$ est d'ordre 1 où $r = r(f)$ l'ordre des pôles ou zéros. Par la théorie des fonctions entières d'ordre finie, on a donc la *factorisation d'Hadamard* suivante pour certaines constantes $a$ et $b$ dépendantes de $f$.

$$(s(s-1))^r\Lambda(f,s) = e^{a+bs}\prod_{\rho\neq 0,1}\left(1-\frac{s}{\rho}\right)e^{s/\rho}$$

où le produit est sur les zéros de $\Lambda(f,s)$ différents de 0 et 1. Ces zéros sont les mêmes que les zéros non triviaux de $L(f,s)$. Par un résultat de la théorie des fonctions entières d'ordre finie et puisque $\Lambda(f,s)$ est d'ordre 1, on a le résultat suivant qui nous sera utile dans la preuve du premier lemme de cette section :

$$\sum_{\rho\neq 0,1}\frac{1}{|\rho|^2} < \infty.$$

La factorisation d'Hadamard nous permet d'obtenir une importante formule en y appliquant la dérivée logarithmique et en utilisant la définition de $\Lambda(f,s)$ :

$$-\frac{L'}{L}(f,s) = \frac{1}{2}\log q + \frac{\gamma'}{\gamma}(f,s) - b + \frac{r}{s} + \frac{r}{s-1} - \sum_{\rho\neq 0,1}\left(\frac{1}{s-\rho}+\frac{1}{\rho}\right). \tag{4.1}$$

Également, une série de Dirichlet est obtenue pour la dérivée logarithmique de $L(f,s)$ à partir du produit d'Euler

$$-\frac{L'}{L}(f,s) = \sum_{n\geq 1}\frac{\Lambda_f(n)}{n^s}.$$

On exprime les facteurs $(1-\alpha_i(p)p^{-s})^{-1}$ du produit d'Euler de $L(f,s)$ en série géométrique et on applique la dérivée logarithmique pour obtenir les informations suivantes sur les coefficients $\Lambda_f(n)$. On a que

$$\Lambda_f(p^k) = \sum_{j=1}^{d}(\alpha_j(p))^k\log p \Rightarrow \Lambda_f(p) = \lambda_f(p)\log p$$

et $\Lambda_f(n) = 0$ si $n$ n'est pas une puissance d'un nombre premier. Montrons maintenant un premier lemme en vue de prouver notre théorème principal.

**Lemme 4.1.** *La fonction $m(T, f)$ comptant le nombre de zéros $\rho = \beta + i\gamma$ tels que $|\gamma - T| \leq 1$ est bornée par*

$$m(T, f) \ll \log \mathfrak{q}(f, iT).$$

*Preuve.* L'idée sera de prendre la partie réelle de (4.1) et de borner chaque terme sauf la somme sur $1/(s-\rho)$ afin d'obtenir une estimation sur celle-ci. Fixons $T \geq 2$ et $s = 3 + iT$. On borne facilement les coefficients $|\Lambda_f(n)| \leq dn \log n$ afin d'obtenir

$$\left|\frac{L'}{L}(f, s)\right| \leq \sum_{n\geq 1} \frac{dn \log n}{|n^{3+iT}|} \leq d\frac{\zeta'(3)}{\zeta(3)} \ll \log \mathfrak{q}(f, s)$$

car $3^d \leq \mathfrak{q}_\infty(s)$ et $\log(3^d) = d \log(3)$. On borne ensuite directement par Stirling

$$\frac{\gamma'}{\gamma}(f, s) \ll \log \mathfrak{q}(f, s).$$

Trivialement,

$$\frac{1}{2}\log q + \frac{r}{s} + \frac{r}{s-1} \ll \log \mathfrak{q}(f, s).$$

Alors, si l'on prend la partie réelle de (4.1), nos bornes nous montrent que la somme sur $\mathrm{Re}(1/(s-\rho) + 1/\rho)$ est absolument convergente. D'ailleurs, on affirme que

$$\mathrm{Re}(b) = -\sum_\rho \mathrm{Re}(1/\rho).$$

On peut donc séparer la somme de (4.1) en deux et la constante $b$ annulera la somme sur $1/\rho$ lorsqu'on prendra la partie réelle. La preuve de cette égalité sera donnée après notre argument. On se retrouve donc avec

$$\sum_\rho \mathrm{Re}\left(\frac{1}{s-\rho}\right) \ll \log \mathfrak{q}(f, s).$$

Par notre choix de $s$, un calcul direct montre que

$$\frac{2}{9 + (T-\gamma)^2} < \mathrm{Re}\left(\frac{1}{s-\rho}\right) < \frac{3}{4 + (T-\gamma)^2},$$

car un zéro $\rho = \beta + i\gamma$ aura $0 \leq \beta \leq 1$. Alors,

$$\mathrm{Re}\left(\frac{1}{s-\rho}\right) \asymp \frac{1}{1 + (T-\gamma)^2}$$

$$\Rightarrow \sum_\rho \frac{1}{1 + (T-\gamma)^2} \ll \log \mathfrak{q}(f, s).$$

Dans cette somme, tous les zéros tels que $|\gamma - T| \leq 1$ contribuent au moins de $1/2$. Aussi, par notre choix de $s$ et la restriction $T \geq 2$, nous avons que $|s| \asymp |iT|$ alors

$$m(T, f) \ll \log \mathfrak{q}(f, iT).$$

Il ne reste qu'à prouver l'égalité reliant la constante $b$ et la somme sur $1/\rho$. On part de l'équation fonctionnelle et on retire les pôles

$$(s(1-s))^r \Lambda(f,s) = \epsilon(f)(s(1-s))^r \Lambda(\overline{f}, 1-s)$$

On applique la factorisation de Hadamard des deux côtés et on calcule la dérivée logarithmique pour obtenir

$$b(f) + b(\overline{f}) = -\sum_{\rho \neq 0,1} \left( \frac{1}{s-\rho} + \frac{1}{1-s-\overline{\rho}} + \frac{1}{\rho} + \frac{1}{\overline{\rho}} \right)$$

car si $\rho$ sont les zéros de $\Lambda(f,s)$, alors $\overline{\rho}$ sont les zéros de $\Lambda(\overline{f},s)$. On a que $b(\overline{f}) = \overline{b(f)}$ par la symétrie $\Lambda(\overline{f},s) = \overline{\Lambda(f,\overline{s})}$ qui découle de $\lambda_{\overline{f}}(n) = \overline{\lambda_f(n)}$. Donc

$$b(f) + b(\overline{f}) = 2 \cdot \mathrm{Re}(b(f)).$$

On peut briser la somme en deux, car les sommes

$$\sum_{\rho \neq 0,1} \left( \frac{1}{s-\rho} + \frac{-1}{s-(1-\overline{\rho})} \right) \quad \text{et} \quad \sum_{\rho \neq 0,1} \left( \frac{1}{\rho} + \frac{1}{\overline{\rho}} \right)$$

sont absolument convergentes, car en mettant sur un dénominateur commun,

$$\frac{1}{s-\rho} + \frac{1}{1-s-\overline{\rho}} \ll_s \frac{1}{|\rho|^2} \quad \text{et} \quad \frac{1}{\rho} + \frac{1}{\overline{\rho}} \ll \frac{1}{|\rho|^2}$$

où la constante implicite de la première borne dépend de $s$. Maintenant, la première somme vaut 0 puisque les termes s'annulent entre eux, car $\rho$ et $1-\overline{\rho}$ sont tous deux zéros de $\Lambda(f,s)$ et on somme sur tous les zéros. L'intérieur de la somme restante vaut $2 \cdot \mathrm{Re}(1/\rho)$, on a donc le résultat qui était désiré afin de conclure la preuve. □

Un dernier résultat utile avant d'aborder la preuve du théorème principal de cette section est obtenu ainsi. On commence avec l'équation triviale suivante où $s = \sigma + it$ et $-1/2 \leq \sigma \leq 2$,

$$-\frac{L'}{L}(f,s) = -\frac{L'}{L}(f,s) + \frac{L'}{L}(f,3+it) - \frac{L'}{L}(f,3+it).$$

Dans le côté droit de l'équation, on utilise (4.1) pour les deux premiers termes et on borne le troisième terme par $\log \mathfrak{q}(f,s)$ comme dans la preuve précédente. Plusieurs termes sont absorbés par le terme d'erreur et la somme des $1/\rho$ disparait, car elle est sommée et soustraite. On obtient

$$-\frac{L'}{L}(f,s) = \frac{\gamma'}{\gamma}(f,s) + \frac{r}{s} + \frac{r}{s-1} - \sum_{\rho} \left( \frac{1}{s-\rho} - \frac{1}{3+it-\rho} \right) + \mathcal{O}(\log \mathfrak{q}(f,s)).$$

Tous les termes de la somme tels que $|s-\rho| \geq 1$ sont absorbés par l'erreur, car dans ce cas

$$\left| \frac{1}{s-\rho} - \frac{1}{3+it-\rho} \right| \leq \frac{3}{1+(t-\gamma)^2}$$

et dans la preuve précédente on a borné une telle somme sur tous les $\rho$ par $\log \mathfrak{q}(f,s)$. Ensuite, Stirling nous permet de remplacer le terme en $\gamma$ par une somme sur les $\kappa_j$, cela à pour avantage de nous donner une formule qui sera reliée aux zéros triviaux.

$$\frac{\gamma'}{\gamma}(f,s) = -\sum_{|s+\kappa_j|<1} \frac{1}{s+\kappa_j} + \mathcal{O}(\log \mathfrak{q}(f,s))$$

Finalement, on peut retirer dans la somme le deuxième terme, car il sera absorbé par l'erreur. Comme on somme seulement les zéros tels que $|s-\rho|<1$, il y aura $m(t,f) \ll \log \mathfrak{q}(f,s)$ termes, et comme les zéros sont situés à l'intérieur de la bande critique, $1/(3+it-\rho) \leq 1/2$. Le résultat est donc le suivant pour tout $s$ à l'intérieur de la bande $-1/2 \leq \sigma \leq 2$,

$$\frac{L'}{L}(f,s) + \frac{r}{s} + \frac{r}{s-1} - \sum_{|s+\kappa_j|<1} \frac{1}{s+\kappa_j} - \sum_{|s-\rho|<1} \frac{1}{s-\rho} \ll \log \mathfrak{q}(f,s). \tag{4.2}$$

Ce résultat important indique que la dérivée logarithmique d'une fonction $L$ croit très lentement si on retire les pôles et que ces derniers ont une influence très locale. On utilisera ce résultat dans les prochaines sections également.

Prouvons maintenant notre théorème principal

$$N(T,f) = \frac{T}{\pi} \log \frac{qT^d}{(2\pi e)^d} + \mathcal{O}(\log \mathfrak{q}(f,iT)).$$

Notons que ce résultat est assez précis étant donné le logarithme de l'erreur. Ce résultat provient du théorème de Cauchy appliqué à la dérivée logarithmique de $\Lambda(f,s)$ qui reliera le nombre de zéros avec une intégrale. L'évaluation de l'intégrale donnera l'estimation voulue.

On dénote notre région d'intégration par $\mathcal{C}$ le rectangle ayant comme sommets $3-i\delta, 3+iT, -2+iT, -2-i\delta$ pour un $\delta>0$ arbitrairement petit. Ce $\delta$ sert à éviter les possibles pôles à $s=0$ et $s=1$. On suppose pour la même raison que $T$ ne coïncide pas avec la partie imaginaire d'un zéro, le résultat pour tout $T$ y découlera puisque l'estimation $m(T,f) \ll \log \mathfrak{q}(f,iT)$ implique que les zéros ne sont pas denses.

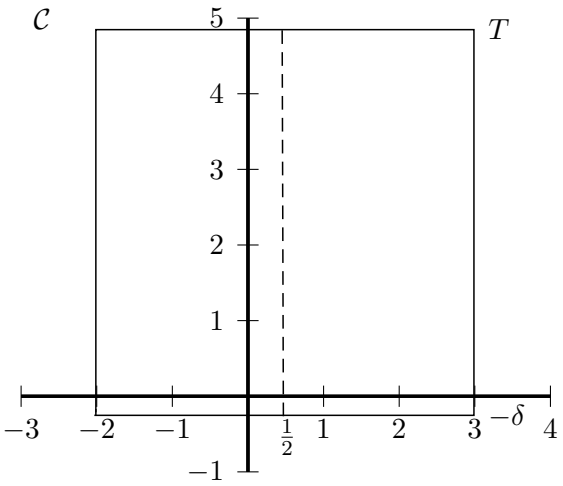
Les zéros triviaux de $L(f,s)$ peuvent possiblement se trouver à l'intérieur de la bande critique et ne seront pas comptés puisqu'on travaille avec $\Lambda$. Il y en aura au plus $d$ alors cette différence sera absorbée par l'erreur.

Notre région d'intégration ne compte que les zéros tels que $\gamma \geq 0$. Les zéros avec partie imaginaire négative sont les zéros avec partie imaginaire positive de $\Lambda(\overline{f},s)$, mais conjugués. Comme la fonction $L$ duale a le même degré, l'estimation pour le nombre de zéros sera la même. On doublera alors notre résultat pour compter les zéros tels que $|\gamma| \leq T$.

Appliquons maintenant le théorème de Cauchy à la dérivée logarithmique de $\Lambda$ afin d'obtenir des pôles aux positions des zéros. La quantité $I(T)$ sera le nombre de zéros à l'intérieur du rectangle.

$$I(T) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{\Lambda'}{\Lambda}(f,s)ds$$

L'intégrale de la dérivée logarithmique de $\Lambda(f,s)$ est $\log \Lambda(f,s)$. En divisant le rectangle en deux aux points $1/2-i\delta$ et $1/2+iT$, la symétrie provenant de l'équation fonctionnelle montre que l'intégrale sur chacune des deux sections aura la même valeur, car $\Lambda'/\Lambda(f,s) = \overline{\Lambda'/\Lambda(f,\overline{1-s})}$ par $\Lambda(\bar{f},s) = \overline{\Lambda(f,\bar{s})}$.




Concentrons-nous donc seulement sur la partie de droite. Un deuxième avantage à la séparation du rectangle est que le logarithme n'est pas bien défini si l'on tourne autour de l'origine. Le calcul de l'intégrale sera alors $\log \Lambda(f, 1/2 + iT) - \log \Lambda(f, 1/2 - i\delta)$.

Pour simplifier les calculs, on remarque que $I(T)$ est un nombre entier, alors l'égalité est encore vraie en prenant la partie réelle de l'équation. Comme on divise par $i$ au côté droit de l'équation, on s'intéresse à la partie imaginaire du log, c'est-à-dire, l'argument de $\Lambda(f, s)$. On calculera séparément la variation de l'argument de $s = 1/2 - i\delta$ à $s = 1/2 + iT$ pour chaque facteur composant $\Lambda$ par additivité de l'argument. On se rappelle que

$$\Lambda(f, s) = \pi^{-ds/2} q^{s/2} \prod_{j=1}^{d} \Gamma\Big(\frac{s + \kappa_j}{2}\Big) L(f, s).$$

Le facteur $\pi^{-ds/2} q^{s/2}$ contribuera de

$$\frac{T}{2\pi} \log \frac{q}{\pi^d} + \mathcal{O}(1),$$

l'erreur provenant de $s = 1/2 - i\delta$. Pour le facteur gamma, Stirling nous donne

$$\operatorname{Im} \log \Gamma(s) = t \log t - t + \mathcal{O}(1) \qquad t \geq 1,$$

alors la contribution sera

$$\frac{1}{\pi}\Big(\frac{dT}{2} \log \frac{T}{2} - \frac{dT}{2}\Big) + \mathcal{O}(\log \mathfrak{q}(f, iT)),$$

où l'erreur provient des $\kappa_j$.

Pour le facteur $L(f, s)$, on estime l'intégrale sur le segment vertical et les deux segments horizontaux. Pour celui vertical, on borne

$$|\log L(f, s)| = \Big| \sum_n \frac{\Lambda_f(n)}{\log n} n^{-s} \Big| \leq \sum_n \Big| \frac{dn \log n}{\log n} n^{-3-it} \Big| \ll \log \mathfrak{q}(f)$$

et donc l'intégrale sur cette partie est $\ll \log \mathfrak{q}(f)$ puisque la borne ne dépend pas de $T$. L'estimation (4.2) et la borne $m(T, f) \ll \log \mathfrak{q}(f, iT)$ nous assure que les parties horizontales

auront une contribution qui sera absorbée par l'erreur.

On obtient le résultat en additionnant nos estimations et en multipliant par 4 puisqu'on a calculé la moitié de $I(T)$ et qu'on a compté que les zéros avec partie imaginaire positive seulement.

## 4.3 Région sans zéro

Nous avons vu avec $\zeta$ que la force des estimations dépendait de la profondeur avec laquelle la région d'intégration traversait la bande critique. Il est donc très utile de calculer une région sans zéro pour les fonctions $L$ en général. Cette section prouvera sous certaines conditions qu'un zéro non trivial $\rho = \beta + i\gamma$ de $L(f,s)$ respectera

$$\beta < 1 - \frac{c}{d^4 \log(\mathfrak{q}(f)(|\gamma|+3))}$$

pour $c > 0$ une certaine constante. On commence par obtenir un résultat sur le nombre maximal de zéros réels proche de $s = 1$ d'une fonction $L$ respectant certains critères.

**Lemme 4.2.** *Soit $L(f,s)$ une fonction $L$ de degré $d$ avec $Re(\Lambda_f(n)) \geq 0$ pour $(n, q(f)) = 1$ et $|\alpha_j(p)| \leq p/2$ pour les nombres premiers ramifiés. Alors $L(f,1) \neq 0$. Également, si $r$ dénote l'ordre du pôle à $s = 1$, il existe $c > 0$ tel que $L(f,s)$ a au plus $r$ zéros réels dans l'intervalle*

$$s \geq 1 - \frac{c}{d(r+1)\log \mathfrak{q}(f)}.$$

*Preuve.* On prouve ceci à partir de la borne (4.2). Soit $1 < s = \sigma \leq 2$ réel, alors en prenant la partie réelle et en retirant quelques termes on obtient l'inégalité

$$\sum_{|s-\rho|<1} \mathrm{Re}\frac{1}{s-\rho} < \frac{r}{\sigma-1} + \mathrm{Re}\frac{L'}{L}(f,\sigma) + \mathcal{O}(\log \mathfrak{q}(f)).$$

L'inégalité restera vraie en retirant des $\rho$ de la somme, car

$$\mathrm{Re}\frac{1}{\sigma-\rho} = \frac{\sigma-\beta}{|\sigma-\rho|^2} \geq 0.$$

On gardera donc que les zéros d'intérêts $\beta_j$ réels tels que $1/2 \leq \beta_j \leq 1$. La prochaine étape est de se débarrasser du terme en $L'/L$. Les termes non ramifiés de sa série de Dirichlet seront tous négatifs par notre hypothèse, car on travaille sur $L'/L$ ici et non sur $-L'/L$. L'inégalité reste donc vraie en les retirant. Pour ce qui est des termes ramifiés, on les obtient par la dérivée logarithmique du produit d'Euler de $L(s,f)$ sur les nombres premiers ramifiés. On borne donc directement en utilisant notre hypothèse sur les nombres premiers ramifiés

$$\mathrm{Re}\sum_{\substack{p|q(f)\\k\geq 1}} \frac{\Lambda_f(p^k)}{p^{k\sigma}} \leq \Big|\sum_{p|q(f)}\sum_{1\leq j\leq d} \frac{\alpha_j(p)p^{-\sigma}\log p}{1-\alpha_j(p)p^{-\sigma}}\Big| \leq d\sum_{p|q(f)} \log p \leq d\log q(f).$$

On a maintenant

$$\sum_j \frac{1}{\sigma-\beta_j} < \frac{r}{\sigma-1} + \mathcal{O}(d\log \mathfrak{q}(f)).$$

Si $r < 0$, c'est-à-dire, $L(f, 1) = 0$, alors il y a un $\beta_j = 1$. En envoyant $r/(\sigma-1)$ de l'autre côté de l'inégalité, on aura une contradiction en faisant tendre $\sigma$ vers 1 car on dépassera certainement notre borne qui elle ne dépend pas de $\sigma$. Alors $r \geq 0$.

Pour le deuxième résultat du lemme, supposons qu'il y a $n$ zéros réels dans l'intervalle $1 - c/(d(r+1)\log \mathfrak{q}(f)) < \beta_j < 1$. Alors en posant $\sigma = 1 + 2c/(d \log \mathfrak{q}(f))$ dans l'inégalité ci-haute on obtient

$$\frac{nd \log \mathfrak{q}(f)}{2c + c/(r+1)} < \left(\frac{r}{2c} + \mathcal{O}(1)\right) d \log \mathfrak{q}(f)$$
$$\Rightarrow n < r + r/2(r+1) + \mathcal{O}(c).$$

Comme la constante implicite est absolue, un $c$ assez petit implique $n \leq r$. □

Nous sommes maintenant en mesure de calculer une région sans zéro. L'idée est de choisir un zéro non trivial $\rho = \beta + it$ d'une fonction $L$ et de construire une nouvelle fonction $L$ qui transformera ce zéro en un zéro réel $\rho = \beta$ d'ordre supérieur au pôle ce qui nous permettra d'utiliser le résultat précédent afin d'obtenir des informations sur $\beta$.

L'argument se servira de $L(f \otimes f, s)$ et de $L(f \otimes \overline{f})$, il est donc nécessaire que ces convolutions existent. La première convolution doit être entière si $f \neq \overline{f}$ et la deuxième doit avoir un pôle simple à $s = 1$. Il faut également avoir aux nombres premiers ramifiés $|\alpha_j(p)|^2 \leq p/2$.

On fixe $t > 0$ et on considère la fonction $L$

$$L(g, s) = \zeta(s)L(f, s + it)^2 L(\overline{f}, s - it)^2 L(f \otimes f, s + 2it)L(\overline{f} \otimes \overline{f}, s - 2it)L(f \otimes \overline{f}, s)^2$$

de degré $(1 + 2d)^2$. Ce choix astucieux de $L(g, s)$ nous permettra d'appliquer le Lemme 4.2. En choisissant $t = \gamma \neq 0$ pour un zéro $\rho = \beta + i\gamma$ de $L(f, s)$ tel que $\beta \geq 1/2$, $L(g, s)$ aura un pôle d'ordre $\leq 3$ à $s = 1$ et $\beta$ sera un zéro réel de $L(g, s)$ d'ordre $\geq 4$. Le pôle provient des facteurs $\zeta(s)L(f \otimes \overline{f}, s)^2$ et le zéro provient des facteurs $L(f, s + it)^2 L(\overline{f}, s - it)^2$. L'ordre du pôle ne peut qu'être réduit et l'ordre du zéro ne peut qu'être augmenté par les possibles zéros des autres facteurs. L'utilité des facteurs $L(f \otimes f, s + 2it)L(\overline{f} \otimes \overline{f}, s - 2it)$ est pour s'assurer que $\mathrm{Re}(\Lambda_g(n)) \geq 0$ pour $(n, q(f)) = 1$ qui était un critère du résultat précédent comme on le prouvera ci-après.

Alors, comme l'ordre du pôle est strictement plus petit que l'ordre du zéro, le résultat précédent nous dit que

$$\beta < 1 - \frac{c}{d^2 \log \mathfrak{q}(g)}$$

pour un certain $c > 0$. On aimerait avoir une relation avec $\mathfrak{q}(f)$ plutôt que $\mathfrak{q}(g)$, les propriétés multiplicatives du conducteur analytique nous donnent

$$\mathfrak{q}(g) \leq \mathfrak{q}(f, it)^4 \mathfrak{q}(f \otimes f, 2it)^2 \mathfrak{q}(f \otimes \overline{f})^2.$$

La borne triviale $\mathfrak{q}(f, s) \leq \mathfrak{q}(f)(|s| + 3)^d$ et le fait que $q(f \otimes f)|q(f)^{2d}$ implique

$$\mathfrak{q}(g) \leq \mathfrak{q}(f)^{4+8d}(|t| + 3)^{6d^2}.$$

Alors

$$\beta < 1 - \frac{c'}{d^4 \log(\mathfrak{q}(f)(|t| + 3))}$$

pour une certaine constante $c' > 0$ absolue. Ce résultat est la région sans zéro de notre fonction $L$ pour les zéros qui ne sont pas réels.

Montrons maintenant que les coefficients de la série de Dirichlet sont positifs aux puissances de nombres premiers non ramifiés. Un facteur du produit d'Euler de $L(g,s)$ à un tel $p$ est de la forme

$$(1-p^{-s})^{-1}\Big(\prod_j(1-\alpha_j p^{it}p^{-s})^{-1}\Big)^2\Big(\prod_j(1-\overline{\alpha_j}p^{-it}p^{-s})^{-1}\Big)^2\prod_j\prod_k(1-\alpha_j\alpha_k p^{2it}p^{-s})^{-1}$$

$$\prod_j\prod_k(1-\overline{\alpha_j\alpha_k}p^{-2it}p^{-s})^{-1}\Big(\prod_j\prod_k(1-\alpha_j\overline{\alpha_k}p^{-s})^{-1}\Big)^2$$

où $\alpha_j = \alpha_j(p)$ sont les $d$ paramètres locaux de $L(f,s)$ à $p$. On rappelle que les coefficients $\Lambda_g(n)$ ne valent pas zéro seulement aux puissances de nombres premiers et que

$$\Lambda_g(p^k) = (\log p)\sum_{j=1}^{(1+2d)^2}(\hat{\alpha_j}(p))^k,$$

où les $\hat{\alpha_j}$ sont les paramètres locaux de $L(g,s)$ explicités ci-haut dans le facteur du produit d'Euler à $p$. Cette somme se factorise ainsi

$$\Big|1+\sum_j\alpha_j^k p^{kit}+\sum_j\bar{\alpha_j}^k p^{-kit}\Big|^2$$

alors $\Lambda_g(p^k)$ est inévitablement positif pour tout $p$ non ramifié et tout $k\in\mathbb{N}$.

Il reste à traiter le cas où $t=0$. L'argument est le même si $\overline{f}\neq f$. Par contre, si $f=\overline{f}$, le pôle de $L(s,g)$ à $s=1$ sera d'ordre 5, il est donc possible qu'un seul zéro réel de $L(s,f)$ ne respecte pas la région sans zéro puisque chaque zéro réel de $L(s,f)$ est un zéro réel de $L(s,g)$ d'ordre au moins 4.

## 4.4 Formule explicite

Lors de la preuve du théorème des nombres premiers, nous avons cherché à éviter les zéros de $\zeta(s)$. Si l'on capture ces zéros dans notre région d'intégration, on obtiendra une formule reliant une somme sur les puissances de nombres premiers et une somme sur les zéros d'une fonction $L$. Cette formule est dite explicite.

On commence encore une fois avec une fonction $\phi:\mathbb{R}\to\mathbb{C}$ de classe $C^\infty$ à support compact afin d'utiliser l'inversion de Mellin et de considérer seulement les nombres premiers jusqu'à $x$ :

$$\sum_n\Lambda_f(n)\phi(n) = \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} -\frac{L'}{L}(f,s)\hat{\phi}(s)ds.$$

L'intégration complexe est effectuée sur la dérivée logarithmique afin de capturer les pôles aux positions des zéros de $L(s,f)$ lorsqu'on appliquera le théorème de Cauchy. Pour intégrer sur un chemin fermé, on considère le rectangle aux sommets $2-iT$, $2+iT$, $-c+iT$ et $-c-iT$ pour $T,c>0$, où $T$ ne correspond pas à l'ordonné d'un zéro et $c$ est suffisamment près de 0. On fera

tendre $T$ vers l'infini pour que le côté droit du rectangle soit notre somme ci-haute. Tous les zéros non triviaux comptés avec multiplicité seront à l'intérieur de notre rectangle avec résidu $-\hat{\phi}(\rho)$, ainsi qu'un pôle de résidu $r\hat{\phi}(1) = r\int_0^\infty \phi(x)dx$ à $s = 1$. Le théorème de Cauchy nous donne

$$\frac{1}{2\pi i}\Big(\int_{2-iT}^{2+iT} + \int_{2+iT}^{-c+iT} + \int_{-c+iT}^{-c-iT} + \int_{-c-iT}^{2-iT}\Big) = -\sum_\rho \hat{\phi}(\rho) + r\hat{\phi}(1)$$

où l'on intègre $-L'/L(f,s)\hat{\phi}(s)$. Les deux intégrales sur les côtés horizontaux tendront vers zéro, car par intégration par partie $\hat{\phi}(s) \ll 1/|s|^2$, la constante implicite dépendante du choix de $\phi$, tandis que $L'/L(s)$ est bornée par une puissance de log par notre estimation (4.2) et du fait qu'il existe un choix possible de $T$ pour chaque intervalle de longueur 1 nous assurant une distance $\gg 1/\log \mathfrak{q}(f, iT)$ d'un zéro par notre estimation sur $m(T, f)$. On utilise une telle suite pour faire tendre $T$ vers l'infini.

L'intégrale de $-c - iT$ à $-c + iT$ est traitée en appliquant l'équation fonctionnelle

$$-\frac{L'}{L}(f,s) = \log q(f) + \frac{\gamma'}{\gamma}(f,s) + \frac{\gamma'}{\gamma}(\bar{f}, 1-s) + \frac{L'}{L}(\bar{f}, 1-s).$$

On a directement par inversion de Mellin que

$$\int_{-c-i\infty}^{-c+i\infty} \log q(f)\hat{\phi}(s) = \log q(f)\phi(1).$$

Pour les facteurs gamma, il est possible de retirer la dépendance à $c$ de la ligne d'intégration en la déplaçant à $\sigma = 1/2$. Les pôles de $\frac{\gamma'}{\gamma}(f,s)$ s'annuleront avec ceux de $\frac{\gamma'}{\gamma}(\bar{f}, 1-s)$. Finalement, la raison pour laquelle nous avons appliqué l'équation fonctionnelle était pour se retrouver dans la région de convergence de la série de Dirichlet de $-L'/L$. On revient sur nos pas en effectuant les étapes inverses de notre point de départ

$$\begin{aligned}\frac{1}{2\pi i}\int_{-c-i\infty}^{-c+i\infty} \frac{L'}{L}(\bar{f}, 1-s)\hat{\phi}(s)ds &= \frac{-1}{2\pi i}\int_{1+c-i\infty}^{1+c+i\infty} -\frac{L'}{L}(\bar{f}, s)\hat{\phi}(1-s)ds \\ &= \sum_{n\geq 1}\Lambda_{\bar{f}}(n)\frac{-1}{2\pi i}\int_{1+c-i\infty}^{1+c+i\infty}\hat{\phi}(1-s)n^{-s}ds \\ &= -\sum_{n\geq 1}\overline{\Lambda_f}(n)\frac{\phi(1/n)}{n},\end{aligned}$$

où nos bornes ci-hautes montrent la convergence absolue de l'intégrale qui justifie l'interversion avec la somme. En rassemblant le tout, on obtient la formule explicite

$$\sum_{n\geq 1}\left(\Lambda_f(n)\phi(n) + \overline{\Lambda_f}(n)\frac{\phi(1/n)}{n}\right) =$$

$$\phi(1)\log q(f) + r\int_0^\infty \phi(x)dx + \frac{1}{2\pi i}\int_{1/2-i\infty}^{1/2+i\infty}\left(\frac{\gamma'}{\gamma}(f,s) + \frac{\gamma'}{\gamma}(\bar{f}, 1-s)\right)\hat{\phi}(s)ds - \sum_\rho \hat{\phi}(\rho)$$

où la somme est sur les zéros triviaux et non triviaux à l'intérieur de la bande critique comptés avec multiplicité.

## 4.5 Théorème des nombres premiers

Nous allons calculer dans cette section une estimation asymptotique sur

$$\psi(f,x) := \sum_{n\leq x} \Lambda_f(n).$$

Le résultat est

$$\psi(f,x) = rx - \frac{x^{\beta_f}}{\beta_f} + \mathcal{O}\Big(\sqrt{\mathfrak{q}(f)}\, x \exp\Big(-\frac{c}{2d^4}\sqrt{\log x}\Big)\Big)$$

où $\beta_f$ est la partie réelle du possible zéro exceptionnel. La preuve sera pratiquement celle donnée au chapitre sur $\zeta$, mais cette fois-ci en toute généralité. C'est pourquoi le terme d'erreur dépend des paramètres de la fonction $L$.

On débute avec une estimation sur une somme entre $x$ et $x+y$ de $|\Lambda_f(n)|$. On a d'abord

$$\sum_{n\leq x} |\Lambda_f(n)|^2 \ll xd^2 \log^2(x\mathfrak{q}(f))$$

où la constante est absolue. Ce résultat découle de (4.2) appliqué à la convolution de Rankin-Selberg $L(f\otimes \bar{f}, s)$. Étant donné qu'on se servira de la région sans zéro, on assume que cette convolution existe avec un pôle simple à $s=1$. Alors, pour $1<\sigma\leq 2$, la borne (4.2) donne

$$-\frac{L'}{L}(f\otimes\bar{f},\sigma) \ll \frac{d^2}{\sigma-1}\log \mathfrak{q}(f)$$

en utilisant comme dans la section sur la région sans zéro la borne sur le conducteur analytique

$$\mathfrak{q}(f\otimes\bar{f},s) \leq \mathfrak{q}(f)^{2d}(|s|+3)^{d^2}.$$

On a que

$$-\frac{L'}{L}(f\otimes\bar{f},\sigma) = \sum_{n\geq 1}\frac{\Lambda_{f\otimes\bar{f}}(n)}{n^\sigma}.$$

Aux nombres premiers non ramifiés, $\Lambda_{f\otimes\bar{f}}(n) = |\Lambda_f(n)|^2$ par la formule exprimant $\Lambda_{f\otimes\bar{f}}(n)$ en somme des paramètres locaux à la puissance $k$. Aux nombres premiers ramifiés, il est aussi possible de remplacer $\Lambda_{f\otimes\bar{f}}(n)$ par $|\Lambda_f(n)|^2$ tout en gardant la même borne puisque la différence sera absorbée. On le voit ainsi par le produit d'Euler en se rappelant l'hypothèse $|\alpha_j(p)|^2\leq p/2$ aux paramètres locaux de $L(f,s)$ que l'on suppose pour la région sans zéro

$$\Big|\sum_{\substack{p|q(f)\\k\geq 1}}\frac{|\Lambda_f(p^k)|^2}{p^{k\sigma}}\Big| = \Big|\sum_{p|q(f)}\sum_{1\leq j\leq d}\sum_{1\leq k\leq d}\frac{\alpha_j(p)\overline{\alpha}_k(p)p^{-\sigma}\log p}{1-\alpha_j(p)\overline{\alpha}_k(p)p^{-\sigma}}\Big| \leq d^2\log q(f).$$

De même pour les termes que l'on remplace,

$$\Big|\sum_{\substack{p|q(f\otimes\bar{f})\\k\geq 1}}\frac{\Lambda_{f\otimes\bar{f}}(p^k)}{p^{k\sigma}}\Big| \leq d^2\log q(f).$$

En posant $\sigma = 1+(\log 3x)^{-1}$ et en multipliant par $x$,

$$\sum_{n\geq 1}\frac{x|\Lambda_f(n)|^2}{n^{1+(\log 3x)^{-1}}} \ll xd^2\log^2(x\mathfrak{q}(f)).$$

On remarque que pour tout $x \geq 1$ et $1 \leq n \leq x$

$$\frac{1}{e} \leq \frac{x}{n^{1+(\log 3x)^{-1}}},$$

car

$$\lim_{x\to\infty} \frac{x}{x^{1+(\log 3x)^{-1}}} = \frac{1}{e}.$$

Alors

$$\sum_{n\leq x} |\Lambda_f(n)|^2 \ll xd^2 \log^2(x\mathfrak{q}(f)).$$

On obtient finalement la borne nécessaire au théorème des nombres premiers

$$\sum_{x<n\leq x+y} |\Lambda_f(n)| \ll d\sqrt{xy}\log(x\mathfrak{q}(f))$$

pour $x \geq y \geq 1$ où la constante implicite est absolue. Cette borne est obtenue simplement en appliquant l'inégalité de Cauchy

$$\left(\sum_{x<n\leq x+y} |\Lambda_f(n)|\right)^2 \leq \left(\sum_{x<n\leq x+y} |\Lambda_f(n)|^2\right)\left(\sum_{x\leq n\leq x+y} 1^2\right).$$

Nous sommes maintenant en mesure de prouver le théorème des nombres premiers. On débute avec la fonction $\phi$ définie à la Section 3.4 pour obtenir

$$\psi(f,x) = \sum_{n\geq 1} \Lambda_f(n)\phi(n) + \mathcal{O}(d\sqrt{xy}\log(x\mathfrak{q}(f))).$$

On utilise l'inversion de Mellin

$$\sum_{n\geq 1} \Lambda_f(n)\phi(n) = \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} -\frac{L'}{L}(f,s)\hat{\phi}(s)ds.$$

La courbe $\mathcal{Z}$ d'intégration sera la frontière de la région

$$\sigma \geq 1 - \frac{c_1}{d^4\log(\mathfrak{q}(f)(|t|+3))}$$

où $c_1$ dépend de l'emplacement du zéro exceptionnel, s'il existe, puisque l'on souhaite s'assurer une distance d'au moins $c/6d^2\log(\mathfrak{q}(f)(|t|+3))$ de tous les zéros afin de borner $-L'/L$ sur $\mathcal{Z}$. La constante $c$ est celle provenant de la région sans zéro. Le segment où le zéro exceptionnel se situe est

$$1 - \frac{c}{d^4\log(3\mathfrak{q}(f))} \leq \beta_f < 1.$$

On divise également ce segment en deux et si le zéro exceptionnel se situe dans la moitié de droite, on pose $c_1 = 2c/3$. On capturera donc ce zéro dans notre région. S'il se trouve dans la moitié de gauche, ou s'il n'existe pas, on pose $c_1 = c/3$.

Par la borne sur $\hat{\phi}(s)$ donnée à la Section 3.4

$$\hat{\phi}(s) \ll \frac{x^\sigma}{|s|}\min\left(1, \frac{x}{|s|y}\right), \qquad \frac{1}{2} \leq \sigma \leq 2$$

et par une borne qu'on donnera immédiatement sur $-L'/L$, l'intégrale des parties horizontales fermant la région tendront vers zéro, alors par le théorème de Cauchy

$$\sum_{n\geq 1}\Lambda_f(n)\phi(n) = r\hat{\phi}(1) - \hat{\phi}(\beta_f) + \frac{1}{2\pi i}\int_{\mathcal{Z}} -\frac{L'}{L}(f,s)\hat{\phi}(s)ds$$

où le terme en $\beta_f$ est présent seulement si le zéro exceptionnel est inclus dans notre région.

Par (4.2), on borne pour $s \in \mathcal{Z}$

$$-\frac{L'}{L}(f,s) \ll d^4\log^2(\mathfrak{q}(f)(|t|+3)).$$

Avec cette borne et celle sur $\hat{\phi}$, on estime l'intégrale sur $\mathcal{Z}$

$$\int_{\mathcal{Z}} -\frac{L'}{L}(f,s)\hat{\phi}(s)ds \ll d^4 x^{\sigma(T)}\log^3(\mathfrak{q}(f)T)$$

où $T = x/y$ et $\sigma(T) = 1 - c_1/d^4\log(\mathfrak{q}(f)T)$, les détails étant exactement les mêmes qu'à la Section 3.4 pour $\zeta$. On évalue $\hat{\phi}(\beta_f)$ ainsi

$$\hat{\phi}(\beta_f) = \int_0^x z^{\beta_f - 1}dz + \mathcal{O}(y) = \frac{x_f^\beta}{\beta_f} + \mathcal{O}(y).$$

De même pour montrer que $\hat{\phi}(1) = x + \mathcal{O}(y)$.

En rassemblant le tout et en posant $T = \exp((1/3)\sqrt{\log x})$ par notre liberté de choisir $y$ entre 1 et $x$, on obtient le théorème des nombres premiers

$$\psi(f,x) = rx - \frac{x^{\beta_f}}{\beta_f} + \mathcal{O}\bigg(x\exp\bigg(\frac{-cd^{-4}\log x}{\sqrt{\log x} + 3\log\mathfrak{q}(f)}\bigg)(d\log x\mathfrak{q}(f))^4\bigg), \qquad x \geq 1.$$

La constante implicite est absolue, et le terme en $\beta_f$ ne sera pas présent si le zéro exceptionnel n'existe pas. Bien que l'on ne capturait pas ce zéro dans la région d'intégration s'il était présent dans la moitié gauche du segment, le terme en $\beta_f$ sera absorbé par l'erreur si c'est le cas. On peut donc dire que le terme en $\beta_f$ est présent dans la formule seulement si le zéro exceptionnel existe, sans se soucier de sa position.

Un fait important à noter est que si

$$x < q(f)^{4c^{-1}d^4\log(d\log q(f))},$$

alors les termes principaux seront absorbés par l'erreur, on doit donc avoir $x$ plus grand que cette quantité.

Afin de simplifier le terme d'erreur, il est possible de l'augmenter pour obtenir

$$\psi(f,x) = rx - \frac{x^{\beta_f}}{\beta_f} + \mathcal{O}\bigg(\sqrt{\mathfrak{q}(f)}\,x\exp\bigg(-\frac{c}{2d^4}\sqrt{\log x}\bigg)\bigg).$$

Il est cru que tous les zéros, y compris ceux triviaux, à l'intérieur de la bande critique $0 < \mathrm{Re}(s) < 1$ sont situés sur la ligne $\mathrm{Re}(s) = 1/2$ pour toutes les fonctions $L$. Cette conjecture est appelée la Grande Hypothèse de Riemann. Sous cette hypothèse, le théorème des nombres premiers prend la forme

$$\psi(f,x) = rx + \mathcal{O}(x^{\frac{1}{2}+\epsilon})$$

pour tout $\epsilon > 0$ où la constante implicite dépend de $f$ et $\epsilon$.

# 5 Théorie des matrices aléatoires

## 5.1 Théorèmes de convergence et d'universalité

La matière de ce chapitre diffèrera substantiellement des chapitres précédents. Pourtant, il semblerait que la distribution des zéros des fonctions $L$ suit les mêmes lois que la distribution des valeurs propres de matrices aléatoires.

Un groupe topologique est un groupe muni d'une topologie telle que la fonction envoyant un élément à son inverse et l'opération du groupe sont continues par rapport à la topologie du groupe et à la topologie produit dans le cas de l'opération. Débutons avec la définition du groupe compact $U(N)$ des matrices unitaires. Les éléments de ce groupe sont les matrices $N \times N$ telles que $AA^* = A^*A = I_N$ avec coefficients en $\mathbb{C}$. L'opération est la multiplication de matrices. La topologie de ce groupe est induite par la topologie de l'espace de toutes les matrices $N \times N$ avec coefficients en $\mathbb{C}$ vu comme étant l'espace $\mathbb{C}^{N^2}$ avec sa topologie usuelle. Le groupe $U(N)$ est compact puisqu'il est fermé et borné dans cet espace. On définit similairement quatre autres sous-groupes compacts de $U(N)$ où leur topologie sera la topologie induite de $U(N)$ :

- $SU(N)$, le groupe des matrices $N \times N$ unitaires de déterminant 1
- $O(N)$, le groupe des matrices $N \times N$ orthogonales unitaires, c'est-à-dire, $A^TA = AA^T = I_N$
- $SO(N)$, le groupe des matrices $N \times N$ orthogonales unitaires de déterminant 1
- $USp(N)$, le groupe des matrices $N \times N$ symplectiques unitaires, où $N$ est pair. Une matrice symplectique est une matrice respectant $A^T\Omega A = \Omega$ où

$$\Omega = \begin{bmatrix} 0 & I_N \\ -I_N & 0 \end{bmatrix}.$$

On travaille avec la représentation unitaire standard de ces groupes puisqu'on effectuera des calculs sur des valeurs propres. On souhaite donc parler d'une matrice lorsqu'on sélectionne un élément d'un de ces groupes. Toutes ces matrices auront leurs valeurs propres sur le cercle unité complexe, ce qui nous permettra d'étudier leur distribution autour du cercle.

Si $\lambda$ est une valeur propre d'une matrice orthogonale ou symplectique, alors $\overline{\lambda}$ l'est également. Aussi, si $A$ est une matrice orthogonale de dimension impaire, 1 ou -1 sera toujours une valeur propre. On distinguera donc les matrices orthogonales de dimensions paires et impaires et on écrira $USp(2N)$, $O(2N)$, $SO(2N)$, $O(2N+1)$ et $SO(2N+1)$ afin de mettre l'emphase sur le fait que $N$ valeurs propres seront sous observation pour une matrice d'un de ces groupes.

Montrons maintenant comment bâtir une mesure sur $\mathbb{R}$ à partir d'une matrice unitaire. Il en sera de même pour les matrices des autres groupes à l'exception que les angles des valeurs propres se situeront dans l'intervalle $[0, \pi]$ pour les groupes ayant une symétrie. Une matrice unitaire $A$ de dimension $N$ est associée à une séquence de $N$ points dans l'intervalle $[0, 2\pi)$ par les angles de ses valeurs propres

$$0 \leq \varphi(1) \leq \varphi(2) \cdots \leq \varphi(N) < 2\pi$$

tels que

$$\det(T \cdot I - A) = \prod_j (T - \exp(i\varphi(j))).$$

On souhaite étudier la distance entre deux valeurs, on définit donc les quantités

$$s_j := (N/2\pi)(\varphi(j+1) - \varphi(j)), \qquad \text{pour } 1 \leq j \leq N-1$$

et

$$s_N := (N/2\pi)(2\pi + \varphi(1) - \varphi(N)).$$

Le facteur $(N/2\pi)$ ne sert qu'à normaliser les distances puisque plus $N$ augmente, plus la distance sera réduite entre deux angles. On garde donc une moyenne de 1 pour ces distances en normalisant ainsi.

On rassemble tous ces $s_j$ sous un seul objet, une mesure $\mu(A, U(N))$ sur $\mathbb{R}$ est établie en attribuant un poids de $1/N$ à chaque $s_j$, c'est-à-dire,

$$\int_{\mathbb{R}} f \, d\mu(A, U(N)) := (1/N) \sum_j f(s_j)$$

pour une fonction $f : \mathbb{R} \to \mathbb{R}$ Borel mesurable.

L'intégration de Weyl par rapport à la mesure de Haar qui sera définie dans la prochaine section nous permet d'obtenir une mesure qui ne dépendra plus d'une matrice en particulier. Cette mesure $\mu(U(N))$ sera en quelque sorte une moyenne des mesures $\mu(A, U(N))$ sur toutes les matrices du groupe. Elle est définie ainsi

$$\int_{\mathbb{R}} f \, d\mu(U(N)) := \int_{U(N)} \Big( \int_{\mathbb{R}} f \, d\mu(A, U(N)) \Big) d \operatorname{Haar}_{U(N)}.$$

Le premier théorème énonce que lorsque $N \to \infty$, la mesure $\mu(U(N))$ converge vers une mesure $\mu(\text{univ})$ dans le sens que pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ Borel mesurable à support compact

$$\int_{\mathbb{R}} f \, d\mu(\text{univ}) = \lim_{N \to \infty} \int_{\mathbb{R}} f \, d\mu(U(N)).$$

De plus, cette mesure a une fonction de densité continue.

Le deuxième théorème justifie la qualification d'universelle de cette mesure. Il démontre que pour n'importe quels groupes compacts énoncés ci-haut, la mesure limite sera toujours la même.

## 5.2 Mesure de Haar

Cette section rendra explicite l'intégration de Weyl sur un groupe compact par rapport à la mesure de Haar. Chaque matrice $A$ du groupe $U(N)$ est associée à un vecteur $X(A)$ dans $\mathbb{R}^N$ donné par les angles des valeurs propres ordonnées

$$X(A) = (\varphi(1), \varphi(2), \ldots, \varphi(N)).$$

Par cette association, les fonctions bornées $g : U(N) \to \mathbb{R}$ Borel mesurables invariantes sous conjugaison sont en bijection avec les fonctions bornées $\tilde{g} : [0, 2\pi)^N \to \mathbb{R}$ Borel mesurables

invariantes sous la permutation des coordonnées étant donné qu'un vecteur $X(A)$ détermine une unique classe de conjugaison de matrices à permutation près de ses coordonnées. Notons que les valeurs propres sont distinctes avec probabilité 1, elles déterminent donc la classe de conjugaison. On définit alors l'intégration de $g$ sur $U(N)$ comme étant l'intégrale de $\tilde{g}$ sur $[0, 2\pi)^N$, mais par rapport à une certaine mesure spéciale dépendante du groupe. Cette mesure sur $[0, 2\pi)^N$ est donnée ainsi pour $U(N)$

$$d\mu_{U(N)} := (1/N!) \left( \prod_{j<k} |e^{ix_j} - e^{ix_k}|^2 \right) \prod_i \frac{dx_i}{2\pi},$$

où les $x_i$ sont les $N$ coordonnées de l'espace $[0, 2\pi)^N$. Alors l'intégration de Weyl de la fonction $g$ sur $U(N)$ est donnée par

$$\int_{U(N)} g \, d\,\text{Haar} := \int_{[0,2\pi)^N} \tilde{g} \, d\mu_{U(N)}.$$

Les autres groupes auront leurs vecteurs $X(A) \in [0, \pi]^N$ puisque l'on considère seulement les valeurs propres dans cet intervalle. Leur mesure sur cet espace est définie ainsi :

- Pour le groupe $USp(2N)$ une fonction bornée $g : USp(2N) \to \mathbb{R}$ Borel mesurable invariante sous la conjugaison d'éléments de $USp(2N)$ est en correspondance avec une fonction bornée $\tilde{g} : [0, \pi]^N \to \mathbb{R}$ Borel mesurable invariante sous la permutation des coordonnées. La mesure est définie comme suit

  $$d\mu_{USp(2N)} := (1/N!) \left( \prod_{i<j} (2\cos(x_i) - 2\cos(x_j)) \right)^2 \prod_i ((2/\pi)\sin^2(x_i) dx_i)$$

  et donc

  $$\int_{USp(2N)} g \, d\,\text{Haar} := \int_{[0,\pi]^N} \tilde{g} \, d\mu_{USp(2N)}.$$

- Pour le groupe $SO(2N + 1)$

  $$d\mu_{SO(2N+1)} := (1/N!) \left( \prod_{i<j} ((1 + 2\cos(x_i)) - (1 + 2\cos(x_j))) \right)^2 \prod_i ((2/\pi)\sin^2(x_i/2) dx_i).$$

- Pour le groupe $SO(2N)$, la fonction $g$ que l'on souhaite intégrer doit être invariante sous la conjugaison des éléments de $O(2N) \supset SO(2N)$. On a

  $$d\mu_{SO(2N)} := (2/N!) \left( \prod_{i<j} (2\cos(x_i) - 2\cos(x_j)) \right)^2 \prod_i (dx_i/2\pi).$$

- On définit l'ensemble $O_-(2N+2)$ comme étant les éléments de $O(2N+2)$ avec déterminant $-1$. On note que cet ensemble n'est pas un groupe. On a la réunion disjointe suivante

  $$O(2N + 2) = SO(2N + 2) \cup O_-(2N + 2).$$

  L'intégration sur cet ensemble se fait par la même mesure que pour le cas $USp(2N)$

  $$d\mu_{O_-(2N+2)} := d\mu_{USp(2N))}$$

  et la fonction $g$ doit être invariante par rapport à la conjugaison par $O(2N + 2)$.

## 5.3 Continuité des valeurs propres

Il est nécessaire de montrer la continuité des valeurs propres afin que les fonctions utilisées dans les sections précédentes soient bien Borel mesurables et donc intégrables. On débutera par définir une topologie sur les classes de conjugaison de $U(N)$. Cette définition sera équivalente à la topologie quotient obtenue par la topologie de $U(N)$. Il sera ensuite nécessaire de montrer la Borel mesurabilité de $\varphi(n)$ pour $1 \leq n \leq N$ afin que la fonction $\tilde{g}$ définie par $\tilde{g}(X(A)) = g(A)$ soit Borel mesurable (avec la topologie usuelle de $[0, 2\pi)^N$) pour une fonction $g$ Borel mesurable. Finalement, la fonction $g$ utilisée par le théorème de convergence était

$$g(A) := \int_{\mathbb{R}} f d\mu(A, U(N))$$

pour une fonction $f : \mathbb{R} \to \mathbb{R}$ Borel mesurable. On devra montrer que $g$ est aussi Borel mesurable.

Une matrice $A \in U(N)$ est associée à un polynôme unitaire de degré $N$ ayant ses racines sur le cercle unité complexe par son polynôme caractéristique. Nous définirons donc une topologie sur l'espace des polynômes unitaire qui ont leurs racines sur le cercle unité complexe.

Plus généralement, soit $\mathcal{P}_{N,Z}$ l'ensemble des polynômes unitaires de degré $N$ ayant toutes leurs racines à l'intérieur de $Z \subset \mathbb{C}$. On définit une topologie sur $\mathcal{P}_{N,\mathbb{C}}$ en voyant cet espace comme étant l'espace $\mathbb{C}^N$ des coefficients des polynômes et en utilisant la topologie produit de la topologie usuelle de $\mathbb{C}$. Alors la topologie sur $\mathcal{P}_{N,Z}$ est induite de celle de $\mathcal{P}_{N,\mathbb{C}}$. On peut donc voir l'ensemble des classes de conjugaison de $U(N)$ comme étant l'espace $\mathcal{P}_{N,S^1}$. Le but sera de montrer qu'une petite modification des coefficients entraîne une petite modification des zéros.

La preuve utilisera le critère de la limite afin de montrer la continuité. Soit $N$ fixé et $(P_n)_{n=1}^{\infty}$ une suite dans $\mathcal{P}_{N,S^1}$ qui converge vers $P \in \mathcal{P}_{N,S^1}$. Par la topologie, la convergence signifie que pour tout $\epsilon > 0$, il existe un $N$ tel que pour tout $n \geq N$, $P_n$ est à distance $< \epsilon$ de $P$ coefficient par coefficient. Nous voulons montrer que les racines de $P_n$ convergent vers les racines de $P$.

D'abord, il est possible de s'assurer que toutes les racines de $P_n$ soient distinctes. En effet, pour chaque $n$ on peut déplacer les racines de $< 1/n$ donnant une nouvelle suite $Q_n$. On aura $Q_n - P_n \to 0$ puisque leurs coefficients sont des polynômes en les racines et donc continus en celles-ci.

Soit $d$ la plus petite distance entre deux racines de $P$ distinctes. Notons que l'on parle ici de distance sur le plan complexe et non de la distance circulaire. Dénotons par $\rho_{n,j}$ les $N$ racines de $P_n$ et $\rho_j$ les $N$ racines de $P$ avec multiplicité. Nous avons sur $S^1$, $\lim_{n\to\infty} \|P_n - P\|_{\infty,S^1} = 0$. Pour tout $0 < \epsilon < d$, il existe $N_\epsilon$ tel que pour tout $n \geq N_\epsilon$, $\|P_n - P\|_{\infty,S^1} < \epsilon^N$. En particulier, pour tout $j$, $|P(\rho_{n,j})| < \epsilon^N$. Puisque $P(x) = (x - \rho_1) \dots (x - \rho_N)$, il existe un $j'$ tel que $|\rho_{n,j} - \rho_{j'}| < \epsilon$. Alors chaque racine de $P_n$ converge vers une racine de $P$. Elles doivent correspondre une à une car sinon il est immédiat que $P_n$ ne converge pas vers $P$ coefficient par coefficient. Alors $\rho_{n,j}$ converge bien vers $\rho_j$ après une possible permutation de la numérotation.

Nous venons donc de prouver que les angles $\varphi(n)$ des valeurs propres sont des fonctions Borel mesurables de $U(N)$. Il ne reste qu'un détail à prouver, on doit montrer que la fonction utilisée pour l'intégration de Weyl lors du théorème de convergence est Borel mesurable et bornée.

Cette fonction était

$$A \mapsto \int_{\mathbb{R}} f \, d\mu(A, U(N))$$

pour une fonction $f : \mathbb{R} \to \mathbb{R}$ Borel mesurable bornée quelconque. On revient à la définition de l'intégration par rapport à la mesure $\mu(A, U(N))$

$$\int_{\mathbb{R}} f \, d\mu(A, U(N)) := (1/N) \sum_j f((N/2\pi)(\varphi(j+1) - \varphi(j))).$$

Il est immédiat que l'intégrale sera une fonction Borel mesurable de $U(N)$ par la Borel mesurabilité de $\varphi(j)$ et de $f$. Elle sera bornée puisque $f$ est bornée. L'intégration de Weyl est donc bien définie pour notre théorème de convergence.

## 5.4 La distribution des zéros des fonctions $L$

À partir de l'image des zéros non triviaux d'une fonction $L$, une mesure sur $\mathbb{R}$ peut également être définie. Cette section construira cette mesure dans le cas de $\zeta$ afin de simplifier les choses, car seule la normalisation des zéros sera différente. Par la Section 4.3, les zéros non triviaux $\rho_j = \beta_j + i\gamma_j$ de $\zeta$ respectent

$$\#\{j : 0 \leq \gamma_j \leq T\} \sim \frac{T \log T}{2\pi}, \qquad \text{lorsque } T \to \infty$$

puisque le conducteur de $\zeta$ et son degré valent 1. La normalisation est donc

$$\hat{\gamma}_j = \frac{\gamma_j \log \gamma_j}{2\pi}.$$

La mesure associée à ces zéros est définie par la distribution de l'espacement entre deux zéros

$$s_j = \hat{\gamma}_{j+1} - \hat{\gamma}_j.$$

On construit alors une mesure pour $L(f, s)$ une fonction $L$ ainsi

$$\mu_{f,N}([a, b]) = \frac{\#\{1 \leq j \leq N \mid s_j \in [a, b]\}}{N}.$$

L'idée sera de faire tendre $N$ à l'infini. Afin de comparer cette mesure avec la mesure universelle, on utilise la distance suivante introduite par Kolmogorov et Smirnov calculant le degré de similarité entre deux mesures. Cette fonction prend évidemment ses valeurs entre 0 et 1.

$$D(\nu_1, \nu_2) = \sup\{|\nu_1(I) - \nu_2(I)| : I \subset \mathbb{R} \text{ un intervalle}\}.$$

La conjecture suggérée par plusieurs expériences numériques est

$$\lim_{N \to \infty} D(\mu_{f,N}, \mu(\text{univ})) = 0$$

pour toutes les fonctions $L$! D'autres manières d'étudier la distribution des zéros sont de considérer les distances $\hat{\gamma}_{j+k} - \hat{\gamma}_j$ pour $k \geq 1$. Plus généralement, on peut définir par le même principe des mesures sur $\mathbb{R}^n$ en considérant les vecteurs

$$(\hat{\gamma}_{j+c(1)} - \hat{\gamma}_j, \hat{\gamma}_{j+c(2)} - \hat{\gamma}_{j+c(1)}, \ldots, \hat{\gamma}_{j+c(n)} - \hat{\gamma}_{j+c(n-1)})$$

pour un certain vecteur d'incrémentation $c = (c(1), \ldots, c(n))$ tel que

$$0 < c(1) < c(2) < \cdots < c(n).$$

Encore une fois, pour un $c$ fixé, les mesures sur les groupes compacts construites à partir de cette incrémentation convergent toutes vers une mesure universelle $\mu_c(\text{univ})$ de $\mathbb{R}^n$ où $n$ est la dimension de $c$. Pour ce même $c$, on construit par le même principe $\mu_{f,c,N}$ en utilisant la distribution des zéros non triviaux de $L(f,s)$. La conjecture plus générale est

$$\lim_{N\to\infty} D(\mu_{f,c,N}, \mu_c(\text{univ})) = 0$$

pour tout $c$ et toute fonction $L$! Comme la distribution de points sur une ligne est totalement déterminée par ces données, la distribution des zéros des fonctions $L$ respecte les mêmes lois que celle des valeurs propres de matrices aléatoires. Ce phénomène est appelé la *loi de Montgomery-Odlyzko*. C'est en 1974 que Montgomery a prouvé une version restrictive de ce phénomène sur $\zeta$. Odlyzko a par la suite effectué d'intensifs calculs numériques qui ont soutenu la véracité de cette loi. Cependant, la loi demeure toujours une conjecture.

Précisément, Montgomery a montré le résultat suivant. On débute par définir la corrélation des paires de valeurs propres normalisées de $A$ sous forme de mesure

$$R_2(A)[a,b] := \frac{\#\{j \neq k : \frac{N}{2\pi}(\varphi_j - \varphi_k) \in [a,b]\}}{N}.$$

Par l'intégration de Weyl et en faisant tendre $N \to \infty$, on obtient une mesure $R_2(\text{univ})$ encore une fois universelle, c'est-à-dire, ne dépendant pas du groupe compact choisi. On appelle $r_2(\text{univ})$ sa fonction de densité, c'est-à-dire

$$R_2(\text{univ})[a,b] = \int_a^b r_2(\text{univ})(x)dx.$$

Une formule simple existe pour cette fonction:

$$r_2(\text{univ})(x) = 1 - \left(\frac{\sin \pi x}{\pi x}\right)^2.$$

Le théorème de Montgomery est le suivant. Soit $\phi \in \mathcal{S}(\mathbb{R})$ telle que $\text{supp}(\hat{\phi}) \subset (-1,1)$, alors

$$\lim_{N\to\infty} \frac{1}{N} \sum_{1\leq j\neq k\leq N} \phi(\hat{\gamma}_j - \hat{\gamma}_k) = \int_{\mathbb{R}} \phi(x) r_2(\text{univ})(x)dx$$

où les $\hat{\gamma}_i$ sont les zéros non triviaux de $\zeta$ normalisés. La preuve est basée sur une version modifiée de la formule explicite donnée à la Section 4.4 tout comme les calculs de Young que nous allons voir en détail au Chapitre 7. Rudnick et Sarnak ont généralisé cette preuve à toutes les fonctions $L$ et à de plus hautes corrélations, toujours sous la restriction du support de la transformée de Fourier de $\phi$. La conjecture est que cette égalité est vraie sous aucune restriction du support. Les détails de ces preuves n'illuminent malheureusement guère l'explication de ce phénomène.

Malgré le fait que les lois asymptotiques d'espacements entre les valeurs propres ne dépendent pas du groupe compact, la densité asymptotique de la position de la $j^{\text{ième}}$ valeur propre diffère selon le type de groupe. Plus précisément, on définit pour une matrice $A$

$$\Delta(A)[a,b] := \#\{\varphi : e^{i\varphi} \text{ est une valeur propre de } A \text{ et } (N/2\pi)\varphi \in [a,b]\}$$

une mesure comptant le nombre de valeurs propres de $A$ normalisées dans l'intervalle $[a, b]$. On applique l'intégration de Weyl pour obtenir une mesure moyenne de toutes les mesures $\Delta(A)$ de $G(N)$. Par abus de notation,

$$W(G(N)) := \int_{G(N)} \Delta(A)\, d\text{Haar}_{G(N)}$$

où $G(N)$ désigne un groupe compact. Voir Section 5.1 pour le sens précis d'une moyenne de mesures par l'intégration de Weyl. Lorsque $N \to \infty$, cette mesure converge vers une mesure $W(G)$. Les fonctions de densité de cette mesure selon le groupe $G$ sont données par

$$w(G)(x) = \begin{cases} 1 & \text{si } G = U \text{ ou } SU \\ 1 + \frac{1}{2}\delta_0(x) & \text{si } G = O \\ 1 + \frac{\sin 2\pi x}{2\pi x} & \text{si } G = SO\,(\text{pair}) \\ 1 + \delta_0(x) - \frac{\sin 2\pi x}{2\pi x} & \text{si } G = SO\,(\text{impair}) \\ 1 - \frac{\sin 2\pi x}{2\pi x} & \text{si } G = USp \end{cases}$$

où $\delta_0$ est la distribution de Dirac et

$$W(G)[a, b] = \int_a^b w(G)(x)dx.$$

Par nos conjectures, on s'attendrait à observer le même phénomène pour les fonctions $L$. Pour $L(f, s)$ une fonction $L$, on analyse ses zéros non triviaux proches du point central $s = 1/2$ par une fonction test $\phi \in \mathcal{S}(\mathbb{R})$ ainsi

$$\Delta(f, \phi) := \sum_{\gamma_j} \phi\left(\frac{\gamma_j \log q_f}{2\pi}\right)$$

où la somme est sur les zéros non triviaux de $L(f, s)$ et $q_f$ est le conducteur de $L(f, s)$. La fonction test nous permet de sélectionner les zéros en distance $\mathcal{O}(1/\log t)$ de $1/2$. L'intervalle des zéros considérés est trop court et la position de ces zéros est trop irrégulière pour que $\Delta(f, \phi)$ ait une quelconque signification. En d'autres mots, il y a un manque d'information pour que le groupe dictant la distribution soit révélé. C'est la raison pour laquelle on considère un ensemble infini de fonctions $L$, autrement appelé *famille de fonctions* $L$, et que l'on calcule $\Delta(f, \phi)$ en moyenne sur celle-ci. L'idée de famille $\mathcal{F}$ de fonctions $L$ est également naturelle en soi, puisque des fonctions $L$ peuvent être intimement liées entre elles. On définit donc

$$W(X, \mathcal{F}, \phi) = \frac{1}{\#\mathcal{F}_X} \sum_{f \in \mathcal{F}_X} \Delta(f, \phi)$$

où $\mathcal{F}_X = \{f \in \mathcal{F} \,|\, q_f \leq X\}$ toutes les fonctions $L$ de $\mathcal{F}$ ayant conducteur $\leq X$. La philosophie de Katz et Sarnak prédit qu'asymptotiquement

$$\lim_{X \to \infty} W(X, \mathcal{F}, \phi) = \int_{\mathbb{R}} \phi(x) w(G)(x) dx$$

pour un certain groupe compact $G$. On appelle ce groupe le *type de symétrie de la famille* $\mathcal{F}$. Il est en général difficile de prouver ce résultat pour des fonctions tests sans restriction sur le support de leur transformée de Fourier. Les preuves découlent de la formule explicite. Évidemment, une conjecture est que l'équation est valide sans aucune restriction sur le support.

Il est d'intérêt de prouver ce résultat avec un support permis plus large que $[-1,1]$ puisque la transformée de Fourier de $w(G)$ est donnée par

$$\hat{w}(G)(x) = \begin{cases} \delta_0(x) & \text{si } G = U \text{ ou } SU \\ \frac{1}{2} + \delta_0(x) & \text{si } G = O \\ \delta_0(x) + \frac{1}{2}\eta(x) & \text{si } G = SO \text{ (pair)} \\ 1 + \delta_0(x) - \frac{1}{2}\eta(x) & \text{si } G = SO \text{ (impair)} \\ \delta_0(x) - \frac{1}{2}\eta(x) & \text{si } G = USp \end{cases}$$

où

$$\eta(x) := \begin{cases} 1 & \text{si } |x| < 1 \\ \frac{1}{2} & \text{si } |x| = 1 \\ 0 & \text{si } |x| > 1. \end{cases}$$

Nous voyons que $\hat{w}(O)(x)$, $\hat{w}(SO(\text{pair}))(x)$ et $\hat{w}(SO(\text{impair}))(x)$ sont tous égaux pour $|x| < 1$ et diffèrent pour $|x| > 1$. Le théorème de Plancherel nous donne

$$\int_{\mathbb{R}} \phi(x) w(G)(x) dx = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{\phi}(x) \hat{w}(G)(x) dx$$

d'où notre désir de pouvoir utiliser des $\hat{\phi}$ avec support dépassant $[-1,1]$ afin de différencier ces symétries si nécessaire.

Remarquons que la normalisation $\log q_f / 2\pi$ utilisée pour les zéros près du point central est due à la formule pour $N(T,f)$. Il n'est pas nécessaire de multiplier par $\log \gamma_j$ puisqu'ici on cherche à normaliser la densité des zéros dans notre intervalle et non à normaliser les distances entre deux zéros. Un autre problème technique est que $\gamma_j \sim 0$ et donc $\log \gamma_j \sim -\infty$. Seul le conducteur qui varie à travers notre famille aura un impact sur cette densité, d'où la normalisation qui neutralise cet effet.

Les évidences étant déjà abondantes à soutenir la philosophie de Katz et Sarnak, il existe un type de fonctions analogues aux fonctions $L$ tel que la source de la symétrie de leurs zéros est connue. Ces fonctions sont définies à partir d'une extension finie $k \supset \mathbb{F}_q(t)$, c'est-à-dire, une extension finie du corps des fonctions rationnelles en $t$ avec coefficients en $\mathbb{F}_q$, le corps fini avec $q$ éléments. La définition est

$$\zeta(T,k) := \prod_v (1 - T^{\deg(v)})^{-1}$$

où le produit est sur tous les éléments premiers $v$ de $k$. La lettre $\zeta$ est choisie afin de souligner la similarité avec la fonction zêta de Riemann. Une version de l'hypothèse de Riemann existe pour ces fonctions: tous les zéros sont situés sur le cercle $|T| = 1/\sqrt{q}$. Contrairement aux fonctions $L$, cette hypothèse est prouvée, la première démonstration ayant été donnée par Weil. Également, un prolongement analytique sur tout le plan complexe et une équation fonctionnelle existent pour ces fonctions.

Un point de vue géométrique existe pour ces fonctions. Soit $C$ une courbe projective non singulière donnée par l'équation

$$C : f(X_1, X_2, X_3) = 0$$

où $f$ est un polynôme homogène avec coefficients en $\mathbb{Z}$. Alors la fonction $\zeta(T,k)$ définie avec $k$ le corps des fonctions de la courbe réduite en $\mathbb{F}_q$ est égale à

$$\zeta(T, C/\mathbb{F}_q) := \exp\left(\sum_{n\geq 1} \frac{N_n T^n}{n}\right)$$

où $N_n$ est le nombre de solutions projectives du polynôme dans $\mathbb{F}_{q^n}$, le corps fini avec $q^n$ éléments. L'interprétation spectrale des zéros provient de ce point de vue. Les quantités $N_n$ correspondent au nombre de points fixés par la $n^{\text{ième}}$ itération de l'endomorphisme de Frobenius. Cet endomorphisme met à la puissance $q$ les coordonnées de la courbe sur $\bar{\mathbb{F}}_q$. Le corps $\mathbb{F}_q \subset \bar{\mathbb{F}}_q$ est donc fixé par cet endomorphisme. La $n^{\text{ième}}$ itération fixe le corps $\mathbb{F}_{q^n}$ d'où le lien avec $N_n$. Ceci permet d'interpréter les zéros de $\zeta(T, V/\mathbb{F}_q)$ comme étant les réciproques des valeurs propres de l'endomorphisme de Frobenius qui agit comme opérateur sur le premier groupe de cohomologie de la variété.

Les zéros de ces fonctions étant distribués autour d'un cercle s'analysent de la même manière que pour les valeurs propres de matrices aléatoires. Pour une courbe donnée, $\zeta(T, V/\mathbb{F}_q)$ aura $2g$ zéros où $g$ est le genre de la courbe. En considérant une famille de courbes assez large, ces zéros respectent en moyenne la loi de Montgomery-Odlyzko. Les zéros près du point central respectent encore un certain type de symétrie. La source de ce comportement est bien connue et provient de la représentation monodromique $\pi_1$ qui associe une courbe de la famille à une matrice d'un certain groupe compact. Les zéros de la fonction zêta de cette courbe sont directement associés aux valeurs propres de cette matrice. Ainsi, il est possible de révéler le type de symétrie de la famille par cette représentation, car en variant les courbes à travers la famille, les matrices associées deviennent uniformément réparties à l'intérieur du groupe compact.

Donnons un exemple afin de mieux comprendre la relation entre ces fonctions et les fonctions $L$. Soit $H_n(\mathbb{F}_q)$ l'ensemble de tous les polynômes unitaires sans carré de degré $n$ avec coefficients en $\mathbb{F}_q$. Considérons les extensions finies $k_\Delta = \mathbb{F}_q(t)(\sqrt{\Delta})$ où $\Delta \in H_n(\mathbb{F}_q)$. La famille des fonctions $\zeta(T, k_\Delta)$ a une symétrie symplectique. L'analogue de cette famille est la famille de fonctions $L(s,\chi)$ où $\chi$ est un caractère primitif réel, car un caractère réel primitif est associé à une extension $\mathbb{Q}(\sqrt{m})$ où $m$ est sans carré. La symétrie de cette famille est également symplectique. Par contre, on ne sait pas encore si ce partage de symétrie est généralisé pour toute famille ayant une famille analogue.

En conclusion, une explication possible pour le lien entre les fonctions $L$ et les matrices aléatoires serait une interprétation spectrale des zéros des fonctions $L$. Une fonction $L$ serait associée à un opérateur d'un espace quelconque et ses zéros seraient liés aux valeurs propres de cet opérateur. Lorsque la fonction $L$ varie dans sa famille, les opérateurs associés seraient uniformément distribués dans leur ensemble d'une symétrie donnée. Cette hypothèse expliquerait le phénomène observé et les fonctions zêta d'une extension finie de $\mathbb{F}_q(t)$ soutiennent ce point de vue par leur grande similarité.

# 6 Courbes elliptiques

## 6.1 Équation de Weierstrass

Pour le reste du chapitre, fixons $k = \mathbb{Q}$. Les espaces affines et projectifs seront sur $\mathbb{C} \supset \bar{\mathbb{Q}}$ au lieu de $\bar{\mathbb{Q}}$. Une variété de dimension 1 est appelée une courbe, nous utiliserons donc les espaces $\mathbb{A}^2(\mathbb{C})$ et $\mathbb{P}^2(\mathbb{C})$. Une courbe elliptique est une courbe régulière de genre 1 avec un point de base spécifié. Le genre est intuitivement une mesure de complexité d'une courbe ou variété. Un fait important est que pour toute courbe elliptique, il existe une courbe isomorphe donnée par les zéros du polynôme homogène

$$f(X, Y, Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3$$

où $a_1, \ldots, a_6 \in \mathbb{C}$ et $\Delta \neq 0$. La quantité $\Delta$ est appelée le *discriminant* de la courbe et sera calculée ci-après. La courbe est singulière si et seulement si $\Delta = 0$. On appelle équation de Weierstrass l'équation suivante respectée par les zéros de $f$

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3.$$

Le point $O = [0, 1, 0]$ est un zéro de cette courbe et correspond au point de base de la courbe. Il est plus simple d'étudier une courbe elliptique sur l'espace affine en gardant toujours en tête le point à l'infini $O$ ainsi retiré. L'équation de Weierstrass devient donc

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

Si $a_1, \ldots, a_6 \in \mathbb{Q}$, on dit que la courbe elliptique est définie sur $\mathbb{Q}$ et on écrit $E/\mathbb{Q}$ pour désigner la courbe. Des équations de Weierstrass différentes peuvent correspondre à des courbes elliptiques isomorphes. Seules les substitutions

$$(x, y) \mapsto (u^2x + r, u^3y + u^2sx + t)$$

pour $u \neq 0, r, s, t \in \mathbb{C}$ envoient une équation de Weierstrass sur une autre tout en préservant le point $O$ où les courbes correspondantes sont isomorphes.

Il est possible de simplifier une équation de Weierstrass par ces substitutions. Pour une équation de Weierstrass donnée, on effectue la substitution

$$y \mapsto \frac{1}{2}(y - a_1x - a_3)$$

pour arriver à l'équation

$$y^2 = 4x^3 + b_2x^2 + 2b_4x + b_6$$

où $b_2 = a_1^2 + 4a_4$, $b_4 = 2a_4 + a_1a_3$ et $b_6 = a_3^2 + 4a_6$. On remarque que la substitution ne respecte pas la forme ci-haute, la raison étant qu'une autre substitution sera effectuée et que la composition des deux respectera la forme. Définissons les quantités

$$\begin{aligned}
b_8 &= a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2, \\
c_4 &= b_2^2 - 24b_4, \\
c_6 &= -b_2^3 + 36b_2b_4 - 216b_6, \\
\Delta &= -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6, \\
j &= c_4^3/\Delta.
\end{aligned}$$

La quantité $j$ est appelée le *$j$-invariant* de la courbe elliptique, car elle est invariante par rapport aux substitutions ci-hautes. En fait, deux courbes elliptiques ont le même $j$-invariant si et seulement si elles sont isomorphes. Si $\Delta = 0$, la courbe est singulière et n'est pas une courbe elliptique. Deux types de singularité sont alors possibles. Un noeud est un point où la courbe se croise en elle-même, il y aura donc deux droites tangentes possibles. Le deuxième type n'admet qu'une seule droite tangente et est appelé point de rebroussement. Un critère pour différencier les deux est donné par la valeur de $c_4$. Si $c_4 = 0$, la singularité est un point de rebroussement, sinon, elle est un noeud.

On applique maintenant la substitution

$$(x, y) \mapsto \left(\frac{x - 3b_2}{36}, \frac{y}{108}\right)$$

pour obtenir l'équation de Weierstrass

$$y^2 = x^3 - 27c_4x - 54c_6.$$

Alors chaque courbe elliptique est isomorphe à une courbe elliptique donnée par

$$y^2 = x^3 + ax + b.$$

Le discriminant de cette équation est $\Delta = -16(4a^3+27b^2)$ et le $j$-invariant vaut $j = -1728(4a)^3/\Delta$. Les substitutions qui garderont cette forme sont pour $u \in \mathbb{C}^*$

$$(x, y) \mapsto (u^2x, u^3y)$$

qui donneront l'équation $y^2 = x^3+a'x+b'$ où $a' = a/u^4$, $b' = b/u^6$ et $\Delta' = \Delta/u^{12}$. On peut donc s'assurer que $a, b \in \mathbb{Z}$ pour les courbes elliptiques définies sur $\mathbb{Q}$. Une équation de Weierstrass sous cette forme avec $a, b \in \mathbb{Z}$ est dite globalement minimale si les nombres premiers divisant $\Delta$ ne peuvent pas être réduits en puissance par de telles substitutions. La courbe est donc globalement minimale si et seulement s'il n'existe pas de $p > 3$ tel que $p^4|a$ et $p^6|b$. Cette forme détermine de manière unique une classe d'isomorphisme de courbes elliptiques.

## 6.2 Le groupe des points rationnels

Une structure de groupe existe sur les points d'une courbe elliptique non singulière où l'identité est le point à l'infini $O = [0, 1, 0]$. L'opération est définie en étudiant les droites traversant les points de la courbe. Les calculs seront effectués sur le plan affine où l'équation de la courbe sera $y^2 = x^3 + ax + b$ par souci de simplicité, car nous travaillerons implicitement dans le monde projectif où le point à l'infini existe. On homogénéise par $x = X/Z$ et $y = Y/Z$ afin d'étudier concrètement ce point à l'infini. Du point de vue projectif, on appelle la ligne à l'infini les points avec $Z = 0$ (car les coordonnées affines diviseront par zéro) engendrant pour la courbe elliptique l'équation $X^3 = 0$. Le point à l'infini est donc un zéro triple de la courbe elliptique situé sur la ligne à l'infini.

Une ligne $x = c$ sur $\mathbb{A}^2(\mathbb{C})$ où $c$ est une constante s'homogénéise à $X = cZ$. Sur $\mathbb{P}^2$, le point $[0, 1, 0]$ est solution de cette équation homogène. Alors, encore une fois de manière implicite, les lignes verticales $x = c$ du plan affine traverse le point à l'infini $O$.

Montrons maintenant comment calculer l'opération de groupe $\oplus$ entre deux points $P$ et $Q$ d'une courbe elliptique $E : y^2 = x^3 + ax + b$. Il s'agit de relier par une droite les points $P$ et $Q$ et de trouver le troisième point d'intersection $P * Q$ entre la droite et $E$. On relie ensuite par une droite le point $P * Q$ et $O$ et le troisième point d'intersection sera $P \oplus Q$, le résultat de l'opération de groupe. La droite reliant $P * Q$ et $O$ sera une droite verticale par le paragraphe précédent, et par la symétrie d'une courbe elliptique par rapport à l'axe des $x$, si $P * Q = (x, y)$, alors $P \oplus Q = (x, -y)$. Il est évident que $\oplus$ est commutatif. Le groupe est donc abélien.

Débutons par l'axiome d'existence d'un inverse. Montrons que le point $P$ a comme inverse le point $-P$ défini comme étant la réflexion de $P$ par rapport à l'axe des $x$. Le troisième point d'intersection de la droite reliant $P$ et $-P$ est $O$. Le troisième point d'intersection de la droite entre $O$ et $O$ est $O$, car la droite reliant $O$ et $O$ est la ligne à l'infini et $O$ est un zéro triple de la courbe. Pour le cas $P = O$, il est immédiat que $O \oplus O = O$ et donc que $-O = O$. Par le même principe, il est évident que $O$ respecte les axiomes de l'identité.

Effectuons maintenant les calculs afin de montrer l'existence d'un troisième point d'intersection et de nous donner un moyen d'obtenir explicitement $P_1 \oplus P_2$. Soit $P_1 = (x_1, y_1)$ et $P_2 = (x_2, y_2)$ deux points de la courbe $E$. La droite $y = \lambda x + \nu$ reliant $P_1$ et $P_2$ aura

$$\lambda = \frac{y_2 - y_1}{x_2 - x_1} \qquad \text{et} \qquad \nu = y_1 - \lambda x_1 = y_2 - \lambda x_2.$$

En substituant $y$ dans l'équation de $E$ pour cette droite, on obtient

$$x^3 + (-\lambda^2)x^2 + (a - 2\lambda\nu)x + (b - \nu^2) = 0.$$

Ce polynôme de degré 3 se factorise ainsi $(x-x_1)(x-x_2)(x-x_3)$, d'où l'existence d'un troisième point d'intersection. Les racines $x_1$ et $x_2$ sont déjà connues, il est donc possible d'isoler $x_3$ afin d'obtenir

$$P_1 \oplus P_2 = (x_3, -y_3) = (\lambda^2 - x_1 - x_2, -\lambda x_3 - \nu)$$

où la quantité $y_3$ est donnée par la droite et est multipliée par $-1$ par réflexivité.

Il reste un important détail à considérer. Cette construction ne fonctionne pas si $P_1 = P_2$ car il ne sera pas possible de calculer la pente $\lambda$ par $\frac{y_2-y_1}{x_2-x_1}$. Intuitivement, lorsque $P_2$ approche $P_1$, la droite reliant les deux points approche la tangente au point $P_1$. Alors $\lambda$ sera donnée par la pente de la tangente au point $P_1$. On calcule cette pente par différentiation implicite

$$\lambda = \left.\frac{dy}{dx}\right|_{P_1} = \frac{f'(x_1)}{2y_1}$$

où l'équation de $E$ est sous la forme $y^2 = f(x)$.

L'associativité de $\oplus$ se vérifie par de fastidieux calculs à partir de ces formules et plusieurs cas doivent être considérés.

L'ensemble $E(\mathbb{Q})$ des points rationnels $P = (x, y)$ tels que $x, y \in \mathbb{Q}$ d'une courbe elliptique $E/\mathbb{Q}$ avec $O$ forment un sous-groupe. L'important théorème de Mordell-Weil affirme que ce sous-groupe est de type fini avec rang $< \infty$. Le tout se généralise pour une courbe elliptique $E/k$ où $k$ est une extension des rationnels. Les points rationnels $E(k)$ de la courbe sont les $P = (x, y)$ tels que $x, y \in k$ et forment un sous-groupe de type fini également.

## 6.3 La fonction $L$ de Hasse-Weil

Cette section définira la fonction $L(s, E)$ de Hasse-Weil construite à partir d'une courbe elliptique $E/\mathbb{Q}$. Les informations concernant le nombre de solutions de la courbe elliptique réduite modulo une puissance d'un nombre premier seront contenues par $L(s, E)$. Également, la fameuse conjecture de Birch et Swinnerton-Dyer tisse un lien entre le rang de $E(\mathbb{Q})$ et l'ordre du zéro de $L(s, E)$ à $s = 1$ en affirmant que ces deux quantités sont égales. La fonction $L(s, E)$ respectera une équation fonctionnelle où la ligne de symétrie sera $\sigma = 1$. Il est donc nécessaire de renormaliser à $L(s + 1/2, E)$ si nous souhaitons respecter la définition donnée au Chapitre 4 d'une fonction $L$.

Détaillons d'abord la réduction d'une courbe elliptique $E/\mathbb{Q}$ vers $\mathbb{F}_p$. Soit $y^2 = x^3 + ax + b$ l'équation reliée à la courbe elliptique où $a, b \in \mathbb{Z}$. On obtient une courbe $E/\mathbb{F}_p$ par l'équation $y^2 = x^3 + \tilde{a}x + \tilde{b}$ où $\tilde{\cdot}$ dénote la réduction modulo $p$. Pour $r > 0$, on définit $\#E(\mathbb{F}_{p^r})$ comme étant le nombre de zéros du polynôme $f(x, y) = y^2 - x^3 - \tilde{a}x - \tilde{b}$ dans $\mathbb{A}^2(\mathbb{F}_{p^r})$ additionné du point à l'infini. Le point à l'infini est compté puisque l'on travaille implicitement sur $\mathbb{P}^2(\mathbb{F}_{p^r})$. Il est trivial que $\#E(\mathbb{F}_{p^r}) < \infty$, le plan $\mathbb{A}^2(\mathbb{F}_{p^r})$ ne comptant que $p^{2r}$ points.

La courbe elliptique peut devenir singulière une fois réduite. On dit alors que la réduction est mauvaise, sinon elle est dite bonne. Si la singularité est un noeud, la réduction à $p$ est dite multiplicative. De plus, une réduction multiplicative est soit séparée ou non séparée selon si les pentes des deux droites tangentes sont éléments de $\mathbb{F}_p$ ou non respectivement. Le cas où la singularité est un point de rebroussement, la réduction est dite additive. Il est possible de montrer que $\#E(\mathbb{F}_p) = p - 1$, $p + 1$, ou $p$, si $E$ a une réduction multiplicative séparée, multiplicative non séparée, ou additive, respectivement.

Une notion de conducteur existe pour une courbe elliptique. Les nombres premiers divisant le conducteur correspondent exactement aux nombres premiers où la réduction est mauvaise. Par le critère $\Delta = 0$ si et seulement si la courbe est singulière, la courbe réduite sera singulière si et seulement si $p|\Delta$ pour le $\Delta$ globalement minimal de la courbe. Le conducteur sera alors composé des mêmes facteurs premiers. On le définit ainsi pour le $\Delta$ globalement minimal

$$N = \prod_{p|\Delta} p^{f_p}$$

où

$$f_p = \begin{cases} 1 & \text{si } p \nmid c_4 \text{ ($E$ est réduit multiplicativement à $p$)}, \\ 2 & \text{si } p \mid c_4 \text{ ($E$ est réduit additivement à $p$)} \end{cases}$$

pour $p > 3$. Lorsque $p = 2$ ou 3, la définition de $f_p$ est plus compliquée. On a toujours que $f_p \leq 8$ et ce fait sera suffisant pour les calculs du prochain chapitre. Si nécessaire, le conducteur peut être calculé par la formule d'Ogg et l'algorithme de Tate.

La construction de la fonction $L$ liée à la courbe elliptique se servira de la fonction zêta d'une courbe définie au Chapitre 5 pour $p$ fixé. La définition est, pour $E/\mathbb{F}_p$ la réduction des coefficients de $E$ modulo $p$

$$\zeta(T, E/\mathbb{F}_p) := \exp\left(\sum_{n\geq 1} \frac{N_n T^n}{n}\right).$$

Les coefficients $N_n$ seront alors égaux à $\#E(\mathbb{F}_{p^n})$ par leur définition. Il est connu que ces fonctions se représentent en une forme rationnelle pour un $p$ de bonne réduction

$$\zeta(T, E/\mathbb{F}_p) = \frac{1 - a_p T + pT^2}{(1-T)(1-pT)}$$

où $a_p$ est un entier dépendant de $E$ et $p$. Ce fait est généralisé pour les fonctions zêtas d'une variété $V$ non singulière où le numérateur sera un polynôme $P(T)$ de degré $2g$ avec coefficients entiers où $g$ est le genre de la variété. Les zéros de $\zeta(T, V/\mathbb{F}_p)$ seront alors les zéros de $P(T)$ provenant de l'endomorphisme de Frobenius tel que discuté au Chapitre 5.

Notons qu'avec la factorisation $1 - a_p T + pT^2 = (1 - \alpha T)(1 - (p/\alpha)T)$, la comparaison des deux formes de $\zeta(T, E/\mathbb{F}_p)$ en appliquant la dérivée logarithmique nous donne

$$N_n = p^n + 1 - \alpha^n - (p/\alpha)^n.$$

En particulier, $N_1 = p + 1 - a_p$. Alors la connaissance de $\#E(\mathbb{F}_p)$ permet de calculer $a_p$, $\alpha$ et ainsi toutes les valeurs $N_n$.

Pour les $p$ de mauvaise réduction, seul le numérateur de $\zeta(T, E/\mathbb{F}_p)$ sous la forme rationnelle diffèrera et sera donné par $1 - a_p T$ où

$$a_p = \begin{cases} 1 & \text{si } E \text{ a une réduction multiplicative séparée à } p \\ -1 & \text{si } E \text{ a une réduction multiplicative non séparée à } p \\ 0 & \text{si } E \text{ a une réduction additive à } p. \end{cases}$$

Nous avons la relation

$$a_p = p + 1 - \#E(\mathbb{F}_p)$$

pour tous autres $p$.

La définition de $L(s, E)$ est donnée par le produit des $\zeta(T, E/\mathbb{F}_p)$ pour tous les $p$ en posant $T = p^{-s}$. Précisément,

$$L(s, E) := \frac{\zeta(s)\zeta(s-1)}{\prod_p \zeta(p^{-s}, E/\mathbb{F}_p)}$$

où $\zeta(s)$ est la fonction zêta de Riemann. On remarque que le produit des dénominateurs des fonctions $\zeta(p^{-s}, E/\mathbb{F}_p)$ équivaut au produit d'Euler de $\zeta(s)\zeta(s-1)$, on élimine donc ces dénominateurs. Notre connaissance des numérateurs nous permet d'affirmer

$$L(s, E) = \prod_{p \nmid N} (1 - a_p p^{-s} + p^{1-2s})^{-1} \prod_{p|N} (1 - a_p p^{-s})^{-1}.$$

Afin de respecter la forme donnée dans la définition d'une fonction $L$, on factorise

$$(1 - a_p p^{-s} + p^{1-2s})^{-1} = (1 - \alpha(p)p^{-s})^{-1}(1 - (p/\alpha(p))p^{-s})^{-1}$$

où $\alpha(p)$ est solution de $\alpha(p)^2 - a_p \alpha(p) + p = 0$. Le choix d'une des deux solutions ne fait que permuter $\alpha(p)$ et $p/\alpha(p)$. La série de Dirichlet est alors donnée par le développement du produit d'Euler. Remarquons que les coefficients à $p$ de la série de Dirichlet vaudront $a_p$ pour tout $p$. Le conducteur de $L(s, E)$ correspond à $N$ et son degré est 2.

L'hypothèse de Riemann pour les fonctions zêta de variétés sur les corps finis montrée par Hasse, Weil et Deligne implique $|\alpha(p)| = \sqrt{p}$. On a alors $a_p < 2\sqrt{p}$ appelée *borne de Hasse*. Le produit converge alors uniformément et absolument pour $\mathrm{Re}(s) > 3/2$. Par les travaux de Wiles menant à la preuve du dernier théorème de Fermat et les travaux subséquents de Breuil, Conrad, Diamond et Taylor, nous savons qu'il existe une forme modulaire $f_E$ telle que $L(s + 1/2, E) = L(f_E, s)$ pour toutes courbes elliptiques $E/\mathbb{Q}$. Par conséquent, il existe un prolongement analytique pour $L(s, E)$ et l'équation fonctionnelle suivante est respectée.

$$\Lambda(s, E) := \left(\frac{\sqrt{N}}{2\pi}\right)^s \Gamma(s)L(s, E) = \epsilon\Lambda(2 - s, E)$$

où $\epsilon = \pm 1$ est le signe de l'équation fonctionnelle. Le changement de variable $s \to s + 1/2$ déplace la ligne de symétrie de $s = 1$ à $s = 1/2$ afin de respecter le cadre défini pour les fonctions $L$.

# 7 La symétrie de la famille des courbes elliptiques sur $\mathbb{Q}$

## 7.1 Introduction

Ce chapitre donnera les détails du calcul du type de symétrie de la famille de toutes les courbes elliptiques sur $\mathbb{Q}$. La grande hypothèse de Riemann sera assumée pour les résultats de ce chapitre. Les calculs présentés seront ceux effectués par Young en 2004 dans sa thèse de doctorat. Le type de symétrie ayant déjà été calculé par Brumer en 1992, l'avantage des travaux de Young réside dans l'augmentation du support permis pour $\hat{\phi}$, la transformée de Fourier d'une fonction test. La dernière section de ce chapitre motivera la recherche d'un meilleur support. Plus le support permis est large, mieux est notre borne supérieure sur la moyenne du rang analytique des fonctions $L$ de notre famille, le rang analytique étant l'ordre du zéro à $s = 1$. Aucune restriction sur le support implique une moyenne de 1/2.

La famille considérée est donc $\mathcal{F} = \{E_{a,b}\}$ où $E_{a,b} : y^2 = x^3 + ax + b$ pour $a, b \in \mathbb{N}$. On considèrera alors la famille de fonctions $L(s, E_{a,b})$ associée à $\mathcal{F}$ afin de calculer la densité moyenne des zéros près du point central. Bien sûr, il y aura plusieurs courbes isomorphes entre elles dans cette famille. La Section 7.3 sur la restriction de minimalité montrera qu'aucun impact significatif n'intervient en considérant tous les $a, b \in \mathbb{N}$. Les calculs sont ainsi beaucoup plus plaisants.

Montrons d'abord une formule pour les coefficients $a_p$ de $L(s, E_{a,b})$ qui sera utile pour effectuer des calculs. La formule est donnée par une somme de symboles de Legendre

$$a_p = - \sum_{x \,(\mathrm{mod}\ p)} \left( \frac{x^3 + ax + b}{p} \right).$$

Si pour un $x \in \mathbb{Z}/p\mathbb{Z}$, $x^3 + ax + b$ est un carré modulo $p$, alors il existe $y$ tel que $y^2 = x^3 + ax + b \pmod p$ et ainsi deux solutions $(x, y)$ et $(x, -y)$ sont comptées par $\#E(\mathbb{F}_p)$ si $p \nmid x^3 + ax + b$ et une seule $(x, 0)$ si $p \mid x^3 + ax + b$. Dénotons par $n_1$ le nombre de $x \in \mathbb{Z}/p\mathbb{Z}$ tels que $x^3 + ax + b$ est un carré mod $p$ et $p \nmid x^3 + ax + b$, $n_2$ le nombre de $x$ tels que $x^3 + ax + b$ est un carré mod $p$ et $p \mid x^3 + ax + b$ et $n_3$ le nombre de $x$ tels que $x^3 + ax + b$ n'est pas un carré mod $p$. Alors $\#E(\mathbb{F}_p) = 2n_1 + n_2 + 1$ en additionnant le point à l'infini. Aussi, la somme sur $x \pmod p$ vaut $n_3 - n_1$. Par $n_1 + n_2 + n_3 = p$ et $a_p = p + 1 - \#E(\mathbb{F}_p)$, la formule est vérifiée.

Notre but sera d'appliquer les idées de la Section 5.4, c'est-à-dire, de calculer

$$\lim_{X \to \infty} W(X, \mathcal{F}, \phi).$$

Le problème dans notre cas est qu'il est très difficile de sélectionner exactement les courbes de notre famille ayant un conducteur $\leq X$. Pour remédier à cette situation, nous nous servirons du discriminant. En fait, il est suffisant que les courbes sélectionnées par rapport à $X$ aient un conducteur $N$ tels que $\log N$ soit asymptotiquement $\log X$ en moyenne. La restriction de minimalité assure que $a < 12$ pour $p^a \| \Delta$ et donc que cette condition est respectée. Nous allons donc prendre des courbes ayant majoritairement $|\Delta| \asymp X$ lorsque $X$ augmente. Rappelons que les facteurs premiers de $N$ sont ceux du discriminant minimal $\Delta$ et que $N | \Delta$. Afin de sélectionner de telles courbes, on introduit une fonction lisse $\omega(x, y)$ à support compact sur $\mathbb{R}^2$. On pose ensuite

$$\omega_X(E_{a,b}) := \omega(a/X^{1/3}, b/X^{1/2}).$$

Alors les courbes $E \in \mathcal{F}$ telles que $\omega_X(E) \neq 0$ auront $|\Delta| \ll X$ et seront les courbes considérées par rapport à $X$. Les courbes non minimales seront également sélectionnées par simplicité, mais la section sur la minimalité montrera qu'elles ont peu d'influence.

Définissons maintenant

$$D(E, \phi) := \sum_{\gamma_E} \phi\left(\frac{\gamma_E}{2\pi} \log X\right)$$

étant l'équivalent de $\Delta(f, \phi)$ défini à la Section 5.4 où les zéros sont comptés avec multiplicité. Rappelons que cette somme représente la densité des zéros de $L(s, E)$ près du point central. La lettre $D$ est choisie ici puisque le symbole $\Delta$ est déjà utilisé pour le discriminant. Nous avons remplacé $\log N_E$ par $\log X$ dans la définition puisqu'il est difficile de travailler directement avec le conducteur, c'est pourquoi nous voulions que $\log N_E$ soit en moyenne asymptotiquement $\log X$ sur les courbes considérées. Pour reprendre le principe de $W(X, \mathcal{F}, \phi)$, on définit

$$\mathcal{D}(\omega_X, \mathcal{F}, \phi) := \sum_{E \in \mathcal{F}} D(E, \phi)\omega_X(E).$$

C'est ici que $\omega_X$ est utilisé afin de sélectionner les courbes telles que $|\Delta| \asymp X$. Au lieu de calculer la moyenne en divisant par le nombre de courbes considérées, on divise par la somme totale des poids

$$W_X(\mathcal{F}) = \sum_{E \in \mathcal{F}} \omega_X(E).$$

La philosophie de Katz et Sarnak prédit donc que

$$\lim_{X \to \infty} \frac{\mathcal{D}(\omega_X, \mathcal{F}, \phi)}{W_X(\mathcal{F})} = \int_{\mathbb{R}} \phi(t) w(G)(t) dt$$

où les $w(G)$ sont données à la Section 5.4 pour les différents types de symétrie. Les calculs de la prochaine section montreront que

$$\frac{\mathcal{D}(\omega_X, \mathcal{F}, \phi)}{W_X(\mathcal{F})} \sim \hat{\phi}(0) + \frac{1}{2}\phi(0) \qquad \text{lorsque } X \to \infty$$

pour $\phi \in \mathcal{S}(\mathbb{R})$ tel que $\text{supp}(\hat{\phi}) \subset (-7/9, 7/9)$. La symétrie de notre famille est donc orthogonale dans la limite de ce que nous pouvons savoir sans avoir un support plus large que $(-1, 1)$.

## 7.2 Calcul du type de symétrie

Le point de départ nous permettant d'effectuer des calculs sur $D(E, \phi)$ est une forme particulière de la formule explicite pour une fonction $L$ d'une forme modulaire (voir (4.25) de [11]). La modularité des courbes elliptiques sur $\mathbb{Q}$ nous permet donc d'utiliser cette formule dans notre cas. La formule est

$$D(E, \phi) = \hat{\phi}(0)\frac{\log N_E}{\log X} + \frac{1}{2}\phi(0) - P(E, \phi) + \mathcal{O}\left(\frac{\log\log|\Delta|}{\log X}\right)$$

où

$$P(E, \phi) = \sum_{p > 3} \lambda_E(p)\hat{\phi}\left(\frac{\log p}{\log X}\right)\frac{2\log p}{p \log X}.$$

Les $\lambda_E(p)$ sont les coefficients de la série de Dirichlet pour $L(s,E)$ à $p$, alors

$$\lambda_E(p) = a_p = -\sum_{x \pmod p} \left(\frac{x^3+ax+b}{p}\right).$$

L'erreur de la formule explicite provient des termes de la somme $P(E,\phi)$ sur les puissances de nombres premiers $p^k$, $k>1$. Les termes où $k \geq 3$ sont directement bornés. Les termes avec $k=2$ sont bornés en utilisant la formule explicite pour $L(s,\mathrm{Sym}^2 E) := L(s, E\otimes E)$ en assumant l'hypothèse de Riemann. Son prolongement analytique et l'équation fonctionnelle furent prouvés par Shimura. Brièvement, en exprimant

$$L(s+\frac{1}{2},E) = \prod_p (1-\alpha(p)p^{-s})^{-1}(1-\beta(p)p^{-s})^{-1},$$

nous avons que $\lambda_E(p^2) = \alpha^2(p)+\alpha(p)\beta(p)+\beta^2(p)$. D'autre part, après la renormalisation appropriée,

$$-\frac{L'}{L}(s+1,\mathrm{Sym}^2 E) = \sum_{k\geq 1}\sum_p \frac{(\alpha(p)^{2k}+\alpha(p)^k\beta(p)^k+\beta(p)^{2k})\log p}{p^{ks}}$$

par définition. Les termes en $\lambda_E(p^2)$ que nous souhaitons borner seront donc présents dans la formule explicite de $L(s+1,\mathrm{Sym}^2 E)$. La borne suit en estimant chaque autre terme.

La quantité qui nous intéresse est $\mathcal{D}(\omega_X,\mathcal{F},\phi)$, on somme alors sur toutes les courbes de notre famille pour obtenir

$$\mathcal{D}(\omega_X,\mathcal{F},\phi) = \hat{\phi}(0)\sum_{E\in\mathcal{F}}\frac{\log N_E}{\log X}\omega_X(E) + \frac{1}{2}\phi(0)W_X(\mathcal{F}) - \mathcal{P}(\omega_X,\mathcal{F},\phi) + \mathcal{O}\left(\frac{W_X(\mathcal{F},\Delta)}{\log X}\right),$$

où

$$\mathcal{P}(\omega_X,\mathcal{F},\phi) := \sum_{E\in\mathcal{F}} P(E,\phi)\omega_X(E)$$

et

$$W_X(\mathcal{F},\Delta) := \sum_{E\in\mathcal{F}} |\omega_X(E)|\log\log|\Delta_E|.$$

On retire la dépendance aux discriminants par l'estimation $\log\log\Delta \ll \log\log X$. Le terme d'erreur ne dépend que de $X$ maintenant et vaut

$$\mathcal{O}\left(\frac{\log\log X}{\log X}W_X(\mathcal{F})\right).$$

Remarquons qu'il tend vers 0 lorsque $X\to\infty$ après avoir divisé par $W_X(\mathcal{F})$. Prouver notre théorème revient alors à montrer que

$$\sum_{E\in\mathcal{F}}\frac{\log N_E}{\log X}\omega_X(E) \sim W_X(\mathcal{F}) \qquad \text{lorsque } X\to\infty$$

appelée *la condition du conducteur* et à montrer que

$$\mathcal{P}(\omega_X,\mathcal{F},\phi) = \mathcal{O}\left(\frac{\log\log X}{\log X}W_X(\mathcal{F})\right)$$

appelée *l'estimation centrale*. Le support de $\hat{\phi}$ contrôle les nombres premiers sommés. Nous ne sommes toujours pas en mesure de montrer cette borne sans restreindre l'étendue des nombres premiers sommés. La restriction du support provient donc de l'estimation centrale.

*Preuve de la condition du conducteur.* Le but sera d'utiliser le discriminant afin d'étudier le comportement du conducteur à travers notre famille. On contrôlera ensuite les nombres premiers composant les discriminants. On élimine d'abord la présence du conducteur en introduisant $\Delta$ trivialement

$$W_X(\mathcal{F}) - \sum_{E\in\mathcal{F}} \frac{\log N_E}{\log X}\omega_X(E) = \sum_{E\in\mathcal{F}} \frac{\log(X/|\Delta_E|)}{\log X}\omega_X(E) + \sum_{E\in\mathcal{F}} \frac{\log(|\Delta_E|/N_E)}{\log X}\omega_X(E).$$

La première somme est $\ll |W_X(\mathcal{F})|/\log X$ car les courbes de $\mathcal{F}$ sélectionnées par $\omega_X$ ont en moyenne $|\Delta| \asymp X$. De manière plus formelle, on sépare d'abord les courbes sélectionnées en deux groupes. On considère les courbes ayant $X^{1/3}/\log X \ll a \ll X^{1/3}$ et $X^{1/2}/\log X \ll b \ll X^{1/2}$. Les autres courbes formeront donc un petit sous-ensemble de proportion $1/\log^2 X$. Ces courbes sont celles ayant un discriminant ne pouvant être mieux borné inférieurement que par $\gg 1$. Leur petite quantité compense alors et la somme sur celles-ci est $\ll |W_X(\mathcal{F})|(\log X)^{-1}$. Les courbes de notre autre sous-ensemble auront $\Delta \gg X/\log^2 X$. La somme sur celles-ci est alors $\ll |W_X(\mathcal{F})|(\log\log X)/\log X$. Cette borne n'est pas tout à fait celle désirée par la présence du facteur $\log\log X$, mais seule la dominance par rapport à $|W_X(\mathcal{F})|$ est nécessaire.

La deuxième somme vaut ensuite

$$\frac{1}{\log X}\sum_{E\in\mathcal{F}}\sum_{\substack{p^\alpha\|\Delta_E\\ p^\beta\|N_E}} \omega_X(E)\log p^{\alpha-\beta}$$

par additivité de logarithme. Pour les termes ayant $\beta > 0$, on sépare ceux avec $\alpha > \beta$ de ceux avec $\alpha < \beta$.

$$\ll \frac{1}{\log X}\sum_{E\in\mathcal{F}}\sum_{\substack{p^\alpha\|\Delta_E\\ a>0}} |\omega_X(E)|\log p^{\alpha-1} + \frac{1}{\log X}\sum_{E\in\mathcal{F}}\sum_{\substack{p\|\Delta_E\\ p^2|N_E\\ p>3}} |\omega_X(E)|\log p + \mathcal{O}\left(\frac{|W_X(\mathcal{F})|}{\log X}\right).$$

Rappelons que $\beta \le 8$ et $\beta \le 2$ pour $p > 3$. La deuxième double somme disparaît par le fait que $N_E|\Delta_E$. L'erreur provient des termes où $p = 2$ ou 3 avec $\alpha < \beta$. Maintenant, les termes avec $\beta = 0$ sont les $p$ tels que $p \,|\, N_E$ et $p \nmid \Delta_E$. Donc $\Delta_E$ n'est pas minimal et $\alpha \ge 12$. Ces termes sont alors absorbés par la première somme, car ils y sont présents puisque $\alpha - 1 > 0$.

Alors, notre but est maintenant de montrer

$$\sum_{E\in\mathcal{F}}\sum_{\substack{p^\alpha\|\Delta_E\\ \alpha>0}} |\omega_X(E)|\log p^{\alpha-1} \ll |W_X(\mathcal{F})|$$

ce qui implique la condition du conducteur.

Par définition, cette somme vaut

$$\sum_a\sum_b \sum_{\substack{p^\alpha\|16(4a^3+27b^2)\\ \alpha>0}} \omega\left(\frac{a}{A},\frac{b}{B}\right)\log p^{\alpha-1}$$

où $A = X^{1/3}$ et $B = X^{1/2}$ afin d'alléger les formules. Changeons l'ordre de sommation, on sommera alors sur les $p^\alpha \ll X$, $\alpha > 0$ où la constante implicite ne dépend que du support de $\omega$. Concentrons-nous sur les $p > 3$, les cas $p = 2$ et 3 sont traités essentiellement de la même manière.

On définit pour chaque $p$ la quantité $\gamma$ qui est telle que $p^\gamma \| a$. Il suffira de traiter trois cas. Considérons en premier lieu les termes où $\alpha < 3\gamma$. Alors $p^\alpha \| b^2$ et donc $\alpha$ est pair. Ces termes sont donc

$$\ll \sum_{\substack{p^\alpha \ll X \\ \alpha > 0}} \sum_{\alpha/3 < \gamma \ll \log A} \sideset{}{'}\sum_{a'} \sideset{}{'}\sum_{b'} \omega\left(\frac{p^\gamma a'}{A}, \frac{p^{\alpha/2} b'}{B}\right) \log p^{\alpha-1}$$

où l'apostrophe à droite des sommes signifie la restriction $(a', p) = (b', p) = 1$. On a ensuite

$$\sideset{}{'}\sum_{a'} \sideset{}{'}\sum_{b'} \omega\left(\frac{p^\gamma a'}{A}, \frac{p^{\alpha/2} b'}{B}\right) \ll \left(1 + \frac{A}{p^\gamma}\right)\left(1 + \frac{B}{p^{\alpha/2}}\right)$$

en comptant les points $(x, y)$ tels que $x \leq 1$ et $y \leq 1$ par le support compact de $\omega$. La constante implicite dépend donc de l'étendue de ce support et de $\|\omega\|_\infty$. Le 1 est ajouté par la borne $\lfloor x \rfloor = x + \mathcal{O}(1)$. En développant ce produit et évaluant la somme sur $\gamma$, on a

$$\ll \sum_{\substack{p^\alpha \ll X \\ \alpha > 0}} \left(\log A + \frac{A}{p^{\lfloor \alpha/3 \rfloor + 1}} + \frac{B \log A}{p^{\alpha/2}} + \frac{AB}{p^{\alpha/2 + \lfloor \alpha/3 \rfloor + 1}}\right) \log p^{\alpha-1}$$

$$\ll \sqrt{X} \log A + A \log X + B \log^2 X + AB \ll X^{5/6}$$

par

$$\sum_{\substack{p^\alpha \leq Z \\ \alpha > 0}} \log p^{\alpha-1} \ll \sqrt{Z} \quad \text{et} \quad \sum_{\substack{p^\alpha \\ \alpha > r}} \frac{\log p^{\alpha-1}}{p^{\alpha/r}} \ll 1.$$

Traitons maintenant les termes où $\alpha > 3\gamma$. On aura alors $p^{3\gamma} \| b^2$ et donc $\gamma$ est pair. La somme sur ces termes est alors

$$\ll \sum_{\substack{p^\alpha \ll X \\ \alpha > 0}} \sum_{0 \leq \gamma < \alpha/3} \sideset{}{'}\sum_{\substack{a' \\ p^{\alpha - 3\gamma} \| 4a'^3 + 27b'^2}} \sideset{}{'}\sum_{b'} \omega\left(\frac{p^\gamma a'}{A}, \frac{p^{3\gamma/2} b'}{B}\right) \log p^{\alpha-1}.$$

Puisque $(a', p) = (b', p) = 1$, pour $a'$ fixé il y a au plus deux solutions à la congruence $4a'^3 + 27b'^2 \equiv 0 \pmod{p^{\alpha-3\gamma}}$. On divise donc par $p^{\alpha-3\gamma}$ le nombre de $b'$ sommés afin de respecter cette congruence. Alors la borne est maintenant

$$\begin{aligned}
&\ll \sum_{\substack{p^\alpha \ll X \\ \alpha > 0}} \sum_{0 \leq \gamma < \alpha/3} \left(1 + \frac{A}{p^\gamma}\right)\left(1 + \frac{1}{p^{\alpha-3\gamma}} \cdot \frac{B}{p^{3\gamma/2}}\right) \log p^{\alpha-1} \\
&\ll \sum_{\substack{p^\alpha \ll X \\ \alpha > 0}} \left(\alpha + A + \frac{B}{p^{\alpha - \frac{3}{2}\lfloor \alpha/3 \rfloor}} + \frac{AB}{p^{\alpha - \frac{1}{2}\lfloor \alpha/3 \rfloor}}\right) \log p^{\alpha-1} \\
&\ll X^{5/6}
\end{aligned}$$

par le même principe que pour $\alpha < 3\gamma$.

Les termes $\alpha = 3\gamma$ sont similairement bornés par $\ll X^{5/6}$. Il est facile de voir que $|W_X(\mathcal{F})| \asymp X^{5/6}$, ce qui termine la preuve. □

Montrons d'abord une proposition et trois lemmes avant de se lancer dans les détails techniques en vue d'obtenir un meilleur support. On obtient assez aisément un support admis contenu à l'intérieur de $(-5/9, 5/9)$ par la proposition, le premier lemme et en bornant trivialement. Brumer a d'abord montré ce résultat. Par la suite, Young a amélioré ce support à $(-7/9, 7/9)$ en montrant qu'il y avait beaucoup d'annulations entre les termes de la somme. Les deux derniers lemmes seront utilisés dans cette poursuite d'un meilleur support afin d'alléger les étapes.

**Proposition 7.1.** *Soit $\omega \in \mathcal{S}(\mathbb{R})$ l'espace de Schwartz, $D \in \mathbb{R}$ positif et $a \in \mathbb{N}$. Alors*

$$\sum_{\substack{d\in\mathbb{Z} \\ d\equiv a \pmod{l}}} \omega\left(\frac{d}{D}\right) = \frac{D}{l}\sum_{h\in\mathbb{Z}} e\left(\frac{ha}{l}\right)\hat{\omega}\left(\frac{hD}{l}\right).$$

*Preuve.* Il s'agit de la sommation de Poisson pour les progressions arithmétiques. La preuve est similaire à celle donnée dans le prochain chapitre. □

**Lemme 7.2.** *Soit $p > 2$ un nombre premier et $\bar{k}$ l'inverse de $k$ modulo $p$ ($k\bar{k} \equiv 1 \pmod{p}$) si $(k, p) = 1$ et $\bar{0} = 0$. Alors pour*

$$T(h, k; p) := \sum_{\alpha \pmod{p}} \sum_{\beta \pmod{p}} \lambda_{\alpha,\beta}(p) e\left(\frac{\alpha h + \beta k}{p}\right)$$

*où*

$$\lambda_{\alpha,\beta}(p) := -\sum_{x \pmod{p}} \left(\frac{x^3 + ax + b}{p}\right),$$

*nous avons l'égalité suivante*

$$T(h, k; p) = -\epsilon_p p^{3/2} \left(\frac{k}{p}\right) e\left(\frac{-h^3\bar{k}^2}{p}\right)$$

*où $(\cdot|p)$ est le symbole de Legendre et $\epsilon_p$ est le signe de la somme de Gauss suivante*

$$\sum_{\beta \pmod{p}} \left(\frac{\beta}{p}\right) e\left(\frac{\beta k}{p}\right).$$

*Preuve.* Par définition,

$$T(h, k; p) = -\sum_{x \pmod{p}} \sum_{\alpha \pmod{p}} \sum_{\beta \pmod{p}} \left(\frac{x^3 + \alpha x + \beta}{p}\right) e\left(\frac{\alpha h + \beta k}{p}\right).$$

Appliquons le changement de variable $\beta \to \beta - x^3 - \alpha x$ qui ne fait que permuter les termes de la somme sur $\beta$. On obtient

$$T(h, k; p) = -\sum_{x \pmod{p}} e\left(\frac{-x^3 k}{p}\right) \sum_{\alpha \pmod{p}} e\left(\frac{\alpha(h - xk)}{p}\right) \sum_{\beta \pmod{p}} \left(\frac{\beta}{p}\right) e\left(\frac{\beta k}{p}\right).$$

La somme sur $\alpha$ vaut $p$ si $x = h\bar{k}$ et 0 sinon. La somme sur $\beta$ est une somme de Gauss valant $\epsilon_p p^{1/2}(k|p)$. Alors

$$T(h,k;p) = -\epsilon_p p^{3/2}\left(\frac{k}{p}\right) e\left(\frac{-h^3\bar{k}^2}{p}\right).$$

□

Seules des esquisses de preuves seront données pour les deux lemmes suivants, leur preuve étant très technique. Voir les annexes A et B de [9] pour plus de détails sur les deux lemmes suivants respectivement.

**Lemme 7.3.** *Soit $F(u,v)$ une fonction lisse respectant*

$$F^{(\alpha_1,\alpha_2)}(u,v)u^{\alpha_1}v^{\alpha_2} \leq C(\alpha_1,\alpha_2)\left(1+\frac{|u|}{U}\right)^{-2}\left(1+\frac{|v|}{V}\right)^{-2}$$

*pour tout $\alpha_1, \alpha_2 \geq 0$ où $F^{(\alpha_1,\alpha_2)}$ signifie la $\alpha_1^{ième}$ dérivée par rapport à $u$ et la $\alpha_2^{ième}$ dérivée par rapport à $v$ où $U, V \in \mathbb{R}$ et $C(\alpha_1,\alpha_2)$ des constantes. Alors, sous l'hypothèse de Riemann généralisée,*

$$\sum_u \sum_v \chi(u)\psi(v)\Lambda(v)e\left(\frac{u^c}{vq}\right)F(u,v) \ll_\epsilon U^{1/2}V^{1/2}\left(1+\frac{U^c}{V|q|}\right)^{1/2}(l_1 l_2 UV)^\epsilon$$

*où $\chi$ et $\psi$ sont des caractères de Dirichlet non principaux modulo $l_1$ et $l_2$ respectivement, $\Lambda$ est la fonction de von Mangoldt, $q$ est un nombre rationnel $\neq 0$, $c$ un entier positif et $\epsilon > 0$. La constante implicite dépend seulement de $c$, $\epsilon$ et $C(\alpha_1,\alpha_2)$. Si $\chi$ est principal, un facteur $U^{1/2}$ s'ajoute à la borne. Si $\psi$ est principal, un facteur $V^{1/2}$ s'ajoute.*

*Esquisse de la preuve.* Par inversion de Mellin,

$$\sum_u \sum_v \chi(u)\psi(v)\Lambda(v)e\left(\frac{u^c}{vq}\right)F(u,v)$$

$$= -\left(\frac{1}{2\pi i}\right)^2 \int_{(1/2+\epsilon)}\int_{(1/2+\epsilon)} L(s_1,\chi)\frac{L'}{L}(s_2,\psi)H(s_1,s_2)ds_1 ds_2$$

où

$$H(s_1,s_2) := \int_0^\infty \int_0^\infty e\left(\frac{u^c}{vq}\right)F(u,v)u^{s_1}v^{s_2}\frac{dudv}{uv}$$

et la notation $\int_{(c)}$ signifie une intégration de $c - i\infty$ à $c + i\infty$. L'intégration se fait sur la ligne Re $s_j = 1+\epsilon$ si le caractère est principal afin d'éviter le pôle à $s = 1$. Les facteurs ainsi ajoutés seront expliqués par ce changement et la borne que nous allons donner sur $H(s_1,s_2)$.

L'hypothèse de Riemann implique la borne sur notre ligne d'intégration

$$L(s_1,\chi)\frac{L'}{L}(s_2,\psi) \ll_\epsilon |s_1 s_2 l_1 l_2|^\epsilon$$

et pour $H(s_1,s_2)$ nous avons la borne

$$H(s_1,s_2) \ll \frac{U^{\sigma_1}V^{\sigma_2}}{|(s_1/c+s_2)(s_1/c+s/2+1)||s_2|^{1+\epsilon}}\left(1+\frac{U^c}{V|q|}\right)^{1/2}.$$

Ces deux bornes impliquent notre lemme.

Essentiellement, on obtient la borne pour $H(s_1, s_2)$ en éliminant d'abord la variable $u$ par le changement de variables $t = u^c/vq$. La preuve utilise fortement l'intégration par partie, car chaque application introduit un facteur de convergence et les dérivées de $F((tvq)^{1/c}, v)$ respectent les mêmes conditions que $F$.

On se concentre ensuite sur la l'intégration par rapport à la variable $t$, car l'annulation provient de l'exponentielle qui aura la forme $e^{iL(t)}$ pour une certaine fonction $L(t)$. La variation de $L(t)$ sera alors responsable de l'annulation. Il y aura un $t$ où $L'(t) = 0$. La contribution principale correspondra donc à l'intervalle proche de ce point. On sépare alors l'intégrale en trois parties et un analyse soigneuse de chaque région permet d'obtenir la borne voulue. □

Dans la preuve de l'estimation centrale, ce lemme introduira quatre cas à considérer. Nous aurons besoin du lemme suivant dans le quatrième cas.

**Lemme 7.4.** *Soient $c_n$ des nombres complexes tels que $|c_n| \leq 1$. Alors*

$$S := \sum_{P \leq p < 2P} \left| \sum_{N \leq n < 2N} e\left(\frac{n^3 d_0^3}{pk^2}\right) c_n \right| \ll N^{1/2}P + N^{1/4}P^{5/4}k^{1/2}d_0^{-3/4}.$$

*Esquisse de la preuve.* On étend d'abord la somme sur $p$ sur tous les entiers entre $P$ et $2P$, utilisons donc la lettre $m$ et posons $M = P$. Une fonction $F_\epsilon(x)$ lisse et positive prenant la valeur 1 entre $1 \leq x \leq 2$ est ensuite introduite telle que son support est inclus dans l'intervalle $(1 - \epsilon, 2 + \epsilon)$. Ceci nous permet de sommer sur tous les $m \in \mathbb{Z}$ en multipliant par $F_\epsilon(m/M)$ à l'intérieur de la somme. On applique l'inégalité de Cauchy pour obtenir

$$S^2 \ll M \sum_{N \leq n_1 < 2N} \sum_{N \leq n_2 < 2N} c_{n_1}\overline{c_{n_2}} \sum_{m \in \mathbb{Z}} F_\epsilon\left(\frac{m}{M}\right) e\left(\frac{(n_1^3 - n_2^3)d_0^3}{mk^2}\right).$$

Les termes diagonaux ($n_1 = n_2$) contribuent trivialement $\ll NM^2$. La sommation de Poisson implique

$$\sum_{m \in \mathbb{Z}} F_\epsilon\left(\frac{m}{M}\right) e\left(\frac{(n_1^3 - n_2^3)d_0^3}{mk^2}\right) = \sum_{r \in \mathbb{Z}} \int_{\mathbb{R}} F_\epsilon\left(\frac{u}{M}\right) e\left(\frac{(n_1^3 - n_2^3)d_0^3}{uk^2} - ru\right) du.$$

On borne trivialement l'intégrale après deux intégrations par partie, la somme sur presque tous les $r$ sera alors $\ll 1$. Les $r$ problématiques seront $|r| \asymp (|n_1^3 - n_2^3|d_0^3)/(M^2k^2)$. La méthode de la phase stationnaire est utilisée afin de borner ceux-ci. □

*Preuve de l'estimation centrale.* Nous voulons montrer que

$$\mathcal{P}(\omega_X, \mathcal{F}, \phi) \ll \frac{X^{5/6}}{\log X}$$

puisque $|W_X(\mathcal{F})| \asymp X^{5/6}$. Par définition

$$\begin{aligned}\mathcal{P}(\omega_X, \mathcal{F}, \phi) &= \sum_a \sum_b P(E, \phi)\omega\left(\frac{a}{A}, \frac{b}{B}\right) \\ &= \sum_{p>3} \frac{2\log p}{p \log X} \hat{\phi}\left(\frac{\log p}{\log X}\right) \sum_a \sum_b \lambda_{a,b}(p)\omega\left(\frac{a}{A}, \frac{b}{B}\right)\end{aligned}$$

où $A = X^{1/3}$ et $B = X^{1/2}$. On applique la Proposition 7.1 à la double somme intérieure en regroupant les $a$ et $b$ selon leur congruence modulo $p$. On obtient alors

$$\sum_a \sum_b \lambda_{a,b}(p)\omega\left(\frac{a}{A}, \frac{b}{B}\right) = \frac{AB}{p^2} \sum_h \sum_k \sum_{\substack{\alpha \pmod p \\ \beta \pmod p}} \sum \lambda_{\alpha,\beta} e\left(\frac{\alpha h + \beta k}{p}\right) \hat{\omega}\left(\frac{hA}{p}, \frac{kB}{p}\right)$$

en changeant l'ordre de sommation et par le fait que $\lambda_{a,b}$ ne dépend que de la classe d'équivalence de $a$ et $b$ modulo $p$ par sa définition. Le Lemme 7.2 nous donne alors

$$\mathcal{P}(\omega_X, \mathcal{F}, \phi) = -\frac{AB}{\log X} \sum_{p>3} \epsilon_p \frac{2\log p}{p^{3/2}} \hat{\phi}\left(\frac{\log p}{\log X}\right) \sum_h \sum_k \left(\frac{k}{p}\right) e\left(\frac{-h^3\bar{k}^2}{p}\right) \hat{w}\left(\frac{hA}{p}, \frac{kB}{p}\right). \quad (7.1)$$

En estimant trivialement la partie de cette somme avec $p \leq P$

$$\ll \frac{AB}{\log X} \sum_{p \leq P} \frac{\log p}{p^{3/2}} \left(1 + \frac{p}{A}\right)\left(1 + \frac{p}{B}\right) \ll \frac{AB + BP^{1/2} + P^{3/2}}{\log X} \ll \frac{X^{5/6}}{\log X}$$

si $P \leq X^{5/9}$. Le support de $\hat{\phi}$ nous assure que les $p$ sommés sont $\leq P$ par la présence du logarithme. Nous voyons maintenant d'où provient la restriction du support. Ce résultat fut obtenu par Brumer.

Montrons maintenant comment obtenir un support plus large par les travaux de Young. On élimine la variation de $\epsilon_p$ en séparant les termes où $p \equiv 1 \pmod 4$ et $p \equiv 3 \pmod 4$

$$\epsilon_p = \begin{cases} 1 & \text{si } p \equiv 1 \pmod 4 \\ i & \text{si } p \equiv 3 \pmod 4. \end{cases}$$

Il sera suffisant de remplacer $\epsilon_p$ par $\psi_4(p)$ un caractère de Dirichlet modulo 4 afin de faciliter les calculs. L'idée est maintenant de partitionner les termes de la somme (7.1) ainsi par une partition de l'unité

$$S(H, K, P) := \sum_{\substack{H \leq h < 2H \\ K \leq k < 2K \\ P \leq p < 2P}} \sum \sum \frac{\log p}{p^{3/2}} \psi_4(p) \left(\frac{k}{p}\right) e\left(\frac{-h^3\bar{k}^2}{p}\right) \hat{\phi}\left(\frac{\log p}{\log X}\right) \hat{w}\left(\frac{hA}{p}, \frac{kB}{p}\right) g(h, k, p)$$

où $g$ est une fonction lisse à support compact limitant les $h, k$ et $p$ à ceux prescrits par la somme provenant de la partition de l'unité. Nous assumons la redondance, car les restrictions sous la somme triple ne servent qu'à expliciter le support de $g$.

La somme (7.1) est composée de $\ll \log X$ telles partitions. Le support de $\hat{\phi}$ limite en premier lieu la valeur maximale de $P$. En pratique, nous obtiendrons $P \leq X^{7/9}$. Ensuite, la borne $\hat{\omega}(x, y) \ll_M (1+|x|)^{-M}(1+|y|)^{-M}$ nous permet d'assumer $H \ll_\epsilon (P/A)^{1+\epsilon}$ et $K \ll_\epsilon (P/B)^{1+\epsilon}$.

Le but est alors de montrer que toutes sommes $S(H, K, P) \ll X^{-\epsilon}$ où la constante implicite ne dépend que de $\omega$ et $g$ car ainsi, la somme (7.1) sans le facteur $-AB/\log X$ sera $\ll 1$ et la borne voulue sera immédiate. Le cas $k = 0$ n'est pas considéré grâce au symbole de Legendre et le cas $h = 0$ est trivialement insignifiant.

La première étape afin d'analyser ces sommes est de rendre $h$ et $k$ relativement premier. Pour y parvenir, définissons $d = (h^3, k^2)$ et $d_0$ comme étant le plus petit entier positif tel que $d|d_0^3$. Il est ainsi possible de factoriser $h = h_0 d_0$. La condition $(h^3, k^2) = d$ est équivalente à $(h_0, k^2/d) = 1$ et $(d_0^3/d, k^2/d) = 1$. En modifiant alors la somme sur $h$, l'exponentielle devient

$$\sum_h \sum_k e\left(\frac{-h^3\bar{k}^2}{p}\right) = \sum_k \sum_{\substack{d|k^2 \\ (h_0,k^2/d)=1 \\ (d_0^3/d,k^2/d)=1}} \sum_{h_0} e\left(\frac{-h_0^3(d_0^3/d)\overline{(k^2/d)}}{p}\right).$$

Le symbole de Legendre dans la somme (7.1) nous assure que $k$ est inversible modulo $p$ et donc que $k^2/d$ s'inverse aussi. L'avantage est que les sommes sur $h$ et $k$ sont courtes lorsque $d$ est grand, on souhaite donc sommer sur $d$ d'abord. Maintenant, cette modification nous permet de transformer l'exponentielle par la formule élémentaire suivante

$$-\frac{\bar{u}}{v} \equiv \frac{\bar{v}}{u} - \frac{1}{uv} \pmod 1$$

où $u, v, \bar{u}$ et $\bar{v}$ sont des entiers tels que $(u, v) = 1$, $u\bar{u} \equiv 1 \pmod v$ et $v\bar{v} \equiv 1 \pmod u$.

En posant $u = k^2/d$, $v = p$ et en multipliant la formule élémentaire par $h_0^3(d_0^3/d)$, la somme est maintenant

$$S(H,K,P) = \sum_{\substack{K\le k<2K \\ P\le p<2P}} \sum_{d|k^2} \sum_{\substack{H/d_0\le h_0<2H/d_0 \\ (h_0,k^2/d)=1 \\ (d_0^3/d,k^2/d)=1}} \left(\frac{k}{p}\right) e\left(\frac{h_0^3(d_0^3/d)\bar{p}}{k^2/d}\right) e\left(\frac{-h_0^3 d_0^3}{pk^2}\right) U(h_0, d_0, k, p)$$

où

$$U(h_0, d_0, k, p) := g(h_0 d_0, k, p)\hat{\omega}\left(\frac{h_0 d_0 A}{p}, \frac{kB}{p}\right) \frac{2\log p}{p^{3/2}\log X}\hat{\phi}\left(\frac{\log p}{\log X}\right).$$

En changeant ainsi les rôles de $p$ et $k$ à l'intérieur de l'exponentielle, nous pouvons maintenant utiliser la somme de Gauss

$$e\left(\frac{h_0^3(d_0^3/d)\bar{p}}{k^2/d}\right) = \frac{1}{\phi(k^2/d)} \sum_{\chi (\mathrm{mod}\ k^2/d)} \tau(\chi)\overline{\chi}(h_0^3(d_0^3/d)\bar{p}).$$

La somme est maintenant

$$S(H,K,P) = \sum_{K\le k<2K} \sum_{d|k^2} \frac{1}{\phi(k^2/d)} \sum_{\chi (\mathrm{mod}\ k^2/d)} \tau(\chi)\bar{\chi}(d_0^3/d)Q(d,k,\chi) \tag{7.2}$$

où

$$Q(d,k,\chi) := \sum_{P\le p<2P} \sum_{H/d_0\le h_0<2H/d_0} \psi_4(p)\chi(p)\left(\frac{k}{p}\right)\bar{\chi}^3(h_0) e\left(\frac{-h_0^3 d_0^3}{pk^2}\right) U(h_0, d_0, k, p).$$

On se concentre maintenant à borner $Q(d, k, \chi)$. La somme $S(H, K, P)$ sera ensuite trivialement bornée. On applique le Lemme 7.3 en définissant

$$F(u,v) = (v/P)^{-3/2} g(ud_0, k, v)\hat{\omega}\left(\frac{ud_0 A}{v}, \frac{kB}{v}\right)\hat{\phi}\left(\frac{\log v}{\log X}\right).$$

Des calculs directs montrent que les conditions du lemme sont respectées où $U = \min(H/d_0, P/d_0A)$ et $V = P$. Pour les autres paramètres du lemme, $\chi(u)$ sera dans notre cas $\bar{\chi}^3(u)$ et $\psi(v)$ sera $\chi(v)\psi_4(v)(k|v)$. Alors $l_1 = k^2/d$ et $l_2 = \text{ppcm}(k^2/d, 4, k^*)$ où $k^*$ est le conducteur de $(k|v)$. Si $k$ n'est pas un multiple du discriminant d'une extension quadratique des rationnels, il faut le multiplier par 4 afin que $(k|\cdot)$ soit un caractère. Nous avons aussi $q = -k^2/d_0^3$ et $c = 3$. La somme se fera sur $u = h_0$ et $v = p$. Par contre, il est nécessaire d'étendre la sommation sur les puissances de nombres premiers par la présence de $\Lambda(v)$ dans le Lemme 7.3. Remarquons que $P \leq p^n < 2P \Rightarrow P^{1/n} \leq p < (2P)^{1/n}$. La nouvelle somme avec ces termes ajoutés sera donc $\sim Q(d, k, \chi)$ lorsque $X \to \infty$. Alors

$$Q(d,k,\chi) \ll P^{-1}\left(\frac{H}{d_0}\right)^{1/2}\left(1+\frac{H^{3/2}}{P^{1/2}k}\right)X^{\epsilon}$$

si $\chi\psi_4(k|\cdot)$ et $\bar{\chi}^3$ sont non principaux. Si $\chi\psi_4(k|\cdot)$ est principal, un facteur $P^{1/2}$ s'ajoute et si $\bar{\chi}^3$ est principal, un facteur $(H/d_0)^{1/2}$ s'ajoute.

Il y aura donc quatre cas à considérer en utilisant cette borne dans $S(H, K, P)$. Séparons alors $S(H, K, P) = S_1 + S_2 + S_3 + S_4$. Montrons que chaque $S_i \ll X^{-\epsilon}$.

**Cas 1.** Les deux caractères sont non principaux. Alors, par (7.2)

$$S_1 \ll \sum_{K\leq k<2K}\sum_{d|k^2}\frac{1}{\varphi(k^2/d)}\sum_{\chi(\text{mod } k^2/d)}P^{-1}\left(\frac{H}{d_0}\right)^{1/2}\left(1+\frac{H^{3/2}}{P^{1/2}k}\right)X^{\epsilon}|\tau(\chi)|.$$

Par $|\tau(\chi)| \leq k/\sqrt{d}$ et le fait qu'il existe $\varphi(k^2/d)$ caractères de Dirichlet

$$\begin{aligned}&\ll H^{1/2}P^{-1}\left(1+\frac{H^{3/2}}{P^{1/2}K}\right)X^{\epsilon}\sum_{K\leq k<2K}k\sum_{d|k^2}\frac{1}{(dd_0)^{1/2}}\\&\ll H^{1/2}P^{-1}(K^2+H^{3/2}KP^{-1/2})X^{\epsilon}.\end{aligned}$$

Rappelons que $H \ll (P/A)^{1+\epsilon}$ et $K \ll (P/B)^{1+\epsilon}$, alors $S_1 \ll X^{-\epsilon}$ pour $P \ll X^{7/9-\epsilon}$. La restriction du support de $\hat{\phi}$ provient de ce fait.

La rareté des caractères principaux compensera les facteurs ajoutés dans les autres cas. Il y aura $\ll k^{\epsilon}$ caractères $\chi$ modulo $k^2/d$ tels que $\bar{\chi}^3$ soit principal. Le caractère $\chi\psi_4(k|\cdot)$ est principal seulement pour $\chi = \psi_4(k|\cdot)\chi_0$. Alors, si les deux caractères sont principaux, $\bar{\chi}^3 = \overline{(\psi_4(k|\cdot)\chi_0)}^3$ est principal. Donc $\psi_4(k|\cdot)$ doit être principal et alors $\chi = \chi_0$. Il y a donc une seule possibilité et $k$ doit être un carré.

**Cas 2.** Seul $\bar{\chi}^3$ est principal. Le facteur $\varphi(k^2/d) \asymp k^2/d$ est gagné par la rareté de tels caractères. Par le même principe que pour $S_1$,

$$S_2 \ll HP^{-1}\left(1+\frac{H^{3/2}}{P^{1/2}K}\right)X^{\epsilon}\sum_{K\leq k<2K}k^{-1}\sum_{d|k^2}\frac{\sqrt{d}}{d_0}.$$

La double somme interne est $\ll 1$. En se rappelant de la définition de $d_0$, il suffit de voir que

$$\sum_{d|k^2}\frac{\sqrt{d}}{d_0} = \prod_{p^v\|k}\sum_{0\leq j\leq 2v}\frac{p^{j/2}}{p^{3\lceil j/3\rceil}} = \prod_{p^v\|k}\left(1+\mathcal{O}\left(\frac{1}{p^{3/2}}\right)\right) \ll 1$$

Par les bornes usuelles sur $H$, $K$ et $K \gg 1$, $S_2 \ll X^{-\epsilon}$ pour $P \ll X^{5/6-\epsilon}$.

**Cas 3.** Les deux caractères sont principaux. Alors $k = \square$ et $|\tau(\chi)| = 1$. Utilisons plutôt la borne triviale $Q(d,k,\chi) \ll HP^{-1/2}/d_0$. On a

$$S_3 \ll HP^{-1/2} \sum_{\substack{k=\square \\ K\leq k<2K}} k^{-2} \sum_{d|k^2} \frac{d}{d_0} = HP^{-1/2} \sum_{K^{1/2}\leq l<(2K)^{1/2}} l^{-4} \sum_{d|l^4} \frac{d}{d_0}.$$

Similairement au cas 2,

$$\sum_{d|l^4} \frac{d}{d_0} = \prod_{p^v\|l} \sum_{0\leq j\leq 4v} \frac{p^j}{p^{\lceil j/3\rceil}} \ll l^\epsilon \prod_{p^v\|l} p^{4v-2} = l^{4+\epsilon} \prod_{p|l} p^{-2}.$$

Donc

$$\begin{aligned} S_3 &\ll HP^{-1/2}X^\epsilon \sum_{K^{1/2}\leq l<(2K)^{1/2}} \prod_{p|l} p^{-2} \\ &\ll HP^{-1/2}K^{-1/2}X^\epsilon \ll X^{-\epsilon} \end{aligned}$$

lorsque $K \gg PX^{-2/3+\epsilon}$.

Il reste à borner lorsque $K \ll PX^{-2/3+\epsilon}$. Revenons à la définition de $S(H,K,P)$. La méthode de Weyl [10] nous donne

$$\sum_{P\leq p<2P} \left| \sum_{H\leq h<2H} e\left(\frac{h^3\bar{k}^2}{p}\right) \right| \ll (H^{3/4}P + HP^{3/4} + H^{1/4}P^{5/4})(HKP)^\epsilon.$$

On applique par sommation partielle cette estimation à la définition de $S(H,K,P)$ pour obtenir

$$S(H,K,P) \ll (H^{3/4}P^{-1/2} + HP^{-3/4} + H^{1/4}P^{-1/4})KX^\epsilon.$$

Dans notre cas, puisque $k = \square$,

$$\begin{aligned} S_3 &\ll (H^{3/4}P^{-1/2} + HP^{-3/4} + H^{1/4}P^{-1/4})K^{1/2}X^\epsilon \\ &\ll X^{-\epsilon} \end{aligned}$$

lorsque $P \ll X^{7/9-\epsilon}$ par notre borne sur $K$ et celle usuelle sur $H$.

**Cas 4.** Seul $\chi\psi_4(k|\cdot)$ est principal. Donc, le conducteur de $\chi$ est égal au conducteur de $\psi_4(k|\cdot)$ valant $k^*$ la partie sans carré de $k$ à un facteur 2 ou 4 près. On a donc $k^*|k^2d^{-1}$ et alors $d|k^2(k^*)^{-1}$. La borne dans ce cas est

$$S_4 \ll H^{1/2}P^{-1/2}\left(1 + \frac{H^{3/2}}{P^{1/2}K}\right) X^\epsilon \sum_{K\leq k<2K} k^{-1} \sum_{d|k^2(k^*)^{-1}} \left(\frac{d}{d_0}\right)^{1/2}.$$

Encore une fois

$$\sum_{d|k^2(k^*)^{-1}} \left(\frac{d}{d_0}\right)^{1/2} \asymp \prod_{p^v\|k} \sum_{j=0}^{v} \left(\frac{p^j}{p^{\lceil j/3\rceil}}\right)^{1/2} \ll k^\epsilon k^{1/2} \prod_{p|k} p^{-1/2}.$$

Alors

$$S_4 \ll H^{1/2}P^{-1/2}X^{\epsilon}\left(1+\frac{H^{3/2}}{P^{1/2}K}\right).$$

Cette borne est $\ll X^{-\epsilon}$ si $H^2K^{-1} \ll PX^{-\epsilon}$. Nous allons voir l'utilité de cette condition à la fin de ce cas.

Trouvons une autre borne pour $Q(d,k,\chi)$ qui vaut dans notre cas

$$Q(d,k,\chi) = \sum_{P\leq p<2P}\sum_{H/d_0\leq h_0<2H/d_0} e\left(\frac{-h_0^3d_0^3}{pk^2}\right)\bar{\chi}^3(h_0)U(h_0,d_0,k,p).$$

On applique le Lemme 7.4 à $Q(d,k,\chi)$ par sommation partielle pour obtenir

$$Q(d,k,\chi) \ll \left(\frac{H}{d_0}\right)^{1/2}P^{-1/2} + P^{-1/4}H^{1/4}k^{1/2}d_0^{-1}X^{\epsilon}.$$

Par cette borne

$$S_4 \ll \sum_{K\leq k<2K}\sum_{d|k^2} d^{1/2}k^{-1}\left(\frac{H^{1/2}}{d_0^{1/2}P^{1/2}} + \frac{H^{1/4}k^{1/2}}{P^{1/4}d_0}X^{\epsilon}\right).$$

Par les mêmes techniques utilisées dans les autres cas, on estime

$$S_4 \ll \frac{H^{1/2}K^{1/2}}{P^{1/2}}X^{\epsilon} + \frac{H^{1/4}K^{1/2}}{P^{1/4}}X^{\epsilon}.$$

Les bornes usuelles sur $H$ et $K$ montrent que le premier terme est $\ll X^{-\epsilon}$ pour $P \ll X^{5/6-\epsilon}$. On peut assumer $H^2K^{-1} \gg PX^{-\epsilon}$ par notre résultat ci-haut. Avec $H \ll (P/A)^{1+\epsilon}$ la borne usuelle, le second terme est $\ll X^{-\epsilon}$ lorsque $P \ll X^{5/6-\epsilon}$.

Nous avons donc prouvé que $S(H,K,P) \ll X^{-\epsilon}$ pour toutes les $\ll \log X$ parties composant la somme (7.1) ce qui termine la preuve. □

## 7.3 Restriction de minimalité

S'il n'y a aucun nombre premier $q > 3$ tel que $q^4|a$ et $q^6|b$, la courbe $y^2 = x^3 + ax + b$ sera minimale. Les courbes non minimales sont indésirables, car elles sont isomorphes à des courbes déjà sélectionnées. Leur discriminant peut également être beaucoup plus large que leur conducteur dans le sens que la condition $\log N \asymp \log X$ ne sera pas respectée en moyenne sur ces courbes. Montrons que leur impact est négligeable en les considérant par simplicité dans notre famille.

Premièrement, la densité asymptotique des nombres sans puissance de 12 est $1/\zeta(12) \approx 0,9998$. S'il n'existe aucun $p > 3$ tel que $p^{12}|\Delta$, la courbe est minimale (l'implication inverse n'est pas vraie). Puisqu'il n'y a aucune raison que le déterminant favorise des nombres divisibles par $p^{12}$, on s'attend à ce qu'une proportion positive de nos courbes soient minimales. Ceci implique que le nombre de courbes minimales sélectionnées par rapport à $X$ sera $\asymp X^{5/6}$ qui est la taille de notre famille. Donnons maintenant un argument plus rigoureux en effectuant les calculs à nouveau sous cette restriction.

La condition du conducteur est directe par nos calculs déjà faits sans restriction. Calculons l'estimation centrale

$$\sum_{\substack{a \\ q^4|a \Rightarrow q^6 \nmid b}} \sum_{b} P(E,\phi)\omega\left(\frac{a}{A},\frac{b}{B}\right) = \sum_{p>3} \frac{2\log p}{p\log X}\hat{\phi}\left(\frac{\log p}{\log X}\right) \sum_{\substack{a \\ q^4|a \Rightarrow q^6 \nmid b}} \sum_{b} \lambda_{a,b}(p)\omega\left(\frac{a}{A},\frac{b}{B}\right)$$

par définition. Le truc afin d'éliminer la condition sur les sommes est d'introduire la fonction de Möbius.

$$= \sum_{p>3} \frac{2\log p}{p\log X}\hat{\phi}\left(\frac{\log p}{\log X}\right) \sum_{d} \mu(d) \sum_{a} \sum_{b} \lambda_{ad^4,bd^6}(p)\omega\left(\frac{ad^4}{A},\frac{bd^6}{B}\right)$$

Le nombre de termes comptés par $\omega$ lorsque $\log X < d \leq X^{1/12}$ est en $o(AB)$. Ils sont donc négligeables asymptotiquement par rapport aux $\asymp AB\log X$ termes lorsque $d < \log X$. On considère donc

$$\sum_{p>3} \frac{2\log p}{p\log X}\hat{\phi}\left(\frac{\log p}{\log X}\right) \sum_{\substack{d \leq \log X \\ (d,p)=1}} \mu(d) \sum_{a} \sum_{b} \lambda_{ad^4,bd^6}(p)\omega\left(\frac{ad^4}{A},\frac{bd^6}{B}\right).$$

On applique la Proposition 7.1

$$= AB \sum_{p>3} \sum_{\substack{d \leq \log x \\ (d,p)=1}} \frac{\mu(d)}{d^{10}p^2} \sum_{h} \sum_{k} \sum_{\substack{\alpha \pmod p \\ \beta \pmod p}} \sum \lambda_{\alpha d^4,\beta d^6}(p) e\left(\frac{\alpha h+\beta k}{p}\right) \Phi(h,k,p)$$

où

$$\Phi(h,k,p) = \frac{2\log p}{p\log X}\hat{\phi}\left(\frac{\log p}{\log X}\right)\hat{\omega}\left(\frac{hA}{d^4p},\frac{kB}{d^6p}\right).$$

On peut appliquer le Lemme 7.2 après le changement de variables $\alpha \to \alpha d^{-4}$ et $\beta \to \beta d^{-6}$ pour obtenir

$$-\frac{AB}{\log X}\sum_{p>3}\epsilon_p \frac{2\log p}{p^{3/2}}\hat{\phi}\left(\frac{\log p}{\log X}\right) \sum_{\substack{d \leq \log X \\ (d,p)=1}} \frac{\mu(d)}{d^{10}} \sum_{h} \sum_{k} \left(\frac{k}{p}\right) e\left(\frac{-h^3\bar{k}^2}{p}\right)\hat{\omega}\left(\frac{hA}{d^4p},\frac{kB}{d^6p}\right).$$

On inverse l'ordre de sommation pour avoir la somme sur $d$ à l'extérieur et on enlève la restriction $(d,p)=1$. Une différence maintenant est que $A$ et $B$ seront plus petits, ce qui n'a aucun effet sur nos bornes supérieures. La somme sur $d$ introduira au plus une puissance de $\log X$. Comme nous avons montré que $S(H,K,P) \ll X^{-\epsilon}$ pour tout $\epsilon > 0$ et que $\log^n X \ll X^{\epsilon}$, ces facteurs logarithmiques n'auront aucun effet asymptotiquement.

## 7.4 Application

Une borne sur le rang analytique moyen de $L(s,E)$ est calculée sur notre famille par rapport au support permis de $\hat{\phi}$. On calcule la densité des courbes avec rang $m$ ainsi

$$p_m(X) := \frac{1}{W_X(\mathcal{F})} \sum_{\substack{E \in \mathcal{F} \\ \operatorname{ord}_{s=1} L(s,E)=m}} \omega_X(E).$$

La moyenne du rang analytique sur notre famille est alors

$$r = \lim_{X\to\infty} \sum_{m=1}^{\infty} mp_m(X)$$

si la limite existe. Par une fonction test positive telle que $\phi(0) = 1$ et $\mathrm{supp}(\hat{\phi}) \subset [-\nu, \nu]$, notre théorème principal sur le type de symétrie de notre famille implique l'inégalité

$$\sum_{m\geq 1} mp_m(X) \leq g + o(X)$$

où

$$g = \int_{\mathbb{R}} \hat{\phi}(y)\hat{w}(O)(y)dy.$$

On voit cette inégalité en remarquant que

$$\sum_{m\geq 1} mp_m(X) \leq \frac{\mathcal{D}(\omega_X, \mathcal{F}, \phi)}{W_X(\mathcal{F})}$$

par positivité de $\phi$ puisque la somme sur $m$ ne somme que les zéros au point central.

Le choix judicieux de $\phi$ est

$$\phi(t) = \left(\frac{\sin(\pi\nu t)}{\pi\nu t}\right)^2, \qquad \hat{\phi}(y) = \frac{1}{\nu}\left(1 - \frac{|y|}{\nu}\right).$$

Alors $g = 1/\nu + 1/2$. En utilisant $\nu = 7/9 - \epsilon$ par les travaux de Young, on obtient $r \leq 25/14 + \epsilon$ pour tout $\epsilon > 0$. Aucune restriction sur le support impliquerait une moyenne de $1/2$. Notons qu'une borne de $1/2$ implique que la majorité des courbes ont rang analytique 0 ou 1.

Par les travaux de Gross-Zagier et Kolyvagin, si le rang analytique est 0 ou 1 alors le rang algébrique est 0 ou 1 respectivement pour les courbes elliptiques sur $\mathbb{Q}$. Comme $25/14 < 2$, les travaux de Young montrent qu'une proportion positive de courbes satisfassent la conjecture de Birch et Swinnerton-Dyer sous la grande hypothèse de Riemann. Bhargava et Shankar ont récemment montré inconditionnellement que la moyenne du rang algébrique de la famille de toutes les courbes $E/\mathbb{Q}$ est $\leq 0,885$.

# 8 La symétrie de la famille des courbes elliptiques sur $\mathbb{Q}(i)$

Nous prouverons dans cette section le théorème suivant.

**Théorème 8.1.** *Pour $\phi \in \mathcal{S}(\mathbb{R})$ tel que supp$(\hat{\phi}) \subset (-5/9, 5/9)$, nous avons*

$$\frac{\mathcal{D}(\omega_X, \mathcal{F}, \phi)}{W_X(\mathcal{F})} \sim \hat{\phi}(0) + \frac{1}{2}\phi(0) \quad \text{lorsque } X \to \infty$$

*pour $\mathcal{F}$ la famille des courbes elliptiques sur $\mathbb{Q}(i)$.*

La symétrie de notre famille est donc orthogonale dans la mesure de ce que nous pouvons calculer. Le début de la preuve sera exactement le même. Il faudra donc vérifier la condition du conducteur et calculer l'estimation centrale.

## 8.1 Le corps $\mathbb{Q}(i)$

Le corps $\mathbb{Q}(i)$ est défini ainsi

$$\mathbb{Q}(i) = \{a + bi \,|\, a, b \in \mathbb{Q}\},$$

où $i$ respecte $i^2 = -1$. L'anneau $\mathbb{Z}[i] = \{a + bi \,|\, a, b \in \mathbb{Z}\} \subset \mathbb{Q}(i)$ joue le rôle des entiers de ce corps et correspond aux éléments de $\mathbb{Q}(i)$ qui sont une solution d'un polynôme unitaire avec coefficients en $\mathbb{Z}$. Un nombre premier peut se factoriser à l'intérieur de $\mathbb{Z}[i]$

$$p = (\lambda + \kappa i)(\lambda - \kappa i) = \lambda^2 + \kappa^2.$$

Un résultat de la théorie des nombres élémentaire est qu'un nombre premier $p$ est une somme de deux carrés si et seulement si $p \equiv 1 \pmod 4$. Il y aura donc deux types de nombres premiers. Si $p \equiv 3 \pmod 4$, $p$ est premier à l'intérieur de $\mathbb{Z}[i]$. Si $p \equiv 1 \pmod 4$, $p$ se factorise en deux nombres premiers distincts $\lambda + \kappa i$ et $\lambda - \kappa i$. Tous les nombres premiers de $\mathbb{Z}[i]$ proviennent d'un nombre premier de $\mathbb{Z}$. Pour le cas spécial $p = 2$, nous avons $2 = (1+i)(1-i)$. Par contre, $1 - i = -i(1 + i)$. Comme $i$ et $-1$ sont des unités dans cet anneau,$1 + i$ et $1 - i$ correspondent au même nombre premier.

En analogie avec $\mathbb{Z}/p\mathbb{Z}$, on considère les quotients $\mathbb{Z}[i]/\mathfrak{p}$ où $\mathfrak{p}$ est un idéal premier généré par un nombre premier de $\mathbb{Z}[i]$. Dans le cas où $\mathfrak{p} = (p)$ lorsque $p \equiv 3 \pmod 4$, un choix possible de représentants est $\{m + ni \,|\, 0 \leq m, n \leq p - 1\}$ et

$$a + bi \equiv \tilde{a} + \tilde{b}i \pmod{\mathfrak{p}}$$

où $\tilde{a}, \tilde{b}$ est la réduction de $a$ et $b$ modulo $p$. Dans le cas $\mathfrak{p} = (\lambda + \kappa i)$ lorsque $p = \lambda^2 + \kappa^2$, un choix possible de représentants est $\{m \,|\, 0 \leq m \leq p - 1\}$.

Pour $\mathfrak{p}$ un idéal premier de $\mathbb{Z}[i]$, on a $\mathfrak{p} \cap \mathbb{Z} = (p)$ où $p$ est le nombre premier associé au générateur de $\mathfrak{p}$.

La norme d'un idéal premier $\mathfrak{p}$ est définie comme étant $N\mathfrak{p} := (\mathfrak{p} \cap \mathbb{Z})^r = (p^r)$ où $r$ est le degré de l'extension $\mathbb{Z}[i]/\mathfrak{p} \supset \mathbb{Z}/(\mathfrak{p} \cap \mathbb{Z}) = \mathbb{Z}/p\mathbb{Z}$. En utilisant plutôt les générateurs de ces idéaux, la norme peut être définie sur les nombres premiers de $\mathbb{Z}[i]$ vers les puissances de nombres premiers de $\mathbb{Z}$. En d'autres mots, soit $P$ un nombre premier de $\mathbb{Z}[i]$, alors $NP = |\mathbb{Z}[i]/(P)|$. Dans notre cas, les nombres premiers $\lambda + \kappa i$ provenant de la factorisation d'un nombre premier $p$ ont norme $p$. Les nombres premiers demeurant stables ont norme $p^2$.

## 8.2 La fonction $L$ de Hasse-Weil pour $E/\mathbb{Q}(i)$

Les définitions utilisées dans le cas de $E/\mathbb{Q}$ se généralisent naturellement. Une courbe elliptique $E/\mathbb{Q}(i)$ se représente par une courbe isomorphe $E : y^2 = x^3 + ax + b$ où $a, b \in \mathbb{Z}[i]$ par substitutions. Cette forme est globalement minimale si $P^{12} \nmid \Delta$ pour tout nombre premier $P$ de $\mathbb{Z}[i]$.

Le principe de réduction sur un corps fini est le même. Soit $P \in \mathbb{Z}[i]$ un nombre premier et $\mathfrak{p} = (P)$, on dénote par $\#E(\mathbb{Z}[i]/\mathfrak{p})$ le nombre de solutions projectives lorsque la courbe est réduite modulo $P$. La courbe réduite modulo $P$ est singulière si et seulement si $P|\Delta$.

On définit le conducteur similairement si $E$ est sous la forme minimale

$$N = \prod_{P|\Delta} P^{f_P}.$$

On aura toujours $N|\Delta$ pour tout $P$. Le conducteur est en fait défini comme étant l'idéal généré par $N$. La Section 8.4 donnera plus de détails sur le conducteur et comment y attacher une valeur entière.

La fonction $L$ de Hasse-Weil est donnée par

$$L(s, E) = \prod_{P|N} (1 - a_P NP^{-s})^{-1} \prod_{P\nmid N} (1 - a_P NP^{-s} + NP^{1-2s})^{-1}$$

où

$$a_P = NP + 1 - \#E(\mathbb{Z}[i]/(P)).$$

En développant le produit en série de Dirichlet, les coefficients à $p$ vaudront

$$\lambda_E(p) = \begin{cases} a_P + a_{\bar{P}} & \text{si } p \equiv 1 \pmod 4 \\ 0 & \text{si } p \equiv 3 \pmod 4 \end{cases}$$

où $p = P\bar{P}$ la factorisation de $p$. La factorisation est en fait unique à unité près, mais $a_P$ ne change pas par la multiplication de $P$ par une unité. Les $p \equiv 3 \pmod 4$ ont $NP = p^2$ ce qui explique pourquoi ils n'ont aucune contribution aux coefficients à $p$. Comme les calculs s'effectuent seulement sur les $\lambda_E(p)$, nous nous concentrons maintenant sur les $p \equiv 1 \pmod 4$. Le conducteur de la fonction $L$ sera la norme de l'idéal généré par le conducteur défini ci-haut.

Puisque $\mathbb{Z}[i]/\mathfrak{p} \cong \mathbb{Z}/p\mathbb{Z}$ pour $\mathfrak{p} = (P)$ et que $a_P$ ne dépend pas du choix des représentants, on a encore une fois

$$a_P = -\sum_{x \in \mathbb{Z}/p\mathbb{Z}} \left( \frac{x^3 + \tilde{a}x + \tilde{b}}{p} \right)$$

où $\tilde{a}$ et $\tilde{b}$ sont la réduction modulo $\mathfrak{p}$ en choisissant les représentants $\{0, 1, \ldots, p-1\}$ pour $\mathbb{Z}[i]/\mathfrak{p}$. On peut calculer efficacement la réduction de $a = \kappa + i\lambda$ par $\tilde{a} = \kappa + \rho\lambda$ (réduit modulo $p$) où $\rho \equiv -1 \pmod p$.

## 8.3 Sommation de Poisson

La principale différence calculatoire entre notre famille et celle des courbes elliptiques sur $\mathbb{Q}$ sera la sommation de Poisson. Au lieu de sommer sur les $m \in \mathbb{Z}$ tels que $m \equiv l \pmod p$, nous

sommerons sur les $m, n \in \mathbb{Z}$ tels que $m + ni \equiv l \pmod{\mathfrak{p}}$ où $l \in \{0, 1, \ldots, p-1\}$. Ainsi, nous pourrons séparer la somme sur notre famille de courbes en progression arithmétique, car $a_P$ ne dépend que de la congruence de $a$ et $b$ modulo $\mathfrak{p}$. Définissons d'abord

$$\wedge_{\mathfrak{p}}^{l} := \{(m, n) \in \mathbb{Z}^2 \,|\, m + ni \equiv l \bmod \mathfrak{p}\}.$$

**Proposition 8.2** (Poisson) *Soit $\omega \in \mathcal{S}(\mathbb{R}^2)$ l'espace de Schwartz, $M, N \in \mathbb{R}$ positifs et $l \in \mathbb{N}$. Alors*

$$\sum_{(m,n)\in\wedge_{\mathfrak{p}}^{l}} \omega\left(\frac{m}{M}, \frac{n}{N}\right) = \frac{MN}{p} \sum_{(r,s)\in\mathbb{Z}^2} e\left(\frac{l}{p}(r\kappa - s\lambda)\right) \hat{\omega}\left(\frac{M}{p}(r\kappa - s\lambda), \frac{N}{p}(r\lambda + s\kappa)\right)$$

*où $p \equiv 1$ (mod 4), $p = \kappa^2 + \lambda^2$ et $\mathfrak{p} = (\kappa + \lambda i)$.*

*Preuve.* Le résultat sans restriction de congruence est connu (voir Serre [14] Proposition 15).

$$\sum_{(m,n)\in\mathbb{Z}^2} \omega(m, n) = \sum_{(r,s)\in\mathbb{Z}^2} \hat{\omega}(r, s)$$

Les éléments de $\mathbb{Z}[i]$ congrus à $l$ modulo $\mathfrak{p}$ sont $(\kappa+\lambda i)(m+ni)+l = (m\kappa-n\lambda+l)+i(m\lambda+n\kappa)$ pour tout $m, n \in \mathbb{Z}$. Par la substitution

$$\omega(m, n) \to \omega\left(\frac{m\kappa - n\lambda + l}{M}, \frac{m\lambda + n\kappa}{N}\right)$$

nous avons qu'à calculer

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \omega\left(\frac{x\kappa - y\lambda + l}{M}, \frac{x\lambda + y\kappa}{N}\right) e(-xr)e(-ys)dxdy.$$

On effectue le changement de variable

$$x \to \frac{1}{\kappa}(Mx + y\lambda - l)$$

pour obtenir

$$\frac{M}{\kappa}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \omega\left(x, \frac{1}{\kappa N}(\lambda Mx + py - l\lambda)\right) e\left(-\frac{1}{\kappa}(Mx + y\lambda - l)r\right) e(-ys)dxdy.$$

On change l'ordre d'intégration et on effectue le changement de variable

$$y \to \frac{1}{p}(\kappa Ny - \lambda Mx + l\lambda)$$

pour obtenir

$$\frac{MN}{p}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \omega(x, y)e\left(-\frac{r}{\kappa}(Mx + \frac{\lambda}{p}(\kappa Ny - \lambda Mx + l\lambda) - l)\right) e\left(\frac{-(\kappa Ny - \lambda Mx + l\lambda)s}{p}\right) dydx.$$

Le résultat suit en rassemblant les termes en $x$, les termes en $y$ et les termes sans dépendance de l'exponentielle. □

## 8.4 Condition du conducteur

Notre famille $\mathcal{F}$ sera composée de courbes elliptiques de la forme

$$E_{a,b,c,d} : y^2 = x^3 + (a+bi)x + (c+di)$$

pour $a, b, c$ et $d \in \mathbb{Z}$. On utilisera donc une fonction lisse $\omega(x, y, z, w)$ à support compact de $\mathbb{R}^4 \to \mathbb{R}$ afin de sélectionner les courbes. Son discriminant vaut $\Delta_E = -16(4(a+bi)^3 + 27(b+ci)^2)$. Notons que $\Delta_E \in \mathbb{Z}[i]$, nous allons voir comment y associer une valeur entière pour les calculs. Définissons d'abord *l'idéal du discriminant minimal de* $E/\mathbb{Q}(i)$. Pour chaque nombre premier $v$ de $\mathbb{Z}[i]$, il existe un modèle pour $E$ tel que la puissance de $v$ divisant $\Delta_E$ est minimale. Dénotons cette puissance $\mathrm{ord}_v(\Delta_E^{\min,v}) \geq 0$. Alors, l'idéal du discriminant minimal est donné par

$$\mathcal{D}_E := \prod_v \mathfrak{p}_v^{\mathrm{ord}_v(\Delta_E^{\min,v})}$$

où le produit est sur tous les nombres premiers $v$ de $\mathbb{Z}[i]$ et $\mathfrak{p}_v$ est l'idéal premier associé à $v$. Le conducteur de $E$ sera également un idéal de $\mathbb{Z}[i]$ et vaut

$$\mathcal{N}_E := \prod_v \mathfrak{p}_v^{f_v}$$

où

$$f_v = \begin{cases} 0 & E \text{ est de bonne réduction à } v \\ 1 & E \text{ est de réduction multiplicative à } v \\ 2 & E \text{ est de réduction additive à } v \end{cases}$$

pour les $v$ tels que $\mathrm{char}(\mathbb{Z}[i]/(\mathfrak{p}_v)) \neq 2, 3$. La définition de $f_v$ pour $v = 1 \pm i$, les seuls nombres premiers tels que le quotient n'aura pas la caractéristique voulue (char=2), est plus compliquée. Il peut être calculé par la formule de Ogg-Saito. Cette formule nous montre également que les idéaux premiers composant $\mathcal{N}_E$ correspondent exactement à ceux de $\mathcal{D}_E$ et que $\mathcal{N}_E | \mathcal{D}_E$ en idéaux.

Le conducteur de la fonction $L(s, E/\mathbb{Q}(i))$ doit être un nombre entier, on introduit alors [13] la *norme du conducteur*. Il s'agit simplement de prendre la norme absolue de l'idéal $\mathcal{N}_E$. On le dénotera par $N(\mathcal{N}_E)$. Notons que $N(\mathcal{N}_E) | N(\mathcal{D}_E)$ et qu'ils sont composés des mêmes nombres premiers. Pour cette raison, nous n'allons pas sélectionner les courbes ayant $|\Delta| \asymp X$, mais plutôt celles ayant $N(\mathfrak{i}_\Delta) \asymp X$ où $\mathfrak{i}_\Delta = (\Delta)$ et $|\cdot|$ est la norme complexe usuelle. Une autre raison est que les facteurs premiers de $|\Delta|$ ne sont pas nécessairement à des puissances entières, tandis que c'est le cas pour $N(\mathfrak{i}_\Delta)$. Par la relation $|\Delta|^2 = N(\mathfrak{i}_\Delta)$, la fonction à support compact devra être

$$\omega_X(E_{a,b,c,d}) := \omega\left(\frac{a}{X^{1/6}}, \frac{b}{X^{1/6}}, \frac{c}{X^{1/4}}, \frac{d}{X^{1/4}}\right)$$

qui nous assure $|\Delta| \asymp X^{1/2} \Rightarrow N(\mathfrak{i}_\Delta) \asymp X$. Le nombre de courbes considérées par rapport à $X$ est donc

$$|W_X(\mathcal{F})| = \left|\sum_{E \in \mathcal{F}} \omega_X(E)\right| \asymp X^{5/6}.$$

La modularité des courbes elliptiques de notre famille est assumée ainsi que la grande hypothèse de Riemann. Nous pouvons donc utiliser le même point de départ qu'au Chapitre 7.

Premièrement, il faut montrer la condition du conducteur.

$$\sum_{E\in\mathcal{F}} \frac{\log N(\mathcal{N}_E)}{\log X}\omega_X(E) \sim W_X(\mathcal{F}) \qquad \text{lorsque } X\to\infty$$

Le début de la preuve est exactement le même, aucun ajustement n'est nécessaire. Nous devons alors montrer que

$$\sum_{E\in\mathcal{F}} \log\left(\frac{N(\mathfrak{i}_{\Delta_E})}{N(\mathcal{N}_E)}\right)\omega_X(E) \ll |W_X(\mathcal{F})|.$$

Nous avons que

$$\log\left(\frac{N(\mathfrak{i}_{\Delta_E})}{N(\mathcal{N}_E)}\right) = 2\log\left|\frac{\Delta_E}{N_E}\right|$$

$$\leq 2\sum_{\substack{v^\alpha\|\Delta_E\\ \alpha>0}} \log|v|^{\alpha-1} \asymp \sum_{\substack{v^\alpha\|\Delta_E\\ \alpha>0}} \log p_v^{\alpha-1}$$

où $p_v$ est le nombre premier sous $v$. Alors, par définition, il faut montrer

$$\sum_{a,b,c,d}\sum_{\substack{v^\alpha\|\Delta_E\\ \alpha>0}} \omega\left(\frac{a}{A},\frac{b}{B},\frac{c}{C},\frac{d}{D}\right)\log p_v^{\alpha-1} \ll |W_X(\mathcal{F})|$$

où $A=B=X^{1/6}$ et $C=D=X^{1/4}$. Nous souhaitons appliquer la même stratégie que pour la famille $E/\mathbb{Q}$. Dans notre cas, il faudra définir $\gamma$ de cette manière analogue: $v^\gamma\|a+bi$, où $v$ est un nombre premier de $\mathbb{Z}[i]$. L'idée sera encore de contrôler les diviseurs premiers des discriminants.

On commence par changer l'ordre de sommation. Limitons-nous aux $v$ tels que $p_v>3$. On estime d'abord les termes où $\alpha<3\gamma$. On aura $v^\alpha\|(c+di)^2$. Notre somme est donc

$$\leq \sum_{\substack{N(v^\alpha)\ll X\\ \alpha>0}} \sum_{\alpha/3<\gamma\ll\log A}\sum_{a+bi}\sum_{c+di} \tilde{\omega}\left(\frac{v^\gamma(a+bi)}{A},\frac{v^{\alpha/2}(c+di)}{C}\right)\log p_v^{\alpha-1}$$

où $\tilde{\omega}(a+bi,c+di):=\omega(a,b,c,d)$. On sépare cette somme en deux selon la congruence modulo 4 du nombre premier sous $v$. Débutons par les $v$ tels que $p_v\equiv 1 \pmod 4$. Nous pouvons nous concentrer sur un seul $v$ au-dessus de $p_v$, car la somme sur les conjugués complexes sera de même ordre. Maintenant, la somme sur $\tilde{\omega}$ compte le nombre de courbes telles que

$$\left|\frac{v^\gamma(a+bi)}{A}\right| \asymp 1,$$

$$\left|\frac{v^{\alpha/2}(c+di)}{C}\right| \asymp 1.$$

Nous avons que $|v^\gamma| = p_v^{\gamma/2}$, le nombre de choix pour $(a,b,c,d)$ est donc

$$\ll \left(1+\frac{A}{p_v^{\gamma/2}}\right)^2\left(1+\frac{C}{p_v^{\alpha/4}}\right)^2.$$

La condition $N(v^\alpha) \ll X$ implique $p_v^\alpha \ll X$. En élargissant la somme sur tous les nombres premiers, notre somme est alors

$$\ll \sum_{\substack{p^\alpha \ll X \\ \alpha>0}} \sum_{\alpha/3<\gamma\ll \log A} \left(1+\frac{A}{p^{\gamma/2}}\right)^2 \left(1+\frac{C}{p^{\alpha/4}}\right)^2 \log p^{\alpha-1}.$$

Cette somme se borne exactement comme au Chapitre 7, le terme principal proviendra de

$$\sum_{\substack{p^\alpha \ll X \\ \alpha>0}} \frac{A^2C^2}{p^{\lfloor \alpha/3\rfloor+\alpha/2+1}} \log p^{\alpha-1} \ll X^{5/6} \asymp |W_X(\mathcal{F})|.$$

La somme sur les nombres premiers $p \equiv 3 \pmod 4$ est traitée de la même manière. Elle sera mieux bornée par le fait que $\alpha$ sera toujours pair et qu'il y aura moins de choix pour $(a,b,c,d)$ car $|v^\gamma| = p_v^\gamma$ dans ce cas-ci ($v = p_v$).

Montrons maintenant la borne pour les termes avec $\alpha > 3\gamma$. Concentrons-nous seulement sur le cas plus difficile, c'est-à-dire, les $v$ provenant de $p \equiv 1 \pmod 4$. Les autres $v$ sont traités de la même manière. Nous avons que $v^{3\gamma} \| (c+di)^2$, la somme à borner est donc

$$\sum_{\substack{N(v^\alpha)\ll X \\ \alpha>0}} \sum_{0\le\gamma<\alpha/3} {\sum_{a+bi}}' {\sum_{c+di}}' \tilde{\omega}\left(\frac{v^\gamma(a+bi)}{A}, \frac{v^{3\gamma/2}(c+di)}{C}\right) \log p_v^{\alpha-1}.$$
$$v^{\alpha-3\gamma} \| 4(a+bi)^3+27(c+di)^2$$

Puisque $(v, a+bi) = (v, c+di) = 1$, pour chaque $a+bi$ fixé, il y aura au plus 2 solutions $c+di$ (réduites modulo $(v^{\alpha-3\gamma})$) à la relation $4(a+bi)^3 + 27(c+di)^2 \equiv 0 \pmod{(v^{\alpha-3\gamma})}$. Ceci est dû à $\mathbb{Z}[i]/(v^\alpha) \cong \mathbb{Z}/p_v^\alpha\mathbb{Z}$. On divise donc par $p_v^{\alpha-3\gamma}$ le nombre de $c+di$ considéré. On obtient alors

$$\ll \sum_{\substack{p^\alpha \ll X \\ \alpha>0}} \sum_{0\le\gamma<\alpha/3} \left(1+\frac{A}{p^{\gamma/2}}\right)^2 \left(1+\frac{C}{p^{3\gamma/4+(\alpha-3\gamma)/2}}\right)^2 \log p^{\alpha-1}.$$

On revient essentiellement à la même somme que dans le cas $E/\mathbb{Q}$, la borne est alors $X^{5/6}$.

Il reste à traiter $\alpha = 3\gamma$ et les nombres premiers $v$ au-dessus de 2 et 3. Ces cas sont bornés de la même manière avec de mineures modifications. Les calculs ci-haut ont montré l'essentiel du problème, c'est-à-dire, comment revenir aux sommes que nous avons déjà bornées dans le cas de $E/\mathbb{Q}$.

## 8.5 L'estimation centrale

Appliquons les techniques utilisées par Brumer du chapitre précédent. Le but est de montrer

$$\mathcal{P}(\omega_X, \mathcal{F}, \phi) \ll \frac{X^{5/6}}{\log X}.$$

Par définition,

$$\mathcal{P}(\omega_X, \mathcal{F}, \phi) = \sum_a \sum_b \sum_c \sum_d P(E,\phi)\omega\left(\frac{a}{A}, \frac{b}{B}, \frac{c}{C}, \frac{d}{D}\right)$$

$$= \sum_{\substack{p>3 \\ p\equiv 1 (\mathrm{mod}\ 4)}} \frac{2\log p}{p\log X}\hat{\phi}\left(\frac{\log p}{\log X}\right)\sum_a\sum_b\sum_c\sum_d \lambda_E(p)\omega\left(\frac{a}{A},\frac{b}{B},\frac{c}{C},\frac{d}{D}\right)$$

où $A = B = X^{1/6}$ et $C = D = X^{1/4}$. Comme $\lambda_E(p) = a_P + a_{\bar{P}}$, la somme se séparera en deux. Il sera suffisant d'en considérer qu'une car les deux seront asymptotiquement les mêmes. Il est important de noter que $a_P$ dépend de $E$. Concentrons-nous maintenant sur les sommes en $a, b, c$ et $d$ où $p$ sera fixe. On a $p = \lambda^2 + \kappa^2$. Dénotons $P = \lambda + i\kappa$ et $\mathfrak{p} = (P)$. On sépare maintenant les $a + bi$ et $c + di$ en progressions arithmétiques.

$$\sum_a\sum_b\sum_c\sum_d a_P\,\omega\left(\frac{a}{A},\frac{b}{B},\frac{c}{C},\frac{d}{D}\right) = \sum_{\alpha\in\mathbb{Z}[i]/\mathfrak{p}}\sum_{\beta\in\mathbb{Z}[i]/\mathfrak{p}}\sum_{(a,b)\in\wedge_{\mathfrak{p}}^{\alpha}}\sum_{(c,d)\in\wedge_{\mathfrak{p}}^{\beta}} a_P\,\omega\left(\frac{a}{A},\frac{b}{B},\frac{c}{C},\frac{d}{D}\right)$$

La quantité $a_P$ sort des deux sommes sur $(a, b)$ et $(c, d)$ car elle ne dépend que de la congruence de $a + bi$ et $c + di$ modulo $\mathfrak{p}$. On peut donc appliquer la Proposition 8.2. Rappelons que $a_P$ s'exprime en somme de symboles de Legendre

$$a_P = -\sum_{x\in\mathbb{Z}/p\mathbb{Z}}\left(\frac{x^3+\tilde{\alpha}x+\tilde{\beta}}{p}\right),$$

où $\alpha = a + bi$ et $\beta = c + di$ et $\tilde{\cdot}$ est le représentant modulo $p$. La Proposition 8.2 appliquée deux fois nous donne alors

$$\sum_{\alpha\in\mathbb{Z}/p\mathbb{Z}}\sum_{\beta\in\mathbb{Z}/p\mathbb{Z}} a_P\frac{ABCD}{p^2}\sum_{h,k,r,s} e\left(\frac{\alpha}{p}(h\kappa-k\lambda)+\frac{\beta}{p}(r\kappa-s\lambda)\right)$$

$$\cdot\,\hat{\omega}\left(\frac{A}{p}(h\kappa-k\lambda),\frac{B}{p}(h\lambda+k\kappa),\frac{C}{p}(r\kappa-s\lambda),\frac{D}{p}(r\lambda+s\kappa)\right).$$

En changeant l'ordre de sommation, le Lemme 7.2 s'applique.

$$\frac{ABCD}{p^2}\sum_{h,j,r,s}\hat{\omega}(\dots)\left[-\epsilon_p p^{3/2}\left(\frac{r\kappa-s\lambda}{p}\right)e\left(\frac{-(h\kappa-j\lambda)^3\overline{(r\kappa-s\lambda)}^2}{p}\right)\right]$$

où $(\cdot|p)$ est le symbole de Legendre. On sort le facteur $p^{3/2}$ des crochets et on borne trivialement ce qui reste par 1.

La tâche est maintenant d'estimer le nombre de $h, j, r$ et $s$ tels que les paramètres de $\hat{\omega}$ ont tous $|\cdot| \leq 1$. La somme sera estimée par cette quantité par les propriétés de $\omega$. Par exemple, pour le premier paramètre, nous désirons avoir

$$-1 \leq \frac{A}{p}(h\kappa - j\lambda) \leq 1.$$

On pose alors pour la borne $\leq 1$ par exemple,

$$\frac{A}{p}(h\kappa - j\lambda) = 1 \Rightarrow h = \frac{p}{A\kappa} + \frac{\lambda}{\kappa}j.$$

Les deux premiers paramètres de $\hat{\omega}$ nous donneront ainsi un parallélogramme dans le plan $(h, j)$ contenant les points que l'on souhaite compter. Les deux derniers paramètres donneront

pareillement un parallélogramme dans le plan $(r, s)$. Le nombre total de points importants sera alors le produit du nombre de points des deux parallélogrammes. Dans le plan $(h, j)$, l'intersection des droites nous donne les sommets du parallélogramme.

$$s_1 = \left(\frac{-\kappa - \lambda}{A}, \frac{\kappa - \lambda}{A}\right)$$
$$s_2 = \left(\frac{\lambda - \kappa}{A}, \frac{-\kappa - \lambda}{A}\right)$$
$$s_3 = \left(\frac{\kappa - \lambda}{A}, \frac{\kappa + \lambda}{A}\right)$$

On estime le nombre de points discrets à l'intérieur du parallélogramme par son aire. Dans le cas où $\lambda$ et $\kappa$ sont très près l'un de l'autre, le parallélogramme sera trop mince pour que l'aire soit considérable. On ajoute donc la circonférence à l'estimation afin de régler ce problème. En calculant $\mathrm{dist}(s_1, s_2) = 2\sqrt{p}/A$ et $\mathrm{dist}(s_2, s_3) = \sqrt{8p}/A$, on borne l'aire par $\ll p/A^2$ et la circonférence par $\ll \sqrt{p}/A$ où les constantes implicites sont absolues. Le nombre de points sera alors $\ll 1 + \sqrt{p}/A + p/A^2$.

Revenons à notre somme principale. On retire la restriction $p \equiv 1 \pmod 4$ par positivité. Alors

$$\mathcal{P}(\omega_X, \mathcal{F}, \phi) \ll \frac{ABCD}{\log X} \sum_{p \leq P} \frac{\log p}{p^{3/2}} \left(1 + \frac{p}{A^2} + \frac{\sqrt{p}}{A}\right)\left(1 + \frac{p}{C^2} + \frac{\sqrt{p}}{C}\right)$$

$$\ll \frac{1}{\log X}\Big(ABCD + P^{1/2}AB + ABD\log P + P^{1/2}CD$$

$$+P^{3/2} + PD + BCD\log P + PB + P^{1/2}BD\Big).$$

Cette dernière borne est obtenue comme suit. Si $p$ est à une puissance positive, on borne trivialement par $p \leq P$ et on multiplie par $P$ le nombre de termes. Si $p$ est à une puissance négative, on borne par l'intégrale. Si $P \ll X^{5/9}$, on a

$$\mathcal{P}(\omega_X, \mathcal{F}, \phi) \ll \frac{X^{5/6}}{\log X}.$$

Le terme $P^{3/2}$ est le plus restrictif. Si $P = X^a$, il faut que $X^{3a/2} \leq X^{5/6}$ d'où le $a = 5/9$. Alors le support de $\hat{\phi}$ doit être contenu dans l'intervalle $(-5/9, 5/9)$ afin de satisfaire cette restriction sur les nombres premiers sommés. Par la Section 7.4, la moyenne du rang analytique de notre famille est $\leq 23/10$.

## 8.6 Directions futures de recherche

En conclusion, les techniques de Brumer afin de calculer le type de symétrie se sont assez bien ajustées à notre famille. La force du résultat, c'est-à-dire le support permis pour $\hat{\phi}$, n'a pas été compromise par cet ajustement. Ces calculs sont donc prometteurs en vue d'une généralisation vers une extension des rationnels arbitraire. Il serait également intéressant de se pencher sur la manière d'ajuster les travaux de Young afin d'augmenter le support.

Il serait d'intérêt d'imposer la restriction de minimalité afin d'éliminer des courbes isomorphes à l'intérieur de notre famille. Par contre, la Section 7.3 sur la minimalité semble montrer

que les résultats demeureraient les mêmes.

Un problème est qu'il n'est toujours pas su si les courbes sur une extension des rationnels sont modulaires. La formule explicite utilisée comme point de départ pourrait donc ne pas s'appliquer. Les courbes sur $\mathbb{Q}(i)$ sont probablement modulaires par la conjecture de Langlands qui dit que toutes les fonctions $L$ proviennent de formes automorphes qui sont une généralisation des formes modulaires.

Il est définitivement d'intérêt de calculer les symétries pour les autres familles de courbes elliptiques sur diverses extensions des rationnels $k \supset \mathbb{Q}$ afin d'obtenir des bornes sur la moyenne du rang analytique. La conjecture de Birch et Swinnerton-Dyer se généralise pour ces courbes, ces bornes pourraient donc se traduire en bornes sur le rang algébrique moyen. On pourrait espérer obtenir un support permis d'au moins $(-5/9, 5/9)$ pour les autres extensions, malgré le fait que cette étendue est insuffisante afin de distinguer le type de symétrie.

Les calculs de l'estimation centrale devraient se généraliser sans trop de problèmes. La sommation de Poisson devrait se traduire aisément pour $k$ en utilisant la même stratégie de preuve. Aussi, la fonction $L$ de Hasse-Weil se généralise naturellement pour tout $k$. Ensuite, seuls les nombres premiers de $k$ ayant norme 1 contribueront aux coefficients à $p$ de la série de Dirichlet de $L(s, E)$. Dans ce cas, $\mathcal{O}_k/\mathfrak{p} \cong \mathbb{Z}/p\mathbb{Z}$ et on pourra exprimer ses coefficients en somme de symboles de Legendre. Les théorèmes utilisés dans le cas $E/\mathbb{Q}$ s'appliquent donc directement comme nous avons pu le constater pour $E/\mathbb{Q}(i)$. Une difficulté résidera dans l'estimation du nombre de points comptés par la fonction test. Il s'agira d'estimer le nombre de points dans un polytope. La condition du conducteur semble toutefois être un plus grand défi à généraliser.

# Références

[1] H. Davenport, *Multiplicative Number Theory*, Third edition. Revised and with a preface by Hugh L. Montgomery. Graduate Texts in Mathematics, 74. Springer-Verlag, New York, 2000.

[2] D. Koukoulopoulos, *The distribution of prime numbers*, http://www.dms.umontreal.ca/~koukoulo/documents/notes/primes.pdf, Université de Montréal, 2015.

[3] H. Iwaniec and E. Kowalski, *Analytic number theory.* American Mathematical Society Colloquium Publications, 53. American Mathematical Society, Providence, RI, 2004.

[4] N. M. Katz and P. Sarnak, *Random Matrices, Frobenius Eigenvalues and Monodromy.* American Mathematical Society Colloquium Publications, 45. American Mathematical Society, Providence, RI, 1999.

[5] N. M. Katz and P. Sarnak, *Zeroes of Zeta Functions and Symmetry*, Bull. Amer. Math. Soc. **36**, 1-26, 1999.

[6] J. H. Silverman, *The Arithmetic of Elliptic Curves*, Second edition. Graduate Texts in Mathematics, 106. Springer-Verlag, New York, 2009.

[7] N. Koblitz, *Introduction to Elliptic Curves and Modular forms*, Second edition. Graduate Texts in Mathematics, 97. Springer-Verlag, New York, 1993.

[8] J. H. Silverman and J. T. Tate, *Rational Points on Elliptic Curves*, Second edition. Undergraduate Texts in Mathematics. Springer International Publishing, 2015.

[9] M. P. Young, *Low-Lying Zeros of Families of Elliptic Curves*, J. Amer. Math. Soc. **19**, no.1, 205-250, 2006.

[10] H. Weyl, *Ueber die Gleichverteilung von Zahlen mod. Eins*, Math. Ann. , **77** (1916) pp. 313352

[11] H. Iwaniec, W. Luo and P. Sarnak, *Low lying zeros of families of L-functions*, Inst. Hautes Études Sci. Publ. Math. no. 91, 55-131 (2001). MR1828743 (2002h:11081)

[12] M. Bhargava and A. Shankar, *The average size of the 5-Selmer group of elliptic curves is 6, and the average rank is less than 1*, arXiv:1312.7859 [math.NT]

[13] J. Bober, A. Deines, A. Klages-Mundt, B. Lévêque, R. A. Ohana, A. Rabindranath, P. Sharaba and W. Stein, *A database of elliptic curves over* $\mathbb{Q}(\sqrt{5})$ — *First report*, arXiv:1202.6612v2 [math.NT] 9 Jul 2012

[14] J.-P. Serre, *A Course in Arithmetic*, Graduate Texts in Mathematics, vol. 7, Springer-Verlag (1973).

[15] G. Harris and C. Martin, *The roots of a polynomial vary continuously as a function of the coefficients*. Proceedings of the American Mathematical Society, Volume 100, Number 2, June 1987.