Journées mathématiques X-UPS

Année 2002

La fonction zêta

Pierre COLMEZ

Arithmétique de la fonction zêta

Journées mathématiques X-UPS (2002), p. 41-196.
https://doi.org/10.5802/xups.2002-02

Les Éditions de l'École polytechnique
Route de Saclay
F-91128 PALAISEAU CEDEX
https://www.editions.polytechnique.fr

Centre de mathématiques Laurent Schwartz
CMLS, École polytechnique, CNRS,
Institut polytechnique de Paris
F-91128 PALAISEAU CEDEX
https://portail.polytechnique.edu/cmls/

Publication membre du
Centre Mersenne pour l'édition scientifique ouverte
www.centre-mersenne.org

Journées mathématiques X-UPS
2002, p. 41–196 doi: 10.5802/xups.2002-02

ARITHMÉTIQUE DE LA FONCTION ZÊTA

par

Pierre Colmez

Table des matières

Publication originelle dans Journées X-UPS 2002. La fonction zêta. Éditions de l'École polytechnique, 2002.

Introduction

La fonction zêta de Riemann est définie pour $\mathrm{Re}(s) > 1$ par la série $\zeta(s) = \sum_{n=1}^{+\infty} n^{-s}$ et elle possède un prolongement à tout le plan complexe avec une équation fonctionnelle reliant s à $1 - s$. Cette fonction et ses généralisations (fonctions zêta de Dedekind, de Hasse-Weil, fonctions L de Dirichlet, de formes modulaires...) jouent un rôle central en arithmétique et leurs valeurs aux entiers (valeurs spéciales, dans la terminologie en vigueur) contiennent une multitude de renseignements concernant l'arithmétique des objets attachés à ces fonctions. Un des slogans à la mode est d'ailleurs « les fonctions L savent tout ; à nous de les faire parler ».

Dans ce texte, nous nous concentrerons sur les valeurs aux entiers de la fonction ζ. Les valeurs aux entiers négatifs de la fonction ζ sont des nombres rationnels (par exemple, $\zeta(0) = -1/2$, $\zeta(-1) = -1/12$, $\zeta(-2) = 0$, ...) et on doit à Kummer le premier résultat concernant les renseignements cachés dans ces nombres. Il a en effet démontré que, si p est un nombre premier $\geqslant 3$ ne divisant pas les numérateurs de $\zeta(-1), \zeta(-3), \ldots, \zeta(2-p)$, alors p ne divise pas le nombre de classes d'idéaux du corps $\mathbf{Q}(e^{2i\pi/p})$ et en a déduit le théorème de Fermat pour un tel nombre premier. Ce n'est que récemment que Mazur et Wiles ont donné une formule donnant la puissance de p divisant exactement le numérateur de $\zeta(k)$, si k est un nombre entier $\leqslant 0$ impair (on a $\zeta(k) = 0$ si k est un entier pair $\leqslant -2$). Cette formule fait intervenir les groupes des classes d'idéaux des corps cyclotomiques $\mathbf{Q}(e^{2i\pi/p^n})$, pour $n \in \mathbf{N}$.

Les valeurs aux entiers positifs ont gardé leur mystère plus longtemps. On sait depuis Euler que l'on a $\zeta(2) = \pi^2/6$, $\zeta(4) = \pi^4/90$ et, plus généralement, que $\pi^{-2k}\zeta(2k)$ est un nombre rationnel si k est un entier $\geqslant 1$; ce nombre rationnel peut d'ailleurs s'exprimer en terme de $\zeta(1-2k)$ grâce à l'équation fonctionnelle. Ce n'est que très récemment que l'on a réussi, grâce aux travaux de Beilinson et de Bloch et Kato, à décoder l'information arithmétique contenue dans les valeurs aux entiers impairs positifs. La réponse fait intervenir des objets trop compliqués pour être donnée de manière précise, mais elle fournit un lien entre les valeurs aux entiers positifs de la fonction zêta et les valeurs aux nombres rationnels des fonctions polylogarithmes (le k-logarithme est défini par la série $\sum_{n=1}^{+\infty} z^n/n^k$, si $|z| < 1$) que nous avons explicité dans le texte.

Le nombre π étant transcendant sur $\mathbf{Q}$, les nombres $\zeta(2k)$, k entier $\geqslant 1$ sont tous irrationnels. Bien que tout le monde soit persuadé qu'il doive en être de même pour les $\zeta(2k+1)$, il a fallu attendre 1978 pour qu'Apéry démontre que $\zeta(3)$ est irrationnel. Entre 1978 et 2000, il n'y a eu aucun progrès sur la question, mais Rivoal a démontré en 2000 qu'il existe une infinité de $\zeta(2k+1)$, k entier $\geqslant 1$, qui sont irrationnels. Ce résultat est purement existentiel ; on ne peut en déduire l'irrationalité d'aucun $\zeta(2k+1)$ particulier.

La démonstration du théorème de Kummer auquel il a été fait allusion ci-dessus repose sur trois ingrédients : la formule analytique du nombre de classes qui donne une formule pour le nombre de classes d'idéaux du corps $\mathbf{Q}(e^{2i\pi/p})$ en termes du comportement en $s = 1$ (ou $s = 0$) de la fonction zêta de Dedekind de ce corps (illustration du principe selon lequel les fonctions L savent tout !), une factorisation de cette fonction zêta de Dedekind en produit de fonctions L de Dirichlet et une congruence modulo p entre la valeur en $s = 0$ de ces fonctions L de Dirichlet et les valeurs de la fonction ζ aux entiers négatifs. Ces congruences modulo p se généralisent en des congruences modulo p^n, pour tout $n \in \mathbf{N}$, entre les valeurs aux entiers négatifs de la fonction ζ. D'un point de vue moderne, ces congruences, découvertes elles-aussi par Kummer, sont équivalentes à l'existence d'une fonction continue sur l'anneau $\mathbf{Z}_p$ des entiers p-adiques et dont les

valeurs aux entiers négatifs sont liées à celles de la fonction zêta. Les zéros de cette fonction zêta p-adique (aussi appelée fonction zêta de Kubota-Leopoldt) sont bien compris, contrairement à ceux de la fonction zêta de Riemann... Le théorème de Mazur et Wiles mentionné ci-dessus en fournit une description en termes d'action du groupe de Galois sur les groupes de classes d'idéaux.

Les valeurs de la fonction zêta de Riemann apparaissent dans le terme constant du développement de Fourier de certaines formes modulaires appelées Séries d'Eisenstein. Ce lien représente un des outils les plus puissants que l'on ait pour étudier l'arithmétique de ces valeurs. Par exemple, Serre a donné une construction de la fonction zêta de Kubota-Leopoldt utilisant le fait que tous les termes, sauf le terme constant, du développement de Fourier des séries d'Eisenstein sont, de manière visible, des fonctions p-adiquement continues et donc qu'il en est de même du terme constant. Ce « donc » demande un lourd arsenal de géométrie algébrique pour être justifié : il faut partir d'objets définis analytiquement (formes modulaires) et aboutir dans un monde où on peut parler de congruence modulo p, ce qui représente un très long trajet. Cet arsenal est d'ailleurs fondamental dans la démonstration du résultat de Mazur et Wiles et dans la démonstration de Wiles du théorème de Fermat. Pour illustrer, de manière plus élémentaire, l'intérêt de l'utilisation des formes modulaires dans l'étude de la fonction zêta, nous avons inclus une démonstration (inédite ?) de l'équation fonctionnelle utilisant les séries d'Eisenstein.

Ce texte est divisé en gros en deux parties, la première traitant de la théorie complexe et la seconde de la théorie p-adique.

La première partie comporte trois chapitres. Dans le premier, on explore en détail les propriétés de la fonction ζ et les propriétés de rationalité de ses valeurs aux entiers ; tout ce qui est démontrable est démontré. Le lecteur intéressé par les polylogarithmes est invité à consulter l'exposé « Polylogarithmes » d'Oesterlé au séminaire Bourbaki. Ce chapitre comporte aussi un § sur la fonction zêta de Dedekind d'un corps de nombres sans aucune démonstration (pour en savoir plus, on pourra consulter le livre *Théorie des nombres* de Borevitch

et Chafarevitch). Dans le second, on donne une démonstration complète du théorème de Rivoal (pour des compléments sur les nombres polyzêtas, nous renvoyons au texte « Valeurs zêta multiples. Une introduction » de Waldschmidt). Le troisième est consacré aux formes modulaires et leur lien avec les fonctions L de l'arithmétique (le lecteur consultera avec profit les livres *Cours d'arithmétique* de Serre, *Introduction to modular forms* de Lang, *Introduction to elliptic curves and modular forms* de Koblitz ou encore *Invitation aux mathématiques de Fermat-Wiles* de Hellegouarch et *Automorphic forms and representations* de Bump, pour des compléments intéressants).

La seconde partie comporte aussi trois chapitres. Comme les nombres p-adiques ne font pas partie du bagage de tout mathématicien (ils ne sont pas enseignés en classe préparatoire, ce qui est un peu dommage : je suis sûr que les taupins apprécieraient de travailler dans un monde où une série converge si et seulement si son terme général tend vers 0), nous avons consacré un chapitre à leur construction (la construction du corps $\mathbf{C}_p$ est un peu plus longue que celle de $\mathbf{C}$, mais pas beaucoup) et un chapitre à l'analyse sur $\mathbf{Z}_p$. Le dernier chapitre est, quant à lui, consacré à la construction de la fonction zêta p-adique. Le second chapitre contient beaucoup plus de choses que ce qui est nécessaire pour cette construction (les mesures suffiraient, mais les distributions deviennent indispensables pour construire des fonction L p-adiques plus générales, par exemple celles attachées aux formes modulaires). Pour d'autres points de vue sur les nombres p-adiques, le lecteur pourra consulter les livres *p-adic numbers, p-adic analysis, and zeta-functions* et *p-adic analysis : a short course on recent work* de Koblitz ou *An introduction to G-functions* de Dwork, Gerotto et Sullivan. L'énoncé précis du théorème de Mazur et Wiles n'est pas donné dans le texte, pas plus que l'application au théorème de Kummer et nous renvoyons le lecteur intéressé aux livres *Introduction to cyclotomic fields* de Washington ou *Cyclotomic fields I and II* de Lang (pour le théorème de Kummer, on peut aussi lire la démonstration dans l'ouvrage de Borevich et Chafarevitch déjà cité).

Chapitre I. la fonction zêta de Riemann

I.1. Prolongement analytique et valeurs aux entiers négatifs

Soit $\zeta(s) = \sum_{n=1}^{+\infty} n^{-s} = \prod_p (1-p^{-s})^{-1}$ la fonction zêta de Riemann. Soit $\Gamma(s) = \displaystyle\int_{t=0}^{+\infty} e^{-t} t^s \frac{dt}{t}$ la fonction Γ d'Euler. Cette fonction est holomorphe pour $\mathrm{Re}(s) > 0$ et satisfait l'équation fonctionnelle $\Gamma(s+1) = s\Gamma(s)$, ce qui permet de la prolonger en une fonction méromorphe sur $\mathbf{C}$ tout entier.

Lemme I.1.1. *Si* $\mathrm{Re}(s) > 1$, *alors*

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^{+\infty} \frac{1}{e^t - 1} t^s \frac{dt}{t}.$$

Démonstration. Il suffit d'écrire $\dfrac{1}{e^t-1}$ sous la forme $\sum_{n=1}^{+\infty} e^{-nt}$ et d'utiliser la formule $\displaystyle\int_0^{+\infty} e^{-nt} t^s \frac{dt}{t} = \frac{\Gamma(s)}{n^s}$.

Proposition I.1.2. *Si f est une fonction $\mathscr{C}^\infty$ sur $\mathbf{R}_+$ à décroissance rapide à l'infini, alors la fonction*

$$\mathrm{L}(f,s) = \frac{1}{\Gamma(s)} \int_0^{+\infty} f(t) t^s \frac{dt}{t}$$

définie pour $\mathrm{Re}(s) > 0$ *admet un prolongement holomorphe à* $\mathbf{C}$ *tout entier et, si* $n \in \mathbf{N}$, *alors* $\mathrm{L}(f,-n) = (-1)^n f^{(n)}(0)$.

Démonstration. Soit φ une fonction $\mathscr{C}^\infty$ sur $\mathbf{R}_+$, valant 1 sur $[0,1]$ et 0 sur $[2,+\infty[$. On peut écrire f sous la forme $\varphi f + (1-\varphi)f$ et $\mathrm{L}(f,s)$ sous la forme $\mathrm{L}(\varphi f, s) + \mathrm{L}((1-\varphi)f, s)$ et comme $(1-\varphi)f$ est nulle dans un voisinage de 0 et à décroissance rapide à l'infini, l'intégrale

$$\int_0^{+\infty} f(t) t^s \frac{dt}{t}$$

définit une fonction holomorphe sur $\mathbf{C}$ tout entier. Comme de plus, $1/\Gamma(s)$ s'annule aux entiers négatifs, on a $\mathrm{L}((1-\varphi)f, -n) = 0$ si $n \in \mathbf{N}$. On voit donc que, quitte à remplacer f par $(1-\varphi)f$, on peut supposer f à support compact. Une intégration par partie nous

fournit alors la formule $\mathrm{L}(f,s) = -\mathrm{L}(f', s+1)$ si $\mathrm{Re}(s) > 1$, ce qui permet de prolonger $\mathrm{L}(f,s)$ en une fonction holomorphe sur $\mathbf{C}$ tout entier. D'autre part, on a

$$\mathrm{L}(f,-n) = (-1)^{n+1}\mathrm{L}(f^{(n+1)},1) = (-1)^{n+1}\int_0^{+\infty} f^{(n+1)}(t)dt = (-1)^n f^{(n)}(0),$$

ce qui termine la démonstration.

On peut en particulier appliquer cette proposition à $f_0(t) = \dfrac{t}{e^t-1}$. Soit $\sum_{n=0}^{+\infty} \mathrm{B}_n t^n/n!$ le développement de Taylor de f_0 en 0. Les B_n sont des nombres rationnels appelés nombres de Bernoulli et qu'on retrouve dans toutes les branches des mathématiques. On a en particulier

$$\mathrm{B}_0 = 1, \quad \mathrm{B}_1 = -\frac{1}{2}, \quad \mathrm{B}_2 = \frac{1}{6}, \quad \mathrm{B}_4 = \frac{-1}{30}, \quad \ldots, \quad \mathrm{B}_{12} = \frac{-691}{2730},$$

et comme $f_0(t) - f_0(-t) = -t$, la fonction f_0 est presque paire et $\mathrm{B}_{2k+1} = 0$ si $k \geqslant 1$. Un test presque infaillible pour savoir si une suite de nombres a un rapport avec les nombres de Bernoulli est de regarder si 691 apparaît dans les premiers termes de cette suite.

Théorème I.1.3

(i) *La fonction ζ a un prolongement méromorphe à $\mathbf{C}$ tout entier, holomorphe en dehors d'un pôle simple en $s = 1$ de résidu 1.*

(ii) *Si $n \in \mathbf{Q}$, alors $\zeta(-n) = (-1)^n \dfrac{\mathrm{B}_{n+1}}{n+1}$; en particulier, $\zeta(-n) \in \mathbf{Q}$.*

Démonstration. On a $\zeta(s) = \frac{1}{s-1}\mathrm{L}(f_0, s-1)$ comme on le constate en utilisant la formule $\Gamma(s) = (s-1)\Gamma(s-1)$; on en déduit le résultat.

I.2. Valeurs aux entiers positifs pairs

Proposition I.2.1. *Si $z \in \mathbf{C} - \mathbf{Z}$, alors*

$$\frac{1}{z} + \sum_{n=1}^{+\infty}\Big(\frac{1}{z+n} + \frac{1}{z-n}\Big) = \pi \mathrm{cotg}\, \pi z.$$

Démonstration. Notons $\mathrm{F}(z)$ la fonction

$$\mathrm{F}(z)=\frac{1}{z}+\sum_{n=1}^{+\infty}\Big(\frac{1}{z+n}+\frac{1}{z-n}\Big)=\frac{1}{z}+\sum_{n=1}^{+\infty}\frac{2z}{z^2-n^2}.$$

La convergence absolue de la série dans le second membre montre que $\mathrm{F}(z)$ est une fonction méromorphe sur $\mathbf{C}$, holomorphe sur $\mathbf{C}-\mathbf{Z}$ avec des pôles simples de résidu 1 en les entiers, que F est impaire et périodique de période 1. La fonction $\mathrm{G}(z)=\mathrm{F}(z)-\pi\mathrm{cotg}\,\pi z$ est donc holomorphe sur $\mathbf{C}$, impaire et périodique de période 1.

Si $-1/2\leqslant x\leqslant 1/2$, on a $|z^2-n^2|\geqslant y^2+n^2-1/4$ et $|z|\leqslant y+1/2$. On obtient donc

$$\begin{aligned}\Big|\sum_{n=1}^{+\infty}\frac{2z}{z^2-n^2}\Big|&\leqslant\sum_{n=1}^{+\infty}\Big|\frac{2z}{z^2-n^2}\Big|\leqslant\sum_{n=1}^{+\infty}\frac{2y+1}{y^2-\frac{1}{4}+n^2}\\&\leqslant\int_0^{+\infty}\frac{2y+1}{y^2-\frac{1}{4}+x^2}dx=\frac{\pi}{2}\cdot\frac{2y+1}{\sqrt{y^2-\frac{1}{4}}}.\end{aligned}$$

Comme la fonction $\pi\mathrm{cotg}\,\pi z$ est bornée sur $|\mathrm{Im}(z)|\geqslant 1$, il existe $c>0$ tel que $|\mathrm{G}(z)|\leqslant c$ si $z=x+iy$ et $-1/2\leqslant x\leqslant 1/2,\ y\geqslant 1$. De plus, G étant holomorphe, il existe $c'>0$ tel que $|\mathrm{G}(z)|\leqslant c'$ si $z=x+iy$ et $-1/2\leqslant x\leqslant 1/2,\ 0\leqslant y\leqslant 1$ et G étant impaire et périodique de période 1, on a alors $|\mathrm{G}(z)|\leqslant\sup(c,c')$ quel que soit $z\in\mathbf{C}$. La fonction G est donc bornée sur $\mathbf{C}$ tout entier, donc constante et nulle car impaire. Ceci permet de conclure.

Maintenant, on a d'une part

$$\frac{1}{z}+\sum_{n=1}^{+\infty}\frac{2z}{z^2-n^2}=\frac{1}{z}-\sum_{n=1}^{+\infty}2z\sum_{k=0}^{+\infty}\frac{z^{2k}}{n^{2k+2}}=\frac{1}{z}-2\sum_{k=1}^{+\infty}\zeta(2k)z^{2k-1},$$

et d'autre part,

$$\pi\mathrm{cotg}\,\pi z=i\pi\frac{e^{2i\pi z}+1}{e^{2i\pi z}-1}=i\pi+\frac{2i\pi}{e^{2i\pi z}-1}=i\pi+\frac{1}{z}\sum_{n=0}^{+\infty}\mathrm{B}_n\frac{(2i\pi z)^n}{n!}.$$

On en déduit le résultat suivant.

Théorème I.2.2. *Si k est un entier $\geqslant 1$, alors*

$$\zeta(2k)=-\frac{1}{2}\mathrm{B}_{2k}\frac{(2i\pi)^{2k}}{(2k)!}.$$

En particulier, $\pi^{-2k}\zeta(2k)$ est un nombre rationnel.

I.3. Polylogarithmes et valeurs aux entiers positifs de la fonction zêta

1. Polylogarithmes

Si k est un entier $\geqslant 1$, on note $\mathrm{Li}_k(z)$ la fonction définie, pour $|z| < 1$, par la formule

$$\mathrm{Li}_k(z) = \sum_{n=1}^{+\infty} \frac{z^n}{n^k}.$$

Ces fonctions Li_k, appelées *polylogarithmes* (Li_2 est le *dilogarithme*, Li_3 le *trilogarithme*...), ont été introduites par Leibniz, mais n'ont commencé à jouer un rôle important que depuis une vingtaine d'années ; on s'est aperçu récemment qu'elles apparaissaient naturellement dans de multiples questions (volumes de variétés hyperboliques, valeurs aux entiers des fonctions zêtas, cohomologie de $\mathbf{GL}_n(\mathbf{C})$...).

On a

$$\mathrm{Li}_1(z) = -\log(1-z), \quad \text{et} \quad \frac{d}{dz}\mathrm{Li}_k(z) = z^{-1}\mathrm{Li}_{k-1}(z),$$

ce qui permet de prolonger analytiquement $\mathrm{Li}_k(z)$, par récurrence sur k, en une fonction holomorphe multivaluée sur $\mathbf{C} - \{0,1\}$. (On est obligé de supprimer 0 bien que la série $\sum_{n=1}^{+\infty} z^n/n^k$ n'ait pas de singularité en 0 car, après un tour autour de 1, la fonction $\mathrm{Li}_1(z) = -\log(1-z)$ augmente de $2i\pi$ et donc vaut $2i\pi$ en 0, ce qui fait que $\mathrm{Li}_2(z) = \int z^{-1}\mathrm{Li}_1(z)$ a une singularité logarithmique en 0.) Pour prolonger analytiquement $\log z$ et les $\mathrm{Li}_k(z)$, $k \geqslant 1$, il faut choisir un chemin γ_z dans $\mathbf{C} - \{0,1\}$ reliant un point fixe (nous prendrons $1/2$ comme point de base de tous nos chemins) à z et poser

$$\log z = -\log 2 + \int_{\gamma_z} \frac{dt}{t},$$

$$\mathrm{Li}_1(z) = \log 2 + \int_{\gamma_z} \frac{dt}{1-t}$$

et

$$\mathrm{Li}_k(z) = \sum_{n=1}^{+\infty} \frac{2^{-n}}{n^k} + \int_{\gamma_z} \mathrm{Li}_{k-1}(t)\frac{dt}{t}.$$

Si on change le chemin γ_z, il existe $b, a_1, a_2, \ldots \in \mathbf{Q}$ (avec $a_k \in \frac{1}{(k-1)!}\mathbf{Z}$) tels que $\log z$, $\mathrm{Li}_1(z)$, $\mathrm{Li}_2(z)$ et $\mathrm{Li}_k(z)$ deviennent respectivement $\log z + b{\cdot}2i\pi$,

$$\mathrm{Li}_1(z) + a_1{\cdot}2i\pi, \quad \mathrm{Li}_2(z) + a_1{\cdot}2i\pi \log z + a_2 \cdot (2i\pi)^2$$

et $$\mathrm{Li}_k(z) + \sum_{j=1}^{k} a_j{\cdot}(2i\pi)^j \frac{\log^{k-j} z}{(k-j)!}.$$

2. La version analytique réelle des fonctions polylogarithmes

La fonction polylogarithme Li_k a un petit frère analytique réel P_k qui a le bon goût d'être monovalué ; cette fonction joue, pour Li_k, le rôle que joue $z \mapsto \log z\overline{z}$ pour la fonction logarithme.

Si $k \in \mathbf{N} - \{0\}$, soit $\mathrm{C}_k \in \mathbf{Q}[\mathrm{X}]$ l'unique polynôme vérifiant

$$\mathrm{C}_k(\mathrm{X}+1) - \mathrm{C}_k(\mathrm{X}) = \frac{\mathrm{X}^{k-1}}{(k-1)!} \quad \text{et} \quad \int_0^1 \mathrm{C}_k(t)dt = 0.$$

Le polynôme C_k est un multiple du k-ième polynôme de Bernoulli ; on a $c_k = \mathrm{C}_k(0) = \mathrm{B}_k/k!$, où B_k est le k-ième nombre de Bernoulli, et $\mathrm{C}_k(\mathrm{X}) = \sum_{i=0}^{k} c_i \mathrm{X}^{k-i}/(k-i)!$ (ceci peut se démontrer en remarquant que $\mathrm{C}_k' = \mathrm{C}_{k-1}$).

Proposition I.3.1. *Si k est un entier $\geqslant 1$, la fonction*

$$\mathrm{P}_k(z) = \sum_{\ell=0}^{k-1} c_\ell \cdot (\log z\overline{z})^\ell \cdot \Big(\mathrm{Li}_{k-\ell}(z) + (-1)^{k-1}\overline{\mathrm{Li}_{k-\ell}(z)}\Big)$$

est une fonction monovaluée sur $\mathbf{C} - \{0,1\}$, analytique réelle, à valeurs dans $\mathbf{R}$ (resp. $i\mathbf{R}$) si k est impair (resp. si k est pair).

Démonstration. La seule chose non évidente est le fait que $\mathrm{P}_k(z)$ ne dépend pas du choix de γ_z. Notons L la fonction $\log z$. D'après la discussion ci-dessus, si on change γ_z, la nouvelle valeur de $\mathrm{P}_k(z)$ diffère de l'ancienne par

$$\sum_{\ell=0}^{k-1} c_\ell{\cdot}(\mathrm{L}+\overline{\mathrm{L}})^\ell{\cdot}\Bigg(\sum_{j=1}^{k-\ell} a_j{\cdot}(2i\pi)^j{\cdot}\Big(\frac{\mathrm{L}^{k-\ell-j}}{(k-\ell-j)!} + (-1)^{k-j-1}\frac{\overline{\mathrm{L}}^{k-\ell-j}}{(k-\ell-j)!}\Big)\Bigg),$$

où $a_1, \dots, a_k$ sont des rationnels sans relation entre eux. Fixant j, et posant $m = k - j$, on est ramené à démontrer la relation

$$\sum_{\ell=0}^{m} \frac{c_\ell}{(m-\ell)!} \cdot (\mathrm{L} + \overline{\mathrm{L}})^\ell \cdot \big(\mathrm{L}^{m-\ell} + (-1)^{m-1}\overline{\mathrm{L}}^{m-\ell}\big) = 0.$$

Pour cela, divisons le tout par $\overline{\mathrm{L}}^m$ et posons $\mathrm{X} = \overline{\mathrm{L}}^{-1}\mathrm{L}$. On obtient

$$\sum_{\ell=0}^{m} c_\ell (\mathrm{X}+1)^\ell \frac{\mathrm{X}^{m-\ell} + (-1)^{m-1}}{(m-\ell)!}$$
$$= (\mathrm{X}+1)^m \Big(\mathrm{C}_m\Big(\frac{\mathrm{X}}{\mathrm{X}+1}\Big) + (-1)^{m-1}\mathrm{C}_m\Big(\frac{1}{\mathrm{X}+1}\Big)\Big) = 0$$

car $\mathrm{C}_m(1-\mathrm{Y}) = (-1)^m \mathrm{C}_m(\mathrm{Y})$ (cela découle immédiatement de la définition de C_m). Ceci permet de conclure.

Remarque I.3.2. Comme on peut faire des combinaisons linéaires des $\mathrm{P}_{k-\ell}(z)(\log z\overline{z})^\ell$ pour obtenir d'autres fonctions monovaluées à partir des $\mathrm{Li}_k(z)$, le choix des constantes c_ℓ peut sembler un peu arbitraire. Leur justification demanderait d'introduire la notion de *variation de structures de Hodge rationnelles*, ce qui nous entraînerait un peu loin.

Proposition I.3.3. *P_k se prolonge en une fonction continue sur $\mathbf{C}$ (resp. $\mathbf{C} - \{1\}$) si $k \geqslant 2$ (resp. si $k = 1$), avec $\mathrm{P}_k(0) = 0$ et $\mathrm{P}_k(1) = (1 + (-1)^{k-1})\zeta(k)$.*

Démonstration. Comme on a le choix de la détermination des $\mathrm{Li}_k(z)$, on peut, pour étudier le comportement de P_k en $z = 0$, utiliser la détermination donnée par la série $\sum_{n=1}^{+\infty} z^n/n^k$, auquel cas, le zéro de cette série en $z = 0$ contrebalance largement la croissance de $(\log|z|^2)^\ell$. En ce qui concerne le comportement en $z = 1$, remarquons que la singularité de Li_k en ce point est de la forme $-\log(1-z)\dfrac{z^{k-1}}{(k-1)!}$, et donc que la fonction Li_k admet un prolongement par continuité en $z = 1$ si $k \geqslant 2$; on en déduit le résultat car $(\log|z|^2)^{k-1}$ tend vers 0 assez vite quand $z \to 1$ pour tuer la singularité de Li_1.

3. Équations fonctionnelles des polylogarithmes

Si r est un entier $\geqslant 1$, la somme des puissances n-ièmes des racines r-ièmes de l'unité vaut r (resp. 0) si n est un multiple de r (resp. si n n'est pas un multiple de r). On en déduit la relation de distribution suivante pour la fonction P_k :

$$\sum_{\varepsilon^r=1} \mathrm{P}_k(\varepsilon z) = r^{1-k}\mathrm{P}_k(z^r).$$

Les fonctions $\mathrm{P}_1(z) = -\log|1-z|^2$, P_2 et P_3 vérifient les équations fonctionnelles suivantes

$$\mathrm{P}_1(z_1+z_2-z_1z_2) - \mathrm{P}_1(z_1) - \mathrm{P}_1(z_2) = 0$$

$$\mathrm{P}_2\Big(\frac{z_1z_2}{(1-z_1)(1-z_2)}\Big) - \mathrm{P}_2\Big(\frac{z_2}{1-z_1}\Big) - \mathrm{P}_2\Big(\frac{z_1}{1-z_2}\Big) + \mathrm{P}_2(z_1) + \mathrm{P}_2(z_2) = 0$$

$$\mathrm{P}_3\Big(\frac{z_1(1-z_2)^2}{z_2(1-z_1)^2}\Big) + \mathrm{P}_3(z_1z_2) + \mathrm{P}_3\Big(\frac{z_1}{z_2}\Big) - 2\mathrm{P}_3\Big(\frac{z_1(1-z_2)}{z_2(1-z_1)}\Big) - 2\mathrm{P}_3\Big(\frac{z_1(1-z_2)}{(z_1-1)}\Big) - 2\mathrm{P}_3\Big(\frac{z_2(1-z_1)}{(z_2-1)}\Big) - 2\mathrm{P}_3\Big(\frac{1-z_2}{1-z_1}\Big) - 2\mathrm{P}_3(z_1) - 2\mathrm{P}_3(z_2) + 2\mathrm{P}_3(1) = 0.$$

La première de ces équations fonctionnelles suit de l'équation fonctionnelle pour le logarithme ; les deux autres, découvertes au 19-ième siècle, se démontrent en dérivant. Comme c'est assez fastidieux, nous nous contenterons de vérifier l'équation fonctionnelle de P_2. Notons $h(z_1,z_2)$ la fonction dont on cherche à démontrer la nullité, et remarquons que $h(0,0) = 0$; il suffit donc de prouver que les dérivées partielles de h sont nulles, et comme h est symétrique par rapport à z_1 et z_2, il suffit de considérer les dérivées partielles par rapport à z_1 et $\overline{z}_1$. D'autre part, un calcul immédiat nous fournit les formules

$$\frac{\partial}{\partial z}\mathrm{P}_2(z) = -\frac{\log|z|}{1-z} - \frac{\log|1-z|}{z} \quad \text{et} \quad \frac{\partial}{\partial \overline{z}}\mathrm{P}_2(z) = -\overline{\frac{\partial}{\partial z}\mathrm{P}_2(z)},$$

ce qui fait que $\dfrac{\partial}{\partial z_1}h = 0$ entraîne $\dfrac{\partial}{\partial \overline{z}_1}h = 0$.

Si f est une fonction de z_1, z_2, soient $\varphi(f) = \dfrac{\partial}{\partial z_1}\log f$ et $\psi(f) = \log|f|$. On a

$$\varphi(fg) = \varphi(f) + \varphi(g),$$
$$\psi(fg) = \psi(f) + \psi(g)$$

et
$$\frac{\partial}{\partial \overline{z}_1}\mathrm{P}_2(f) = \varphi(1-f)\psi(f) - \varphi(f)\psi(1-f).$$

Posons $u_1 = z_1$, $u_2 = z_2$, $u_3 = 1 - z_1$, $u_4 = 1 - z_2$ et $u_5 = 1 - z_1 - z_2$, et donc

$$1 - \frac{z_1 z_2}{(1-z_1)(1-z_2)} = \frac{u_5}{u_3 u_4}, \quad 1 - \frac{z_2}{1-z_1} = \frac{u_5}{u_3}, \quad 1 - \frac{z_1}{1-z_2} = \frac{u_5}{u_4}.$$

Finalement, soient $v_i = \varphi(u_i)$ et $w_i = \psi(u_i)$, si $1 \leqslant i \leqslant 5$. On obtient

$$\begin{aligned}\frac{\partial}{\partial \overline{z}_1}h = {} & [(v_5 - v_3 - v_4)(w_1 + w_2 - w_3 - w_4) \\ & \qquad - (v_1 + v_2 - v_3 - v_4)(w_5 - w_3 - w_4)] \\ & - [(v_5 - v_3)(w_2 - w_3) - (v_2 - v_3)(w_5 - w_3)] \\ & - [(v_5 - v_4)(w_1 - w_4) - (v_1 - v_4)(w_5 - w_4)] \\ & + [v_3 w_1 - v_1 w_3] + [v_4 w_2 - v_2 w_4]\end{aligned}$$

Un calcul sans mystère permet de montrer que cette dernière quantité est nulle, ce qui permet de conclure.

Il est possible que tous les P_k vérifient des équations fonctionnelles du type ci-dessus, mais on ne sait pas comment en exhiber de manière systématique. L'exemple de P_2 montre qu'on a intérêt à minimiser le nombre de « facteurs premiers » de l'ensemble des f et $1 - f$ qui interviennent. La proposition I.3.4 ci-dessous « décrit » les équations fonctionnelles satisfaites par P_k.

Soit $\ell \geqslant 1$. Soit $\boldsymbol{e} = (e_1, \ldots, e_r)$, où les $e_i \in \mathbf{C}(z_1, \ldots, z_\ell)^*$ sont libres sur $\mathbf{Z}$ (*i.e.* $e_1^{i_1} \cdots e_r^{i_r} = 1$, avec $i_1, \ldots, i_r \in \mathbf{Z}$, implique $i_1 = \cdots = i_r = 0$). Soit $\mathrm{X}(\boldsymbol{e})$ le sous-groupe de $\mathbf{C}(z_1, \ldots, z_\ell)^*$ engendré par $e_1, \ldots, e_r$ et les racines de l'unité. Tout élément f de $\mathrm{X}(\boldsymbol{e})$ peut alors s'écrire de manière unique sous la forme $\varepsilon(f) e_1^{v_1(f)} \cdots e_r^{v_r(f)}$, où $\varepsilon(f)$ est une racine de l'unité et $v_1(f), \ldots, v_r(f) \in \mathbf{Z}$. Finalement, soit $\mathrm{A}(\boldsymbol{e})$ l'ensemble des $f \in \mathrm{X}(\boldsymbol{e})$ tels que $1 - f \in \mathrm{X}(\boldsymbol{e})$.

Proposition I.3.4. *Soit $k \geqslant 2$. Soient $f_1, \dots, f_n \in \mathrm{A}(\boldsymbol{e})$ et $a_1, \dots, a_n \in \mathbf{Z}$. Pour que la fonction $\sum_{i=1}^n a_i \mathrm{P}_k(f_i)$ soit constante sur $\mathbf{C}^\ell$ privé des pôles des f_i, il faut et il suffit que, quels que soient $j_1, \dots, j_k \in \{1, \dots, r\}$, on ait*

$$\sum_{i=1}^{n} a_i \cdot v_{j_1}(f_i) \cdots v_{j_{k-2}}(f_i) \cdot \big(v_{j_{k-1}}(f_i) v_{j_k}(1-f_i) - v_{j_{k-1}}(1-f_i) v_{j_k}(f_i)\big) = 0.$$

Démonstration. Cette proposition se démontre par récurrence sur k, en dérivant (ce n'est pas totalement évident).

Exemple I.3.5. Si on prend $\ell = 1$, $e_1 = z$ et $e_2 = 1 - z$, alors $\mathrm{A}(\boldsymbol{e})$ contient z, $1 - z$ et z^{-1}, et on obtient les équations fonctionnelles

$$\mathrm{P}_k(z^{-1}) + (-1)^k \mathrm{P}_k(z) = 0 \quad \text{et} \quad \mathrm{P}_k(z) + \mathrm{P}_k(1-z) = \mathrm{P}_k(1).$$

Remarque I.3.6. On connaît des équations fonctionnelles pour P_k en deux variables (*i.e.* $\ell = 2$) pour $k \leqslant 6$ et en une variable (autres que celles ci-dessus) pour $k \leqslant 7$. Pour aller plus loin, le problème est que le nombre de conditions à satisfaire est de l'ordre de k^r, où r est le cardinal de $\boldsymbol{e}$, et que pour être raisonnablement sûr d'avoir une solution pour $a_1, \dots, a_n$, il faut que le cardinal de $\mathrm{A}(\boldsymbol{e})$ soit plus grand que le nombre de conditions, ce qui n'est pas très facile à réaliser.

4. Valeurs aux entiers positifs de la fonction zêta

La formule $\mathrm{P}_k(1) = (1 + (-1)^{k-1})\zeta(k)$ fournit un lien (trivial) entre les fonctions polylogarithmes et les valeurs aux entiers de la fonction zêta ; le théorème ci-dessous en fournit un autre, beaucoup plus profond.

Si $x \in \mathbf{Q}^*$ et p est un nombre premier, on note $v_p(x)$ la valuation de x en p [si x est un entier, $v_p(x)$ est le plus grand entier n tel que p^n divise x ; dans le cas général, $v_p(b^{-1}a) = v_p(a) - v_p(b)$].

Théorème I.3.7. *Soit $k \geqslant 2$ un entier impair. Si $x_1, \dots, x_n \in \mathbf{Q} - \{0, 1\}$ et si $\lambda_1, \dots, \lambda_n \in \mathbf{Q}$ sont tels que, quels que soient les nombres*

premiers $p_1, \ldots, p_k$ (pas forcément distincts), on ait

$$\sum_{i=1}^{n} \lambda_i \cdot v_{p_1}(x_i) \cdots v_{p_{k-1}}(x_i) \cdot v_{p_k}(1 - x_i) = 0,$$

alors $\sum_{i=1}^{k} \mathrm{P}_k(x_i)$ est un multiple rationnel de $\zeta(k)$.

Démonstration. La démonstration actuelle de ce théorème est un véritable trou noir. La définition des objets qu'elle met en jeu (K-théorie, cohomologie motivique, motifs mixtes, régulateur de Borel...) est très loin de tenir dans la marge de ces pages, et mettre le doigt dedans expose l'individu qui s'y risque à disparaître corps et âme pour une très longue période. D'autre part, il ne semble pas que cette démonstration permette de borner le dénominateur du rationnel obtenu (et donc de le déterminer numériquement). Un des points de départ de la démonstration est le calcul du volume de $\mathbf{SL}_n(\mathbf{R})/\mathbf{SL}_n(\mathbf{Z})$.

5. Volume de $\mathbf{SL}_n(\mathbf{R})/\mathbf{SL}_n(\mathbf{Z})$

On note $g_n = (x_{i,j})_{1 \leqslant i,j \leqslant n}$ un élément général de $\mathbf{SL}_n(\mathbf{R})$ et on munit $\mathbf{SL}_n(\mathbf{R})$ de la mesure de Haar (à droite et à gauche)

$$dg_n = \prod_{(i,j) \neq (n,n)} dx_{i,j}.$$

On munit l'espace homogène $\mathscr{R}_n = \mathbf{SL}_n(\mathbf{R})/\mathbf{SL}_n(\mathbf{Z})$ de la mesure quotient et on a le résultat suivant :

Théorème I.3.8. *Le volume de $\mathscr{R}_n$ est fini et est donné par la formule*

$$\mathrm{Vol}(\mathscr{R}_n) = \prod_{k=2}^{n} \zeta(k).$$

Démonstration. La démonstration se fait par récurrence sur n, le cas $n = 1$ étant évident (un produit vide vaut 1 par convention). Pour passer de $n-1$ à n, on dévisse $\mathbf{SL}_n(\mathbf{R})$ en introduisant le sous-groupe

$\mathrm{H}_n(\mathbf{R}) \subset \mathbf{SL}_n(\mathbf{R})$ des matrices de la forme

$$h_n = \begin{pmatrix} 1 & v_{n-1} \\ 0 & g_{n-1} \end{pmatrix},$$
$$g_{n-1} = (x_{i,j})_{2 \leqslant i,j \leqslant n} \in \mathbf{SL}_{n-1}(\mathbf{R}),$$
$$v_{n-1} = (x_{1,2}, \dots, x_{1,n}) \in \mathbf{R}^{n-1},$$

et l'application $\pi_n : \mathbf{SL}_n(\mathbf{R}) \to \mathbf{R}^n - \{0\}$ envoyant g_n sur le vecteur $w_n = (x_{1,1}, \dots, x_{n,1})$ formé par les coefficients de la première colonne de g_n. Cette application induit une bijection de l'espace homogène $\mathbf{SL}_n(\mathbf{R})/\mathrm{H}_n(\mathbf{R})$ sur $\mathbf{R}^n - \{0\}$. Si on munit $\mathbf{R}^{n-1}$ de la forme volume $dv_{n-1} = dx_{1,2} \cdots dx_{1,n}$, $\mathrm{H}_n(\mathbf{R})$ de la forme volume $dh_n = dv_{n-1} dg_{n-1}$, $\mathbf{R}^n$ de la forme volume $dw_n = dx_{1,1} \cdots dx_{n,1}$, alors la forme dh_n est invariante par translation (à gauche ou à droite) par un élément de $\mathrm{H}_n(\mathbf{R})$ et $dg_n = dw_n dh_n$.

Notons c_{n-1} le volume de $\mathscr{R}_{n-1}$. Comme $\mathrm{H}_n(\mathbf{R})$ (resp. $\mathrm{H}_n(\mathbf{Z})$) est le produit semi-direct de $\mathbf{SL}_{n-1}(\mathbf{R})$ (resp. $\mathbf{SL}_{n-1}(\mathbf{Z})$) et de $\mathbf{R}^{n-1}$ (resp. $\mathbf{Z}^{n-1}$), on a aussi

$$\begin{aligned} \mathrm{Vol}(\mathrm{H}_n(\mathbf{R})/\mathrm{H}_n(\mathbf{Z})) &= \mathrm{Vol}(\mathbf{SL}_{n-1}(\mathbf{R})/\mathbf{SL}_{n-1}(\mathbf{Z})) \cdot \mathrm{Vol}(\mathbf{R}^{n-1}/\mathbf{Z}^{n-1}) \\ &= c_{n-1}. \end{aligned}$$

Pour calculer le volume de $\mathscr{R}_n$, considérons une fonction φ positive sur $\mathbf{R}^n$. On a, d'une part

$$\int_{\mathbf{R}^n - \{0\}} \varphi(w_n) dw_n = \int_{\mathbf{R}^n} \varphi(w_n) dw_n,$$

car $\{0\}$ est de mesure nulle et, d'autre part, le théorème de Fubini nous donne

$$\begin{aligned} \mathrm{Vol}(\mathrm{H}_n(\mathbf{R})/\mathrm{H}_n(\mathbf{Z})) \int_{\mathbf{R}^n - \{0\}} \varphi(w_n) dw_n &= \int_{\mathbf{SL}_n(\mathbf{R})/\mathrm{H}_n(\mathbf{Z})} \varphi(\pi_n(g_n)) dg_n \\ &= \int_{\mathscr{R}_n} \Big(\sum_{\gamma \in \mathbf{SL}_n(\mathbf{Z})/\mathrm{H}_n(\mathbf{Z})} \varphi(\pi_n(g_n\gamma)) \Big) dg. \end{aligned}$$

L'application qui, à $g_n \in \mathbf{SL}_n(\mathbf{R})$ associe le sous-$\mathbf{Z}$-module Λ_{g_n} de $\mathbf{R}^n$, engendré par les colonnes de g_n, induit une bijection de l'espace homogène $\mathscr{R}_n = \mathbf{SL}_n(\mathbf{R})/\mathbf{SL}_n(\mathbf{Z})$ sur l'ensemble des réseaux de volume 1 de $\mathbf{R}^n$. Si $\gamma \in \mathbf{SL}_n(\mathbf{Z})$, alors $\pi_n(g_n\gamma)$ est une combinaison

linéaire, à coefficients entiers *premiers entre eux* des colonnes de g_n ; c'est donc un élément *primitif* du réseau Λ_{g_n} (un élément λ d'un réseau Λ est dit primitif si on ne peut pas l'écrire sous la forme $a \cdot \lambda'$, avec $a \in \mathbf{N} - \{0, 1\}$ et $\lambda' \in \Lambda$). L'application $\gamma \mapsto \pi_n(g_n\gamma)$ induit une bijection de $\mathbf{SL}_n(\mathbf{R})/\mathrm{H}_n(\mathbf{Z})$ sur l'ensemble Λ'_{g_n} des éléments primitifs de Λ_{g_n}. On obtient donc

$$c_{n-1}\int_{\mathbf{R}^n}\varphi(w_n)dw_n = \int_{\mathscr{R}_n}\Big(\sum_{\lambda\in\Lambda'_{g_n}}\varphi(\lambda)\Big)dg_n.$$

Pour passer de Λ'_{g_n} à Λ_{g_n}, on utilise la formule

$$\int_{\mathbf{R}^n}\varphi(aw_n)dw_n = a^{-n}\int_{\mathbf{R}^n}\varphi(w_n)dw_n,$$

ce qui nous donne, en sommant sur $a \in \mathbf{N} - \{0\}$,

$$\begin{aligned}\zeta(n)c_{n-1}\int_{\mathbf{R}^n}\varphi(w_n)dw_n &= c_{n-1}\sum_{a=1}^{+\infty}\int_{\mathbf{R}^n}\varphi(aw_n)dw_n\\ &= \sum_{a=1}^{+\infty}\int_{\mathscr{R}_n}\Big(\sum_{\lambda\in\Lambda'_{g_n}}\varphi(a\lambda)\Big)dg_n = \int_{\mathscr{R}_n}\Big(\sum_{\lambda\in\Lambda_{g_n}-\{0\}}\varphi(\lambda)\Big)dg_n.\end{aligned}$$

(Modulo la formule $\mathrm{Vol}(\mathscr{R}_n) = \zeta(n)c_{n-1}$ que l'on cherche à obtenir, cette égalité peut s'exprimer, après division par $\mathrm{Vol}(\mathscr{R}_n)$, sous la forme « l'intégrale sur $\mathbf{R}^n$ d'une fonction φ est la moyenne sur l'ensemble des réseaux de volume 1 de $\mathbf{R}^n$ de la somme des valeurs de φ en les points du réseau ».)

D'après le lemme de Minkowski (un des résultats les plus utiles en théorie des nombres), un convexe de $\mathbf{R}^n$, de volume $\geqslant 2^n$, symétrique par rapport à l'origine, contient un élément non nul de tout réseau de $\mathbf{R}^n$ de volume 1. Si on prend pour φ la fonction caractéristique d'un tel convexe, alors $\sum_{\lambda\in\Lambda_{g_n}-\{0\}}\varphi(\lambda) \geqslant 1$ quel que soit $g_n \in \mathbf{SL}_n(\mathbf{Z})$; on en déduit l'inégalité $c_n = \int_{\mathscr{R}_n} dg_n \leqslant 2^n\zeta(n)c_{n-1}$; en particulier, c_n est fini.

Maintenant, si φ est une fonction de Schwartz sur $\mathbf{R}^n$ (*i.e.* φ est $\mathscr{C}^\infty$ et à décroissance rapide ainsi que toute ses dérivées), et si $\widehat{\varphi}$

désigne la transformée de Fourier

$$\widehat{\varphi}(y) = \int_{\mathbf{R}^n} e^{-2i\pi\, {}^t x\cdot w_n} dw_n$$

de φ, on dispose de la formule de Poisson (où l'on pose $g'_n = {}^t g_n^{-1}$)

$$\sum_{\lambda\in\Lambda_{g_n}} \varphi(\lambda) = \sum_{\lambda\in\Lambda_{g'_n}} \widehat{\varphi}(\lambda).$$

Ceci nous donne

$$\begin{aligned}\zeta(n)c_{n-1}\widehat{\varphi}(0) + c_n\varphi(0) &= \int_{\mathscr{R}_n} \Big(\sum_{\lambda\in\Lambda_{g_n}} \varphi(\lambda)\Big) dg_n \\ &= \int_{\mathscr{R}_n} \Big(\sum_{\lambda\in\Lambda_{g'_n}} \widehat{\varphi}(\lambda)\Big) dg_n = \zeta(n)c_{n-1}\widehat{\widehat{\varphi}}(0) + c_n\widehat{\varphi}(0).\end{aligned}$$

Cette formule étant valable quel que soit φ, la formule d'inversion de Fourier $\widehat{\widehat{\varphi}}(0) = \varphi(0)$ nous permet de d'obtenir l'égalité

$$c_n = \zeta(n)c_{n-1}$$

qui permet de conclure.

Remarque I.3.9. En voyant la formule du théorème, on se dit qu'il doit exister une « décomposition » de $\mathbf{SL}_n$ en morceaux reflétant la factorisation naturelle du volume de $\mathbf{SL}_n(\mathbf{R})/\mathbf{SL}_n(\mathbf{Z})$. C'est cette vague intuition qui mène à la K-théorie et aux régulateurs de Borel (*cf.* démonstration du théorème I.3.7).

I.4. Équation fonctionnelle de la fonction zêta

Soit $\xi(s) = \pi^{-s/2}\Gamma(s/2)\zeta(s)$.

Théorème I.4.1. *La fonction $\xi(s)$ admet un prolongement méromorphe à $\mathbf{C}$ tout entier, holomorphe en dehors de pôles simples de résidus respectifs -1 et 1 en $s = 0$ et $s = 1$, et vérifie l'équation fonctionnelle*

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

Démonstration. Il y a un nombre assez conséquent de méthodes pour arriver au résultat. Nous en donnerons deux dans ce n° et une autre au § III.4 (*cf.* th. III.4.5).

1. Première méthode : la fonction thêta

Si $t > 0$, on a

$$\int_{-\infty}^{+\infty} e^{-\pi t x^2} e^{-2i\pi xy} dx = e^{-\pi t^{-1} y^2} \int_{-\infty}^{+\infty} e^{-\pi t(x+it^{-1}y)^2} dx$$
$$= t^{-1/2} e^{-\pi t^{-1} y^2}.$$

(C'est classique ; faire le changement de variables $u = \sqrt{t}(x + it^{-1}y)$, utiliser le théorème des résidus pour revenir sur la droite réelle et la formule $\int_{-\infty}^{+\infty} e^{-\pi u^2} du = 1$ pour conclure.)

Si $t > 0$, soit $\theta(t) = \sum_{n \in \mathbf{Z}} e^{-\pi t n^2}$. La formule de Poisson

$$\sum_{n \in \mathbf{Z}} \varphi(n) = \sum_{n \in \mathbf{Z}} \int_{-\infty}^{+\infty} \varphi(x) e^{-2i\pi nx} dx,$$

utilisée pour $\varphi(x) = e^{-\pi t x^2}$, nous fournit l'équation fonctionnelle

$$\theta(t) = t^{-1/2} \theta(t^{-1}).$$

On a alors

$$\xi(s) = \frac{1}{2} \int_0^{+\infty} (\theta(t) - 1) t^{s/2} \frac{dt}{t}$$
$$= \frac{1}{2} \int_0^1 (t^{-1/2} \theta(t^{-1}) - 1) t^{s/2} \frac{dt}{t} + \frac{1}{2} \int_1^{+\infty} (\theta(t) - 1) t^{s/2} \frac{dt}{t}$$

et on peut changer t en t^{-1} dans la première intégrale pour obtenir

$$\xi(s) = -\frac{1}{s} + \frac{1}{s-1} + \frac{1}{2} \int_1^{+\infty} (\theta(t) - 1)(t^{s/2} + t^{(1-s)/2}) \frac{dt}{t}.$$

On en déduit le résultat car $\theta(t) - 1$ est à décroissance rapide à l'infini, ce qui implique que l'intégrale est une fonction holomorphe de s sur $\mathbf{C}$ tout entier, et le membre de droite est évidemment invariant par $s \mapsto 1 - s$.

2. Deuxième méthode : intégrale sur un contour

Si $c > 0$, soit C_c le contour obtenu en suivant la droite réelle de $+\infty$ à $c\pi$, puis en parcourant le carré de sommets $c\pi(\pm 1 \pm i)$ dans le

sens trigonométrique, et en retournant en $+\infty$ le long de l'axe réel. Soit

$$\mathrm{F}_c(s) = \int_{\mathrm{C}_c} \frac{1}{e^z-1}(-z)^s \frac{dz}{z},$$

où $(-z)^s = \exp(s\log(-z))$ et la branche du logarithme choisie est celle dont la partie imaginaire est comprise entre $-\pi$ et π ; en particulier, on a $(-z)^s = e^{-i\pi s}z^s$ de $+\infty$ à $c\pi$ et $(-z)^s = e^{i\pi s}z^s$ de $c\pi$ à $+\infty$ (après avoir parcouru le carré). Comme $\frac{1}{e^z-1}$ est à décroissance rapide à l'infini, la fonction $\mathrm{F}_c(s)$ est holomorphe sur $\mathbf{C}$ pour tout c qui n'est pas un entier pair (pour éviter les pôles de $\frac{1}{e^z-1}$). D'autre part, le théorème des résidus montre que $\mathrm{F}_c(s)$ ne dépend pas de c si c reste dans un intervalle du type $]2\mathrm{N}, 2\mathrm{N}+2[$, avec $\mathrm{N} \in \mathbf{N}$. En particulier, on a $\mathrm{F}_1(s) = \mathrm{F}_c(s)$ quel que soit $c \in]0,2[$. Si $\mathrm{Re}(s) > 1$, quand c tend vers 0, l'intégrale sur le carré de sommets $c\pi(\pm 1 \pm i)$ tend vers 0, et on obtient, en passant à la limite

$$\begin{aligned}\mathrm{F}_1(s) &= e^{-i\pi s}\int_{+\infty}^{0} \frac{1}{e^z-1} z^s \frac{dz}{z} + e^{i\pi s}\int_0^{+\infty} \frac{1}{e^z-1} z^s \frac{dz}{z} \\ &= 2i \cdot \sin \pi s \cdot \Gamma(s) \cdot \zeta(s).\end{aligned}$$

Maintenant, quand N tend vers $+\infty$, la fonction $\mathrm{F}_{2\mathrm{N}+1}(s)$ tend vers 0 quand $\mathrm{Re}(s) < 0$ car $\frac{1}{e^z-1}$ est majorée, indépendamment de N, sur $\mathrm{C}_{2\mathrm{N}+1}$. La différence entre $\mathrm{F}_1(s)$ et $\mathrm{F}_{2\mathrm{N}+1}(s)$ peut se calculer grâce au théorème des résidus. La fonction $\frac{1}{e^z-1}(-z)^{s-1}$ a des pôles en $z = \pm 2i\pi, \pm 4i\pi, \ldots, \pm 2\mathrm{N}i\pi$ dans le contour délimité par la différence entre $\mathrm{C}_{2\mathrm{N}+1}$ et C_1. Si $k \in \{1, \ldots, \mathrm{N}\}$, le résidu de $\frac{1}{e^z-1}(-z)^{s-1}$ est $(2k\pi)^{s-1}e^{-i\pi(s-1)/2}$ en $2ik\pi$ et $(2k\pi)^{s-1}e^{i\pi(s-1)/2}$ en $-2ik\pi$, ce qui nous donne

$$\frac{1}{2i\pi}(\mathrm{F}_{2\mathrm{N}+1}(s) - \mathrm{F}_1(s)) = 2\cos\pi\frac{s-1}{2} \cdot \sum_{k=1}^{\mathrm{N}} (2k\pi)^{s-1}.$$

En passant à la limite, on obtient donc $\mathrm{F}_1(s) = 4i\pi \cdot \cos(\pi\frac{s-1}{2}) \cdot (2\pi)^{s-1} \cdot \zeta(1-s)$, si $\mathrm{Re}(s) < 0$. On en déduit l'équation fonctionnelle

$$\sin \pi s \cdot \Gamma(s) \cdot \zeta(s) = \cos\pi\frac{s-1}{2} \cdot (2\pi)^s \cdot \zeta(1-s).$$

Pour passer de cette équation fonctionnelle à celle de ξ, il faut utiliser les formules classiques

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s},\ \Gamma\big(\frac{s}{2}\big)\Gamma\big(\frac{s+1}{2}\big) = 2^{1-s}\Gamma\big(\frac{1}{2}\big)\Gamma(s),\ \Gamma\big(\frac{1}{2}\big) = \sqrt{\pi}.$$

I.5. Fonctions L de Dirichlet

1. Caractères de Dirichlet et sommes de Gauss

Si d est un entier, on appelle caractère de Dirichlet modulo d un morphisme de groupes de $(\mathbf{Z}/d\mathbf{Z})^*$ dans $\mathbf{C}^*$. L'image d'un caractère de Dirichlet est bien évidemment incluse dans le groupe des racines de l'unité.

Si d' est un diviseur de d et χ est un caractère de Dirichlet modulo d', on peut aussi voir χ comme un caractère de Dirichlet modulo d en composant χ avec la projection $(\mathbf{Z}/d\mathbf{Z})^* \to (\mathbf{Z}/d'\mathbf{Z})^*$. On dit que χ est de conducteur d si on ne peut pas trouver de diviseur d' de d distinct de d, tel que χ provienne d'un caractère modulo d'. De manière équivalente, χ est de conducteur d si quel que soit d' diviseur de d distinct de d, la restriction de χ au noyau de la projection $(\mathbf{Z}/d\mathbf{Z})^* \to (\mathbf{Z}/d'\mathbf{Z})^*$ n'est pas triviale.

Si χ est un caractère de Dirichlet modulo d, on note χ^{-1} le caractère de Dirichlet modulo d défini par $\chi^{-1}(n) = (\chi(n))^{-1}$ si $n \in (\mathbf{Z}/d\mathbf{Z})^*$.

Si χ est un caractère de Dirichlet modulo d, on considère aussi souvent χ comme une fonction périodique sur $\mathbf{Z}$ de période d en composant χ avec la projection naturelle de $\mathbf{Z}$ sur $\mathbf{Z}/d\mathbf{Z}$ et en étendant χ par 0 sur les entiers non premiers à d. On a donc $\chi^{-1}(n) = (\chi(n))^{-1}$ si $(n,d) = 1$, mais $\chi^{-1}(n) = 0$ si $(n,d) \neq 1$.

Si d est un entier, χ un caractère de Dirichlet de conducteur d et si $n \in \mathbf{Z}$, on définit la somme de Gauss tordue $\mathrm{G}(\chi, n)$ par la formule

$$\mathrm{G}(\chi, n) = \sum_{a \bmod d} \chi(a) e^{2i\pi na/d}$$

et on pose $\mathrm{G}(\chi) = \mathrm{G}(\chi, 1)$.

Lemme I.5.1

(i) *Si* $n \in \mathbf{N}$, *alors* $\mathrm{G}(\chi, n) = \chi^{-1}(n)\mathrm{G}(\chi)$

(ii) $\mathrm{G}(\chi)\mathrm{G}(\chi^{-1}) = \chi(-1)d$.

Démonstration. Si $(n,d)=1$, alors n est inversible dans $(\mathbf{Z}/d\mathbf{Z})^*$, ce qui permet d'écrire

$$\begin{aligned}\mathrm{G}(\chi,n)&=\sum_{a \bmod d}\chi(a)e^{2i\pi na/d}=\chi^{-1}(n)\sum_{an \bmod d}\chi(an)e^{2i\pi na/d}\\&=\chi^{-1}(n)\mathrm{G}(\chi).\end{aligned}$$

Si $(n,d)=e>1$, on peut écrire $d=ed'$ et $n=en'$. Soit U le noyau de la projection de $(\mathbf{Z}/d\mathbf{Z})^*$ sur $(\mathbf{Z}/d'\mathbf{Z})^*$. Si on choisit un système S de représentants de $(\mathbf{Z}/d'\mathbf{Z})^*$ dans $(\mathbf{Z}/d\mathbf{Z})^*$, on a

$$\mathrm{G}(\chi,n)=\sum_{a\in\mathrm{S}}\sum_{u\in\mathrm{U}}\chi(au)e^{2i\pi nau/d}.$$

Si $u\in\mathrm{U}$ et $a\in\mathbf{Z}$, alors $nau-na=n'ea(u-1)\equiv 0 \bmod d$ et donc $e^{2i\pi nau/d}=e^{2i\pi na/d}$. On obtient donc

$$\mathrm{G}(\chi,n)=\sum_{a\in\mathrm{S}}\chi(a)e^{2i\pi na/d}\Big(\sum_{u\in\mathrm{U}}\chi(u)\Big)=0$$

car $\sum_{u\in\mathrm{U}}\chi(u)=0$ puisque χ est un caractère non trivial de U (sinon χ serait de conducteur d'). Ceci démontre le (i).

Utilisant ce qui précède, on obtient

$$\begin{aligned}\mathrm{G}(\chi)\mathrm{G}(\chi^{-1})&=\sum_{b \bmod d}e^{2i\pi b/d}\chi^{-1}(b)\mathrm{G}(\chi)\\&=\sum_{b \bmod d}e^{2i\pi b/d}\Big(\sum_{a \bmod d}\chi(a)e^{2i\pi ab/d}\Big)\\&=\sum_{a \bmod d}\chi(a)\Big(\sum_{b \bmod d}e^{2i\pi(a+1)b/d}\Big)\end{aligned}$$

et comme $\sum\limits_{b \bmod d}e^{2i\pi(a+1)b/d}=\begin{cases}d & \text{si } a=-1\\ 0 & \text{sinon}\end{cases}$, on en tire la formule $\mathrm{G}(\chi)\mathrm{G}(\chi^{-1})=\chi(-1)d$.

2. Les fonctions L de Dirichlet

Soient d un entier et χ un caractère de Dirichlet de conducteur d. Soit

$$\mathrm{L}(\chi,s)=\sum_{n=1}^{+\infty}\frac{\chi(n)}{n^s}=\prod_{p \text{ premier}}(1-\chi(p)p^{-s})^{-1}$$

la fonction L de Dirichlet attachée à χ. Si on utilise l'identité

$$\chi(n) = \frac{1}{\mathrm{G}(\chi^{-1})} \sum_{b \bmod d} \chi^{-1}(b) e^{2i\pi nb/d},$$

on obtient

$$\mathrm{L}(\chi, s) = \frac{1}{\mathrm{G}(\chi^{-1})} \sum_{b \bmod d} \chi^{-1}(b) \sum_{n=1}^{+\infty} \frac{e^{2i\pi nb/d}}{n^s}$$

Utilisant la formule $\displaystyle\int_0^{+\infty} e^{-nt} t^s \frac{dt}{t} = \frac{\Gamma(s)}{n^s}$, on obtient, ayant posé $\varepsilon_d = e^{2i\pi/d}$,

$$\begin{aligned}\mathrm{L}(\chi, s) &= \frac{1}{\mathrm{G}(\chi^{-1})} \frac{1}{\Gamma(s)} \sum_{b \bmod d} \chi^{-1}(b) \int_0^{+\infty} \sum_{n=1}^{+\infty} \varepsilon_d^{nb} e^{-nt} \\ &= \frac{1}{\mathrm{G}(\chi^{-1})} \frac{1}{\Gamma(s)} \int_0^{+\infty} \sum_{b \bmod d} \frac{\chi^{-1}(b)}{\varepsilon_d^{-b} e^t - 1} t^s \frac{dt}{t}.\end{aligned}$$

En particulier, la proposition I.1.2 implique que $\mathrm{L}(\chi, s)$ admet un prolongement holomorphe à $\mathbf{C}$ tout entier et que, si $n \in \mathbf{N}$, alors $\mathrm{L}(\chi, -n)$ est $(-1)^n\times$ la dérivée n-ième de la fonction

$$\frac{1}{\mathrm{G}(\chi^{-1})} \sum_{b \bmod d} \frac{\chi^{-1}(b)}{\varepsilon_d^{-b} e^t - 1}$$

prise en $t = 0$. Pour supprimer le $(-1)^n$, on peut changer t en $-t$, utiliser l'identité

$$\frac{1}{\varepsilon_d^{-b} e^{-t} - 1} = -1 - \frac{1}{\varepsilon_d^{b} e^{t} - 1}$$

et le fait que $\displaystyle\sum_{b \bmod d} \chi^{-1}(b) = 0$ et on obtient $\mathrm{L}(\chi, -n) = (\frac{d}{dt})^n \mathscr{L}_\chi(t)_{|0}$, où l'on a posé

$$\mathscr{L}_\chi(t) = \frac{-1}{\mathrm{G}(\chi^{-1})} \sum_{b \bmod d} \frac{\chi^{-1}(b)}{\varepsilon_d^{b} e^{t} - 1}.$$

Remarque I.5.2. On peut démontrer, par exemple à partir de la formule ci-dessus pour $\mathrm{L}(\chi, s)$, que $\mathrm{L}(\chi, s)$ vérifie l'équation fonctionnelle

$\Lambda(\chi^{-1}, 1-s) = (-i)^{a(\chi)} d^{-1/2} \mathrm{G}(\chi^{-1}) \Lambda(\chi, s)$, où l'on a posé

$$\Lambda(\chi, s) = \frac{\Gamma(\frac{s+a(\chi)}{2})}{\pi^{(s+a(\chi))/2}} \cdot d^{s/2} \cdot \mathrm{L}(\chi, s),$$

avec $a(\chi) = 0$ si $\chi(-1) = 1$ et $a(\chi) = 1$ si $\chi(-1) = -1$.

I.6. La fonction zêta de Dedekind d'un corps de nombres

Un *corps de nombres* est une extension finie de $\mathbf{Q}$. Par exemple, les *corps quadratiques* (de la forme $\mathbf{Q}(\sqrt{\mathrm{D}})$, avec $\mathrm{D} \in \mathbf{Q}^*$) ou les *corps cyclotomiques* (de la forme $\mathbf{Q}(e^{2i\pi/m})$, m entier $\geqslant 3$) sont des corps de nombres.

D'après le théorème de l'élément primitif, si F est un corps de nombres de degré n sur $\mathbf{Q}$, alors il existe $\alpha \in \mathrm{F}$ engendrant F en tant que $\mathbf{Q}$-algèbre. Si $\mathrm{P} \in \mathbf{Q}[\mathrm{X}]$ est le polynôme minimal de α, alors P est de degré n et l'application $\mathrm{A} \mapsto \mathrm{A}(\alpha)$ induit un isomorphisme de $\mathbf{Q}[\mathrm{X}]/\mathrm{P}$ sur F. On note r_1 le nombre de racines réelles de P et $2r_2$ le nombre de racines non réelles de P. On a donc $r_1 + 2r_2 = n$ et on dit que F est *totalement réel* si $r_2 = 0$. Si $\alpha_1, \ldots, \alpha_{r_1}, \alpha_{r_1+1}, \ldots, \alpha_{r_1+r_2}, \overline{\alpha_{r_1+1}}, \ldots, \overline{\alpha_{r_1+r_2}}$ sont les racines de P, on note σ_i (resp. $\overline{\sigma_i}$), si $1 \leqslant i \leqslant r_1 + r_2$ (resp. $r_1 + 1 \leqslant i \leqslant r_1 + r_2$) le plongement de F dans $\mathbf{C}$ envoyant α sur α_i (resp. $\overline{\alpha_i}$).

On dit que $x \in \mathrm{F}$ est *entier* si son polynôme minimal est à coefficients dans $\mathbf{Z}$. L'ensemble des entiers de F forme un anneau $\mathscr{O}_\mathrm{F}$ appelé *anneau des entiers* de F. Par exemple, si $\mathrm{D} \in \mathbf{Z}$ n'est divisible par le carré d'aucun nombre premier, alors l'anneau des entiers de $\mathbf{Q}(\sqrt{\mathrm{D}})$ est $\mathbf{Z}[\sqrt{\mathrm{D}}]$ (resp. $\mathbf{Z}[\frac{1+\sqrt{\mathrm{D}}}{2}]$) si D $\equiv 2$ ou 3 modulo 4 (resp. si D $\equiv 1$ modulo 4). L'anneau des entiers de $\mathbf{Q}(e^{2i\pi/m})$ est $\mathbf{Z}[e^{2i\pi/m}]$.

L'anneau $\mathscr{O}_\mathrm{F}$ est un $\mathbf{Z}$-module libre de rang n (c'est un réseau du $\mathbf{Q}$-espace vectoriel F qui est de dimension n). Si $e_1, \ldots, e_n$ forment une base de $\mathscr{O}_\mathrm{F}$ sur $\mathbf{Z}$, on note Δ_F le *discriminant* de F ; c'est la valeur absolue du déterminant de la matrice $n \times n$ dont les coefficients sont les $(\mathrm{Tr}_{\mathrm{F}/\mathbf{Q}}(e_i e_j))_{1 \leqslant i,j \leqslant n}$. Le discriminant de $\mathbf{Q}(\sqrt{\mathrm{D}})$ est $4|\mathrm{D}|$ (resp. $|\mathrm{D}|$) D $\equiv 2$ ou 3 modulo 4 (resp. si D $\equiv 1$ modulo 4). L'application $\mathrm{A} \mapsto (\mathrm{A}(\alpha_1), \ldots, \mathrm{A}(\alpha_{r_1+r_2}))$ induit un isomorphisme de $\mathbf{R}[\mathrm{X}]/\mathrm{P}$ sur $\mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$; l'image de $\mathscr{O}_\mathrm{F} \subset \mathrm{F} \cong \mathbf{Q}[\mathrm{X}]/\mathrm{P}$ par cet isomorphisme est un réseau du $\mathbf{R}$-espace vectoriel $\mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$ et le volume de $\mathscr{O}_\mathrm{F}$ pour la

forme volume naturelle (dx sur $\mathbf{R}$ et $|dz \wedge d\bar{z}| = 2\,dx\,dy$ sur $\mathbf{C}$) est égal à $\sqrt{\Delta_{\mathrm{F}}}$.

On note $\mathrm{U_F}$ le *groupe des unités* de $\mathscr{O}_{\mathrm{F}}$; c'est aussi l'ensemble des éléments u de $\mathscr{O}_{\mathrm{F}}$ vérifiant $\mathrm{N}_{\mathrm{F}/\mathbf{Q}}(u) = \pm 1$. D'après un théorème de Dirichlet, ce groupe est de la forme $\boldsymbol{\mu}_{\mathrm{F}} \times \mathbf{Z}^{r_1+r_2-1}$, où $\boldsymbol{\mu}_{\mathrm{F}}$ est le groupe des racines de l'unité contenues dans F ; c'est un groupe fini dont on note w_{F} le cardinal. On note $\ell : \mathrm{F}^* \to \mathbf{R}^{r_1+r_2}$ l'application

$$x \longmapsto (\log(|\sigma_1(x)|, \dots, \log|\sigma_{r_1}(x)|, 2\log|\sigma_{r_1+1}(x)|, \dots, 2\log|\sigma_{r_1+r_2}(x)|).$$

L'image de $\mathrm{U_F}$ par ℓ est un réseau de l'hyperplan d'équation

$$\lambda_1 + \cdots + \lambda_{r_1+r_2} = 0.$$

On définit le *régulateur* $\mathrm{R_F}$ de F comme le volume de ce réseau : si $\varepsilon_1, \dots, \varepsilon_{r_1+r_2-1}$ est une base du $\mathbf{Z}^{r_1+r_2-1}$ contenu dans $\mathrm{U_F}$, alors $\mathrm{R_F}$ est la valeur absolue du déterminant de la matrice $(r_1 + r_2 - 1) \times (r_1 + r_2 - 1)$ de coefficients $(a_i \log|\sigma_i(\varepsilon_j)|)_{1 \leqslant i,j \leqslant r_1+r_2-1}$, avec $a_i = 1$ (resp. $a_i = 2$) si $i \leqslant r_1$ (resp. si $i \geqslant r_1 + 1$).

L'anneau $\mathscr{O}_{\mathrm{F}}$ n'est pas forcément principal, mais c'est un *anneau de Dedekind* ce qui signifie que tout idéal non nul $\mathfrak{a}$ de $\mathscr{O}_{\mathrm{F}}$ peut se factoriser de manière unique sous la forme $\prod_{i=1}^{r} \mathfrak{p}_i^{k_i}$, où les $\mathfrak{p}_i$ sont des idéaux maximaux distincts de $\mathscr{O}_{\mathrm{F}}$ et les k_i des éléments de $\mathbf{N}$. On dit que deux idéaux non nuls $\mathfrak{a}$ et $\mathfrak{b}$ de $\mathscr{O}_{\mathrm{F}}$ sont équivalents ($\mathfrak{a} \simeq \mathfrak{b}$) s'il existe $\gamma, \delta \in \mathscr{O}_{\mathrm{F}} - \{0\}$ tels que $(\gamma)\mathfrak{a} = (\delta)\mathfrak{b}$ (*i.e.* s'ils diffèrent d'un idéal principal). Comme $\mathfrak{a}\mathfrak{a}' \simeq \mathfrak{b}\mathfrak{b}'$ si $\mathfrak{a} \simeq \mathfrak{b}$ et $\mathfrak{a}' \simeq \mathfrak{b}'$, la multiplication des idéaux induit une multiplication sur l'ensemble $\mathrm{Pic}(\mathscr{O}_{\mathrm{F}})$ des classes d'équivalences, et $\mathrm{Pic}(\mathscr{O}_{\mathrm{F}})$ est un groupe fini, appelé *groupe des classes d'idéaux* de F, dont on note h_{F} le cardinal (c'est le *nombre de classes* de F). Le groupe $\mathrm{Pic}(\mathscr{O}_{\mathrm{F}})$ mesure la « non principalité » de $\mathscr{O}_{\mathrm{F}}$.

La finitude de $\mathrm{Pic}(\mathscr{O}_{\mathrm{F}})$ et la structure de $\mathrm{U_F}$ sont des théorèmes profonds que le 19-ième siècle nous a légués et les quantités h_{F} et $\mathrm{R_F}$ sont des invariants subtils du corps F qui sont très délicats à calculer. Un des seuls outils dont on dispose pour les étudier est la *formule analytique du nombre de classes* (*cf.* (ii) du théorème I.6.1 ci-dessous).

Si $\mathfrak{a}$ est un idéal non nul de $\mathscr{O}_{\mathrm{F}}$, l'anneau $\mathscr{O}_{\mathrm{F}}/\mathfrak{a}$ est un anneau fini dont on note $\mathrm{N}(\mathfrak{a})$ le cardinal (si $\mathrm{F}=\mathbf{Q}$ et $a\in\mathbf{Z}$, on a $\mathrm{N}((a))=|a|$). Ceci nous permet de définir la *fonction zêta de Dedekind* ζ_{F} de F par la formule

$$\zeta_{\mathrm{F}}(s)=\sum_{\mathfrak{a}\subset\mathscr{O}_{\mathrm{F}}}\mathrm{N}(\mathfrak{a})^{-s}=\prod_{\mathfrak{p}\subset\mathscr{O}_{\mathrm{F}}}(1-\mathrm{N}(\mathfrak{p})^{-s})^{-1},$$

où la somme porte sur les idéaux non nuls de $\mathscr{O}_{\mathrm{F}}$ et le produit sur les idéaux maximaux de $\mathscr{O}_{\mathrm{F}}$. Si $\mathrm{F}=\mathbf{Q}$, on retombe sur la fonction ζ de Riemann et la plupart des propriétés de ζ sont partagées par ζ_{F} ; les démonstrations sont toutefois nettement plus difficiles.

Théorème I.6.1

(i) *La série* $\sum_{\mathfrak{a}\subset\mathscr{O}_{\mathrm{F}}}\mathrm{N}(\mathfrak{a})^{-s}$ *converge absolument pour* $\mathrm{Re}(s)>1$ *et* ζ_{F} *admet un prolongement méromorphe à* $\mathbf{C}$ *tout entier, holomorphe en dehors d'un pôle simple en* $s=1$.

(ii) *Le résidu en* $s=1$ *de* ζ_{F} *est donné par la formule analytique du nombre de classes*

$$\lim_{s\to 1}(s-1)\zeta_{\mathrm{F}}(s)=\frac{2^{r_1}\cdot(2\pi)^{r_2}\cdot h_{\mathrm{F}}\cdot \mathrm{R}_{\mathrm{F}}}{w_{\mathrm{F}}\cdot\sqrt{\Delta_{\mathrm{F}}}}.$$

(iii) *La fonction* ζ_{F} *vérifie l'équation fonctionnelle* $\xi_{\mathrm{F}}(s)=\xi_{\mathrm{F}}(1-s)$, *avec*

$$\xi_{\mathrm{F}}(s)=\Big(\frac{\Gamma(s/2)}{\pi^{s/2}}\Big)^{r_1}\cdot\Big(\frac{\Gamma(s)}{(2\pi)^s}\Big)^{r_2}\cdot\Delta_{\mathrm{F}}^{s/2}\cdot\zeta_{\mathrm{F}}(s).$$

En ce qui concerne les propriétés de rationalité des valeurs aux entiers, on a les résultats suivants :

Théorème I.6.2

(i) *Si* $k\leqslant 0$, *on a* $\zeta_{\mathrm{F}}(k)\in\mathbf{Q}$.

(ii) *Si* F *est totalement réel, alors* $\pi^{-2k[\mathrm{F}:\mathbf{Q}]}\zeta_{\mathrm{F}}(2k)\in\sqrt{\Delta_{\mathrm{F}}}\mathbf{Q}$, *si* k *est un entier* $\geqslant 1$.

Commentaire. Si F n'est pas totalement réel ou si $k\leqslant 0$ est pair ($k\leqslant -2$ si $\mathrm{F}=\mathbf{Q}$), l'équation fonctionnelle implique que l'on a $\zeta_{\mathrm{F}}(k)=0$.

Le (ii) du théorème ci-dessus est une généralisation du théorème d'Euler. Dans le cas général, on dispose d'une conjecture de Zagier

disant que, si $k \geqslant 2$, alors $\zeta_{\mathrm{F}}(k)$ doit pouvoir s'exprimer comme un déterminant de k-logarithmes évalués en les conjugués d'éléments de F (remarquons que la formule analytique du nombre de classes fournit une telle expression pour $k = 1$). Si F est un corps cyclotomique, cette conjecture est un théorème et, dans le cas général, on a le théorème I.6.3 ci-dessous, dû à Deligne et Beilinson, et qui se démontre de la même manière que le théorème I.3.7.

Notons $\mathscr{B}_k(\mathrm{F})$ l'ensemble des combinaisons formelles du type

$$\sum_{j=1}^{m} a_j \cdot \{x_j\}_k,$$

où $a_j \in \mathbf{Z}$ et $x_j \in \mathrm{F} - \{0, 1\}$, telles que, quels que soient les morphismes de groupes $v_1, \ldots, v_k$ de F^* dans $\mathbf{Z}$, on ait

$$\sum_{j=1}^{m} a_j \cdot v_1(x_j) \cdots v_{k-1}(x_j) \cdot v_k(1 - x_j) = 0.$$

Si σ est un plongement de F dans $\mathbf{C}$ et

$$x = \sum_{j=1}^{m} a_j \cdot \{x_j\}_k \in \mathscr{B}_k(\mathrm{F}),$$

on note $\mathrm{P}_k(\sigma(x))$ la quantité $\sum_{j=0}^{m} a_j \cdot \mathrm{P}_k(\sigma(x_j))$. Remarquons que l'on a $\mathrm{P}_k(\sigma_i(x)) = 0$ si k est pair et $1 \leqslant i \leqslant r_1$ puisque P_k est identiquement nul sur $\mathbf{R}$ si k est pair.

Théorème I.6.3. *Si $k \geqslant 2$ est un entier pair (respectivement impair) et si $x^{(j)}$ pour $r_1 + 1 \leqslant j \leqslant r_1 + r_2$ (resp. $1 \leqslant j \leqslant r_1 + r_2$) sont des éléments de $\mathscr{B}_k(\mathrm{F})$, alors le déterminant de la matrice $r_2 \times r_2$ (respectivement $(r_1 + r_2) \times (r_1 + r_2)$) de coefficients $\mathrm{P}_k(\sigma_i(x^{(j)}))$, pour $r_1 \leqslant i, j \leqslant r_1 + r_2$ (respectivement $1 \leqslant i, j \leqslant r_1 + r_2$) est un multiple rationnel de $\pi^{-(r_1+r_2)k}\sqrt{\Delta_{\mathrm{F}}}\zeta_{\mathrm{F}}(k)$ (resp. $\pi^{-r_2 k}\sqrt{\Delta_{\mathrm{F}}}\zeta_{\mathrm{F}}(k)$).*

Pour démontrer la conjecture de Zagier, il suffirait donc de prouver que l'on peut trouver des éléments $x^{(j)}$ de $\mathscr{B}_k(\mathrm{F})$ tels que le déterminant correspondant soit non nul.

Chapitre II. Propriétés diophantiennes des valeurs de la fonction zêta aux entiers positifs

II.1. Le théorème de Rivoal

Ce § est consacré à la démonstration du résultat de Rivoal mentionné dans l'introduction. Plus précisément, nous allons démontrer le théorème suivant.

Théorème II.1.1

(i) *La dimension du sous-*$\mathbf{Q}$*-espace vectoriel de* $\mathbf{R}$ *engendré par les* $\zeta(2n+1)$, $n \in \mathbf{N} - \{0\}$ *est infinie.*

(ii) *La dimension du sous-*$\mathbf{Q}$*-espace vectoriel de* $\mathbf{R}$ *engendré par les* $\zeta(2n)$, $n \in \mathbf{N} - \{0\}$ *est infinie.*

Remarque II.1.2. Comme $\pi^{-2n}\zeta(2n)$ est rationnel, le (ii) est équivalent à la transcendance de π, ce qui nous fournit une nouvelle démonstration du théorème de Lindemann (et donc de l'impossibilité de la quadrature du cercle) et montre que les $\zeta(2n)$, $n \in \mathbf{N} - \{0\}$, sont en fait linéairement indépendants sur $\mathbf{Q}$.

Pour démontrer que des nombres réels sont linéairement indépendants sur $\mathbf{Q}$, il « suffit » de produire des combinaisons linéaires à coefficients entiers entre ces nombres qui sont non nulles et « petites ». Par exemple, pour démontrer que deux nombres v_1, v_2 sont linéairement indépendants sur $\mathbf{Q}$, il suffit de produire deux suites d'entiers $(a_{n,i})_{n\in\mathbf{N}}$, pour $i = 1, 2$, telles que l'on ait $a_{n,1}v_1 + a_{n,2}v_2 \neq 0$, pour n assez grand, et $\lim_{n\to+\infty} a_{n,1}v_1 + a_{n,2}v_2 = 0$. Dans le cas général, on dispose du critère suivant, dû à Nesterenko, et dont la démonstration (assez technique) est donnée au n° 6.

Théorème II.1.3 **(Critère de Nesterenko).** *Soient* $v_1, \dots, v_b$ *des nombres réels. On suppose qu'il existe* $\mathrm{B} > 1$, $\mathrm{A} > 0$ *et* b *suites d'entiers* $(a_{n,j})_{n\in\mathbf{N}}$, $j \in \{1, \dots, b\}$ *telles que l'on ait* :

(i) $\sup_{1\leqslant j\leqslant b} |a_{n,j}| \leqslant \mathrm{B}^{n+o(n)}$;

(ii) $|a_{n,1}v_1 + \cdots + a_{n,b}v_b| = \mathrm{A}^{n+o(n)}$.

Alors la dimension du sous-$\mathbf{Q}$*-espace vectoriel de* $\mathbf{R}$ *engendré par* $v_1, \dots, v_b$ *est* $\geqslant 1 - \frac{\log \mathrm{A}}{\log \mathrm{B}}$.

Dans le cas qui nous intéresse, on dispose d'une machine (voir la proposition II.1.4 ci-dessous) à fabriquer des relations linéaires à coefficients rationnels entre les $\zeta(n)$, $n \geqslant 2$, en partant de fonctions rationnelles n'ayant des pôles qu'aux entiers $\leqslant 0$; le problème est alors de bien choisir ces fonctions rationnelles pour que les combinaisons linéaires ainsi obtenues soient petites.

1. Génération de combinaisons linéaires entre les $\zeta(n)$

Soit $a \in \mathbf{N}$, et soit $\mathrm{F} \in \mathbf{Q}(\mathrm{X})$, de degré $\leqslant -2$, n'ayant des pôles qu'aux entiers $\leqslant 0$, ces pôles étant d'ordre $\leqslant a$. Soient $\alpha_{j,k}$, pour $j \geqslant 0$, $1 \leqslant k \leqslant a$ et α_k pour $1 \leqslant k \leqslant a$, les rationnels définis via la décomposition en éléments simples de $\mathrm{F}(\mathrm{X})$ par

$$\mathrm{F}(\mathrm{X}) = \sum_{j=0}^{+\infty} \sum_{k=1}^{a} \frac{\alpha_{j,k}}{(\mathrm{X}+j)^k} \quad \text{et} \quad \alpha_k = \sum_{j=0}^{+\infty} \alpha_{j,k}.$$

Remarquons que l'on a $\alpha_{j,k} = 0$ sauf pour un nombre fini de couples (j,k) et donc que les séries ci-dessus sont en fait des sommes finies et, d'autre part, que $\alpha_1 = 0$ car on a supposé F de degré $\leqslant -2$.

Proposition II.1.4. *La série $\sum_{m=1}^{+\infty} \mathrm{F}(m)$ converge absolument, et on a*

$$\sum_{m=1}^{+\infty} \mathrm{F}(m) = \sum_{k=2}^{a} \alpha_k \zeta(k) - \sum_{k=1}^{a} \sum_{j=0}^{+\infty} \alpha_{j,k} \Big(\sum_{u=1}^{j} \frac{1}{u^k} \Big).$$

Démonstration.

La convergence absolue découle de l'hypothèse $\deg \mathrm{F} \leqslant -2$. Si $\mathrm{N} \in \mathbf{N}$ est tel que $\alpha_{j,k} = 0$ si $j \geqslant \mathrm{N}+1$, on a

$$\sum_{m=1}^{\mathrm{M}} \sum_{j=0}^{\mathrm{N}} \frac{\alpha_{j,1}}{m+j} = \Big(\sum_{j=0}^{\mathrm{N}} \alpha_{j,1} \Big) \Big(\sum_{u=1}^{\mathrm{N+M}} \frac{1}{u} \Big) - \sum_{j=0}^{\mathrm{N}} \alpha_{j,1} \Big(\sum_{u=1}^{j} \frac{1}{u} + \sum_{u=\mathrm{M}+j+1}^{\mathrm{N+M}} \frac{1}{u} \Big),$$

et comme $\sum_{j=0}^{\mathrm{N}} \alpha_{j,1} = 0$, on obtient

$$\sum_{m=1}^{+\infty} \sum_{j=0}^{\mathrm{N}} \frac{\alpha_{j,1}}{m+j} = -\sum_{j=0}^{\mathrm{N}} \alpha_{j,1} \Big(\sum_{u=1}^{j} \frac{1}{u} \Big).$$

D'autre part, on a

$$\sum_{m=1}^{+\infty}\sum_{j=0}^{\mathrm{N}}\sum_{k=2}^{a}\frac{\alpha_{j,k}}{(m+j)^k}$$
$$=\sum_{k=2}^{a}\sum_{j=0}^{\mathrm{N}}\sum_{u=j+1}^{+\infty}\frac{\alpha_{j,k}}{u^k}=\sum_{k=2}^{a}\sum_{j=0}^{\mathrm{N}}\alpha_{j,k}\Big(\zeta(k)-\sum_{u=1}^{j}\frac{1}{u^k}\Big)$$
$$=\sum_{k=2}^{a}\alpha_k\zeta(k)-\sum_{k=2}^{a}\sum_{j=0}^{\mathrm{N}}\alpha_{j,k}\Big(\sum_{u=1}^{j}\frac{1}{u^k}\Big),$$

et il n'y a plus qu'à faire la somme des deux expressions ci-dessus pour conclure.

2. Un choix judicieux de fonction rationnelle

Soient a et r deux entiers vérifiant $2r\leqslant a$ et $r\geqslant 1$. (On cherche des combinaisons linéaires à coefficients entiers entre 1 et les $\zeta(k)$, $k\leqslant a$, et r est un paramètre que l'on ajustera de manière à ce que ces combinaisons linéaires soient les plus petites possibles.) Soit

$$\mathrm{F}_n(\mathrm{X})=(n!)^{a-2r}\frac{(\mathrm{X}-rn)(\mathrm{X}-rn+1)\cdots(\mathrm{X}+(r+1)n)}{(\mathrm{X}(\mathrm{X}+1)\cdots(\mathrm{X}+n))^{a+1}}$$
$$=(n!)^{a-2r}\frac{(\mathrm{X}-rn)\cdots(\mathrm{X}-1)(\mathrm{X}+n+1)\cdots(\mathrm{X}+(r+1)n)}{(\mathrm{X}(\mathrm{X}+1)\cdots(\mathrm{X}+n))^{a}}.$$

Soient aussi $\alpha_{j,k}^{(n)}$, pour $0\leqslant j\leqslant n$, $1\leqslant k\leqslant a$ et $\alpha_k^{(n)}$ pour $1\leqslant k\leqslant a$, les rationnels définis via la décomposition en éléments simples de $\mathrm{F}_n(\mathrm{X})$ par

$$\mathrm{F}_n(\mathrm{X})=\sum_{j=0}^{n}\sum_{k=1}^{a}\frac{\alpha_{j,k}^{(n)}}{(\mathrm{X}+j)^k}\quad\text{et}\quad\alpha_k^{(n)}=\sum_{j=0}^{n}\alpha_{j,k}^{(n)}.$$

Cette fraction rationnelle a un certain nombre de propriétés intéressantes. — F_n est de degré $(2r+1)n+1-(n+1)(a+1)=(2r-a)n-a\leqslant -a\leqslant -2$ et a des pôles d'ordre a en $0,-1,\dots,-n$; la série $\mathrm{S}_n=\sum_{m=1}^{+\infty}\mathrm{F}_n(m)$ converge donc absolument et on a

$$\mathrm{S}_n=\beta^{(n)}+\sum_{k=2}^{a}\alpha_k^{(n)}\zeta(k)\quad\text{avec }\beta^{(n)}=-\sum_{k=1}^{a}\sum_{j=0}^{n}\alpha_{j,k}^{(n)}\Big(\sum_{u=1}^{j}\frac{1}{u^k}\Big).$$

— $\mathrm{F}_n(1) = \cdots = \mathrm{F}_n(rn) = 0$, ce qui montre que la somme définissant S_n ne commence qu'à $m = rn+1$ [*i.e.* $\mathrm{S}_n = \sum_{m=0}^{+\infty} \mathrm{F}_n(rn+m+1)$] et assure que S_n est petit.

— $\mathrm{F}_n(m) \geqslant 0$ si $m \geqslant 1$, ce qui assure que S_n est non nul.

— F_n vérifie l'équation fonctionnelle $\mathrm{F}_n(-n-\mathrm{X}) = (-1)^{(n+1)a}\mathrm{F}(\mathrm{X})$, ce qui nous fournit les relations $\alpha^{(n)}_{n-j,k} = (-1)^{k+(n+1)a}\alpha^{(n)}_{j,k}$, si $1 \leqslant k \leqslant a$ et $0 \leqslant j \leqslant n$; ces relations impliquent que $\alpha^{(n)}_k = 0$ si $k+(n+1)a$ est impair, ce qui est le point crucial pour arriver à séparer les valeurs aux entiers pairs des valeurs aux entiers impairs. (Pour ne garder que les valeurs aux entiers pairs, il suffit de prendre a pair, et pour ne garder que les valeurs aux entiers impairs, il suffit de prendre a impair et n pair.)

Pour pouvoir appliquer le critère de Nesterenko (*cf.* n° 5), il s'agit alors d'évaluer précisément S_n (*cf.* n° 4), majorer les coefficients $\alpha^{(n)}_k$ et le p.p.c.m. de leurs dénominateurs (*cf.* n° 3) car on a besoin de combinaisons linéaires à coefficients *entiers*.

3. Propriétés archimédiennes et arithmétiques des $\alpha^{(n)}_k$

Notons d_n le p.p.c.m. de $1, 2, \ldots, n$.

Proposition II.1.5. *Si $1 \leqslant k \leqslant a$ et $0 \leqslant j \leqslant n$, alors*

$$d_n^{a-k}\alpha^{(n)}_{k,j} \in \mathbf{Z} \quad \text{et} \quad |\alpha^{(n)}_{k,j}| \leqslant (2n)^{a-1}(a-1)!2^{na}(r+1)^{2(r+1)n}.$$

Vu la relation entre $\beta^{(n)}$, les $\alpha^{(n)}_k$ et les $\alpha^{(n)}_{k,j}$, on en déduit le résultat suivant :

Corollaire II.1.6

(i) *$d_n^a\beta^{(n)} \in \mathbf{Z}$ et $d_n^{a-k}\alpha^{(n)}_k \in \mathbf{Z}$ si $n \in \mathbf{N}$ et $k \in \{2, \ldots, a\}$.*

(ii) *$|\beta^{(n)}| \leqslant (2^a(r+1)^{2r+2})^{n+o(n)}$ et $|\alpha^{(n)}_k| \leqslant (2^a(r+1)^{2r+2})^{n+o(n)}$, si $k \in \{2, \ldots, a\}$.*

Pour démontrer la proposition II.1.5, écrivons $\mathrm{F}_n(\mathrm{X})$ sous la forme

$$\mathrm{F}_n(\mathrm{X}) = \prod_{i=1}^{a} \frac{n!\,\mathrm{P}_i(\mathrm{X})}{\mathrm{X}(\mathrm{X}+1)\cdots(\mathrm{X}+n)},$$

où $\mathrm{P}_i(\mathrm{X})$ est le polynôme défini par

$$\mathrm{P}_i(\mathrm{X}) = \begin{cases} \binom{\mathrm{X}-1-(i-1)n}{n} & \text{si } 1 \leqslant i \leqslant r, \\ \binom{\mathrm{X}+(i-r+1)n}{n} & \text{si } r+1 \leqslant i \leqslant 2r, \\ 1 & \text{si } 2r+1 \leqslant i \leqslant a, \end{cases}$$

et $\binom{\mathrm{X}}{n}$ est le polynôme de degré n défini par $\binom{\mathrm{X}}{0} = 1$, et $\binom{\mathrm{X}}{n} = \frac{\mathrm{X}(\mathrm{X}-1)\cdots(\mathrm{X}-n+1)}{n!}$ si $n \geqslant 1$. Nous aurons besoin d'un certain nombre de résultats préparatoires.

Lemme II.1.7. *Soit $\mathrm{Q} \in \mathbf{Q}[\mathrm{X}]$, de degré $\leqslant n$ prenant des valeurs entières aux entiers et soient $(\beta_j(\mathrm{Q}))_{0\leqslant j\leqslant n}$ les rationnels définis par la décomposition en éléments simples de la fraction rationnelle*

$$\mathrm{F}(\mathrm{X}) = \frac{n!\,\mathrm{Q}(\mathrm{X})}{\mathrm{X}(\mathrm{X}+1)\cdots(\mathrm{X}+n)} = \sum_{j=0}^{n} \frac{\beta_j(\mathrm{Q})}{\mathrm{X}+j}$$

et $\mathrm{B}(\mathrm{Q}) = \sup_{0\leqslant j\leqslant n} |\beta_j(\mathrm{Q})|$. Alors $\beta_j \in \mathbf{Z}$ quel que soit $j \in \{0,\dots,n\}$ et $\mathrm{B}(\mathrm{Q}) \leqslant 2^n \sup_{0\leqslant j\leqslant n} |\mathrm{Q}(-j)|$.

Démonstration. On a

$$\beta_j(\mathrm{Q}) = \lim_{\mathrm{X}\to -j} (\mathrm{X}+j)\mathrm{F}(\mathrm{X}) = (-1)^j \binom{n}{j} \mathrm{Q}(-j).$$

Lemme II.1.8. *Soient Q_i pour $1 \leqslant i \leqslant a$ des polynômes de degrés $\leqslant n$ prenant des valeurs entières aux entiers. Soient $\alpha_{j,k}$ pour $0 \leqslant j \leqslant n$ et $1 \leqslant k \leqslant a$, définis grâce à la décomposition en élément simples de*

$$\mathrm{F}(\mathrm{X}) = \prod_{i=1}^{a} \frac{n!\,\mathrm{Q}_i(\mathrm{X})}{\mathrm{X}(\mathrm{X}+1)\cdots(\mathrm{X}+n)} = \sum_{j=0}^{n} \sum_{k=1}^{a} \frac{\alpha_{j,k}}{(\mathrm{X}+j)^k}.$$

Alors $d_n^{a-k}\alpha_{j,k} \in \mathbf{Z}$ et $|\alpha_{j,k}| \leqslant (2n)^{a-1}(a-1)! \prod_{i=1}^{a} \mathrm{B}(\mathrm{Q}_i)$, quels que soient $0 \leqslant j \leqslant n$ et $1 \leqslant k \leqslant a$.

Démonstration. La démonstration se fait par récurrence sur a, le cas $a = 1$ étant contenu dans le lemme II.1.7. Si $a \geqslant 2$, l'hypothèse de récurrence permet d'écrire

$$\prod_{i=1}^{a-1} \frac{n!\,\mathrm{Q}_i(\mathrm{X})}{\mathrm{X}(\mathrm{X}+1)\cdots(\mathrm{X}+n)} = \sum_{j=0}^{n} \sum_{k=1}^{a-1} \frac{\gamma_{j,k}}{(\mathrm{X}+j)^k},$$

avec $d_n^{a-1-k}\gamma_{j,k} \in \mathbf{Z}$ et $|\gamma_{j,k}| \leqslant (2n)^{a-2}(a-2)!\prod_{i=1}^{a-1}\mathrm{B}(\mathrm{Q}_i)$.

Maintenant, si $j_1 \neq j_2$, on a

$$\frac{1}{(\mathrm{X}+j_1)(\mathrm{X}+j_2)^\ell} = \frac{1}{(j_2-j_1)^\ell(\mathrm{X}+j_1)} - \sum_{k=1}^{\ell}\frac{1}{(j_2-j_1)^{\ell+1-k}(\mathrm{X}+j_2)^k}.$$

On en déduit la formule

$$\alpha_{j,k} = \begin{cases} \beta_j(\mathrm{P}_a)\gamma_{j,k-1} - \displaystyle\sum_{\ell\geqslant k}\sum_{j'\neq j}\frac{\beta_{j'}(\mathrm{P}_a)\gamma_{j,\ell}}{(j-j')^{\ell+1-k}} & \text{si } k \geqslant 2, \\ \displaystyle\sum_{\ell\geqslant 1}\sum_{j'\neq j}\frac{\beta_j(\mathrm{P}_a)\gamma_{j',\ell} - \beta_{j'}(\mathrm{P}_a)\gamma_{j,\ell}}{(j-j')^\ell} & \text{si } k = 1. \end{cases}$$

La somme dans le membre de droite comporte au plus $2n(a-1)$ termes et chacun de ces termes est de valeur absolue

$$\leqslant \mathrm{B}(\mathrm{P}_a)(2n)^{a-2}(a-2)!\prod_{i=1}^{a-2}\mathrm{B}(\mathrm{P}_i);$$

on en tire la majoration voulue pour $|\alpha_{j,k}|$.

Finalement, on a

$$d_n^{a-k}\frac{\beta_{j'}(\mathrm{P}_a)\gamma_{j,\ell}}{(j-j')^{\ell+1-k}} = d_n^{a-1-\ell}\gamma_{j,\ell}\cdot\frac{d_n^{\ell+1-k}}{(j-j')^{\ell+1-k}}\cdot\beta_{j'}(\mathrm{P}_a) \in \mathbf{Z}$$

et
$$d_n^{a-k}\gamma_{j,k-1} = d_n^{a-1-(k-1)}\gamma_{j,k-1} \in \mathbf{Z},$$

et tous les termes intervenant dans le calcul de $\alpha_{j,k}$ sont entiers ; il en est donc de même de $\alpha_{j,k}$, ce qui termine la démonstration.

Lemme II.1.9. *Si $1 \leqslant i \leqslant a$, alors P_i est de degré $\leqslant n$ et prend des valeurs entières aux entiers. De plus, on a*

$$\mathrm{B}(\mathrm{P}_i) \leqslant \begin{cases} 2^n\dfrac{((i+1)n)!}{(in)!n!} & \textit{si } 1 \leqslant i \leqslant r, \\ 2^n\dfrac{((i-r+1)n)!}{((i-r)n)!n!} & \textit{si } r+1 \leqslant i \leqslant 2r, \\ 2^n & \textit{si } 2r+1 \leqslant i \leqslant a. \end{cases}$$

Démonstration. On a $\binom{0}{n} \in \mathbf{Z}$ de manière évidente, et $\binom{\mathrm{X}+1}{n+1} - \binom{\mathrm{X}}{n+1} = \binom{\mathrm{X}}{n}$, ce qui montre par une récurrence double que les polynômes $\binom{\mathrm{X}}{n}$ prennent des valeurs entières aux entiers ; il en est donc de même des P_i. La majoration de $\mathrm{B}(\mathrm{P}_i)$, quant à elle, s'obtient en en majorant

le coefficient binomial $\binom{n}{j}$ par 2^n et en constatant que le maximum de $|\mathrm{P}_i(-j)|$ pour $0 \leqslant j \leqslant n$ est atteint en $j = 0$ (resp. $j = n$) si $1 \leqslant i \leqslant r$ (resp. $r+1 \leqslant i \leqslant 2r$).

La proposition II.1.5 est une conséquence des lemmes II.1.9 et II.1.8 et de la majoration

$$\prod_{i=1}^{a} \mathrm{B}(\mathrm{P}_i) \leqslant 2^{na}\Big(\frac{((r+1)n)!}{(n!)^{r+1}}\Big)^2 \leqslant 2^{na}(r+1)^{2(r+1)n},$$

la première inégalité s'obtenant en multipliant les majorations obtenues dans le lemme II.1.9 et la seconde s'obtenant en utilisant la majoration des coefficients multinomiaux $\frac{m!}{m_1!\cdots m_c!} \leqslant c^m$ si $m_1+\cdots+m_c=m$.

4. Évaluation de S_n

Une expression de S_n sous forme d'une intégrale (prop. II.1.12) va nous permettre d'étudier le comportement asymptotique de S_n et, en particulier, de démontrer le résultat suivant.

Proposition II.1.10

(i) *Il existe* $\mathrm{A}_0 > 0$ *tel que* $\lim_{n\to+\infty} \mathrm{S}_n^{1/n} = \mathrm{A}_0$.

(ii) *On a* $\mathrm{A}_0 \leqslant (2r+1)^{2r+1} r^{2r-a}$.

Lemme II.1.11. *Si* $a, b \in \mathbf{N}$, *alors* $\displaystyle\int_0^1 x^a(1-x)^b\,dx = \frac{a!\ b!}{(a+b+1)!}$.

Démonstration. C'est un cas particulier de la formule $\displaystyle\int_0^1 x^{s-1}(1-x)^{t-1}\,dx = \frac{\Gamma(s)\Gamma(t)}{\Gamma(s+t)}$ d'Euler. (On peut aussi utiliser la formule $\displaystyle\int_0^1 x^a(1-x)^b\,dx = \frac{b}{a+1}\int_0^1 x^{a+1}(1-x)^{b-1}\,dx$.)

Proposition II.1.12. *On a*

$$\mathrm{S}_n = \frac{((2r+1)n+1)!}{(n!)^{2r+1}} \int_{[0,1]^{a+1}} \frac{\prod_{\ell=1}^{a+1} x_\ell^{nr}(1-x_\ell)^n\,dx_\ell}{(1-x_1\cdots x_{a+1})^{(2r+1)n+2}}.$$

Démonstration. Si $k \geqslant 1$ et $|x| < 1$, on a

$$(1-x)^{-k} = \sum_{m=0}^{+\infty} \binom{m+k-1}{k-1} x^m.$$

Comme toutes les fonctions que l'on considère sont positives, on peut intervertir somme et intégrale pour obtenir

$$\begin{aligned}\int_{[0,1]^{a+1}} & \frac{\prod_{\ell=1}^{a+1} \bigl(x_\ell^{nr}(1-x_\ell)^n\,dx_\ell\bigr)}{(1-x_1\cdots x_{a+1})^k} \\ &= \sum_{m=0}^{+\infty} \binom{m+k-1}{k-1} \Bigl(\int_0^1 x^{rn+m}(1-x)^n\,dx\Bigr)^{a+1} \\ &= \sum_{m=0}^{+\infty} \frac{(m+1)\cdots(m+k-1)}{(k-1)!} \Bigl(\frac{(rn+m)!n!}{((r+1)n+m+1)!}\Bigr)^{a+1}.\end{aligned}$$

Pour $k = (2r+1)n+2$, on a

$$\begin{aligned}\frac{((2r+1)n+1)!}{(n!)^{2r+1}} \frac{(m+1)\cdots(m+k-1)}{(k-1)!} \Bigl(\frac{(rn+m)!n!}{((r+1)n+m+1)!}\Bigr)^{a+1} \\ = \mathrm{F}_n(rn+m+1).\end{aligned}$$

On en tire le résultat car $\mathrm{F}_n(m) = 0$ si $1 \leqslant m \leqslant rn$.

Lemme II.1.13. *On a*

$$\lim_{n\to+\infty} \Bigl(\frac{((2r+1)n+1)!}{(n!)^{2r+1}}\Bigr)^{1/n} = (2r+1)^{2r+1}.$$

Démonstration. C'est une conséquence du comportement asymptotique de $n!$ donné par la formule de Stirling :

$$n! = n^n e^{-n}\sqrt{2\pi n}\,(1+\mathrm{O}(1/n)).$$

Lemme II.1.14. *Soient* K *un compact, f une fonction continue positive sur* K *et μ une mesure sur* K *telle que la mesure de tout ouvert soit finie et > 0. Alors* $\lim_{n\to+\infty} |\int_{\mathrm{K}} f^n\,d\mu|^{1/n} = \sup_{x\in\mathrm{K}} f(x)$.

Démonstration. Exercice.

Passons à la démonstration de la proposition II.1.10. Le lemme II.1.13 et le lemme II.1.14 utilisé pour $\mathrm{K} = [0,1]^{a+1}$,

$$f(x_1,\dots,x_{a+1}) = \frac{\prod_{\ell=1}^{a+1} x_\ell^r(1-x_\ell)}{(1-x_1\cdots x_{a+1})^{2r+1}} \quad \text{et} \quad \mu = \frac{\prod_{\ell=1}^{a+1} dx_\ell}{(1-x_1\cdots x_{a+1})^2},$$

nous permettent de démontrer que l'on a

$$\lim_{n\to+\infty} \mathrm{S}_n^{1/n} = \mathrm{A}_0,$$

avec $$\mathrm{A}_0 = (2r+1)^{2r+1} \sup_{(x_1,\dots,x_{a+1})\in[0,1]^{a+1}} f(x_1,\dots,x_{a+1}).$$

D'autre part, on a $1-x_1\cdots x_{a+1} \geqslant 1-x_\ell$ pour tout $\ell \in \{1,\dots,a+1\}$. On a donc aussi $1-x_1\cdots x_{a+1} \geqslant \prod_{\ell=1}^{a+1}(1-x_\ell)^{1/(a+1)}$, ce qui nous fournit la majoration

$$f(x_1,\dots,x_{a+1}) \leqslant \prod_{\ell=1}^{a+1} \Big(x_\ell^r (1-x_\ell)^{(a-2r)/(a+1)}\Big).$$

Le maximum de $x \to x^r(1-x)^{(a-2r)/(a+1)}$ sur $[0,1]$ est atteint en

$$\rho = \Big(1+\frac{a-2r}{r(a+1)}\Big)^{-1};$$

le maximum de f sur $[0,1]^{a+1}$ est donc inférieur ou égal à

$$\rho^{r(a+1)}(1-\rho)^{a-2r} \leqslant (1-\rho)^{a-2r},$$

et on termine la démonstration de la proposition II.1.10 en utilisant la majoration

$$1-\rho = 1-\frac{1}{1+\frac{a-2r}{r(a+1)}} = \frac{\frac{a-2r}{r(a+1)}}{1+\frac{a-2r}{r(a+1)}} \leqslant \frac{a-2r}{r(a+1)} \leqslant \frac{1}{r}.$$

5. Utilisation du critère de Nesterenko

5.1. *Le plus petit commun multiple des n premiers entiers.* Pour appliquer le critère de Nesterenko, nous aurons besoin d'estimer la taille de d_n ; c'est l'objet de la proposition suivante qui est une conséquence du théorème des nombres premiers.

Proposition II.1.15. *Si d_n désigne le p.p.c.m de $1,2,\dots,n$, alors $d_n = e^{n+o(n)}$.*

Démonstration. Si p est un nombre premier $\leqslant n$, on a $v_p(d_n) = \big[\frac{\log n}{\log p}\big]$, où $[x]$ désigne la partie entière de x, si $x \in \mathbf{R}$. On obtient donc, en notant $\pi(n)$ le nombre de nombres premiers $\leqslant n$,

$$n - \log d_n = \sum_{p\leqslant n} \Big(\log n - \Big[\frac{\log n}{\log p}\Big]\log p\Big) + n - \pi(n)\log n.$$

D'après le théorème des nombres premiers, on a $n-\pi(n)\log n = o(n)$, et il s'agit de prouver que, si on note s_n la somme $\sum_{p\leqslant n} u_p$, avec $u_p = \log n - \left[\frac{\log n}{\log p}\right]\log p$, alors $s_n = o(n)$. Pour cela, choisissons $\varepsilon \in]0, \frac{1}{2}[$. Si $p \leqslant n^{1-\varepsilon}$, on peut utiliser la majoration triviale $|u_p| \leqslant \log n$, et, si $n^{1-\varepsilon} < p \leqslant n$, on a $\left[\frac{\log n}{\log p}\right] = 1$ et $|u_p| \leqslant \varepsilon \log n$. On en déduit la majoration

$$s_n \leqslant \log n\big(\pi(n^{1-\varepsilon}) + (\pi(n) - \pi(n^{1-\varepsilon}))\varepsilon\big) \leqslant \varepsilon n + o(n).$$

Ceci étant vrai quel que soit $\varepsilon \in]0, \frac{1}{2}[$, on en déduit le résultat.

Remarque II.1.16. Pour démontrer le théorème de Rivoal, on a juste besoin de savoir qu'il existe $\gamma > 0$ tel que $d_n = \mathrm{O}(\gamma^{n+o(n)})$, et on peut montrer (exercice) que l'on peut prendre $\gamma = 4$ en utilisant le fait que le produit des nombres premiers compris entre $m+1$ et $2m$ divise $\binom{2m}{m} \leqslant 2^{2m}$, ce qui permet de se passer du théorème des nombres premiers. Le prix à payer est que les constantes sont un peu moins bonnes (remarque II.1.18).

5.2. *Minoration de la dimension du* $\mathbf{Q}$*-espace vectoriel engendré par les* $\zeta(2j+1)$

Nos efforts sont récompensés par la proposition suivante qui fournit une minoration non triviale de la dimension du $\mathbf{Q}$-espace vectoriel engendré par les valeurs de la fonction zêta aux entiers pairs ou impairs.

Proposition II.1.17. *Soit $a \geqslant 3$ un entier impair. Soit $\delta_{\mathrm{pair}}(a)$ (resp. $\delta_{\mathrm{impair}}(a)$) la dimension du sous-$\mathbf{Q}$-espace vectoriel de $\mathbf{R}$ engendré par 1 et les $\zeta(k)$, k pair (resp. k impair), $2 \leqslant k \leqslant a$. Alors, quel que soit $r \leqslant a/2$, les dimensions $\delta_{\mathrm{pair}}(a)$ et $\delta_{\mathrm{impair}}(a)$ sont minorées par*

$$1 + \frac{(a-2r)\log r - a - (2r+1)\log(2r+1)}{a + a\log 2 + (2r+2)\log(r+1)}.$$

Remarque II.1.18. En prenant $r = \dfrac{a}{(\log a)^2} + \mathrm{O}(1)$, on voit que $\delta_{\mathrm{pair}}(a)$ et $\delta_{\mathrm{impair}}(a)$ sont minorées par $1 + \dfrac{\log a}{1+\log 2} + o(1)$; en particulier,

$\delta_{\mathrm{pair}}(a)$ et $\delta_{\mathrm{impair}}(a)$ tendent vers $+\infty$, ce qui termine la démonstration du théorème II.1.1 (modulo la démonstration de la proposition). Si on veut se passer du théorème des nombres premiers et utiliser une majoration du type $d_n \leqslant \mathrm{O}(\gamma^{n+o(n)})$, les arguments ci-dessous mènent à une minoration de $\delta_{\mathrm{pair}}(a)$ et $\delta_{\mathrm{impair}}(a)$ par $1 + \dfrac{\log a}{\log(2\gamma)} + o(1)$.

Démonstration. La démonstration pour $\delta_{\mathrm{pair}}(a)$ est la même (en un peu plus simple) que la démonstration pour $\delta_{\mathrm{impair}}(a)$; nous nous contenterons donc de traiter le cas de $\delta_{\mathrm{impair}}(a)$.

Soit $b = (a+1)/2$. Soit $v_1 = 1$, et $v_j = \zeta(2j-1)$ si $2 \leqslant j \leqslant b$. Soit $a_{n,1} = d_{2n+1}^a \beta^{(2n+1)}$ et $a_{n,j} = d_{2n+1}^a \alpha_{2j-1}^{(2n+1)}$ si $2 \leqslant j \leqslant b$. D'après la proposition II.1.6, les $a_{n,j}$ sont des entiers, et la conjonction de la proposition II.1.15 et du (ii) du corollaire II.1.6 montre que l'on a

$$\sup_{1\leqslant j\leqslant b} |a_{n,j}| \leqslant \mathrm{B}^{n+o(n)}, \quad \text{avec } \mathrm{B} = \left(e^a 2^a (r+1)^{2r+2}\right)^2.$$

D'autre part, comme a est impair, on a $\alpha_k^{(2n+1)} = 0$ si k est pair, et donc $a_{n,1}v_1 + \cdots + a_{n,b}v_b = d_{2n+1}^a \mathrm{S}_{2n+1}$. La conjonction des propositions II.1.10 et II.1.15 montrent qu'il existe $\mathrm{A} > 0$ tel que

$$|a_{n,1}v_1 + \cdots + a_{n,b}v_b| = \mathrm{A}^{n+o(n)} \quad \text{avec } \mathrm{A} \leqslant \left(e^a (2r+1)^{2r+1} r^{2r-a}\right)^2.$$

Le critère de Nesterenko (th. II.1.3) permet de conclure.

6. Démonstration du critère de Nesterenko

La démonstration du critère de Nesterenko qui suit est adaptée de notes de F. Amoroso ; elle va demander un peu de préparation.

6.1. *Hauteur* $\mathrm{H}(\mathrm{M})$ *d'une matrice* M *à coefficients entiers*

Soit $s \leqslant r$ deux entiers et soit $\mathrm{M} \in \mathbf{M}_{s\times r}(\mathbf{Z})$ une matrice à s lignes et r colonnes et à coefficients entiers $(a_{i,j})_{1\leqslant i\leqslant s,\ 1\leqslant j\leqslant r}$. Si J est une partie à s éléments de $\{1,\dots,r\}$, on note M_{J} la matrice $s \times s$ de coordonnées $(a_{i,j})_{1\leqslant i\leqslant s,\ j\in \mathrm{J}}$, et on pose

$$\mathrm{H}(\mathrm{M}) = \sup_{\mathrm{J}} |\det \mathrm{M}_{\mathrm{J}}|,$$

où J décrit les parties à s éléments de $\{1,\dots,r\}$. Soit

$$w = (w_1,\dots,w_r) \in \mathbf{R}^r.$$

On note $\mathrm{M}w$ le vecteur colonne dont la i-ième coordonnée est $\sum_{j=1}^{r} a_{i,j}w_j$. Si J' est une partie à $s-1$ éléments de $\{1,\dots,r\}$, on note $\mathrm{M}_{\mathrm{J}',w}$ la matrice $s\times s$ obtenue en rajoutant à la matrice $s\times(s-1)$ de coefficients $(a_{i,j})_{1\leqslant i\leqslant s,\ j\in\mathrm{J}'}$, le vecteur $\mathrm{M}w$, et on pose

$$\Delta(\mathrm{M}) = \sup_{\mathrm{J}'} |\det \mathrm{M}_{\mathrm{J}',w}|,$$

où J' décrit les parties à $s-1$ éléments de $\{1,\dots,r\}$.

Lemme II.1.19

(i) *Si* $s=1$ *et* $\mathrm{M}=(a_1,\dots,a_r)\in\mathbf{M}_{1\times r}(\mathbf{Z})$, *alors*

$$\mathrm{H}(\mathrm{M}) = \sup_{1\leqslant i\leqslant r} |a_i| \quad \textit{et} \quad \Delta(\mathrm{M}) = \left|\sum_{i=1}^{r} a_i w_i\right|.$$

(ii) *Si* $\mathrm{M}\in\mathbf{M}_{r\times r}(\mathbf{Z})$, *alors*

$$\mathrm{H}(\mathrm{M}) = |\det \mathrm{M}| \quad \textit{et} \quad \Delta(\mathrm{M}) = |\det \mathrm{M}|(\sup_{1\leqslant i\leqslant r} |w_i|).$$

(iii) *Si* $s\leqslant r$ *et* $\mathrm{M}\in\mathbf{M}_{s\times r}(\mathbf{Z})$, *et si* $\Delta(\mathrm{M})\neq 0$, *alors* $\mathrm{H}(\mathrm{M})\neq 0$.

(iv) *Si* $s\leqslant r$ *et si* $(\mathrm{M}_n)_{n\in\mathbf{N}}$ *est une suite d'éléments de* $\mathbf{M}_{s\times r}(\mathbf{Z})$ *vérifiant* $\Delta(\mathrm{M}_n)\neq 0$ *et* $\lim_{n\to+\infty}\Delta(\mathrm{M}_n)=0$, *alors*

$$\lim_{n\to+\infty} \mathrm{H}(\mathrm{M}_n) = +\infty.$$

(v) *Si* $s\leqslant r-1$, *si* $\mathrm{M}\in\mathbf{M}_{s\times r}(\mathbf{Z})$, *si* $\mathrm{L}\in\mathbf{M}_{1\times r}(\mathbf{Z})$, *et si* $\mathrm{M}\oplus\mathrm{L}$ *désigne l'élément de* $\mathbf{M}_{(s+1)\times r}(\mathbf{Z})$ *obtenu en accolant verticalement* M *et* L, *alors*

$$\mathrm{H}(\mathrm{M}\oplus\mathrm{L}) \leqslant (s+1)\mathrm{H}(\mathrm{L})\mathrm{H}(\mathrm{M}),$$
$$\mathrm{H}(\mathrm{M})\Delta(\mathrm{L}) - s\mathrm{H}(\mathrm{L})\Delta(\mathrm{M}) \leqslant \Delta(\mathrm{M}\oplus\mathrm{L}) \leqslant \mathrm{H}(\mathrm{M})\Delta(\mathrm{L}) + s\mathrm{H}(\mathrm{L})\Delta(\mathrm{M}).$$

Démonstration. Le (i) et le (ii) sont des évidences. Pour démontrer les (iii) et (iv), remarquons que, si J' est une partie à $s-1$ éléments de $\{1,\dots,r\}$, alors

$$\det \mathrm{M}_{\mathrm{J}',w} = \sum_{j\notin\mathrm{J}'} \varepsilon_j w_j \det \mathrm{M}_{\mathrm{J}'\cup\{j\}},$$

où $\varepsilon_j\in\{\pm 1\}$ est un signe dépendant de j. En particulier, si tous les $\det \mathrm{M}_{\mathrm{J}}$ sont nuls, alors il en est de même des $\det \mathrm{M}_{\mathrm{J}',w}$ ce qui démontre le (iii), et si $\mathrm{H}(\mathrm{M}_n)$ est borné, alors $\Delta(\mathrm{M}_n)$ ne peut prendre

qu'un nombre fini de valeurs (car les $\det \mathrm{M}_\mathrm{J}$ sont des entiers), et donc ne peut pas tendre vers 0, ce qui démontre le (iv). Finalement, le (v) se démontre en développant par rapport à la dernière ligne les déterminants $(s+1)\times(s+1)$ qui apparaissent.

6.2. *Reformulation du critère de Nesterenko.* Reprenons les notations du théorème II.1.3. Choisissons une base $w = (w_1, \dots, w_r)$ du sous-$\mathbf{Z}$-module de $\mathbf{R}$ engendré par $v_1, \dots, v_b$, notons C le maximum des valeurs absolues des coordonnées de $v_1, \dots, v_b$ dans cette base, notons $\mathrm{L}_n \in \mathbf{M}_{1\times r}(\mathbf{Z})$ la forme linéaire définie par $\mathrm{L}_n(w) = \sum_{j=1}^b a_{n,j}v_j$, et posons $\mathrm{Q}_n = \sup(\mathrm{Q}_{n-1}, \mathrm{B}^n, \mathrm{C}\sup_{1\leqslant j\leqslant b}|a_{n,j}|)$ et $\lambda = -\frac{\log \mathrm{A}}{\log \mathrm{B}}$. Par construction, Q_n est une suite croissante, tendant vers $+\infty$, et $\mathrm{H}(\mathrm{L}_n) \leqslant \mathrm{Q}_n$. Les hypothèses du théorème II.1.3 se traduisent par

$$\mathrm{Q}_{n+1} = \mathrm{Q}_n^{1+o(1)} \quad \text{et} \quad |\mathrm{L}_n(w)| = \mathrm{Q}_n^{-\lambda+o(1)}.$$

On est donc ramené à démontrer le résultat suivant.

Théorème II.1.20. *Soient $w = (w_1, \dots, w_r) \in \mathbf{R}^r$. On suppose qu'il existe une suite croissante $(\mathrm{Q}_n)_{n\in\mathbf{N}}$ d'éléments de $\mathbf{R}_+^*$, tendant vers $+\infty$ et vérifiant $\mathrm{Q}_{n+1} = \mathrm{Q}_n^{1+o(1)}$, et une suite de formes linéaires à coefficients entiers, $(\mathrm{L}_n)_{n\in\mathbf{N}}$, où $\mathrm{L}_n \in \mathbf{M}_{1\times r}(\mathbf{Z})$, telles que l'on ait*

$$\mathrm{H}(\mathrm{L}_n) \leqslant \mathrm{Q}_n \quad \text{et} \quad |\mathrm{L}_n(w)| = \mathrm{Q}_n^{-\lambda+o(1)}.$$

Alors $\lambda \leqslant r-1$.

Démonstration. La démonstration se fait par l'absurde. Supposons $\lambda > r-1$ et choisissons $\mu \in]r-1, \lambda[$. Nous allons construire, par récurrence sur $s \in \{1, \dots, r\}$, une suite $(\mathrm{M}_n^{(s)})_{n\in\mathbf{N}}$ d'éléments de $\mathbf{M}_{s\times n}(\mathbf{Z})$ vérifiant les conditions suivantes :

(i) $\lim_{n\to+\infty} \mathrm{H}(\mathrm{M}_n^{(s)}) = +\infty$,

(ii) $\Delta(\mathrm{M}_n^{(s)}) \neq 0$ si n est assez grand,

(iii) $\Delta(\mathrm{M}_n^{(s)})^s \mathrm{H}(\mathrm{M}_n^{(s)})^{\mu-s+1}$ tend vers 0 quand n tend vers $+\infty$.

Pour $s = r$, ceci implique que $\Delta(\mathrm{M}_n^{(s)})$ tend vers 0, et donc, d'après le (ii) du lemme II.1.19, que $\det(\mathrm{M}_n^{(s)})$ tend vers 0, ce qui est absurde puisque $\det(\mathrm{M}_n^{(s)})$ est un entier non nul pour n assez grand.

Pour $n = 1$, les hypothèses du théorème II.1.20 font que l'on peut prendre $\mathrm{M}_n^{(1)} = \mathrm{L}_n$. Soit $s \leqslant r-1$, et supposons la suite $(\mathrm{M}_n^{(s)})_{n\in\mathbf{N}}$ construite. Posons $h_n = \mathrm{H}(\mathrm{M}_n^{(s)})$ et $\delta_n = \Delta(\mathrm{M}_n^{(s)})$. Soit $\eta_s \in \,]s\frac{\lambda-\mu}{\mu+1}, (s+1)\frac{\lambda-\mu}{\mu+1}[$, et soit $\varepsilon_s > 0$ tel que l'on ait

$$1 - \frac{\lambda+\varepsilon_s}{s+\eta_s} > \frac{1}{s+\eta_s} - \frac{\mu-s+1}{s}$$

et
$$(s+1)\Big(1 - \frac{\lambda-\varepsilon_s}{s+\eta_s}\Big) + (\mu-s)\Big(1 + \frac{1}{s+\eta_s}\Big) < 0.$$

(Le lemme II.1.21 ci-dessous garantit, par continuité, l'existence d'un tel ε_s.)

Soit $\varphi(n)$ le plus petit entier tel que $h_n^{1/(s+\eta_s)} < \mathrm{Q}_{\varphi(n)+1}$. On a alors $\mathrm{H}(\mathrm{L}_{\varphi(n)}) \leqslant h_n^{1/(s+\eta_s)}$ et, si n est assez grand,

$$h_n^{-(\lambda+\varepsilon_s)/(s+\eta_s)} \leqslant \Delta(\mathrm{L}_{\varphi(n)}) \leqslant h_n^{(-\lambda+\varepsilon_s)/(s+\eta_s)}.$$

Vérifions que la suite $(\mathrm{M}_n^{(s+1)})_{n\in\mathbf{N}}$ définie par $\mathrm{M}_n^{(s+1)} = \mathrm{M}_n^{(s)} \oplus \mathrm{L}_{\varphi(n)}$ répond à nos besoins. Utilisant le (v) du lemme II.1.19, on obtient l'encadrement suivant pour $\Delta(\mathrm{M}_n^{(s+1)})$:

$$h_n^{1-\frac{\lambda+\varepsilon_s}{s+\eta_s}} - s h_n^{\frac{1}{s+\eta_s}}\delta_n \leqslant \Delta(\mathrm{M}_n^{(s+1)}) \leqslant h_n^{1-\frac{\lambda-\varepsilon_s}{s+\eta_s}} + s h_n^{\frac{1}{s+\eta_s}}\delta_n.$$

Par ailleurs, comme $\delta_n^s h_n^{\mu-s+1}$ tend vers 0, on a $\delta_n = o(h_n^{-(\mu-s+1)/s})$, et comme $1 - \frac{\lambda+\varepsilon_s}{s+\eta_s} > \frac{1}{s+\eta_s} - \frac{\mu-s+1}{s}$, le second terme du membre de gauche de la première inégalité est négligeable devant le premier ; on en déduit la non nullité de $\Delta(\mathrm{M}_n^{(s+1)})$ pour n assez grand. Pour les mêmes raisons, on a $\Delta(\mathrm{M}_n^{(s+1)}) = \mathrm{O}(h_n^{1-\frac{\lambda-\varepsilon_s}{s+\eta_s}})$, et comme

$$\mathrm{H}(\mathrm{M}_n^{(s+1)}) \leqslant (s+1)\mathrm{H}(\mathrm{M}_n^{(s)})\mathrm{H}(\mathrm{L}_{\varphi(n)}) = \mathrm{O}(h_n^{1+\frac{1}{s+\eta_s}}),$$

on obtient

$$\Delta(\mathrm{M}_n^{(s+1)})^{s+1}\mathrm{H}(\mathrm{M}_n^{(s+1)})^{\mu-s} = \mathrm{O}\Big(h_n^{(s+1)(1-\frac{\lambda-\varepsilon_s}{s+\eta_s})+(\mu-s)(1+\frac{1}{s+\eta_s})}\Big).$$

Comme on a choisi ε_s et η_s de telle sorte que l'exposant de h_n soit < 0, et comme h_n tend vers $+\infty$, on a

$$\lim_{n\to+\infty} \Delta(\mathrm{M}_n^{(s+1)})^{s+1}\mathrm{H}(\mathrm{M}_n^{(s+1)})^{\mu-s} = 0.$$

Pour terminer la vérification, constatons que, comme $\mu - s > 0$, on a aussi $\lim_{n\to+\infty} \Delta(\mathrm{M}_n^{(s+1)}) = 0$, et donc $\lim_{n\to+\infty} \mathrm{H}(\mathrm{M}_n^{(s+1)}) = +\infty$.

Lemme II.1.21. *Si* $s\frac{\lambda-\mu}{\mu+1} < \eta < (s+1)\frac{\lambda-\mu}{\mu+1}$, *alors*

$$1 - \frac{\lambda}{s+\eta} > \frac{1}{s+\eta} - \frac{\mu-s+1}{s} \tag{II.1.1}$$

et
$$(s+1)\Big(1 - \frac{\lambda}{s+\eta}\Big) + (\mu - s)\Big(1 + \frac{1}{s+\eta}\Big) < 0.$$

Démonstration. Les équivalences suivantes sont immédiates, si on a $\lambda, \mu, \delta > 0$ et $\mu < \lambda$:

$$\begin{aligned} 1 - \frac{\lambda}{s+\eta} > \frac{1}{s+\eta} - \frac{\mu-s+1}{s} &\iff \frac{\mu+1}{s} > \frac{\lambda+1}{s+\eta} \\ &\iff (\mu+1)(s+\eta) > (\lambda+1)s \\ &\iff \eta > s\frac{\lambda-\mu}{\mu+1} \end{aligned}$$

$$\begin{aligned} (s+1)\Big(1 - \frac{\lambda}{s+\eta}\Big) + (\mu - s)\Big(1 + \frac{1}{s+\eta}\Big) < 0 \\ &\iff (s+1)(s+\eta-\lambda) + (\mu-s)(s+\eta+1) < 0 \\ &\iff (s+1)(\mu-\lambda) + \eta(\mu+1) < 0 \\ &\iff \eta < (s+1)\frac{\lambda-\mu}{\mu+1}. \end{aligned}$$

II.2. Nombres Polyzêtas

1. Définition

Le sous-$\mathbf{Q}$-espace vectoriel de $\mathbf{R}$ engendré par 1 et les $\zeta(a)$, a entier $\geqslant 2$ n'est pas une sous-algèbre de $\mathbf{R}$ (du moins il ne devrait pas l'être si les énoncés d'indépendance linéaire sur $\mathbf{Q}$ auxquels on croit sont vrais). Pour remédier à cette situation, on considère un ensemble de nombres un peu plus gros contenant les $\zeta(a)$, à savoir l'ensemble des nombres polyzêtas. Ces nombres polyzêtas ont été introduits par Euler, et viennent de faire un retour tonitruant en mathématique après plus de deux siècles d'oubli ; ils apparaissent naturellement dans un certain nombre de questions à la frontière de la théorie des nombres et de la physique théorique.

Lemme II.2.1. *Si $a_1, \dots, a_k$ sont des entiers $\geqslant 1$, la série*

$$\sum_{n_1>n_2>\cdots>n_k\geqslant 1} \frac{1}{n_1^{a_1}\cdots n_k^{a_k}}$$

converge si et seulement si $a_1 \geqslant 2$.

Démonstration. La démonstration se fait par récurrence sur k en remarquant que $\sum_{n=\mathrm{N}}^{+\infty} \frac{1}{n^a} = \mathrm{O}(\frac{1}{\mathrm{N}^{a-1}})$ si $a > 1$.

Soit A l'ensemble des suites d'entiers $\geqslant 1$ de longueur finie, et $\mathrm{A}_0 \subset \mathrm{A}$ l'ensemble de ces suites dont le premier terme est $\geqslant 2$. Si $\boldsymbol{a} = (a_1, \dots, a_k) \in \mathrm{A}$, on définit la *longueur* de $\boldsymbol{a}$ comme l'entier $d(\boldsymbol{a}) = k$ et son *poids* $|\boldsymbol{a}|$ par la formule $|\boldsymbol{a}| = a_1 + \cdots + a_k$. Si $\boldsymbol{a} = (a_1, \dots, a_k) \in \mathrm{A}_0$, on définit le nombre polyzêta $\zeta(\boldsymbol{a})$ par la formule

$$\zeta(\boldsymbol{a}) = \sum_{n_1>n_2>\cdots>n_k\geqslant 1} \frac{1}{n_1^{a_1}\cdots n_k^{a_k}}.$$

En particulier, si $\boldsymbol{a} = (a)$ n'a qu'un élément, on retombe sur les valeurs de la fonction zêta aux entiers $\geqslant 2$.

Si $\boldsymbol{a} \in \mathrm{A}_0$, on peut aussi écrire $\zeta(\boldsymbol{a})$ sous la forme

$$\sum_{m_1=1}^{+\infty} \cdots \sum_{m_k=1}^{+\infty} \frac{1}{(m_1+\cdots+m_k)^{a_1}(m_2+\cdots+m_k)^{a_2}\cdots m_k^{a_k}}.$$

D'autre part, notant Δ_r l'ensemble des $(t_1, \dots, t_r) \in \mathbf{R}^r$ vérifiant $1 > t_1 > \cdots > t_r > 0$, on a

$$\frac{1}{(m_1+\cdots+m_k)^{a_1}(m_2+\cdots+m_k)^{a_2}\cdots m_k^{a_k}} = \int_{\Delta_{|\boldsymbol{a}|}} \prod_{i=1}^{k} \Big(t_{a_1+\cdots+a_i}^{m_i} \prod_{\ell=a_1+\cdots+a_{i-1}+1}^{a_1+\cdots+a_i} \frac{dt_\ell}{t_\ell} \Big).$$

Sommant alors sur $m_1, \dots, m_k \geqslant 1$, et intervertissant somme et intégrale, ce qui ne pose pas de problème, tous les termes étant positifs, on obtient

$$\zeta(\boldsymbol{a}) = \int_{\Delta_{|\boldsymbol{a}|}} \prod_{i=1}^{k} \Big(\frac{dt_{a_1+\cdots+a_i}}{1-t_{a_1+\cdots+a_i}} \prod_{\ell=a_1+\cdots+a_{i-1}+1}^{a_1+\cdots+a_i-1} \frac{dt_\ell}{t_\ell} \Big).$$

Cette écriture étant relativement pénible, on introduit les formes différentielles $\omega_1(t) = \dfrac{dt}{1-t}$ et $\omega_0(t) = \dfrac{dt}{t}$, ce qui nous donne

$$\zeta(\boldsymbol{a}) = \int_{\Delta_{|\boldsymbol{a}|}} \prod_{\ell=1}^{|\boldsymbol{a}|} \omega_{a_i^*}(t_\ell),$$

où $\boldsymbol{a}^* = (a_1^*, \dots, a_{|\boldsymbol{a}|}^*)$ est défini par $a_i^* = 1$ si $i \in \{a_1, a_1+a_2, \dots, a_1+\cdots+a_k\}$, et $a_i^* = 0$ sinon. D'autre part, on note A^* l'ensemble des suites de longueur finie constituées de 0 et de 1, et se terminant par un 1, et A_0^* le sous-ensemble de ces suites commençant par un 0. L'application $\boldsymbol{a} \mapsto \boldsymbol{a}^*$ induit une bijection de A sur A^* et de A_0 sur A_0^*.

2. Relations quadratiques entre les nombres polyzêtas

2.1. *Relations de type I*

Si r est un entier $\geqslant 1$, notons $\mathrm{D}_r \subset (\mathbf{N} - \{0\})^r$ l'ensemble des r-uplets $(n_1, \dots, n_r)$ d'entiers vérifiant $n_1 > \cdots > n_k \geqslant 1$. Si r et s sont deux entiers, on peut écrire $\mathrm{D}_r \times \mathrm{D}_s$ comme une réunion disjointe d'ensembles de la forme D_d. De manière précise, soit $\Sigma_{r,s}$ l'ensemble des applications $\varphi : \{1, \dots, r+s\} \to \mathbf{N} - \{0\}$ dont l'image est un intervalle contenant 1 (*i.e.* est de la forme $\{1, \dots, d_\varphi\}$) et telles que l'on ait $\varphi(1) < \cdots < \varphi(r)$ et $\varphi(r+1) < \cdots < \varphi(r+s)$. Si $\varphi \in \Sigma_{r,s}$, on définit le sous-ensemble D_φ de $(\mathbf{N} - \{0\})^{r+s}$ comme étant l'ensemble des suites $(n_1, \dots, n_{r+s})$ vérifiant $n_i = n_j$, si $\varphi(i) = \varphi(j)$, et $n_i > n_j$, si $\varphi(i) < \varphi(j)$.

Lemme II.2.2

(i) *On a* $\mathrm{D}_r \times \mathrm{D}_s = \coprod_{\varphi \in \Sigma_{r,s}} \mathrm{D}_\varphi$.

(ii) *L'application* $(m_1, \dots, m_{d(\varphi)}) \mapsto (n_1, \dots, n_{r+s})$, *où* $n_i = m_{\varphi(i)}$, *est une bijection de* $\mathrm{D}_{d(\varphi)}$ *sur* D_φ.

Démonstration. Le (i) s'obtient en ordonnant les coordonnées d'un élément $(n_1, \dots, n_{r+s})$ de $\mathrm{D}_r \times \mathrm{D}_s$. Le (ii) est évident.

Maintenant, si $\boldsymbol{a}$ et $\boldsymbol{b}$ sont deux éléments de A_0, on note $\mathbf{c}$ l'élément de A_0 de longueur $d(\boldsymbol{a}) + d(\boldsymbol{b})$ obtenu en accolant $\boldsymbol{b}$ à $\boldsymbol{a}$. Si $\boldsymbol{a} = (a_1, \dots, a_{d(\boldsymbol{a})})$ et $\boldsymbol{b} = (b_1, \dots, b_{d(\boldsymbol{b})})$, on a $\mathbf{c} = (c_1, \dots, c_{d(\boldsymbol{a})+d(\boldsymbol{b})})$, avec

$c_i = a_i$ si $i \leqslant d(\boldsymbol{a})$ et $c_i = b_{i-d(\boldsymbol{b})}$ si $i \geqslant d(\boldsymbol{a}) + 1$. Si $\varphi \in \Sigma_{d(\boldsymbol{a}),d(\boldsymbol{b})}$, on note $\varphi(\boldsymbol{a}, \boldsymbol{b}) = (c_{\varphi,1}, \ldots, c_{\varphi,d(\varphi)})$ l'élément de A_0 défini par $c_{\varphi,i} = \sum_{\varphi(j)=i} c_j$. (Remarquons que les hypothèses mises sur φ font que l'équation $\varphi(j) = i$ a une ou deux solutions.) Le lemme II.2.2 nous fournit alors les relations quadratiques suivantes, dites *de type I*, entre les nombres polyzêtas.

Proposition II.2.3. *Si $\boldsymbol{a}$ et $\boldsymbol{b}$ sont deux éléments de A_0, alors*

$$\zeta(\boldsymbol{a})\zeta(\boldsymbol{b}) = \sum_{\varphi \in \Sigma_{d(\boldsymbol{a}),d(\boldsymbol{b})}} \zeta(\varphi(\boldsymbol{a}, \boldsymbol{b})).$$

Corollaire II.2.4. *Le sous-$\mathbf{Q}$-espace vectoriel de $\mathbf{R}$ engendré par les nombres polyzêtas est une sous-$\mathbf{Q}$-algèbre de $\mathbf{R}$.*

2.2. *Relations de type II.* Si r est un entier $\geqslant 1$, et si σ est une permutation de $\{1, \ldots, r\}$, soit $\Delta_r(\sigma)$ l'ensemble de r-uplets $(t_1, \ldots, t_r)$ de réels vérifiant $1 > t_{\sigma(1)} > \cdots > t_{\sigma(r)} > 0$. Si r et s sont deux entiers $\geqslant 1$, soit $\mathrm{S}_{r,s}$ l'ensemble des permutations σ de $\{1, \ldots, r+s\}$ telles que

$$\sigma(1) < \sigma(2) < \cdots < \sigma(r) \quad \text{et} \quad \sigma(r+1) < \cdots < \sigma(r+s).$$

Lemme II.2.5. *Si r et s sont des entiers $\geqslant 1$, alors $\Delta_r \times \Delta_s$ est, à des sous-ensembles de mesure nulle près (en fait des faces de codimension 1), la réunion disjointe des $\Delta_{r+s}(\sigma)$, $\sigma \in \mathrm{S}_{r,s}$.*

Démonstration. Il suffit d'ordonner les coordonnées $(t_1, \ldots, t_{r+s})$ de $\Delta_r \times \Delta_s$.

Soient alors $\boldsymbol{a}$ et $\boldsymbol{b}$ deux éléments de A_0, et soit $\mathbf{c}$ l'élément de A_0 obtenu, comme ci-dessus, en accolant $\boldsymbol{b}$ à $\boldsymbol{a}$; alors $\mathbf{c}^*$ s'obtient en accolant $\boldsymbol{b}^*$ à $\boldsymbol{a}^*$. Si $\sigma \in \mathrm{S}_{|\boldsymbol{a}|,|\boldsymbol{b}|}$, on note $\sigma(\boldsymbol{a}, \boldsymbol{b})$ l'élément de A_0 défini par

$$(\sigma(\boldsymbol{a}, \boldsymbol{b}))^* = (c^*_{\sigma(1)}, \ldots, c^*_{\sigma(|\boldsymbol{a}|+|\boldsymbol{b}|)}).$$

On obtient, utilisant le lemme ci-dessus, une seconde série de relations quadratiques, dites *de type II*, en les nombres polyzêtas :

$$\zeta(\boldsymbol{a})\zeta(\boldsymbol{b}) = \int_{\Delta_{|\boldsymbol{a}|}\times\Delta_{|\boldsymbol{b}|}} \prod_{\ell=1}^{|\boldsymbol{a}|+|\boldsymbol{b}|} \omega_{c_i^*}(t_\ell) = \sum_{\sigma\in \mathrm{S}_{|\boldsymbol{a}|,|\boldsymbol{b}|}} \int_{\Delta_{|\boldsymbol{a}|+|\boldsymbol{b}|}(\sigma)} \prod_{\ell=1}^{|\boldsymbol{a}|+|\boldsymbol{b}|} \omega_{c_i^*}(t_\ell)$$
$$= \sum_{\sigma\in \mathrm{S}_{|\boldsymbol{a}|,|\boldsymbol{b}|}} \int_{\Delta_{|\boldsymbol{a}|+|\boldsymbol{b}|}} \prod_{\ell=1}^{|\boldsymbol{a}|+|\boldsymbol{b}|} \omega_{c^*_{\sigma(i)}}(t_\ell) = \sum_{\sigma\in \mathrm{S}_{|\boldsymbol{a}|,|\boldsymbol{b}|}} \zeta(\sigma(\boldsymbol{a},\boldsymbol{b})).$$

3. Relations linéaires entre les nombres polyzêtas

Soit $\boldsymbol{a} = (a_1,\dots,a_k) \in \mathrm{A}_0$. Si on écrit formellement les deux expressions obtenues pour le produit divergent $\zeta(\boldsymbol{a})\zeta(1)$, et qu'on égale les deux expressions que l'on obtient en supprimant les termes apparaissant des deux cotés, on se retrouve avec l'égalité

$$\sum_{j=1}^{k}\sum_{i=1}^{a_j-1} \zeta(a_1,\dots,a_{j-1},i+1,a_j-i,a_{j+1},\dots,a_k) = \sum_{j=1}^{k} \zeta(a_1,\dots,a_{j-1},a_j+1,a_{j+1},\dots,a_k).$$

Dans cette égalité, tous les termes ont un sens (avec les conventions évidentes si $j=1$ ou $j=k$), et cette égalité est, en fait, une vraie égalité, ce qui nous fournit une nouvelle famille de relations, linéaires cette fois, dites *de type III*, en les nombres polyzêtas. Pour transformer ce qui précède en une démonstration, on introduit les fonctions polypolylogarithmes (en une variable)

$$\mathrm{Li}_{\boldsymbol{a}}(z) = \sum_{n_1>\cdots>n_k\geqslant 1} \frac{z^{n_1}}{n_1^{a_1}\cdots n_k^{a_k}} = \int_{z>t_1>\cdots>t_{|\boldsymbol{a}|}>0} \prod_{\ell=1}^{|\boldsymbol{a}|} \omega_{a_i^*}(t_\ell).$$

On a bien évidemment $\mathrm{Li}_{\boldsymbol{a}}(1) = \zeta(\boldsymbol{a})$ si $\boldsymbol{a}\in\mathrm{A}_0$, mais l'intérêt est que $\mathrm{Li}_{\boldsymbol{a}}(z)$ est défini, si $|z|<1$, pour tout $\boldsymbol{a}\in\mathrm{A}$, et pas seulement pour $\boldsymbol{a}\in\mathrm{A}_0$. La méthode qui nous a permis de démontrer les relations de type I et II conduit aux formules suivantes, quels que soient $\boldsymbol{a},\boldsymbol{b}\in\mathrm{A}$:

$$\mathrm{Li}_{\boldsymbol{a}}(z)\mathrm{Li}_{\boldsymbol{b}}(z) = \sum_{\sigma\in \mathrm{S}_{|\boldsymbol{a}|,|\boldsymbol{b}|}} \mathrm{Li}_{\sigma(\boldsymbol{a},\boldsymbol{b})}(z)$$
$$\mathrm{Li}_{\boldsymbol{a}}(z)\mathrm{Li}_{\boldsymbol{b}}(z) = \sum_{\varphi\in\Sigma_{d(\boldsymbol{a}),d(\boldsymbol{b})}} \sum_{n_1>\cdots>n_{d(\varphi)}\geqslant 1} \frac{z^{n_\varphi(1)}}{n_1^{c_{\varphi,1}}\cdots n_{d(\varphi)}^{c_{\varphi,d(\varphi)}}}.$$

On peut utiliser les formules précédentes pour $\boldsymbol{a} = (1)$ et $\boldsymbol{b} = (a_1, \dots, a_k) \in \mathrm{A}_0$. Tous les termes qui apparaissent convergent en $z = 1$ vers le polyzêta correspondant, à l'exception des termes correspondant à $\sigma = \varphi = \mathrm{id}$. La différence de ces deux termes est alors égale à

$$\sum_{n_1 > n_2 > \cdots > n_{k+1} \geqslant 1} \frac{z^{n_1} - z^{n_1+n_2}}{n_1 n_2^{a_1} \cdots n_{k+1}^{a_{k+1}}}.$$

Comme $a_1 \geqslant 1$, et comme on peut majorer $z^{n_1} - z^{n_1+n_2}$ par

$$n_2(1-z)z^{n_1} \leqslant n_2(1-z)z^{(n_1+\cdots+n_{k+1})/(k+1)},$$

cela permet de majorer cette différence par $(1-z)|\log(1-z^{1/(k+1)})|^{k+1}$ qui tend vers 0 quand z tend vers 1. On en déduit les relations de type III.

4. L'algèbre engendrée par les nombres polyzêtas

On vient d'obtenir trois types de relations algébriques entre les nombres polyzêtas et on peut se demander si ce sont « les seules », ce qui se traduit de la manière suivante :

Question II.2.6. *Soit* Z *la* $\mathbf{Q}$*-algèbre de polynômes en les variables* $\mathrm{Z}_{\boldsymbol{a}}$*, pour* $\boldsymbol{a} \in \mathrm{A}_0$*. Soit* $\widehat{\zeta} : \mathrm{Z} \to \mathbf{R}$ *le morphisme d'algèbres défini par* $\widehat{\zeta}(\mathrm{Z}_{\boldsymbol{a}}) = \zeta(\boldsymbol{a})$*. Est-il vrai que* $\ker \widehat{\zeta}$ *est l'idéal de* Z *engendré par les éléments*

$$\mathrm{Z}_{\boldsymbol{a}}\mathrm{Z}_{\boldsymbol{b}} - \sum_{\sigma \in \mathrm{S}_{|\boldsymbol{a}|,|\boldsymbol{b}|}} \mathrm{Z}_{\sigma(\boldsymbol{a},\boldsymbol{b})}, \quad \mathrm{Z}_{\boldsymbol{a}}\mathrm{Z}_{\boldsymbol{b}} - \sum_{\varphi \in \Sigma_{d(\boldsymbol{a}),d(\boldsymbol{b})}} \mathrm{Z}_{\varphi(\boldsymbol{a},\boldsymbol{b})}, \quad \boldsymbol{a}, \boldsymbol{b} \in \mathrm{A}_0,$$

et

$$\sum_{j=1}^{k} \sum_{i=1}^{a_j - 1} \mathrm{Z}_{(a_1,\dots,a_{j-1},i+1,a_j-i,a_{j+1},\dots,a_k)} - \sum_{j=1}^{k} \mathrm{Z}_{(a_1,\dots,a_{j-1},a_j+1,a_{j+1},\dots,a_k)},$$

$(a_1, \dots, a_k) \in \mathrm{A}_0$?

Remarque II.2.7. Les relations de type I, II et III ne font intervenir que des nombres polyzêtas de même poids. Par ailleurs, $\zeta(a)$ est de poids a si $a \geqslant 2$; en particulier, 1 et les $\zeta(a)$, $a \geqslant 2$ sont tous de poids différents et une réponse positive à la question ci-dessus impliquerait qu'ils doivent tous être linéairement indépendants sur $\mathbf{Q}$. Le seul résultat que l'on ait à ce sujet est le théorème de Rivoal qui est

un petit pas. D'un autre côté, le fait d'avoir exhibé des relations multiplicatives entre les nombres polyzêtas peut se révéler une aide précieuse pour attaquer cette question : après tout, on peut déduire du théorème de Rivoal l'indépendance linéaire sur $\mathbf{Q}$ de tous les $\zeta(2a)$, et ceci grâce aux relations multiplicatives vérifiées par ces nombres.

Chapitre III. Formes modulaires

III.1. $\mathbf{SL_2(R)}$ et le demi-plan de Poincaré

Si A est un anneau commutatif, on note $\mathbf{SL}_2(\mathrm{A})$ le groupe des matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, avec $a, b, c, d \in \mathrm{A}$ de déterminant $ad - bc = 1$.

Si $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbf{SL}_2(\mathbf{R})$ et $z \in \mathbf{C} - \{-d/c\}$, on pose

$$\begin{aligned} \gamma z = \frac{az+b}{cz+d} &= \frac{(az+b)(c\overline{z}+d)}{|cz+d|^2} \\ &= \frac{ac|z|^2 + bd + (a+c)x + iy(ad-bc)}{|cz+d|^2}. \end{aligned}$$

En particulier, on a

$$\mathrm{Im}(\gamma z) = \frac{\mathrm{Im}(z)}{|cz+d|^2}, \tag{III.1.1}$$

et le demi-plan de Poincaré $\mathscr{H} = \{z = x + iy,\ y > 0\}$ est stable par $z \mapsto \gamma z$.

Lemme III.1.1. *Si $\gamma_1, \gamma_2 \in \mathbf{SL}_2(\mathbf{R})$ et $z \in \mathscr{H}$, alors $\gamma_1\gamma_2 \cdot z = \gamma_1 \cdot \gamma_2 z$.*

Démonstration. Si $\gamma_1 = \left(\begin{smallmatrix} a_1 & b_1 \\ c_1 & d_1 \end{smallmatrix}\right)$ et $\gamma_2 = \left(\begin{smallmatrix} a_2 & b_2 \\ c_2 & d_2 \end{smallmatrix}\right)$, alors

$$\begin{aligned} \gamma_1(\gamma_2 z) = \frac{a_1\gamma_2 z + b_1}{c_1\gamma_2 z + d_1} &= \frac{a_1 \frac{a_2 z + b_2}{c_2 z + d_2} + b_1}{c_1 \frac{a_2 z + b_2}{c_2 z + d_2} + d_1} \\ &= \frac{(a_1a_2 + b_1c_2)z + (a_1b_2 + b_1d_2)}{(c_1a_2 + d_1c_2)z + (c_1b_2 + d_1d_2)} \end{aligned}$$

et $$\gamma_1\gamma_2 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}\begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1a_2 + b_1c_2 & a_1b_2 + b_1d_2 \\ c_1a_2 + d_1c_2 & c_1b_2 + d_1d_2 \end{pmatrix},$$

ce qui permet de conclure.

Théorème III.1.2

(i) *Si* $\gamma \in \mathbf{SL}_2(\mathbf{R})$, *l'application* $z \mapsto \gamma z$ *appartient au groupe* (pour la composition) $\mathrm{Aut}(\mathscr{H})$ *des bijections holomorphes de* $\mathscr{H}$ *dans* $\mathscr{H}$.

(ii) *L'application qui à* $\gamma \in \mathbf{SL}_2(\mathbf{R})$ *associe l'élément* $z \mapsto \gamma z$ *de* $\mathrm{Aut}(\mathscr{H})$ *est un morphisme de groupes.*

(iii) *L'action de* $\mathbf{SL}_2(\mathbf{R})$ *sur* $\mathscr{H}$ *ainsi définie est transitive et le stabilisateur de* i *est le sous-groupe des rotations* $\mathbf{SO}_2(\mathbf{R})$ *de* $\mathbf{SL}_2(\mathbf{R})$.

(iv) *Si* $\varphi : \mathscr{H} \to \mathscr{H}$ *est holomorphe bijectif, alors il existe* $\gamma \in \mathbf{SL}_2(\mathbf{R})$ *tel que* $\varphi(z) = \gamma z$ *; autrement dit, le morphisme* $\mathbf{SL}_2(\mathbf{R}) \to \mathrm{Aut}(\mathscr{H})$ *défini ci-dessus est surjectif.*

Démonstration. L'holomorphie de $z \mapsto \gamma z$ est une évidence et les points (i) et (ii) sont des conséquences immédiates du lemme précédent.

Pour démontrer la transitivité de l'action de $\mathbf{SL}_2(\mathbf{R})$ sur $\mathscr{H}$, il suffit de prouver que l'on peut envoyer i sur n'importe quel point de $\mathscr{H}$, ce qui suit de la formule

$$\begin{pmatrix} y^{1/2} & y^{-1/2}x \\ 0 & y^{-1/2} \end{pmatrix} \cdot i = \frac{y^{1/2}i + y^{-1/2}x}{y^{-1/2}} = x + iy.$$

Maintenant, si $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ vérifie $\gamma i = i$, alors $ai + b = i(ci + d)$ et donc $c = -b$ et $a = d$. Comme de plus, $ad - bc = 1$, on a $a^2 + b^2 = 1$, et γ est la matrice d'une rotation. Ceci termine la démonstration du (iii). Pour démontrer le (iv), nous aurons besoin du lemme suivant.

Lemme III.1.3. *Soit* $\mathrm{D} = \{z \in \mathbf{C},\ |z| < 1\}$. *Si* $\psi : \mathrm{D} \to \mathrm{D}$ *est une bijection holomorphe vérifiant* $\psi(0) = 0$, *alors il existe* $\lambda \in \mathbf{C}$, $|\lambda| = 1$ *tel que* $\psi(z) = \lambda z$ *quel que soit* $z \in \mathrm{D}$.

Démonstration. On a $\psi'(0) = \frac{1}{2i\pi} \int_{|z|=r} z^{-2}\psi(z)dz$ quel que soit r tel que $0 < r < 1$, et donc $|\psi'(0)| \leqslant r^{-2} \sup_{|z|=r} |\psi(z)| \leqslant r^{-2}$ quel que soit r tel que $0 < r < 1$. Faisant tendre r vers 1, on obtient $|\psi'(0)| \leqslant 1$.

Le même raisonnement appliqué à la bijection holomorphe de D dans D, réciproque de ψ, montre que $|\psi'(0)| \geqslant 1$, et donc $|\psi'(0)| = 1$.

Considérons alors la fonction $g : \mathrm{D} \to \mathrm{D}$ définie par $g(z) = z^{-1}\psi(z)$ si $z \neq 0$ et $g(0) = \psi'(0)$. C'est une fonction holomorphe sur D et le

maximum de $|g|$ sur le disque de centre 0 et de rayon r est atteint sur le cercle de centre 0 et de rayon r ; il est donc $\leqslant r^{-1}\sup_{|z|=r}|\psi(z)|\leqslant r^{-1}$. Faisant tendre r vers 1, on obtient $\sup_{z\in \mathrm{D}}|g(z)|\leqslant 1$. Comme par ailleurs $|g(0)|=1$, la fonction holomorphe g atteint son maximum en un point intérieur à D ; elle est donc constante, ce qui permet de conclure.

Revenons à la démonstration du (iv). Soit $\varphi : \mathscr{H}\to\mathscr{H}$ holomorphe bijective. D'après le (iii), il existe $\gamma\in\mathbf{SL}_2(\mathbf{R})$ tel que $\gamma i=\varphi(i)$. Quitte à remplacer φ par $\gamma^{-1}\circ\varphi$, on peut donc supposer que $\varphi(i)=i$.

Soit h la fonction définie par $h(z)=\frac{i-z}{i+z}$. Un petit calcul montre que h est une bijection holomorphe de $\mathscr{H}$ dans D dont la réciproque h^{-1} est donnée par la formule $h^{-1}(z)=i\frac{1-u}{1+u}$. Considérons alors la composée $\psi=h\circ\varphi\circ h^{-1}$; c'est une bijection holomorphe de D dans D vérifiant $\psi(0)=0$. D'après le lemme précédent, il existe $\lambda\in\mathbf{C}$ de module 1 tel que $\psi(z)=\lambda z$. On a alors

$$\varphi(z)=h^{-1}\circ\psi\circ h(z)=h^{-1}\circ\psi\Big(\frac{i-z}{i+z}\Big)=h^{-1}\Big(\lambda\frac{i-z}{i+z}\Big)=i\frac{1-\lambda\frac{i-z}{i+z}}{1+\lambda\frac{i-z}{i+z}},$$

et, écrivant λ sous la forme $\lambda=e^{2i\theta}$, un petit calcul nous donne

$$\varphi(z)=\frac{\cos\theta\cdot z+\sin\theta}{-\sin\theta\cdot z+\cos\theta},$$

ce qui permet de conclure.

Corollaire III.1.4. *L'application $\gamma\mapsto\gamma i$ induit une bijection de $\mathbf{SL}_2(\mathbf{R})/\mathbf{SO}_2(\mathbf{R})$ sur $\mathscr{H}$.*

Démonstration. C'est une réécriture du (iii) du théorème.

Corollaire III.1.5. *Tout élément γ de $\mathbf{SL}_2(\mathbf{R})$ peut s'écrire de manière unique sous la forme $\gamma=\mathrm{UAR}$, où $\mathrm{U}=\left(\begin{smallmatrix}1&u\\0&1\end{smallmatrix}\right)$ est une matrice unipotente, $\mathrm{A}=\left(\begin{smallmatrix}a&0\\0&a^{-1}\end{smallmatrix}\right)$ est une matrice diagonale avec $a>0$, et $\mathrm{R}\in\mathbf{SO}_2(\mathbf{R})$ est une matrice de rotation.*

Démonstration. Si $\gamma i=x+iy$, on doit poser $a=\sqrt{y}$, $u=x$ et $\mathrm{R}=(\mathrm{UA})^{-1}\gamma$ appartient au stabilisateur de i.

Remarque III.1.6. La décomposition précédente est connue sous le nom de *décomposition d'Iwasawa.*

Proposition III.1.7. *La mesure hyperbolique* $\dfrac{dx\,dy}{y^2}$ *est invariante sous l'action de* $\mathbf{SL}_2(\mathbf{R})$.

Démonstration. Si

$$dz = dx + i\,dy, \qquad \frac{\partial f}{\partial z} = \frac{1}{2}\Big(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}\Big),$$
$$d\overline{z} = dx - i\,dy, \qquad \frac{\partial f}{\partial \overline{z}} = \frac{1}{2}\Big(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\Big),$$

alors

$$dx \wedge dy = \frac{i}{2}\,dz \wedge d\overline{z} \quad \text{et} \quad df = \frac{\partial f}{\partial z}dz + \frac{\partial f}{\partial \overline{z}}d\overline{z},$$

si f est une fonction sur un ouvert de $\mathbf{C}$.

Si $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbf{SL}_2(\mathbf{R})$ et $u = \gamma z$, on a

$$du = \frac{dz}{(cz+d)^2}, \quad d\overline{u} = \frac{d\overline{z}}{(c\overline{z}+d)^2} \quad \text{et} \quad \mathrm{Im}(u) = \frac{\mathrm{Im}(z)}{|cz+d|^2},$$

et donc

$$\frac{du \wedge d\overline{u}}{\mathrm{Im}(u)^2} = \frac{\frac{dz}{(cz+d)^2} \wedge \frac{d\overline{z}}{(c\overline{z}+d)^2}}{\big(\frac{\mathrm{Im}(z)}{|cz+d|^2}\big)^2} = \frac{dz \wedge d\overline{z}}{\mathrm{Im}(z)^2},$$

ce qui permet de conclure.

III.2. Formes automorphes et formes modulaires

1. Facteur d'automorphie

Si $k \in \mathbf{Z}$, si $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbf{SL}_2(\mathbf{R})$, et si $f : \mathscr{H} \to \mathbf{C}$ est une fonction $\mathscr{C}^\infty$, on définit la fonction $f_{|_k}\gamma$ par la formule

$$f_{|_k}\gamma(z) = (cz+d)^{-k} f(\gamma z).$$

Si $\gamma_1 = \left(\begin{smallmatrix} a_1 & b_1 \\ c_1 & d_1 \end{smallmatrix}\right)$ et $\gamma_2 = \left(\begin{smallmatrix} a_2 & b_2 \\ c_2 & d_2 \end{smallmatrix}\right)$ sont deux éléments de $\mathbf{SL}_2(\mathbf{R})$, on a

$$\big(f_{|_k}\gamma_1\big)_{|_k}\gamma_2(z) = (c_2 z + d_2)^{-k}\Big(c_1\frac{a_2 z + b_2}{c_2 z + d_2} + d_1\Big)^{-k} f(\gamma_1\gamma_2 z)$$
$$= \big((c_1a_2 + d_1c_2)z + (c_1b_2 + d_1d_2)\big)^{-k} f(\gamma_1\gamma_2 z) = f_{|_k}\gamma_1\gamma_2(z),$$

et donc $\big(f_{|_k}\gamma_1\big)_{|_k}\gamma_2 = f_{|_k}\gamma_1\gamma_2$.

2. Sous-groupes d'indice fini de $\mathbf{SL}_2(\mathbf{Z})$

Nous allons définir la notion de forme automorphe ou de forme modulaire de poids k, caractère χ pour un sous-groupe Γ d'indice fini de $\mathbf{SL}_2(\mathbf{Z})$, où $k \in \mathbf{Z}$ et $\chi : \Gamma \to \mathbf{C}^*$ est un caractère d'ordre fini (*i.e.* on a $\chi(\gamma_1\gamma_2) = \chi(\gamma_1)\chi(\gamma_2)$ si $\gamma_1, \gamma_2 \in \Gamma$ et il existe $d \in \mathbf{N}$ tel que $\chi(\gamma)^d = 1$ si $\gamma \in \Gamma$).

En liaison avec l'arithmétique, les sous-groupes d'indice fini de $\mathbf{SL}_2(\mathbf{Z})$ que l'on rencontre le plus fréquemment sont $\Gamma_0(\mathrm{N})$, $\Gamma_1(\mathrm{N})$ et $\Gamma(\mathrm{N})$, où $\mathrm{N} \geqslant 1$ est un entier, et

$$\Gamma_0(\mathrm{N}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbf{Z}),\ c \equiv 0 \bmod \mathrm{N} \right\};$$

$$\Gamma_1(\mathrm{N}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbf{Z}),\ c \equiv 0 \bmod \mathrm{N}, a \equiv d \equiv 1 \bmod \mathrm{N} \right\};$$

$$\Gamma(\mathrm{N}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbf{Z}),\ b \equiv c \equiv 0 \bmod \mathrm{N},\ a \equiv d \equiv 1 \bmod \mathrm{N} \right\}.$$

Une manière naturelle de fabriquer un caractère d'ordre fini de $\Gamma_0(\mathrm{N})$ est de partir d'un caractère de Dirichlet modulo N (*i.e.* une application $\widetilde{\chi} : (\mathbf{Z}/\mathrm{N}\mathbf{Z})^* \to \mathbf{C}^*$ vérifiant $\widetilde{\chi}(ab) = \widetilde{\chi}(a) = \widetilde{\chi}(b)$ que l'on peut aussi voir comme une application multiplicative de $\mathbf{Z}$ dans $\mathbf{C}^*$, périodique de période N, telle que $\widetilde{\chi}(n) = 0$ si n n'est pas premier à N), et de poser $\chi(\gamma) = \widetilde{\chi}(d)$ si $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(\mathrm{N})$.

Lemme III.2.1. *Si Γ est un sous-groupe d'indice fini de $\mathbf{SL}_2(\mathbf{Z})$ et si $\chi : \Gamma \to \mathbf{C}^*$ est un caractère d'ordre fini, alors il existe $\mathrm{N} \in \mathbf{N} - \{0\}$ tel que $\gamma = \left(\begin{smallmatrix} 1 & \mathrm{N} \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma$ et $\chi(\gamma) = 1$.*

Démonstration. Soit $\mathrm{U} = \left\{ \left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right),\ n \in \mathbf{Z} \right\}$; c'est un sous-groupe de $\mathbf{SL}_2(\mathbf{Z})$. Considérons l'action de U sur $\mathrm{X} = \mathbf{SL}_2(\mathbf{Z})/\Gamma$ par translation à gauche. Si $\gamma \in \mathrm{U}$ agit trivialement, il respecte en particulier la classe à gauche de Γ et donc appartient à Γ. L'ensemble X étant fini, il existe un sous-groupe U' d'indice fini dans U qui agit trivialement et $\mathrm{U}_1 = \Gamma \cap \mathrm{U}$ est d'indice fini dans U. En particulier, U_1 est un groupe infini et, l'ensemble des valeurs prises par χ étant fini, le noyau U_2 de la restriction de χ à U_1 est un groupe infini, ce qui permet de conclure.

3. Définition des formes modulaires

Définition III.2.2. Si Γ est un sous-groupe d'indice fini de $\mathbf{SL}_2(\mathbf{Z})$ et si $k \in \mathbf{Z}$, une fonction $f : \mathscr{H} \to \mathbf{C}$ est dite *automorphe de poids k pour* Γ, si elle est $\mathscr{C}^\infty$, et si l'on a :

(i) $f_{|_k}\gamma = f$ quel que soit $\gamma \in \Gamma$;

(ii) f est à croissance lente à l'infini (voir ci-dessous).

Plus généralement, si $\chi : \Gamma \to \mathbf{C}^*$ est un caractère d'ordre fini, une fonction $f : \mathscr{H} \to \mathbf{C}$ est dite *automorphe de poids k et caractère χ pour* Γ, si elle est $\mathscr{C}^\infty$, à croissance lente à l'infini, et si l'on a :

(i′) $f_{|_k}\gamma = \chi(\gamma)f$ quel que soit $\gamma \in \Gamma$.

Une forme automorphe holomorphe est dite *modulaire* et on note $\mathrm{M}_k(\Gamma)$ (resp. $\mathrm{M}_k(\Gamma, \chi)$) le $\mathbf{C}$-espace vectoriel des formes modulaires de poids k (resp. de poids k et caractère χ) pour Γ.

Remarque III.2.3. Si Γ contient $-\mathrm{I} = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, et si $\chi(-\mathrm{I}) \neq (-1)^k$, alors une forme automorphe de poids k, caractère χ pour Γ est identiquement nulle.

La condition « f est à croissance lente à l'infini » s'exprime de la manière suivante : quel que soit $\gamma \in \mathbf{SL}_2(\mathbf{Z})$, quels que soient $a < b \in \mathbf{R}$, il existe $\mathrm{C} \in \mathbf{R}$ tel que l'on ait, au voisinage de $y = +\infty$,

$$\sup_{z \in [a+iy, b+iy]} \left| f_{|_k}\gamma(z) \right| = \mathrm{O}(y^{\mathrm{C}}).$$

La relation d'automorphie $f_{|_k}\gamma = \chi(\gamma)f$ permet de réduire beaucoup le nombre de vérifications à faire : on peut se contenter de prendre γ dans un système de représentants de $\mathbf{SL}_2(\mathbf{Z})/\Gamma$ qui, par hypothèse, est un ensemble fini. La fonction $f_{|_k}\gamma$ est automorphe pour le groupe $\gamma\Gamma\gamma^{-1}$ qui est contenu dans $\mathbf{SL}_2(\mathbf{Z})$ de même indice que Γ, et le lemme III.2.1 montre qu'il existe $\mathrm{N} \in \mathbf{N} - \{0\}$ (dépendant de γ) tel que f soit périodique de période N, ce qui permet de ne considérer que le cas $a = 0$ et $b = \mathrm{N}$ au lieu de a et b quelconques. En résumé, il n'y a qu'un nombre fini de vérifications à faire.

4. Développement de Fourier des formes automorphes et modulaires

Soit f une forme automorphe de poids k, caractère χ pour Γ, où Γ est d'indice fini dans $\mathbf{SL}_2(\mathbf{Z})$ et $\chi : \Gamma \to \mathbf{C}^*$ est d'ordre fini. D'après la

discussion précédente, il existe $\mathrm{N} \in \mathbf{N} - \{0\}$ tel que l'on ait $f(z+\mathrm{N}) = f(z)$ si $z \in \mathscr{H}$, et comme f est $\mathscr{C}^\infty$, elle est somme de sa série de Fourier

$$f(z) = \sum_{n \in \mathbf{Z}} a_n(f,y) e^{2i\pi nx/\mathrm{N}},$$

avec
$$a_n(f,y) = \frac{1}{\mathrm{N}} \int_0^{\mathrm{N}} f(x+iy) e^{-2i\pi nx/\mathrm{N}} dx.$$

Comme on a supposé que f est à croissance lente à l'infini, la formule donnant $a_n(f,y)$ montre qu'il existe $\mathrm{C} \in \mathbf{R}$ tel que $a_n(f,y) = \mathrm{O}(y^{\mathrm{C}})$ au voisinage de $y = +\infty$.

Maintenant, si f est modulaire, la fonction $\widetilde{f}(q_{\mathrm{N}}) = f(\mathrm{N}\frac{\log q_{\mathrm{N}}}{2i\pi})$ est bien définie (la périodicité de f montre que $\widetilde{f}(q_{\mathrm{N}})$ ne dépend pas de la détermination de $\log q_{\mathrm{N}}$) et est une fonction holomorphe sur $\mathrm{D}^* = \{0 < |q_{\mathrm{N}}| < 1\}$. On peut donc écrire f sous la forme

$$\sum_{n \in \mathbf{Z}} a_n(f) q_{\mathrm{N}}^n, \quad \text{avec} \quad q_{\mathrm{N}} = e^{2i\pi z/\mathrm{N}}.$$

Si on identifie les coefficients de Fourier, on obtient la relation $a_n(f,y) = a_n(f) e^{-2\pi ny/\mathrm{N}}$ et la croissance lente des coefficients de Fourier montre que l'on a $a_n(f) = 0$ si $n < 0$. La fonction $\widetilde{f}$ se prolonge donc en une fonction holomorphe sur $\mathrm{D} = \{|q_{\mathrm{N}}| < 1\}$ en posant $\widetilde{f}(0) = a_0(f)$. On dit que f est *holomorphe en* $i\infty$ et on pose $f(i\infty) = a_0(f)$.

On dit que f est *parabolique* ou *cuspidale* si $f_{|_k}\gamma$ est nulle en $i\infty$ quel que soit $\gamma \in \mathbf{SL}_2(\mathbf{Z})$, et on note $\mathrm{S}_k(\Gamma) \subset \mathrm{M}_k(\Gamma)$ (resp. $\mathrm{S}_k(\Gamma,\chi) \subset \mathrm{M}_k(\Gamma,\chi)$) le sous-$\mathbf{C}$-espace vectoriel des formes paraboliques[(1)].

Les espaces vectoriels $\mathrm{S}_k(\Gamma,\chi)$ et $\mathrm{M}_k(\Gamma,\chi)$ sont de dimension finie et on dispose de formules générales donnant la dimension de ces espaces vectoriels.

[(1)]Une forme parabolique est une « Spitzenform » en allemand, ce qui explique le S et « Spitz » veut dire « pointe » et se traduit par « cusp » en anglais...

III.3. $\mathbf{SL_2(Z)}$

1. Les éléments S et T

Soient S et T les éléments de $\mathbf{SL}_2(\mathbf{Z})$ définis par

$$\mathrm{T} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad \mathrm{S} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

On a $\mathrm{S}^2 = -\mathrm{I}$. Si $z \in \mathscr{H}$, on a $\mathrm{S}z = -\frac{1}{z}$ et $\mathrm{T}^n z = z + n$ si $n \in \mathbf{Z}$.

Théorème III.3.1. *Le groupe* $\mathbf{SL}_2(\mathbf{Z})$ *est engendré par* S *et* T.

Démonstration. On a

$$\mathrm{S}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -c & -d \\ a & b \end{pmatrix}$$

$$\mathrm{T}^n \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a + nc & b + nd \\ c & d \end{pmatrix} \quad \text{si } n \in \mathbf{Z}.$$

Nous allons montrer, par récurrence sur $|c|$, que $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbf{SL}_2(\mathbf{Z})$ appartient au sous-groupe engendré par S et T. Si $|c| = 0$, on a $a = d = \pm 1$ et γ est de la forme T^b ou $-\mathrm{T}^{-b} = \mathrm{S}^2\mathrm{T}^{-b}$. Si $|c| > |a|$, on peut appliquer l'hypothèse de récurrence à $\mathrm{S}\gamma$, et si $|c| \leqslant |a|$, on choisit $n \in \mathbf{Z}$ tel que $|a + nc| \leqslant \frac{1}{2}|c|$, et on applique l'hypothèse de récurrence à $\mathrm{ST}^n\gamma$.

Corollaire III.3.2. *Une fonction* $f : \mathscr{H} \to \mathbf{C}$*, de classe* $\mathscr{C}^\infty$*, à croissance lente à l'infini, est automorphe de poids* k *pour* $\mathbf{SL}_2(\mathbf{Z})$*, si et seulement si*

$$f(z+1) = f(z) \quad \text{et} \quad f(-1/z) = z^k f(z).$$

Exercice III.3.3. Soit G le sous-groupe de $\mathbf{SL}_2(\mathbf{Q})$ engendré par $\Gamma_0(4)$ et $\mathrm{S}_2 = \left(\begin{smallmatrix} 0 & -1/2 \\ 2 & 0 \end{smallmatrix}\right)$.

(i) Montrer que $\Gamma_0(4)$ est d'indice 2 dans G.

(ii) Montrer que G est engendré par S_2 et T.

2. Domaine fondamental pour l'action de $\mathbf{SL}_2(\mathbf{Z})$

Rappelons que, si G est un groupe agissant sur un ensemble X, un *domaine fondamental* Δ pour l'action de G sur X est un système de

représentants de G\X ; autrement dit Δ est un sous-ensemble de X vérifiant la condition suivante :

quel que soit $x \in \mathrm{X}$, il existe un *unique* élément y de Δ tel qu'il existe $g \in \mathrm{G}$ vérifiant $x = gy$.

Théorème III.3.4. *L'ensemble Δ constitué de la réunion de l'ouvert*

$$\{z \in \mathscr{H},\ |z| > 1 \text{ et } -1/2 < \mathrm{Re}(z) < 1/2\},$$

de la demi-droite

$$\{z \in \mathscr{H},\ \mathrm{Re}(z) = 1/2 \text{ et } \mathrm{Im}(z) \geqslant 1\}$$

et de l'arc de cercle

$$\{z \in \mathscr{H},\ |z| = 1 \text{ et } 0 \leqslant \mathrm{Re}(z) \leqslant 1/2\},$$

est un domaine fondamental pour l'action de $\mathbf{SL}_2(\mathbf{Z})$ sur $\mathscr{H}$.

Démonstration. Soit $z_0 \in \mathscr{H}$. Si $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, on a $\mathrm{Im}(\gamma z_0) = \frac{\mathrm{Im}(z_0)}{|cz_0+d|^2}$ et, comme $|cz_0 + d|$ tend vers $+\infty$ quand c ou d tend vers l'infini, la fonction $\gamma \mapsto \mathrm{Im}(\gamma z_0)$ ne prend qu'un nombre fini de valeurs $\geqslant \mathrm{Im}(z_0)$; elle atteint donc son maximum pour un certain γ_0. Maintenant, il existe $n \in \mathbf{Z}$ (unique) tel que $-1/2 < \mathrm{Re}(\gamma_0 z_0) + n \leqslant 1/2$; posons $\gamma_1 = \mathrm{T}^n\gamma_0$. On a alors

$$\mathrm{Im}(\gamma_1 z_0) = \mathrm{Im}(\gamma_0 z_0) \geqslant \mathrm{Im}(\mathrm{S}\gamma_1 z_0) = \frac{\mathrm{Im}(\gamma_1 z_0)}{|\gamma_1 z_0|^2},$$

et donc $|\gamma_1 z_0| \geqslant 1$. Finalement, soit $\gamma = \gamma_1$ (resp. $\gamma = \mathrm{S}\gamma_1$) si $|\gamma_1 z_0| > 1$ ou si $|\gamma_1 z_0| = 1$ et $\mathrm{Re}(\gamma_1 z_0) \geqslant 0$ (resp. si $|\gamma_1 z_0| = 1$ et $\mathrm{Re}(\gamma_1 z_0) \leqslant 0$), de telle sorte que $\gamma z_0 \in \Delta$, ce qui prouve que Δ contient un domaine fondamental.

Pour terminer la démonstration, il reste à vérifier que si z_1 et z_2 sont deux éléments de Δ tels qu'il existe $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbf{SL}_2(\mathbf{Z})$ tel que $z_1 = \gamma z_2$, alors $z_1 = z_2$. Par symétrie, on peut supposer que $\mathrm{Im}(z_2) \geqslant \mathrm{Im}(z_1) = \frac{\mathrm{Im}(z_2)}{|cz_2+d|^2}$, ce qui implique que $|cz_2 + d| \leqslant 1$. Comme $\mathrm{Im}(z) \geqslant \frac{\sqrt{3}}{2}$ si $z \in \Delta$, l'inégalité étant stricte sauf si $z = \rho = \frac{1}{2} + i\frac{\sqrt{3}}{2}$, on en déduit l'inégalité $|c| \leqslant 1$. Au signe près, il suffit de traiter les cas $c = 0$ et $c = 1$. Dans le cas $c = 0$, γ agit par translation par un entier et comme Δ ne contient pas deux éléments dont la différence des parties réelles est un entier non nul, on doit avoir $\gamma = \pm\mathrm{I}$ et

$z_2 = z_1$. Dans le cas $c = 1$ et $d = 0$, on a $z_2 = -1/z_1$ et, z_1, z_2 étant de module $\geqslant 1$, cela implique $|z_1| = |z_2| = 1$ et $z_1 = z_2 = i$ car on a supposé $z_1, z_2 \in \Delta$. Finalement, si $|d| \geqslant 1$, on a $|d + \mathrm{Re}(z_2)| \geqslant 1/2$ avec inégalité stricte si $\mathrm{Re}(z_2) \neq 1/2$ ou si $d \neq -1$; on doit donc avoir $d = -1$, $z_2 = \rho$, et $z_1 = a - \frac{1}{\rho - 1} = a + \rho$; comme $z_1 \in \Delta$, cela implique $a = 0$ et donc $z_1 = \rho = z_2$.

3. Le produit scalaire de Petersson

Si f et g sont deux formes automorphes de poids k pour $\mathbf{SL}_2(\mathbf{Z})$, la forme

$$\frac{i}{2}\, \overline{f} g\, y^{k-2} dz \wedge d\overline{z}$$

est invariante sous l'action de $\mathbf{SL}_2(\mathbf{Z})$ comme le montre un petit calcul utilisant la proposition III.1.7 et la formule (III.1.1). L'intégrale

$$\langle f, g \rangle = \int_{\mathbf{SL}_2(\mathbf{Z}) \backslash \mathscr{H}} \frac{i}{2}\, \overline{f} g\, y^{2k-2} dz \wedge d\overline{z}$$

ne dépend donc, si elle converge, pas du domaine fondamental de $\mathscr{H}$ sous l'action de $\mathbf{SL}_2(\mathbf{Z})$ que l'on choisit pour effectuer le calcul. En particulier, on peut choisir le domaine Δ construit ci-dessus.

On note simplement M_k (resp. S_k), si k est un entier, l'espace vectoriel des formes modulaires (resp. paraboliques) de poids k pour $\mathbf{SL}_2(\mathbf{Z})$. Comme $-\mathrm{I} \in \mathbf{SL}_2(\mathbf{Z})$, on a $\mathrm{M}_k = 0$ si k est impair.

Si $f \in \mathrm{S}_k$, on a $f = \mathrm{O}(e^{-2\pi y})$ au voisinage de $y = +\infty$ comme le montre l'existence du développement de Fourier. Ceci permet de montrer que, si f et g sont deux formes *paraboliques* de poids k, alors $\langle f, g \rangle$ est bien défini et la forme $(f, g) \mapsto \langle f, g \rangle$ est une forme hermitienne sur S_k. Comme $\langle f, f \rangle = \int_\Delta |f|^2 y^{k-2}\, dx\, dy$, cette forme hermitienne est définie positive et définit donc un produit scalaire sur S_k appelé *produit scalaire de Petersson.*

4. Séries d'Eisenstein holomorphes

On utilise la notation $\sum'_{(m,n)}$ pour indiquer que l'on somme sur tous les couples d'entiers relatifs $(m, n) \neq (0, 0)$.

Proposition III.3.5. *Si $k \geqslant 3$ est un entier pair, et si $z \in \mathscr{H}$, la série*

$$\mathrm{E}_k(z) = \frac{1}{2}\sum_{(m,n)}{}' \frac{1}{(mz+n)^k}$$

converge normalement et la fonction E_k appartient à M_k.

Démonstration. La partie imaginaire de $mz+n$ est my et donc $|mz+n| \geqslant |m|y$ et, d'autre part, le trinôme $\mathrm{X}^2|z|^2+2nx\mathrm{X}+n^2$ admet $n^2y^2/|z|^2$ comme minimum ; on a donc

$$|mz+n| \geqslant \inf(y, y/|z|)\sup(|m|,|n|).$$

Si $\mathrm{N} \geqslant 1$, il y a $(2\mathrm{N}+1)^2-(2\mathrm{N}-1)^2 = 8\mathrm{N}$ couples (m,n) vérifiant $\sup(|m|,|n|) = \mathrm{N}$, et on peut majorer (en module) la série par

$$\frac{1}{2}\inf\big(y, \frac{y}{|z|}\big)^{-k}\sum_{(m,n)}{}' \frac{1}{\sup(|m|,|n|)^k} = \frac{1}{2}\inf\big(y, \frac{y}{|z|}\big)^{-k}\sum_{\mathrm{N}=1}^{+\infty}\frac{8\mathrm{N}}{\mathrm{N}^k},$$

ce qui permet de prouver que la série définissant $\mathrm{E}_k(z)$ converge normalement sur tout compact de $\mathscr{H}$ et donc définit une fonction holomorphe sur $\mathscr{H}$. Finalement, si $\gamma = \left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in \mathbf{SL}_2(\mathbf{Z})$, on a

$$\begin{aligned}(\mathrm{E}_k)_{|_k}\gamma(z) &= (cz+d)^{-k}\mathrm{E}_k\Big(\frac{az+b}{cz+d}\Big)\\ &= \frac{1}{2}(cz+d)^{-k}\sum_{(m,n)}{}' \frac{1}{\big(m\frac{az+b}{cz+d}+n\big)^k}\\ &= \frac{1}{2}\sum_{(m,n)}{}' \frac{1}{\big((am+cn)z+(bm+dn)\big)^k},\end{aligned}$$

et, l'application

$$(m,n) \longmapsto (am+cn, bm+dn)$$

étant une bijection de $\mathbf{Z}^2-\{(0,0)\}$, on obtient la formule $(\mathrm{E}_k)_{|_k}\gamma(z) = \mathrm{E}_k(z)$ qui permet de conclure.

Exercice III.3.6. Soit k un entier pair $\geqslant 3$. Soit Γ_∞ le sous-groupe $\Big\{\left(\begin{smallmatrix} 1 & n \\ 0 & 1\end{smallmatrix}\right),\ n \in \mathbf{Z}\Big\}$ de $\mathbf{SL}_2(\mathbf{Z})$.

(i) Montrer que, si $m \in \mathbf{N}$ et $\gamma \in \mathbf{SL}_2(\mathbf{Z})$, la quantité

$$e^{2i\pi m\gamma z}\big(d\gamma z/dz\big)^{k/2}$$

ne dépend que de l'image de γ dans $\Gamma_\infty\backslash\mathbf{SL}_2(\mathbf{Z})$.

(ii) Montrer que la série de Poincaré

$$\mathrm{P}_{k,m}(z) = \frac{1}{2} \sum_{\gamma \in \Gamma_\infty \backslash \mathbf{SL}_2(\mathbf{Z})} e^{2i\pi m\gamma z} \Big(\frac{d\gamma z}{dz} \Big)^{k/2}$$

converge absolument si $z \in \mathscr{H}$ et que $\mathrm{P}_{k,m} \in \mathrm{M}_k$.

(iii) Montrer que l'application $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mapsto (c,d)$ induit une bijection de $\Gamma_\infty \backslash \mathbf{SL}_2(\mathbf{Z})$ sur l'ensemble des couples (c,d), avec c et d premiers entre eux ; établir l'identité

$$\mathrm{P}_{k,0} = \frac{1}{\zeta(2k)} \mathrm{E}_k.$$

(iv) Montrer que $\mathrm{P}_{k,m} \in \mathrm{S}_k$ si $m \geqslant 1$ et, si $f = \sum_{n=1}^{+\infty} a_n(f) q^n \in \mathrm{S}_k$, calculer le produit scalaire de Petersson $\langle \mathrm{P}_{k,m}, f \rangle$.

(v) Montrer que les $(\mathrm{P}_{k,m})_{m \geqslant 1}$ forment une famille génératrice de S_k.

Nous allons maintenant déterminer le développement de Fourier des séries d'Eisenstein. On note $q = q_1 = e^{2i\pi z}$ le « paramètre local en $i\infty$ », Γ la fonction Gamma d'Euler, ζ la fonction zêta de Riemann et, si N est un entier et $s \in \mathbf{C}$, $\sigma_s(\mathrm{N})$ la somme des puissances s-ième des diviseurs $\geqslant 1$ de N.

Proposition III.3.7. *Si $k \geqslant 3$ est un entier pair, le développement de Fourier de E_k est donné par la formule*

$$\frac{\Gamma(k)}{(-2i\pi)^k} \mathrm{E}_k(z) = \frac{\Gamma(k)}{(-2i\pi)^k} \zeta(k) + \sum_{\mathrm{N}=1}^{+\infty} \sigma_{k-1}(\mathrm{N}) q^{\mathrm{N}}.$$

Démonstration. On remarque que les couples (m,n) et $(-m,-n)$ contribuent de la même manière, ce qui permet de ne garder que les couples $(0,n)$ avec $n \geqslant 1$ et (m,n) avec $m \geqslant 1$. On obtient alors

$$\mathrm{E}_k(z) = \zeta(k) + \sum_{m \geqslant 1} \mathrm{A}_k(mz),$$

où l'on a posé

$$\mathrm{A}_k(z) = \sum_{n \in \mathbf{Z}} \frac{1}{(z+n)^k}.$$

On peut calculer la transformée de Fourier de la fonction $x \mapsto 1/(x+iy)^k$ grâce à la formule des résidus ; on obtient

$$\int_{-\infty}^{+\infty} \frac{e^{-2i\pi tx}}{(x+iy)^k} dx = \begin{cases} 0 & \text{si } t \leqslant 0, \\ \dfrac{(-2i\pi)^k}{(k-1)!} t^{k-1} e^{-2\pi ty} & \text{si } t > 0 ; \end{cases}$$

et la formule de Poisson (valable par exemple si $\varphi \in \mathrm{L}^1(\mathbf{R})$ est deux fois dérivable et $\varphi'' \in \mathrm{L}^1(\mathbf{R})$),

$$\sum_{n\in\mathbf{Z}} \varphi(x+n) = \sum_{n\in\mathbf{Z}} \Big(\int_{-\infty}^{+\infty} \varphi(x) e^{-2i\pi nx} dx \Big) e^{2i\pi nx},$$

nous donne

$$\frac{\Gamma(k)}{(-2i\pi)^k} \mathrm{A}_k(z) = \sum_{n\geqslant 1} n^{k-1} q^n.$$

On obtient donc

$$\frac{\Gamma(k)}{(-2i\pi)^k} \mathrm{E}_k(z) = \frac{\Gamma(k)}{(-2i\pi)^k} \zeta(k) + \sum_{m\geqslant 1} \sum_{n\geqslant 1} n^{k-1} q^{mn},$$

et le résultat s'en déduit en posant $\mathrm{N} = mn$ et en faisant le changement de sommation

$$\sum_{m\geqslant 1} \sum_{n\geqslant 1} = \sum_{\mathrm{N}\geqslant 1} \sum_{n|\mathrm{N}\, n\geqslant 1} .$$

III.4. Prolongement analytique de séries d'Eisenstein

Dans le développement de Fourier des séries d'Eisenstein holomorphes, tous les coefficients sont trivialement rationnels à part le terme constant ; on peut utiliser cette remarque pour fabriquer une démonstration du théorème d'Euler selon lequel $\zeta(k)/\pi^k \in \mathbf{Q}$, si k est un entier pair $\geqslant 2$. Plus généralement, on peut utiliser le fait que $\zeta(2k)$ apparaît dans le terme constant du développement de Fourier des séries d'Eisenstein holomorphes pour étudier les propriétés arithmétiques de $\zeta(2k)$ (divisibilités, congruences…). Nous allons en faire de même avec la fonction ζ elle-même en la faisant apparaître dans le terme constant du développement de Fourier d'une série d'Eisenstein non holomorphe.

1. Séries d'Eisenstein non holomorphes

Proposition III.4.1. *Si $z = x + iy \in \mathscr{H}$, la série*

$$\mathrm{E}(s,z) = \frac{1}{2} \cdot \frac{\Gamma(s)}{\pi^s} \sum_{(m,n)}{}' \frac{y^s}{|mz+n|^{2s}}$$

converge normalement si $\mathrm{Re}(s) > 2$ *et la fonction* $z \mapsto \mathrm{E}(s,z)$ *est une forme automorphe de poids* 0 *pour* $\mathbf{SL}_2(\mathbf{Z})$

Démonstration. La convergence de la série résulte de la majoration

$$|mz+n| \geqslant \inf(y, y/|z|) \sup(|m|,|n|)$$

obtenue au cours de la démonstration de la proposition III.3.5, qui permet aussi de majorer $|\mathrm{E}(s,z)|$ par $4\zeta(2\mathrm{Re}(s))y^s$ au voisinage de $y = +\infty$, ce qui prouve que $\mathrm{E}(s,z)$ est à croissance lente à l'infini. Par ailleurs, la fonction $z \mapsto \mathrm{E}(s,z)$ est $\mathscr{C}^\infty$ car la série des dérivées (de n'importe quel ordre) converge mieux que la série elle-même. Finalement, si $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbf{SL}_2(\mathbf{Z})$, on a

$$\begin{aligned}
\frac{2\cdot\pi^s}{\Gamma(s)}\mathrm{E}(s,\gamma z) &= \sum_{(m,n)}{}' \frac{\mathrm{Im}(\gamma z)^s}{|m\gamma z+n|^{2s}} \\
&= \sum_{(m,n)}{}' \frac{\frac{y^s}{|cz+d|^{2s}}}{\left|m\frac{az+b}{cz+d}+n\right|^{2s}} \\
&= \sum_{(m,n)}{}' \frac{y^s}{|(am+cn)z+(bm+dn)|^{2s}}
\end{aligned}$$

et le résultat suit de ce que $(m,n) \mapsto (am+cn, bm+dn)$ est une bijection de $\mathbf{Z}^2$.

2. La transformée de Fourier de $x \mapsto 1/(x^2+y^2)^s$

La fonction $\mathrm{E}(s,z)$ est somme de sa série de Fourier

$$\mathrm{E}(s,z) = \sum_{n\in\mathbf{Z}} a_n(s,y) e^{2i\pi nx}.$$

Le calcul des coefficients de Fourier de $\mathrm{E}(s,z)$ est parallèle à celui des séries d'Eisenstein holomorphes ; la seule différence est que l'on rencontre la transformée de Fourier de la fonction $x \mapsto 1/|x+iy|^{2s}$

qui ne peut plus s'évaluer au moyen de la méthode des résidus comme dans le cas $s = 0$.

Si $u > 0$ et $s \in \mathbf{C}$, soit

$$\mathrm{K}_s(u) = \int_0^{+\infty} e^{-u(t+t^{-1})} t^s \frac{dt}{t} = \int_1^{+\infty} e^{-u(t+t^{-1})} (t^s + t^{-s}) \frac{dt}{t}.$$

Lemme III.4.2

(i) *Si $u > 0$ est fixé, la fonction $s \mapsto \mathrm{K}_s(u)$ est holomorphe sur $\mathbf{C}$ et vérifie l'équation fonctionnelle $\mathrm{K}_s(u) = \mathrm{K}_{-s}(u)$;*

(ii) *Si $\mathrm{M} \subset \mathbf{C}$ est compact, il existe $\mathrm{C}(\mathrm{M}) > 0$ tel que l'on ait $|\mathrm{K}_s(u)| \leqslant \mathrm{C}(\mathrm{M}) e^{-u}$ pour tout $s \in \mathrm{M}$ et tout $u > 0$.*

(iii) *On a $\mathrm{K}_s(u) \sim \sqrt{\pi} u^{-1/2} e^{-2u}$ au voisinage de $+\infty$; en particulier, si s est fixé, la fonction $u \mapsto \mathrm{K}_s(u)$ n'est pas identiquement nulle.*

Démonstration. Les deux premiers points sont des exercices ; démontrons le troisième. Effectuant le changement de variable $t = 1 + v/\sqrt{u}$, on peut écrire $\sqrt{u} e^{2u} \mathrm{K}_s(u)$ sous la forme $\int_0^{+\infty} f_s(v) dv$, avec

$$f_s(v) = e^{-v^2/(1+v/\sqrt{u})} \Big(\Big(1 + \frac{v}{\sqrt{u}}\Big)^{s-1} + \Big(1 + \frac{v}{\sqrt{u}}\Big)^{-s-1} \Big).$$

On conclut en utilisant le théorème de convergence dominée.

Lemme III.4.3. *Si $\mathrm{Re}(s) > 1/2$, la transformée de Fourier de la fonction $x \mapsto \dfrac{\Gamma(s)}{\pi^s} \dfrac{1}{(x^2+y^2)^s}$ est donnée par la formule*

$$\frac{\Gamma(s)}{\pi^s} \int_{-\infty}^{+\infty} \frac{e^{-2i\pi tx}}{|x+iy|^{2s}} dx = \begin{cases} \dfrac{\Gamma(s-1/2)}{\pi^{s-1/2}} \cdot y^{1-2s} & \text{si } t = 0 \,; \\ \Big|\dfrac{t}{y}\Big|^{s-1/2} \mathrm{K}_{s-1/2}(\pi|t|y) & \text{si } t \neq 0. \end{cases}$$

Démonstration. On a

$$\int_{-\infty}^{+\infty} \frac{dx}{|x+iy|^{2s}} = 2 \int_0^{+\infty} \frac{dx}{|x+iy|^{2s}}$$

et les changements de variables $x = y\sqrt{u}$ et $u = v/(1-v)$ nous donnent

$$\int_{-\infty}^{+\infty} \frac{dx}{|x+iy|^{2s}} = y^{1-2s} \int_0^{+\infty} \frac{du}{u^{1/2}(u+1)^s}$$
$$= y^{1-2s} \int_0^1 (1-v)^{s-3/2} v^{-1/2},$$

et cette dernière intégrale s'évalue grâce à la formule d'Euler

$$\int_0^1 v^{a-1}(1-v)^{b-1} dv = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

On en tire le résultat pour $t = 0$ en utilisant la formule $\Gamma(1/2) = \sqrt{\pi}$.

Si $t \neq 0$, on utilise la formule

$$\frac{\Gamma(s)}{\pi^s} \frac{1}{|x+iy|^{2s}} = \int_0^{+\infty} e^{-\pi u(x^2+y^2)} u^s \frac{du}{u}.$$

L'intégrale à calculer devient, en utilisant Fubini,

$$\int_0^{+\infty} \Big(\int_{-\infty}^{+\infty} e^{-\pi u x^2} e^{-2i\pi t x} dx \Big) e^{-\pi u y^2} u^s \frac{du}{u}.$$

Comme, d'autre part,

$$\int_{-\infty}^{+\infty} e^{-\pi u x^2} e^{-2i\pi t x} dx = \frac{1}{\sqrt{u}} e^{-\pi t^2/u},$$

on obtient le résultat en faisant le changement de variable $u = |t/y|v$.

3. Développement de Fourier des séries d'Eisenstein

Si $s \in \mathbf{C}$ et $n \in \mathbf{Z} - \{0\}$, soit $\sigma_s(n) = \sum_{c|n,\ c\geqslant 1} \big|n/c^2\big|^s$. La bijection $c \mapsto \big|n/c\big|$ montre que l'on a $\sigma_s(n) = \sigma_{-s}(n)$ quels que soient $s \in \mathbf{C}$ et $n \in \mathbf{Z} - \{0\}$.

Soit aussi $\xi(s) = \dfrac{\Gamma(s/2)}{\pi^{s/2}} \zeta(s)$.

Proposition III.4.4. *Si* $\mathrm{Re}(s) > 2$, *les coefficients de Fourier de* $\mathrm{E}(s,z)$ *sont donnés par*

$$a_n(s,y) = \begin{cases} \xi(2s) y^s + \xi(2s-1) y^{1-s} & \text{si } n = 0\,; \\ y^{1/2} \sigma_{s-1/2}(n) \mathrm{K}_{s-1/2}(\pi |n| y) & \text{si } n \neq 0. \end{cases}$$

Démonstration. On remarque que les couples (c,d) et $(-c,-d)$ contribuent de la même manière, ce qui permet de ne garder que les couples $(0,d)$ avec $d\geqslant 1$ et (c,d) avec $c\geqslant 1$. On obtient alors

$$\mathrm{E}(s,z)=\xi(2s)y^s+\sum_{c=1}^{+\infty}y^s\mathrm{A}_s(cz),$$

où l'on a posé

$$\mathrm{A}_s(z)=\frac{\Gamma(s)}{\pi^s}\sum_{n\in\mathbf{Z}}\frac{1}{|z+n|^{2s}}.$$

La formule de Poisson et le calcul de la transformée de Fourier de $\dfrac{\Gamma(s)}{\pi^s}\dfrac{1}{|x+iy|^{2s}}$ effectué ci-dessus nous donnent

$$\mathrm{A}_s(z)=\frac{\Gamma(s-1/2)}{\pi^{s-1/2}}y^{1-2s}+\sum_{n\in\mathbf{Z}-\{0\}}\Big(\frac{|n|}{y}\Big)^{s-1/2}\mathrm{K}_{s-1/2}(\pi|n|y)e^{2i\pi nx}.$$

On obtient donc

$$\sum_{c=1}^{+\infty}y^s\mathrm{A}_s(cz)=\frac{\Gamma(s-1/2)}{\pi^{s-1/2}}y^s\sum_{c\geqslant 1}(cy)^{1-2s}$$
$$+y^s\Big(\sum_{c\geqslant 1}\sum_{n\in\mathbf{Z}-\{0\}}\Big(\frac{|n|}{cy}\Big)^{s-1/2}\mathrm{K}_{s-1/2}(\pi|n|cy)e^{2i\pi cnx}\Big).$$

Le résultat s'en déduit en posant $m=cn$ (et donc $n=m/c$) et en faisant le changement de sommation

$$\sum_{c\geqslant 1}\sum_{n\in\mathbf{Z}-\{0\}}=\sum_{m\in\mathbf{Z}-\{0\}}\sum_{c|m,\ c\geqslant 1}.$$

4. Prolongement analytique des séries d'Eisenstein

Les fonctions $\mathrm{K}_s(u)$ et $\sigma_s(n)$ possédant un prolongement holomorphe à $\mathbf{C}$ tout entier et une équation fonctionnelle reliant s et $-s$, on en déduit, pour tout $n\neq 0$, l'existence d'un prolongement holomorphe de $a_n(s,y)$ vérifiant l'équation fonctionnelle $a_n(s,y)=a_n(1-s,y)$. Nous allons voir que ces propriétés s'étendent au terme constant.

Théorème III.4.5

(i) *$\xi(s)$ admet un prolongement méromorphe à $\mathbf{C}$ tout entier, holomorphe en dehors de pôles simples en $s = 0$ et $s = 1$.*

(ii) *Si $z \in \mathscr{H}$, alors $\mathrm{E}(s, z)$ admet un prolongement méromorphe à $\mathbf{C}$ tout entier, holomorphe en dehors de pôles simples en $s = 0$ et $s = 1$. De plus, quels que soient $z \in \mathscr{H}$ et $s \neq 0, 1$, on a $\mathrm{E}(s, \frac{az+b}{cz+d}) = \mathrm{E}(s, z)$ si $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbf{SL}_2(\mathbf{Z})$.*

(iii) *Les fonctions $\xi(s)$ et $\mathrm{E}(s, z)$ vérifient les équations fonctionnelles*

$$\mathrm{E}(s, z) = \mathrm{E}(1 - s, z) \quad \text{et} \quad \xi(s) = \xi(1 - s).$$

Démonstration. Le lemme III.4.2 permet de montrer que, si $z \in \mathscr{H}$ est fixé, la série $\mathrm{R}(s, z) = \sum_{n \neq 0} a_n(s, y) e^{2i\pi nx}$ converge absolument sur tout compact ; elle définit donc une fonction holomorphe de s sur $\mathbf{C}$ tout entier. Par ailleurs, si $\mathrm{Re}(s) > 2$, on a $\mathrm{E}(s, z) = \mathrm{E}(s, -1/z)$ et donc, la fonction

$$\begin{aligned} \xi(2s)\Big(y^s - \frac{y^s}{(x^2+y^2)^s}\Big) + \xi(2s-1)\Big(y^{1-s} - \frac{y^{1-s}}{(x^2+y^2)^{1-s}}\Big) & \\ = a_0(s, y) - a_0\Big(s, \frac{y}{x^2+y^2}\Big) & \\ = -\mathrm{R}(s, z) + \mathrm{R}\Big(s, -\frac{1}{z}\Big) & \end{aligned}$$

admet un prolongement holomorphe à $\mathbf{C}$ tout entier, quel que soit $x + iy \in \mathscr{H}$. Prenant deux valeurs de z, on obtient un système permettant d'exprimer $\xi(2s)$ et $\xi(2s - 1)$ comme quotient de fonctions holomorphes. On peut en particulier, si $a < 1$, prendre $z = z_a$ ou z_{a^2}, où $z_a = \sqrt{a^{-1} - 1} + i$ est tel que $\mathrm{Im}(z_a) = 1$ et $\mathrm{Im}(-1/z_a) = a$. On obtient alors, en notant $\mathrm{F}_a(s)$ la fonction holomorphe $-\mathrm{R}(s, z_a) + \mathrm{R}(s, -1/z_a)$,

$$\begin{aligned} \xi(2s) &= \frac{(1 - a^{2-2s})\mathrm{F}_a(s) - (1 - a^{1-s})\mathrm{F}_{a^2}(s)}{(1 - a^s)(1 - a^{2-2s}) - (1 - a^{2s})(1 - a^{1-s})} \\ &= \frac{(1 + a^{1-s})\mathrm{F}_a(s) - \mathrm{F}_{a^2}(s)}{a^{2s} - a^s + a^{1-s} - a}. \end{aligned}$$

Les seules valeurs de s annulant, quel que soit $0 < a < 1$, le dénominateur de cette fraction sont $s = 0$ et $s = 1/2$ et le zéro en $s = 0$ ou $s = 1/2$ est un zéro simple. On en déduit le (i).

Comme $a_0(s, y) = \xi(2s)y^s + \xi(2s-1)y^{1-s}$, le (i) implique que $a_0(s, y)$ (et donc aussi $\mathrm{E}(s, z) = a_0(s, y) + \mathrm{R}(s, z)$) admet un prolongement méromorphe à $\mathbf{C}$ tout entier, holomorphe en dehors de pôles simples en $s = 0, 1/2, 1$. D'autre part, on a $\mathrm{E}(s, \frac{az+b}{cz+d}) = \mathrm{E}(s, z)$ si $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbf{SL}_2(\mathbf{Z})$, si $z \in \mathscr{H}$ et si $\mathrm{Re}(s) > 2$. Par prolongement analytique, cette formule reste vraie pour tout $s \in \mathbf{C}$ pour lequel elle a un sens (*i.e.* au moins si $s \neq 0, 1/2, 1$).

Maintenant, soit $\mathrm{F}(s, z) = \mathrm{E}(s, z) - \mathrm{E}(1-s, z)$. Du fait de l'égalité $a_n(s, y) = a_n(1-s, y)$ si $n \neq 0$, on a

$$\mathrm{F}(s, z) = a_0(s, y) - a_0(1-s, y).$$

Comme, par ailleurs, $\mathrm{F}(s, \frac{az+b}{cz+d}) = \mathrm{F}(s, z)$ si $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbf{SL}_2(\mathbf{Z})$, on en déduit, vu la forme de $a_0(s, y)$, que $\mathrm{F}(s, z) = 0$ (autrement dit $\mathrm{E}(s, z) = \mathrm{E}(1-s, z)$, et $\mathrm{E}(s, z)$ n'a pas de pôle en $s = 1/2$, car ce pôle doit être d'ordre pair au vu de la symétrie $s \mapsto 1-s$) et $a_0(s, y) - a_0(1-s, y) = 0$, quels que soient $y > 0$ et $s \in \mathbf{C}$ pour lequel tout est bien défini. L'identification des coefficients de y^s et y^{1-s} dans $a_0(s, y)$ et $a_0(1-s, y)$ nous fournit les équations fonctionnelles $\xi(2s) = \xi(1-2s)$ et $\xi(2s-1) = \xi(2-2s)$ qui se résument à $\xi(s) = \xi(1-s)$. Ceci termine la démonstration du théorème.

5. Non annulation sur la droite $\mathrm{Re}(s) = 1$

Soit $\Gamma_\infty = \left\{ \pm \left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right),\ n \in \mathbf{Z} \right\}$. C'est un sous-groupe de $\mathbf{SL}_2(\mathbf{Z})$ et on a $\mathrm{Im}(\gamma z) = \mathrm{Im}(z)$ si $z \in \mathscr{H}$ et $\gamma \in \Gamma_\infty$.

Lemme III.4.6. *Si* $\mathrm{Re}(s) > 2$ *et* $z \in \mathscr{H}$*, alors*

$$\mathrm{E}(s, z) = \xi(2s) \sum_{g \in \Gamma_\infty \backslash \mathbf{SL}_2(\mathbf{Z})} \mathrm{Im}(gz)^s.$$

Démonstration. Commençons par constater que la somme dans le second membre a bien un sens puisque $\mathrm{Im}(\gamma g z) = \mathrm{Im}(gz)$ si $\gamma \in \Gamma_\infty$. Maintenant, si $(c, d) \neq (0, 0)$, on peut écrire (c, d) de manière unique

sous la forme (ec', ed'), où $e \geqslant 1$ et $(c', d') = 1$ (e est donc le p.g.c.d. de c et d). Ceci nous permet d'écrire $\mathrm{E}(s, z)$ sous la forme

$$\begin{aligned}\mathrm{E}(s,z) &= \frac{1}{2}\frac{\Gamma(s)}{\pi^{2s}}\Big(\sum_{e\geqslant 1}\frac{1}{e^{2s}}\Big)\Big(\sum_{(c',d')=1}\frac{y^s}{|c'z+d'|^{2s}}\Big)\\ &= \xi(2s)\Big(y^s + \sum_{(c',d')=1\ c'\geqslant 1}\frac{y^s}{|c'z+d'|^{2s}}\Big).\end{aligned}$$

Le lemme de Bézout permet de montrer que, si $c' \geqslant 1$ et d' est premier à c', il existe $a', b' \in \mathbf{Z}$ uniques tels que $a'd' - b'c' = 1$ et $-1/2 < a'/c' \leqslant 1/2$. D'autre part, tout élément de $\mathbf{SL}_2(\mathbf{Z})$ n'appartenant pas à Γ_∞ peut s'écrire de manière unique sous la forme $\gamma\left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)$, avec $\gamma \in \Gamma_\infty$, $c' \geqslant 1$ et $-1/2 < a'/c' \leqslant 1/2$. On en déduit la formule

$$\sum_{g\in\Gamma_\infty\backslash\mathbf{SL}_2(\mathbf{Z})}\mathrm{Im}(gz)^s = y^s + \sum_{(c',d')=1\ c'\geqslant 1}\frac{y^s}{|c'z+d'|^{2s}},$$

ce qui permet de conclure.

Proposition III.4.7. *Soit $f : \mathscr{H} \to \mathbf{C}$ une fonction faiblement modulaire telle qu'il existe $\delta > 0$ et $\mathrm{C} > 0$ tels que l'on ait $|f(x+iy)| \leqslant \mathrm{C}e^{-\delta y}$ si $y \geqslant 1$, et soit $f(x+iy) = \sum_{n\in\mathbf{Z}} b_n(y)e^{2i\pi x}$ le développement de Fourier de f. Alors, quel que soit $s \neq 0, 1/2$, on a*

$$\int_{\mathrm{Y}}\overline{f(z)}\mathrm{E}(s,z)\frac{dx\,dy}{y^2} = \xi(2s)\int_0^{+\infty} y^{s-2}\overline{b_0(y)}dy.$$

Démonstration. Commençons par constater que la condition de décroissance à l'infini que l'on a imposée à f implique que l'intégrale

$$\int_{\mathrm{Y}}\overline{f(z)}\mathrm{E}(s,z)\frac{dx\,dy}{y^2}$$

converge (du moins si $\mathrm{E}(s, z)$ est défini, c'est-à-dire si $s \neq 0, 1/2$) et que la fonction

$$s \longmapsto \int_{\mathrm{Y}}\overline{f(z)}\mathrm{E}(s,z)\frac{dx\,dy}{y^2}$$

est une fonction holomorphe sur $\mathbf{C} - \{0, 1/2\}$. Comme par ailleurs, on a $b_0(y) = \int_0^1 f(x+iy)dx$, la fonction $b_0(y)$ est à décroissance rapide

en l'infini et en 0 grâce à l'équation fonctionnelle $f(-1/z) = f(z)$. Ceci implique que

$$s \longmapsto \int_0^{+\infty} y^{s-2}\overline{b_0(y)}dy$$

est une fonction holomorphe sur $\mathbf{C}$ tout entier. Pour vérifier l'égalité ci-dessus, il suffit donc de la vérifier pour $\mathrm{Re}(s) > 2$, le cas général s'en déduisant par prolongement analytique. Or on a

$$\int_0^{+\infty} y^{s-2}\overline{b_0(y)}dy = \int_{\mathrm{B}} \overline{f(z)}y^{s-2}dxdy,$$

où $\mathrm{B} =]0,1] \times]0,+\infty[\subset \mathscr{H}$. Maintenant, B est un domaine fondamental de $\mathscr{H}$ modulo l'action de Γ_∞ et, si on utilise l'invariance de f et de la mesure $\dfrac{dx\,dy}{y^2}$ sous l'action de $\mathbf{SL}_2(\mathbf{Z})$ et le lemme III.4.6, on obtient

$$\begin{aligned}\int_{\Gamma_\infty\backslash\mathscr{H}} \overline{f(z)}y^{s-2}dxdy &= \sum_{\gamma\in\Gamma_\infty\backslash\mathbf{SL}_2(\mathbf{Z})} \int_{\mathbf{SL}_2(\mathbf{Z})\backslash\mathscr{H}} \overline{f(\gamma z)}\mathrm{Im}(\gamma z)^s \frac{dx\,dy}{y^2} \\ &= \int_{\mathbf{SL}_2(\mathbf{Z})\backslash\mathscr{H}} \sum_{\gamma\in\Gamma_\infty\backslash\mathbf{SL}_2(\mathbf{Z})} \overline{f(z)}\mathrm{Im}(\gamma z)^s \frac{dx\,dy}{y^2} \\ &= \int_{\mathbf{SL}_2(\mathbf{Z})\backslash\mathscr{H}} \overline{f(z)}\frac{\mathrm{E}(s,z)}{\xi(2s)}\frac{dx\,dy}{y^2},\end{aligned}$$

ce qui permet de conclure.

Théorème III.4.8. *La fonction ζ ne s'annule pas sur la droite* $\mathrm{Re}(s)=1$

Démonstration

Si s_0 vérifie $\zeta(s_0) = 0$ et $\mathrm{Re}(s_0) = 1$, on a aussi $\zeta(s_0-1) = 0$ à cause de l'équation fonctionnelle $\zeta(s) = \zeta(1-s)$ et de ce que $\zeta(\overline{s}) = \overline{\zeta(s)}$ (on a $s_0 - 1 = \overline{1-s_0}$). On en déduit le fait que $a_0(s_0/2, y)$ est nul quel que soit y et donc que $\mathrm{E}(s_0/2, z) = \mathrm{O}(e^{-(1-\varepsilon)2\pi y})$ au voisinage de $+\infty$, quel que soit $\varepsilon > 0$. On peut donc calculer le produit scalaire de Petersson de $\mathrm{E}(s_0/2, z)$ avec $\mathrm{E}(s, z)$ pour tout $s \in \mathbf{C} - \{0, 1\}$ et la proposition III.4.7 montre que ce produit scalaire est nul. On en déduit, en utilisant ce résultat pour $s = s_0/2$, que $\mathrm{E}(s_0/2, z)$ est identiquement nulle, ce qui est absurde car la fonction $\mathrm{K}_{s_0/2}$ n'est pas identiquement nulle d'après le lemme III.4.2 et le coefficient de Fourier $a_1(s_0/2, y)$ n'est donc pas identiquement nul.

Remarque III.4.9. Ce théorème est une des étapes importantes de la démonstration d'Hadamard du théorème des nombres premiers.

III.5. Opérateurs de Hecke

1. Généralités

Soient $\Gamma \subset \mathrm{G}$ deux groupes. Si $x \in \mathrm{G}$, on note $x\Gamma$ (resp. Γx) le sous-ensemble de G des éléments de la forme $x\gamma$ (resp. γx), avec $\gamma \in \Gamma$.

Si X est un sous-ensemble de G, on note encore X la fonction caractéristique de X ; on a donc $\mathrm{X}(x) = 1$ (resp. $\mathrm{X}(x) = 0$) si $x \in \mathrm{X}$ (resp. si $x \notin \mathrm{X}$).

Si K est un corps, soit $\mathrm{K}[\Gamma\backslash\mathrm{G}/\Gamma]$ l'espace vectoriel des fonctions $\varphi : \mathrm{G} \to \mathrm{K}$ *bi-invariantes* (*i.e.* $\varphi(\gamma x) = \varphi(x)$ et $\varphi(x\gamma) = \varphi(x)$ quels que soient $x \in \mathrm{G}$ et $\gamma \in \Gamma$), et *à support fini* (*i.e.* il existe un ensemble fini I tel que $\varphi = \sum_{i\in\mathrm{I}} \lambda_i \cdot \Gamma x_i$). Remarquons que, si les x_i ont des images distinctes dans $\Gamma\backslash\mathrm{G}$, les λ_i sont uniquement déterminés.

Proposition III.5.1

(i) *Si* $\varphi = \sum_{i\in\mathrm{I}} \lambda_i \cdot \Gamma x_i$ *et* $\varphi' = \sum_{j\in\mathrm{J}} \mu_j \cdot \Gamma y_j$ *sont deux éléments de* $[\Gamma\backslash\mathrm{G}/\Gamma]$, *alors*

$$\varphi \star \varphi' = \sum_{(i,j)\in\mathrm{I}\times\mathrm{J}} \lambda_i\mu_j \cdot \Gamma x_i y_j$$

est un élément de $\mathrm{K}[\Gamma\backslash\mathrm{G}/\Gamma]$ *qui ne dépend pas du choix des* x_i, $i \in \mathrm{I}$ *et des* y_j, $j \in \mathrm{J}$.

(ii) $\mathrm{K}[\Gamma\backslash\mathrm{G}/\Gamma]$, *muni de la structure de* K*-espace vectoriel et de la multiplication* $\star$ *ainsi définie, est une algèbre associative admettant* Γ *comme unité.*

Démonstration. On peut supposer que les x_i ont des images distinctes dans $\Gamma\backslash\mathrm{G}$ (si x_i et $x_{i'}$ ont même image dans $\Gamma\backslash\mathrm{G}$, on a $\Gamma x_i = \Gamma x_{i'}$ en tant qu'ensemble et donc aussi en tant que fonction et on peut regrouper ces deux termes). D'autre part, si $\gamma \in \Gamma$, on a $\Gamma x_i(x\gamma^{-1}) =$

$\Gamma x_i\gamma(x)$; ceci permet, en utilisant l'identité

$$\sum_{i\in \mathrm{I}} \lambda_i \cdot \Gamma x_i\gamma(x) = \sum_{i\in \mathrm{I}} \lambda_i \cdot \Gamma x_i(x\gamma^{-1}) \\ = \varphi(x\gamma^{-1}) = \varphi(x) = \sum_{i\in \mathrm{I}} \lambda_i \cdot \Gamma x_i(x),$$

de montrer qu'il existe une permutation $\sigma : \mathrm{I} \to \mathrm{I}$ telle que, si $i \in \mathrm{I}$, alors $\lambda_i = \lambda_{\sigma(i)}$ et il existe $\gamma_i \in \Gamma$ tel que $x_i\gamma = \gamma_i x_{\sigma(i)}$.

Passons à la démonstration du (i). Le choix des x_i n'influe, de manière évidente, pas sur le résultat, et on peut changer y_j en $\gamma_j y_j$, avec $\gamma_j \in \Gamma$. D'après la discussion précédente, si $j \in \mathrm{J}$, il existe une permutation $\sigma_j : \mathrm{I} \to \mathrm{I}$ telle que, si $i \in \mathrm{I}$, alors $\lambda_i = \lambda_{\sigma_j(i)}$ et il existe $\gamma_{j,i} \in \Gamma$ tel que $x_i\gamma_j = \gamma_{j,i}x_{\sigma_j(i)}$. Notons $(i', j') \in \mathrm{I} \times \mathrm{J}$ le couple $(\sigma_j(i), j)$; l'application $(i, j) \mapsto (i', j')$ est une bijection de $\mathrm{I} \times \mathrm{J}$. On a alors

$$\sum_{(i,j)\in \mathrm{I}\times \mathrm{J}} \lambda_i\mu_j \cdot \Gamma x_i\gamma_j y_j = \sum_{(i,j)\in \mathrm{I}\times \mathrm{J}} \lambda_i\mu_j \cdot \Gamma x_{\sigma_j(i)} y_j = \sum_{(i,j)\in \mathrm{I}\times \mathrm{J}} \lambda_{i'}\mu_{j'} \cdot \Gamma x_{i'} y_{j'},$$

ce qui permet de montrer que le choix des y_j n'influe pas sur le résultat.

Comme $\varphi \star \varphi'(\gamma x) = \varphi \star \varphi'(x)$ de manière évidente, il ne reste plus qu'à démontrer que l'on a aussi $\varphi \star \varphi'(x\gamma) = \varphi \star \varphi'(x)$. Comme précédemment, il existe une permutation $\tau : \mathrm{J} \to \mathrm{J}$ telle que, si $j \in \mathrm{J}$, alors $\mu_j = \mu_{\tau(j)}$ et il existe $\gamma_j \in \Gamma$ tel que $y_j\gamma^{-1} = \gamma_j y_{\tau(j)}$. On a alors

$$\varphi\star\varphi'(x\gamma) = \sum_{(i,j)\in \mathrm{I}\times \mathrm{J}} \lambda_i\mu_j \cdot \Gamma x_i y_j\gamma^{-1}(x) = \sum_{(i,j)\in \mathrm{I}\times \mathrm{J}} \lambda_i\mu_j \cdot \Gamma x_i\gamma_j y_{\tau(j)}(x),$$

et comme $\mu_j = \mu_{\tau(j)}$, on peut faire le changement $j' = \tau(j)$, pour réécrire cette dernière formule sous la forme

$$\sum_{(i,j)\in \mathrm{I}\times \mathrm{J}} \lambda_i\mu_j \cdot \Gamma x_i\gamma_j y_j,$$

et utiliser l'invariance par rapport au choix des y_j pour conclure.

Le (ii) est plus ou moins évident ; par exemple, l'associativité découle immédiatement de l'associativité de la multiplication dans G.

2. Opérateurs de Hecke

Nous allons appliquer les résultats précédents à $\mathrm{G} = \mathbf{GL}_2(\mathbf{Q})^+$, sous groupe de $\mathbf{GL}_2(\mathbf{Q})$ des matrices de déterminant > 0, et à $\Gamma = \mathbf{SL}_2(\mathbf{Z})$.

Lemme III.5.2. *Si $g \in \mathrm{G}$ est à coefficients entiers, il existe $\gamma \in \Gamma$ tel que $\gamma g = \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)$, avec $a \geqslant 1$. De plus, a et d sont complètement déterminés par g, ainsi que la classe de b modulo d.*

Démonstration. Si $g = \left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)$, on peut trouver $u, v \in \mathbf{Z}$ premiers entre eux, tels que $ua' + vc' = 0$. On peut alors, grâce au théorème de Bézout, trouver $x, y \in \mathbf{Z}$ tels que $xv - uy = 1$ et prendre pour $\gamma = \pm\left(\begin{smallmatrix} x & y \\ u & v \end{smallmatrix}\right)$, le signe étant choisi de telle sorte que a soit positif.

Maintenant, si $\gamma_1 g = \left(\begin{smallmatrix} a_1 & b_1 \\ 0 & d_1 \end{smallmatrix}\right)$ et $\gamma_2 g = \left(\begin{smallmatrix} a_2 & b_2 \\ 0 & d_2 \end{smallmatrix}\right)$, alors

$$\gamma_1\gamma_2^{-1} = (\gamma_1 g)(\gamma_2 g)^{-1} = \begin{pmatrix} a_1 & b_1 \\ 0 & d_1 \end{pmatrix}\begin{pmatrix} a_2 & b_2 \\ 0 & d_2 \end{pmatrix}^{-1} = \begin{pmatrix} a_1/a_2 & \frac{a_2b_1 - a_1b_2}{a_2d_2} \\ 0 & d_1/d_2 \end{pmatrix}$$

est à coefficients entiers. Ceci implique que a_2 divise a_1, d_2 divise d_1, et, comme $a_1d_1 = a_2d_2 = \det g$ et a_1, a_2, d_1, d_2 sont $\geqslant 1$, on obtient les égalités $a_1 = a_2$ et $d_1 = d_2$. Finalement, $b_1 - b_2$ est divisible par d_2, ce qui termine la démonstration.

Lemme III.5.3. *Si $n \geqslant 1$, soient T_n l'ensemble des éléments de $\mathbf{M}_2(\mathbf{Z})$ de déterminant n et $\mathrm{R}_n = \Gamma\left(\begin{smallmatrix} n & 0 \\ 0 & n \end{smallmatrix}\right)$. Alors T_n et R_n appartiennent à $\mathbf{Q}[\Gamma\backslash\mathrm{G}/\Gamma]$.*

Démonstration. Le résultat est immédiat pour R_n car $\left(\begin{smallmatrix} n & 0 \\ 0 & n \end{smallmatrix}\right)$ commute à tout. La bi-invariance de T_n est une conséquence de la multiplicativité du déterminant et la formule suivante, conséquence immédiate du lemme précédent, montre que T_n est à support fini, ce qui permet de conclure

$$\mathrm{T}_n = \sum_{\substack{a \geqslant 1 \\ ad = n \\ b \bmod d}} \Gamma\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}.$$

On note $\mathbb{T}_{\mathbf{Q}}$ la sous-$\mathbf{Q}$-algèbre de $\mathbf{Q}[\Gamma\backslash\mathrm{G}/\Gamma]$ engendrée par les T_n et les R_n, pour $n \geqslant 1$.

Théorème III.5.4

(i) *On a*

a) $\mathrm{R}_n\mathrm{R}_m = \mathrm{R}_{nm}$ *quels que soient* $n, m \geqslant 1$ *;*

b) $\mathrm{R}_n\mathrm{T}_m = \mathrm{T}_m\mathrm{R}_n$ *quels que soient* $n, m \geqslant 1$ *;*

c) $\mathrm{T}_n\mathrm{T}_m = \mathrm{T}_{nm}$ *quels que soient* $n, m \geqslant 1$ *premiers entre eux ;*

d) $\mathrm{T}_{p^r}\mathrm{T}_p = \mathrm{T}_{p^{r+1}} + p\mathrm{R}_p\mathrm{T}_{p^{r-1}}$ *si* p *est un nombre premier et* $r \geqslant 1$.

(ii) $\mathbb{T}_{\mathbf{Q}}$ *est une algèbre commutative*

Démonstration. Commençons par montrer comment déduire le (ii) du (i). Le d) permet de montrer, par récurrence sur r, que T_{p^r} est un polynôme à coefficients entiers en T_p et R_p ; comme R_p et T_p commutent d'après le b), cela permet de montrer que T_{p^r} et T_{p^s} commutent si $r, s \in \mathbf{N}$. Les autres commutations sont immédiates en vertu des a), b) et c).

Passons à la vérification du (i). Les points a) et b) découlent simplement de ce que $\left(\begin{smallmatrix} n & 0 \\ 0 & n \end{smallmatrix}\right)$ commute à tout élément de G. Passons au c). On a

$$\mathrm{T}_n\mathrm{T}_m = \sum_{\substack{a\geqslant 1 \\ ad=n \\ b \bmod d}} \sum_{\substack{a'\geqslant 1 \\ a'd'=m \\ b' \bmod d'}} \Gamma\begin{pmatrix} aa' & ab'+bd' \\ 0 & dd' \end{pmatrix}.$$

Si m et n sont premiers entre eux, il en est, a fortiori, de même de a et d', et on est ramené à prouver que, si a et d' sont premiers entre eux, si b (resp. b') parcourt un système de représentants modulo d (resp. d'), alors $ab'+bd'$ parcourt un système de représentants modulo dd'. Comme il y a le bon nombre d'éléments, il suffit de vérifier que l'application $(b, b') \mapsto ab'+bd'$ modulo dd' est injective. En regardant modulo d', et en utilisant le fait que a est inversible modulo d', on voit que, si $ab'_1+b_1d' = ab'_2+b_2d' \bmod. dd'$, alors $b'_1-b'_2$ est divisible par d' et donc $b'_1 = b'_2$, puis que b_1-b_2 est divisible par d et donc que $b_1 = b_2$, ce qui permet de conclure.

Reste le point d). On a $\mathrm{T}_{p^r} = \sum_{i=0}^{r} \sum_{b \bmod p^i} \Gamma\left(\begin{smallmatrix} p^{r-i} & b \\ 0 & p^i \end{smallmatrix}\right)$ et, en particulier,

$$\mathrm{T}_p = \Gamma\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} + \sum_{c \bmod p} \Gamma\begin{pmatrix} 1 & c \\ 0 & p \end{pmatrix};$$

on obtient donc

$$\mathrm{T}_{p^r}\mathrm{T}_p = \sum_{i=0}^{r} \sum_{b \bmod p^i} \Gamma \begin{pmatrix} p^{r+1-i} & b \\ 0 & p^i \end{pmatrix} + \sum_{i=0}^{r} \sum_{b \bmod p^i} \sum_{c \bmod p} \Gamma \begin{pmatrix} p^{r-i} & pb + p^{r-i}c \\ 0 & p^{i+1} \end{pmatrix}.$$

Si on regroupe dans cette somme le premier terme et le morceau du second terme correspondant à $i = r$, on retrouve l'opérateur $\mathrm{T}_{p^{r+1}}$ car $pb + c$ décrit un système de représentants modulo p^{r+1} si b (resp. c) parcourt un système de représentants modulo p^r (resp. p).

On peut alors mettre R_p en facteur dans ce qui reste, ce qui permet de le mettre sous la forme

$$\mathrm{R}_p \sum_{c \bmod p} \Big(\sum_{i=0}^{r-1} \sum_{b \bmod p^i} \Gamma \begin{pmatrix} p^{r-i-1} & b + p^{r-i-1}c \\ 0 & p^i \end{pmatrix} \Big).$$

Comme $b + p^{r-i-1}c$ parcourt un système de représentants modulo p^i quand b parcourt un système de représentants modulo p^i, on voit que la somme entre parenthèses est égale à $\mathrm{T}_{p^{r-1}}$ pour tout c, et donc que la quantité ci-dessus est aussi égale à $p\mathrm{R}_p\mathrm{T}_{p^{r-1}}$. Ceci permet de conclure.

3. Action des opérateurs de Hecke sur les formes modulaires

On définit l'action *en poids* k de $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{G}$ sur une fonction $f : \mathscr{H} \to \mathbf{C}$ par la formule

$$f_{|_k}\gamma(z) = \frac{(\det \gamma)^{k-1}}{(cz+d)^k} f(\gamma z).$$

(La puissance de $\det\gamma$ que l'on introduit est destinée à supprimer les dénominateurs dans l'action des opérateurs de Hecke ; un autre choix naturel consisterait à prendre la puissance $k/2$-ième pour rendre l'action de G unitaire.)

Si $\varphi = \sum_{i\in\mathrm{I}} \lambda_i \cdot \Gamma g_i \in \mathbf{Q}[\Gamma\backslash\mathrm{G}/\Gamma]$ et $f : \mathscr{H} \to \mathbf{C}$ vérifie $f_{|_k}\gamma = f$ pour tout $\gamma \in \Gamma$, on peut définir la fonction $f_{|_k}\varphi$ par la formule

$$f_{|_k}\varphi(z) = \sum_{i\in\mathrm{I}} \lambda_i \cdot f_{|_k} g_i(z),$$

ce qui ne dépend pas du choix des g_i.

Proposition III.5.5

(i) *Si* $f \in \mathrm{M}_k$ (resp. $f \in \mathrm{S}_k$) *et* $\varphi \in \mathbf{Q}[\Gamma\backslash \mathrm{G}/\Gamma]$, *alors* $f_{|_k}\varphi \in \mathrm{M}_k$ (resp. $f_{|_k}\varphi \in \mathrm{S}_k$) *;*

(ii) *Si* $f \in \mathrm{M}_k$ *et* $\varphi_1, \varphi_2 \in \mathbf{Q}[\Gamma\backslash \mathrm{G}/\Gamma]$, *alors*

$$f_{|_k}(\varphi_1+\varphi_2) = f_{|_k}\varphi_1 + f_{|_k}\varphi_2 \quad \text{et} \quad f_{|_k}(\varphi_1 \star \varphi_2) = (f_{|_k}\varphi_1)_{|_k}\varphi_2.$$

Démonstration. L'invariance de $f_{|_k}\varphi$ par Γ et le (ii) se démontrent comme la proposition III.5.1. L'holomorphie de $f_{|_k}\varphi$ étant une évidence, il ne reste plus qu'à considérer le comportement au voisinage de $i\infty$. Si $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{G}$, et si $a/c = a_1/c_1$ avec $a_1, c_1 \in \mathbf{Z}$ premiers entre eux ($(a_1, c_1) = (1, 0)$ si $c = 0$), soient $b_1, d_1 \in \mathbf{Z}$ tels que $\gamma_1 = \left(\begin{smallmatrix} a_1 & b_1 \\ c_1 & d_1 \end{smallmatrix}\right) \in \Gamma$. La matrice $\gamma_1^{-1}\gamma$ est alors de la forme $\left(\begin{smallmatrix} a_2 & b_2 \\ 0 & d_2 \end{smallmatrix}\right)$ et on a

$$f_{|_k}\gamma(z) = f_{|_k}\gamma_1^{-1}\gamma(z) = a_2^{k-1}d_2^{-1}f\Big(\frac{a_2 z + b_2}{d_2}\Big).$$

On en déduit le fait que, si f est à croissance lente à l'infini (resp. est nulle à l'infini), alors $f_{|_k}\gamma$ est à croissance lente à l'infini (resp. est nulle à l'infini), ce qui termine la démonstration.

Lemme III.5.6. *Si* $f \in \mathrm{M}_k$ *et* $n \geqslant 1$, *alors*

$$f_{|_k}\mathrm{R}_n = n^{k-2}f;$$
$$f_{|_k}\mathrm{T}_n = n^{k-1}\sum_{\substack{a\geqslant 1,\ ad=n \\ b \bmod d}} d^{-k} f(\frac{az+b}{d}).$$

Démonstration. La première identité est évidente et la seconde suit du lemme III.5.2.

Proposition III.5.7. *Si* $f \in \mathrm{M}_k$, *si* $n \geqslant 1$ *et si* $m \in \mathbf{N}$, *alors*

$$c_m(f_{|_k}\mathrm{T}_n) = \sum_{\substack{a\geqslant 1 \\ a|(n,m)}} a^{k-1}c_{nm/a^2}(f).$$

Démonstration. On a $\sum_{b \bmod d} e^{2i\pi mb/d} = \begin{cases} d & \text{si } d|m, \\ 0 & \text{sinon.} \end{cases}$. On en déduit la formule

$$\sum_{b \bmod d} d^{-k} f(\frac{az+b}{d}) = d^{1-k} \sum_{\ell=0}^{+\infty} c_{d\ell}(f)\, q^{a\ell}.$$

Si $a \geqslant 1$ divise n, la contribution de $\sum_{b \bmod d} d^{-k} f(\frac{az+b}{d})$ à $c_m(f_{|_k} \mathrm{T}_n)$ est donc

$$\begin{cases} n^{k-1} d^{1-k} c_{dm/a}(f) = a^{k-1} c_{nm/a^2}(f) & \text{si } a \text{ divise } m, \\ 0 & \text{si } a \text{ ne divise pas } m. \end{cases}$$

On en tire le résultat.

La proposition précédente admet plusieurs cas particuliers intéressants

Corollaire III.5.8. *Si $f \in \mathrm{M}_k$, si $n \geqslant 1$, alors*

(i) $c_0(f_{|_k} \mathrm{T}_n) = \sigma_{k-1}(n) c_0(f)$;

(ii) $c_1(f_{|_k} \mathrm{T}_n) = c_n(f)$;

(iii) *Si $f \in \mathrm{M}_k$, si p est un nombre premier et si $m \in \mathbf{N}$, alors*

$$c_m(f_{|_k} \mathrm{T}_p) = \begin{cases} c_{pm}(f) & \text{si } p \text{ ne divise pas } m, \\ c_{pm}(f) + p^{k-1} c_{m/p}(f) & \text{si } p \text{ divise } m. \end{cases}$$

Théorème III.5.9. *Soit $f \in \mathrm{M}_k - \{0\}$ vecteur propre pour tous les opérateurs T_n, $n \geqslant 1$ de valeur propre λ_n. Alors,*

(i) $c_1(f) \neq 0$;

(ii) *si on a normalisé f de telle sorte que $c_1(f) = 1$, alors $\lambda_n = c_n(f)$ pour tout $n \geqslant 1$; de plus les $c_n(f)$ vérifient les relations*

(a) $c_n(f) c_m(f) = c_{nm}(f)$ *si n et m sont premiers entre eux ;*

(b) $c_{p^{r+1}}(f) - c_p(f) c_{p^r}(f) + p^{k-1} c_{p^{r-1}}(f) = 0$ *si p est un nombre premier et $r \geqslant 1$.*

Démonstration. D'après le (i) du corollaire III.5.8, on a $c_n(f) = \lambda_n c_1(f)$ si $n \geqslant 0$. En particulier, si $c_1(f) = 0$, alors f est constante, ce qui est en contradiction avec l'appartenance de f à $\mathrm{M}_k - \{0\}$.

Maintenant, si $c_1(f) = 1$, on a $c_n(f) = \lambda_n$ et le reste du théorème est une traduction du théorème III.5.4.

Corollaire III.5.10 **(Théorème de multiplicité 1).** *Si f et g sont deux formes modulaires de poids k, vecteurs propres des T_n avec la même valeur propre pour tout $n \geqslant 1$, et si f et g sont normalisées ($c_1(f) = c_1(g) = 1$), alors $f = g$.*

Démonstration. Il suffit d'appliquer le (i) du théorème précédent à $f - g$.

Proposition III.5.11. *La série d'Eisenstein $\mathrm{G}_{2k}(z) = \frac{\Gamma(2k)}{(-2i\pi)^{2k}}\mathrm{E}_{2k}(z)$ est une forme propre normalisée pour tous les T_n, et on a $(\mathrm{G}_{2k})_{|_{2k}}\mathrm{T}_n = \sigma_{2k-1}(n)\mathrm{G}_{2k}$.*

Démonstration. On a $\mathrm{G}_{2k}(z) = \frac{\Gamma(2k)}{(-2i\pi)^{2k}}\zeta(2k) + \sum_{m=1}^{+\infty}\sigma_{2k-1}(m)q^m$. En particulier, $c_1(\mathrm{G}_{2k}) = 1$ et $c_m(\mathrm{G}_{2k}) = \sigma_{2k-1}(m)$ si $m \geqslant 1$. Pour démontrer la proposition, il suffit donc de prouver que G_{2k} est vecteur propre de T_p pour la valeur propre $\sigma_{2k-1}(p)$, ce qui se ramène à vérifier que, si p est un nombre premier, on a

$$(1+p^s)\sum_{d|m} d^s = \begin{cases} \sum_{d|pm} d^s & \text{si } p \text{ ne divise pas } m, \\ \sum_{d|pm} d^s + p^s \sum_{d|\frac{m}{p}} d^s & \text{si } p \text{ divise } m. \end{cases}$$

Ceci ne pose aucune difficulté.

Théorème III.5.12. *Si $n \geqslant 1$, l'opérateur T_n, agissant sur S_k muni du produit scalaire de Petersson, est hermitien.*

Démonstration. Cela résulte, par un calcul un peu pénible, de l'invariance de la mesure hyperbolique sous l'action de $\mathbf{SL}_2(\mathbf{Z})$. (Pour se simplifier la vie, on peut se contenter de traiter le cas p premier.)

Théorème III.5.13. *S_k et M_k admettent une base de formes propres pour tous les T_n, $n \geqslant 1$; de plus, si on normalise ces formes propres, alors une telle base est unique à permutation près de ses éléments.*

Démonstration. L'unicité d'une telle base (à permutation près et modulo son existence) est une conséquence du théorème de multiplicité 1. De plus, comme $\mathrm{M}_k = \mathbf{C}\mathrm{E}_k \oplus \mathrm{S}_k$, et E_k est une forme propre pour tous les T_n, il suffit de traiter le cas de S_k. Les T_n étant hermitiens, ils sont diagonalisables (dans une base orthonormée), et,

comme ils commutent deux à deux, on peut trouver une base dans laquelle ils agissent tous de façon diagonale, ce qui permet de conclure.

III.6. Fonctions L des formes modulaires

1. La transformée de Mellin

Proposition III.6.1

(i) *Si* $\varphi : \mathbf{R}_+^* \to \mathbf{C}$ *une fonction continue telle qu'il existe* $\mathrm{A} > \mathrm{B}$ *tels que* $\varphi(t) = \mathrm{O}(t^{\mathrm{A}})$ *au voisinage de* $t = 0$ *et* $\varphi(t) = \mathrm{O}(t^{\mathrm{B}})$ *au voisinage de* $+\infty$*, la transformée de Mellin* $\mathrm{Mel}(\varphi, s)$ *de* φ*, définie par*

$$\mathrm{Mel}(\varphi, s) = \int_0^{+\infty} \varphi(t) t^s \frac{dt}{t},$$

est holomorphe sur la bande $-\mathrm{A} < \mathrm{Re}(s) < -\mathrm{B}$ *et bornée sur toute bande de la forme* $a < \mathrm{Re}(s) < b$*, avec* $-\mathrm{A} < a < b < -\mathrm{B}$*.*

(ii) *Si* φ *est de classe* $\mathscr{C}^r$ *et si* $\varphi^{(r)}(t) = \mathrm{O}(t^{\mathrm{A}-r})$ *au voisinage de* $t = 0$ *et* $\varphi^{(r)}(t) = \mathrm{O}(t^{\mathrm{B}-r})$ *au voisinage de* $+\infty$*, alors* $\mathrm{Mel}(\varphi, s)$ *est* $\mathrm{O}(|s|^{-r})$ *sur toute bande de la forme* $a < \mathrm{Re}(s) < b$*, avec* $-\mathrm{A} < a < b < -\mathrm{B}$*.*

(iii) *Si de plus,* $r \geqslant 2$*, on peut retrouver* φ *à partir de sa transformée de Mellin par la formule d'inversion*

$$\varphi(x) = \frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} \mathrm{Mel}(\varphi, s) x^{-s} ds,$$

où c *est un élément quelconque de* $]-\mathrm{A}, -\mathrm{B}[$*.*

Démonstration. Le (i) est immédiat et le (ii) est une conséquence de la formule

$$\mathrm{Mel}(\varphi, s) = \frac{(-1)^r}{s(s+1)\cdots(s+r-1)} \mathrm{Mel}(\varphi^{(r)}, s+r)$$

qui s'obtient par intégration par partie. Finalement, si on note $\psi_c : \mathbf{R} \to \mathbf{C}$ la fonction définie par $\psi_c(x) = \varphi(e^x)e^{cx}$, alors

$$\mathrm{Mel}(\varphi, c+iu) = \widehat{\psi_c}(u) = \int_{-\infty}^{+\infty} \psi_c(x) e^{iux} dx,$$

est la transformée de Fourier de ψ_c. On déduit alors le (iii) de la formule d'inversion de Fourier

$$\psi_c(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \widehat{\psi_c}(u)e^{-iux}du,$$

l'hypothèse $r \geqslant 2$ suffisant à garantir que l'on est dans les conditions d'application de cette formule d'inversion : on obtient

$$\begin{aligned}\frac{1}{2i\pi}\int_{c-i\infty}^{c+i\infty} \mathrm{Mel}(\varphi,s)x^{-s}ds &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \widehat{\psi_c}(u)e^{-c\log x - iu\log x}du \\ &= e^{-c\log x}\psi_c(\log x) = \varphi(x).\end{aligned}$$

Exercice III.6.2. Calculer $\int_{c-i\infty}^{c+i\infty}\Gamma(s)t^{-s}ds$ par la méthode des résidus ; retrouver la formule d'inversion de Mellin pour e^{-x}.

2. Transformée de Mellin des formes modulaires

Si $f = \sum_{m=0}^{+\infty} c_m(f)q^m \in \mathrm{M}_{2k}$, on définit la fonction L de f par la formule

$$\mathrm{L}(f,s) = \sum_{m=1}^{+\infty} c_m(f)m^{-s}, \quad \text{et on pose} \quad \Lambda(f,s) = \frac{\Gamma(s)}{(2\pi)^s}\mathrm{L}(f,s).$$

Exemple III.6.3. La fonction L de la série d'Eisenstein G_{2k} est donnée par la formule

$$\mathrm{L}(\mathrm{G}_{2k},s) = \zeta(s)\zeta(s-2k+1).$$

Démonstration. On a $\mathrm{G}_{2k}(z) = \frac{\Gamma(2k)}{(-2i\pi)^{2k}}\zeta(2k) + \sum_{m=1}^{+\infty}\sigma_{2k-1}(m)q^m$, et donc

$$\begin{aligned}\mathrm{L}(\mathrm{G}_{2k},s) &= \sum_{m=1}^{+\infty}\Big(\sum_{\substack{d\geqslant 1\\ ad=m}} d^{2k-1}\Big)(ad)^{-s} \\ &= \sum_{a=1}^{+\infty}\sum_{d=1}^{+\infty} a^{-s}d^{2k-1-s} = \zeta(s)\zeta(s-2k+1).\end{aligned}$$

Théorème III.6.4

(i) *La série définissant* $\mathrm{L}(f,s)$ *converge absolument pour* $\mathrm{Re}(s) > 2k$ *et définit une fonction holomorphe sur le demi-plan* $\mathrm{Re}(s) > 2k$.

(ii) *La fonction* $\Lambda(f,s)$ *vérifie les propriétés suivantes :*

(a) *elle admet un prolongement méromorphe à* $\mathbf{C}$ *tout entier ;*
(b) *elle satisfait à l'équation fonctionnelle*

$$\Lambda(f,s) = (-1)^{-k}\Lambda(f,2k-s);$$

(c) *elle est holomorphe en dehors de pôles simples en* $s=0$*, de résidu* $-c_0(f)$*, et en* $s=2k$*, de résidu* $(-1)^k c_0(f)$ *;*
(d) *elle tend vers* 0 *à l'infini dans toute bande de la forme* $a<\mathrm{Re}(s)<b$*.*

Démonstration. Le (i) suit de la majoration $c_m(f) = \mathrm{O}(m^{2k-1})$ que nous n'avons pas démontrée...

Pour démontrer le (ii), considérons la fonction $\varphi : \mathbf{R}_+^* \to \mathbf{C}$ définie par $\varphi(t) = f(it) - c_0(f)$. Cette fonction est $\mathscr{C}^\infty$ sur $\mathbf{R}_+^*$, vérifie l'équation fonctionnelle

$$\varphi(t^{-1}) = (-1)^k t^{2k}\varphi(t) + (-1)^k t^{2k} c_0(f) - c_0(f),$$

et est $\mathrm{O}(e^{-2\pi t})$ au voisinage de $+\infty$. Par ailleurs, on a

$$\int_0^{+\infty} e^{-at} t^s \frac{dt}{t} = \Gamma(s) a^{-s},$$

si $a>0$ et $\mathrm{Re}(s)>0$. Ceci permet d'écrire $\Lambda(f,s)$ comme la transformée de Mellin de φ, (du moins si $\mathrm{Re}(s)>k$, de manière que l'on puisse intervertir somme et intégrale). Coupant l'intégrale de 0 à $+\infty$ en une intégrale de 0 à 1 et une de 1 à $+\infty$, et faisant le changement de variable $t \mapsto t^{-1}$ sur le premier morceau nous permet, en utilisant l'équation fonctionnelle ci-dessus, d'obtenir

$$\begin{aligned}\Lambda(f,s) &= \int_1^{+\infty} \varphi(t^{-1}) t^{-s} \frac{dt}{t} + \int_1^{+\infty} \varphi(t) t^s \frac{dt}{t} \\ &= (-1)^k \int_1^{+\infty} \varphi(t) t^{2k-s} \frac{dt}{t} + c_0(f)\Big(\frac{(-1)^k}{s-2k} - \frac{1}{s}\Big) + \int_1^{+\infty} \varphi(t) t^s \frac{dt}{t}.\end{aligned}$$

On en déduit les points (a), (b) et (c). Le point (d) s'obtient en remarquant que l'on a

$$\int_1^{+\infty} \varphi(t) t^s \frac{dt}{t} = -\frac{\varphi(1)}{s} - \frac{1}{s}\int_1^{+\infty} \varphi'(t) t^{s+1} \frac{dt}{t},$$

et que $\int_1^{+\infty} \varphi'(t) t^{s+1} \frac{dt}{t}$ est borné dans toute bande de la forme $a < \mathrm{Re}(s) < b$.

Le théorème précédent admet une réciproque que voici :

Théorème III.6.5. *Soit $(c_m)_{m\in\mathbf{N}}$ une suite de nombres complexes. Si la série de Dirichlet $\mathrm{L}(s)=\sum_{n=1}^{+\infty} c_m m^{-s}$ converge absolument dans un demi-plan de la forme $\mathrm{Re}(s)>\mathrm{A}$, et si la fonction $\Lambda(s)=\frac{\Gamma(s)}{(2\pi)^s}\mathrm{L}(s)$ vérifie les propriétés suivantes :*

(a) *elle admet un prolongement méromorphe à $\mathbf{C}$ tout entier ;*

(b) *elle satisfait à l'équation fonctionnelle*

$$\Lambda(s)=(-1)^{-k}\Lambda(2k-s);$$

(c) *elle est holomorphe en dehors de pôles simples en $s=0$ et en $s=2k$, le résidu en $s=0$ étant $-c_0$;*

(d) *elle tend vers 0 à l'infini dans toute bande de la forme $a<\mathrm{Re}(s)<b$;*

alors $\sum_{m=0}^{+\infty} c_m q^m\in \mathrm{M}_{2k}$.

Démonstration. Soit $f(z)=\sum_{m=0}^{+\infty} c_m q^m$. L'existence d'un demi-plan de convergence pour la série de Dirichlet montre que f est holomorphe sur le demi-plan de Poincaré. Pour prouver que $f\in\mathrm{M}_{2k}$, il suffit, car f est invariante par $z\to z+1$, de prouver que la fonction $g(z)=z^{-2k}f(\frac{-1}{z})-f(z)$ est identiquement nulle sur $\mathscr{H}$. Par prolongement analytique, il suffit de prouver que g est identiquement nulle sur $i\mathbf{R}_+^*$. Pour cela, considérons la fonction $\varphi(t)=f(it)-c_0=\sum_{m=1}^{+\infty} c_m e^{-2\pi m t}$. La fonction $\Lambda(s)$ est alors la transformée de Mellin de φ et la formule d'inversion de Mellin nous donne, si $c>\mathrm{A}$,

$$\varphi(t)-\frac{(-1)^k}{t^{2k}}\varphi(t^{-1})$$
$$=\frac{1}{2i\pi}\Big(\int_{c-i\infty}^{c+i\infty}\Lambda(s)t^{-s}ds-\frac{(-1)^k}{t^{2k}}\int_{c-i\infty}^{c+i\infty}\Lambda(s)t^{s}ds\Big).$$

Par ailleurs, l'équation fonctionnelle $\Lambda(s)=(-1)^k\Lambda(2k-s)$ nous donne

$$\frac{(-1)^k}{t^{2k}}\int_{c-i\infty}^{c+i\infty}\Lambda(s)t^s ds=\int_{c-i\infty}^{c+i\infty}\Lambda(2k-s)t^{s-2k}\,ds$$
$$=\int_{2k-c-i\infty}^{2k-c+i\infty}\Lambda(s)t^{-s}\,ds.$$

Maintenant, la formule des résidus appliquée à l'intégrale de $\Lambda(s)t^{-s}$ sur le rectangle de sommets $c-i\mathrm{T}$, $c+i\mathrm{T}$, $2k-c+i\mathrm{T}$ et $2k-c-i\mathrm{T}$, et un passage à la limite en faisant tendre T vers $+\infty$ [c'est là que l'on utilise l'hypothèse (d)], nous fournissent la formule

$$\begin{aligned}\varphi(t)-\frac{(-1)^k}{t^{2k}}\varphi(t^{-1}) &= \mathrm{Res}_{s=0}\Lambda(s)t^{-s}+\mathrm{Res}_{s=2k}\Lambda(s)t^{-s}\\ &= -c_0+\frac{(-1)^k}{t^{2k}}c_0,\end{aligned}$$

ce qui permet de conclure. (L'équation fonctionnelle

$$\Lambda(s)=(-1)^k\Lambda(2k-s)$$

montrant que le résidu en $s=2k$ de $\Lambda(s)$ est $(-1)^k c_0$ si le résidu de $\Lambda(s)$ en 0 est $-c_0$.)

3. Opérateurs de Hecke et produits eulériens

Théorème III.6.6. *Si $f\in \mathrm{M}_{2k}$ est un vecteur propre pour tous les T_n, $n\geqslant 1$, et est normalisée, alors $\mathrm{L}(f,s)$ admet une factorisation en un produit de facteurs d'Euler*

$$\mathrm{L}(f,s)=\prod_{p\text{ premier}}\frac{1}{1-c_p(f)p^{-s}+p^{2k-1-2s}}.$$

Démonstration. Comme f est vecteur propre des T_n, on a $c_m(f)=\prod_i c_{p_i^{r_i}}(f)$ si $n=\prod_i p_i^{r_i}$ est la décomposition en facteurs premiers de m. On obtient donc

$$\mathrm{L}(f,s)=\prod_{p\text{ premier}}(1+c_p(f)p^{-s}+c_{p^2}(f)p^{-2s}+\cdots).$$

D'autre part, les $c_{p^r}(f)$ satisfont à une relation de récurrence linéaire de degré 2, et, si α_p et β_p sont les deux racines du polynôme

$$\mathrm{X}^2-c_p(f)\mathrm{X}+p^{2k-1}=0,$$

alors $c_{p^r}(f) = \frac{\alpha_p^{p^{r+1}} - \beta_p^{p^{r+1}}}{\alpha_p - \beta_p}$. On obtient donc

$$\begin{aligned} 1 + c_p(f)p^{-s} + c_{p^2}(f)p^{-2s} + \cdots &= \sum_{r=0}^{+\infty} \frac{\alpha_p^{p^{r+1}} - \beta_p^{p^{r+1}}}{\alpha_p - \beta_p} p^{-rs} \\ &= \frac{1}{\alpha_p - \beta_p}\Big(\frac{\alpha_p}{1-\alpha_p p^{-s}} - \frac{\beta_p}{1-\beta_p p^{-s}}\Big) \\ &= \frac{1}{(1-\alpha_p p^{-s})(1-\beta_p p^{-s})} = \frac{1}{1 - c_p(f)p^{-1} + p^{2k-1-2s}}, \end{aligned}$$

ce qui permet de conclure.

Le théorème III.6.6 permet de démontrer une forme forte du « théorème de multiplicité 1 ».

Corollaire III.6.7. *Si $f, g \in \mathrm{M}_{2k}$ vérifient les conditions :*

(a) $c_1(f) = c_1(g) = 1$;

(b) *quel que soit p premier, f (resp. g) est vecteur propre de T_p pour la valeur propre λ_p (resp. μ_p) ;*

(c) $\lambda_p = \mu_p$ *si p est assez grand ;*

alors $f = g$.

Démonstration. Soit I l'ensemble fini de nombres premiers tels que $\lambda_p \neq \mu_p$. La fonction

$$h(s) = \frac{\mathrm{L}(f,s)}{\mathrm{L}(g,s)} = \prod_{p \in \mathrm{I}} \frac{1 - \mu_p p^{-s} + p^{2k-1-2s}}{1 - \lambda_p p^{-s} + p^{2k-1-2s}}$$

vérifie l'équation fonctionnelle $h(2k - s) = h(s)$ d'après le théorème III.6.4. Par ailleurs chacun des termes u_p du produit vérifie l'équation fonctionnelle $u_p(2k-1-s) = u_p(s)$, et donc $h(s)$ est périodique de période 1. Comme h est une série de Dirichlet $\sum_{n=1}^{+\infty} a_n n^{-s}$, cela se traduit par $na_n = a_n$, pour tout n, ce qui implique que h est constante. On en déduit le résultat.

Lemme III.6.8. *Soit $\mathrm{P}_{p,i}$, $1 \leqslant i \leqslant d$, p premier, une famille de polynômes vérifiant les conditions suivantes*

(a) $\mathrm{P}_{p,i}(0) = 1$ *quel que soient i et p ;*

(b) *si $i \neq j$, il existe, quel que soit $\mathrm{B} > 0$, un nombre premier $p \geqslant \mathrm{B}$, tel que $\mathrm{P}_{p,i}$ et $\mathrm{P}_{p,j}$ sont premiers entre eux ;*

(c) *il existe des nombres complexes* λ_i, $1 \leqslant i \leqslant d$ *et des séries entières* $\mathrm{R}_p = a_{p,0} + a_{p,1}\mathrm{X} + \cdots \in \mathbf{C}[\![\mathrm{X}]\!]$ *avec* $a_{p,0} = 1$ *pour* p *assez grand, tels que l'on ait*

$$\sum_{i=1}^{d} \frac{\lambda_i}{\prod_p \mathrm{P}_{p,i}(p^{-s})} = \prod_p \mathrm{R}_p(p^{-s}),$$

alors il existe A *tel que l'on ait* $\mathrm{P}_{p,i}(\mathrm{X}) = \mathrm{R}_p(\mathrm{X})^{-1}$ *quel que soit* $p \geqslant \mathrm{A}$ *et* $1 \leqslant i \leqslant d$.

Démonstration. Soient

$$\frac{1}{\prod_p \mathrm{P}_{p,i}(p^{-s})} = \sum_{n=1}^{+\infty} a_{i,n} n^{-s} \quad \text{et} \quad \prod_p \mathrm{R}_p(p^{-s}) = \sum_{n=1}^{+\infty} b_n n^{-s}.$$

La démonstration se fait par récurrence sur d. Si $\sum_{i=1}^{d} \frac{\lambda_i}{\prod_p \mathrm{P}_{p,i}(p^{-s})} = 0$, on peut faire passer un des termes de l'autre coté et utiliser l'hypothèse de récurrence. Sinon, il existe n tel que $b_n \neq 0$.

Si on fixe p premier ne divisant pas n et tel que $\mathrm{R}_p(0) = 1$, on a alors

$$\sum_{i=1}^{d} \lambda_i a_{i,n} n^{-s} \frac{1}{\mathrm{P}_{p,i}(p^{-s})} = \sum_{i=1}^{d} \lambda_i \sum_{k=0}^{+\infty} a_{i,np^k} (np^k)^{-s} = b_n n^{-s} \mathrm{R}_p(p^{-s}).$$

En particulier, R_p est une fraction rationnelle.

Choisissons alors ℓ assez grand pour que $\mathrm{R}_\ell(0) = 1$ et qu'il existe $j' \neq j$ tels que $\mathrm{P}_{\ell,j}$ et $\mathrm{P}_{\ell,j'}$ soient premiers entre eux. Soit $\mathrm{P}_{\ell,j} = \prod_k \mathrm{Q}_k^{n_k}$ la décomposition de $\mathrm{P}_{\ell,j}$ en facteurs premiers dans $\mathbf{C}[\mathrm{X}]$. Soit α_i (resp. β) le coefficient de $\mathrm{Q}_1^{-n_1}$ dans la décomposition en éléments simples de $\frac{1}{\mathrm{P}_{\ell,i}(\mathrm{X})}$ (resp. $\mathrm{R}_\ell(\mathrm{X})$). Comme $\mathrm{P}_{\ell,j'}$ est premier à $\mathrm{P}_{\ell,j}$, on a $\alpha_{j'} = 0$ et on obtient

$$\sum_{i \neq j'} \frac{\lambda_i \alpha_i}{\prod_{p \neq \ell} \mathrm{P}_{p,i}(p^{-s})} = \beta \prod_{p \neq \ell} \mathrm{R}_p(p^{-s}),$$

ce qui permet d'utiliser l'hypothèse de récurrence pour conclure.

Théorème III.6.9. *Soit* $f \in \mathrm{M}_{2k}$. *Si* $\mathrm{L}(f,s)$ *admet une factorisation du type* $\mathrm{L}(f,s) = \prod_{p \text{ premier}} \mathrm{L}_p(f,s)$, *où* $\mathrm{L}_p(f,s) = 1 + c_{p,1}p^{-s} + c_{p,2}p^{-2s} + \cdots$ *est une série de Dirichlet ne faisant intervenir que des*

puissances de p, alors f est vecteur propre pour tous les T_n, $n \geqslant 1$, et est normalisée.

Démonstration. C'est une conséquence du lemme précédent, en écrivant f sous la forme $\sum_i \lambda_i f_i$, où les f_i sont des vecteurs propres normalisés de tous les T_n.

4. Torsion par un caractère de Dirichlet

Définition III.6.10. Si f est une forme modulaire de poids $2k$ pour $\mathrm{SL}_2(\mathbf{Z})$ et χ est un caractère de Dirichlet de conducteur m, on appelle tordue de f par le caractère χ la fonction f_χ donnée par la formule $f_\chi(z) = \sum_{n=1}^{+\infty} \chi(n) c_n q^n$.

Lemme III.6.11. $\mathrm{G}(\chi) f_{\chi^{-1}}(z) = \sum\limits_{a \bmod m} \chi(a) f(z + \frac{a}{m})$.

Démonstration. Un calcul immédiat nous donne

$$\sum_{a \bmod m} \chi(a) f(z + \frac{a}{m}) = \sum_{n=1}^{+\infty} \mathrm{G}(\chi, n) q^n,$$

ce qui, utilisant le lemme I.5.1, permet de conclure.

Proposition III.6.12. $\mathrm{G}(\chi) f_{\chi^{-1}}(-1/m^2 z) = (mz)^{2k} \chi(-1) \mathrm{G}(\chi^{-1}) f_\chi(z)$.

Démonstration. Soit a premier à m. On part de la formule

$$f\Big(\frac{-1}{m^2 z} + \frac{a}{m}\Big) = f\Big(\frac{maz - 1}{m^2 z}\Big) = (mz)^{2k} \Big(f|_{2k} \begin{pmatrix} ma & -1 \\ m^2 & 0 \end{pmatrix}\Big)(z).$$

D'autre part, on a

$$\begin{pmatrix} ma & -1 \\ m^2 & 0 \end{pmatrix} = \begin{pmatrix} a & -\frac{ay+1}{m} \\ m & -y \end{pmatrix} \begin{pmatrix} m & y \\ 0 & m \end{pmatrix}$$

et si on a choisit $y \in \mathbf{Z}$ tel que $ay + 1 \equiv 0 \bmod m$ (ce qui est possible car $(a, m) = 1$), la matrice $\left(\begin{smallmatrix} a & -\frac{ay+1}{m} \\ m & -y \end{smallmatrix}\right)$ appartient à $\mathrm{SL}_2(\mathbf{Z})$ et donc

$$\Big(f|_{2k} \begin{pmatrix} ma & -1 \\ m^2 & 0 \end{pmatrix}\Big)(z) = \Big(f|_{2k} \begin{pmatrix} m & y \\ 0 & m \end{pmatrix}\Big)(z) = f(z + y/m).$$

On obtient donc, en remarquant que $\chi(a) \neq 0$ implique a premier à m et que $\chi(a) = \chi^{-1}(-y)$ si $ay + 1 \equiv 1 \bmod. m$,

$$\begin{aligned} \mathrm{G}(\chi) f_{\chi^{-1}}(\frac{-1}{m^2 z}) &= \sum_{a \bmod m} \chi(a) f\Big(\frac{-1}{m^2 z} + \frac{a}{m}\Big) \\ &= (mz)^{2k} \sum_{y \bmod m} \chi^{-1}(-y) f(z + y/m) \\ &= (mz)^{2k} \chi(-1) \mathrm{G}(\chi^{-1}) f_\chi(z), \end{aligned}$$

ce qu'il fallait démontrer.

Exercice III.6.13. Montrer que $f_\chi \in \mathrm{S}_{2k}(m^2, \chi^2)$

Si χ est un caractère de Dirichlet de conducteur m, on pose

$$\mathrm{L}(f, \chi, s) = \sum_{n=1}^{+\infty} \frac{\chi(n) c_n}{n^s} \quad \text{et} \quad \Lambda(f, \chi, s) = m^s \frac{\Gamma(s)}{(2\pi)^s} \mathrm{L}(f, \chi, s).$$

Si f est une forme propre pour les opérateurs de Hecke et normalisée, alors

$$\mathrm{L}(f, \chi, s) = \prod_p \frac{1}{1 - c_p(\chi(p) p^{-s}) + p^{2k-1} (\chi(p) p^{-s})^2}.$$

Proposition III.6.14. *$\Lambda(f, \chi, s)$ et donc $\mathrm{L}(f, \chi, s)$ a un prolongement analytique à $\mathbf{C}$ tout entier et vérifie l'équation fonctionnelle*

$$\frac{\Lambda(f, \chi, s)}{\mathrm{G}(\chi)} = (-1)^k \chi(-1) \frac{\Lambda(f, \chi^{-1}, 2k - s)}{\mathrm{G}(\chi^{-1})}.$$

Démonstration. On a $\Lambda(f, \chi, s) = \int_0^{+\infty} f_\chi(i \frac{t}{m}) t^s \frac{dt}{t}$ et le prolongement analytique s'obtient en utilisant la décroissance rapide de f au voisinage de 0 et de $i\infty$. D'autre part,

$$\int_0^{+\infty} f_\chi(i \frac{t}{m}) t^s \frac{dt}{t} = \int_0^{+\infty} \chi(-1) \frac{\mathrm{G}(\chi)}{\mathrm{G}(\chi^{-1})} \frac{1}{(it)^{2k}} f_{\chi^{-1}}(\frac{-1}{itm}) t^s \frac{dt}{t}$$

et l'équation fonctionnelle s'obtient en changeant t en $1/t$ dans l'intégrale.

III.7. Formes de niveau supérieur

Ce § est un résumé, sans démonstration, de la théorie en niveau $\mathrm{N} \geqslant 1$. Les énoncés sont assez similaires à ceux que l'on rencontre en niveau 1. On ne s'intéresse qu'au cas du caractère trivial, mais les énoncés pour $\mathrm{S}_k(\Gamma_0(\mathrm{N}), \chi)$, $\chi : (\mathbf{Z}/\mathrm{N}\mathbf{Z})^* \to \mathbf{C}^*$ sont juste plus visuellement compliqués.

1. Opérateurs de Hecke et d'Atkin-Lehner

Si $(n, \mathrm{N}) = 1$, on définit l'opérateur T_n sur $\mathrm{S}_k(\Gamma_0(\mathrm{N}))$ par la formule

$$f_{|_k}\mathrm{T}_n(z) = n^{k-1} \sum_{\substack{a\geqslant 1,\ ad=n \\ b \bmod d}} d^{-k} f\Big(\frac{az+b}{d}\Big).$$

Si p est un nombre premier divisant N, on définit un opérateur $\mathrm{U}_p : \mathrm{S}_k(\Gamma_0(\mathrm{N})) \to \mathrm{S}_k(\Gamma_0(\mathrm{N}))$ par la formule

$$f_{|_k}\mathrm{U}_p(z) = \frac{1}{p} \sum_{b \bmod p} f\Big(\frac{az+b}{p}\Big)$$

et une involution $w_{\mathrm{N}} : \mathrm{S}_k(\Gamma_0(\mathrm{N})) \to \mathrm{S}_k(\Gamma_0(\mathrm{N}))$ par

$$f_{|_k}w_{\mathrm{N}}(z) = \mathrm{N}^{-k/2} z^{-k} f\Big(\frac{-1}{\mathrm{N}z}\Big).$$

Lemme III.7.1. *Si $d\mathrm{M}|\mathrm{N}$ et si $f \in \mathrm{S}_k(\Gamma_0(\mathrm{M}))$, alors $f_d \in \mathrm{S}_k(\Gamma_0(\mathrm{N}))$, où $f_d : \mathscr{H} \to \mathbf{C}$ est la fonction définie par $f_d(z) = f(dz)$.*

Démonstration. Calcul immédiat.

On dit que f est *nouvelle de niveau* N si $f \in \mathrm{S}_k(\Gamma_0(\mathrm{N}))$ est orthogonale, pour le produit scalaire de Petersson, à toute forme du type g_d, où $g \in \mathrm{S}_k(\Gamma_0(\mathrm{M}))$ et $d\mathrm{M}|\mathrm{N}$, $\mathrm{M} \neq \mathrm{N}$. On note $\mathrm{S}_k^{\mathrm{new}}(\Gamma_0(\mathrm{N}))$ le sous-espace des formes nouvelles de niveau N. On dit que $f \in \mathrm{S}_k(\Gamma_0(\mathrm{N}))$ est *primitive* si

— f est nouvelle de niveau N,
— f est vecteur propre de tous les T_n, $(n, \mathrm{N}) = 1$,
— f est normalisée (*i.e.* $c_1(f) = 1$).

Théorème III.7.2

(i) *La sous-algèbre de* $\mathrm{End}(\mathrm{S}_k(\Gamma_0(\mathrm{N})))$ *engendrée par les* T_n, $(n, \mathrm{N}) = 1$, *les* U_p, *p premier divisant* N, *et* w_{N} *est commutative.*

(ii) *Les* T_n, $(n,\mathrm{N}) = 1$ *agissent de manière hermitienne sur* $\mathrm{S}_k(\Gamma_0(\mathrm{N}))$.

(iii) $\mathrm{S}_k^{\mathrm{new}}(\Gamma_0(\mathrm{N}))$ *admet une base formée de formes primitives. De plus, si* $f = \sum_{m=1}^{+\infty} c_m(f)q^m$ *est un élément de cette base, alors* f *est aussi vecteur propre des* U_p, *pour* $p|\mathrm{N}$, *et de* w_{N}, *et on a*

(a) $f_{|_k}\mathrm{T}_n = c_n(f)f$ *et* $c_m(f)c_n(f) = c_{nm}(f)$ *si* $(n,m) = 1$;

(b) $c_{p^{r+1}}(f) - c_p(f)c_{p^r}(f) + p^{k-1}c_{p^{r-1}}(f) = 0$ *si* p *est un nombre premier ne divisant pas* N *et* $r \geqslant 1$;

(c) $f_{|_k}\mathrm{U}_p = c_p(f)f$ *et* $c_{p^r}(f) = (c_p(f))^r$ *si* p *est un nombre premier divisant* N ;

(d) $f_{|k}w_{\mathrm{N}} = \varepsilon_f f$, *avec* $\varepsilon_f \in \{\pm 1\}$.

(iv) *Si* $f \in \mathrm{S}_k(\Gamma_0(\mathrm{N}))$ *et* $g \in \mathrm{S}_k(\Gamma_0(\mathrm{M}))$ *sont primitives et si* $c_p(f) = c_p(g)$ *en dehors d'un nombre fini de nombres premiers, alors* $\mathrm{N} = \mathrm{M}$ *et* $f = g$.

(v) *Si* $\mathrm{M} \in \mathbf{N}$, *soit* $f_{\mathrm{M},i}$, $i \in \mathrm{I}_{\mathrm{M}}$ *une base de* $\mathrm{S}_k^{\mathrm{new}}(\Gamma_0(\mathrm{M}))$ *formée de formes primitives. Alors on obtient une base de* $\mathrm{S}_k(\Gamma_0(\mathrm{N}))$ *formée de vecteurs propres pour tous les* T_n, $(n,\mathrm{N}) = 1$, *en prenant les fonctions de la forme* $f_{\mathrm{M},i}(dz)$, *où* M *parcourt les diviseurs de* N *et* d *les entiers* $\geqslant 1$ *tels que* $d\mathrm{M}|\mathrm{N}$.

2. Fonctions L

Théorème III.7.3. *Soit* $f = \sum_{m=1}^{+\infty} c_m(f)q^m$ *une forme primitive de niveau* N *et* $\varepsilon_f \in \{\pm 1\}$ *tel que* $f_{|_k}w_{\mathrm{N}} = \varepsilon_f f$. *Soient*

$$\mathrm{L}(f,s) = \sum_{m=1}^{+\infty} c_m(f)m^{-s} \quad \text{et} \quad \Lambda(f,s) = \Gamma(s)\Big(\frac{\sqrt{\mathrm{N}}}{2\pi}\Big)^s \mathrm{L}(f,s).$$

(i) $\mathrm{L}(f,s)$ *admet une factorisation en produit de facteurs d'Euler*

$$\mathrm{L}(f,s) = \prod_{p|\mathrm{N}} \frac{1}{1 - c_p(f)p^{-s}} \prod_{(p,\mathrm{N})=1} \frac{1}{1 - c_p(f)p^{-s} + p^{k-1-2s}}.$$

(ii) *La fonction* $\Lambda(f,s)$ *admet un prolongement analytique à tout le plan complexe, tend vers* 0 *à l'infini dans toute bande verticale* $a < \mathrm{Re}(s) < b$ *et vérifie l'équation fonctionnelle*

$$\Lambda(f,s) = i^{-k}\varepsilon_f \Lambda(f,k-s).$$

Plus généralement, si $(\mathrm{D},\mathrm{N})=1$, et si χ est un caractère de Dirichlet de conducteur D, alors

(i) $$f_\chi=\sum_{m=1}^{+\infty}\chi(m)c_m(f)q^m\in \mathrm{S}_k(\Gamma_0(\mathrm{ND}^2),\chi^2);$$

(ii) $$\chi(-\mathrm{N})\mathrm{G}(\chi)f_{\chi^{-1}}\Big(\frac{-1}{\mathrm{ND}^2z}\Big)=(\mathrm{D}\sqrt{\mathrm{N}}z)^k\varepsilon_f\mathrm{G}(\chi^{-1})f_\chi(z),$$

(iii) $$\mathrm{L}(f_\chi,s)=\prod_{p|\mathrm{N}}\frac{1}{1-\chi(p)c_p(f)p^{-s}}\times\prod_{(p,\mathrm{N})=1}\frac{1}{1-\chi(p)c_p(f)p^{-s}+\chi(p)^2p^{k-1-2s}}$$

(iv) $$\Lambda(f_\chi,s)=\Gamma(s)\Big(\frac{\mathrm{D}\sqrt{\mathrm{N}}}{2\pi}\Big)^s\mathrm{L}(f_\chi,s)$$

vérifie l'équation fonctionnelle

$$\chi(-\mathrm{N})\frac{\Lambda(f_\chi,s)}{\mathrm{G}(\chi)}=i^{-k}\varepsilon_f\frac{\Lambda(f_{\chi^{-1}},s)}{\mathrm{G}(\chi^{-1})}.$$

Réciproquement, si $(c_m)_{m\geqslant 1}$ est une suite de nombres complexes telle que la série $\sum_{m=1}^{+\infty}|c_m|m^{-s}$ converge absolument dans un demi-plan $\mathrm{Re}(s)>\mathrm{A}$, et vérifie les propriétés ci-dessus, pour tout caractère de Dirichlet de niveau D premier à N, alors $\sum_{m=1}^{+\infty}a_mq^m$ est une forme primitive de niveau N.

III.8. Fonctions zêta de Hasse-Weil

Soit Λ une algèbre de type fini sur $\mathbf{Z}$; il existe donc $d\in\mathbf{N}$ et un idéal I de $\mathbf{Z}[\mathrm{X}_1,\dots,\mathrm{X}_d]$ tels que l'on ait $\Lambda\cong\mathbf{Z}[\mathrm{X}_1,\dots,\mathrm{X}_d]/\mathrm{I}$.

Proposition III.8.1. *Si $\mathfrak{m}$ est un idéal maximal de Λ, alors $\Lambda/\mathfrak{m}$ est un corps fini.*

La proposition précédente permet de considérer les séries de Dirichlet

$$\zeta_\Lambda(s)=\prod_{\mathfrak{m}}(1-|\Lambda/\mathfrak{m}|^{-s})^{-1}$$

et $$\zeta_{\Lambda,p}(s)=\prod_{\mathfrak{m}\ni p}(1-|\Lambda/\mathfrak{m}|^{-s})^{-1},\quad \text{si } p \text{ est un nombre premier.}$$

Ces séries sont des produits de séries de Dirichlet à coefficients positifs et la fonction $\zeta_\Lambda(s)$ se factorise en produit de facteurs d'Euler sous la forme

$$\zeta_\Lambda(s) = \prod_p \zeta_{\Lambda,p}(s).$$

1. Nombre de points des variétés sur les corps finis

L'anneau $\mathbf{Z}[\mathrm{X}_1, \ldots, \mathrm{X}_d]$ étant noethérien, l'idéal I possède un système fini de générateurs. Si $f_1, \ldots, f_n$ est un tel système, on peut s'intéresser aux solutions du système $f_1(x) = \cdots = f_n(x) = 0$ dans différents corps. Comme nous le verrons plus loin, la fonction $\zeta_\Lambda(s)$ est fabriquée à partir du nombre de solutions de ce système dans les différents corps finis. Considérons en particulier les variétés algébriques V_∞ et V_p, pour p premier, définie par

$$\mathrm{V}_\infty = \{x = (x_1, \ldots, x_d) \in \mathbf{C}^d,\ f_1(x) = \cdots = f_n(x) = 0\},$$
$$\mathrm{V}_p = \{x = (x_1, \ldots, x_d) \in \overline{\mathbf{F}_p}^d,\ f_1(x) = \cdots = f_n(x) = 0\}.$$

Si $r \geqslant 1$, soit

$$\mathrm{N}_{p,r} = |\mathrm{V}_p(\mathbf{F}_{p^r})| = |\{(x_1, \ldots, x_d) \in \mathrm{V}_p,\ x_i \in \mathbf{F}_{p^r} \text{ si } 1 \leqslant i \leqslant d\}|.$$

On définit la fonction zêta de V_p par la formule

$$\mathrm{Z}_{\mathrm{V}_p}(\mathrm{T}) = \exp\Big(\sum_{r=1}^{+\infty} \frac{\mathrm{N}_{p,r}}{r}\mathrm{T}^r\Big).$$

Théorème III.8.2

(i) $\mathrm{Z}_{\mathrm{V}_p}(\mathrm{T}) \in \mathbf{Q}(\mathrm{T})$

(ii) *Si* $\mathrm{I} \cap \mathbf{Z} = 0$, *il existe* $\mathrm{M} \geqslant 0$, *tel que, si* $p \geqslant \mathrm{M}$, *alors on peut écrire* $\mathrm{Z}_{\mathrm{V}_p}(\mathrm{T})$ *sous la forme* $\prod_{i=0}^{2d} \frac{\mathrm{P}_{i,p}(\mathrm{T})}{\mathrm{Q}_{i,p}(\mathrm{T})}$, *où, si* $0 \leqslant i \leqslant 2d$, $\mathrm{P}_{i,p}$ *et* $\mathrm{Q}_{i,p}$ *sont des polynômes à coefficients entiers dont le terme constant est égal à* 1, *et dont tous les zéros sont de valeur absolue* $p^{-i/2}$. *De plus,* $\deg \mathrm{P}_{i,p} - \deg \mathrm{Q}_{i,p}$ *ne dépend que de la géométrie de* V_∞ *; en particulier,* $\deg \mathrm{P}_{i,p} - \deg \mathrm{Q}_{i,p}$ *ne dépend pas de* $p \geqslant \mathrm{M}$.

Commentaire. Le théorème ci-dessus est une version imprécise des conjectures de Weil (1949) sur le nombre de points des variétés algébriques sur les corps finis. Le (i) a été démontré par Dwork (1962) par une méthode très astucieuse et « élémentaire ». L'existence d'une décomposition naturelle de $\mathrm{Z}_{\mathrm{V}_p}(\mathrm{T})$ sous la forme $\prod_{i=0}^{2d} \frac{\mathrm{P}_{i,p}(\mathrm{T})}{\mathrm{Q}_{i,p}(\mathrm{T})}$ telle que $\deg \mathrm{P}_{i,p} - \deg \mathrm{Q}_{i,p}$ ne dépende que de la géométrie de V_∞ a été démontrée par Grothendieck (1964) comme aboutissement d'un énorme programme qui a totalement révolutionné les mathématiques. Finalement, le fait que les zéros de $\mathrm{P}_{i,p}$ et $\mathrm{Q}_{i,p}$ sont de valeur absolue $p^{-i/2}$, « hypothèse de Riemann sur les corps finis », a été démontré par Deligne (1973).

2. La conjecture de Hasse-Weil

Pour faire le lien entre $\mathrm{Z}_{\mathrm{V}_p}(\mathrm{T})$ et le facteur d'Euler $\zeta_{\Lambda,p}$ de ζ_Λ en p, nous aurons besoin du résultat suivant qui est une des formes du « théorème des zéros de Hilbert ». Si $\mathfrak{m}$ est un idéal maximal de Λ tel que la caractéristique du corps $\Lambda/\mathfrak{m}$ est p, soit

$$\mathrm{V}_{\mathfrak{m}} = \{x \in \mathrm{V}_p,\ f(x) = 0 \text{ quel que soit } f \in \mathfrak{m}\}.$$

Proposition III.8.3. *Si p est un nombre premier, l'application $\mathfrak{m} \mapsto \mathrm{V}_{\mathfrak{m}}$ induit une bijection de l'ensemble des idéaux maximaux de Λ contenant p sur l'ensemble des orbites de V_p sous l'action de $\mathrm{G}_p = \mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$. De plus, si $\mathfrak{m}$ est un idéal maximal de Λ contenant p, alors le cardinal de $\mathrm{V}_{\mathfrak{m}}$ est égal au degré du corps résiduel $\Lambda/\mathfrak{m}$ sur $\mathbf{F}_p$.*

Proposition III.8.4. *On a*

$$\zeta_{\Lambda,p}(s) = \mathrm{Z}_{\mathrm{V}_p}(p^{-s}).$$

Démonstration

Si $x = (x_1, \ldots, x_d) \in \mathrm{V}_p$, soit $d(x) = [\mathbf{F}_p[x_1, \ldots, x_d] : \mathbf{F}_p]$; c'est le degré du *corps de définition* de x.

Si $r \in \mathbf{N}$, soit $\mathrm{N}'_{p,r} = |\{x \in \mathrm{V}_p,\ d(x) = r\}|$. Comme les sous-corps de $\mathbf{F}_{p^e}$ sont les $\mathbf{F}_{p^r}$, avec $r|e$, on a

$$\mathrm{N}_{p,e} = \sum_{r|e} \mathrm{N}'_{p,e}.$$

Maintenant, comme les idéaux maximaux de Λ contenant p sont en bijection avec les idéaux maximaux de Λ_p, et donc aussi avec les orbites de V_p sous l'action de G_p, et comme l'orbite d'un point $x \in \mathrm{V}_p$ contient $d(x)$ éléments, on a

$$\begin{aligned}\zeta_{\Lambda,p}(s) &= \prod_{x\in \mathrm{V}_p/\mathrm{G}_p} (1-(p^{d(x)})^{-s})^{-1} \\ &= \prod_{x\in \mathrm{V}_p} (1-p^{-d(x)s})^{-1/d(x)} = \prod_{r=1}^{+\infty}(1-p^{-rs})^{-\mathrm{N}'_{p,r}/r} \\ &= \exp\Big(\sum_{r=1}^{+\infty}\frac{\mathrm{N}'_{p,r}}{r}\sum_{d=1}^{+\infty}\frac{p^{-drs}}{d}\Big) = \exp\Big(\sum_{e=1}^{+\infty}\frac{p^{-es}}{e}\Big(\sum_{r|e}\mathrm{N}'_{p,r}\Big)\Big)\end{aligned}$$

ce qui permet de conclure.

Comme le nombre de points de $\mathrm{V}_p(\mathbf{F}_{p^r})$ est $\leqslant p^{dr}$, cela permet de démontrer que le produit définissant $\zeta_\Lambda(s)$ converge absolument pour $\mathrm{Re}(s) > d+1$ et donc définit une fonction holomorphe sur le demi-plan $\mathrm{Re}(s) > d+1$.

Conjecture III.8.5

(i) *$\zeta_\Lambda(s)$ admet un prolongement méromorphe à tout le plan complexe, dont les zéros et les pôles se situent aux entiers $\leqslant d+1$ et sur les droites $\mathrm{Re}(s) = i/2$, $0 \leqslant i \leqslant 2d+1$.*

(ii) *Plus précisément, si on fixe i, on peut compléter $\prod_{p\geqslant \mathrm{M}} \frac{\mathrm{P}_{i,p}(p^{-s})}{\mathrm{Q}_{i,p}(p^{-s})}$ en une fonction $\mathrm{L}_{\Lambda,i}(s)$ en la multipliant par un produit de facteurs d'Euler pour $p < \mathrm{M}$ et par un produit de facteurs Gamma [i.e. un produit de facteurs du type $\pi^{-(s+k)/2}\Gamma(\frac{s+k}{2})$] ne dépendant que de la géométrie de V_∞, de telle sorte que*

a) *$\mathrm{L}_{\Lambda,i}(s)$ ait un prolongement méromorphe à tout le plan complexe et une équation fonctionnelle du type*

$$\mathrm{L}_{\Lambda,i}(s) = \pm \mathrm{N}_i^{-s}\mathrm{L}_{\Lambda,i}(i+1-s),$$

où N_i est un entier dont tous les diviseurs premiers sont $\leqslant \mathrm{M}$,

b) *les zéros et les pôles de $\mathrm{L}_{\Lambda,i}(s)$ sont tous sur la droite $\mathrm{Re}(s) = \frac{i+1}{2}$ (et en $\frac{i}{2}$, $\frac{i}{2}+1$ si i est pair).*

Commentaire. Cette conjecture est presque un théorème si $\mathrm{I}\cap\mathbf{Z}\neq\{0\}$ d'après le théorème III.8.2 et la proposition III.8.4. Par contre, dans le cas où $\mathrm{I}\cap\mathbf{Z}=\{0\}$, la conjecture (ii) a) n'a été démontrée que dans des cas très particuliers, le plus célèbre étant celui correspondant à la conjecture de Taniyama-Weil. Quant au (ii) b), « hypothèse de Riemann généralisée », il n'a été démontré dans aucun cas !

3. Exemples

Si $\Lambda=\mathbf{Z}$, on retombe sur la fonction zêta de Riemann $\zeta(s)$. Plus généralement, si F est un corps de nombres et Λ est l'anneau des entiers de F, alors $\zeta_\Lambda(s)$ est la fonction zêta de Dedekind de F.

Si $\Lambda=\mathbf{Z}[\mathrm{X}_1,\dots,\mathrm{X}_d]$, on a $\mathrm{I}=0$ et $\mathrm{V}_p=\overline{\mathbf{F}_p}^d$, et donc $\mathrm{V}_p(\mathbf{F}_{p^r})=\mathbf{F}_{p^r}^d$ et $\mathrm{N}_{p,r}=p^{dr}$. Ceci nous donne

$$\mathrm{Z}_{\mathrm{V}_p}(\mathrm{T})=\exp\Big(\sum_{r=1}^{+\infty}\frac{p^{dr}}{r}\mathrm{T}^r\Big)=(1-p^d\mathrm{T})^{-1},$$

et donc $\zeta_\Lambda(s)=\zeta(s-d)$.

Si $\Lambda=\mathbf{Z}[\mathrm{X}]/(2\mathrm{X}-1)=\mathbf{Z}[\frac{1}{2}]$, les idéaux maximaux de Λ sont ceux de $\mathbf{Z}$ moins l'idéal (2) puisqu'on a rendu 2 inversible. On obtient donc $\zeta_\Lambda(s)=(1-2^{-s})\zeta(s)$.

Si $\Lambda=\mathbf{Z}[\mathrm{X}]/(2\mathrm{X})$, on a $\mathrm{V}_p=\{0\}$ et $\zeta_{\Lambda,p}(s)=(1-p^{-s})^{-1}$ si $p\neq 2$, et $\mathrm{V}_2=\overline{\mathbf{F}_2}$ et $\zeta_{\Lambda,2}(s)=(1-2^{1-s})^{-1}$, ce qui nous donne $\zeta_\Lambda(s)=\frac{1-2^{-s}}{1-2^{1-s}}\zeta(s)$.

Si $\mathrm{P}(\mathrm{X})=\mathrm{X}(\mathrm{X}-\mathrm{N})$ et $\Lambda=\mathbf{Z}[\mathrm{X}]/\mathrm{P}$, alors $\mathrm{V}_p=\{0,\mathrm{N}\}\subset\mathbf{F}_p$ a deux éléments et $\zeta_{\Lambda,p}(s)=(1-p^{-s})^{-2}$ si p ne divise pas N, et un seul et $\zeta_{\Lambda,p}(s)=(1-p^{-s})^{-1}$, si p divise N, on obtient donc $\zeta_\Lambda(s)=\zeta(s)^2\prod_{p|\mathrm{N}}(1-p^{-s})$.

Plus généralement, si $\mathrm{P}\in\mathbf{Z}[\mathrm{X}]$ est un polynôme non constant, si $\mathrm{P}=\mathrm{P}_1^{k_1}\cdots\mathrm{P}_r^{k_r}$ est sa décomposition en facteurs irréductibles dans $\mathbf{Q}[\mathrm{X}]$, si F_i est le corps de nombres $\mathbf{Q}[\mathrm{X}]/\mathrm{P}_i$ et si $\Lambda=\mathbf{Z}[\mathrm{X}]/\mathrm{P}$, alors $\zeta_\Lambda(s)$ ne diffère de $\prod_{i=1}^r\zeta_{\mathrm{F}_i}(s)$ que par un nombre fini de facteurs d'Euler.

Si Λ est monogène sur $\mathbf{Z}$ (*i.e.* est un quotient de $\mathbf{Z}[\mathrm{X}]$), le théorème I.6.1 permet donc de démontrer une bonne partie de la conjecture III.8.5. C'est très loin d'être le cas dans le cas non monogène.

Par exemple, si $\Lambda = \mathbf{Z}[\mathrm{X},\mathrm{Y}]/\mathrm{P}$, la fonction $\zeta_\Lambda(s)$ s'exprime en termes de fonctions zêta de Dedekind si $\deg \mathrm{P} \leqslant 2$, mais ce n'est plus le cas (en général), si $\deg \mathrm{P} \geqslant 3$. Si $\deg \mathrm{P} \geqslant 4$, on ne sait traiter que des cas particuliers très spéciaux et le cas $\deg \mathrm{P} = 3$ vient en grande partie d'être résolu par Wiles et ses successeurs (Breuil, Conrad, Diamond et Taylor) qui ont démontré la conjecture de Taniyama-Weil : de manière (im)précise, on a le résultat suivant.

Théorème III.8.6. *Si* $\alpha, \beta, \gamma \in \mathbf{Z}$, *si* $\mathrm{P}(\mathrm{X},\mathrm{Y}) = \mathrm{Y}^2 - \mathrm{X}^3 - \alpha\mathrm{X}^2 - \beta\mathrm{X} - \gamma$, *et si* $\Lambda = \mathbf{Z}[\mathrm{X},\mathrm{Y}]/\mathrm{P}$, *alors il existe un entier* $\mathrm{N_P}$ (« explicite ») *et une forme modulaire* f_P *parabolique primitive de poids* 2 *pour* $\Gamma_0(\mathrm{N_P})$ *tels que* $\zeta_\Lambda(s)$ *ne diffère de* $\mathrm{L}(f_\mathrm{P}, s)^{-1}\zeta(s-1)$ *que par un nombre fini de facteurs d'Euler.*

Ce théorème, couplé avec la « conjecture ε » de Serre, démontrée par Ribet, a des retombées assez spectaculaires. La « conjecture ε » décrit les congruences que l'on peut avoir entre les coefficients de Fourier de deux formes paraboliques primitives pour des poids et des niveaux différents (les coefficients de Fourier d'une forme parabolique primitive sont des entiers d'une extension finie de $\mathbf{Q}$). Si $p \geqslant 5$ est un nombre premier, si a, b, c sont des entiers non nuls tels que $a^p + b^p = c^p$, si $\mathrm{P}(\mathrm{X},\mathrm{Y}) = \mathrm{Y}^2 - \mathrm{X}(\mathrm{X} - a^p)(\mathrm{X} + b^p)$, et si $f_\mathrm{P} = \sum_{n=1}^{+\infty} a_n q^n$, alors la « conjecture ε » garantit l'existence d'une forme modulaire parabolique $g = \sum_{n=1}^{+\infty} b_n q^n$, de poids 2 pour $\Gamma_0(2)$ ou $\Gamma_0(1)$, telle que, si F est le corps de nombres obtenu en adjoignant les b_n à $\mathbf{Q}$, alors $\mathrm{N}_{\mathrm{F}/\mathbf{Q}}(a_n - b_n)$ est divisible par p quel que soit $n \geqslant 1$ (plus précisément, il existe un idéal premier $\mathfrak{p}$ de $\mathscr{O}_\mathrm{F}$ divisant p tel que a_n soit congru à b_n modulo $\mathfrak{p}$ pour tout $n \geqslant 1$). Comme $a_1 = 1$ et comme une forme parabolique de poids 2 pour $\Gamma_0(2)$ ou $\Gamma_0(1)$ est identiquement nulle, cela conduit à une contradiction ; on en déduit le théorème de Fermat.

Chapitre IV. Les nombres p-adiques

IV.1. Généralités sur les corps normés

1. Normes et valuations

Définition IV.1.1. Soit K un corps. Une *norme* sur K est une application $x \mapsto |x|$ de K dans $\mathbf{R}_+$ vérifiant les 3 propriétés suivantes

(i) $|x| = 0 \Leftrightarrow x = 0$.

(ii) $|xy| = |x| \cdot |y|$.

(iii) $|x + y| \leqslant |x| + |y|$

Si $|\ |$ vérifie la condition (iii′) $|x + y| \leqslant \sup(|x|, |y|)$ plus forte que la condition (iii), on dit que la norme est *ultramétrique*.

Définition IV.1.2. Si K est un corps, une *valuation* v sur K est une application $v : \mathrm{K} \to \mathbf{R}_+ \cup \{+\infty\}$ vérifiant les trois conditions suivantes :

(i) $v(x) = +\infty \Leftrightarrow x = 0$.

(ii) $v(xy) = v(x) + v(y)$.

(iii) $v(x + y) \geqslant \inf(v(x), v(y))$.

Remarque IV.1.3

(i) Si K est un corps muni d'une norme ultramétrique $|\ |$ et si $\lambda < 0$, alors $v : \mathrm{K} \to \mathbf{R}_+ \cup \{+\infty\}$ défini par $v(x) = \lambda \log |x|$ est une valuation sur K.

(ii) Réciproquement, si v est une valuation sur K et $0 < a < 1$, alors $|x| = a^{v(x)}$ est une norme ultramétrique sur K.

Lemme IV.1.4. *Si $|\ |$ est ultramétrique et $|x| \neq |y|$, alors $|x + y| = \sup(|x|, |y|)$.*

Démonstration. Quitte à permuter x et y, on peut supposer $|x| > |y|$. On a alors

$$|x + y| \leqslant |x| = |(x + y) - y| \leqslant \sup(|x + y|, |y|)$$

et comme $|y| < |x|$, on en déduit l'égalité $\sup(|x + y|, |y|) = |x + y|$, ce qui permet de conclure.

Proposition IV.1.5. *Si* K *est un corps et $|\ |$ est une norme ultramétrique sur* K*, alors $\mathscr{O}_\mathrm{K} = \{x \in \mathrm{K} \mid |x| = 1\}$ est un anneau local d'idéal maximal $\mathfrak{m} = \{x \in \mathrm{K} \mid |x| < 1\}$.*

Démonstration. Le fait que $\mathscr{O}_{\mathrm{K}}$ soit un anneau et $\mathfrak{m}$ un idéal est une conséquence immédiate de l'inégalité ultramétrique. D'autre part, si x est un élément de $\mathscr{O}_{\mathrm{K}}$ n'appartenant pas à $\mathfrak{m}$, alors $|x| = 1$ et x^{-1} est un élément de K de norme 1 donc appartient à $\mathscr{O}_{\mathrm{K}}$, ce qui prouve que $\mathscr{O}_{\mathrm{K}} - \mathfrak{m}$ n'est autre que le groupe des unités $\mathscr{O}_{\mathrm{K}}^{*}$ de $\mathscr{O}_{\mathrm{K}}$ et permet de conclure.

Définition IV.1.6. L'anneau $\mathscr{O}_{\mathrm{K}}$ s'appelle l'*anneau des entiers* de K et le corps $\mathscr{O}_{\mathrm{K}}/\mathfrak{m}$ le *corps résiduel* de K.

2. Topologie associée à une norme

Si K est un corps muni d'une norme $|\ |$, et x, y sont deux éléments de K, on pose $d(x, y) = |x - y|$. La propriété (iii) des normes assure que d est une distance sur K et donc définit une topologie sur K

Remarque IV.1.7. Si $|\ |$ est ultramétrique, alors

(i) Tout triangle est isocèle.

(ii) Tout point d'une boule en est « le » centre.

(iii) Deux boules sont soit disjointes soit l'une est contenue dans l'autre (comme des billes de mercure).

(iv) Les boules sont à la fois ouvertes et fermées.

(v) La topologie est totalement discontinue.

Démonstration. Le point (i) est une conséquence immédiate du lemme IV.1.4. Si $x_1 \in \mathrm{B}(x_0, r)$ et $y \in \mathrm{B}(x_1, r)$, alors

$$d(x_0, y) \leqslant \sup(d(x_0, x_1), d(x_1, y)) \leqslant r$$

et donc $\mathrm{B}(x_1, r) \subset \mathrm{B}(x_0, r)$. L'inclusion dans l'autre sens s'obtient en échangeant les rôles de x_0 et x_1, ce qui permet de démontrer le (ii). D'après le (ii), si deux boules ont une intersection non vide, tout élément de l'intersection est le centre des deux boules ce qui démontre le (iii). Le (v) est une conséquence immédiate du (iv) et si B est une boule ouverte de rayon r, le complémentaire de B contient la boule de rayon $r/2$ autour de chacun de ses points d'après le (iii), ce qui montre que ce complémentaire est ouvert et donc que B est fermée. Si B est une boule fermée, alors B est un voisinage de chacun de ses points puisque ceux-ci en sont « le » centre.

Définition IV.1.8. Deux normes sur un corps K sont dites *équivalentes* si elle définissent la même topologie.

Proposition IV.1.9. *Deux normes $|\ |_1$ et $|\ |_2$ sont équivalentes si et seulement si il existe $s \in \mathbf{R}_+^*$ tel que l'on ait $|x|_1 = |x|_2^s$ quel que soit $x \in \mathrm{K}^*$.*

Démonstration. Si x est une norme sur un corps K, alors $|x| < 1$ si et seulement si la suite de terme général x^n tend vers 0 quand n tend vers $+\infty$. On en déduit le fait que si $|\ |_1$ et $|\ |_2$ sont équivalentes, alors

$$\{x \in \mathrm{K} \mid |x|_1 < 1\} = \{x \in \mathrm{K} \mid |x|_2 < 1\}.$$

Si ce dernier ensemble est réduit à $\{0\}$, on a $|x|_1 = |x|_2 = 1$ quel que soit $x \in \mathrm{K}^*$. Sinon, soit $x \in \mathrm{K}^*$ vérifiant $|x|_1 < 1$. Si $y \in \mathrm{K}^*$, $a \in \mathbf{Z}$ et $b \in \mathbf{N}$, alors

$$|y^b x^{-a}|_1 < 1 \Longleftrightarrow |y^b x^{-a}|_2 < 1.$$

On en déduit le fait que

$$\left\{r \in \mathbf{Q} \mid r < \frac{\log|y|_1}{\log|x|_1}\right\} = \left\{r \in \mathbf{Q} \mid r < \frac{\log|y|_2}{\log|x|_2}\right\}$$

et donc que les réels $\log|y|_1/\log|x|_1$ et $\log|y|_2/\log|x|_2$ définissent la même coupure de Dedekind et sont donc égaux, ce qui montre que la fonction $y \mapsto \log|y|_1/\log|y|_2$ est constante sur K^* et permet de conclure.

3. Complétion

Si K est un corps normé, on note $\widetilde{\mathrm{K}}$ l'ensemble des suites de Cauchy à valeurs dans K, c'est-à-dire l'ensemble des suites $(a_n)_{n\in\mathbf{N}}$ telles que

$$\forall\, \varepsilon > 0,\ \ \exists\, \mathrm{N} \in \mathbf{N},\ \ \forall\, n \geqslant \mathrm{N} \text{ et } k \in \mathbf{N},\quad |a_n - a_{n+k}| < \varepsilon.$$

Soit $\mathrm{I} \subset \widetilde{\mathrm{K}}$ l'ensemble des suites tendant vers 0.

Lemme IV.1.10

(i) *Si $(a_n)_{n\in\mathbf{N}} \in \widetilde{\mathrm{K}}$, alors la suite $(|a_n|)_{n\in\mathbf{N}}$ converge dans $\mathbf{R}_+$.*

(ii) *Si on suppose de plus que $|\ |$ est ultramétrique et que $a \notin \mathrm{I}$, alors la suite $(|a_n|)_{n\in\mathbf{N}}$ est constante à partir d'un certain rang.*

(iii) *Si* $a = (a_n)_{n\in\mathbf{N}}$ *et* $b = (b_n)_{n\in\mathbf{N}}$ *sont deux éléments de* $\widetilde{\mathrm{K}}$ *différant par un élément de* I, *alors* $\lim_{n\to+\infty} |a_n| = \lim_{n\to+\infty} |b_n|$.

Démonstration

(i) L'inégalité triangulaire implique $||a_{n+k}| - |a_n|| \leqslant |a_{n+k} - a_n|$ quels que soient n et k et la suite de terme général $|a_n|$ est de Cauchy, ce qui permet de conclure.

(ii) Si $a = (a_n)_{n\in\mathbf{N}} \in \widetilde{\mathrm{K}} - \mathrm{I}$, il existe $\delta > 0$ tel que l'on ait $|a_n| \geqslant \delta$ pour une infinité de n et la limite de la suite $|a_n|$ est donc supérieure ou égale à δ. Il existe donc $\mathrm{N} \in \mathbf{N}$ tel que si $n \geqslant \mathrm{N}$, alors $|a_n| > \frac{2}{3}\delta$ et $|a_{n+k} - a_n| < \delta/2$ quel que soit $k \in \mathbf{N}$. Ceci implique $|a_{n+k} - a_n| < |a_n|$ et donc, comme $|\ |$ est supposée ultramétrique, $|a_{n+k}| = |a_n|$ quel que soit $k \in \mathbf{N}$; d'où le résultat.

(iii) On a $||a_n| - |b_n|| \leqslant |a_n - b_n|$ et l'hypothèse implique que cette dernière suite tend vers 0 d'où le résultat.

Lemme IV.1.11. $\widetilde{\mathrm{K}}$ *est un anneau et* I *est un idéal maximal de* $\widetilde{\mathrm{K}}$.

Démonstration. Le fait que $\widetilde{\mathrm{K}}$ est un anneau et I un idéal est immédiat. L'élément unité de $\widetilde{\mathrm{K}}$ est la suite constante 1 dont tous les termes sont égaux à 1. Si $a = (a_n)_{n\in\mathbf{N}} \in \widetilde{\mathrm{K}} - \mathrm{I}$, d'après ce qui précède, il existe $\delta > 0$ et $\mathrm{N} \in \mathbf{N}$ tels que l'on ait $|a_n| \geqslant \delta$ si $n \geqslant \mathrm{N}$. La suite $b = (b_n)_{n\in\mathbf{N}}$ définie par $b_n = 0$ si $n < \mathrm{N}$ et $b_n = a_n^{-1}$ si $n \geqslant \mathrm{N}$ est de Cauchy et $ab - 1$ est élément de I, ce qui montre que a est inversible dans $\widetilde{\mathrm{K}}/\mathrm{I}$ et permet de conclure au fait que I est maximal.

Il résulte du lemme précédent que $\widehat{\mathrm{K}} = \widetilde{\mathrm{K}}/\mathrm{I}$ est un corps et du lemme IV.1.10, que $|\ |$ s'étend à $\widehat{\mathrm{K}}$.

Proposition IV.1.12. $|\ |$ *est une norme sur* $\widehat{\mathrm{K}}$ *et* $\widehat{\mathrm{K}}$ *est complet pour cette norme et contient* K *comme sous-corps dense.*

Démonstration. La multiplicativité de la norme et l'inégalité triangulaire (resp. ultramétrique) passent à la limite. D'autre part, $|a| = 0 \Leftrightarrow \lim_{n\to+\infty} |a_n| = 0 \Leftrightarrow a \in \mathrm{I}$ et donc $|\ |$ est une norme sur $\widehat{\mathrm{K}}$ qui est ultramétrique si $|\ |$ est ultramétrique sur K. Il est clair que, si $a = (a_n)_{n\in\mathbf{N}} \in \widetilde{\mathrm{K}}$, alors $a = \lim_{n\to+\infty} a_n$ dans $\widehat{\mathrm{K}}$ et donc que K est dense dans $\widetilde{\mathrm{K}}$. Finalement, si $(a_n)_{n\in\mathbf{N}}$ est une suite de Cauchy

dans $\widehat{\mathrm{K}}$, comme K est dense dans $\widehat{\mathrm{K}}$, on peut trouver pour chaque n un élément b_n de K tel que l'on ait $|a_n - b_n| \leqslant 2^{-n}$ et la suite b_n est de Cauchy dans K donc converge dans $\widehat{\mathrm{K}}$ vers une limite qui est aussi celle de la suite $(a_n)_{n\in\mathbf{N}}$; ce qui prouve que $\widehat{\mathrm{K}}$ est complet.

Remarque IV.1.13

(i) Si K est ultramétrique, il résulte du (ii) du lemme IV.1.10 que $|\widehat{\mathrm{K}}^*| = |\mathrm{K}^*|$ et donc que $|\widehat{\mathrm{K}}^*|$ est un sous-groupe de $\mathbf{R}_+^*$ qui n'a aucune raison d'être complet.

(ii) L'inégalité ultramétrique montre qu'une suite $(u_n)_{n\in\mathbf{N}}$ est de Cauchy si et seulement si $|u_{n+1} - u_n|$ tend vers 0. Ceci implique que, si K est un corps ultramétrique complet, une suite $(u_n)_{n\in\mathbf{N}}$ converge dans K si et seulement si $u_{n+1} - u_n$ tend vers 0. De même, une série converge si et seulement si son terme général tend vers 0.

Définition IV.1.14. Le corps $\widehat{\mathrm{K}}$ (muni de la norme $|\ |$) s'appelle le *complété* de K pour la norme $|\ |$.

4. Le lemme de Hensel

Soit K un corps complet pour une norme ultramétrique. La proposition suivante (lemme de Hensel) est très utile pour localiser les zéros des polynômes à coefficients dans K.

Proposition IV.1.15. *Si $f \in \mathscr{O}_{\mathrm{K}}[\mathrm{X}]$ et $\alpha \in \mathscr{O}_{\mathrm{K}}$ vérifie $|f(\alpha)| < |f'(\alpha)|^2$, alors il existe $\widetilde{\alpha} \in \mathscr{O}_{\mathrm{K}}$ unique vérifiant les conditions $|\widetilde{\alpha} - \alpha| < |f(\alpha)/f'(\alpha)|$ et $f(\widetilde{\alpha}) = 0$.*

Démonstration. Soit $\varphi : \mathrm{K} \to \mathrm{K}$ la fonction définie par $\varphi(x) = x - f(x)/f'(\alpha)$ et soit $\delta = |f(\alpha)/f'(\alpha)|$. Nous allons montrer que φ est une application contractante de la boule fermée $\mathrm{B}(\alpha, \delta)$ dans elle-même, ce qui permet de conclure car, cette boule étant complète, l'unique point fixe de φ est le $\widetilde{\alpha}$ que l'on cherche. Commençons par remarquer que, si $g \in \mathscr{O}_{\mathrm{K}}[\mathrm{X}]$ et $x, y \in \mathscr{O}_{\mathrm{K}}$, alors

$$\Big|g(y) - \sum_{i=0}^{k} \frac{g^{(i)}(x)}{i!}(y-x)^i\Big| \leqslant |y-x|^{k+1}$$

d'après la formule de Taylor pour les polynômes car le polynôme $g^{(i)}(x)/i!$ est à coefficients dans $\mathscr{O}_{\mathrm{K}}$ quel que soit $i \in \mathbf{N}$ et sa valeur

absolue est donc $\leqslant 1$ en un point de $\mathscr{O}_{\mathrm{K}}$. En particulier, on a

$$|f(x)-f(\alpha)| \leqslant \sup(|f'(\alpha)|\cdot|x-\alpha|, |x-\alpha|^2) \leqslant |f'(\alpha)|\cdot|x-\alpha|,$$

si $|x-\alpha| \leqslant \delta < |f'(\alpha)|$; ceci permet de montrer que φ laisse stable $\mathrm{B}(\alpha,\delta)$. De plus, on a

$$\begin{aligned}
&|f'(\alpha)(\varphi(x)-\varphi(y))| \\
&\quad = |f(y)-f(x)-f'(x)(y-x)+(f'(x)-f'(\alpha))(y-x)| \\
&\quad \leqslant \sup(|f(y)-f(x)-f'(x)(y-x)|, |(f'(x)-f'(\alpha))(y-x)|) \\
&\quad \leqslant \sup(|y-x|^2, |x-\alpha|\cdot|y-x|) \leqslant \delta\cdot|y-x|,
\end{aligned}$$

si $x, y \in \mathrm{B}(\alpha,\delta)$; comme $\delta < |f'(\alpha)|$, cela permet de conclure.

Remarque IV.1.16. La démonstration ci-dessus est la version ultramétrique de l'algorithme de Newton.

Corollaire IV.1.17. *Soit $f \in \mathscr{O}_{\mathrm{K}}[\mathrm{X}]$ un polynôme unitaire et soit $\overline{f}$ la réduction de f modulo $\mathfrak{m}$. Si $\overline{\alpha}$ est une racine simple de $\overline{f}$ dans k, alors il existe $\widetilde{\alpha} \in \mathscr{O}_{\mathrm{K}}$ unique, dont la réduction modulo $\mathfrak{m}$ est $\overline{\alpha}$ et qui vérifie $f(\widetilde{\alpha}) = 0$*

Démonstration. Soit $\alpha \in \mathscr{O}_{\mathrm{K}}$ dont la réduction modulo $\mathfrak{m}$ est $\overline{\alpha}$. L'hypothèse selon laquelle $\overline{\alpha}$ est racine simple de $\overline{f}$ se traduit par $|f(\alpha)| < 1$ et $|f'(\alpha)| = 1$, ce qui permet d'utiliser la proposition IV.1.15 pour conclure.

IV.2. Construction du corps $\mathbf{C}_p$

1. Normes sur Q

Si p est un nombre premier, on note $v_p : \mathbf{Q}^* \to \mathbf{Z}$ la valuation p-adique : si a est un entier, $v_p(a)$ est le plus grand entier n tel que p^n divise a, et $v_p(b^{-1}a) = v_p(a) - v_p(b)$ si a et b sont des entiers. Il n'est pas difficile de vérifier que v_p est bien définie et est une valuation sur $\mathbf{Q}$ (en posant $v_p(0) = +\infty$). On définit la norme p-adique $|\ |_p$ sur $\mathbf{Q}$ par la formule $|x|_p = p^{-v_p(x)}$.

Théorème IV.2.1 **(Ostrowski).** *Une norme non triviale sur $\mathbf{Q}$ est équivalente à la norme usuelle $|\ |_\infty$ ou à la norme p-adique pour un nombre premier p.*

Démonstration. Commençons par supposer qu'il existe $k \in \mathbf{N}$ tel que $\|k\| > 1$. Comme $\|1\| = 1$, l'inégalité triangulaire implique $\|k\| \leqslant k$ et il existe $\alpha \in]0,1]$ tel que l'on ait $\|k\| = k^\alpha$. Soit $m \in \mathbf{N}$. On peut écrire m en base k sous la forme $m = \sum_{i=0}^{n} a_i k^i$ avec $a_i \in \{0,1,\dots,k-1\}$ et $a_n \neq 0$ de telle sorte que l'on a $m \geqslant k^n$. Comme $\|a_i\| \leqslant a_i \leqslant k-1$ et $\|k^i\| = \|k\|^i$, on obtient la majoration

$$\|m\| \leqslant (k-1)\sum_{i=1}^{n} k^{i\alpha} = \frac{k-1}{k^\alpha - 1}(k^{(n+1)\alpha}-1) \leqslant \frac{k^\alpha(k-1)}{k^\alpha-1}k^{n\alpha} \leqslant \mathrm{C}m^\alpha,$$

où $\mathrm{C} = k^\alpha(k-1)/(k^\alpha-1)$ est indépendant de m. On peut appliquer cette inégalité à m^n, ce qui nous donne $\|m\|^n \leqslant \mathrm{C}m^{n\alpha}$ et, prenant la racine n-ième de cette égalité et passant à la limite, l'inégalité $\|m\| \leqslant m^\alpha$. On a donc

$$\frac{\log\|m\|}{\log m} \leqslant \frac{\log\|k\|}{\log k}$$

quel que soit $m \in \mathbf{N}$. Par symétrie, on en déduit le fait que si $\|m\| > 1$, alors cette inégalité est une égalité.

Maintenant, si $m \in \mathbf{N} - \{0\}$ est quelconque, il existe $n \in \mathbf{N}$ tel que l'on ait $\|k^n m\| > 1$. On a alors $\|m\| = \|k^n\|^{-1}\|k^n m\| = k^{-\alpha n}(k^n m)^\alpha = m^\alpha$, ce qui montre que l'on a égalité quel que soit $m \in \mathbf{N}$ puis, utilisant la multiplicativité de la norme et le fait que $\|-1\| = 1$, que $\|x\| = |x|_\infty^\alpha$ quel que soit $x \in \mathbf{Q}$. On en déduit que s'il existe $k \in \mathbf{N}$ tel que $\|k\| > 1$, alors $\|\ \|$ est équivalente à la norme usuelle.

Dans le cas contraire, on a $\|\ell\| \leqslant 1$ pour tout nombre premier ℓ. Comme on a supposé $\|\ \|$ non triviale, il existe au moins un nombre premier p tel que $\|p\| < 1$. Si il en existe un autre q, alors quel que soit $n \in \mathbf{N}$, on peut, d'après le théorème de Bézout, trouver $u_n, v_n \in \mathbf{Z}$ tels que l'on ait $u_n p^n + v_n q^n = 1$. On obtient donc

$$1 = \|1\| = \|u_n p^n + v_n q^n\| \leqslant \|u_n\| \cdot \|p\|^n + \|v_n\| \cdot \|q\|^n \leqslant \|p\|^n + \|q\|^n,$$

ce qui est impossible pour n assez grand. Il existe donc un et un seul nombre premier p tel que $\|p\| < 1$ et $\|\ \|$ est équivalente à la norme p-adique. Ceci termine la démonstration.

Le résultat suivant est trivial (c'est une conséquence immédiate de l'unicité de la décomposition d'un entier en produit de facteurs premiers), mais très important ; c'est ce qui justifie la normalisation utilisée pour $|\ |_p$

***Théorème IV.2.2* (Formule du produit).** *Si $x \in \mathbf{Q}^*$, alors*

$$|x|_\infty \cdot \prod_{p \text{ premier}} |x|_p = 1.$$

2. Le corps $\mathbf{Q}_p$ et l'anneau $\mathbf{Z}_p$

On note $\mathbf{Q}_p$ le complété de $\mathbf{Q}$ pour la norme p-adique et $\mathbf{Z}_p$ l'anneau de ses entiers. Son idéal maximal est $p\mathbf{Z}_p$ car $|\mathbf{Q}_p^*|_p = |\mathbf{Q}^*|_p = p^{\mathbf{Z}}$ et donc, si $|x|_p < 1$, alors $|x|_p \leqslant p^{-1}$. Il ressort de la discussion générale que $\mathbf{Q}_p$ est un corps ultramétrique complet et qu'une série converge dans $\mathbf{Q}_p$ si et seulement si son terme général tend vers 0.

Lemme IV.2.3. *L'application naturelle de $\mathbf{Z}/p^n\mathbf{Z}$ dans $\mathbf{Z}_p/p^n\mathbf{Z}_p$ est un isomorphisme.*

Démonstration. Si x est un élément de $\mathbf{Z} \cap p^n\mathbf{Z}_p$, on a $|x|_p \leqslant p^{-n}$, ce qui signifie que $v_p(x) \geqslant n$ et donc que x est divisible par p^n dans $\mathbf{Z}$. On en déduit l'injectivité. Prouvons sa surjectivité. Soit $\overline{x} \in \mathbf{Z}_p/p^n\mathbf{Z}_p$ et $x \in \mathbf{Z}_p$ ayant pour image $\overline{x}$ modulo p^n. Comme $\mathbf{Q}$ est dense dans $\mathbf{Q}_p$, il existe $r \in \mathbf{Q}$ vérifiant $|x - r|_p \leqslant p^{-n}$; en particulier $|r|_p \leqslant 1$. Écrivons r sous la forme a/b avec $a, b \in \mathbf{Z}$. Comme $|r|_p \leqslant 1$, on a $v_p(b) \leqslant v_p(a)$ et quitte à tout diviser par $p^{v_p(b)}$, on peut supposer $(b, p) = 1$. Soit $\overline{c}$ l'inverse de b dans $\mathbf{Z}/p^n\mathbf{Z}$ et $c \in \mathbf{Z}$ dont la réduction modulo p^n est $\overline{c}$. On a alors $|r - ac|_p = |a|_p|1 - bc|_p \leqslant p^{-n}$ et donc $|x - ac|_p \leqslant p^{-n}$, ce qui prouve que ac a pour image $\overline{x}$ dans $\mathbf{Z}_p/p^n\mathbf{Z}_p$ et permet de conclure.

On déduit de ce lemme une autre description, plus algébrique, de l'anneau $\mathbf{Z}_p$. On aurait d'ailleurs pu partir de cette construction, et inverser p pour obtenir $\mathbf{Q}_p$; les deux approches ont leurs avantages.

Corollaire IV.2.4. *L'application qui à $x \in \mathbf{Z}_p$ associe la suite $(x_n)_{n\in\mathbf{N}}$ de ses images modulo p^n, $n \in \mathbf{N}$ induit un isomorphisme d'anneaux topologiques de $\mathbf{Z}_p$ sur la limite projective des $\mathbf{Z}/p^n\mathbf{Z}$* (l'ensemble des

suites $(x_n)_{n\in\mathbf{N}}$, où $x_n \in \mathbf{Z}/p^n\mathbf{Z}$ et x_{n+1} a pour image x_n par l'application naturelle (de réduction modulo p^n) de $\mathbf{Z}/p^{n+1}\mathbf{Z}$ dans $\mathbf{Z}/p^n\mathbf{Z}$).

Théorème IV.2.5

(i) $\mathbf{Z}_p$ *est compact et* $\mathbf{Q}_p$ *est localement compact.*

(ii) *Tout élément de* $\mathbf{Z}_p$ *peut s'écrire de manière unique sous la forme* $\sum_{n=0}^{+\infty} p^n a_n$ *où* $a_n \in \{0, 1, \ldots, p-1\}$.

(iii) $\mathbf{N}$ *est dense dans* $\mathbf{Z}_p$ *; plus généralement, si* $b \in \mathbf{Z}$ *est premier à* p *et* $a \in \mathbf{Z}$, $a + b\mathbf{N}$ *est dense dans* $\mathbf{Z}_p$.

Démonstration. Le (i) suit du corollaire précédent : $\mathbf{Z}_p$ est un fermé d'un produit de compacts (et même d'ensembles finis) ; il est donc compact et les boules de $\mathbf{Q}_p$ sont isomorphes à $\mathbf{Z}_p$, donc compactes.

Il n'est pas difficile de voir que, si $x \in \mathbf{Z}_p$, et si x_n est l'unique élément de $\{0, \ldots, p^n - 1\}$ ayant même image que x dans $\mathbf{Z}/p^n\mathbf{Z}$, alors $\sum_{i=0}^{n} a_i p^i$ est le développement de x_n en base p (écrit dans le sens opposé à celui dont on a l'habitude...) ; on en déduit le (ii).

Le (iii) est une conséquence du (ii) (x est la limite de la suite de terme général $\sum_{i=0}^{n} p^i a_i$ dont tous les termes sont dans $\mathbf{N}$) et du fait que $x \to bx + a$ est une isométrie de $\mathbf{Z}_p$ si $(b, p) = 1$.

3. Le corps $\mathbf{C}_p$

Proposition IV.2.6. *Si* F *est une extension finie de degré* d *de* $\mathbf{Q}_p$*, alors il existe une unique norme* $|\ |$ *sur* F *dont la restriction à* $\mathbf{Q}_p$ *est la norme* p*-adique et on a* $|x| = \left|\mathrm{N}_{\mathrm{F}/\mathbf{Q}_p}(x)\right|^{1/d}$*, si* $x \in \mathrm{F}$.

Démonstration. Commençons par prouver l'unicité d'une telle norme. Supposons que l'on en ait deux $|\ |_1$ et $|\ |_2$. Comme $\mathbf{Q}_p$ est localement compact et F est un $\mathbf{Q}_p$-espace vectoriel de dimension finie, les normes $|\ |_1$ et $|\ |_2$ sont équivalentes. D'après la proposition IV.1.9, il existe $s \in \mathbf{R}_+$ tel que l'on ait $|x|_2 = |x|_1^s$ quel que soit $x \in \mathrm{F}$ et, prenant $x = p$, on obtient $s = 1$ et donc $|\ |_1 = |\ |_2$.

Pour démontrer l'existence de $|\ |$, choisissons une base $(e_1, \ldots, e_d)$ de F sur $\mathbf{Q}_p$. Si $u \in \mathrm{End}_{\mathbf{Q}_p}\mathrm{F}$ est un endomorphisme $\mathbf{Q}_p$-linéaire du $\mathbf{Q}_p$-espace vectoriel F, notons $\|u\|$ le maximum des normes p-adiques des coefficients de la matrice de u dans la base $(e_1, \ldots, e_d)$. On a

$\|u\| = 0$ si et seulement si $u = 0$ et, la norme p-adique étant ultramétrique,

$$\|u+v\| \leqslant \sup(\|u\|, \|v\|) \quad \text{et} \quad \|uv\| \leqslant \|u\| \cdot \|v\| \quad \text{si } u, v \in \mathrm{End}_{\mathbf{Q}_p}\mathrm{F}.$$

La suite de terme général $\|u^n\|^{1/n}$ converge vers sa limite inf. que l'on appelle la norme spectrale $\|u\|_{\mathrm{sp}}$ de u (c'est bien connu, et cela résulte facilement de l'inégalité $\alpha_{n+m} \leqslant \alpha_n^{n/(n+m)}\alpha_m^{m/(n+m)}$ si $\alpha_n = \|u^n\|^{1/n}$).

On peut considérer $x \in \mathrm{F}$ comme un élément de $\mathrm{End}_{\mathbf{Q}_p}(\mathrm{F})$ en associant à x la multiplication par x. Nous allons vérifier que l'on peut poser $|x| = \|x\|_{\mathrm{sp}}$.

— Si $x \in \mathbf{Q}_p$, on a $\|x\| = |x|_p$ car la matrice de la multiplication par x est la matrice diagonale avec des x sur la diagonale ; ceci permet de montrer que $|\ |$ coïncide avec $|\ |_p$ sur $\mathbf{Q}_p$.

— Comme x et y commutent, on a $\|(xy)^n\| = \|x^n y^n\| \leqslant \|x^n\| \cdot \|y^n\|$; on en déduit l'inégalité $|xy| \leqslant |x| \cdot |y|$.

— De même (grâce à l'ultramétricité de $|\ |_p$), on a

$$\|(x+y)^n\| \leqslant \sup_{0 \leqslant i \leqslant n} \|x^i\| \cdot \|y^{n-i}\| ;$$

on en déduit l'inégalité $|x+y| \leqslant \sup(|x|, |y|)$.

— L'ensemble S des $x \in \mathrm{F}$ tel que $\|x\| = 1$ est borné dans F (considérer par exemple l'action de x sur 1) et donc est compact. L'application $(x, y) \mapsto \|xy\|$ de $\mathrm{S} \times \mathrm{S}$ dans $\mathbf{R}$ est continue ; elle atteint donc son minimum C qui est strictement positif car $\|u\| = 0$ si et seulement si $u = 0$. Comme par ailleurs, on a $\|\lambda x\| = |\lambda|_p \cdot \|x\|$ si $\lambda \in \mathbf{Q}_p$ et $x \in \mathrm{F}$, on en déduit l'inégalité $\|xy\| \geqslant \mathrm{C} \cdot \|x\| \cdot \|y\|$ quels que soient $x, y \in \mathrm{F}$. En particulier, quel que soit $n \in \mathbf{N}$, on a $\|(xy)^n\| \geqslant \mathrm{C} \cdot \|x^n\| \cdot \|y^n\|$, ce qui, élevant le tout à la puissance $1/n$, nous fournit l'inégalité $|xy| \geqslant |x| \cdot |y|$.

— Finalement, l'équivalence « $|x| = 0 \Leftrightarrow x = 0$ » est une conséquence des égalités $|xy| = |x| \cdot |y|$ et $|1| = 1$.

Il reste à vérifier que l'on a $|x| = \left|\mathrm{N}_{\mathrm{F}/\mathbf{Q}_p}(x)\right|_p^{1/d}$. Pour cela, commençons par supposer que F est une extension galoisienne de $\mathbf{Q}_p$. Soit $\sigma \in \mathrm{Gal}(\mathrm{F}/\mathbf{Q}_p)$. L'application $x \mapsto |\sigma(x)|$ est une norme sur F dont la restriction à $\mathbf{Q}_p$ est la norme p-adique ; elle coïncide donc avec la

norme | | par unicité d'une telle norme. On obtient donc

$$\big|\mathrm{N}_{\mathrm{F}/\mathbf{Q}_p}(x)\big|_p = \big|\mathrm{N}_{\mathrm{F}/\mathbf{Q}_p}(x)\big| = \prod_{\sigma\in\mathrm{Gal}(\mathrm{F}/\mathbf{Q}_p)} |\sigma(x)| = |x|^d\,;$$

d'où le résultat dans le cas galoisien. Si F n'est pas galoisienne sur $\mathbf{Q}_p$, il suffit alors de choisir une extension finie galoisienne K de $\mathbf{Q}_p$ contenant F et de remarquer que la norme sur F est la restriction de la norme sur K par unicité et que l'on a $\mathrm{N}_{\mathrm{K}/\mathbf{Q}_p}(x) = \mathrm{N}_{\mathrm{F}/\mathbf{Q}_p}(x)^{[\mathrm{K}:\mathrm{F}]}$, si $x \in \mathrm{F}$.

Corollaire IV.2.7. *Si $\overline{\mathbf{Q}}_p$ est une clôture algébrique de $\mathbf{Q}_p$, Il existe une unique manière de prolonger | |$_p$ à $\overline{\mathbf{Q}}_p$. Si $x \neq 0$, alors $|x|_p \in p^{\mathbf{Q}}$.*

Exemple IV.2.8. Le polynôme $\mathrm{P}_n(\mathrm{X}) = \frac{(\mathrm{X}+1)^{p^n}-1}{(1+\mathrm{X})^{p^{n-1}}-1}$ est irréductible sur $\mathbf{Q}_p$; ses racines sont les $\varepsilon - 1$, ε parcourant les racines primitives p^n-ièmes de l'unité et on a $v_p(\varepsilon - 1) = 1/(p-1)p^{n-1}$.

Démonstration. On a $\mathrm{P}_n(\mathrm{X}) = \mathrm{X}^{(p-1)p^{n-1}}$ modulo p et $\mathrm{P}_n(0)$ est le quotient des dérivées de $(\mathrm{X}+1)^{p^n} - 1$ et $(1+\mathrm{X})^{p^{n-1}} - 1$ en $\mathrm{X} = 0$. On a donc $\mathrm{P}_n(0) = p$ et P_n vérifie le critère d'Eisenstein.

Définition IV.2.9. On note $\mathbf{C}_p$ le complété de $\overline{\mathbf{Q}}_p$ pour la norme | |$_p$. Si $x \in \mathbf{C}_p$, on note $v_p(x)$ l'élément de $\mathbf{Q} \cup \{+\infty\}$ tel que l'on ait $|x|_p = p^{-v_p(x)}$.

Théorème IV.2.10. *Si* K *est un corps ultramétrique algébriquement clos de caractéristique* 0*, son complété* $\widehat{\mathrm{K}}$ *est algébriquement clos.*

Démonstration. Soit $\mathrm{P}(\mathrm{X}) = \mathrm{X}^n + a_{n-1}\mathrm{X}^{n-1} + \cdots + a_0$ un polynôme unitaire irréductible de $\widehat{\mathrm{K}}[\mathrm{X}]$. Quitte à changer X en Y/α, ce qui multiplie a_i par α^{n-i}, on peut, en prenant $\alpha \in \widehat{\mathrm{K}}$ de norme assez petite, supposer que P est à coefficients entiers. Comme K est de caractéristique 0, les polynômes P et P$'$ sont premiers entre eux et il existe des polynômes U et V tels que l'on ait $\mathrm{UP} + \mathrm{VP}' = 1$.

Si $\mathrm{Q} \in \mathrm{K}[\mathrm{X}]$, notons $|\mathrm{Q}|_{\mathrm{G}}$ la norme de Gauss de Q, c'est-à-dire le maximum des valeurs absolues de ses coefficients. Soit $\varepsilon < \inf(1, 1/|\mathrm{U}|_{\mathrm{G}}, 1/|\mathrm{V}|_{\mathrm{G}}^2)$ et, si $0 \leqslant i \leqslant n-1$, soit $b_i \in \mathrm{K}$ tels que l'on ait $|b_i - a_i| \leqslant \varepsilon$. Soit $x_0 \in \mathrm{K}$ une racine du polynôme

$\mathrm{Q}(\mathrm{X}) = \mathrm{X}^n + b_{n-1}\mathrm{X}^{n-1} + \cdots + b_0$. On a $|x_0| \leqslant 1$ à cause de l'inégalité ultramétrique car $|b_i| \leqslant 1$ quel que soit i. D'autre part, on a $|\mathrm{U}(x_0)\mathrm{P}(x_0)| \leqslant \varepsilon|\mathrm{U}|_{\mathrm{G}} < 1$ et donc $|\mathrm{P}'(x_0)||\mathrm{V}(x_0)| = 1$, ce qui implique $|\mathrm{P}'(x_0)| \geqslant 1/|\mathrm{V}|_{\mathrm{G}}$, et donc $|\mathrm{P}(x_0)| \leqslant \varepsilon < |\mathrm{P}'(x_0)|^2$. Le lemme de Hensel permet de conclure au fait que l'équation $\mathrm{P}(x) = 0$ a une solution dans $\widehat{\mathrm{K}}$.

Corollaire IV.2.11. *$\mathbf{C}_p$ est algébriquement clos.*

Le corps $\mathbf{C}_p$ est donc un corps complet algébriquement clos qui jouera le rôle de $\mathbf{C}$ en p-adique. $\mathbf{C}$ et $\mathbf{C}_p$ sont deux corps algébriquement clos de même degré de transcendance sur $\mathbf{Q}$; ils sont donc abstraitement isomorphes en tant que corps (mais pas en tant que corps topologiques). Ce corps n'est pas encore assez gros (il est pourtant de dimension non dénombrable sur $\mathbf{Q}_p$) : Tate a démontré qu'il ne contenait pas d'analogue naturel de $2i\pi$. (Une indication de cette non existence de $2i\pi$ dans $\mathbf{C}_p$ est fournie par le corollaire IV.2.20.)

4. Représentants de Teichmüller

Lemme IV.2.12. *Soient* K *un corps ultramétrique complet et* L *une extension finie de* K*, alors k_{L} est une extension algébrique de k_{K} de degré $\leqslant [\mathrm{L} : \mathrm{K}]$.*

Démonstration. On a $\mathscr{O}_{\mathrm{K}} \cap \mathfrak{m}_{\mathrm{L}} = \mathfrak{m}_{\mathrm{K}}$ et donc k_{K} s'injecte dans k_{L}. Soient $\overline{\alpha}_1, \ldots, \overline{\alpha}_d$ des éléments de k_{L} formant une famille libre sur k_{K}. Choisissons pour chaque $i \in \{1, \ldots, d\}$ un élément α_i de $\mathscr{O}_{\mathrm{L}}$ dont l'image dans k_{L} est $\overline{\alpha}_i$. Supposons que les α_i forment une famille liée sur K et soit $(\lambda_1, \ldots, \lambda_d)$ une famille d'éléments non tous nuls de K tels que l'on ait $\lambda_1\alpha_1 + \cdots + \lambda_d\alpha_d = 0$. Quitte à diviser tous les λ_i par celui qui a la plus grande norme, on peut supposer qu'ils sont tous éléments de $\mathscr{O}_{\mathrm{K}}$ et que l'un d'entre eux est égal à 1, ce qui conduit à une contradiction quand on réduit modulo $\mathfrak{m}_{\mathrm{L}}$. Ceci permet de conclure.

Lemme IV.2.13. *Si* K *est un corps ultramétrique algébriquement clos, alors k_{K} est algébriquement clos.*

Démonstration. Soit $\overline{\mathrm{P}}(\mathrm{X}) \in k_{\mathrm{K}}[\mathrm{X}]$ unitaire de degré $n \geqslant 1$ et soit $\mathrm{P}(\mathrm{X}) \in \mathscr{O}_{\mathrm{K}}[\mathrm{X}]$ unitaire de degré n relevant $\overline{\mathrm{P}}$. Soit $\alpha \in \mathrm{K}$ une racine de P. On a $|\alpha| \leqslant 1$ à cause de l'inégalité ultramétrique et donc $\alpha \in \mathscr{O}_{\mathrm{K}}$ et l'image de α dans k_{K} est une racine de $\overline{\mathrm{P}}$, ce qui permet de conclure.

Lemme IV.2.14. *Si* K *est un corps ultramétrique et* $\widehat{\mathrm{K}}$ *dénote son complété, alors* $k_{\mathrm{K}} = k_{\widehat{\mathrm{K}}}$.

Démonstration. $\mathscr{O}_{\mathrm{K}} \cap \mathfrak{m}_{\widehat{\mathrm{K}}} = \{x \in \mathscr{O}_{\mathrm{K}} \mid |x| < 1\} = \mathfrak{m}_{\mathrm{K}}$ et donc l'application naturelle de k_{K} dans $k_{\widehat{\mathrm{K}}}$ est injective. D'autre part, comme $\mathscr{O}_{\mathrm{K}}$ est dense dans $\mathscr{O}_{\widehat{\mathrm{K}}}$, cette application est surjective, ce qui permet de conclure.

Corollaire IV.2.15. *Le corps résiduel de* $\mathbf{C}_p$ *est* $\overline{\mathbf{F}}_p$ *une clôture algébrique de* $\mathbf{F}_p$.

Démonstration. Les trois lemmes précédents montrent que ce corps résiduel est le même que celui de $\overline{\mathbf{Q}}_p$ et donc est algébriquement clos et algébrique sur $\mathbf{F}_p$, ce qui permet de conclure.

Proposition IV.2.16. *Si* $x \in \overline{\mathbf{F}}_p^*$, *il existe dans* $\mathscr{O}_{\mathbf{C}_p}$ *une unique racine de l'unité* $[x]$ *d'ordre premier à* p *dont l'image dans* $\overline{\mathbf{F}}_p$ *est* x.

Démonstration. Il existe n tel que x appartienne à $\mathbf{F}_{p^n}$. Les éléments de $\mathbf{F}_{p^n}^*$ sont solutions de l'équation $\mathrm{P}(\mathrm{X}) = \mathrm{X}^{p^n-1} - 1 = 0$. Or le polynôme $\mathrm{X}^{p^n-1} - 1$ a un discriminant premier à p [son discriminant est au signe près $\prod_{\eta^{p^n-1}=1} \mathrm{P}'(\eta) = \pm(p^n-1)^{p^n-1}$], ce qui signifie que toutes les images de ses racines sont distinctes modulo $\mathfrak{m}_{\mathbf{C}_p}$; on a donc une injection d'un ensemble à p^n-1 éléments dans un ensemble à p^n-1 éléments et donc la réduction modulo $\mathfrak{m}_{\mathbf{C}_p}$ est une bijection, ce qui permet de conclure.

Remarque IV.2.17. L'unicité de $[x]$ implique que $[xy] = [x] \cdot [y]$. En posant $[0] = 0$, on a fabriqué un système multiplicatif de représentants, appelés *représentants de Teichmüller*, de $\overline{\mathbf{F}}_p$ dans $\mathscr{O}_{\mathbf{C}_p}$.

Choisissons un morphisme de groupes de $\mathbf{Q}$ dans $\mathbf{C}_p^*$ envoyant 1 sur p. On notera p^r l'image de r par ce morphisme. Si $x \in \mathscr{O}_{\mathbf{C}_p}^*$,

on note $\omega(x)$ l'unique racine de l'unité d'ordre premier à p telle que $|x-\omega(x)|_p<1$. Si $\overline{x}$ désigne l'image de x dans $\overline{\mathbf{F}}_p$, on a $\omega(x)=[\overline{x}]$.

Proposition IV.2.18. *Les applications*

$$x\longmapsto p^{v_p(x)},\quad x\longmapsto\widetilde{\omega}(x)=\omega(xp^{-v_p(x)})$$

et

$$x\longmapsto\langle x\rangle=xp^{-v_p(x)}\widetilde{\omega}(x)^{-1}$$

sont des morphismes de groupes de $\mathbf{C}_p^*$ dans $p^{\mathbf{Q}}$, le groupe des racines de l'unité d'ordre premier à p et $\mathrm{B}(1,1^-)$ respectivement, et on a $x=p^{v_p(x)}\widetilde{\omega}(x)\langle x\rangle$.

Démonstration. Évident.

5. La fonction logarithme

On définit la fonction logarithme dans la boule $\mathrm{B}(1,1^-)$ par la formule $\log(x)=\sum_{n=1}^{+\infty}\frac{(-1)^{n-1}}{n}(x-1)^n$.

Lemme IV.2.19. *Si $|x|_p<1$ et $|y|_p<1$, alors*

$$\log((1+x)(1+y))=\log(1+x)+\log(1+y).$$

Démonstration. En tant que série formelle,

$$\log((1+\mathrm{X})(1+\mathrm{Y}))=\log(1+\mathrm{X})+\log(1+\mathrm{Y})$$

(il suffit de dériver). D'autre part, la série triple

$$\sum_{i_1+i_2+i_3\geqslant 1}\frac{(-1)^{i_1+i_2+i_3}(i_1+i_2+i_3)!}{(i_1+i_2+i_3)i_1!i_2!i_3!}x^{i_1+i_3}y^{i_2+i_3}$$

converge car le terme général tend vers 0 quand $i_1+i_2+i_3$ tend vers $+\infty$ et on peut réordonner les termes comme on veut, ce qui permet de conclure.

Corollaire IV.2.20. *Si ε est une racine de l'unité d'ordre une puissance de p, alors $\log(\varepsilon)=0$. Réciproquement, si $x\in\mathrm{B}(1,1^-)$ vérifie $\log x=0$, alors x est une racine de l'unité d'ordre une puissance de p.*

Démonstration. On a $\log x = 0 \Leftrightarrow x = 1$ si $|x-1|_p < p^{-1/(p-1)}$ car alors le seul terme de valuation maximale est le premier. D'autre part, si $x \in \mathrm{B}(1,1^-)$, la fonction $s \to x^s = \sum_{n=0}^{+\infty} \binom{s}{n}(x-1)^n$ est continue. On en déduit le fait que x^{p^n} tend vers 1 quand n tend vers $+\infty$, ce qui permet de conclure.

Proposition IV.2.21. *La fonction* log *a un unique prolongement noté* $\log_p$ (et appelé logarithme d'Iwasawa) *à* $\mathbf{C}_p^*$ *vérifiant les trois conditions suivantes.*

(i) $\log_p(xy) = \log_p(x) + \log_p(y)$;
(ii) $\log_p$ *est donné par la série sur* $\mathrm{B}(1,1^-)$;
(iii) $\log_p p = 0$.

Démonstration. On a vu que si l'on se fixe un morphisme $r \to p^r$ de $\mathbf{Q}$ dans $\mathbf{C}_p^*$, l'on pouvait écrire tout élément de $\mathbf{C}_p^*$ de manière unique sous la forme $p^{v_p(x)}\omega u$, où ω est une racine de l'unité d'ordre premier à p et $u \in \mathrm{B}(1,1^-)$; on doit donc avoir $\log x = \log u$, ce qui prouve l'unicité. L'existence résulte du fait que $x \to u$ est un morphisme de groupes.

Remarque IV.2.22. La condition $\log_p p = 0$ est naturelle d'un point de vue arithmétique à cause de la formule du produit, mais on pourrait fixer arbitrairement la valeur de $\log p$. (on dit que l'on fixe une branche du logarithme).

6. La fonction exponentielle

Notons $[x]$ la partie entière de x si $x \in \mathbf{R}$ (ne pas confondre avec un représentant de Teichmüller !).

Proposition IV.2.23. *Si* $n \in \mathbf{N}$, *alors*

$$v_p(n!) = \sum_{k=1}^{+\infty} \left[\frac{n}{p^k}\right] = \frac{n - \mathrm{S}_p(n)}{p-1},$$

où $\mathrm{S}_p(n)$ *désigne la somme des chiffres de l'écriture de* n *en base* p.

Démonstration. Soit a_k (resp. b_k) le cardinal de l'ensemble des entiers i vérifiant $1 \leqslant i \leqslant n$ et $v_p(i) = k$ (resp. $v_p(i) \geqslant k$). On a $b_k = \sum_{\ell \geqslant k} a_\ell$ et

$$v_p(n!) = \sum_{k=1}^{+\infty} k a_k = \sum_{k=1}^{+\infty} b_k = \sum_{k=1}^{+\infty} \Big[\frac{n}{p^k}\Big],$$

ce qui nous fournit la première égalité. La seconde est laissée en exercice.

Proposition IV.2.24. *La série* $\exp x = \sum_{n=0}^{+\infty} x^n/n!$ *converge si et seulement si* $|x|_p < p^{-1/(p-1)}$ *et induit un isomorphisme de groupes de la boule ouverte* $\mathrm{B}(0, p^{-1/(p-1)})$ *munie de l'addition sur la boule ouverte* $\mathrm{B}(1, p^{-1/(p-1)})$ *munie de la multiplication, inverse de l'application* log.

Démonstration. La détermination du rayon de convergence vient de ce qu'il y a une infinité de n tels que $\mathrm{S}_p(n) = 1$ (à savoir, les puissances de p). Le reste de la proposition est laissé en exercice.

Chapitre V. Fonctions d'une variable *p*-adique

V.1. Espaces de Banach *p*-adiques

Définition V.1.1. Un espace de Banach p-adique est un $\mathbf{Q}_p$-espace vectoriel E muni d'une norme ultramétrique $\| \ \|$ pour laquelle il est complet.

Hypothèse V.1.2. On dit que E vérifie l'hypothèse (N) si quel que soit $x \in \mathrm{E}$, il existe $\lambda \in \mathbf{Q}_p$ tel que $\|x\| = |\lambda|_p$.

Lemme V.1.3. *Si* $(\mathrm{E}, \| \ \|)$ *est un espace de Banach p-adique, alors on peut trouver une norme* $\| \ \|_1$ *sur* E *qui est équivalente à* $\| \ \|$ *et qui vérifie l'hypothèse* (N).

Démonstration. Si $x \in \mathrm{E}$, soit $v_p(x)$ l'élément de $\mathbf{R} \cup \{+\infty\}$ défini par $\|x\| = p^{-v_p(x)}$. et soit $\|x\|_1 = p^{-[v_p(x)]}$, où, si $v \in \mathbf{R}$, $[v]$ désigne la partie entière de v. Alors $\| \ \|_1$ est une norme ultramétrique sur E et on a de plus $\frac{1}{p}\|x\|_1 \leqslant \|x\| \leqslant \|x\|_1$.

Exemple V.1.4

(i) $\mathbf{C}_p$ muni de la norme p-adique.

(ii) Si I est un ensemble et E est un espace de Banach p-adique, soit $l_\infty(\mathrm{I},\mathrm{E})$ l'ensemble des suites bornées $(a_i)_{i\in\mathrm{I}}$ d'éléments de E. On munit $l_\infty(\mathrm{I},\mathrm{E})$ de la norme $\|\ \|_\infty$ définie par $\|(a_i)_{i\in\mathrm{I}}\|_\infty = \sup_{i\in\mathrm{I}}\|a_i\|$, ce qui en fait un espace de Banach p-adique.

(iii) Soit $l^0_\infty(\mathrm{I},\mathrm{E})$ le sous-espace de $l_\infty(\mathrm{I},\mathrm{E})$ des suites $(a_i)_{i\in\mathrm{I}}$ d'éléments de E tendant vers 0 suivant le filtre complémentaire des parties finies. C'est un espace de Banach p-adique comme sous-espace fermé d'un espace de Banach p-adique. C'est aussi l'adhérence dans $l_\infty(\mathrm{I},\mathrm{E})$ de l'espace des suites n'ayant qu'un nombre fini de termes non nuls. On note δ_i la suite dont tous les termes sont nuls sauf celui d'indice i qui est égal à 1.

(iv) Si X est un espace topologique compact et E est un espace de Banach p-adique, l'espace $\mathscr{C}(\mathrm{X},\mathrm{E})$ des applications continues de X dans E muni de la norme du sup. est un espace de Banach p-adique.

La théorie des espaces de Banach p-adiques est très loin d'être aussi riche que son homologue archimédienne ; elle se rapproche plutôt de celle des espaces de Hilbert réels. En particulier, la notion suivante remplace celle de base hilbertienne dans un espace de Hilbert.

Définition V.1.5. Soit E un espace de Banach p-adique. On dit qu'une famille bornée $(e_i)_{i\in\mathrm{I}}$ est une *base de Banach* de E si l'application de $l^0_\infty(\mathrm{I},\mathbf{Q}_p)$ dans E qui à $(a_i)_{i\in\mathrm{I}}$ associe $\sum_{i\in\mathrm{I}} a_i e_i$ est une isométrie. On dit que c'est une *pseudo-base de Banach* si cette application est seulement un isomorphisme d'espaces de Banach p-adiques.

Autrement dit, une famille $(e_i)_{i\in\mathrm{I}}$ est une base de Banach de E si et seulement si

(i) tout élément x de E peut s'écrire de manière unique sous la forme d'une série convergente $x = \sum_{i\in\mathrm{I}} a_i e_i$, où les a_i sont des éléments de $\mathbf{Q}_p$ tendant vers 0 suivant le filtre complémentaire des parties finies,

(ii) $\|x\| = \sup_{i\in\mathrm{I}} |a_i|$

C'est une pseudo-base de Banach si elle est bornée et vérifie la condition (i), ce qui implique, d'après le théorème de l'image ouverte, la propriété suivante.

(ii') il existe une constante $C \geqslant 1$ telle que l'on ait

$$C^{-1} \sup_{i\in I} |a_i| \leqslant \|x\| \leqslant C \sup_{i\in I} |a_i|.$$

Exemple V.1.6. Par définition, ou presque, les δ_i, pour $i \in I$, forment une base de Banach de $l^0_\infty(I, \mathbf{Q}_p)$.

Proposition V.1.7

(i) *Tout espace de Banach p-adique possède des pseudo-bases de Banach.*

(ii) *Un espace de Banach p-adique possède des bases de Banach si et seulement si il vérifie l'hypothèse* (N). *De plus, sous cette hypothèse, si on note* $E^0 = \{x \in E \mid \|x\| \leqslant 1\}$, *alors* $(e_i)_{i\in I}$ *est une base de Banach de* E *si et seulement si* $(\overline{e}_i)_{i\in I}$ *est une base algébrique du* $\mathbf{F}_p$*-espace vectoriel* $\overline{E} = E^0/pE^0$.

Démonstration. Le lemme V.1.3 implique que le (i) est une conséquence du (ii). Supposons donc que E vérifie l'hypothèse (N) et montrons que $(e_i)_{i\in I}$ est une base de Banach de E si et seulement si $(\overline{e}_i)_{i\in I}$ est une base algébrique du $\mathbf{F}_p$-espace vectoriel $\overline{E}$.

Soit $(e_i)_{i\in I}$ une famille d'éléments de E^0 telle que la famille $(\overline{e}_i)_{i\in I}$ soit une base du $\mathbf{F}_p$-espace vectoriel $\overline{E}$. Soient S un système de représentants de $\mathbf{F}_p$ dans $\mathbf{Z}_p$ contenant 0 et $s : \mathbf{F}_p \to S$ l'inverse de la réduction modulo p. Si $x \in E^0$, on peut écrire $\overline{x}$ comme une somme finie $\sum_{i\in I} a_i \overline{e}_i$, où les a_i sont des éléments de $\mathbf{F}_p$ presque tous nuls. Soit $s(x) = \sum_{i\in I} s(a_i) e_i$. Par construction, on a $x - s(x) \in pE^0$.

Si $x \in E^0$, définissons par récurrence une suite x_n d'éléments de E_0 par $x_0 = x$ et $x_{n+1} = \frac{1}{p}(x_n - s(x_n))$. On a alors $x = \sum_{n=0}^{k} p^n s(x_n) + p^{k+1} x_{k+1}$ quel que soit $k \in \mathbf{N}$. On peut écrire $s(x_n) = \sum_{i\in I} s_{n,i} e_i$, où les $s_{n,i}$ sont des éléments de S presque tous nuls, ce qui montre que si on pose $a_i = \sum_{n=0}^{+\infty} p^n s_{n,i}$, alors la suite des a_i tend vers 0 suivant le filtre complémentaire des parties finies. Ceci montre que l'application de $l^0_\infty(I, \mathbf{Q}_p)$ dans E qui à $(a_i)_{i\in I}$ associe $\sum_{i\in I} a_i e_i$ est surjective. Si $(a_i)_{i\in I}$ est un élément non nul de $l^0_\infty(I, \mathbf{Q}_p)$, alors quitte

à multiplier a par une puissance de p, on peut supposer $\|a\|_\infty = 1$ et le fait que les $(\overline{e}_i)_{i\in \mathrm{I}}$ forment une base de $\overline{\mathrm{E}}$ implique que $\sum_{i\in \mathrm{I}} a_i e_i \neq 0$ modulo p et donc que $1 \geqslant \|\sum_{i\in \mathrm{I}} a_i e_i\| > p^{-1}$. Comme on a supposé que E vérifie (N), ceci implique $\|\sum_{i\in \mathrm{I}} a_i e_i\| = 1$ et on en tire le fait que l'application qui à $(a_i)_{i\in \mathrm{I}}$ associe $\sum_{i\in \mathrm{I}} a_i e_i$ est une isométrie de $l^0_\infty(\mathrm{I}, \mathbf{Q}_p)$ sur E.

La réciproque est immédiate.

V.2. Mesures sur $\mathbf{Z}_p$

1. Fonctions continues sur $\mathbf{Z}_p$

Soit E un espace de Banach p-adique et $\mathscr{C}^0(\mathbf{Z}_p, \mathrm{E})$ l'ensemble des fonctions continues de $\mathbf{Z}_p$ dans E. Comme $\mathbf{Z}_p$ est compact, une fonction continue sur $\mathbf{Z}_p$ est bornée. Ceci permet de munir $\mathscr{C}^0(\mathbf{Z}_p, \mathrm{E})$ de la norme $\|\ \|_\infty$ définie par $\|f\|_\infty = \sup_{x\in\mathbf{Z}_p} |f(x)|_p$, ce qui en fait un espace de Banach p-adique.

Si $n \in \mathbf{N}$, soit $\binom{x}{n}$ le polynôme défini par

$$\binom{x}{n} = \begin{cases} 1 & \text{si } n = 0 \\ \dfrac{x(x-1)\dots(x-n+1)}{n!} & \text{si } n \geqslant 1. \end{cases}$$

Proposition V.2.1. $\|\binom{x}{n}\|_\infty = 1$.

Démonstration. On a $\binom{n}{n} = 1$ et donc $\|\binom{x}{n}\|_\infty \geqslant 1$. D'autre part, $\binom{n+k}{n}$ est le nombre de manière de choisir n objets parmi $n+k$ et est donc entier. On en déduit le fait que $|\binom{n+k}{n}|_p \leqslant 1$ quel que soit $k \in \mathbf{N}$ et comme $n + \mathbf{N}$ est dense dans $\mathbf{Z}_p$, cela implique que $|\binom{x}{n}|_p \leqslant 1$ quel que soit $x \in \mathbf{Z}_p$ et permet de conclure.

On définit la n-ième dérivée discrète $f^{[n]}$ d'une fonction f par récurrence à partir des formules $f^{[0]} = f$ et $f^{[n+1]}(x) = f^{[n]}(x+1) - f^{[n]}(x)$, et on pose $a_n(f) = f^{[n]}(0)$. On a aussi

$$f^{[n]}(x) = \sum_{i=0}^{n} (-1)^i \binom{n}{i} f(x+n-i)$$

et

$$a_n(f) = \sum_{i=0}^{n} (-1)^i \binom{n}{i} f(n-i).$$

Théorème V.2.2 **(Mahler)**

(i) *Si* $f \in \mathscr{C}^0(\mathbf{Z}_p, \mathbf{C}_p)$, *alors*

a) $\lim_{n\to+\infty} a_n(f) = 0$,

b) $\sum_{n=0}^{+\infty} a_n(f)\binom{x}{n} = f(x)$ *quel que soit* $x \in \mathbf{Z}_p$.

(ii) *L'application* $f \mapsto a(f) = (a_n(f))_{n\in\mathbf{N}}$ *est une isométrie de* $\mathscr{C}^0(\mathbf{Z}_p, \mathbf{C}_p)$ *sur* $l^0_\infty(\mathbf{N}, \mathbf{C}_p)$.

Corollaire V.2.3. *Les* $\binom{x}{n}$ *pour* $n \in \mathbf{N}$ *forment une base de Banach de* $\mathscr{C}^0(\mathbf{Z}_p, \mathbf{Q}_p)$.

Démonstration. Le corollaire est immédiat. Pour démontrer le théorème V.2.2, commençons par remarquer que l'on a $\|a(f)\|_\infty \leqslant \sup_{k\in\mathbf{N}} |f(k)|_p \leqslant \|f\|_\infty$, et que l'application $f \mapsto a(f)$ est injective, car $a(f) = 0$ implique $f(k) = 0$ quel que soit $k \in \mathbf{N}$, et $\mathbf{N}$ est dense dans $\mathbf{Z}_p$.

Si $a = (a_n)_{n\in\mathbf{N}} \in l^0_\infty(\mathbf{N}, \mathbf{C}_p)$, la série $\sum_{n=0}^{+\infty} a_n\binom{x}{n}$ converge normalement dans $\mathscr{C}^0(\mathbf{Z}_p, \mathbf{C}_p)$ en vertu de la proposition V.2.1 et de l'ultramétricité de la norme p-adique ; on note f_a la somme de cette série et on a $\|f_a\|_\infty \leqslant \|a\|_\infty$. D'autre part, comme $\binom{x+1}{n+1} - \binom{x}{n+1} = \binom{x}{n}$, une récurrence immédiate nous fournit la formule $f_a^{[k]}(x) = \sum_{n=0}^{+\infty} a_{n+k}\binom{x}{n}$ et donc $a(f_a) = a$. Il ressort de la discussion précédente que $a \to f_a$ est une isométrie de $l^0_\infty(\mathbf{N}, \mathbf{C}_p)$ sur le sous-espace B de $\mathscr{C}^0(\mathbf{Z}_p, \mathbf{C}_p)$ des fonctions continues f vérifiant $\lim_{n\to+\infty} a_n(f) = 0$. Pour démontrer le théorème, il suffit donc de prouver le (i) a) ou, autrement dit, $\mathrm{B} = \mathscr{C}^0(\mathbf{Z}_p, \mathbf{C}_p)$. Nous allons donner deux démonstrations de ce fait.

Première démonstration. Nous aurons besoin du lemme suivant.

Lemme V.2.4. *Si* k *est un entier* $\geqslant 1$, *alors* $\binom{p^k}{i}$ *est divisible par* p *si* $1 \leqslant i \leqslant p^k - 1$.

Démonstration. Dériver $(1 + \mathrm{X})^{p^k}$.

Lemme V.2.5. *Si* $f \in \mathscr{C}^0(\mathbf{Z}_p, \mathbf{C}_p)$, *il existe* $k \in \mathbf{N}$ *tel que* $\|f^{[p^k]}\|_\infty \leqslant p^{-1}\|f\|_\infty$.

Démonstration. Comme $\mathbf{Z}_p$ est compact, f est uniformément continue et il existe $k \in \mathbf{N}$ tel que l'on ait $|f(x+p^k)-f(x)|_p \leqslant p^{-1}\|f\|_\infty$ quel que soit $x \in \mathbf{Z}_p$. Maintenant, on a

$$f^{[p^k]}(x) = f(x+p^k) - f(x) + \Big(\sum_{i=1}^{p^k-1}(-1)^i\binom{p^k}{i}f(x+p^k-i)\Big) + (1+(-1)^{p^k})f(x).$$

D'après le lemme V.2.4, tous les termes de la somme $\sum_{i=1}^{p^k-1}$ ont une norme $\leqslant p^{-1}\|f\|_\infty$ et $(1+(-1)^{p^k})f(x)$ est nul si $p \neq 2$ et de norme $\leqslant p^{-1}\|f\|_\infty$ si $p=2$. Comme on a choisi k de telle sorte que $|f(x+p^k)-f(x)|_p \leqslant p^{-1}\|f\|_\infty$ quel que soit $x \in \mathbf{Z}_p$, on a $\|f^{[p^k]}\|_\infty \leqslant p^{-1}\|f\|_\infty$, ce qui permet de conclure.

Une utilisation répétée du lemme précédent permet de montrer que, si $\varepsilon > 0$ et si $f \in \mathscr{C}^0(\mathbf{Z}_p, \mathbf{C}_p)$, alors il existe $k \in \mathbf{N}$ tel que $\|f^{[p^k]}\|_\infty \leqslant \varepsilon$. Comme $|a_n(f)| \leqslant \|f^{[p^k]}\|_\infty$ si $n \geqslant p^k$, cela montre que $a_n(f)$ tend vers 0 quand n tend vers $+\infty$.

Deuxième démonstration. Si $x \in \mathbf{C}_p$ vérifie $|x-1|_p < 1$, la série $f_x(s) = \sum_{n=0}^{+\infty}\binom{s}{n}(x-1)^n$ converge normalement d'après la proposition V.2.1 et définit donc une fonction continue $f_x(s)$ de $s \in \mathbf{Z}_p$. D'autre part, si $k \in \mathbf{N}$, on a $f_x(k) = x^k$ comme on le constate facilement, ce qui nous permet de noter de manière plus parlante $s \mapsto x^s$ la fonction $s \mapsto f_x(s)$. On a $x^{s+t} = x^s x^t$ quels que soient $s, t \in \mathbf{Z}_p$ car cette formule est vraie si $s, t \in \mathbf{N}$ et $\mathbf{N}^2$ est dense dans $\mathbf{Z}_p^2$. Si on regarde ce que donne cette formule sur les coefficients binomiaux (en écrivant x^s, x^t et x^{s+t} comme des séries en $x-1$), on obtient la formule $\binom{s+t}{n} = \sum_{i=0}^{n}\binom{s}{i}\binom{t}{n-i}$, formule que l'on peut aussi démontrer directement.

Exemple V.2.6. On a $(\sum_{n=0}^{+\infty}\binom{1/2}{n}x^n)^2 = 1+x$ si $|x|_p < 1$ et $p \neq 2$. Par exemple, si $x = 7/9$, la série ci-dessus converge dans $\mathbf{R}$ et dans $\mathbf{Q}_7$, mais dans $\mathbf{R}$ la limite est la racine positive de $16/9$ soit $4/3$ alors que dans $\mathbf{Q}_7$, c'est la racine de $16/9$ congrue à 1 modulo 7, à savoir $-4/3$.

Un cas particulièrement utile est celui où x est une racine de l'unité d'ordre une puissance p. Si $x^{p^n} = 1$, on a $x^{s+p^n k} = x^s$ quel que

soit $k \in \mathbf{N}$ et donc $x^s = x^t$ si $t \in s + p^n\mathbf{Z}_p$, ce qui fait que la fonction x^s est localement constante. D'autre part, on a $\sum_{\varepsilon^{p^m}=1} \varepsilon^x = \begin{cases} p^m & \text{si } x \in p^n\mathbf{Z}_p \\ 0 & \text{sinon} \end{cases}$ et la fonction caractéristique de $a + p^m\mathbf{Z}_p$ est donc

$$\frac{1}{p^m} \sum_{\varepsilon^{p^m}=1} \varepsilon^{x-a} = \sum_{n=0}^{+\infty} \Big(\frac{1}{p^m} \sum_{\varepsilon^{p^m}=1} \varepsilon^{-a}(\varepsilon - 1)^n \Big) \binom{x}{n}$$

Lemme V.2.7. *Si f est localement constante, alors $f \in \mathrm{B}$.*

Démonstration. Si f est localement constante, cela signifie que quel que soit $a \in \mathbf{Z}_p$, il existe $n_a \in \mathbf{N}$ tel que f soit constante sur $a + p^{n_a}\mathbf{Z}_p$. Comme $\mathbf{Z}_p$ est compact, on peut extraire du recouvrement de $\mathbf{Z}_p$ par les $a + p^{n_a}\mathbf{Z}_p$ un sous-recouvrement fini et il existe donc $m \in \mathbf{N}$ tel que f soit constante sur $a + p^m\mathbf{Z}_p$ quel que soit $a \in \mathbf{Z}_p$. On a donc $f(x) = \sum_{i=0}^{p^m-1} f(i)\mathbf{1}_{i+p^m\mathbf{Z}_p}$ et on peut se ramener par linéarité au cas où f est la fonction caractéristique de $i + p^m\mathbf{Z}_p$. La formule ci-dessus nous donne $a_n(\mathbf{1}_{i+p^m\mathbf{Z}_p}) = \frac{1}{p^m} \sum_{\varepsilon^{p^m}=1} \varepsilon^{-i}(\varepsilon - 1)^n$ et comme $|\varepsilon - 1|_p < 1$ si $\varepsilon^{p^m} = 1$, on en déduit le fait que $a_n(\mathbf{1}_{i+p^m\mathbf{Z}_p})$ tend vers 0 quand n tend vers $+\infty$, ce qui permet de conclure.

Lemme V.2.8. *Les fonctions localement constantes sont denses dans $\mathscr{C}^0(\mathbf{Z}_p, \mathbf{C}_p)$.*

Démonstration. $\mathbf{Z}_p$ étant compact, une fonction continue sur $\mathbf{Z}_p$ est uniformément continue. Soient $f \in \mathscr{C}^0(\mathbf{Z}_p, \mathbf{C}_p)$ et $\eta > 0$. Il existe $m \in \mathbf{N}$ tel que si $|x - y|_p \leqslant p^{-m}$, alors $|f(x) - f(y)|_p \leqslant \eta$. Soit f_m la fonction localement constante définie par $f_m(x) = \sum_{i=0}^{p^m-1} f(i)\mathbf{1}_{i+p^m\mathbf{Z}_p}$. Si $x \in \mathbf{Z}_p$, il existe $i \in \{0, \ldots, p^m - 1\}$ tel que $x \in i + p^m\mathbf{Z}_p$ et $|f(x) - f_m(x)|_p = |f(x) - f(i)|_p \leqslant \eta$ par construction de m ; on a donc $\|f - f_m\|_\infty \leqslant \eta$, ce qui permet de conclure.

L'application $f \mapsto a(f)$ étant continue, B est fermé dans l'espace $\mathscr{C}^0(\mathbf{Z}_p, \mathbf{C}_p)$. Par ailleurs, B contient l'espace des fonctions localement constantes d'après le lemme V.2.7 et comme cet espace est dense dans $\mathscr{C}^0(\mathbf{Z}_p, \mathbf{C}_p)$ d'après le lemme V.2.8, on en déduit l'égalité $\mathrm{B} = \mathscr{C}^0(\mathbf{Z}_p, \mathbf{C}_p)$ que l'on cherchait à établir.

2. Mesures sur $\mathbf{Z}_p$

Définition V.2.9. Une mesure sur $\mathbf{Z}_p$ à valeurs dans $\mathbf{C}_p$ est une forme linéaire continue μ sur $\mathscr{C}^0(\mathbf{Z}_p, \mathbf{C}_p)$. On écrira $\mu(f)$ sous la forme plus parlante $\int_{\mathbf{Z}_p} f(x)\mu(x)$ ou simplement sous la forme $\int_{\mathbf{Z}_p} f\mu$. Si μ est une mesure, on note $\|\mu\|_\infty$ la norme de μ en tant qu'opérateur, c'est-à-dire le sup. de $|\int_{\mathbf{Z}_p} f\mu|_p$, avec $\|f\|_\infty \leqslant 1$.

A une mesure, on associe deux séries formelles

$$\mathscr{A}_\mu(\mathrm{T}) = \sum_{n=0}^{+\infty} \mathrm{T}^n \int_{\mathbf{Z}_p} \binom{x}{n} \mu(x) \quad \text{et} \quad \mathscr{L}_\mu(t) = \sum_{n=0}^{+\infty} \frac{t^n}{n!} \int_{\mathbf{Z}_p} x^n \mu(x)$$

appelées respectivement *transformée d'Amice* et *transformée de Laplace* de μ. On a $\mathscr{L}_\mu(t) = \mathscr{A}_\mu(e^t - 1)$ et formellement

$$\mathscr{A}_\mu(\mathrm{T}) = \int_{\mathbf{Z}_p} (1+\mathrm{T})^x \mu(x) \quad \text{et} \quad \mathscr{L}_\mu(t) = \int_{\mathbf{Z}_p} e^{tx} \mu(x).$$

Lemme V.2.10. *Si $|z|_p < 1$, alors $\int_{\mathbf{Z}_p} (1+z)^x \mu(x) = \mathscr{A}_\mu(z)$*

Démonstration. La suite de fonctions $\sum_{n=0}^{\mathrm{N}} \binom{x}{n} z^n$ converge uniformément sur $\mathbf{Z}_p$ vers $(1+z)^x$ (le reste est plus petit en norme $\|\ \|_\infty$ que $|z|_p^{n+1}$) ; on peut donc échanger intégration et somme, d'où le résultat.

Théorème V.2.11. *L'application qui à une mesure μ associe sa transformée d'Amice est une isométrie de l'espace des mesures muni de la norme ci-dessus sur l'espace des séries formelles à coefficients bornés muni de la norme du sup. des normes des coefficients.*

Démonstration. Si μ est une mesure, alors $\mathscr{A}_\mu(\mathrm{T})$ est à coefficients bornés car $|\int_{\mathbf{Z}_p} \binom{x}{n} \mu(x)| \leqslant \|\mu\|_\infty \|\binom{x}{n}\|_\infty \leqslant \|\mu\|_\infty$ et la réciproque est une conséquence immédiate du théorème de Mahler : à une suite $(b_n)_{n\in\mathbf{N}}$ bornée, on associe la mesure μ définie par $\mu(f) = \sum_{n=0}^{+\infty} b_n a_n(f)$ si f est une fonction continue sur $\mathbf{Z}_p$.

Remarque V.2.12. Le théorème précédent permet de construire une mesure à partir d'une série entière dont les coefficients sont bornés. On peut aussi utiliser le fait que les fonctions localement constantes sont denses dans les fonctions continues, ce qui permet de définir une

mesure en ne connaissant que les intégrales $\int_{\mathbf{Z}_p} \mathbf{1}_{a+p^n\mathbf{Z}_p}\mu(x)$ pour $a \in \mathbf{Z}_p$ et $n \in \mathbf{N}$. L'intégrale $\int_{\mathbf{Z}_p} \mathbf{1}_{a+p^n\mathbf{Z}_p}\mu(x)$ sera notée $\mu(a+p^n\mathbf{Z}_p)$ et appelée la mesure de $a+p^n\mathbf{Z}_p$. Comme $a+p^n\mathbf{Z}_p$ est la réunion disjointe des $a+jp^n+p^{n+1}\mathbf{Z}_p$ pour $0 \leqslant j \leqslant p-1$, on a $\mu(a+p^n\mathbf{Z}_p) = \sum_{j=0}^{p-1} \mu(a+jp^n+p^{n+1}\mathbf{Z}_p)$. De plus, comme $\|\mathbf{1}_{a+p^n\mathbf{Z}_p}\|_\infty = 1$, les $\mu(a+p^n\mathbf{Z}_p)$ sont bornés. Si f est une fonction continue sur $\mathbf{Z}_p$, alors $\int_{\mathbf{Z}_p} f(x)\mu(x)$ est donné par la formule

$$\int_{\mathbf{Z}_p} f(x)\mu(x) = \lim_{n\to+\infty} \sum_{a=0}^{p^n-1} f(a)\mu(a+p^n\mathbf{Z}_p),$$

formule qui ressemble beaucoup à une somme de Riemann.

Remarque V.2.13. Il y a 2 topologies naturelles que l'on peut mettre sur $\mathbf{Z}_p[\![\mathrm{T}]\!]$:

— la topologie donnée par le sup. des normes des coefficients,

— la topologie de la limite projective $\mathbf{Z}_p[\![\mathrm{T}]\!] = \varprojlim \mathbf{Z}_p[\mathrm{T}]/(p,\mathrm{T})^n$ qui en fait un anneau compact, limite projective d'anneaux finis.

Proposition V.2.14. *La première correspond à la topologie de la convergence forte sur les mesures et la seconde à la topologie de la convergence faible.* (*i.e.* une suite $(\mu_n)_{n\in\mathbf{N}}$ de mesures converge si pour toute fonction continue f, la suite $\int_{\mathbf{Z}_p} f(x)\mu_n(x)$ converge).

Démonstration. Exercice.

3. Exemples de mesures et opérations sur les mesures

i) La mesure de Haar : On cherche une mesure invariante par translation sur $\mathbf{Z}_p$, ce qui implique $\mu(a+p^n\mathbf{Z}_p) = \frac{1}{p^n}\mu(\mathbf{Z}_p)$ quels que soient $a \in \mathbf{Z}_p$ et $n \in \mathbf{N}$. Ceci n'est possible que si $\mu(\mathbf{Z}_p) = 0$ (et donc si $\mu = 0$) car les $\mu(a+p^n\mathbf{Z}_p)$ doivent être bornés en norme par $\|\mu\|_\infty$. Il n'y a donc pas de mesure de Haar en p-adique et donc aucun moyen canonique d'associer une mesure à une fonction.

ii) Masses de Dirac : si $a \in \mathbf{Z}_p$, soit δ_a la masse de Dirac en a, c'est-à-dire la mesure qui à f associe $f(a)$. Sa transformée d'Amice est $(1+\mathrm{T})^a$ et sa transformée de Laplace est e^{at}.

iii) Multiplication d'une mesure par une fonction continue : si μ est une mesure et $g \in \mathscr{C}^0(\mathbf{Z}_p, \mathbf{Q}_p)$, on définit la mesure $g\mu$ par la formule $\int_{\mathbf{Z}_p} f(x)(g\mu(x)) = \int_{\mathbf{Z}_p} fg(x)\mu(x)$.

— Multiplication par x. On a

$$x \cdot \binom{x}{n} = ((x-n)+n)\binom{x}{n} = (n+1)\binom{x}{n+1} + n\binom{x}{n}.$$

On en déduit la formule

$$\mathscr{A}_{x\mu}(\mathrm{T}) = (1+\mathrm{T})\frac{d}{d\mathrm{T}}\mathscr{A}_\mu(\mathrm{T}).$$

— Multiplication par z^x si $|z-1|_p < 1$. D'après le lemme V.2.10, si $|y-1|_p < 1$, et si λ est une mesure, alors $\int_{\mathbf{Z}_p} y^x \lambda(x) = \mathscr{A}_\lambda(y-1)$. Appliquant ceci à $\lambda = z^x\mu$, on obtient $\mathscr{A}_\lambda(y-1) = \mathscr{A}_\mu(yz-1)$ quel que soit $y \in \mathrm{B}(1, 1^-)$. On en déduit la formule

$$\mathscr{A}_{z^x\mu}(\mathrm{T}) = \mathscr{A}_\mu((1+\mathrm{T})z - 1).$$

iv) Restriction d'une mesure à un ouvert compact : Si X est un ouvert compact de $\mathbf{Z}_p$ (réunion finie d'ouvert du type $a + p^n\mathbf{Z}_p$), la fonction caractéristique 1_{X} de X est continue sur $\mathbf{Z}_p$. Si μ est une mesure sur $\mathbf{Z}_p$, la mesure $1_{\mathrm{X}}\mu$ est appelée la restriction de μ à X et est notée $\mathrm{Res}_{\mathrm{X}}(\mu)$. On écrira indifféremment $\int_{\mathbf{Z}_p} f\mathrm{Res}_{\mathrm{X}}(\mu)$ ou $\int_{\mathbf{Z}_p} 1_{\mathrm{X}}(x)f(x)\mu(x)$ ou en général $\int_{\mathrm{X}} f(x)\mu(x)$. Comme la fonction caractéristique de $a+p^n\mathbf{Z}_p$ est $\frac{1}{p^n}\sum_{\varepsilon^{p^n}=1}\varepsilon^{x-a}$, on déduit de la formule ci-dessus, l'identité

$$\mathscr{A}_{\mathrm{Res}_{a+p^n\mathbf{Z}_p}(\mu)}(\mathrm{T}) = \frac{1}{p^n}\sum_{\varepsilon^{p^n}=1}\varepsilon^{-a}\mathscr{A}_\mu((1+\mathrm{T})\varepsilon - 1).$$

v) Convolution des mesures :

Lemme V.2.15. *Si f est une fonction continue sur $\mathbf{Z}_p$ et si μ est une mesure sur $\mathbf{Z}_p$, la fonction $\mu * f$ définie par $\int_{\mathbf{Z}_p} f(x+y)\mu(x)$ est une fonction continue de $y \in \mathbf{Z}_p$ et l'application $f \to \mu * f$ est une application linéaire continue de $\mathscr{C}^0(\mathbf{Z}_p, \mathbf{Q}_p)$ dans lui-même.*

Démonstration. La continuité de $\mu * f$ est une conséquence de l'uniforme continuité de f et d'autre part, $\|\mu * f\|_\infty \leqslant \|\mu\|\|f\|_\infty$ d'où la continuité de $f \to \mu * f$.

Ceci nous permet, si λ et μ sont deux mesures sur $\mathbf{Z}_p$ de définir leur convolée $\lambda * \mu$ par la formule $\lambda * \mu(f) = \lambda(\mu * f)$ ou encore

$$\int_{\mathbf{Z}_p} f(x)\lambda * \mu(x) = \int_{\mathbf{Z}_p} \Big(\int_{\mathbf{Z}_p} f(x+y)\mu(x) \Big) \lambda(y).$$

Un calcul immédiat nous donne $\delta_a * \delta_b = \delta_{a+b}$.

D'autre part, si $|z|_p < 1$, alors

$$\int_{\mathbf{Z}_p} \Big(\int_{\mathbf{Z}_p} (1+z)^{x+y}\mu(x) \Big) \lambda(y) = \mathscr{A}_\mu(z) \int_{\mathbf{Z}_p} (1+z)^y \lambda(y) = \mathscr{A}_\mu(z)\mathscr{A}_\lambda(z);$$

On en déduit la formule

$$\mathscr{A}_{\lambda*\mu}(\mathrm{T}) = \mathscr{A}_\lambda(\mathrm{T})\mathscr{A}_\mu(\mathrm{T})$$

ce qui prouve que les mesures munies de la convolution forment une algèbre commutative et associative et que $\mu \to \mathscr{A}_\mu$ est un isomorphisme d'algèbres de l'espace des mesures sur celui des séries entières bornées.

4. Autre point de vue sur les mesures

Soit Γ un groupe commutatif profini (c'est-à-dire que l'on peut trouver une suite de sous-groupes ouverts $\Gamma_n \supset \Gamma_{n+1}$ de Γ tels que Γ/Γ_n soit un groupe fini et l'application naturelle de Γ dans la limite projective des Γ/Γ_n soit un isomorphisme de groupes topologiques. Le groupe Γ est alors compact et les Γ_n forment une base de voisinages de 0). Exemples de base $\mathbf{Z}_p = \varprojlim \mathbf{Z}/p^n\mathbf{Z}$ ou $\mathbf{Z}_p^* = \varprojlim (\mathbf{Z}/p^n\mathbf{Z})^*$.

Comme $\Gamma_{n+1} \subset \Gamma_n$, on a une application naturelle de l'algèbre de groupe $\mathbf{Z}_p[\Gamma/\Gamma_{n+1}]$ sur $\mathbf{Z}_p[\Gamma/\Gamma_n]$ et on peut donc considérer l'algèbre de groupe complétée $\mathbf{Z}_p[\![\Gamma]\!]$ limite projective des $\mathbf{Z}_p[\Gamma/\Gamma_n]$. Par construction, on dispose d'une application $\pi_n : \mathbf{Z}_p[\![\Gamma]\!] \to \mathbf{Z}_p[\Gamma/\Gamma_n]$.

Soit $\mu \in \mathbf{Z}_p[\![\Gamma]\!]$. Si f est une fonction localement constante sur Γ, comme Γ est compact, il existe $n \in \mathbf{N}$ tel que f soit constante modulo Γ_n. On vérifie facilement que si $\pi_n(\mu) = \sum_{g\in\Gamma/\Gamma_n} a_g\delta_g$ et si f est constante modulo Γ_n, alors $\sum_{g\in\Gamma/\Gamma_n} a_g f(g)$ ne dépend pas du choix de n, ce qui nous permet de voir μ comme une forme linéaire sur l'espace des fonctions localement constantes à valeurs dans $\mathbf{Q}_p$. Comme $\pi_n(\mu) \in \mathbf{Z}_p[\Gamma/\Gamma_n]$, on a $|\mu(f)|_p \leqslant \|f\|_\infty$ et donc μ peut se prolonger au complété de l'espace des fonctions localement

constantes sur Γ pour la norme $\|\ \|_\infty$, et comme les fonctions localement constantes sont continues et Γ est compact, ce complété est l'espace des fonctions continues sur Γ à valeurs dans $\mathbf{Q}_p$ et μ peut donc être vu comme une mesure sur Γ.

Réciproquement, si μ est une mesure sur Γ de norme 1, on voit μ comme un élément de $\mathbf{Z}_p[\![\Gamma]\!]$ en posant

$$\pi_n(\mu) = \sum_{g\in\Gamma/\Gamma_n} \Big(\int_{g+\Gamma_n} \mu\Big)\delta_g \in \mathbf{Z}_p[\Gamma/\Gamma_n].$$

La convolution des mesures correspondant au produit dans l'algèbre de groupe complétée.

Par exemple, dans le cas $\Gamma = \mathbf{Z}_p$ et $\Gamma_n = p^n\mathbf{Z}_p$, la transformée d'Amice induit un isomorphisme de $\mathbf{Z}_p[\![\Gamma]\!]$ sur $\mathbf{Z}_p[\![\mathrm{T}]\!]$ et elle envoie δ_a sur $(1+\mathrm{T})^a$ et donc

$$\mathbf{Z}_p[\Gamma/\Gamma_n] = \mathbf{Z}_p[\![\Gamma]\!]/(\delta_{p^n}-\delta_0) \cong \mathbf{Z}_p[\![\mathrm{T}]\!]/((1+\mathrm{T})^{p^n}-1).$$

V.3. Distributions sur $\mathbf{Z}_p$

1. Fonctions localement analytiques sur $\mathbf{Z}_p$

Si $h\in\mathbf{N}$, soit LA_h l'espace des fonctions de $\mathbf{Z}_p$ dans $\mathbf{Q}_p$ analytiques sur $a+p^h\mathbf{Z}_p$ quel que soit $a\in\mathbf{Z}_p$. Si $f\in\mathrm{LA}_h$, alors, quel que soit $x_0\in\mathbf{Z}_p$, on peut développer f sous la forme

$$f(x) = \sum_{k=0}^{+\infty} a_k(x_0)\big(\frac{x-x_0}{p^h}\big)^k,$$

où $a_k(x_0)$ est une suite d'éléments de $\mathbf{Q}_p$ tendant vers 0 quand k tend vers $+\infty$. Si $f\in\mathrm{LA}_h$ et $x_0\in\mathbf{Z}_p$, on pose $\|f\|_{h,x_0} = \max_k(|a_k(x_0)|_p)$ et $\|f\|_{\mathrm{LA}_h} = \sup_{x_0\in\mathbf{Z}_p}\|f\|_{h,x_0}$. Il résulte de la théorie des séries formelles que l'on a aussi

$$\|f\|_{h,x_0} = \sup_{z\in\mathscr{O}_{\mathbf{C}_p}} |f(x_0+p^hz)|_p$$

et donc que $\|f\|_{h,x_0} = \|f\|_{h,x_1}$ si $x_1-x_0\in p^h\mathbf{Z}_p$ et que, si S est un système de représentants de $\mathbf{Z}_p$ modulo $p^h\mathbf{Z}_p$, alors $\|f\|_{\mathrm{LA}_h} = \sup_{x\in\mathrm{S}}\|f\|_{h,x}$.

Lemme V.3.1. *Si $f \in \mathrm{LA}_h$, alors la dérivée n-ième de f appartient à LA_h et de plus, $\|p^{nh} f^{(n)}/n!\|_{\mathrm{LA}_h} \leqslant \|f\|_{\mathrm{LA}_h}$ quel que soit $n \in \mathbf{N}$ et la suite de terme général $\|p^{nh} f^{(n)}/n!\|_{\mathrm{LA}_h}$ tend vers 0 quand n tend vers $+\infty$.*

Démonstration. Soit S un système de représentants de $\mathbf{Z}_p$ modulo $p^h\mathbf{Z}_p$. Si $x_0 \in \mathbf{Z}_p$, alors $p^{nh} f^{(n)}(x)/n! = \sum_{k=0}^{+\infty} \binom{n+k}{n} a_{n+k}(x_0)(\frac{x-x_0}{p^h})^k$. Comme $\binom{n+k}{n}$ est entier, on en déduit l'inégalité

$$\Big\| \frac{p^{nh} f^{(n)}}{n!} \Big\|_{h,x_0} \leqslant \sup_{k \geqslant n} |a_k(x_0)|_p \leqslant \sup_{k \geqslant 0} |a_k(x_0)|_p \leqslant \|f\|_{h,x_0}.$$

On en déduit la majoration que l'on cherche et le fait que la suite de terme général $\|p^{nh} f^{(n)}/n!\|_{\mathrm{LA}_h}$ tend vers 0 est une conséquence du fait que la suite $a_k(x_0)$ tend vers 0.

On peut écrire tout entier n de manière unique sous la forme $n = mp^h - i$ avec $1 \leqslant i \leqslant p^h$ et $m \geqslant 1$. Soit alors $e_n(x)$ la fonction $\mathbf{1}_{-i+p^h\mathbf{Z}_p}(x)(\frac{x+i}{p^h})^{m-1}$

Lemme V.3.2. *Les e_n pour $n \in \mathbf{N}$ forment une base de Banach de LA_h.*

Démonstration. Il suffit de revenir à la définition de la norme sur LA_h et d'utiliser le fait que $\{-i,\ 1 \leqslant i \leqslant p^h\}$ est un système de représentants de $\mathbf{Z}_p$ modulo $p^h\mathbf{Z}_p$.

Théorème V.3.3. *Les $[\frac{n}{p^h}]!\binom{x}{n}$ pour $n \in \mathbf{N}$ forment une base de Banach de LA_h.*

Démonstration. Posons $g_n(x) = [\frac{n}{p^h}]!\binom{x}{n}$. Si $j \in \{1, \ldots, p^h\}$, soit $g_{n,j}(x) = g_n(-j + p^h x)$. Par définition, on a

$$g_{n,j}(x) = [\frac{n}{p^h}]!\frac{1}{n!}\prod_{k=0}^{n-1}(-j-k+p^h x).$$

Pour simplifier les formules (transformer des produits de p^x en sommes), on définit $v_p(g_{n,j})$ comme étant l'inf. des valuations p-adiques des coefficients de $g_{n,j}$, ce qui fait que l'on a

$$\|g_n\|_{h,-j} = p^{-v_p(g_{n,j})},$$

comme on le voit en revenant à la définition. On a aussi $v_p(g_{n,j}) = \inf_{x\in\mathscr{O}_{\mathbf{C}_p}} v_p(g_{i,j}(x))$. Si $v_p(j+k) < h$, alors $v_p(-j-k+p^h x) = v_p(j+k)$ quel que soit $x \in \mathscr{O}_{\mathbf{C}_p}$ et si $v_p(j+k) \geqslant h$, alors $v_p(-j-k+p^h x) \geqslant h$ quel que soit $x \in \mathscr{O}_{\mathbf{C}_p}$ avec égalité si l'image de x dans $\overline{\mathbf{F}}_p$ n'appartient pas à $\mathbf{F}_p$. On en déduit la formule

$$v_p(g_{n,j}) = v_p([\frac{n}{p^h}]!) - v_p(n!) + \sum_{k=0}^{n-1} \inf(v_p(j+k), h)$$
$$= \sum_{\ell=1}^{h} \Big([\frac{n+j-1}{p^\ell}] - [\frac{j-1}{p^\ell}] - [\frac{n}{p^\ell}] \Big).$$

Comme $[x+y] \geqslant [x]+[y]$, chacun des termes de la somme est positif ou nul et donc $\|g_n\|_{\mathrm{LA}_h} \leqslant 1$ et d'autre part, le cas $j = 1$ implique $\|g_n\|_{\mathrm{LA}_h} = 1$. Pour continuer la démonstration, nous aurons besoin du lemme suivant.

Lemme V.3.4. *Soit $n = mp^h - i$ avec $1 \leqslant i \leqslant p^h$ et, si $j \in \{1, \dots, p^h\}$, soit $\overline{g}_{n,j}$ la réduction de $g_{n,j}$ modulo p. Alors*
(i) $\overline{g}_{n,j} = 0$ *si* $j > i$
(ii) $\deg(\overline{g}_{n,i}) = m-1$
(iii) $\deg(\overline{g}_{n,j}) \leqslant m-1$ *si* $j < i$.

Démonstration. Développons $g_{n,j}$ sous la forme $\sum_{k=0}^{n} a_k x^k$. D'après la discussion précédente, on a $a_k \in \mathbf{Z}_p$. D'autre part, les zéros de $g_{n,j}$ sont les $\frac{j+k}{p^h}$ pour $k \in \{0,1,\dots,n-1\}$. Le nombre de zéros de $g_{n,j}$ dans $\mathscr{O}_{\mathbf{C}_p}$ est donc égal au nombre d'éléments de $\{0,1,\dots,n-1\}$ tels que $v_p(j+k) \geqslant h$; on en déduit le fait que $g_{n,j}$ a $[\frac{n+j-1}{p^h}] - [\frac{j-1}{p^h}] = [\frac{n+j-1}{p^h}]$ zéros dans $\mathscr{O}_{\mathbf{C}_p}$. Si on utilise alors la relation entre les coefficients d'un polynôme et les fonctions symétriques de ses racines, on montre que $v_p(a_k)$ atteint son minimum pour $k = k_0 = [\frac{n+j-1}{p^h}]$ et que l'on a $v_p(a_k) > v_p(a_{k_0})$ si $k > k_0$. Le degré de $\overline{g}_{n,j}$ est donc inférieur ou égal à $[\frac{n+j-1}{p^h}]$, l'égalité se produisant si et seulement si $v_p(g_{n,j}) = 0$ (autrement, $\overline{g}_{n,j} = 0$). On en tire le (iii) et le (i) car si $j > i$, alors $[\frac{n+j-1}{p^h}] - [\frac{j-1}{p^h}] - [\frac{n}{p^h}] = 1$ et $v_p(g_{n,j}) \geqslant 1$.

Si $x \in \mathbf{R}$, soit $\{x\} = x - [x]$. On a $\{\frac{n+i-1}{p^k}\} = 1 - \frac{1}{p^k}$ si $k \leqslant h$ et, comme de plus $\{\frac{i-1}{p^k}\} \leqslant 1 - \frac{1}{p^k}$ et $\{\frac{n}{p^k}\} \leqslant 1 - \frac{1}{p^k}$, on en déduit

l'inégalité

$$0 \leqslant \Big[\frac{n+i-1}{p^k}\Big] - \Big[\frac{i-1}{p^k}\Big] - \Big[\frac{n}{p^k}\Big] = \Big\{\frac{i-1}{p^k}\Big\} + \Big\{\frac{n}{p^k}\Big\} - \Big\{\frac{n+i-1}{p^k}\Big\} \leqslant 1 - \frac{1}{p^k} < 1.$$

On en déduit $v_p(g_{n,i}) = 0$ et $\deg(\overline{g}_{n,i}) = [\frac{n+i-1}{p^h}] = [\frac{mp^h-1}{p^h}] = m-1$, ce qui permet de conclure.

Corollaire V.3.5. *La matrice exprimant les $\overline{g}_n$ dans la base des $\overline{e}_n$ est triangulaire supérieure et ses coefficients diagonaux sont inversibles.*

Ce corollaire permet de terminer la démonstration du théorème V.3.3 en vertu de la proposition V.1.7.

Corollaire V.3.6. *Soit* LA *l'espace des fonctions localement analytiques sur $\mathbf{Z}_p$. Alors $f \in$ LA si et seulement si* $\liminf \frac{1}{n} v_p(a_n(f)) > 0$.

Démonstration. $\mathbf{Z}_p$ étant compact, si f est localement analytique sur $\mathbf{Z}_p$, alors elle appartient à LA_h pour un certain h et alors $\liminf \frac{1}{n} v_p(a_n(f)) \geqslant \frac{1}{(p-1)p^h}$ d'après le théorème V.3.3. Réciproquement, si $\liminf \frac{1}{n} v_p(a_n(f)) > \frac{1}{(p-1)p^h}$, alors $([\frac{n}{p^h}]!)^{-1} a_n(f)$ tend vers 0 et $f \in \mathrm{LA}_h$.

2. Fonctions k-fois uniformément dérivables

On dit qu'une fonction f sur $\mathbf{Z}_p$ est de classe $\mathscr{C}_u^k$ si les fonctions $f^{[i]}(x, h_1, \ldots, h_i)$ définies pour $0 \leqslant i \leqslant k$ sur $\mathbf{Z}_p \times (\mathbf{Z}_p - \{0\})^i$ par récurrence grâce à la formule $f^{[0]}(x) = f(x)$ et

$$f^{[i]}(x, h_1, \ldots, h_i) = \frac{1}{h_i}(f^{[i-1]}(x + h_i, h_1, \ldots, h_{i-1}) - f^{[i-1]}(x, h_1, \ldots, h_{i-1}))$$

se prolongent en des fonctions continues sur $\mathbf{Z}_p^{i+1}$.

Sur $\mathbf{R}$, les notions de classe $\mathscr{C}_u^k$ et de classe $\mathscr{C}^k$ coïncident car on a

$$f^{[i]}(x, h_1, \ldots, h_i) = \int_{[0,1]^i} f^{(i)}(x + t_1 h_1 + \cdots + t_i h_i)\, dt_1 \ldots dt_i.$$

Sur $\mathbf{Z}_p$, l'exemple suivant montre qu'il n'en est rien. Comme on l'a vu, tout élément x de $\mathbf{Z}_p$ peut s'écrire de manière unique sous la forme $\sum_{n=0}^{+\infty} p^n a_n(x)$, avec $a_n(x) \in \{0, \dots, p-1\}$, ce qui permet de définir une fonction $f : \mathbf{Z}_p \to \mathbf{Z}_p$ grâce à la formule $f(x) = \sum_{n=0}^{+\infty} p^{2n} a_n(x)$. On a alors $|f(x) - f(y)|_p \leqslant |x-y|_p^2$ quels que soient $x, y \in \mathbf{Z}_p$, ce qui montre que f est dérivable, de dérivée nulle, en tout point (elle est donc aussi de classe $\mathscr{C}^\infty$) bien que f ne soit localement constante au voisinage d'aucun point. D'autre part, si $(x, h_1, h_2) = (0, p^n, p^n)$ (resp. $(x, h_1, h_2) = ((p-1)p^n, p^n, p^n)$), on a $f^{[2]}(x, h_1, h_2) = 0$ (resp. $f^{[2]}(x, h_1, h_2) = p - p^2$), ce qui montre que $f^{[2]}$ ne peut se prolonger par continuité en $(0, 0, 0)$.

La fonction f est aussi un contre-exemple à un autre énoncé naturel puisqu'elle est de classe $\mathscr{C}^\infty$ mais n'a de développement limité d'ordre 2 en aucun point.

On munit $\mathscr{C}^k_u(\mathbf{Z}_p, \mathbf{Q}_p)$ de la norme naturelle $\| \ \|$ définie par

$$\|f\| = \sup_{0 \leqslant i \leqslant k} \ \sup_{(x, h_1, \dots, h_i) \in \mathbf{Z}_p^{i+1}} |f^{[i]}(x, h_1, \dots, h_i)|_p$$

qui en fait un espace de Banach p-adique. Soit

$$\mathrm{L}(n, k) = \max\{v_p(n_1) + \cdots + v_p(n_i) \mid i \leqslant k, \ n_1 + \cdots + n_i = n, \ n_j \geqslant 1\}.$$

Théorème V.3.7. *Les $p^{\mathrm{L}(n,k)}\binom{x}{n}$ forment une base de Banach de $\mathscr{C}^k_u(\mathbf{Z}_p, \mathbf{Q}_p)$.*

Démonstration. Soit $g_{\mathrm{T}}(x) = (1+\mathrm{T})^x$. On a

$$g_{\mathrm{T}}^{[i]}(x, h_1, \dots, h_k) = (1+\mathrm{T})^x \prod_{j=1}^{i} \frac{(1+\mathrm{T})^{h_j} - 1}{h_j},$$

ce qui nous donne, notant P_n le polynôme $\binom{x}{n}$, identifiant les termes de degré n de chaque coté, et utilisant l'identité $\frac{1}{x}\binom{x}{n} = \frac{1}{n}\binom{x-1}{n-1}$, la formule

$$\mathrm{P}_n^{[i]}(x, h_1, \dots, h_i) = \sum_{\substack{n_0 + n_1 + \cdots + n_i = n \\ n_1, \dots, n_i \geqslant 1}} \frac{1}{n_1 \cdots n_i} \binom{x}{n_0} \binom{h_1 - 1}{n_1 - 1} \cdots \binom{h_i - 1}{n_i - 1}.$$

Si y est une variable, soit $\partial_y^{[1]}$ l'opérateur qui à $f(z,y)$ associe $\partial_y^{[1]}(z,y) = f(z,y+1) - f(z,y)$ et, si $k \in \mathbf{N}$, soit $\partial_y^{[k]}$ l'itéré k-ième de $\partial_y^{[1]}$. Soit $g_i(x,h_1,\dots,h_i) = f^{[i]}(x,h_1+1,\dots,h_i+1)$. On déduit de la formule précédente, l'identité

$$\partial_x^{[n_0]}\partial_{h_1}^{[n_1-1]}\dots\partial_{h_i}^{[n_i-1]}g_i(0,0,\dots,0) = \frac{a_{n_0+\dots+n_i}(f)}{n_1\cdots n_i}$$

et donc, utilisant le théorème de Mahler à $i+1$ variables, que si g_i (et donc $f^{[i]}$) est continue, alors la suite de terme général $\frac{a_{n_0+\dots+n_i}(f)}{n_1\cdots n_i}$ tend vers 0 quand $n_0+\dots+n_i$ tend vers $+\infty$ et que, réciproquement, si cette suite tend vers 0, alors la série

$$\sum_{n_0=0}^{+\infty}\sum_{n_1=1}^{+\infty}\dots\sum_{n_i=1}^{+\infty}\frac{a_{n_0+\dots+n_i}(f)}{n_1\cdots n_i}\binom{x}{n_0}\binom{h_1-1}{n_1-1}\dots\binom{h_i-1}{n_i-1}$$

définit une fonction continue sur $\mathbf{Z}_p^{i+1}$ qui coïncide avec $f^{[i]}$ sur $\mathbf{N}\times(\mathbf{N}-\{0\})^i$ et donc sur $\mathbf{Z}_p\times(\mathbf{Z}_p-\{0\})^i$, et que l'on a de plus

$$\|f^{[i]}\|_\infty = \|g_i\|_\infty = \sup_{n_0,\dots,n_i}\frac{1}{|n_1\cdots n_i|_p}|a_{n_0+\dots+n_i}(f)|_p\,;$$

d'où le résultat.

Lemme V.3.8. *Il existe* $\mathrm{C}_k > 0$ *tel que l'on ait* $k\frac{\log n}{\log p} - \mathrm{C}_k \leqslant \mathrm{L}(n,k) \leqslant k\frac{\log n}{\log p}$.

Démonstration. Si $n_i \leqslant n$, alors $v_p(n_i) \leqslant \frac{\log n}{\log p}$. D'autre part, si $k \leqslant p^r$ et $u = [\frac{\log n}{\log p}]$, on peut prendre $(n_1,\dots,n_k) = (p^{u-r},\dots,p^{u-r})$ ce qui implique $\mathrm{L}(n,k) \geqslant k(\frac{\log n}{\log p} - 1 - r)$ et permet de conclure.

Corollaire V.3.9. *$f \in \mathscr{C}_u^k$ si et seulement si la suite de terme général $n^k|a_n(f)|_p$ tend vers 0 et la norme $\|f\|_{\mathscr{C}_u^k} = \sup_n(n+1)^k|a_n(f)|_p$ est équivalente à la norme naturelle sur $\mathscr{C}_u^k$.*

3. Distributions continues

On appelle distribution continue sur $\mathbf{Z}_p$ une forme linéaire continue sur LA, c'est-à-dire une forme linéaire sur LA dont la restriction à chaque LA_h est continue. On note $\mathscr{D}_{\mathrm{cont}}$ l'ensemble de ces distributions.

Si $h \in \mathbf{N}$, on pose $\rho_h = p^{-1/(p-1)p^h}$. On a aussi $\rho_h = |\varepsilon - 1|_p$ si ε est une racine primitive p^{h+1}-ième de l'unité (*cf.* exemple IV.2.8).

Transformées d'Amice et Laplace :

$$\mathscr{A}_\mu(\mathrm{T}) = \int_{\mathbf{Z}_p} (1+\mathrm{T})^x \mu(x) = \sum_{n=0}^{+\infty} \mathrm{T}^n \int_{\mathbf{Z}_p} \binom{x}{n} \mu(x)$$

$$\mathscr{L}_\mu(t) = \int_{\mathbf{Z}_p} e^{tx} \mu(x) = \sum_{n=0}^{+\infty} t^n \int_{\mathbf{Z}_p} \frac{x}{n!} \mu(x).$$

Lemme V.3.10. *Si $z \in \mathrm{B}(0,1^-)$, alors $\int_{\mathbf{Z}_p} (1+z)^x \mu(x) = \mathscr{A}_\mu(z)$.*

Démonstration. Si $|z|_p < \rho_h$, alors la série $\sum_{n=0}^{+\infty} \binom{x}{n} z^n$ converge vers $(1+z)^x$ dans $\mathrm{LA}_h(\mathbf{Z}_p, \mathbf{C}_p)$ car $z^n/[n/p^h]!$ tend vers 0.

Si $\mathrm{F}(\mathrm{T}) = \sum_{n=0}^{+\infty} a_n \mathrm{T}^n$ est de rayon de convergence $\geqslant 1$ et si $\rho < 1$, on note $\|\mathrm{F}\|_\rho$ le maximum de $|a_n|_p \rho^n$ pour $n \in \mathbf{N}$. C'est aussi le maximum de $|\mathrm{F}(x)|_p$ pour $x \in \mathbf{C}_p$ vérifiant $|x|_p \leqslant \rho$. On a $\|\mathrm{F}\|_{\rho_1} \leqslant \|\mathrm{F}\|_{\rho_2}$ si $\rho_1 \leqslant \rho_2$ et une adaptation de la démonstration du lemme de Gauss montre que l'on a $\|\mathrm{FG}\|_\rho \leqslant \|\mathrm{F}\|_\rho \|\mathrm{G}\|_\rho$ si F et G sont de rayon de convergence $\geqslant 1$ et $\rho < 1$.

Théorème V.3.11. *L'application qui à une distribution associe sa transformée d'Amice est un isomorphisme d'espaces vectoriels topologiques de $\mathscr{D}_{\mathrm{cont}}$ sur l'espace des séries formelles convergeant sur $\mathrm{B}(0,1^-)$. De plus, on a*

$$\|\mathscr{A}_\mu\|_{\rho_h} \leqslant \|\mu\|_{\mathrm{LA}_h} \leqslant p \|\mathscr{A}_\mu\|_{\rho_{h+1}}.$$

Démonstration. Soit μ une distribution et

$$\mathscr{A}_\mu(\mathrm{T}) = \sum_{n=0}^{+\infty} b_n \mathrm{T}^n$$

sa transformée d'Amice. D'après le théorème V.3.3, quel que soit $h \in \mathbf{N}$, la suite de terme général $[n/p^h]! b_n$ est bornée et

$$\|\mu\|_{\mathrm{LA}_h} = \sup_n |[n/p^h]! b_n|_p.$$

On en déduit le fait que $\mathscr{A}_\mu$ converge sur $\mathrm{B}(0,\rho_h^-)$ quel que soit h (et donc converge sur $\mathrm{B}(0,1^-)$) et que de plus,

$$\|\mathscr{A}_\mu(\mathrm{T})\|_{\rho_h} = \sup_n |b_n|_p \rho_h^n \leqslant \|\mu\|_{\mathrm{LA}_h}$$

car $\rho_h^n \leqslant |[n/p^h]!|_p$.

Réciproquement, si F converge sur $\mathrm{B}(0,1^-)$, alors la suite de terme général $[n/p^h]!b_n$ tend vers 0 et $\sup_n |[n/p^h]!b_n|_p \leqslant p\|\mathrm{F}(\mathrm{T})\|_{p^{-1/p^{h+1}}}$ (car $v_p([n/p^h]!) \geqslant n/p^{h+1} - 1$) et comme $\rho_{h+1} \geqslant p^{-1/p^{h+1}}$, on en déduit la majoration du théorème, ce qui termine la démonstration.

Remarque V.3.12. On montre plus généralement que quel que soit $\varepsilon > 0$, il existe $\mathrm{C}_{h,\varepsilon} > 0$ tel que l'on ait $\|\mu\|_{\mathrm{LA}_h} \leqslant \mathrm{C}_{h,\varepsilon}\|\mathscr{A}_\mu\|_{\rho_h+\varepsilon}$.

4. Opérations sur les distributions

Les opérations sur les distributions sont en gros les mêmes que celles que nous avons définies sur les mesures et les démonstrations sont très semblables.

i) Dérivée d'une distribution : Si μ est une distribution, on définit sa dérivée $\frac{d}{dx}\mu$ par la formule

$$\int_{\mathbf{Z}_p} f(x)\frac{d}{dx}\mu(x) = \int_{\mathbf{Z}_p} f'(x)\mu(x).$$

La transformée d'Amice de $\frac{d}{dx}\mu$ s'obtient à partir de celle de μ par multiplication par $\log(1+\mathrm{T})$; sa transformée de Laplace par multiplication par t. Contrairement au cas réel, on ne peut en général pas définir la primitive d'une distribution puisque cela revient au niveau des transformées d'Amice à diviser par la série entière $\log(1+\mathrm{T})$ qui a beaucoup de zéros dans $\mathrm{B}(0,1^-)$.

ii) Multiplication par une fonction localement analytique :

$$\int_{\mathbf{Z}_p} f(x)(g(x)\mu(x)) = \int_{\mathbf{Z}_p} (f(x)g(x))\mu(x).$$

– Multiplication par x : $\mathscr{A}_{x\mu}(\mathrm{T}) = (1+\mathrm{T})\frac{d}{d\mathrm{T}}\mathscr{A}_\mu(\mathrm{T})$ et $\mathscr{L}_{x\mu}(t) = \frac{d}{dt}\mathscr{L}_\mu(t)$.

– Multiplication par z^x avec $|z-1|_p<1$. Le lemme V.3.10 montre que l'on a

$$\mathscr{A}_{z^x\mu}(\mathrm{T})=\mathscr{A}_\mu(z(1+\mathrm{T})-1).$$

iii) Division par x. Si μ est une distribution, l'équation $x\lambda=\mu$ a une infinité de solutions différant 2 à 2 par un multiple de la masse de Dirac en 0. En effet, on peut écrire n'importe quel élément de LA sous la forme $f=f(0)\mathbf{1}_{\mathbf{Z}_p}+xg$, où $g\in\mathrm{LA}$ et les solutions de l'équation $x\lambda=\mu$ sont donc de la forme

$$\int_{\mathbf{Z}_p}f(x)\lambda(x)=af(0)+\int_{\mathbf{Z}_p}g(x)\mu(x).$$

La transformée d'Amice de $x^{-1}\mu$ est une primitive (déterminée à constante près, ce qui correspond à l'indétermination de $x^{-1}\mu$ à un multiple près de la masse de Dirac) de $(1+\mathrm{T})^{-1}\mathscr{A}_\mu(\mathrm{T})$.

iv) Restriction à $a+p^n\mathbf{Z}_p$:

$$\mathscr{A}_{\mathrm{Res}_{a+p^n\mathbf{Z}_p}(\mu)}(\mathrm{T})=\frac{1}{p^n}\sum_{\varepsilon^{p^n}=1}\varepsilon^{-a}\mathscr{A}_\mu((1+\mathrm{T})\varepsilon-1).$$

v) Convolution : Si $f\in\mathrm{LA}_h$ et $\mu\in\mathscr{D}_{\mathrm{cont}}$, alors

$$\mu*f\in\mathrm{LA}_h\quad\text{et}\quad\|\mu*f\|_{\mathrm{LA}_h}\leqslant\|\mu\|_{\mathrm{LA}_h}\|f\|_{\mathrm{LA}_h}.$$

En effet, si $|y-y_0|_p\leqslant p^{-h}$, alors

$$f(x+y)=\sum_{n=0}^{+\infty}\frac{p^{nh}f^{(n)}(x+y_0)}{n!}\Big(\frac{y-y_0}{p^h}\Big)^n$$

et le lemme V.3.1 montre que

$$\Big\|\frac{p^{nh}f^{(n)}(x+y_0)}{n!}\Big\|_{\mathrm{LA}_h}\leqslant\|f\|_{\mathrm{LA}_h}$$

quel que soit $n\in\mathbf{N}$, et la suite de terme général $p^{nh}f^{(n)}(x+y_0)/n!$ tend vers 0 dans LA_h.

Ceci permet de définir la convolution de deux distributions continues et on a

$$\mathscr{A}_{\lambda*\mu}(\mathrm{T})=\mathscr{A}_\lambda(\mathrm{T})\mathscr{A}_\mu(\mathrm{T}).$$

5. Distributions tempérées

Définition V.3.13. Soit $r \in \mathbf{R}$. Une distribution continue μ sur $\mathbf{Z}_p$ est dite d'ordre r si la suite de terme général $p^{-nr}\|\mu\|_{\mathrm{LA}_n}$ est bornée. On note $\mathscr{D}_r$ l'ensemble des distributions d'ordre r que l'on munit de la norme $\| \ \|_r$ définie par $\|\mu\|_r = \sup_{n\in\mathbf{N}} p^{-nr}\|\mu\|_{\mathrm{LA}_n}$. Une distribution est dite tempérée s'il existe $r \in \mathbf{R}_+$ telle qu'elle soit d'ordre r. On note $\mathscr{D}_{\mathrm{temp}}$ l'espace des distributions tempérées.

Remarque V.3.14

(i) Comme on a $\|f\|_{\mathrm{LA}_{n+1}} \leqslant \|f\|_{\mathrm{LA}_n}$ si $f \in \mathrm{LA}_n$, la suite $\|\mu\|_{\mathrm{LA}_n}$ est une suite croissante de n ; une distribution d'ordre < 0 est donc nulle.

(ii) Si f est constante modulo $p^n\mathbf{Z}_p$, alors $\|f\|_{\mathrm{LA}_n} = \|f\|_\infty$. On en déduit le fait qu'une distribution d'ordre 0 est continue sur l'espace des fonctions localement constantes muni de la norme du sup et donc est une mesure.

(iii) Si r et r' sont deux éléments de $\mathbf{R}$ vérifiant $r \leqslant r'$, toute distribution d'ordre r est aussi d'ordre r'.

(iv) si μ est d'ordre r, alors la suite de terme général

$$\sup_{a\in\mathrm{X}}\sup_{j\in\mathbf{N}} p^{-nr}\Big\| \int_{a+p^n\mathbf{Z}_p} (\frac{x-a}{p^n})^j \mu\Big\| = p^{-nr}\|\mu\|_{\mathrm{LA}_n} \leqslant p^{-rn}\|\mathscr{A}_\mu\|_{\rho_n}$$

est bornée.

Nous allons maintenant caractériser les distributions d'ordre r en termes de leur transformée d'Amice.

Définition V.3.15. Une série entière F de rayon de convergence 1 sera dite d'ordre r si la suite de terme général $p^{-nr}\|\mathrm{F}\|_{\rho_n}$ est bornée. L'espace des séries entières d'ordre r muni de la norme $\sup_n p^{-nr}\|\mathrm{F}\|_{\rho_n}$ est un espace de Banach.

Remarque V.3.16

(i) Comme la suite $\|\mathrm{F}\|_{\rho_n}$ est croissante, une série entière d'ordre r pour $r < 0$, est identiquement nulle.

(ii) Une série entière d'ordre 0 est à coefficients bornés.

(iii) Comme $\|\mathrm{FG}\|_\rho = \|\mathrm{F}\|_\rho\|\mathrm{G}\|_\rho$, on en déduit le fait que si F est d'ordre r et G est d'ordre s, alors FG est d'ordre $r + s$.

Lemme V.3.17. *Si* $\mathrm{F}\log(1+\mathrm{T})$ *est d'ordre* r*, alors* F *est d'ordre* $r-1$*.*

Démonstration. $\|\log(1+\mathrm{T})\|_{\rho_n} = p^n p^{-1/(p-1)}$. (Le maximum de $|1/k|_p \rho_n^k$ est atteint pour $k=p^n$ et $k=p^{n+1}$.)

Proposition V.3.18. *L'application qui à une distribution associe sa transformée d'Amice induit un isomorphisme d'espaces de Banach de l'espace des distributions d'ordre* r *sur celui des séries entières d'ordre* r*.*

Démonstration. On a $\|\mathscr{A}_\mu\|_{\rho_n} \leqslant \|\mu\|_{\mathrm{LA}_n} \leqslant p\|\mathscr{A}_\mu\|_{\rho_{n+1}}$ d'après le théorème V.3.11. On en déduit le résultat.

Le lemme suivant nous fournira une caractérisation plus commode des séries d'ordre r et donc des distributions d'ordre r.

Lemme V.3.19. *Soit* $\mathrm{F}(\mathrm{T}) = \sum_{n=0}^{+\infty} a_n \mathrm{T}^n$ *une série convergeant sur* $\mathrm{B}(0,1^-)$ *et* $r \in \mathbf{R}_+$*. Les deux conditions suivantes sont équivalentes :*

(i) *la suite de terme général* $p^{-nr}\|\mathrm{F}\|_{\rho_n}$ *est bornée,*

(ii) *la suite de terme général* $n^{-r}|a_n|_p$ *est bornée.*

Démonstration

(i)⇒(ii) Soit $u_m = \sup_{n\geqslant m} p^{-nr}\|\mathrm{F}\|_{\rho_n}$. Quel que soit $k \in \mathbf{N}$, on a $|a_k|_p \rho_n^k \leqslant u_m p^{nr}$ si $n \geqslant m$. En particulier, si $k \geqslant p^{m+1}$, on peut appliquer cette inégalité à $n = [\frac{\log k}{\log p}]$ de telle sorte que l'on a $\frac{\log k}{\log p} - 1 \leqslant n \leqslant \frac{\log k}{\log p}$, ce qui nous donne

$$|a_k|_p \leqslant u_m p^{nr} \rho_n^{-k} \leqslant u_m k^r p^{-k\frac{\log \rho_n}{\log p}}$$

et comme

$$-\log \rho_n = \frac{\log p}{(p-1)p^n} \leqslant \frac{\log p}{(p-1)p^{\frac{\log k}{\log p}-1}},$$

on obtient finalement $|a_k|_p \leqslant p^{p/(p-1)} u_m k^r$ si $k \geqslant p^{m+1}$, ce qui permet de montrer que si la suite de terme général u_m est bornée, alors celle de terme général $n^{-r}|a_n|_p$ est bornée, et permet de conclure.

(ii)⇒(i) Soit $v_m = \sup_{k\geqslant m} k^{-r}|a_k|_p$. Un petit calcul montre que si $r>0$ et $a<1$, alors la fonction $x^r a^x$ atteint son maximum sur $\mathbf{R}_+$

en $-r/\log a$ et que ce maximum vaut $e^{-r}(-r/\log a)^r$. On en déduit la majoration

$$\|\mathrm{F}\|_{\rho_n} = \sup_k |a_k|_p \rho_n^k \leqslant \max\Big(\sup_{k\leqslant m} |a_k|_p \rho_n^k, v_m e^{-r}\Big(\frac{-r}{\log \rho_n}\Big)^r\Big)$$

et comme ρ_n tend vers 1, le terme $\sup_{k\leqslant m} |a_k|_p \rho_n^k$ est majoré par une constante ne dépendant pas de n et comme

$$\frac{-1}{\log \rho_n} = \frac{(p-1)p^n}{\log p},$$

on obtient la majoration

$$p^{-nr}\|\mathrm{F}\|_{\rho_n} \leqslant e^{-r} r^r \Big(\frac{p-1}{\log p}\Big)^r v_m$$

si n est assez grand. On en déduit le fait que si la suite de terme général v_m est bornée, alors celle de terme général $p^{-nr}\|\mathrm{F}\|_{\rho_n}$ est bornée, ce qui permet de conclure.

Corollaire V.3.20. *Si $r \in \mathbf{R}_+$, on peut munir $\mathscr{D}_r$ des trois normes suivantes qui sont équivalentes.*

(i) $\|\mu\|_r = \sup_{n\in\mathbf{N}} p^{-nr}\|\mu\|_{\mathrm{LA}_n}$.
(ii) $\|\mu\|_{r,1} = \sup_{n\in\mathbf{N}} p^{-nr}\|\mathscr{A}_\mu\|_{\rho_n}$
(iii) $\|\mu\|_{r,2} = \sup_{n\in\mathbf{N}} (1+n)^{-r} |\int_{\mathbf{Z}_p} \binom{x}{n} \mu(x)|_p$.

Corollaire V.3.21. *Si k est un entier, $\mathscr{D}_k$ est le dual de $\mathscr{C}_u^k(\mathbf{Z}_p, \mathbf{Q}_p)$.*

Démonstration. Cela suit immédiatement de l'équivalence entre les normes $\|\ \|$ et $\|\ \|_{r,2}$ et de la caractérisation de $\mathscr{C}_u^k(\mathbf{Z}_p, \mathbf{Q}_p)$ donnée au corollaire V.3.9.

Remarque V.3.22. C'est ce corollaire qui justifie le nom de distribution tempérée pour les distributions d'ordre fini.

6. Une autre caractérisation des distributions tempérées

Dans le paragraphe précédent, on a caractérisé les distributions tempérées en termes de leurs transformées d'Amice ce qui permet de construire une distribution tempérée à partir d'une série entière de rayon de convergence 1 vérifiant des conditions de croissance. La connaissance de la transformée d'Amice d'une distribution est

équivalente à la connaissance des $\int_{\mathbf{Z}_p} x^i\mu(x)$ pour $i \in \mathbf{N}$. Dans ce paragraphe, nous montrons comment construire (théorème V.3.23 et corollaire V.3.24) une distribution d'ordre fini en ne connaissant que les intégrales du type $\int_{a+p^n\mathbf{Z}_p} x^i\mu(x)$ pour $a \in \mathbf{Z}_p$, $n \in \mathbf{N}$ et $0 \leqslant i \leqslant \mathrm{N}$. Cette construction généralise celle des mesures donnée dans la remarque V.2.12 et est très importante pour les applications arithmétiques.

Si I est une partie de $\mathbf{N}$, notons LP^{I} le $\mathbf{Q}_p$-espace vectoriel des fonctions localement polynomiales sur $\mathbf{Z}_p$ dont les degrés appartiennent à I, c'est-à-dire les fonctions qui s'écrivent localement sous la forme $\sum_{i\in\mathrm{I}} a_i x^i$, où les a_i sont des éléments de $\mathbf{Q}_p$ presque tous nuls.

On note $\mathscr{D}^{\mathrm{I}}_{\mathrm{alg}}$ l'ensemble des distributions algébriques sur $\mathbf{Z}_p$ (de degré $\in$ I), c'est-à-dire l'ensemble des formes linéaires sur LP^{I}. Un élément μ de $\mathscr{D}^{\mathrm{I}}_{\mathrm{alg}}$ est donc équivalent à la donnée des valeurs $\int_{a+p^n\mathbf{Z}_p} x^i\mu(x)$ pour $i \in \mathrm{I}$, $a \in \mathbf{Z}_p$ et $n \in \mathbf{N}$ avec les relations de compatibilité évidentes

$$\int_{a+p^n\mathbf{Z}_p} x^i\mu(x) = \sum_{k=0}^{p-1} \int_{a+kp^n+p^{n+1}\mathbf{Z}_p} x^i\mu(x).$$

On a une application naturelle de $\mathscr{D}_{\mathrm{cont}}$ dans $\mathscr{D}^{\mathrm{I}}_{\mathrm{alg}}$ quel que soit $\mathrm{I} \subset \mathbf{N}$.

Si $r \in \mathbf{R}$,on note $\mathscr{D}_r^{[0,\mathrm{N}]}$ le sous-espace des $\mu \in \mathscr{D}^{[0,\mathrm{N}]}_{\mathrm{alg}}$ tels que la suite de terme général $\sup\limits_{a\in\mathbf{Z}_p} \sup\limits_{0\leqslant i\leqslant \mathrm{N}} p^{-nr} \| \int_{a+p^n\mathbf{Z}_p} (\frac{x-a}{p^n})^i\mu \|$ est bornée et on munit $\mathscr{D}_r^{[0,\mathrm{N}]}$ de la norme $\| \ \|_{r,[0,\mathrm{N}]}$ définie par

$$\|\mu\|_{r,[0,\mathrm{N}]} = \sup_{n\in\mathbf{N}} \Big(\sup_{a\in\mathbf{Z}_p} \sup_{0\leqslant i\leqslant \mathrm{N}} p^{-nr} \| \int_{a+p^n\mathbf{Z}_p} (\frac{x-a}{p^n})^i\mu \| \Big).$$

On note $\mathrm{E}(r)$ la partie entière de r.

Théorème V.3.23. *Si $r \in \mathbf{R}$ et $\mathrm{N} \geqslant \mathrm{E}(r)$, l'application naturelle de $\mathscr{D}_r$ dans $\mathscr{D}^{[0,\mathrm{N}]}_{\mathrm{alg}}$ induit un isomorphisme d'espaces de Banach p-adiques de $\mathscr{D}_r$ sur $\mathscr{D}_r^{[0,\mathrm{N}]}$.*

Corollaire V.3.24. *Si $\mathrm{N} \geqslant \mathrm{E}(r)$, la norme $\| \ \|_{r,[0,\mathrm{N}]}$ est une norme sur $\mathscr{D}_r$ équivalente à la norme $\| \ \|_r$.*

Démonstration. Le fait que l'image de $\mathscr{D}_r$ dans $\mathscr{D}_{\mathrm{alg}}^{[0,\mathrm{N}]}$ soit incluse dans $\mathscr{D}_r^{[0,\mathrm{N}]}$ est une conséquence de l'inégalité

$$\sup_{a\in\mathbf{Z}_p}\sup_{0\leqslant i\leqslant \mathrm{N}} p^{-nr}\Big\|\int_{a+p^n\mathbf{Z}_p}\Big(\frac{x-a}{p^n}\Big)^i\mu\Big\| \leqslant p^{-nr}\|\mu\|_{\mathrm{LA}_n}$$

qui implique de plus que l'application en question est continue. Il suffit donc, d'après le théorème de l'image ouverte, de prouver que c'est un isomorphisme d'espaces vectoriels.

Commençons par prouver la surjectivité. Nous allons construire une forme linéaire continue sur LA, en approximant $f \in \mathrm{LA}$ par une suite d'éléments de $\mathrm{LP}^{[0,\mathrm{N}]}$. Les suites de fonctions que l'on va considérer vont être obtenues en remplaçant f par sa série de Taylor tronquée à l'ordre N au voisinage d'un système de représentants modulo $p^n\mathbf{Z}_p$ et en faisant tendre n vers $+\infty$. La vérification que le procédé converge repose sur le lemme V.3.26 dont le rôle est de montrer que $\int_{\mathbf{Z}_p} f_n\mu(x)$ ne dépend pas trop du choix du système de représentants modulo $p^n\mathbf{Z}_p$.

Si $h \in \mathbf{N}$ et $\mu \in \mathscr{D}_r^{[0,\mathrm{N}]}$, posons

$$\|\mu\|_{r,h} = \sup_{n\geqslant h}\Big(\sup_{a\in\mathbf{Z}_p}\sup_{0\leqslant j\leqslant \mathrm{N}} p^{-nr}\Big|\int_{a+p^n\mathbf{Z}_p}\Big(\frac{x-a}{p^n}\Big)^j\mu(x)\Big|_p\Big)$$

Remarque V.3.25

(i) La suite $\|\mu\|_{r,h}$ est décroissante.

(ii) Si $g \in \mathrm{LP}^{[0,\mathrm{N}]}\cap\mathrm{LA}_h$, alors $\Big|\int_{\mathbf{Z}_p} g(x)\mu(x)\Big|_p \leqslant p^{rh}\|\mu\|_{r,h}\|g\|_{\mathrm{LA}_h}$.

Soit $f \in \mathrm{LA}_h$. Si $a \in \mathbf{Z}_p$, il existe une suite $c_k(a)$ tendant vers 0 telle que l'on ait $f(x) = \sum_{k=0}^{+\infty} c_k(a)(\frac{x-a}{p^h})^k$ si $x \in a+p^h\mathbf{Z}_p$. Si $a \in \mathbf{Z}_p$ et X est un ouvert compact de $\mathbf{Z}_p$ contenu dans $a+p^h\mathbf{Z}_p$, soit $f_{\mathrm{X},a}$ la fonction définie par la formule $f_{\mathrm{X},a}(x) = \mathbf{1}_{\mathrm{X}}(x)\Big(\sum_{n=0}^{\mathrm{N}} c_k(a)(\frac{x-a}{p^h})^k\Big)$. Nous aurons besoin du lemme suivant.

Lemme V.3.26. *Soient $f \in \mathrm{LA}_h$, $n \geqslant h$, $a \in \mathbf{Z}_p$ et $\mathrm{X} = a + p^{n+1}\mathbf{Z}_p$. Alors, quel que soit $b \in a + p^n\mathbf{Z}_p$, on a la majoration*

$$\|f_{\mathrm{X},a} - f_{\mathrm{X},b}\|_{\mathrm{LA}_{n+1}} \leqslant p^{-(n-h)(\mathrm{N}+1)}\|f\|_{\mathrm{LA}_h}.$$

Démonstration. Si on écrit $f_{\mathrm{X},a}(x) - f_{\mathrm{X},b}(x)$ sous la forme

$$\mathbf{1}_{\mathrm{X}}(x)\Big(\sum_{i=0}^{\mathrm{N}} b_i\Big(\frac{x-a}{p^{n+1}}\Big)^i\Big),$$

on a $\|f_{\mathrm{X},a} - f_{\mathrm{X},b}\|_{\mathrm{LA}_{n+1}} = \sup_{0\leqslant i\leqslant \mathrm{N}} |b_i|_p$. D'autre part,

$$f_{\mathrm{X},a}(x) - f_{\mathrm{X},b}(x) = \mathbf{1}_{\mathrm{X}}(x)\Big(\sum_{n=\mathrm{N}+1}^{+\infty}\Big(c_k(b)\big(\frac{x-b}{p^h}\big)^k - c_k(a)\big(\frac{x-a}{p^h}\big)^k\Big)\Big),$$

ce qui, remplaçant $x-b$ par $(x-a)+(a-b)$ et développant, nous donne, pour $i \leqslant \mathrm{N}$, l'identité

$$p^{-(n+1)i} b_i = p^{-ih} \sum_{k=\mathrm{N}+1}^{+\infty} \binom{k}{i} c_k(b)\big(\frac{a-b}{p^h}\big)^{k-i}.$$

D'où, utilisant la majoration $|c_k(b)|_p \leqslant \|f\|_{\mathrm{LA}_h}$ valable quel que soit $k \in \mathbf{N}$ par définition de $\|f\|_{\mathrm{LA}_h}$, on tire la majoration

$$|b_i|_p \leqslant \|f\|_{\mathrm{LA}_h} p^{-(n+1-h)i} p^{-(n-h)(\mathrm{N}+1-i)} \leqslant \|f\|_{\mathrm{LA}_h} p^{-(n-h)(\mathrm{N}+1)}$$

et le résultat.

Revenons à la démonstration du théorème V.3.23. Fixons pour chaque $n \in \mathbf{N}$ un système R_n de représentants de $\mathbf{Z}_p$ modulo $p^n\mathbf{Z}_p$. Si $f \in \mathrm{LA}_h$ et $n \geqslant h$, soit $f_n \in \mathrm{LP}^{[0,\mathrm{N}]}$ la fonction définie par

$$f_n(x) = \sum_{a\in \mathrm{R}_n} f_{a+p^n\mathbf{Z}_p,a}(x).$$

Si $a \in \mathrm{R}_{n+1}$, soit $\pi(a) \in \mathrm{R}_n$ le représentant de a modulo $p^n\mathbf{Z}_p$. On a

$$f_{n+1} - f_n = \sum_{a\in \mathrm{R}_{n+1}} f_{a+p^{n+1}\mathbf{Z}_p,a} - f_{a+p^{n+1}\mathbf{Z}_p,\pi(a)}$$

et donc

$$\|f_{n+1} - f_n\|_{\mathrm{LA}_{n+1}} \leqslant p^{-(n-h)(\mathrm{N}+1)} \|f\|_{\mathrm{LA}_h}$$

On en déduit la majoration

$$\Big|\int_{\mathbf{Z}_p} (f_{n+1}(x) - f_n(x))\mu(x)\Big|_p \leqslant p^{r(n+1)}\|\mu\|_{r,n+1} p^{-(n-h)(\mathrm{N}+1)}\|f\|_{\mathrm{LA}_h}$$
$$= (p^{-(\mathrm{N}+1-r)(n+1)}\|\mu\|_{r,n+1})(p^{(h+1)(\mathrm{N}+1)}\|f\|_{\mathrm{LA}_h}),$$

et comme la suite de terme général $p^{-(\mathrm{N}+1-r)(n+1)}\|\mu\|_{r,n+1}$ est décroissante et tend vers 0 par hypothèse car $\mathrm{N}+1-r>0$, on en tire la convergence de la suite de terme général $\int_{\mathbf{Z}_p} f_n(x)\mu(x)$. La limite est indépendante du choix des R_n car si on a deux choix $\mathrm{R}_{1,n}$ et $\mathrm{R}_{2,n}$, on en fabrique un troisième en posant $\mathrm{R}_{2n}=\mathrm{R}_{1,2n}$ et $\mathrm{R}_{2n+1}=\mathrm{R}_{2,2n+1}$, ce qui prouve que les trois limites coïncident. On a donc défini de cette manière une forme linéaire sur LA.

Finalement, on a $\|f_h\|_{\mathrm{LA}_h}\leqslant\|f\|_{\mathrm{LA}_h}$ et

$$\begin{aligned}\Big|\int_{\mathbf{Z}_p} f(x)\mu(x)\Big|_p &\leqslant \sup\Big(\Big|\int_{\mathbf{Z}_p} f_h(x)\mu(x)\Big|_p, \sup_{n\geqslant h}\Big|\int_{\mathbf{Z}_p}(f_{n+1}-f_n)\mu\Big|_p\Big)\\ &\leqslant \sup(p^{rh}\|\mu\|_{r,h}\|f\|_{\mathrm{LA}_h}, p^{r(h+1)}\|\mu\|_{r,h}\|f\|_{\mathrm{LA}_h})\\ &= p^{r(h+1)}\|\mu\|_{r,h}\|f\|_{\mathrm{LA}_h},\end{aligned}$$

ce qui permet de montrer que μ est continue sur LA et que si $h\in\mathbf{N}$, alors $\|\mu\|_{\mathrm{LA}_h}\leqslant p^{r(h+1)}\|\mu\|_{r,h}$. On en déduit le fait que la suite de terme général $p^{-nh}\|\mu\|_{\mathrm{LA}_h}$ est bornée et donc que μ est d'ordre r.

On vient donc de montrer que l'application naturelle de $\mathscr{D}_r$ dans $\mathscr{D}_r^{[0,\mathrm{N}]}$ est surjective. Passons à la démonstration de son injectivité. Nous aurons besoin du lemme suivant

Lemme V.3.27. *Soit μ une distribution continue sur $\mathbf{Z}_p$ telle que $\int_{a+p^n\mathbf{Z}_p}\mu(x)=0$ quels que soient $a\in\mathbf{Z}_p$ et $n\in\mathbf{N}$. Il existe alors une unique distribution λ dont μ est la dérivée. De plus, si μ est d'ordre r, alors λ est d'ordre $r-1$.*

Démonstration. Soit $f\in\mathrm{LA}_h\subset\mathrm{LA}_{h+1}$. Soit R_{h+1} un système de représentants de $\mathbf{Z}_p$ modulo $p^{h+1}\mathbf{Z}_p$ et soit $f^{(-1)}$ l'élément de LA_{h+1} dont la dérivée est f et qui s'annule aux points de R_{h+1}. L'hypothèse implique que $\int_{\mathbf{Z}_p} f^{(-1)}(x)\mu(x)$ ne dépend ni du choix de R_{h+1} ni du choix de h tel que $f\in\mathrm{LA}_h$ car deux choix différents aboutissent à deux fonctions $f^{(-1)}$ différant d'une fonction localement constante. Ceci permet de définir une forme linéaire λ sur LA en posant $\int_{\mathbf{Z}_p} f(x)\lambda=\int_{\mathbf{Z}_p} f^{(-1)}(x)\mu(x)$. Un calcul immédiat montre que

l'on a de plus $\|f^{(-1)}\|_{\mathrm{LA}_{h+1}} \leqslant \|f\|_{\mathrm{LA}_h}$ et donc

$$\left|\int_{\mathbf{Z}_p} f(x)\lambda(x)\right|_p \leqslant \|\mu\|_{\mathrm{LA}_{h+1}}\|f\|_{\mathrm{LA}_h},$$

ce qui prouve que λ est continue. D'autre part, on a $\frac{d}{dx}\lambda = \mu$ par construction.

Finalement, on a $\mathscr{A}_\lambda(\mathrm{T})\log(1+\mathrm{T}) = \mathscr{A}_\mu(\mathrm{T})$, ce qui, utilisant le lemme V.3.17, permet de prouver que λ est d'ordre $r-1$ si μ est d'ordre r et termine la démonstration du lemme.

Remarque V.3.28. Traduit en termes de séries formelles, ce lemme dit que si F est une série de rayon de convergence 1 s'annulant en les $\varepsilon-1$, où ε parcourt l'ensemble des racines de l'unité d'ordre une puissance de p, alors F est divisible par $\log(1+\mathrm{T})$.

Revenons à la démonstration de l'injectivité. Soit μ un élément du noyau, c'est-à-dire une distribution d'ordre r vérifiant $\int_{a+p^n\mathbf{Z}_p} x^i\mu(x) = 0$ quels que soient $0 \leqslant i \leqslant \mathrm{N}$, $a \in \mathbf{Z}_p$ et $n \in \mathbf{N}$. D'après le lemme précédent, on peut écrire μ sous la forme $\frac{d}{dx}\mu_1$, où μ_1 est d'ordre $r-1$ et vérifie

$$\int_{a+p^n\mathbf{Z}_p} x^i\mu(x) = i\int_{a+p^n\mathbf{Z}_p} x^{i-1}\mu_1(x) = 0$$

quel que soit $1 \leqslant i \leqslant \mathrm{N}$. On en déduit par récurrence le fait que $\mu = (\frac{d}{dx})^{\mathrm{N}+1}\mu_{\mathrm{N}+1}$, où $\mu_{\mathrm{N}+1}$ est d'ordre $r-\mathrm{N}-1<0$ et donc nulle, ce qui permet de conclure.

Chapitre VI. La fonction zêta p-adique

VI.1. Les congruences de Kummer

Comme nous l'avons mentionné dans l'introduction, Kummer a découvert des congruences modulo p^k entre les valeurs aux entiers négatifs de la fonction zêta. La théorie des mesures sur $\mathbf{Z}_p$ permet de les redémontrer sans douleur et d'aller plus loin en construisant une fonction zêta p-adique.

Si $a \in \mathbf{R}_+^*$, on peut appliquer la proposition I.1.2 à la fonction

$$f_a(t) = \frac{1}{e^t - 1} - \frac{a}{e^{at} - 1}$$

qui est $\mathscr{C}^\infty$ sur $\mathbf{R}_+$ (on s'est débrouillé pour supprimer le pôle en 0) et à décroissance rapide à l'infini.

Corollaire VI.1.1. *Si $a \in \mathbf{R}_+^*$, la fonction $(1 - a^{1-s})\zeta(s) = \mathrm{L}(f_a, s)$ a un prolongement analytique à $\mathbf{C}$ tout entier et, si $n \in \mathbf{N}$, alors $(1 - a^{1+n})\zeta(-n) = (-1)^n f_a^{(n)}(0)$. En particulier, si $a \in \mathbf{Q}$, alors $(1 - a^{1+n})\zeta(-n) \in \mathbf{Q}$.*

Proposition VI.1.2. *Si $a \in \mathbf{Z}_p^*$, il existe une mesure μ_a dont la transformée de Laplace est $f_a(t)$. De plus, $\|\mu_a\|_\infty \leqslant 1$ et, si $n \in \mathbf{N}$, alors $\int_{\mathbf{Z}_p} x^n \mu_a = (-1)^n (1 - a^{1+n})\zeta(-n)$.*

Démonstration. Pour démontrer l'existence de μ_a, il suffit de vérifier que la série obtenue en remplaçant e^t par $1 + \mathrm{T}$ est à coefficients bornés ; ce sera alors la transformée d'Amice de μ_a. Or on peut écrire $(1 + \mathrm{T})^a - 1$ sous la forme $a\mathrm{T}(1 + \mathrm{T}g(\mathrm{T}))$ avec $g(\mathrm{T}) = \sum_{n=2}^{+\infty} \frac{1}{a}\binom{a}{n}\mathrm{T}^{n-2} \in \mathbf{Z}_p[\![\mathrm{T}]\!]$ et donc

$$\frac{1}{\mathrm{T}} - \frac{a}{(1 + \mathrm{T})^a - 1} = \sum_{n=1}^{+\infty} (-\mathrm{T})^{n-1} g^n \in \mathbf{Z}_p[\![\mathrm{T}]\!].$$

Comme on a obtenu une série à coefficients entiers, on obtient en prime la majoration $\|\mu_a\|_\infty \leqslant 1$. Finalement, on a

$$\int_{\mathbf{Z}_p} x^n \mu_a = \mathscr{L}_{\mu_a}^{(n)}(0) = f_a^{(n)}(0).$$

***Proposition VI.1.3* (congruences de Kummer).** *Soit $a \in \mathbf{N} - \{1\}$ premier à p. Soit $k \geqslant 1$. Si n_1 et n_2 sont deux entiers $\geqslant k$ vérifiant $n_1 \equiv n_2$ mod. $(p-1)p^{k-1}$, alors*

$$v_p\big((1 - a^{1+n_1})\zeta(-n_1) - (1 - a^{1+n_2})\zeta(-n_2)\big) \geqslant k.$$

Démonstration. Comme on a supposé $n_1 \geqslant k$ et $n_2 \geqslant k$, on a $|x^{n_1}|_p \leqslant p^{-k}$ et $|x^{n_2}|_p \leqslant p^{-k}$ si $x \in p\mathbf{Z}_p$. D'autre part, comme $(\mathbf{Z}/p^k\mathbf{Z})^*$ est de cardinal $(p-1)p^{k-1}$, et que l'on a supposé $n_1 \equiv n_2 \bmod (p-1)p^{k-1}$, on a $x^{n_1} - x^{n_2} \in p^k\mathbf{Z}_p$ si $x \in \mathbf{Z}_p^*$. En résumé, $|x^{n_1} - x^{n_2}|_p \leqslant p^{-k}$ quel que soit $x \in \mathbf{Z}_p$. Comme $\|\mu_a\|_\infty \leqslant 1$, ceci implique

$$\begin{aligned}\big|(1-a^{1+n_1})\zeta(-n_1) - (1-a^{1+n_2})\zeta(-n_2)\big|_p \\ = \Big|\int_{\mathbf{Z}_p} (x^{n_1} - x^{n_2})\,\mu_a(x)\Big|_p \leqslant p^{-k}\end{aligned}$$

et permet de conclure.

VI.2. Interpolation p-adique

Comme $\mathbf{Q} \subset \mathbf{Q}_p$, on peut voir $n \to \zeta(n)$ comme une application de $-\mathbf{N}$ dans $\mathbf{Q}_p$. On peut se demander s'il est possible de construire une fonction continue sur $\mathbf{Z}_p$ coïncidant avec ζ sur $-\mathbf{N}$. Ce n'est pas possible sous cette forme, mais on a le théorème suivant dont nous donnerons plusieurs démonstrations et quelques généralisations.

Théorème VI.2.1. *Si $i \in \mathbf{Z}/(p-1)\mathbf{Z}$, ($i \in \mathbf{Z}/2\mathbf{Z}$ si $p=2$) il existe une unique fonction $\zeta_{p,i}$ continue sur $\mathbf{Z}_p$ (resp. $\mathbf{Z}_p - \{1\}$) si $i \neq 1$ (resp. si $i=1$) telle que la fonction $(s-1)\zeta_{p,i}(s)$ soit analytique sur $\mathbf{Z}_p$ et que l'on ait $\zeta_{p,i}(-n) = (1-p^n)\zeta(-n)$ si $n \in \mathbf{N}$ vérifie $-n \equiv i$ modulo $p-1$.*

Remarque VI.2.2

(i) La continuité p-adique de la fonction $\zeta_{p,i}$ se traduit par des congruences du type de celles de la proposition VI.1.3.

(ii) Le théorème ci-dessus est dû à Kubota et Leopoldt et la fonction $\zeta_{p,i}$ est appelée la i-ème branche de la fonction zêta de Kubota-Leopoldt. Si i est pair, alors $\zeta_{p,i}$ est identiquement nulle car $\zeta(-n) = 0$ si $n \geqslant 2$ est pair.

(iii) On voit que pour rendre la fonction $n \to \zeta(-n)$ p-adiquement continue, on a été forcé de se restreindre à une classe de congruence modulo $p-1$ et surtout de multiplier $\zeta(-n)$ par le facteur $(1-p^n)$ qui est le facteur d'Euler en p de la fonction ζ. L'explication folklorique de ce phénomène est en général la suivante. On a $\zeta(s) = \prod_\ell \frac{1}{1-\ell^{-s}}$.

Si $\ell \neq p$, on peut, en se restreignant à une classe de congruence modulo $p-1$ (*cf.* n° suivant), prolonger la fonction $n \to \ell^n$ en une fonction continue sur $\mathbf{Z}_p$; par contre, il n'y a rien à faire avec le facteur $(1-p^n)$ qui tend p-adiquement vers 1 quand n tend vers $+\infty$. Il semble donc normal d'être forcé de retirer ce dernier facteur si on veut que le produit soit p-adiquement continu. Malheureusement, cette explication séduisante est un petit peu trop simpliste pour être juste.

1. Interpolation p-adique de la fonction $x \mapsto x^n$

De manière générale, étant donnée une suite $(a_n)_{n\in\mathbf{N}}$ d'éléments de $\mathbf{C}_p$, on peut essayer de les interpoler p-adiquement, c'est-à-dire construire une fonction continue $f : \mathbf{Z}_p \to \mathbf{C}_p$ telle que l'on ait $f(n) = a_n$ quel que soit $n \in \mathbf{N}$. Comme $\mathbf{N}$ est dense dans $\mathbf{Z}_p$, une telle fonction, si elle existe, est unique.

Considérons pour commencer l'exemple simple de la suite $a_n = x^n$, où $x \in \mathbf{Z}_p$. Si $x \in 1 + p\mathbf{Z}_p$, la fonction $f_x(s) = \sum_{n=0}^{+\infty} \binom{s}{n}(x-1)^n$ est une fonction continue de $s \in \mathbf{Z}_p$ et on a $f_x(n) = x^n$, ce qui fait que dans ce cas, la suite de terme général x^n est p-adiquement interpolable par la fonction $f_x(s)$ que nous noterons plutôt $s \to x^s$.

Si $x \in p\mathbf{Z}_p$, la suite de terme général x^n tend vers 0 quand n tend vers $+\infty$ et n'est donc pas interpolable car si $s \in \mathbf{Z}_p$, on devrait avoir $x^s = \lim_{n\to s} x^n = 0$ quel que soit $s \in \mathbf{Z}_p$, ce qui est absurde si $s \in \mathbf{N}$.

Supposons $p \neq 2$. Si $x \in \mathbf{Z}_p^*$, la situation est un peu meilleure car, si $\omega(x)$ est la racine $(p-1)$-ième de l'unité vérifiant $x - \omega(x) \in p\mathbf{Z}_p$, on peut écrire x sous la forme $\omega(x)\langle x\rangle$ et x^n sous la forme $\omega(x)^n\langle x\rangle^n$. Comme $\langle x\rangle \in 1+p\mathbf{Z}_p$, la fonction $n \to \langle x\rangle^n$ se prolonge par continuité et $\omega(x)$ étant une racine $p-1$-ième de l'unité, la fonction $n \to \omega(x)^n$ est périodique de période $p-1$. Ceci fait que, si on fixe $i \in \mathbf{Z}/(p-1)\mathbf{Z}$ et si $x \in \mathbf{Z}_p^*$, la fonction $x \to x^n$ de $i + (p-1)\mathbf{N}$ dans $\mathbf{Z}_p$ se prolonge par continuité en une fonction continue sur $\mathbf{Z}_p$, ce qui permet de voir $x \to x^s$ comme une fonction continue sur $(\mathbf{Z}/(p-1)\mathbf{Z})\times\mathbf{Z}_p$ ou comme une fonction multivaluée sur $\mathbf{Z}_p$.

Ce qui précède marche encore pour $p = 2$, mais on préfère en général modifier un peu les choses pour tenir compte du fait que $\mathbf{Z}_2^*$

contient 2 racines de l'unité (1 et -1) même si celles-ci ont même réduction modulo 2. On a $\mathbf{Z}_2^* = \{\pm 1\} \times (1 + 4\mathbf{Z}_2)$ et on note ω le caractère de $\mathbf{Z}_2^*$ défini par $\omega(x) = 1$ si $x \in 1 + 4\mathbf{Z}_2$ et $\omega(x) = -1$ si $x \in -1 + 4\mathbf{Z}_2$.

Pour avoir des formules uniformes, on pose $q = 4$ si $p = 2$ et $q = p$ si $p \neq 2$, et on note Δ le groupe des racines de l'unité contenues dans $\mathbf{Q}_p^*$, ce qui fait que si φ désigne la fonction indicatrice d'Euler (*i.e.* $\varphi(n) = \mathrm{card}((\mathbf{Z}/n\mathbf{Z})^*)$), alors Δ est le groupe (cyclique) des racines $\varphi(q)$-ième de l'unité et $\mathbf{Z}_p^*$ est la réunion disjointe des $\varepsilon + q\mathbf{Z}_p$ pour $\varepsilon \in \Delta$. Finalement, on note encore ω la fonction sur $\mathbf{Z}_p$ obtenue en prolongeant ω sur $\mathbf{Z}_p^*$ par 0 sur $p\mathbf{Z}_p$.

Proposition VI.2.3. *Si $i \in \mathbf{Z}/\varphi(q)\mathbf{Z}$, la fonction $\omega(x)^i\langle x\rangle^s$ est une fonction localement analytique de $x \in \mathbf{Z}_p$ et de plus,*

$$\omega(x)^i\langle x\rangle^n = x^n \text{ si } n \equiv i\ [\varphi(q)] \text{ et } x \in \mathbf{Z}_p^*,$$
$$\omega(x)^i\langle x\rangle^s = \lim_{\substack{n \to s \\ n \equiv i\ [\varphi(q)]}} x^n \text{quel que soit } x \in \mathbf{Z}_p.$$

Démonstration

L'analyticité locale vient de ce que l'on a $\omega(x)^i\langle x\rangle^s = 0$ sur $p\mathbf{Z}_p$ et

$$\omega(x)^i\langle x\rangle^s = \varepsilon^i\Big(\frac{x}{\varepsilon}\Big)^s = \sum_{n=0}^{+\infty} \binom{s}{n} \varepsilon^{i-n}(x-\varepsilon)^n$$

si $x \in \varepsilon + q\mathbf{Z}_p$ et $\varepsilon \in \Delta$. Le reste de la proposition suit de ce que Δ étant d'ordre $\varphi(q)$, on a $\omega(x)^n = \omega(x)^i$ si $n \equiv i\ [\varphi(q)]$.

2. Transformée de Mellin p-adique et transformée Γ de Leopoldt

Si $i \in \mathbf{Z}/\varphi(q)\mathbf{Z}$, on définit la i-ème branche $\mathrm{Mel}_{i,\mu}$ de la transformée de Mellin d'une distribution continue μ par la formule

$$\mathrm{Mel}_{i,\mu}(s) = \int_{\mathbf{Z}_p} \omega(x)^i\langle x\rangle^s\,\mu(x) = \int_{\mathbf{Z}_p^*} \omega(x)^i\langle x\rangle^s\,\mu(x),$$

la seconde égalité résultant du fait que $\omega(x) = 0$ si $x \in p\mathbf{Z}_p$. D'autre part, on a $\mathrm{Mel}_{i,\mu}(n) = \int_{\mathbf{Z}_p^*} x^n\,\mu$ si $n \equiv i\ [\varphi(q)]$.

On préfère souvent définir la transformée de Mellin d'une distribution μ sur $\mathbf{Z}_p^*$ comme la fonction qui à un caractère localement analytique ψ de $\mathbf{Z}_p^*$ à valeurs dans $\mathbf{C}_p^*$ associe l'intégrale

$$\mathrm{Mel}_\mu(\psi) = \int_{\mathbf{Z}_p^*} \psi(x)\,\mu(x).$$

On retrouve l'autre définition de la transformée de Mellin en évaluant cette transformée de Mellin en le caractère $\omega(x)^i\langle x\rangle^s$ et on a donc la formule

$$\mathrm{Mel}_{i,\mu}(s) = \mathrm{Mel}_\mu(\omega(x)^i\langle x\rangle^s).$$

Soit $\varphi : 1+q\mathbf{Z}_p \to \mathbf{Z}_p$ le morphisme de groupes qui à x associe $\frac{\log_p x}{q}$. Ce morphisme est analytique et son inverse aussi, ce qui fait que si f est une fonction localement analytique (resp. continue) sur $\mathbf{Z}_p$, la fonction φ^*f définie par $\varphi^*f(x) = f(\varphi(x))$ est localement analytique sur $1+q\mathbf{Z}_p$ (resp. continue).

Si μ est une distribution à support dans $1+q\mathbf{Z}_p$, on définit la distribution $\varphi_*\mu$ sur $\mathbf{Z}_p$ par la formule

$$\int_{\mathbf{Z}_p} f(x)\,\varphi_*\mu(x) = \int_{1+q\mathbf{Z}_p} \varphi^*f(y)\,\mu(y).$$

Et comme φ^* transforme une fonction continue sur $\mathbf{Z}_p$ en une fonction continue sur $1+q\mathbf{Z}_p$, l'image d'une mesure par φ_* est encore une mesure. (On montre plus généralement que l'image d'une distribution d'ordre r par φ_* est une distribution d'ordre r.)

Si μ est une distribution et $\alpha \in \mathbf{Z}_p^*$, on définit la distribution $\delta_\alpha \star \mu$ par la formule

$$\int_{\mathbf{Z}_p} f(x)\,\delta_\alpha\star\mu(x) = \int_{\mathbf{Z}_p} f(\alpha x)\,\mu(x).$$

Lemme VI.2.4. *Si* X *est un ouvert compact de* $\mathbf{Z}_p$, $\alpha \in \mathbf{Z}_p^*$ *et* μ *est une distribution continue sur* $\mathbf{Z}_p$, *alors*

$$\mathrm{Res}_{\mathrm{X}}(\delta_\alpha\star\mu) = \delta_\alpha \star \mathrm{Res}_{\alpha^{-1}\mathrm{X}}(\mu).$$

Démonstration. Comme on a $\mathbf{1}_{\mathrm{X}}(\alpha x) = \mathbf{1}_{\alpha^{-1}\mathrm{X}}(x)$ si $\mathrm{X} \subset \mathbf{Z}_p$, on en déduit la formule

$$\begin{aligned}\int_{\mathbf{Z}_p} f(x)\,\mathrm{Res}_{\mathrm{X}}(\delta_\alpha \star \mu) &= \int_{\mathbf{Z}_p} \mathbf{1}_{\mathrm{X}}(x) f(x)\,\delta_\alpha \star \mu \\ &= \int_{\mathbf{Z}_p} \mathbf{1}_{\mathrm{X}}(\alpha \mathrm{X}) f(\alpha x)\,\mu(x) = \int_{\mathbf{Z}_p} f(\alpha x)(1_{\alpha^{-1}\mathrm{X}}(x)\,\mu(x)) \\ &= \int_{\mathbf{Z}_p} f(\alpha x)\,\mathrm{Res}_{\alpha^{-1}\mathrm{X}}(\mu) = \int_{\mathbf{Z}_p} f(x)\,\delta_\alpha \star \mathrm{Res}_{\alpha^{-1}\mathrm{X}}(\mu),\end{aligned}$$

ce qui permet de conclure.

Définition VI.2.5. Si μ est une distribution sur $\mathbf{Z}_p^*$ et $i \in \mathbf{Z}/\varphi(q)\mathbf{Z}$, on définit la i-ième branche $\Gamma_\mu^{(i)}$ de la transformée Γ de μ par la formule

$$\Gamma_\mu^{(i)} = \varphi_* \mathrm{Res}_{1+q\mathbf{Z}_p}\Big(\sum_{\varepsilon\in\Delta} \varepsilon^{-i}\delta_\varepsilon \star \mu\Big) = \varphi_*\Big(\sum_{\varepsilon\in\Delta} \varepsilon^{-i}\delta_\varepsilon \star \mathrm{Res}_{\varepsilon^{-1}+q\mathbf{Z}_p}(\mu)\Big),$$

l'égalité entre les deux définitions résultant du lemme précédent.

Proposition VI.2.6. *Soit $u = e^q \in 1 + q\mathbf{Z}_p$. Si μ est une distribution continue et $i \in \mathbf{Z}/\varphi(q)\mathbf{Z}$, alors*

$$\mathrm{Mel}_{i,\mu}(s) = \int_{\mathbf{Z}_p^*} \omega(x)^i \langle x\rangle^s\,\mu(x) = \int_{\mathbf{Z}_p} u^{sy}\,\Gamma_\mu^{(i)}(y) = \mathscr{A}_{\Gamma_\mu^{(i)}}(u^s - 1).$$

Démonstration. La première (resp. dernière) égalité est une conséquence de la définition de la transformée de Mellin (resp. d'Amice) d'une distribution continue. Si $y = \varphi(x) = \frac{\log_p x}{q}$, on a $u^{sy} = \exp(s\log_p x) = \langle x\rangle^s$ et donc

$$\int_{\mathbf{Z}_p} u^{sy}\,\Gamma_\mu^{(i)}(y) = \int_{1+q\mathbf{Z}_p} \langle x\rangle^s \sum_{\varepsilon\in\Delta} \varepsilon^{-i}\delta_\varepsilon \star \mathrm{Res}_{\varepsilon^{-1}+q\mathbf{Z}_p}(\mu).$$

Utilisant le fait que $\omega(x) = \varepsilon^{-1}$ si $x \in \varepsilon^{-1} + q\mathbf{Z}_p$ et que $\langle \varepsilon x\rangle = \langle x\rangle$, on obtient

$$\int_{\mathbf{Z}_p} u^{sy}\,\Gamma_\mu^{(i)}(y) = \sum_{\varepsilon\in\Delta} \int_{\varepsilon^{-1}+q\mathbf{Z}_p} \omega(x)^i \langle x\rangle^s\,\mu(x),$$

et le résultat suit de ce que $\mathbf{Z}_p^*$ est la réunion disjointe des $\varepsilon + q\mathbf{Z}_p$ pour $\varepsilon \in \Delta$.

Corollaire VI.2.7. *Si μ est une distribution continue et $i \in \mathbf{Z}/\varphi(q)\mathbf{Z}$, la fonction $\mathrm{Mel}_{i,\mu}(s)$ est une fonction analytique de s et même de $u^s - 1$.*

VI.3. Construction de la fonction zêta de Kubota-Leopoldt

1. Première construction

La série formelle $\frac{\log(1+\mathrm{T})}{\mathrm{T}}$ convergeant sur $\mathrm{B}(0,1^-)$, il existe une distribution $\mu_{\mathrm{K-L}}$ dont c'est la transformée d'Amice. La transformée de Laplace de $\mu_{\mathrm{K-L}}$ est $\frac{t}{e^t-1} = f_0(t)$ et

$$\int_{\mathbf{Z}_p} x^n \, \mu_{\mathrm{K-L}} = f_0^{(n)}(0) = (-1)^{n-1} n\zeta(1-n),$$

Comme on le constate en utilisant le théorème I.1.3. Cette distribution ressemble beaucoup à la mesure de Haar (mais n'est pas invariante par translation). On a en effet le lemme suivant.

Lemme VI.3.1. $\displaystyle\int_{a+p^n\mathbf{Z}_p} \mu_{\mathrm{K-L}}(x) = \frac{1}{p^n}.$

Démonstration. $\int_{a+p^n\mathbf{Z}_p} \mu_{\mathrm{K-L}}(x) = \frac{1}{p^n} \sum_{\varepsilon^{p^n}=1} \varepsilon^{-a} \mathscr{A}_{\mu_{\mathrm{K-L}}}(\varepsilon - 1)$ et comme $\log(\varepsilon) = 0$ si ε est une racine de l'unité d'ordre une puissance de p, tous les termes de la somme sont nuls sauf celui correspondant à $\varepsilon = 1$, ce qui donne le résultat.

On a $\mathbf{1}_{\mathbf{Z}_p^*}(x) = 1 - \mathbf{1}_{p\mathbf{Z}_p}(x) = 1 - \frac{1}{p}\sum_{\varepsilon^p=1} \varepsilon^x$ et donc, si λ est une distribution continue sur $\mathbf{Z}_p$, alors

$$\begin{aligned}
\mathscr{A}_{\mathrm{Res}_{\mathbf{Z}_p^*}(\lambda)}(\mathrm{T}) &= \int_{\mathbf{Z}_p^*} (1+\mathrm{T})^x \, \lambda(x) \\
&= \int_{\mathbf{Z}_p} (1+\mathrm{T})^x \lambda(x) - \frac{1}{p} \sum_{\varepsilon^p=1} \int_{\mathbf{Z}_p} ((1+\mathrm{T})\varepsilon)^x \lambda(x) \\
&= \mathscr{A}_\lambda(\mathrm{T}) - \frac{1}{p} \sum_{\varepsilon^p=1} \mathscr{A}_\lambda((1+\mathrm{T})\varepsilon - 1),
\end{aligned}$$

Lemme VI.3.2. $\displaystyle\frac{1}{p} \sum_{\varepsilon^p=1} \frac{1}{\varepsilon z - 1} = \frac{1}{z^p - 1}.$

Démonstration. Les deux membres sont des fractions rationnelles en z et, si $|z|_p < 1$, on a

$$\frac{1}{p}\sum_{\varepsilon^p=1}\frac{1}{\varepsilon z-1}=-\frac{1}{p}\sum_{\varepsilon^p=1}\sum_{n=0}^{+\infty}(\varepsilon z)^n=-\sum_{n\equiv 0\ [p]}z^n=\frac{1}{z^p-1},$$

d'où le résultat.

La transformée d'Amice de $\mathrm{Res}_{\mathbf{Z}_p^*}(\mu_{\mathrm{K-L}})$ est donc

$$\frac{\log(1+\mathrm{T})}{\mathrm{T}}-\frac{1}{p}\sum_{\varepsilon^p=1}\frac{\log((1+\mathrm{T})\varepsilon)}{(1+\mathrm{T})\varepsilon-1}=\frac{\log(1+\mathrm{T})}{\mathrm{T}}-\frac{\log(1+\mathrm{T})}{(1+\mathrm{T})^p-1}$$
$$=\mathscr{A}_{\mu_{\mathrm{K-L}}}(\mathrm{T})-\frac{1}{p}\mathscr{A}_{\mu_{\mathrm{K-L}}}((1+\mathrm{T})^p-1).$$

On en tire les formules

$$\mathscr{L}_{\mathrm{Res}_{\mathbf{Z}_p^*}(\mu_{\mathrm{K-L}})}(t)=\mathscr{L}_{\mu_{\mathrm{K-L}}}(t)-\mathscr{L}_{\mu_{\mathrm{K-L}}}(pt)=f_0(t)-\frac{1}{p}f_0(pt)$$
$$\int_{\mathbf{Z}_p^*}x^n\mu_{\mathrm{K-L}}(x)=(1-p^{n-1})f_0^{(n)}(0)=(-1)^{n-1}n(1-p^{n-1})\zeta(1-n).$$

Pour définir la fonction zêta de Kubota-Leopoldt, il suffit alors de poser, si $i\in\mathbf{Z}/\varphi(q)\mathbf{Z}$,

$$\zeta_{p,i}(s)=\frac{(-1)^{i-1}}{s-1}\mathrm{Mel}_{1-i,\mu_{\mathrm{K-L}}}(1-s)$$
$$=\frac{(-1)^{i-1}}{s-1}\int_{\mathbf{Z}_p^*}\omega(x)^{1-i}\langle x\rangle^{1-s}\mu_{\mathrm{K-L}}(x).$$

Par construction, on a $\zeta_{p,i}(-n)=(1-p^n)\zeta(-n)$ si $n\in\mathbf{N}$ vérifie $-n\equiv i\ [p-1]$. D'autre part, la fonction $\mathrm{Mel}_{1-i,\mu_{\mathrm{K-L}}}(1-s)$ étant analytique sur $\mathbf{Z}_p$, la fonction $(s-1)\zeta_{p,i}(s)$ est analytique sur $\mathbf{Z}_p$ et

$$\lim_{s\to 1}(s-1)\zeta_{p,i}(s)=\int_{\mathbf{Z}_p^*}\omega(x)^{1-i}\mu_{\mathrm{K-L}}(x)$$
$$=\sum_{\alpha\in\mathbf{Z}_p^*/(1+p\mathbf{Z}_p)}\omega(\alpha)^{1-i}\int_{\alpha+p\mathbf{Z}_p}\mu_{\mathrm{K-L}}(x)=\begin{cases}1-1/p & \text{si } i=1,\\ 0 & \text{sinon.}\end{cases}$$

On en déduit le fait que si $i\neq 1$, la fonction $\zeta_{p,i}$ peut se prolonger par continuité en $s=1$ en une fonction analytique sur $\mathbf{Z}_p$.

Remarque VI.3.3. La formule $\lim_{s\to 1}(s-1)\zeta_{p,1}(s) = 1-1/p$ est à rapprocher de la formule analogue pour la fonction zêta de Riemann. La différence entre les deux formules est encore une fois donnée par un facteur d'Euler en p.

2. Deuxième construction

On peut aussi partir de la mesure μ_a introduite au § VI.1.

Lemme VI.3.4. *Les transformées d'Amice et Laplace de la restriction de μ_a à $\mathbf{Z}_p^*$ sont données par les formules*

$$\mathscr{A}_{\mathrm{Res}_{\mathbf{Z}_p^*}(\mu_a)}(\mathrm{T}) = \mathscr{A}_{\mu_a}(\mathrm{T}) - \mathscr{A}_{\mu_a}((1+\mathrm{T})^p - 1)$$
$$\mathscr{L}_{\mathrm{Res}_{\mathbf{Z}_p^*}(\mu_a)}(t) = \mathscr{L}_{\mu_a}(t) - \mathscr{L}_{\mu_a}(pt) = f_a(t) - f_a(pt).$$

Démonstration. C'est le même calcul que celui effectué pour calculer les transformées d'Amice et Laplace de la restriction de $\mu_{\mathrm{K-L}}$ à $\mathbf{Z}_p^*$.

On en déduit la formule

$$\int_{\mathbf{Z}_p^*} x^n \mu_a(x) = (1-p^n) f_a^{(n)}(0) = (-1)^n (1-a^{1+n})(1-p^n)\zeta(-n)$$

qui montre que restreindre à $\mathbf{Z}_p^*$ fait apparaître un facteur d'Euler en p.

Si $i \in \mathbf{Z}/\varphi(q)\mathbf{Z}$, et si $a \in \mathbf{Z}_p^*$ vérifie $\langle a\rangle \neq 1$, définissons une fonction $g_{a,i}$ sur $\mathbf{Z}_p$ par la formule

$$\begin{aligned} g_{a,i}(s) &= \frac{1}{1-\omega(a)^{1-i}\langle a\rangle^{1-s}} \mathrm{Mel}_{-i,\mu_a}(-s) \\ &= \frac{1}{1-\omega(a)^{1-i}\langle a\rangle^{1-s}} \int_{\mathbf{Z}_p^*} \omega(x)^{-i}\langle x\rangle^{-s}\mu_a(x). \end{aligned}$$

D'après le corollaire VI.2.7, $\mathrm{Mel}_{-i,\mu_a}(-s)$ est une fonction analytique de s. D'autre part, si $\omega(a)^{1-i} \neq 1$, la fonction $s \to 1-\omega(a)^{1-i}\langle a\rangle^{1-s}$ est une fonction analytique de s ne s'annulant pas sur $\mathbf{Z}_p$ car $\langle a\rangle^s \in 1+q\mathbf{Z}_p$ et $\omega(a)^{1-i} \in \Delta - \{1\}$ et donc $\omega(a)^{i-1} \notin 1+q\mathbf{Z}_p$; si $\omega(a)^{1-i} = 1$, la fonction $1-\langle a\rangle^{1-s}$ ne s'annule que pour $s=1$. On en déduit le fait que $g_{a,i}$ est une fonction continue sur $\mathbf{Z}_p - \{1\}$ et même sur $\mathbf{Z}_p$ si $\omega(a)^{1-i} \neq 1$.

De plus, si $-n \equiv i\ [\varphi(q)]$, on a $\omega(a)^{1-i} = \omega(a)^{1+n}$ et $\omega(x)^{-i} = \omega(x)^n$ si $x \in \mathbf{Z}_p^*$ et donc

$$g_{a,i}(-n) = \frac{1}{1-\omega(a)^{1+n}\langle a\rangle^{1+n}} \int_{\mathbf{Z}_p^*} \omega(x)^n \langle x\rangle^n \mu_a(x)$$
$$= \frac{1}{1-a^{1+n}} \int_{\mathbf{Z}_p^*} x^n \mu_a(x) = (-1)^n(1-p^n)\zeta(-n)$$

ne dépend pas du choix de a. Si a et a' sont 2 éléments de $\mathbf{Z}_p^*$, la fonction $g_{a,i} - g_{a',i}$ est donc un quotient de fonctions analytiques sur $\mathbf{Z}_p$ s'annulant en un nombre infini de points, ce qui implique qu'elle est identiquement nulle et que la fonction $g_{a,i}$ est indépendante du choix de a. Il suffit donc de poser $\zeta_{p,i} = g_{a,i}$ pour n'importe quel choix de a vérifiant $\langle a\rangle \neq 1$ et $\omega(a)^{1-i} \neq 1$ si $i \neq 1$ pour avoir une construction de la fonction zêta de Kubota-Leopoldt.

Remarque VI.3.5. Si $p \neq 2$, on peut aussi conclure au fait que $g_{a,i} - g_{a',i}$ est identiquement nulle en utilisant le fait que c'est une fonction continue sur $\mathbf{Z}_p - \{1\}$ s'annulant en tous les éléments de $i - (p-1)\mathbf{N}$ qui est un sous-ensemble dense de $\mathbf{Z}_p$.

VI.4. Les zéros de la fonction zêta *p*-adique

Comme nous l'avons signalé dans l'introduction, la fonction zêta p-adique est étroitement reliée aux groupes de classes d'idéaux des corps $\mathbf{Q}(e^{2i\pi/p^n})$, pour $n \in \mathbf{N}$. Le théorème VI.4.1 ci-dessous est une conséquence d'un théorème plus précis de Mazur et Wiles et donne une bonne illustration de ce lien.

L'énoncé de ce théorème va demander un peu de préparation. Tout d'abord, si K/F est une extension finie de corps de nombres, et si $\mathfrak{a}$ est un idéal non nul de l'anneau des entiers $\mathscr{O}_\mathrm{K}$ de K, on définit l'idéal $\mathrm{N}_{\mathrm{K/F}}(\mathfrak{a})$ comme l'idéal de $\mathscr{O}_\mathrm{F}$ engendré par les $\mathrm{N}_{\mathrm{K/F}}(\alpha)$, pour $\alpha \in \mathfrak{a}$. On a $\mathrm{N}_{\mathrm{K/F}}(\mathfrak{ab}) = \mathrm{N}_{\mathrm{K/F}}(\mathfrak{a})\mathrm{N}_{\mathrm{K/F}}(\mathfrak{b})$ et $\mathrm{N}_{\mathrm{K/F}}((\alpha)) = (\mathrm{N}_{\mathrm{K/F}}(\alpha))$, ce qui montre que $\mathrm{N}_{\mathrm{K/F}}$ induit, par passage aux quotients, un morphisme de groupes du groupe des classes d'idéaux de K dans celui des classes d'idéaux de F. Ce morphisme envoie le p-Sylow dans le p-Sylow puisque le p-Sylow d'un groupe abélien fini n'est autre que l'ensemble des éléments d'ordre une puissance de p.

Dans tout ce qui suit, on suppose $p \neq 2$. Si $n \geqslant 1$, on note F_n le corps cyclotomique $\mathbf{Q}(e^{2i\pi/p^n})$ et X_n le p-Sylow du groupe des classes d'idéaux de F_n. D'après la discussion précédente, l'application $\mathrm{N}_{\mathrm{F}_{n+1}/\mathrm{F}_n}$ induit un morphisme de groupes de X_{n+1} dans X_n. On note X la limite projective des X_n relativement aux applications $\mathrm{N}_{\mathrm{F}_{n+1}/\mathrm{F}_n}$. Un élément c de X est donc une suite $(c_n)_{n\geqslant 1}$, avec $c_n \in \mathrm{X}_n$ et $c_n = \mathrm{N}_{\mathrm{F}_{n+1}/\mathrm{F}_n}(c_{n+1})$ quel que soit $n \geqslant 1$. Comme chaque X_n est un p-groupe abélien fini et donc un $\mathbf{Z}_p$-module, X est un $\mathbf{Z}_p$-module compact.

On note F_∞ la réunion des F_n. C'est une extension galoisienne de $\mathbf{Q}$ et son groupe de Galois est canoniquement isomorphe à $\mathbf{Z}_p^*$ via le caractère cyclotomique χ_{cycl} : en effet, si ε est une racine primitive p^n-ième de l'unité, alors tout conjugué de ε est de la forme ε^a, avec $a \in (\mathbf{Z}/p^n\mathbf{Z})^*$, et si $\sigma \in \mathrm{Gal}(\mathrm{F}_\infty/\mathbf{Q})$, alors $\chi_{\mathrm{cycl}}(\sigma)$ est l'unique élément $a \in \mathbf{Z}_p^*$ (rappelons que $\mathbf{Z}_p^*$ est la limite projective des $(\mathbf{Z}/p^n\mathbf{Z})^*$) tel que l'on ait $\sigma(\varepsilon) = \varepsilon^a$ pour toute racine de l'unité d'ordre une puissance de p.

Le groupe $\mathrm{Gal}(\mathrm{F}_\infty/\mathbf{Q})$ laisse stable chaque F_n, respecte l'anneau des entiers, et donc transforme un idéal en un idéal et un idéal principal en un idéal principal et, par suite, agit sur X_n. Comme cette action commute aux applications $\mathrm{N}_{\mathrm{F}_{n+1}/\mathrm{F}_n}$, on obtient une action de $\mathrm{Gal}(\mathrm{F}_\infty/\mathbf{Q})$ sur X.

Théorème VI.4.1. *Si $i \in (\mathbf{Z}/(p-1)\mathbf{Z})^*$ est impair et si $s \in \mathbf{Z}_p$, alors les deux conditions suivantes sont équivalentes* :

(i) $\zeta_{p,i}(s) = 0$;

(ii) *il existe un élément c de* X *qui n'est pas tué par une puissance de p et sur lequel $\sigma \in \mathrm{Gal}(\mathrm{F}_\infty/\mathbf{Q})$ agit via la formule*

$$\sigma(c) = \omega(\chi_{\mathrm{cycl}}(\sigma))^i \langle \chi_{\mathrm{cycl}}(\sigma)\rangle^s \cdot c.$$

Le théorème précédent caractérise les zéros de la fonction zêta p-adique mais n'est pas très explicite : on ne sait, par exemple, pas démontrer l'énoncé suivant qui reste le principal problème ouvert concernant la fonction zêta p-adique.

Conjecture VI.4.2. *Si $i \in (\mathbf{Z}/(p-1)\mathbf{Z})^*$ est impair, et si k est un entier $\geqslant 1$, alors $\zeta_{p,i}(k) \neq 0$.*

On sait démontrer ce résultat pour $k = 1$, mais cela résulte d'un théorème profond sur les formes linéaires de logarithmes de nombres algébriques (*cf.* n° 2 du § VI.5). Pour traiter le cas général, il faudrait disposer d'un résultat analogue pour les polylogarithmes.

VI.5. Fonctions L *p*-adiques attachées aux caractères de Dirichlet

Ce que l'on vient de faire pour la fonction zêta de Riemann s'étend sans douleur aux fonctions L de Dirichlet (à l'exception du théorème de Mazur-Wiles qui s'étend, mais pas sans douleur).

1. Construction

Soit χ un caractère de Dirichlet de conducteur $d > 1$ premier à p. On note ε_d la racine de l'unité $\varepsilon_d = e^{2i\pi/d}$. Si $\chi^{-1}(b) \neq 0$, alors ε_d^b est une racine de l'unité d'ordre premier à p et distincte de 1, ce qui implique $|\varepsilon_d^b - 1|_p = 1$. On en déduit le fait que la série entière

$$\begin{aligned} \mathrm{F}_\chi(\mathrm{T}) &= \frac{-1}{\mathrm{G}(\chi^{-1})} \sum_{b \bmod d} \frac{\chi^{-1}(b)}{(1+\mathrm{T})\varepsilon_d^b - 1} \\ &= \frac{1}{\mathrm{G}(\chi^{-1})} \sum_{b \bmod d} \chi^{-1}(b) \sum_{n=0}^{+\infty} \frac{\varepsilon_d^{nb}}{(\varepsilon_d^b - 1)^{n+1}} \mathrm{T}^n \end{aligned}$$

est à coefficients bornés et donc est la transformée d'Amice d'une mesure μ_χ sur $\mathbf{Z}_p$ dont la transformée de Laplace est $\mathrm{F}_\chi(e^t - 1) = \mathscr{L}_\chi(t)$. On a donc

$$\int_{\mathbf{Z}_p} x^n \mu_\chi = \mathscr{L}_\chi^{(n)}(0) = \mathrm{L}(\chi, -n)$$

d'après le n° 2 du § I.5.

Définition VI.5.1. On définit la fonction-L p-adique associée à χ comme étant la transformée de Mellin de μ_χ et on note cette fonction

$\psi \to \mathrm{L}_p(\chi \otimes \psi)$. Si ψ est un caractère localement analytique sur $\mathbf{Z}_p^*$, on a donc

$$\mathrm{L}_p(\chi \otimes \psi) = \int_{\mathbf{Z}_p^*} \psi(x)\mu_\chi(x).$$

D'autre part, si $i \in \mathbf{Z}/\varphi(q)\mathbf{Z}$, on pose

$$\mathrm{L}_{p,i}(\chi, s) = \mathrm{L}_p(\chi \otimes (\omega^{-i}(x)\langle x\rangle^{-s})) = \int_{\mathbf{Z}_p^*} \omega^{-i}(x)\langle x\rangle^{-s}\mu_\chi(x).$$

Proposition VI.5.2. *Si $i \in \mathbf{Z}/\varphi(q)\mathbf{Z}$, la fonction $\mathrm{L}_{p,i}(\chi, s)$ est une fonction analytique sur $\mathbf{Z}_p$ et on a $\mathrm{L}_{p,i}(\chi, -n) = (1 - \chi(p)p^n)\mathrm{L}(\chi, -n)$ si $n \in \mathbf{N}$ vérifie $-n \equiv i\ [\varphi(q)]$.*

Démonstration. Le fait que $\mathrm{L}_{p,i}(\chi, s)$ soit une fonction analytique sur $\mathbf{Z}_p$ suit des propriétés générales de la transformée de Mellin d'une mesure (corollaire VI.2.7). D'autre part, d'après le lemme VI.3.2, on a

$$\sum_{\eta^p=1} \frac{1}{(1+\mathrm{T})\varepsilon_d^b\eta - 1} = p\,\frac{1}{(1+\mathrm{T})^p\varepsilon_d^{pb} - 1}$$

on en déduit le fait que la transformée d'Amice de la restriction à $\mathbf{Z}_p^*$ de μ_χ est

$$\frac{-1}{\mathrm{G}(\chi^{-1})} \sum_{b \bmod d} \frac{\chi^{-1}(b)}{(1+\mathrm{T})\varepsilon_d^b - 1} - \frac{\chi^{-1}(b)}{(1+\mathrm{T})^p\varepsilon_d^{pb} - 1},$$

ce qui, mettant $\chi^{-1}(b)$ sous la forme $\chi(p)\chi^{-1}(pb)$ et utilisant le fait que $b \to pb$ est une bijection modulo d, peut se réécrire sous la forme $\mathscr{A}_{\mu_\chi}(\mathrm{T}) - \chi(p)\mathscr{A}_{\mu_\chi}((1+\mathrm{T})^p - 1)$. On en déduit les formules

$$\mathscr{L}_{\mathrm{Res}_{\mathbf{Z}_p^*}(\mu_\chi)}(t) = \mathscr{L}_{\mu_\chi}(t) - \chi(p)\mathscr{L}_{\mu_\chi}(pt)$$

et

$$\int_{\mathbf{Z}_p^*} x^n\mu_\chi = (1 - \chi(p)p^n)\mathrm{L}(\chi, -n),$$

si $n \in \mathbf{N}$, et le résultat.

2. Comportement en $s = 1$ des fonctions L de Dirichlet

En reprenant la formule pour $\mathrm{L}(\chi, s)$ donnée au n° 2 du § I.5, on obtient

$$\begin{aligned}\mathrm{L}(\chi, 1) &= \frac{1}{\mathrm{G}(\chi^{-1})} \sum_{b \bmod d} \chi^{-1}(b) \sum_{n=0}^{+\infty} \frac{\varepsilon_d^{nb}}{n} \\ &= \frac{-1}{\mathrm{G}(\chi^{-1})} \sum_{b \bmod d} \chi^{-1}(b) \log(1 - \varepsilon_d^b).\end{aligned}$$

On peut obtenir de plus jolies formules en regroupant les contributions de b et $-b$ et en séparant le cas χ pair ($\chi(-1) = 1$) du cas χ impair ($\chi(-1) = -1$).

Nous allons établir l'analogue p-adique de cette formule. Il s'agit de calculer $\int_{\mathbf{Z}_p^*} x^{-1}\mu_\chi$. Pour ce faire, nous allons calculer la transformée d'Amice de $x^{-1}\mu_\chi$ (qui n'est déterminée qu'à constante près) puis restreindre à $\mathbf{Z}_p^*$, ce qui tue l'indétermination qui correspond à un multiple de la masse de Dirac en 0.

Proposition VI.5.3. *La transformée d'Amice de $x^{-1}\mu_\chi$ est* (à constante près) *donnée par la formule*

$$\mathscr{A}_{x^{-1}\mu_\chi}(\mathrm{T}) = \frac{-1}{\mathrm{G}(\chi^{-1})} \sum_{b \bmod d} \chi^{-1}(b) \log_p((1 + \mathrm{T})\varepsilon_d^b - 1)$$

Démonstration. Si μ est une distribution, les transformées d'Amice de μ et $x^{-1}\mu$ sont reliées par la formule

$$(1 + \mathrm{T})\frac{d}{d\mathrm{T}}\mathscr{A}_{x^{-1}\mu}(\mathrm{T}) = \mathscr{A}_\mu(\mathrm{T}).$$

Appliquons l'opérateur $(1 + \mathrm{T})\frac{d}{d\mathrm{T}}$ au membre de droite de l'identité à vérifier ; on obtient

$$\begin{aligned}&\frac{-1}{\mathrm{G}(\chi^{-1})} \sum_{b \bmod d} \chi^{-1}(b) \frac{(1 + \mathrm{T})\varepsilon_d^b}{(1 + \mathrm{T})\varepsilon_d^b - 1} \\ &\qquad = \frac{-1}{\mathrm{G}(\chi^{-1})} \sum_{b \bmod d} \chi^{-1}(b) \Big(\frac{1}{(1 + \mathrm{T})\varepsilon_d^b - 1} + 1 \Big)\end{aligned}$$

et cette dernière expression est égale à $\mathscr{A}_{\mu_\chi}(\mathrm{T})$ comme on le voit en utilisant le fait que $\sum_{b \bmod d} \chi^{-1}(b) = 0$. On en déduit le fait que les

deux membres ont même image par $(1+\mathrm{T})\frac{d}{d\mathrm{T}}$ et donc qu'ils diffèrent par une fonction localement constante. Pour conclure, il faut encore vérifier que le second membre est bien donné par une série de rayon de convergence 1 ; mais on a

$$\begin{aligned}\log_p((1+\mathrm{T})\varepsilon_d^b - 1) &= \log_p(\varepsilon_d^b - 1) + \log_p(1 + \frac{\varepsilon_d^b \mathrm{T}}{\varepsilon_d^b - 1}) \\ &= \log_p(\varepsilon_d^b - 1) + \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n}\Big(\frac{\varepsilon_d^b \mathrm{T}}{\varepsilon_d^b - 1}\Big)^n\end{aligned}$$

et comme on a supposé $(d,p)=1$, on a $|\varepsilon_d^b - 1|_p = 1$ et la série est bien de rayon de convergence 1. Ceci permet de conclure.

Lemme VI.5.4. *La transformée d'Amice de la restriction de $x^{-1}\mu_\chi$ à $\mathbf{Z}_p^*$ est donnée par la formule*

$$\begin{aligned}\mathscr{A}_{\mathrm{Res}_{\mathbf{Z}_p^*}(x^{-1}\mu_\chi)}(\mathrm{T}) &= \frac{-1}{\mathrm{G}(\chi^{-1})} \sum_{b \bmod d} \chi^{-1}(b) \\ &\quad \times \Big(\log_p((1+\mathrm{T})\varepsilon_d^b - 1) - \frac{1}{p}\log_p((1+\mathrm{T})^p \varepsilon_d^{pb} - 1)\Big) \\ &= \mathscr{A}_{x^{-1}\mu_\chi}(\mathrm{T}) - \frac{\chi(p)}{p}\mathscr{A}_{x^{-1}\mu_\chi}((1+\mathrm{T})^p - 1).\end{aligned}$$

Démonstration. On utilise la formule générale donnant la transformé d'Amice de la restriction à $\mathbf{Z}_p^*$ d'une mesure en fonction de celle de la mesure et l'identité

$$\sum_{\eta^p=1} \log_p((1+\mathrm{T})\varepsilon_d^b \eta - 1) = \log_p((1+\mathrm{T})^p \varepsilon_d^{pb} - 1),$$

ce qui permet de montrer la première des deux égalités ; la seconde se démontre en écrivant $\chi^{-1}(b)$ sous la forme $\chi(p)\chi^{-1}(bp)$ et en utilisant le fait que $b \to pb$ est une bijection modulo d comme nous l'avons déjà fait.

En posant $\mathrm{T}=0$ dans la formule précédente, on obtient

$$\int_{\mathbf{Z}_p^*} x^{-1}\mu_\chi = \frac{-1}{\mathrm{G}(\chi^{-1})}\Big(1 - \frac{\chi(p)}{p}\Big) \sum_{b \bmod d} \chi^{-1}(b) \log_p(\varepsilon_d^b - 1),$$

formule qui ne diffère de la formule complexe que par un facteur d'Euler et le remplacement du logarithme usuel par le logarithme p-adique. Il s'agit d'une illustration d'un phénomène surprenant au premier abord qui fait que les formules p-adiques et les formules complexes continuent à se ressembler beaucoup même en des points où elles n'ont aucune raison de le faire a priori.

3. Torsion par un caractère de conducteur une puissance de p

Notre but dans ce paragraphe est d'étendre les résultats des deux paragraphes précédents pour calculer la fonction L p-adique de χ évaluée en un caractère de la forme $\psi(x)x^n$, où ψ est un caractère de Dirichlet de conducteur une puissance de p (vu comme caractère localement constant de $\mathbf{Z}_p^*$) et n un entier $\geqslant -1$. Nous utiliserons la notation $\chi\otimes\psi$ pour désigner le caractère de Dirichlet modulo dp^k défini par $(\chi\otimes\psi)(a)=\chi(a)\psi(a)$, où χ et ψ sont vus comme des caractères mod dp^k grâce aux projections respectives de $(\mathbf{Z}/dp^k\mathbf{Z})^*$ sur $(\mathbf{Z}/d\mathbf{Z})^*$ et $(\mathbf{Z}/p^k\mathbf{Z})^*$.

Lemme VI.5.5. *Soit $k\geqslant 1$, ψ un caractère de Dirichlet de conducteur p^k et μ une distribution continue sur $\mathbf{Z}_p$. Alors on a*

$$\int_{\mathbf{Z}_p}\psi(x)(1+\mathrm{T})^x\mu(x)=\frac{1}{\mathrm{G}(\psi^{-1})}\sum_{c \bmod p^k}\psi^{-1}(c)\mathscr{A}_\mu((1+\mathrm{T})\varepsilon_{p^k}^c-1).$$

Démonstration. On a

$$\begin{aligned}\int_{\mathbf{Z}_p}\psi(x)(1+\mathrm{T})^x\mu(x)&=\sum_{a \bmod p^k}\psi(a)\int_{a+p^k\mathbf{Z}_p}(1+\mathrm{T})^x\mu_\chi\\&=\sum_{a \bmod p^k}\psi(a)\Big(\frac{1}{p^k}\sum_{\eta^{p^k}=1}\eta^{-a}\mathscr{A}_\mu((1+\mathrm{T})\eta-1)\Big)\\&=\sum_{\eta^{p^k}=1}\mathscr{A}_\mu((1+\mathrm{T})\eta-1)\Big(\frac{1}{p^k}\sum_{a \bmod p^k}\psi(a)\eta^{-a}\Big).\end{aligned}$$

Si on écrit η sous la forme $\varepsilon_{p^k}^c=e^{2i\pi c/p^k}$, on reconnaît dans le terme entre parenthèses une somme de Gauss tordue (divisée par p^k) dont

la valeur est donnée par le lemme I.5.1 et ce terme vaut donc

$$\frac{1}{p^k}\psi^{-1}(-c)\mathrm{G}(\psi) = \frac{\psi^{-1}(c)}{\mathrm{G}(\psi^{-1})},$$

la dernière égalité provenant de la formule $\mathrm{G}(\psi)\mathrm{G}(\psi^{-1}) = \psi(-1)p^k$ (lemme I.5.1). On en tire le résultat.

Proposition VI.5.6. *Si μ est une mesure sur $\mathbf{Z}_p$ dont la transformée d'Amice est de la forme*

$$\mathscr{A}_\mu(\mathrm{T}) = \frac{-1}{\mathrm{G}(\chi^{-1})}\sum_{b \bmod d}\chi^{-1}(b)\mathrm{F}((1+\mathrm{T})\varepsilon_d^b - 1)$$

et si ψ est un caractère de Dirichlet de conducteur p^k avec $k \geqslant 1$, alors

$$\int_{\mathbf{Z}_p}\psi(x)(1+\mathrm{T})^x\mu(x) = \frac{-1}{\mathrm{G}((\chi\otimes\psi)^{-1})}\sum_{a \bmod dp^k}(\chi\otimes\psi)^{-1}(a)\mathrm{F}((1+\mathrm{T})\varepsilon_{dp^k}^a - 1).$$

Démonstration. D'après le lemme précédent, on a

$$\int_{\mathbf{Z}_p}\psi(x)(1+\mathrm{T})^x\mu(x) = \frac{-1}{\mathrm{G}(\chi^{-1})\mathrm{G}(\psi^{-1})}\sum_{b \bmod d}\sum_{c \bmod p^k}\chi^{-1}(b)\psi^{-1}(c)\mathrm{F}((1+\mathrm{T})\varepsilon_d^b\varepsilon_{p^k}^c - 1).$$

Pour mettre cette expression sous une forme un peu plus sympathique, on peut utiliser le fait que tout élément a de $\mathbf{Z}/dp^k\mathbf{Z}$ peut s'écrire de manière unique sous la forme $dc + p^k b$, avec $b \in \mathbf{Z}/d\mathbf{Z}$ et $c \in \mathbf{Z}/p^k\mathbf{Z}$, ce qui donne les formules

$$\begin{aligned}
\varepsilon_{dp^k}^a &= \varepsilon_d^b\varepsilon_{p^k}^c\\
(\chi\otimes\psi)^{-1}(a) &= \chi^{-1}(p^k)\psi^{-1}(d)\chi^{-1}(b)\psi^{-1}(c)\\
\mathrm{G}((\chi\otimes\psi)^{-1}) &= \sum_{a \bmod dp^k}(\chi\otimes\psi)^{-1}(a)\varepsilon_{dp^k}^a\\
&= \chi^{-1}(p^k)\psi^{-1}(d)\Big(\sum_{b \bmod d}\chi^{-1}(b)\varepsilon_d^b\Big)\Big(\sum_{c \bmod p^k}\psi^{-1}(c)\varepsilon_{p^k}^c\Big)\\
&= \chi^{-1}(p^k)\psi^{-1}(d)\mathrm{G}(\chi^{-1})\mathrm{G}(\psi^{-1})
\end{aligned}$$

et permet de conclure.

On peut appliquer la proposition précédente à la distribution $x^{-1}\mu_\chi$ et à la fonction $\mathrm{F}(\mathrm{T}) = \log_p(\mathrm{T})$. On obtient, en évaluant le résultat en $\mathrm{T} = 0$,

$$\begin{aligned}\mathrm{L}_p(\chi \otimes (x^{-1}\psi)) &= \int_{\mathbf{Z}_p^*} \psi(x) x^{-1} \mu_\chi \\ &= \frac{-1}{\mathrm{G}((\chi\otimes\psi)^{-1})} \sum_{x \bmod dp^k} (\chi\otimes\psi)^{-1}(x) \log_p(\varepsilon_{dp^k}^x - 1),\end{aligned}$$

formule qui est à rapprocher de la formule correspondante sur les complexes.

Proposition VI.5.7. *Si ψ est un caractère de Dirichlet non trivial de conducteur une puissance de p et $n \in \mathbf{N}$, alors $\mathrm{L}_p(\chi \otimes (x^n\psi)) = \mathrm{L}(\chi\otimes\psi, -n)$.*

Démonstration. On tire de la proposition précédente et de la formule donnant la transformée d'Amice de μ_χ, le fait que la transformée d'Amice de $\psi(x)\mu_\chi(x)$ est

$$\frac{-1}{\mathrm{G}((\chi\otimes\psi)^{-1})} \sum_{x \bmod dp^k} \frac{(\chi\otimes\psi)^{-1}(x)}{(1+\mathrm{T})\varepsilon_{dp^k}^x - 1}$$

et donc que sa transformée de Laplace est la fonction $\mathscr{L}_{\chi\otimes\psi}(t)$. Le résultat s'en déduit.

Remarque VI.5.8. Il n'y a pas de facteur d'Euler apparaissant dans les deux formules précédentes ; c'est dû au fait que p n'est pas premier au conducteur de $\chi\otimes\psi$ et donc $\chi\otimes\psi(p) = 0$.

Pierre Colmez, Équipe d'arithmétique, Institut de mathématiques, Tour 46-56 - 5ème étage- Boîte 247, 4, place Jussieu, 75252 Paris Cedex 05, France
E-mail : pierre.colmez@imj-prg.fr
Url : https://webusers.imj-prg.fr/~pierre.colmez/