Thèse de Doctorat

de l'Université Pierre et Marie Curie (Paris VI)

Spécialité : Mathématiques

présentée par M. Stéphane Fischler

Contributions à l'étude diophantienne des polylogarithmes et des groupes algébriques

Soutenue le 6 juin 2003, devant le jury composé de :

M. Francesco AMOROSO (Professeur, Caen)
M. Frits BEUKERS (Professeur, Utrecht)
M. Pierre CARTIER (Directeur de Recherches, Paris)
M. Sinnou DAVID (Maître de Conférences Habilité, Paris)
M. David MASSER (Professeur, Bâle - Rapporteur)
M. Michel WALDSCHMIDT (Professeur, Paris - Directeur)

Table des matières

Remerciements

Je tiens à remercier, en premier lieu, mon directeur Michel Waldschmidt. Il a pleinement répondu à mes attentes, simultanées, d'encadrement et d'autonomie. Il s'est toujours montré à l'écoute, intéressé, disponible et efficace. Ses "suggestions" de tout ordre m'ont été très précieuses.

Damien Roy a accepté d'être rapporteur de cette thèse. C'est le reflet d'un lien qui commence à être ancien : j'ai choisi mon premier sujet de recherche suite à un exposé de Damien (voir le chapitre 6). Le chapitre 7 doit aussi beaucoup à sa patience et à sa ténacité. J'espère continuer à travailler avec lui dans le futur avec autant de plaisir.

Merci à David Masser pour avoir accepté d'être rapporteur, et avoir effectué ce travail très rapidement malgré un emploi du temps chargé.

Je suis très heureux que Francesco Amoroso, Frits Beukers, Pierre Cartier et Sinnou David fassent partie du jury. Je les en remercie (et tout particulièrement Frits Beukers, qui a fait le déplacement jusqu'à Paris).

J'ai bénéficié, tout au long de cette thèse, d'un double statut : d'une part doctorant au sein de l'Équipe de Théorie des Nombres de l'Institut de Mathématiques de Chevaleret, d'autre part élève puis agrégé-préparateur ("caïman") au Département de Mathématiques et Applications de l'École Normale Supérieure. Je remercie ces deux structures pour l'atmosphère propice qu'elles m'ont fournie : les enseignants, les chercheurs et les personnels administratifs allient compétence et disponibilité. Je tiens à mentionner également les élèves avec lesquels j'ai été en contact : grâce à eux, enseigner à l'ENS a toujours été un grand plaisir. Je n'oublie pas non plus les nombreuses personnes auxquelles j'ai posé des questions mathématiques, et qui ont ainsi contribué à cette thèse. Parmi elles, je citerai mon co-auteur (pour le chapitre 5), Tanguy Rivoal.

Mes remerciements vont aussi à l'association Math en Jeans[1] ; cette thèse prouve (si besoin est) qu'elle suscite des vocations ! La personnalité de Gilles Godefroy, le premier chercheur que j'aie rencontré, n'y est pas étrangère.

Je tiens à mentionner ici Cédric Villani : c'est notamment suite à de longues discussions avec lui que j'ai choisi d'étudier l'approximation diophantienne (et je ne le regrette pas !).

Finalement, je remercie tous les amis qui m'ont soutenu durant ces années : doctorants ou ex-doctorants (de l'ENS, de Chevaleret ou d'ailleurs), futurs doctorants et (surtout ?) pas doctorants du tout.

Enfin, du fond du cœur, je remercie ma famille, mes parents et ma sœur Nathalie.

[1] http ://www.mjc-andre.org/pages/amej/accueil.htm

Résumé

Cette thèse est consacrée à l'approximation diophantienne.

Dans une première partie, on s'intéresse aux problèmes d'irrationalité liés aux valeurs de polylogarithmes. Après une synthèse des différentes approches connues concernant les valeurs aux entiers (positifs) de la fonction zêta de Riemann, on en vient à des résultats nouveaux. D'une part, on exhibe des changements de variables entre intégrales multiples, qui permettent de généraliser les travaux de Rhin-Viola sur $\zeta(2)$ et $\zeta(3)$ et de relier l'approche de Beukers, Vasilenko, Vasilyev à celle de Sorokin pour la construction de formes linéaires en valeurs de ζ. D'autre part, dans un travail en commun avec Rivoal, on écrit comme solution unique d'un problème d'approximation de Padé une série hypergéométrique très générale. On retrouve comme cas particuliers essentiellement toutes les séries qui apparaissent dans les preuves connues dans ce contexte. Cette méthode permet, notamment, de démontrer que parmi les nombres $\mathrm{Li}_2(1/2) + \log(2)^2$, $\mathrm{Li}_3(1/2) - \frac{1}{2}\log(2)^3$ et $\mathrm{Li}_4(1/2) + \frac{1}{6}\log(2)^4$, l'un au moins est irrationnel (avec $\mathrm{Li}_s(1/2) = \sum_{n=1}^{\infty} 2^{-n} n^{-s}$).

La deuxième partie de cette thèse se situe dans le cadre des groupes algébriques commutatifs. Tout d'abord, on s'intéresse aux puissances du groupe multiplicatif : il s'agit d'étudier les relations polynomiales entre logarithmes de nombres algébriques. La conjecture d'indépendance algébrique de ces logarithmes a été traduite géométriquement par Roy. On suit cette piste, en démontrant cette conjecture pour certaines variétés. Ensuite on s'intéresse, dans un groupe algébrique commutatif G quelconque, à un ingrédient particulier des démonstrations de transcendance : le lemme d'interpolation. On généralise un tel énoncé, dû à Masser, en y incluant des multiplicités. Dans le cas où le groupe G est linéaire, la dualité de Fourier-Borel relie le lemme d'interpolation à un lemme de zéros ; on exprime ces énoncés en termes d'algèbres de Hopf, ce qui permet de mieux percevoir leur symétrie.

Introduction

Cette thèse se situe dans le cadre de l'approximation diophantienne. L'objectif est d'obtenir des résultats d'irrationalité, d'indépendance linéaire, de transcendance voire d'indépendance algébrique (ou bien leurs versions quantitatives : mesures d'irrationalité, ...). Dans la première partie, on s'intéresse à des constructions explicites, dans le cas des polylogarithmes. La seconde partie est consacrée aux groupes algébriques commutatifs ; on utilise alors le principe des tiroirs, et un "lemme de zéros" (ou un "lemme d'interpolation") est nécessaire.

1 Indépendance linéaire et valeurs de polylogarithmes

La première partie de cette thèse est consacrée à des problèmes d'irrationalité, ou d'indépendance linéaire sur $\mathbb{Q}$, de valeurs de polylogarithmes.

Tout d'abord, on présente une synthèse des résultats, et des différentes démonstrations, connus au sujet de l'irrationalité des valeurs de la fonction ζ de Riemann. Ceci permet de mettre en lumière les méthodes variées utilisées pour construire des formes linéaires en valeurs de polylogarithmes. Deux d'entre elles sont étudiées plus particulièrement dans la suite : les intégrales multiples et les problèmes de Padé.

D'une part, on exhibe des changements de variables qui relient différentes intégrales multiples, et qui généralisent ceux utilisés par Beukers [B4] et Rhin-Viola ([R3], [R4]) pour $\zeta(2)$ et $\zeta(3)$. D'autre part, on écrit (dans un travail en commun avec Tanguy Rivoal) une série hypergéométrique très générale comme solution unique d'un problème d'approximation de Padé. On obtient ainsi deux nouveaux résultats diophantiens.

1.1 Irrationalité de valeurs de zêta

La preuve de l'irrationalité de $\zeta(3)$ par Apéry [A10], en 1978, a suscité un grand élan d'intérêt pour les problèmes d'irrationalité de valeurs de ζ. En 2000, Rivoal [R5] et Ball-Rivoal [B2] ont publié une preuve de l'irrationalité (et même de l'indépendance linéaire sur $\mathbb{Q}$) d'une infinité de $\zeta(2n+1)$. L'irrationalité de $\zeta(5)$ est toujours conjecturale, mais Zudilin a démontré [Z6] que parmi $\zeta(5)$, $\zeta(7)$, $\zeta(9)$ et $\zeta(11)$, l'un au moins est irrationnel.

Ces trois résultats sont démontrés dans le chapitre 1 de la thèse ; il s'agit du texte d'un exposé au Séminaire Bourbaki de Novembre 2002. Dans une première partie, on effectue une synthèse des nombreuses preuves connues de l'irrationalité de $\zeta(3)$. A ma connaissance, aucune synthèse de ce genre n'existait dans la littérature. Le point de vue adopté consiste à construire, d'une dizaine de manières, les nombres d'Apéry $\frac{v_n}{u_n}$, et à montrer comment passer d'une construction à l'autre.

La deuxième partie du chapitre 1 donne une preuve des théorèmes de [R5] et [B2]. Il s'agit d'une variante de la preuve originelle. En effet, on considère la série

$$\mathbf{S}_n = 2n!^{a-2r} \sum_{k \geq 1} (k + \frac{n}{2}) \frac{(k-1)(k-2)\dots(k-rn)(k+n+1)(k+n+2)\dots(k+(r+1)n)}{k^a(k+1)^a\dots(k+n)^a},$$

où a et r sont des paramètres entiers (avec $1 \leq r < \frac{a}{2}$) et n est un indice qui tend vers l'infini. C'est le facteur $k + \frac{n}{2}$ (absent dans [R5] et [B2]) qui constitue la différence. Son intérêt est que pour $a = 4$ et $r = 1$ on retrouve les formes linéaires d'Apéry en 1 et $\zeta(3)$. On peut donc espérer que certaines propriétés des nombres d'Apéry se généralisent à ces formes linéaires (par exemple en liaison avec l'aspect modulaire des nombres d'Apéry, mis en évidence par Beukers [B9]).

La troisième partie du chapitre 1 contient notamment une esquisse de la preuve [Z6] que parmi $\zeta(5)$, $\zeta(7)$, $\zeta(9)$ et $\zeta(11)$, l'un au moins est irrationnel. La nouveauté réside dans la traduction : on utilise les notations de [R7], qui sont différentes de celles de [Z6]. La traduction ne présente pas de difficulté théorique, mais compte tenu de la complexité technique elle n'est pas immédiate. Ceci résume bien l'orientation de ce chapitre : donner une présentation, aussi unifiée et cohérente que possible, des résultats connus concernant l'irrationalité de valeurs de ζ.

1.2 Changements de variables dans des intégrales multiples

Une version quantitative du théorème d'Apéry est une majoration de l'exposant d'irrationalité de $\zeta(3)$. La meilleure connue à ce jour est $5,513891$; elle est due à Rhin-Viola [R4]. Le point crucial de leur démonstration est l'action d'un groupe sur une famille d'intégrales triples. Au chapitre 4 (qui est à paraître au Journal de Théorie des Nombres de Bordeaux), on écrit ces intégrales triples comme des intégrales sur une variété semi-algébrique réelle $\Omega_3 \times U_3$, et les transformations utilisées par Rhin-Viola découlent d'automorphismes de la variété $\Omega_3 \times U_3$ (conformément à l'esprit de la "philosophie des périodes"[K3], et en réponse à une question de Pierre Cartier). Le même travail est effectué pour $\zeta(2)$; le chapitre 2, qui est paru aux Publications Mathématiques de l'UFR Sciences et Techniques de Besançon, annonce ce résultat et contient des rappels sur la notion d'exposant d'irrationalité.

La motivation initiale de ce travail était de mieux comprendre l'action du groupe de Rhin-Viola, en l'exprimant géométriquement. Dans le cas de $\zeta(2)$, on obtient une variété dont les équations sont relativement simples et symétriques. Ceci permet d'espérer construire, par analogie, d'autres variétés auxquelles la méthode de Rhin-Viola pourrait s'appliquer. Toutefois, à l'heure actuelle, aucune variété n'a été construite ainsi ; et le cas de $\zeta(3)$, qui mène à des équations nettement plus compliquées, n'incite pas à l'optimisme.

Dans la deuxième moitié du chapitre 4, on construit des généralisations du groupe de Rhin-Viola à des intégrales n-uples (ces auteurs ayant considéré uniquement le cas $n \leq 3$). Les résultats obtenus sont résumés dans une note parue aux Comptes-Rendus de l'Académie des Sciences, qui forme le chapitre 3.

Ces généralisations n'ont pas été obtenues grâce à l'interprétation géométrique mentionnée ci-dessus, mais en exhibant des changements de variables qui préservent la forme d'une intégrale multiple. Précisément, on définit trois familles d'intégrales multiples, et on montre que chacune d'elles est munie d'un groupe de Rhin-Viola ; l'une de ces trois familles donne, pour $n = 2$ et $n = 3$, les groupes introduits par Rhin et Viola. Si on savait démontrer que ces intégrales sont des formes linéaires en valeurs de ζ, avec un dénominateur convenable, on pourrait utiliser la méthode de Rhin-Viola pour améliorer ce dénominateur : ce serait une application possible des résultats de la seconde moitié du chapitre 4.

Mais la recherche de tels changements de variables peut avoir d'autres buts. Par exemple, le changement de variables de $[0,1]^n$ dans $[0,1]^n$ défini par :

$$\begin{aligned} x_p &= y_{n+1-p} \text{ si } p \equiv n \mod 2, \\ \text{et } x_p &= \frac{(1-y_1y_2\dots y_{n-p})y_{n+1-p}}{1-y_1y_2\dots y_{n+1-p}} \text{ si } p \not\equiv n \mod 2 \end{aligned}$$

permet de démontrer le théorème suivant :

Théorème 1.1 *L'intégrale*

$$\int_{[0,1]^n} \frac{\prod_{k=1}^n x_k^N(1-x_k)^N}{(1-x_n(1-x_{n-1}(1-\dots(1-x_1))))^{N+1}} \mathrm{d}x_1\dots\mathrm{d}x_n$$

est égale, pour tout $N \geq 0$, à

$$\int_{[0,1]^n} \frac{\prod_{k=1}^n y_k^N(1-y_k)^N}{\prod_{k\in\{2,\dots,n\} \text{ pair}} (1-y_1\dots y_k)^{N+1}} \mathrm{d}y_1\dots\mathrm{d}y_n$$

si n est pair, et à

$$\int_{[0,1]^n} \frac{\prod_{k=1}^n y_k^N(1-y_k)^N}{(1-y_1\dots y_n)^{N+1}\prod_{k\in\{2,\dots,n\} \text{ pair}} (1-y_1\dots y_k)^{N+1}} \mathrm{d}y_1\dots\mathrm{d}y_n$$

si n est impair.

L'intérêt de ce théorème est de relier (en réponse à une question de Tanguy Rivoal) les approches de Beukers [B4], Vasilenko [V1] et Vasilyev ([V2], [V3]) d'une part, de Sorokin [S7] d'autre part, pour construire des formes linéaires en valeurs de ζ. En outre le même changement de variables permet de démontrer un théorème beaucoup plus général, qui englobe les intégrales considérées par Rhin-Viola ([R3], [R4]), Rivoal ([B2], [R5]) et Sorokin [S6].

1.3 Approximants de Padé et séries hypergéométriques équilibrées

Le chapitre 5 de la thèse est un travail en commun avec Tanguy Rivoal, qui a fait l'objet d'une prépublication de l'Université de Caen. Il est à paraître au Journal de Mathématiques Pures et Appliquées (journal de Liouville). Considérons une série de la forme

$$S(z) = \sum_{k=1}^{\infty} \left(k+\frac{n}{2}\right)^{\varepsilon} \frac{(k-\rho)_\rho(k+n+1)_\sigma}{(k)_{n+1}^a} z^{-k},$$

où $\varepsilon \in \{0,1\}$ et $n, a, \rho, \sigma \geq 0$ vérifient $\rho+\sigma+\varepsilon \leq a(n+1)-1$; si $\varepsilon = 1$, on suppose en outre $\rho = \sigma$. Cette famille de séries englobe celles utilisées dans [N4], [B2], [R5], [R6], … Le résultat principal du chapitre 5 est la construction (en réponse à une question posée dans [R6]) d'un problème de Padé dont une telle série est solution unique. Une généralisation de ce résultat est également donnée, qui permet d'inclure les séries de [N2] et [R7].

L'intérêt d'un tel énoncé est de contribuer à inscrire dans un cadre conceptuel (ici celui des approximants de Padé) les séries hypergéométriques utilisées dans les preuves d'indépendance linéaire. En outre, il établit un lien avec les constructions de Reyssat [R2] et permet d'obtenir de nouveaux résultats d'indépendance linéaire ou d'irrationalité :

Théorème 1.2 *Soient μ un nombre réel, et α un nombre rationnel tel que $0 < \alpha < 1$. L'ensemble des $\mathrm{Li}_s(\alpha) + \mu \frac{\log^s(\alpha)}{(s-1)!}$, où s décrit $\mathbb{N}^*$, contient une infinité de nombres linéairement indépendants sur $\mathbb{Q}$.*

Théorème 1.3 *Au moins l'un des trois nombres $\mathrm{Li}_2(1/2) + \log^2(2)$, $\mathrm{Li}_3(1/2) - \frac{1}{2}\log^3(2)$ et $\mathrm{Li}_4(1/2) + \frac{1}{6}\log^4(2)$ est irrationnel.*

2 Transcendance dans les groupes algébriques

L'un des contextes dans lesquels on sait appliquer les méthodes de transcendance est celui des groupes algébriques. La deuxième partie de cette thèse est consacrée à quelques résultats dans ce domaine.

Tout d'abord, on considère le cas où le groupe algébrique G est une puissance du groupe multiplicatif $\mathbb{G}_m$; il s'agit d'étudier les logarithmes de nombres algébriques. En suivant une piste suggérée par Damien Roy [R13], on tente de décrire les points d'une variété algébrique dont les coordonnées sont des logarithmes de nombres algébriques.

Ensuite, on considère un groupe algébrique commutatif G quelconque, et on se focalise sur l'un des ingrédients essentiels d'une preuve de transcendance : le lemme de zéros. Plus précisément, on étudie les lemmes d'interpolation, qui peuvent remplacer les lemmes de zéros. Un tel énoncé, optimal à constante multiplicative près, a été démontré par Masser [M2]. On généralise ici ce résultat, en y incluant des multiplicités.

Enfin, dans le cas où G est affine, on formule en termes d'algèbres de Hopf les lemmes de zéros, d'interpolation, et la dualité de Fourier-Borel.

2.1 Orbites de groupes algébriques et logarithmes de nombres algébriques

Le chapitre 6 de cette thèse est un article paru (en anglais) dans Acta Arithmetica.

Notons $L = \{\lambda \in \mathbb{C}, \exp(\lambda) \in \bar{\mathbb{Q}}^*\}$ le $\mathbb{Q}$-espace vectoriel des logarithmes de nombres algébriques, et $\mathcal{L}$ le sous-$\bar{\mathbb{Q}}$-espace vectoriel de $\mathbb{C}$ engendré par 1 et L. Conjecturalement, des éléments $\lambda_1, \ldots, \lambda_l$ de $\mathcal{L}$ sont algébriquement indépendants (sur $\bar{\mathbb{Q}}$) si, et seulement si, $1, \lambda_1, \ldots, \lambda_l$ sont linéairement indépendants sur $\bar{\mathbb{Q}}$. Damien Roy a traduit géométriquement cette conjecture :

Conjecture 2.1 ([R13]) *Soit X un fermé algébrique de $\mathbb{C}^n$ défini sur $\bar{\mathbb{Q}}$. Alors*

$$X \cap \mathcal{L}^n = \bigcup_{E \subset X} E \cap \mathcal{L}^n$$

où E décrit l'ensemble des sous-espaces affines de $\mathbb{C}^n$ contenus dans X et définis sur $\bar{\mathbb{Q}}$.

Cette conjecture signifie que les points de X à coordonnées dans $\mathcal{L}$ sont exceptionnels, au sens suivant : leur présence s'explique toujours par celle d'un sous-espace affine E défini sur $\bar{\mathbb{Q}}$ et contenu dans X. Elle est triviale quand la variété X est linéaire ; au chapitre 6, on la démontre quand $X = \mathcal{V}(k,m)$ est le cône affine au-dessus de la variété de Veronese (i.e. l'ensemble des $v \cdot \ldots \cdot v \in \mathrm{Sym}^k(\mathbb{C}^m)$, quand v décrit $\mathbb{C}^m$), sous l'hypothèse $k \geq 3$. Il n'y a essentiellement aucune autre variété X pour laquelle la conjecture 2.1 soit connue. Dans le cas de la variété $\mathcal{V}(k,m)$, la conjecture 2.1 s'énonce comme suit :

Théorème 2.2 *Soient k et m des entiers, avec $k \geq 3$. Soit P un polynôme homogène, de degré k, en m variables, dont les coefficients appartiennent à $\mathcal{L}$. Supposons que P est la* *puissance k-ième d'une forme linéaire (à coefficients complexes). Alors il existe une forme linéaire Φ, à coefficients dans $\bar{\mathbb{Q}}$, et un élément $a \in \mathcal{L}$ tels que $P = a\Phi^k$.*

On considère, en outre, la situation générale suivante. Soit G un groupe algébrique affine défini sur $\bar{\mathbb{Q}}$, qui agit sur un $\mathbb{C}$-espace vectoriel W muni d'une $\bar{\mathbb{Q}}$-structure. On suppose que l'action est donnée par un morphisme de groupes algébriques $\rho : G \to \mathrm{GL}(W)$ défini sur $\bar{\mathbb{Q}}$. On note X une orbite sous cette action ; alors X est une partie localement fermée de W. Sous certaines hypothèses, on démontre que tout point de $X \cap \mathcal{L}^n$ appartient à un hyperplan vectoriel de W défini sur $\bar{\mathbb{Q}}$. On parvient parfois à obtenir des informations plus précises sur les points de $X \cap \mathcal{L}^n$, comme dans l'énoncé suivant :

Théorème 2.3 *Supposons que ρ a une orbite Zariski-dense Y, et qu'on a $2\dim(X) < \dim(W)$. Alors pour tout $x \in X \cap \mathcal{L}^n$ il existe un sous-espace vectoriel $\mathcal{E}$ de W, défini sur $\bar{\mathbb{Q}}$, qui contient x et est disjoint de Y.*

Les théorèmes généraux obtenus dans ce cadre sont déduits d'un théorème de transcendance dû à Damien Roy [R11], qui généralise le théorème des six exponentielles. Ce théorème permet de minorer le rang d'une matrice à coefficients dans $\mathcal{L}$ en fonction des relations linéaires, à coefficients dans $\bar{\mathbb{Q}}$, entre ses lignes ou entre ses colonnes.

2.2 Lemme d'interpolation dans un groupe algébrique commutatif

Soit G un groupe algébrique commutatif connexe sur $\mathbb{C}$, plongé de manière quasiprojective dans $\mathbb{P}^N$. Soient W un sous-espace vectoriel de l'algèbre de Lie TG de G, et Γ un sous-groupe de type fini de $G(\mathbb{C})$. Soient $(\partial_1, \ldots, \partial_d)$ une base de W, et $(\gamma_1, \ldots, \gamma_l)$ une famille génératrice de Γ. Pour $\tau = (\tau_1, \ldots, \tau_d) \in \mathbb{N}^d$ on considère l'opérateur différentiel $\partial^\tau = \partial_1^{\tau_1} \ldots \partial_d^{\tau_d}$, qui commute aux translations. Pour $S \in \mathbb{R}_+$ on note $\Gamma(S)$ l'ensemble des combinaisons linéaires $n_1\gamma_1 + \ldots + n_l\gamma_l$, avec n_j entier et $0 \le n_j < S$ pour tout $j \in \{1, \ldots, l\}$. On suppose que Γ est contenu dans l'ouvert $\{X_0 \neq 0\}$ de $\mathbb{P}^N$.

Considérons l'application linéaire Φ qui à tout polynôme P sur G, homogène de degré D, associe la famille des $\partial^\tau \left(\frac{P}{X_0^D}\right)(\gamma)$ pour $\gamma \in \Gamma(S)$ et $\tau \in \mathbb{N}^d$, $|\tau| = \tau_1 + \ldots + \tau_d < T$. Un *lemme de zéros* est un énoncé qui fournit une condition suffisante (portant sur D, T, S, et en général aussi sur les sous-groupes algébriques de G) pour que Φ soit injective. Ces lemmes de zéros sont un point crucial dans les démonstrations de transcendance dans ce contexte.

La question de la surjectivité de Φ a été moins étudiée. On appelle *lemme d'interpolation* tout énoncé qui fournit une condition suffisante pour que Φ soit surjective. Dans le cas où il n'y a pas de multiplicités, Masser a démontré un lemme d'interpolation, optimal à constante multiplicative près. Le chapitre 7 est consacré à la preuve du résultat suivant, qui est analogue à celui de Masser mais autorise des multiplicités :

Théorème 2.4 *Il existe une constante $c > 0$ ayant la propriété suivante. Soient D, S, T trois entiers strictement positifs tels que, pour tout sous-groupe algébrique H connexe, non nul, de G on ait*

$$S^{\mathrm{rg}(\Gamma\cap H)} T^{\dim(W\cap TH)} < c D^{\dim(H)} \tag{1}$$

Pour tous $\gamma \in \Gamma(S)$ et $\tau \in \mathbb{N}^d$, $|\tau| < T$, soit $a_{\gamma,\tau}$ un nombre complexe. Alors il existe un polynôme P, homogène de degré D, tel que

$$\partial^\tau \left(\frac{P}{X_0^D}\right)(\gamma) = a_{\gamma,\tau}$$

pour tous $\gamma \in \Gamma(S)$ *et* $\tau \in \mathbb{N}^d$, $|\tau| < T$.

La preuve de ce théorème suit celle de Masser. Deux outils nouveaux sont nécessaires : une étude des fonctionnelles sur G et une analyse des sous-groupes obstructeurs.

Le fait que Φ ne soit pas surjective équivaut à l'existence d'une fonctionnelle, combinaison linéaire des $P \mapsto \partial^\tau \left(\frac{P}{X_0^D}\right)(\gamma)$, qui s'annule identiquement sur l'espace des polynômes homogènes de degré D. On définit, sur l'espace des fonctionnelles, des opérations de translation et de dérivation. Cette construction n'est pas évidente a priori pour deux raisons. D'une part, ces fonctionnelles n'opèrent pas sur de vraies fonctions sur G, mais sur des polynômes homogènes. D'autre part, le degré de $\partial^\tau \left(\frac{P}{X_0^D}\right)$ risque de croître avec $|\tau|$; pour éviter ceci, on utilise l'astuce dite de Baker-Coates-Anderson-Chudnovski.

En l'absence de multiplicités, la répartition de Γ dans G est mesurée, classiquement, par un exposant $\mu(\Gamma, G) = \min_{H \neq G} \frac{\mathrm{rg}(\Gamma/\Gamma\cap H)}{\dim(G/H)}$, où le minimum porte sur l'ensemble des sous-groupes algébriques H de G, distincts de G. Cet exposant apparaît dans le lemme de zéros ; Masser définit un exposant analogue qui est utilisé dans son lemme d'interpolation : $\mu^*(\Gamma, G) = \max_{H \neq 0} \frac{\mathrm{rg}(\Gamma\cap H)}{\dim(H)}$. Notons $\bar{H}$ un sous-groupe algébrique non nul de G qui réalise ce maximum ; alors dans le théorème 2.4 (avec $T = 1$), il suffit de vérifier la condition (1) pour $H = \bar{H}$.

Quand on autorise des multiplicités, il ne suffit plus en général de vérifier la condition (1) pour un seul sous-groupe (qui dépendrait seulement de Γ, W et G). Il est nécessaire d'introduire une dépendance en S, T, D. Pour cela, on construit (en fonction de Γ, W et G uniquement) des sous-groupes $H_1, \ldots, H_{p-1}$ et des secteurs angulaires $\mathcal{C}_1, \ldots, \mathcal{C}_{p-1}$ de $\mathbb{R}_+^2$, avec $p \geq 2$, tels que :

(i) Dans le théorème 2.4, il suffit de vérifier la condition (1) quand H est l'un des H_i.

(ii) Étant donnés S, T, D, avec $D > 1$, il existe un indice i tel que $(\frac{\log(S)}{\log(D)}, \frac{\log(T)}{\log(D)})$ appartienne à $\mathcal{C}_i$; alors dans le théorème 2.4 il suffit de vérifier la condition (1) pour $H = H_i$.

(iii) L'ensemble $\{H_1, \ldots, H_{p-1}\}$ est minimal au sens suivant : si on omet l'un des H_i alors la propriété (i) est en défaut.

On définit aussi des sous-groupes $\widetilde{H}_j$ qui jouent un rôle analogue pour le lemme de zéros, et on étudie le lien entre les H_i et les $\widetilde{H}_j$.

Les outils introduits pour démontrer le théorème 2.4 sont susceptibles d'autres applications. Quant à l'énoncé lui-même de ce théorème, il peut permettre d'obtenir de nouvelles preuves de transcendance. En effet, supposons que G et W sont définis sur $\bar{\mathbb{Q}}$, et qu'on a $\Gamma \subset G(\bar{\mathbb{Q}})$. On considère la matrice formée par les $\partial^\tau \left(\frac{P}{X_0^D}\right)(\gamma)$, où P décrit une base de l'espace $\mathbb{C}[G]_D^h$ des polynômes homogènes sur G de degré D, et (γ, τ) décrit $\Gamma(S) \times \{\tau \in \mathbb{N}^d, |\tau| < T\}$. D'habitude, on construit une combinaison linéaire non triviale des lignes, c'est-à-dire un polynôme P non nul, de degré D, qui s'annule en tout point de $\Gamma(S)$ à l'ordre T le long de W. Ce polynôme, appelé *fonction auxiliaire*, est obtenu grâce à un lemme de Siegel, un lemme de Schwarz et une inégalité de Liouville. On conclut la démonstration en appliquant un lemme de zéros.

Une autre méthode consiste, dans la même situation, à faire des combinaisons linéaires sur les colonnes. Ceci revient à construire une fonctionnelle non nulle, combinaison linéaire des évaluations aux points de $\Gamma(S)$ de dérivées le long de W jusqu'à l'ordre T, qui s'annule sur tous les polynômes homogènes de degré D. Cette fonctionnelle, appelée *fonctionnelle auxiliaire*, est construite grâce à un lemme de Siegel, un lemme de Schwarz et une inégalité de Liouville.

On conclut la démonstration en appliquant un lemme d'interpolation (celui de Masser si il n'y a pas de multiplicités, le théorème 2.4 sinon).

Grâce au théorème 2.4, on peut ainsi donner de nouvelles preuves de résultats connus (par exemple [W4] le théorème de Schneider sur la transcendance d'intégrales elliptiques de première espèce). Peut-être cette méthode permet-elle, dans certains cas, d'améliorer des mesures de transcendance ? En outre, les fonctionnelles auxiliaires ont été utilisées (sans multiplicités) dans [W3], pour faire un pas en direction d'une généralisation de la conjecture de Leopoldt sur le rang p-adique du groupe des unités d'un corps de nombres (voir cependant [R12]).

2.3 Lemmes de zéros et algèbres de Hopf

Supposons maintenant que le groupe algébrique commutatif G est affine, mais pas nécessairement connexe ; alors la composante neutre de G est isomorphe au produit d'une puissance de $\mathbb{G}_a$ par une puissance de $\mathbb{G}_m$. La dualité de Fourier-Borel (voir [W5] et [W8]) relie cette situation (i.e. Γ, W et G) à une autre de même nature, avec $\widetilde{\Gamma}$, $\widetilde{W}$ et $\widetilde{G}$. Elle est fondée sur la formule suivante (qui se généralise à plusieurs variables) :

$$\left(\frac{\mathrm{d}}{\mathrm{d}z}\right)^t (z^s e^{xz})_{|z=y} = \left(\frac{\mathrm{d}}{\mathrm{d}z}\right)^s \left(z^t e^{yz}\right)_{|z=x} \tag{2}$$

Cette dualité échange points et facteurs $\mathbb{G}_m$, dérivations et facteurs $\mathbb{G}_a$, fonctions et fonctionnelles. Elle transforme un lemme de zéros en un lemme d'interpolation (dans la situation duale). En particulier, elle permet de déduire le théorème 2.4 ci-dessus (dans le cas où G est linéaire) du lemme de zéros appliqué à la situation duale.

La formule (2) a plusieurs inconvénients ; notamment, elle nécessite de choisir des coordonnées, et n'est pas susceptible de généralisations simples à d'autres groupes G. C'est pourquoi, au chapitre 8, on formule la dualité de Fourier-Borel, ainsi que les lemmes de zéros et d'interpolation, dans un langage plus adapté aux groupes algébriques : celui des algèbres de Hopf. Précisément, on définit deux catégories $\mathcal{C}$ et $\overline{\mathcal{C}}$ comme suit.

Les objets de $\mathcal{C}$ sont les triplets (G, Γ, W) où $G = G_0 \times G_1$ est un groupe algébrique affine commutatif (avec G_0 isomorphe à une puissance de $\mathbb{G}_a$ et G_1 plongé dans une puissance de $\mathbb{G}_m$), Γ est un sous-groupe de $G(\mathbb{C})$ muni d'une famille génératrice (finie) $\gamma_1, \ldots, \gamma_l$, et W est un sous-espace vectoriel de TG, tels qu'aucun sous-groupe algébrique G' de G, distinct de G, ne vérifie $\Gamma \subset G'$ et $W \subset TG'$.

Les objets de $\overline{\mathcal{C}}$ sont les triplets $(\widetilde{H}, H, < \cdot, \cdot >)$ où $\widetilde{H} = \mathbb{C}[\widetilde{G}_0 \times \widetilde{G}_1]$ et $H = \mathbb{C}[G_0 \times G_1]$ sont des algèbres de Hopf de type fini, commutatives et co-commutatives (avec G_0 et $\widetilde{G}_0$ isomorphes à des puissances de $\mathbb{G}_a$, et G_1 et $\widetilde{G}_1$ plongés dans des puissances de $\mathbb{G}_m$), et $< \cdot, \cdot >$ est un produit bilinéaire non dégénéré, de $\widetilde{H} \times H$ dans $\mathbb{C}$, qui permet d'identifier $\widetilde{H}$ à une sous-algèbre de Hopf de H° (où H° est l'algèbre de Hopf duale de H).

Alors les catégories $\mathcal{C}$ et $\overline{\mathcal{C}}$ sont équivalentes. Pour le démontrer, étant donné un objet (G, Γ, W) de $\mathcal{C}$, on note $H = \mathbb{C}[G]$ l'algèbre de Hopf des fonctions polynomiales sur G, et $\widetilde{H} = \mathbb{C}\Gamma \otimes \mathrm{Sym}(W)$, où $\mathbb{C}\Gamma$ est l'algèbre du groupe Γ. On définit un produit bilinéaire non dégénéré de $\widetilde{H} \times H$ dans $\mathbb{C}$ en posant, pour $R \in H$, $\gamma \in \Gamma$ et $\partial^{(1)}, \ldots, \partial^{(k)} \in W$:

$$< \gamma \otimes \left(\partial^{(1)} \cdot \ldots \cdot \partial^{(k)}\right), R >= \left(\partial^{(1)} \ldots \partial^{(k)} R\right)(\gamma) \tag{3}$$

On montre alors que $(\widetilde{H}, H, < \cdot, \cdot >)$ est un objet de $\overline{\mathcal{C}}$, et qu'on a défini une équivalence de catégories.

Dans la catégorie $\overline{\mathcal{C}}$, la transformation de Fourier-Borel prend une forme particulièrement simple : elle échange H et $\widetilde{H}$. On peut aussi traduire dans la catégorie $\overline{\mathcal{C}}$ le lemme de zéros. Pour cela, on note $H_{S,T}$ l'espace des polynômes sur $G = G_0 \times G_1$ de bi-degré au plus (S, T), et on définit $\widetilde{H}_{S,T}$ de même. On obtient alors l'énoncé suivant :

Théorème 2.5 *Pour tout objet* $(\widetilde{H} = \mathbb{C}[\widetilde{G}_0 \times \widetilde{G}_1], H = \mathbb{C}[G_0 \times G_1], < \cdot, \cdot >)$ *de* $\overline{\mathcal{C}}$ *il existe une constante* $c > 0$ *ayant la propriété suivante. Soient* $S, T, \widetilde{S}, \widetilde{T}$ *des entiers strictement positifs, et* $R \in H_{S,T}$ *non nul tel que*

$$< x, R >= 0 \textit{ pour tout } x \in \widetilde{H}_{\widetilde{S},\widetilde{T}}$$

Alors il existe une sous-algèbre de Hopf $H' = \mathbb{C}[G_0/G'_0 \times G_1/G'_1]$ *de* H*, et une algèbre de Hopf quotient* $\widetilde{H}' = \mathbb{C}[\widetilde{G}'_0 \times \widetilde{G}'_1]$ *de* $\widetilde{H}$*, telles que :*

$$T^{\dim(G_0/G'_0)} S^{\dim(G_1/G'_1)} > c\, \widetilde{T}^{\dim(\widetilde{G}'_0)} \widetilde{S}^{\dim(\widetilde{G}'_1)}$$

En outre, l'obstruction est un objet $(\widetilde{H}', H', < \cdot, \cdot >')$ de $\overline{\mathcal{C}}$. Ce théorème traduit un lemme de zéros dans le groupe algébrique G, et un lemme d'interpolation dans le groupe algébrique $\widetilde{G}$.

Le théorème 2.5 peut donner des idées pour l'étude de situations plus générales que celles considérées habituellement : que se passe-t-il si on ne suppose plus H et $\widetilde{H}$ commutatives, co-commutatives et de type fini ? D'autre part, si G est un groupe algébrique commutatif quasiprojectif (non affine), plongé dans $\mathbb{P}^N$ (comme au paragraphe 2.2), dans quelle mesure peut-on généraliser ces méthodes ?

Première partie

Indépendance linéaire et valeurs de polylogarithmes

Séminaire Bourbaki, Novembre 2002
Exposé numéro 910

Chapitre 1

Irrationalité de valeurs de zêta (d'après Apéry, Rivoal, ...)

INTRODUCTION

Cet exposé est consacré aux valeurs aux entiers $s \geq 2$ de la fonction zêta de Riemann, définie par $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$. Quand $s = 2k$ est pair, on sait que $\zeta(2k)\pi^{-2k}$ est un nombre rationnel, lié aux nombres de Bernoulli. Comme π est transcendant (voir l'appendice de [L2] pour une preuve), $\zeta(2k)$ l'est aussi pour tout $k \geq 1$. La nature arithmétique des $\zeta(2k+1)$ est beaucoup moins bien connue. D'un point de vue conjectural, la situation est simple :

Conjecture 0.1 *Les nombres π, $\zeta(3)$, $\zeta(5)$, $\zeta(7)$, …sont algébriquement indépendants sur $\mathbb{Q}$.*

Cette conjecture est un cas particulier d'une conjecture diophantienne sur les polyzêtas (voir [W7] ou [C3]). Elle implique que les $\zeta(2k+1)$ sont tous transcendants, donc irrationnels, et linéairement indépendants sur $\mathbb{Q}$.

Très peu de résultats sont connus en direction de la conjecture 0.1. Le premier d'entre eux a été annoncé par Apéry lors des Journées Arithmétiques de Luminy, en 1978 :

Théorème 0.2 ([A10]) *$\zeta(3)$ est irrationnel.*

Apéry lui-même n'a donné lors de son exposé (voir [M4]), et n'a publié [A10], qu'une esquisse de sa preuve. Les détails (qui sont loin d'être triviaux) ont été publiés par Van Der Poorten [P4] (voir aussi [C6] et [R1]), grâce à des contributions de Cohen et Zagier. Par la suite, plusieurs autres démonstrations du théorème d'Apéry sont parues. La première partie de ce texte est consacrée à une synthèse des différents points de vue qu'on peut adopter pour le démontrer.

La grande percée suivante date de 2000 :

Théorème 0.3 ([R5], [B2]) *Le $\mathbb{Q}$-espace vectoriel engendré par 1, $\zeta(3)$, $\zeta(5)$, $\zeta(7)$, …est de dimension infinie.*

En conséquence, il existe une infinité de k tels que $\zeta(2k+1)$ soit irrationnel. On peut donner des versions effectives de ce dernier énoncé : Rivoal a démontré [R7] que parmi les neuf nombres $\zeta(5)$, $\zeta(7)$, …, $\zeta(21)$, l'un au moins est irrationnel. Ce résultat a été amélioré par Zudilin :

Théorème 0.4 ([Z6], [Z9]) *L'un au moins des quatre nombres $\zeta(5)$, $\zeta(7)$, $\zeta(9)$, $\zeta(11)$ est irrationnel.*

Malgré ces développements récents, il n'existe aucun entier $s \geq 5$ impair pour lequel on sache si $\zeta(s)$ est rationnel ou non.

Ce texte est divisé en trois parties. La première est une synthèse des méthodes connues pour démontrer l'irrationalité de $\zeta(3)$; l'intérêt des différentes approches est qu'elles se généralisent plus ou moins facilement à d'autres situations. La deuxième partie fournit une preuve du théorème 0.3, et de résultats voisins. La troisième est consacrée à des résultats "quantitatifs" : mesure d'irrationalité de $\zeta(3)$ et théorème 0.4.

Remerciements : Je remercie toutes les personnes qui m'ont aidé dans la préparation de ce texte, notamment F. Amoroso, V. Bosser, N. Brisebarre, P. Cartier, G. Christol, P. Colmez, P. Grinspan, L. Habsieger, M. Huttner, C. Krattenthaler, C. Maclean, F. Martin, Yu. Nesterenko, F. Pellarin, A. Pulita, E. Royer, M. Waldschmidt, D. Zagier et W. Zudilin. Je remercie tout particulièrement T. Rivoal pour les nombreuses discussions très instructives que nous avons eues.

1 Irrationalité de $\zeta(3)$

Toutes les preuves connues de l'irrationalité de $\zeta(3)$ ont la même structure. On construit, pour tout $n \geq 0$, des nombres rationnels u_n et v_n ayant les propriétés suivantes :

1. La forme linéaire $I_n = u_n\zeta(3) - v_n$ vérifie
$$\limsup_{n\to\infty} |I_n|^{1/n} \leq (\sqrt{2}-1)^4 = 0,0294372\ldots$$
2. En notant d_n le p.p.c.m. des entiers compris entre 1 et n, les coefficients u_n et v_n vérifient :
$$u_n \in \mathbb{Z} \text{ et } 2d_n^3 v_n \in \mathbb{Z}.$$
3. Pour une infinité d'entiers n, on a $I_n \neq 0$.

La conclusion est alors immédiate : si $\zeta(3)$ était un nombre rationnel p/q, alors $2qd_n^3 I_n$ serait un entier pour tout n, et tendrait vers zéro quand n tend vers l'infini (car $(\sqrt{2}-1)^4 e^3 < 1$, en utilisant [I2] le théorème des nombres premiers sous la forme $\lim_{n\to\infty} \frac{\log(d_n)}{n} = 1$) : cela contredit la troisième assertion.

Remarque 1.1 *Comme $(\sqrt{2}-1)^4 \cdot 3,23^3 < 1$, le théorème des nombres premiers peut être remplacé par l'assertion plus faible $d_n < 3,23^n$ pour n assez grand, qui se démontre en utilisant des arguments élémentaires à la Tchebychev ([N6], §8.1 ; [I2], p. 15).*

Dans la suite, on donne plusieurs constructions (§ 1.1 à 1.10) de u_n, v_n et I_n, à chaque fois notées $u_{i,n}$, $v_{i,n}$ et $I_{i,n}$ (l'indice $i \in \{\mathrm{R}, \mathrm{E}, \mathbb{R}, \Sigma, \mathbb{C}, \mathrm{P}, \mathrm{TB}, \mathrm{M}\}$ fait référence à la construction utilisée). En fait, on construit toujours les mêmes formes linéaires : *a posteriori* on s'aperçoit que $u_{i,n}$, $v_{i,n}$ et $I_{i,n}$ ne dépendent pas de i. La preuve de cette indépendance est le plus souvent directe. Parfois, on montre simplement que $I_{i,n} = I_{j,n}$; les deux autres égalités en découlent en utilisant l'irrationalité de $\zeta(3)$.

Les premières valeurs de u_n et v_n sont :

$$\begin{aligned} (u_n)_{n\geq 0} &= 1, 5, 73, 1445, 33001, 819005, \ldots \\ (v_n)_{n\geq 0} &= 0, 6, \frac{351}{4}, \frac{62531}{36}, \frac{11424695}{288}, \ldots \end{aligned}$$

Cette partie contient l'esquisse de plusieurs preuves de l'irrationalité de $\zeta(3)$, notamment celles d'Apéry [A10] (§ 1.1 et 1.2), de Beukers [B4] par les intégrales multiples (§ 1.3) ou [B9] par les formes modulaires (§ 1.10), de Prevost [P7] (§ 1.1 et 1.2), de Nesterenko [N2] (§ 1.4 et 1.5), de Sorokin [S7] (§ 1.8), et de nombreuses variantes. Certaines preuves sont obtenues en montrant que deux constructions différentes fournissent les mêmes formes linéaires, puis en prouvant le point (2) à l'aide de l'une et les points (1) et (3) à l'aide de l'autre (par exemple en montrant que $\lim_{n\to\infty} |I_n|^{1/n} = (\sqrt{2}-1)^4$).

La plupart des méthodes connues pour démontrer des résultats d'irrationalité sur les valeurs de ζ sont liées aux polylogarithmes, définis pour tout entier $k \geq 1$ par :

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

avec $|z| < 1$ si $k = 1$ et $|z| \leq 1$ si $k \geq 2$. L'idée est de construire des formes linéaires en polylogarithmes, à coefficients polynomiaux, puis de spécialiser en $z = 1$. C'est la méthode employée dans les paragraphes 1.3 à 1.9. Les formes linéaires en polylogarithmes $I_{i,n}(z)$ qu'on utilise ne sont pas toujours les mêmes, mais elles coïncident en $z = 1$, pour donner les formes linéaires d'Apéry.

Les polylogarithmes s'insèrent dans la famille des séries hypergéométriques ${}_{q+1}F_q$ (avec $q \geq 1$), définies par :

$${}_{q+1}F_q\left(\begin{array}{cccc} \alpha_0, & \alpha_1, & \ldots, & \alpha_q \\ & \beta_1, & \ldots, & \beta_q \end{array} \middle| \, z\right) = \sum_{k=0}^{\infty} \frac{(\alpha_0)_k(\alpha_1)_k \cdots (\alpha_q)_k}{k!\,(\beta_1)_k \cdots (\beta_q)_k} z^k\ ,$$

où le symbole de Pochhammer est $(\alpha)_k = \alpha(\alpha+1)\cdots(\alpha+k-1)$. Dans cet exposé, les α_j et les β_j seront des entiers, les β_j étant positifs, et z sera un nombre complexe avec $|z| \leq 1$. On adopte les définitions suivantes ([A9], §3.3 et 3.4) :

- ${}_{q+1}F_q$ est dite *bien équilibrée* si $\alpha_0 + 1 = \alpha_1 + \beta_1 = \cdots = \alpha_q + \beta_q$;
- ${}_{q+1}F_q$ est dite *très bien équilibrée* si elle est bien équilibrée et $\alpha_1 = \frac{1}{2}\alpha_0 + 1$.

1.1 Récurrence linéaire

Définition 1.2 *Soient* $(u_{\mathrm{R},n})_{n\geq 0}$ *et* $(v_{\mathrm{R},n})_{n\geq 0}$ *les suites définies par la relation de récurrence*

$$(n+1)^3 y_{n+1} - (34n^3 + 51n^2 + 27n + 5)y_n + n^3 y_{n-1} = 0 \tag{1.1}$$

et les conditions initiales

$$u_{\mathrm{R},0} = 1\ ,\ u_{\mathrm{R},1} = 5\ ,\ v_{\mathrm{R},0} = 0\ ,\ v_{\mathrm{R},1} = 6.$$

Une récurrence immédiate montre que les suites $(u_{\mathrm{R},n})$ et $(v_{\mathrm{R},n})$ sont croissantes et à termes rationnels, avec $n!^3 u_{\mathrm{R},n} \in \mathbb{Z}$ et $n!^3 v_{\mathrm{R},n} \in \mathbb{Z}$. En fait on verra qu'on peut remplacer $n!^3$ par d_n^3.

Les propriétés asymptotiques des suites vérifiant la récurrence (1.1) sont faciles à déterminer (voir par exemple [G1], Chapitre 5). L'équation caractéristique associée est $X^2 - 34X + 1$; elle a deux racines simples, $(\sqrt{2}+1)^4$ et $(\sqrt{2}-1)^4$. L'espace vectoriel des solutions de (1.1) est de dimension deux, et admet une base formée de suites $(y_n^{(0)})_{n\geq 0}$ et $(y_n^{(1)})_{n\geq 0}$ avec $\lim_{n\to+\infty} \frac{\log|y_n^{(0)}|}{n} = \log((\sqrt{2}+1)^4)$ et $\lim_{n\to+\infty} \frac{\log|y_n^{(1)}|}{n} = \log((\sqrt{2}-1)^4)$. La suite $(y_n^{(1)})$

est uniquement déterminée (à proportionnalité près) par son comportement asymptotique ; toutes les autres solutions de (1.1) se comportent comme $(y_n^{(0)})$. Comme $(u_{\mathrm{R},n})$ et $(v_{\mathrm{R},n})$ sont croissantes, on a :

$$\lim_{n\to\infty} u_{\mathrm{R},n}^{1/n} = \lim_{n\to\infty} v_{\mathrm{R},n}^{1/n} = (\sqrt{2}+1)^4 = 33,9705627\ldots \tag{1.2}$$

Quand on adopte ce point de vue, on a intérêt [P4] à considérer $\Delta_n = \begin{vmatrix} v_{\mathrm{R},n} & v_{\mathrm{R},n-1} \\ u_{\mathrm{R},n} & u_{\mathrm{R},n-1} \end{vmatrix}$ pour $n \geq 1$. La relation de récurrence montre qu'on a $\Delta_n = \frac{6}{n^3}$ pour tout n, ce qui signifie $\frac{v_{\mathrm{R},n}}{u_{\mathrm{R},n}} - \frac{v_{\mathrm{R},n-1}}{u_{\mathrm{R},n-1}} = \frac{6}{n^3 u_{\mathrm{R},n} u_{\mathrm{R},n-1}}$. Donc la suite $(\frac{v_{\mathrm{R},n}}{u_{\mathrm{R},n}})$ est strictement croissante et tend vers une limite finie ℓ, avec $u_{\mathrm{R},n}\ell - v_{\mathrm{R},n} = \sum_{k=n+1}^{\infty} \frac{6u_{\mathrm{R},n}}{k^3 u_{\mathrm{R},k} u_{\mathrm{R},k-1}}$. Ceci prouve que $u_{\mathrm{R},n}\ell - v_{\mathrm{R},n}$ est une solution de (1.1) qui tend vers zéro quand n tend vers l'infini : son comportement asymptotique est nécessairement donné par

$$\lim_{n\to+\infty} \frac{\log|u_{\mathrm{R},n}\ell - v_{\mathrm{R},n}|}{n} = \log((\sqrt{2}-1)^4).$$

Avec cette définition de $u_{\mathrm{R},n}$ et $v_{\mathrm{R},n}$, il n'est pas évident de démontrer que $\ell = \zeta(3)$, et de borner par d_n^3 les dénominateurs de $u_{\mathrm{R},n}$ et $v_{\mathrm{R},n}$. Pour ceci, une possibilité est de faire le lien avec le paragraphe 1.2 : c'est la méthode employée dans les premières preuves détaillées de l'irrationalité de $\zeta(3)$, qui sont parues peu après l'exposé d'Apéry ([R1], [P4], [C6]).

Remarque 1.3 *Le raisonnement ci-dessus montre que* $\frac{v_{\mathrm{R},n}}{u_{\mathrm{R},n}}$ *est la* n*-ième somme partielle de la série* $\zeta(3) = \sum_{k=1}^{\infty} \frac{6}{k^3 u_{\mathrm{R},k} u_{\mathrm{R},k-1}}$.

La définition 1.2 s'interprète en termes de fractions continues généralisées. En effet, considérons la récurrence linéaire

$$Y_{n+1} - (34n^3 + 51n^2 + 27n + 5)Y_n + n^6 Y_{n-1} = 0. \tag{1.3}$$

On passe d'une solution de (1.1) à une solution de (1.3), et réciproquement, en posant $Y_n = n!^3 y_n$. Si $U_{\mathrm{R},n}$ et $V_{\mathrm{R},n}$ sont ainsi associées à $u_{\mathrm{R},n}$ et $v_{\mathrm{R},n}$, alors $\frac{V_{\mathrm{R},n}}{U_{\mathrm{R},n}} = \frac{v_{\mathrm{R},n}}{u_{\mathrm{R},n}}$ est la n-ième réduite de la fraction continue généralisée

$$\zeta(3) = \frac{6|}{|5} - \frac{1|}{|117} - \frac{64|}{|535} - \cdots - \frac{n^6|}{|34n^3 + 51n^2 + 27n + 5} - \cdots .$$

On peut trouver cette formule grâce à un procédé général ([A11], [B3], [Z4]) qui accélère la convergence d'un développement en fraction continue généralisée. Ce procédé s'applique, en particulier, au développement dont les réduites sont les sommes partielles de la série $\sum_{n=1}^{\infty} \frac{1}{f(n)}$, où f est un polynôme sans zéro parmi les entiers strictement positifs.

En utilisant cette méthode d'accélération de convergence, André-Jeannin a démontré [A7] que la somme des inverses des nombres de Fibonacci est irrationnelle (voir aussi [B17] et [P8]).

1.2 Formules explicites

Définition 1.4 *Soient* $(u_{\mathrm{E},n})$ *et* $(v_{\mathrm{E},n})$ *les suites définies par les formules suivantes :*

$$\begin{aligned} u_{\mathrm{E},n} &= \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 \\ v_{\mathrm{E},n} &= \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 \left(\sum_{m=1}^{n} \frac{1}{m^3} + \sum_{m=1}^{k} \frac{(-1)^{m-1}}{2m^3 \binom{n}{m}\binom{n+m}{m}} \right) \end{aligned}$$

Sous cette forme, il est clair que $u_{\mathrm{E},n} \in \mathbb{Z}$ et que $\frac{v_{\mathrm{E},n}}{u_{\mathrm{E},n}}$ tend vers $\zeta(3)$. Pour démontrer ([P4], [C6], [R1]) que $2d_n^3 v_{\mathrm{E},n} \in \mathbb{Z}$, il suffit de démontrer que, pour $1 \leq m \leq k \leq n$,

$$\frac{\binom{n+k}{k} d_n^3}{m^3 \binom{n}{m}\binom{n+m}{m}} = \frac{\binom{n+k}{k-m} d_n^3}{m^3 \binom{n}{m}\binom{k}{m}} \tag{1.4}$$

est entier. Soit p un nombre premier ; la valuation p-adique $\mathrm{v}_p(n!)$ de $n!$ vaut $\sum_{i=1}^{\alpha} [\frac{n}{p^i}]$ avec $\alpha = [\frac{\log(n)}{\log(p)}] = \mathrm{v}_p(d_n)$. Pour $1 \leq i \leq \mathrm{v}_p(m)$ on a $[\frac{n}{p^i}] = [\frac{n-m}{p^i}] + [\frac{m}{p^i}]$ et pour $\mathrm{v}_p(m) < i \leq \mathrm{v}_p(d_n)$ on a $[\frac{n}{p^i}] \leq [\frac{n-m}{p^i}] + [\frac{m}{p^i}] + 1$. On en déduit $\mathrm{v}_p(\binom{n}{m}) \leq \mathrm{v}_p(d_n) - \mathrm{v}_p(m)$ et $\mathrm{v}_p(\binom{k}{m}) \leq \mathrm{v}_p(d_k) - \mathrm{v}_p(m)$. Il en résulte que $\frac{d_n^3}{m^3\binom{n}{m}\binom{k}{m}}$ est un entier, et le quotient (1.4) aussi.

Montrons maintenant ([P4], [C6]) que les suites $(u_{\mathrm{E},n})$ et $(v_{\mathrm{E},n})$ vérifient la récurrence (1.1). On pose $\lambda_{n,k} = \binom{n}{k}^2 \binom{n+k}{k}^2$ pour $k, n \in \mathbb{Z}$, et

$$\mathbf{A}_{n,k} = 4(2n+1)(k(2k+1) - (2n+1)^2)\lambda_{n,k},$$

avec les conventions habituelles (i.e. $\lambda_{n,k} = 0$ si $k < 0$ ou $k > n$). On a alors

$$\mathbf{A}_{n,k} - \mathbf{A}_{n,k-1} = (n+1)^3 \lambda_{n+1,k} - (34n^3 + 51n^2 + 27n + 5)\lambda_{n,k} + n^3 \lambda_{n-1,k}.$$

En sommant sur k, on obtient que la suite $(u_{\mathrm{E},n})$ satisfait à la récurrence (1.1). Pour la suite $(v_{\mathrm{E},n})$, on peut faire de même en utilisant la suite double

$$\mathbf{B}_{n,k} = \mathbf{A}_{n,k} \left(\sum_{m=1}^{n} \frac{1}{m^3} + \sum_{m=1}^{k} \frac{(-1)^{m-1}}{2m^3 \binom{n}{m}\binom{n+m}{m}} \right) + \frac{5(2n+1)k(-1)^{k-1}}{n(n+1)} \binom{n}{k}\binom{n+k}{k}.$$

Ceci démontre qu'on a $u_{\mathrm{E},n} = u_{\mathrm{R},n}$ et $v_{\mathrm{E},n} = v_{\mathrm{R},n}$ pour tout $n \geq 0$. Compte tenu des résultats démontrés au paragraphe 1.1, on obtient une preuve de l'irrationalité de $\zeta(3)$.

La démonstration donnée ci-dessus que $(u_{\mathrm{E},n})$ et $(v_{\mathrm{E},n})$ vérifient la récurrence (1.1) n'est qu'une simple vérification, à condition d'être capable d'exhiber les suites doubles $\mathbf{A}_{n,k}$ et $\mathbf{B}_{n,k}$, ce qui n'a pas été une tâche facile (voir [P4], §7). Motivés par ce problème, plusieurs auteurs (notamment Zeilberger) ont ensuite mis au point des algorithmes permettant d'exhiber de telles suites doubles. On a ainsi un moyen automatique de produire des preuves d'identités (voir [C2], [Z3], [P2]). De plus, ces preuves sont immédiatement vérifiables à la main.

Dans les formules ci-dessus, un rôle central est joué par la suite double $c_{n,k} = \sum_{m=1}^{n} \frac{1}{m^3} + \sum_{m=1}^{k} \frac{(-1)^{m-1}}{2m^3\binom{n}{m}\binom{n+m}{m}}$ (définie pour $0 \leq k \leq n$). Elle tend vers $\zeta(3)$ quand n tend vers l'infini, uniformément en k. On a $c_{n,n} - c_{n-1,n-1} = \frac{5}{2}\frac{(-1)^{n-1}}{n^3\binom{2n}{n}}$ et $\lim_{n\to\infty} c_{n,n} = \zeta(3)$ donc :

$$\zeta(3) = \frac{5}{2}\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^3\binom{2n}{n}}. \tag{1.5}$$

Cette série n'est pas utilisée dans la preuve de l'irrationalité de $\zeta(3)$, mais elle a un intérêt non négligeable puisque les $c_{n,k}$ sont au cœur des formules explicites définissant $u_{\mathrm{E},n}$ et $v_{\mathrm{E},n}$. C'est pourquoi plusieurs auteurs ont cherché des généralisations de (1.5) (voir par exemple [P4], [P6], [C7], [K2], [L3], [B12], [A4]), parmi lesquelles $\zeta(5) = \frac{5}{2}\sum_{n\geq 1} \frac{(-1)^n}{n^3\binom{2n}{n}}\left(\sum_{j=1}^{n-1}\frac{1}{j^2} - \frac{4}{5n^2}\right)$. Mais aucune de ces généralisations n'a permis d'obtenir de nouveau résultat d'irrationalité : la croissance des dénominateurs est trop rapide par rapport à la convergence.

Prévost a montré [P7] comment interpréter les formules explicites données dans ce paragraphe en termes d'approximants de Padé. Posons $\varphi(x) = \sum_{k\geq 1}\frac{1}{(k+x)^3}$, c'est-à-dire $\zeta(3, 1+x)$ où ζ est la fonction zêta d'Hurwitz (voir [W10], Chapitre XIII). Pour tout $n \geq 1$, considérons les polynômes suivants :

$$\begin{aligned} P_n(x) &= \sum_{k=0}^{n}\binom{n}{k}\binom{n+k}{k}\binom{x}{k}\binom{x+k}{k} = {}_4F_3\left(\begin{matrix} -n, & -x, & n+1, & x+1 \\ & 1, & 1, & 1\end{matrix}\,\middle|\, 1\right) \\ \text{et}\quad Q_n(x) &= \sum_{k=0}^{n}\binom{n}{k}\binom{n+k}{k}\binom{x}{k}\binom{x+k}{k}\sum_{m=1}^{k}\frac{(-1)^{m-1}}{2m^3\binom{x}{m}\binom{x+m}{m}}. \end{aligned}$$

Alors P_n est de degré $2n$, Q_n de degré $2n-2$, et on a $P_n(x)\varphi(x) - Q_n(x) = \mathrm{O}(x^{-2n-1})$ quand x tend vers l'infini. Cela signifie que P_n et Q_n sont des approximants de Padé de la fonction φ. Quand x est un entier n, on a $\varphi(n) = \zeta(3) - \sum_{m=1}^{n}\frac{1}{m^3}$ d'où $P_n(n)\varphi(n) - Q_n(n) = u_{\mathrm{E},n}\zeta(3) - v_{\mathrm{E},n}$. On peut en déduire [P7] la majoration $|u_{\mathrm{E},n}\zeta(3) - v_{\mathrm{E},n}| \leq \frac{4\pi^2}{(2n+1)^2 u_{\mathrm{E},n}}$. Pour conclure, on a besoin d'une minoration asymptotique de $u_{\mathrm{E},n}$ comme celle de la formule (1.2). Il suffit donc de vérifier que $u_{\mathrm{E},n}$ satisfait à la récurrence (1.1). On peut utiliser $\mathbf{A}_{n,k}$ comme ci-dessous ; une autre méthode [A12] est d'utiliser des relations de contiguïté entre séries hypergéométriques balancées.

En effet, $u_{\mathrm{E},n}$ s'écrit ${}_4F_3\left(\begin{matrix} -n, & -n, & n+1, & n+1 \\ & 1, & 1, & 1\end{matrix}\,\middle|\, 1\right)$. Une série hypergéométrique ${}_4F_3\left(\begin{matrix} \alpha_0, & \alpha_1, & \alpha_2, & \alpha_3 \\ & \beta_1, & \beta_2, & \beta_3\end{matrix}\,\middle|\, z\right)$ est dite ([S4], §2.1.1) *balancée* (ou *Saalschützienne*) si $1 + \sum_{i=0}^{3}\alpha_i = \sum_{j=1}^{3}\beta_j$. Si on modifie deux des sept paramètres d'une série balancée, en ajoutant ou en retranchant 1 à chacun des deux, on peut obtenir à nouveau une série balancée. Si c'est le cas, on dit que ces deux séries sont *contiguës*. Il y a $2\cdot\binom{7}{2} = 42$ séries balancées qui sont contiguës à une série balancée donnée. Quand α_0 est un entier négatif (ce qui signifie que la série hypergéométrique est en fait un polynôme), il existe des relations linéaires entre les valeurs en 1 de ces 42 séries, dont les coefficients sont des polynômes en les paramètres α_0, $\ldots$, β_3 (voir [A9], §3.7). On peut [A12] déduire de ces relations de contiguïté que la suite $u_{\mathrm{E},n}$ vérifie la récurrence (1.1).

1.3 Intégrale triple réelle

Considérons l'intégrale suivante, qui a été introduite par Beukers [B4] (voir aussi [B6]) :

$$I_{\mathbb{R},n}(z) = \int_0^1 \int_0^1 \int_0^1 \frac{u^n(1-u)^n v^n(1-v)^n w^n(1-w)^n}{((1-w)z+uvw)^{n+1}} \mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w.$$

Cette intégrale converge pour tout $z \in \mathbb{C}\backslash]-\infty,0]$. Voici une esquisse de preuve de l'irrationalité de $\zeta(3)$ qui utilise $I_{\mathbb{R},n}(1)$. Les détails se trouvent dans [B4].

Comme le maximum de la fonction $\frac{u(1-u)v(1-v)w(1-w)}{1-w(1-uv)}$ sur le cube unité vaut $(\sqrt{2}-1)^4$, on a :

$$\lim_{n\to+\infty} \frac{\log(I_{\mathbb{R},n}(1))}{n} = \log((\sqrt{2}-1)^4).$$

Par ailleurs, si on intègre n fois par parties par rapport à v, qu'on change w en $\frac{1-w}{1-w(1-uv)}$, et enfin qu'on intègre n fois par parties par rapport à u, on obtient :

$$I_{\mathbb{R},n}(1) = \int_0^1 \int_0^1 \int_0^1 \frac{P_n(u)P_n(v)}{1-w(1-uv)} \mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w,$$

où $P_n(X) = \frac{1}{n!}(X^n(1-X)^n)^{(n)}$ est le n-ième polynôme de Legendre. En intégrant par rapport à w, il vient $I_{\mathbb{R},n}(1) = \int_0^1 \int_0^1 \frac{-\log(uv)}{1-uv} P_n(u)P_n(v)\mathrm{d}u\,\mathrm{d}v$. Or pour tous $k,l \in \{0,\dots,n\}$ on peut écrire $\int_0^1 \int_0^1 \frac{-\log(uv)}{1-uv} u^k v^l \mathrm{d}u\,\mathrm{d}v = 2a_{k,l}\zeta(3) + b_{k,l}$ avec $a_{k,l} \in \mathbb{Z}$ et $d_n^3 b_{k,l} \in \mathbb{Z}$. On a donc :

$$I_{\mathbb{R},n}(1) = 2(u_{\mathbb{R},n}\zeta(3) - v_{\mathbb{R},n}) \text{ avec } u_{\mathbb{R},n} \in \mathbb{Z} \text{ et } 2d_n^3 v_{\mathbb{R},n} \in \mathbb{Z}.$$

Cela termine la preuve de l'irrationalité de $\zeta(3)$.

1.4 Série de type hypergéométrique

Posons

$$R_n(X) = \frac{(X-1)^2 \dots (X-n)^2}{X^2(X+1)^2 \dots (X+n)^2} = \frac{(X-n)_n^2}{(X)_{n+1}^2} = \frac{\Gamma(X)^4}{\Gamma(X-n)^2\Gamma(X+n+1)^2}, \tag{1.6}$$

où Γ est la fonction Gamma d'Euler, qui vérifie $\Gamma(s+1) = s\Gamma(s)$. En outre, pour $|z| \geq 1$ on pose :

$$I_{\Sigma,n}(z) = -\sum_{k=1}^{\infty} R_n'(k) z^{-k}. \tag{1.7}$$

En suivant [B5], [G4] et [N2] on développe la fraction rationnelle R_n en éléments simples :

$$R_n(X) = \sum_{i=0}^{n} \left(\frac{\alpha_i}{(X+i)^2} + \frac{\beta_i}{X+i} \right), \tag{1.8}$$

avec $\alpha_i = \binom{n}{i}^2 \binom{n+i}{i}^2$ et $\beta_i = 2(-1)^i \binom{n}{i}\binom{n+i}{i} \sum_{j\in\{0,\dots,n\}, j\neq i} \frac{(-1)^j \binom{n}{j}\binom{n+j}{j}}{j-i}$ pour $i \in \{0,\dots,n\}$ (ces formules s'obtiennent en remarquant que $R_n(X) = (\frac{(X-n)_n}{(X)_{n+1}})^2$; voir la démonstration du lemme 2.12 ci-dessous, ou bien [C8], [H1] ou [Z10]). En utilisant (1.8) pour exprimer (1.7) il vient :

$$\begin{aligned} I_{\Sigma,n}(z) &= 2\sum_{i=0}^{n} \alpha_i z^i \sum_{k\geq 1} \frac{z^{-(k+i)}}{(k+i)^3} + \sum_{i=0}^{n} \beta_i z^i \sum_{k\geq 1} \frac{z^{-(k+i)}}{(k+i)^2} \\ &= 2A_n(z)\mathrm{Li}_3(1/z) + B_n(z)\mathrm{Li}_2(1/z) + C_n(z) \end{aligned} \tag{1.9}$$

où les polynômes A_n, B_n et C_n sont définis par :

$$\begin{aligned} A_n(z) &= \sum_{i=0}^{n} \alpha_i z^i = {}_4F_3\left(\begin{matrix} -n, & -n, & n+1, & n+1 \\ & 1, & 1, & 1 \end{matrix} \,\middle|\, z \right) \\ B_n(z) &= \sum_{i=0}^{n} \beta_i z^i \\ C_n(z) &= -\sum_{t=0}^{n-1} z^t \sum_{i=t+1}^{n} \left(\frac{2\alpha_i}{(i-t)^3} + \frac{\beta_i}{(i-t)^2} \right) \end{aligned}$$

Il est clair que les polynômes $A_n(z)$, $d_nB_n(z)$ et $d_n^3C_n(z)$ sont à coefficients entiers. On a $B_n(1) = 0$ car R_n n'a pas de résidu à l'infini. En posant $u_{\Sigma,n} = A_n(1)$ et $v_{\Sigma,n} = -C_n(1)/2$ il vient :

$$I_{\Sigma,n}(1) = 2(u_{\Sigma,n}\zeta(3) - v_{\Sigma,n}) \text{ avec } u_{\Sigma,n} \in \mathbb{Z} \text{ et } 2d_n^3 v_{\Sigma,n} \in \mathbb{Z}. \tag{1.10}$$

Pour démontrer l'irrationalité de $\zeta(3)$, il ne reste plus qu'à estimer $I_{\Sigma,n}(1)$. On peut le faire en transformant $I_{\Sigma,n}(1)$ en une intégrale complexe (voir le paragraphe 1.5) ; c'est ainsi que Nesterenko démontre [N2] le théorème d'Apéry.

On peut démontrer, en utilisant [Z10] l'algorithme de "creative telescoping" ([P2], Chapitre 6), que $I_{\Sigma,n}(1)$, $u_{\Sigma,n}$ et $v_{\Sigma,n}$ satisfont à la relation de récurrence (1.1). Cela démontre en particulier l'identité $v_{\Sigma,n} = v_{\mathrm{E},n}$.

1.5 Intégrale complexe

Soit c un réel, avec $0 < c < n+1$. Pour $z \neq 0$, choisissons une détermination de $\arg(z)$ strictement comprise entre -2π et 2π, et considérons l'intégrale suivante, le long de la droite verticale $\mathrm{Re}(s) = c$ dans $\mathbb{C}$, orientée de bas en haut :

$$\begin{aligned} I_{\mathbb{C},n}(z) &= \frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} \left(\frac{\pi}{\sin(\pi s)} \right)^2 R_n(s) z^{-s} \mathrm{d}s \\ &= \frac{1}{2i\pi} \int_{c-i\infty}^{c+i\infty} \frac{\Gamma(n+1-s)^2\Gamma(s)^4}{\Gamma(n+1+s)^2} z^{-s} \mathrm{d}s, \end{aligned} \tag{1.11}$$

cette dernière égalité provenant directement de (1.6) et de la formule classique $\frac{\pi}{\sin(\pi s)} = \Gamma(s)\Gamma(1-s)$. La valeur de $I_{\mathbb{C},n}(z)$ ne dépend pas du choix de c d'après le théorème des résidus. L'intégrale (1.11) est un exemple de G-fonction de Meijer (voir [L6], §5.2) :

$$I_{\mathbb{C},n}(z) = G_{4,4}^{4,2}\left(\begin{matrix} -n, & -n, & n+1, & n+1 \\ 0, & 0, & 0, & 0 \end{matrix} \,\middle|\, z \right).$$

La méthode du col (voir par exemple [D2], Chapitre IX) permet [N2] d'obtenir une estimation asymptotique très précise :

$$I_{\mathbb{C},n}(1) = \frac{\pi^{3/2}2^{3/4}}{n^{3/2}}(\sqrt{2}-1)^{4n+2}(1+\mathrm{O}(n^{-1})).$$

Quand on déplace le contour d'intégration vers la droite pour faire apparaître les pôles $n+1$, $n+2$, …, le théorème des résidus donne ([G3], [G4]), puisque $(\frac{\pi}{\sin(\pi s)})^2 = \frac{1}{(s-k)^2} + \mathrm{O}(1)$ quand s tend vers un entier k :

$$I_{\mathbb{C},n}(z) = I_{\Sigma,n}(z) + \log(z)\sum_{k=1}^{\infty} R_n(k)z^{-k}. \tag{1.12}$$

En particulier pour $z=1$ on obtient $I_{\mathbb{C},n}(1) = I_{\Sigma,n}(1)$.

Par ailleurs, Nesterenko a démontré [N3] un théorème général qui relie une intégrale multiple réelle à une intégrale complexe ; dans notre cas particulier, ce théorème donne $I_{\mathbb{C},n}(z) = I_{\mathbb{R},n}(z)$.

On peut démontrer [N2] que $I_{\mathbb{C},n}(1)$ vérifie la récurrence (1.1) en utilisant les relations de contiguïté sur les G-fonctions de Meijer. C'est en fait une preuve parallèle à celle du paragraphe 1.2, où on utilisait la contiguïté entre des ${}_4F_3$. En effet ([L6], §5.8), ces ${}_4F_3$ satisfont aux mêmes équations différentielles que les G-fonctions de Meijer correspondantes, donc aux mêmes relations de contiguïté.

1.6 Un problème d'approximation de Padé

Considérons [B5] le problème suivant : trouver quatre polynômes A_n, B_n, C_n et D_n, à coefficients rationnels, de degré au plus n, tels que :

$$\begin{cases} F_n(z) := A_n(z)\mathrm{Li}_2(1/z) + B_n(z)\mathrm{Li}_1(1/z) + D_n(z) = \mathrm{O}(z^{-n-1}) \text{ quand } z \to \infty \\ G_n(z) := 2A_n(z)\mathrm{Li}_3(1/z) + B_n(z)\mathrm{Li}_2(1/z) + C_n(z) = \mathrm{O}(z^{-n-1}) \text{ quand } z \to \infty \\ B_n(1) = 0 \end{cases} \tag{1.13}$$

Une solution à ce problème de Padé est donnée par les polynômes A_n, B_n et C_n du paragraphe 1.4 (et un polynôme D_n convenable). On a alors :

$$\begin{cases} F_n(z) = \sum_{k=1}^{\infty} R_n(k)z^{-k} = \frac{n!^4}{(2n+1)!^2} z^{-n-1}{}_4F_3\left(\left.\begin{matrix} n+1, & n+1, & n+1, & n+1 \\ & 2n+2, & 2n+2, & 1 \end{matrix}\right| z^{-1}\right) \\ G_n(z) = I_{\Sigma,n}(z) = -\sum_{k=1}^{\infty} R_n'(k)z^{-k} \end{cases}$$

En effet, la seconde égalité est simplement une réécriture de (1.7) et (1.9). La première se démontre de manière analogue à (1.9), mais sans dériver (1.8).

L'équation différentielle hypergéométrique sous-jacente aux constructions des paragraphes 1.4 et 1.5 s'écrit $Ly=0$, en posant

$$L = z(\delta+n+1)^2(\delta-n)^2 - \delta^4 \text{ avec } \delta = z\frac{\mathrm{d}}{\mathrm{d}z}.$$

Elle admet au voisinage de l'infini quatre solutions linéairement indépendantes : $F_n(z)$, $I_{\mathbb{C},n}(z) = G_n(z) + F_n(z)\log(z)$, $A_n(z)$ et $B_n(z) - A_n(z)\log(z)$ (voir [L6], §5.1 et 5.8, [H11] et [G4]). Ces solutions sont reliées par la monodromie : en prolongeant analytiquement F_n le long d'un lacet qui entoure le point 1 on fait apparaître $B_n(z)+A_n(z)\log(1/z)$, puis en faisant le tour de l'infini on obtient $A_n(z)$ (voir [O1] pour la monodromie des polylogarithmes).

Ce point de vue permet de démontrer [H11] que le problème de Padé (1.13) a une solution unique (à proportionnalité près). En effet, en partant d'une solution A_n, B_n, C_n, D_n, on montre que F_n vérifie une équation différentielle linéaire fuchsienne d'ordre 4 qu'on détermine explicitement (en calculant ses exposants, et en utilisant la relation de Fuchs) : on trouve que c'est $Ly = 0$.

Pour démontrer l'unicité de la solution de ce problème de Padé, on peut aussi suivre [B5]. On part d'une solution quelconque, avec des polynômes A_n, B_n, C_n, D_n et des fonctions F_n et G_n. On note α_i et β_i les coefficients de A_n et B_n, et on leur associe la fraction rationnelle R_n définie par (1.8). On voit alors que $F_n(z) = \sum_{k=1}^{\infty} R_n(k)z^{-k}$ et $G_n(z) = -\sum_{k=1}^{\infty} R_n'(k)z^{-k}$, donc les deux premières contraintes de (1.13) signifient que R_n et sa dérivée s'annulent aux points 1, 2, ..., n. En outre, le résidu à l'infini de R_n est alors $B_n(1) = 0$: la fraction rationnelle R_n est nécessairement donnée, à constante multiplicative près, par (1.6).

1.7 Polynômes orthogonaux

Considérons ([B13], [A13]) le problème suivant : trouver deux polynômes $\widetilde{A}_n$ et $\widetilde{B}_n$, de degré au plus n, tels que :

$$\begin{cases} \int_0^1 \left(\widetilde{B}_n(x) - \widetilde{A}_n(x)\log(x)\right) x^k \mathrm{d}x = 0 \text{ pour tout } k \in \{0, \dots, n-1\} \\ \int_0^1 \left(\widetilde{B}_n(x) - \widetilde{A}_n(x)\log(x)\right) x^k \log(x) \mathrm{d}x = 0 \text{ pour tout } k \in \{0, \dots, n-1\} \\ \widetilde{B}_n(1) = 0 \end{cases} \tag{1.14}$$

Une solution à ce problème est donnée par les polynômes $\widetilde{A}_n$ et $\widetilde{B}_n$ définis par :

$$\widetilde{B}_n(x) - \widetilde{A}_n(x)\log(x) = \int_x^1 P_n(\frac{x}{t})P_n(t)\frac{\mathrm{d}t}{t}, \tag{1.15}$$

où P_n est le n-ième polynôme de Legendre (comme au paragraphe 1.3). En effet, on a alors $\int_0^1 \left(\widetilde{B}_n(x) - \widetilde{A}_n(x)\log(x)\right) x^k \mathrm{d}x = \left(\int_0^1 P_n(u)u^k \mathrm{d}u\right)^2$ en posant $u = \frac{x}{t}$. La première condition de (1.14) en découle immédiatement ; la deuxième s'obtient après dérivation par rapport à k.

Comme on a $\mathrm{Li}_j(1/z) = \frac{(-1)^{j-1}}{(j-1)!}\int_0^1 \log^{j-1}(x)\frac{\mathrm{d}x}{z-x}$ pour tout entier $j \geq 1$, il vient :

$$2\widetilde{A}_n(z)\mathrm{Li}_3(1/z) + \widetilde{B}_n(z)\mathrm{Li}_2(1/z) = -\int_0^1 \left(\widetilde{B}_n(z) - \widetilde{A}_n(z)\log(x)\right) \frac{\log(x)\,\mathrm{d}x}{z-x}. \tag{1.16}$$

On définit un polynôme $\widetilde{C}_n(z)$ par :

$$\widetilde{C}_n(z) = \int_0^1 \frac{\widetilde{B}_n(z) - \widetilde{B}_n(x)}{z-x}\log(x)\,\mathrm{d}x - \int_0^1 \frac{\widetilde{A}_n(z) - \widetilde{A}_n(x)}{z-x}\log^2(x)\,\mathrm{d}x.$$

Grâce à (1.15) on peut obtenir des formules explicites pour $\widetilde{A}_n$, $\widetilde{B}_n$ et $\widetilde{C}_n$; on trouve les mêmes que pour A_n, B_n et C_n respectivement au paragraphe 1.4. Donc $\widetilde{A}_n$, $d_n\widetilde{B}_n$ et $d_n^3\widetilde{C}_n$ sont à coefficients entiers. On obtient aussi ([B13], Corollaire A.2.3) que tous les zéros de $\widetilde{A}_n(z)$ et de $\frac{\widetilde{B}_n(z)}{z-1}$ sont réels négatifs, et entrelacés. Par ailleurs, on a :

$$2\widetilde{A}_n(z)\mathrm{Li}_3(1/z) + \widetilde{B}_n(z)\mathrm{Li}_2(1/z) + \widetilde{C}_n(z) = -\int_0^1 \left(\widetilde{B}_n(x) - \widetilde{A}_n(x)\log(x)\right) \frac{\log(x)\mathrm{d}x}{z-x}.$$

Quand $z = 1$, le membre de droite se transforme (en utilisant (1.15) et en posant $u = t$, $v = \frac{x}{t}$) en $I_{\mathbb{R},n}(1) = -\int_0^1 \int_0^1 \frac{\log(uv)}{1-uv} P_n(u)P_n(v)\mathrm{d}u\mathrm{d}v$. En appliquant l'estimation asymptotique de $I_{\mathbb{R},n}(1)$ obtenue au paragraphe 1.3, on obtient une démonstration de l'irrationalité de $\zeta(3)$.

En fait un couple $(\widetilde{A}_n, \widetilde{B}_n)$ vérifie (1.14) si, et seulement si, il existe C_n et D_n tels que $(\widetilde{A}_n, \widetilde{B}_n, C_n, D_n)$ soit une solution du problème de Padé (1.13). Plus précisément, la première (resp. la deuxième) assertion de (1.13) équivaut à la première (resp. la deuxième) assertion de (1.14) (il s'agit d'un fait général : voir par exemple [N5], Chapitre 4, §3.4). Démontrons-le pour la deuxième. Soient Γ un chemin qui entoure le segment $[0, 1]$ dans sens direct, et $k \in \{0, \ldots, n-1\}$. On a :

$$\frac{1}{2i\pi}\int_\Gamma z^k \left(2\widetilde{A}_n(z)\mathrm{Li}_3(1/z) + \widetilde{B}_n(z)\mathrm{Li}_2(1/z)\right) \mathrm{d}z = -\int_0^1 \left(\widetilde{B}_n(x) - \widetilde{A}_n(x)\log(x)\right) x^k \log(x)\,\mathrm{d}x,$$

d'après (1.16), en intervertissant les deux signes d'intégration et en appliquant le théorème des résidus.

Il découle de ceci que le problème (1.14) a une solution unique (à proportionnalité près), donnée par $\widetilde{A}_n = A_n$ et $\widetilde{B}_n = B_n$.

1.8 D'autres problèmes d'approximation de Padé

Sorokin [S7] considère le problème de Padé suivant : pour $n \geq 0$, trouver des polynômes T_n, U_n, V_n, W_n de degré au plus n tels qu'on ait :

$$\begin{cases} I_{\mathrm{P},n}(z) := T_n(z)\mathrm{Le}_{2,1}(1/z) + U_n(z)\mathrm{Le}_{1,1}(1/z) + V_n(z)\mathrm{Li}_1(1/z) + W_n(z) = \mathrm{O}(z^{-n-1}) \\ \hfill \text{quand } z \to \infty \\ T_n(z)\mathrm{Li}_2(1-z) + V_n(z) = \mathrm{O}((1-z)^{n+1}) \text{ quand } z \to 1 \\ T_n(z)\mathrm{Li}_1(1-z) + U_n(z) = \mathrm{O}((1-z)^{n+1}) \text{ quand } z \to 1, \end{cases}$$

où pour $s_1, \ldots, s_k \geq 1$ on définit le polylogarithme multiple

$$\mathrm{Le}_{s_1,\ldots,s_k}(z) = \sum_{n_1 \geq \ldots \geq n_k \geq 1} \frac{z^{n_1}}{n_1^{s_1} \ldots n_k^{s_k}},$$

qui vérifie $\mathrm{Le}_{2,1}(1) = 2\zeta(3)$ (voir [W7]).

Sorokin démontre que ce problème de Padé admet une solution unique, et qu'à proportionnalité près elle vérifie (pour $z \in \mathbb{C} \setminus [0, 1[$) :

$$I_{\mathrm{P},n}(z) = z^{n+1} \int_0^1 \int_0^1 \int_0^1 \frac{u^n(1-u)^n v^n(1-v)^n w^n(1-w)^n}{(z-uv)^{n+1}(z-uvw)^{n+1}} \mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w.$$

Avec cette normalisation, T_n est à coefficients entiers (donc aussi d_nU_n, $d_n^2V_n$ et $d_n^3W_n$), d'où :

$$I_{\mathrm{P},n}(1) = 2(u_{\mathrm{P},n}\zeta(3) - v_{\mathrm{P},n}) \text{ avec } u_{\mathrm{P},n} \in \mathbb{Z} \text{ et } 2d_n^3 v_{\mathrm{P},n} \in \mathbb{Z}.$$

De plus l'expression intégrale donne facilement l'estimation asymptotique de $I_{\mathrm{P},n}(1)$; c'est ainsi que Sorokin démontre l'irrationalité de $\zeta(3)$.

Un théorème général de Zlobin [Z5] montre qu'on a

$$I_{\mathrm{P},n}(z) = \int_0^1 \int_0^1 \int_0^1 \frac{u^n(1-u)^n v^n(1-v)^n w^n(1-w)^n}{(z-w(1-uv))^{n+1}} \mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w,$$

d'où $I_{\mathrm{P},n}(1) = I_{\mathbb{R},n}(1)$. On peut obtenir directement ce résultat en appliquant le changement de variables ([F2], §2) défini par $U = 1-w$, $V = \frac{(1-u)v}{1-uv}$ et $W = u$ (et qui vérifie $1 - W(1 - UV) = \frac{(1-u)(1-uvw)}{1-uv}$).

Il existe plusieurs autres problèmes de Padé liés à $\zeta(3)$; l'un d'entre eux [S5] fait apparaître l'intégrale suivante :

$$\int_0^1 \int_0^1 \int_0^1 \frac{u^n(1-u)^n v^n(1-v)^n w^n(1-w)^n}{(z(1-u+uv)-uvw)^{n+1}} \mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w.$$

Le changement de variables qui fixe u et w et change v en $\frac{v}{1-u(1-v)}$ transforme cette intégrale en

$$\int_0^1 \int_0^1 \int_0^1 \frac{u^n(1-u)^n v^n(1-v)^n w^n(1-w)^n}{(1-uv)^{n+1}(z-uvw)^{n+1}} \mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w.$$

Ces différents problèmes de Padé fournissent tous les formes linéaires d'Apéry en 1 et $\zeta(3)$, mais ils correspondent à des combinaisons linéaires différentes de polylogarithmes.

1.9 Série hypergéométrique très bien équilibrée

On pose :

$$\begin{aligned} H_n(X) &= n!^2(2X+n)\frac{(X-1)\dots(X-n)(X+n+1)\dots(X+2n)}{X^4(X+1)^4\dots(X+n)^4} \\ &= n!^2(2X+n)\frac{(X-n)_n(X+n+1)_n}{(X)_{n+1}^4} \end{aligned}$$

et

$$I_{\mathrm{TB},n}(z) = \sum_{k=1}^{\infty} H_n(k) z^{-k}.$$

La série $I_{\mathrm{TB},n}(1)$ a été introduite par K. Ball (voir [R8]) dans le but de répondre à une question de Nesterenko [N2] : trouver une preuve de l'irrationalité de $\zeta(3)$ analogue à celle de Fourier ([F1], Chapitre 2, §1.1) pour l'irrationalité de e. En effet, on peut estimer $I_{\mathrm{TB},n}(1)$ de manière élémentaire ([Z10], Lemme 4 ; [R6], §5.1 ; voir aussi la seconde démonstration du lemme 3 de [B2]) :

$$\lim_{n\to+\infty} \frac{\log(I_{\mathrm{TB},n}(1))}{n} = \log((\sqrt{2}-1)^4),$$

ou bien (voir le paragraphe 2.3) déduire cette estimation d'une représentation intégrale de $I_{\mathrm{TB},n}(z)$ vue comme série hypergéométrique très bien équilibrée :

$$I_{\mathrm{TB},n}(z) = z^{-n-1}\frac{n!^7(3n+2)!}{(2n+1)!^5}\, {}_7F_6\left(\begin{matrix} 3n+2, & \frac{3}{2}n+2, & n+1, & \dots, & n+1 \\ & \frac{3}{2}n+1, & 2n+2, & \dots, & 2n+2 \end{matrix} \,\middle|\, z^{-1} \right).$$

De plus, on a $I_{\mathrm{TB},n}(z) = P_0(z) + \sum_{j=1}^4 P_j(z)\mathrm{Li}_j(1/z)$ avec des polynômes $P_0, \ldots, P_4 \in \mathbb{Q}[z]$ vérifiant $P_j(z) = (-1)^{j+1} z^4 P_j(1/z)$ pour tout $j \in \{1, \ldots, 4\}$, $P_1(1) = 0$ et $d_n^{4-j} P_j(z) \in \mathbb{Z}[z]$ pour tout $j \in \{0, \ldots, 4\}$ (ceci sera généralisé au paragraphe 2.3). En particulier, on en déduit

$$I_{\mathrm{TB},n}(1) = 2(u_{\mathrm{TB},n}\zeta(3) - v_{\mathrm{TB},n}) \text{ avec } 2d_n u_{\mathrm{TB},n} \in \mathbb{Z} \text{ et } 2d_n^4 v_{\mathrm{TB},n} \in \mathbb{Z}. \tag{1.17}$$

Mais ceci ne suffit pas à démontrer l'irrationalité de $\zeta(3)$, car $(\sqrt{2}-1)^4 e^4 > 1$.

Une identité de Bailey ([Z9], Proposition 2 ; [S4], formule (4.7.1.3)) donne $I_{\mathrm{TB},n}(1) = I_{\mathbb{C},n}(1)$. Une telle identité ne peut pas avoir lieu pour tout z, car $\mathrm{Li}_4(1/z)$ apparaît dans la décomposition en polylogarithmes de $I_{\mathrm{TB},n}(z)$ mais pas dans celle de $I_{\mathbb{C},n}(z)$. Par ailleurs Zudilin a démontré une identité générale ([Z8], Théorème 5) qui écrit une série hypergéométrique très bien équilibrée sous la forme d'une intégrale généralisant celles introduites par Beukers [B4], Vasilenko [V1] et Vasilyev ([V2], [V3]). Dans notre cas particulier, cette identité est $I_{\mathrm{TB},n}(1) = I_{\mathbb{R},n}(1)$. Enfin, en utilisant les algorithmes décrits dans [P2] on peut démontrer que $I_{\mathrm{TB},n}(1)$ ([R6], §5.1 ; [Z10]), ainsi que $u_{\mathrm{TB},n}$ et $v_{\mathrm{TB},n}$ [K4], vérifient la relation de récurrence (1.1). On en déduit $u_{\mathrm{TB},n} = u_{\mathrm{E},n}$ et $v_{\mathrm{TB},n} = v_{\mathrm{E},n}$, d'où $u_{\mathrm{TB},n} \in \mathbb{Z}$ et $2d_n^3 v_{\mathrm{TB},n} \in \mathbb{Z}$ (ce qui est plus précis que (1.17)).

1.10 Preuve utilisant des formes modulaires

Dans ce paragraphe, on esquisse une preuve due à Beukers [B9] de l'irrationalité de $\zeta(3)$. Les outils mis en œuvre sont exposés dans [S1] (Chapitre VII) et [Z1].

Pour τ dans le demi-plan de Poincaré $\mathfrak{H}$, posons $q = e^{2i\pi\tau}$ et considérons les séries d'Eisenstein $E_2(\tau) = 1 - 24\sum_{n\geq 1}\sigma_1(n)q^n$ et $E_4(\tau) = 1 + 240\sum_{n\geq 1}\sigma_3(n)q^n$. On pose :

$$\begin{array}{rrcl} & E(\tau) & = & \dfrac{1}{24}\left(-5E_2(\tau) + 2E_2(2\tau) - 3E_2(3\tau) + 30E_2(6\tau)\right) \\ \text{et} & F(\tau) & = & \dfrac{1}{40}\left(E_4(\tau) - 28E_4(2\tau) + 63E_4(3\tau) - 36E_4(6\tau)\right). \end{array}$$

Alors $E(\tau)$, respectivement $F(\tau)$, est une forme modulaire de poids 2, resp. 4, pour $\Gamma_0(6)$. Si $F(\tau) = \sum_{n\geq 1} f_n q^n$ désigne le développement de Fourier de F à l'infini (où elle s'annule), on pose $f(\tau) = \sum_{n\geq 1}\frac{f_n}{n^3}q^n$. On a alors $(\frac{\mathrm{d}}{\mathrm{d}\tau})^3 f(\tau) = (2i\pi)^3 F(\tau)$.

Considérons la fonction modulaire pour $\Gamma_0(6)$ donnée par :

$$t(\tau) = \left(\frac{\Delta(6\tau)\Delta(\tau)}{\Delta(2\tau)\Delta(3\tau)}\right)^{1/2} = q \prod_{\substack{n\geq 1 \\ \mathrm{pgcd}(n,6)=1}} (1-q^n)^{12},$$

avec $\Delta(\tau) = q\prod_{n\geq 1}(1-q^n)^{24}$. Elle n'a ni zéro ni pôle dans $\mathfrak{H}$. Au voisinage de $q = 0$, $t(\tau) = q - 12q^2 + 66q^3 - \ldots$ s'écrit comme une série entière en q, à coefficients entiers, avec un rayon de convergence égal à 1. Elle admet une réciproque locale, notée $q(t) \in \mathbb{Z}[[t]]$. Par composition, on peut donc définir des suites $(u_{\mathrm{M},n})$ et $(v_{\mathrm{M},n})$ par :

$$\begin{array}{rcl} E(q(t)) & = & \displaystyle\sum_{n\geq 0} u_{\mathrm{M},n} t^n \in \mathbb{Z}[[t]] \\ \text{et } E(q(t))f(q(t)) & = & \displaystyle\sum_{n\geq 0} v_{\mathrm{M},n} t^n \in \mathbb{Q}[[t]] \text{ avec } v_{\mathrm{M},0} = 0 \text{ et } d_n^3 v_{\mathrm{M},n} \in \mathbb{Z} \text{ pour tout } n \geq 1. \end{array}$$

Notons, pour $k \in \mathbb{Z}$, w_k l'opérateur d'Atkin-Lehner défini par $(w_k g)(\tau) = 6^{-k/2}\tau^{-k}g(\frac{-1}{6\tau})$. Alors $w_2(E) = -E$ et $w_4(F) = -F$. De cette seconde égalité (et d'un lemme de Hecke : voir [W9], §5) découle la relation $w_{-2}(h) = -h$, en posant $h(\tau) = L(F,3) - f(\tau)$, où $L(F,s)$ est la fonction L de F. Il vient alors $w_0(Eh) = Eh$, c'est-à-dire que la fonction $E(\tau)h(\tau)$ est invariante par la substitution $\tau \mapsto \frac{-1}{6\tau}$.

Considérons maintenant les rayons de convergence. La fonction $t(\tau)$ est ramifiée seulement au-dessus des points $(\sqrt{2}-1)^4$, $(\sqrt{2}+1)^4$ et ∞. Au-dessus de $(\sqrt{2}-1)^4$, le seul point de ramification (modulo $\Gamma_0(6)$) est $\tau = i/\sqrt{6}$; il est d'indice deux, et les deux branches en ce point sont échangées par l'involution $\tau \mapsto \frac{-1}{6\tau}$. Comme $E(\tau)h(\tau)$ est invariante par cette involution, on peut définir Eh comme une fonction de t au voisinage de $t = (\sqrt{2}-1)^4$, et en fait sur tout le disque $|t| < (\sqrt{2}+1)^4$. Cela signifie que la série $\sum_{n\geq 0}(L(F,3)u_{\mathrm{M},n} - v_{\mathrm{M},n})t^n$ a un rayon de convergence supérieur ou égal à $(\sqrt{2}+1)^4$, c'est-à-dire qu'on a :

$$\limsup_{n\to\infty} \frac{\log|L(F,3)u_{\mathrm{M},n} - v_{\mathrm{M},n}|}{n} \leq \log((\sqrt{2}-1)^4).$$

Ceci conclut la démonstration de l'irrationalité de $L(F,3)$. Or on peut calculer explicitement $L(F,s)$. En effet, quand $\mathrm{Re}(s) > 4$ on a, pour tout entier $j \geq 1$:

$$L(E_4(j\tau),s) = 1 + 240\sum_{n\geq 1}\frac{\sigma_3(n)}{(jn)^s} = 1 + 240\sum_{d,e\geq 1}\frac{d^3}{(jde)^s} = 1 + 240\zeta(s)\zeta(s-3)j^{-s}.$$

On en déduit immédiatement $L(F,s) = -2\zeta(s)\zeta(s-3)$, d'où $L(F,3) = \zeta(3)$.

Comme $E(\tau)$ est une forme modulaire de poids 2 et $t(\tau)$ une fonction modulaire, la fonction $E(q(t))$ de la variable t est solution [Z2] (voir aussi [B7], p. 58) d'une équation différentielle linéaire $\mathfrak{D}y = 0$, d'ordre trois. On peut la déterminer explicitement :

$$\mathfrak{D} = (t^4 - 34t^3 + t^2)\frac{\mathrm{d}^3}{\mathrm{d}t^3} + (6t^3 - 153t^2 + 3t)\frac{\mathrm{d}^2}{\mathrm{d}t^2} + (7t^2 - 112t + 1)\frac{\mathrm{d}}{\mathrm{d}t} + (t-5).$$

Cette équation différentielle vérifiée par la série génératrice des $u_{\mathrm{M},n}$ montre qu'ils satisfont à la relation de récurrence (1.1) : on a donc $u_{\mathrm{M},n} = u_{\mathrm{R},n}$ (voir aussi [B10]). En posant $V(t) = E(q(t))f(q(t))$ on montre [Z2] que $\mathfrak{D}V = 5$, d'où $v_{\mathrm{M},n} = v_{\mathrm{R},n}$.

Une base de solutions de l'équation différentielle $\mathfrak{D}y = 0$ est donnée par $E(q(t))$, $\tau(t)E(q(t))$ et $\tau^2(t)E(q(t))$ (voir aussi [B11], Corollaire 2). La seule solution qui soit régulière en 0 est $E(q(t))$ (à proportionnalité près). De plus, la construction de $\mathfrak{D}$ montre [Z2] que c'est un carré symétrique, ce qui peut se vérifier directement (voir [D4]).

Remarque 1.5 *Le point de vue adopté dans ce paragraphe est lié "individuellement" à $\zeta(3)$ (qui est vu comme valeur spéciale d'une fonction L), par opposition aux méthodes utilisées dans les paragraphes 1.3 à 1.9, où $\zeta(3)$ apparaissait comme la valeur en 1 d'un polylogarithme.*

Cette preuve de l'irrationalité de $\zeta(3)$ s'exprime naturellement en termes des séries génératrices $U(t) = \sum_{n\geq 0} u_n t^n$ et $V(t) = \sum_{n\geq 0} v_n t^n$ des approximations rationnelles de $\zeta(3)$ (voir [P5], [B9] et [C5], §5 pour d'autres preuves dans le même esprit). L'aspect arithmétique consiste à démontrer que les coefficients de $U(t)$ sont entiers, et que d_n^3 est un dénominateur commun aux n premiers coefficients de $V(t)$: c'est une majoration p-adique de ces coefficients, pour toute place finie p. L'aspect analytique est une minoration, par $(1+\sqrt{2})^4$, du rayon de convergence (archimédien) de la série entière $\zeta(3)U(t)-V(t)$. En particulier, $U(t)$ et $V(t)$ sont des G-fonctions de Siegel. La série $U(t)$ est une solution de l'équation différentielle $\mathfrak{D}y=0$; la conjecture de Bombieri-Dwork prédit ([D4], [D5] ; voir aussi [A6] et [D6]) que $\mathfrak{D}$ provient de la géométrie.

Or, pour $t \in \mathbb{P}^1(\mathbb{C}) \setminus \{0, 1, (\sqrt{2} \pm 1)^4, \infty\}$, Beukers et Peters construisent [B11] une surface K3 X_t birationnellement équivalente à la surface projective S_t d'équation affine $1-(1-xy)z-txyz(1-x)(1-y)(1-z)=0$. Ils montrent que si ω_t est l'unique 2-forme holomorphe sur X_t (à proportionnalité près), et si τ_t est un certain 2-cycle (constant pour la connexion de Gauss-Manin), alors $U(t)$ est l'intégrale de ω_t sur τ_t. En particulier $\mathfrak{D}y=0$ est l'équation de Picard-Fuchs de cette famille de surfaces : elle provient bien de la géométrie.

1.11 Congruences

De nombreux auteurs ont étudié des propriétés de congruence sur les nombres d'Apéry u_n. Par exemple, Chowla, Cowles et Cowles [C4] ont conjecturé $u_p \equiv 5 \mod p^3$ pour tout $p \geq 5$ premier. Cette conjecture a été démontrée par plusieurs auteurs (voir par exemple [G2], [S9], …). De nombreuses autres congruences ont été prouvées, pour les nombres d'Apéry et certaines de leurs généralisations.

Notons $\sum_{n\geq 1} \gamma_n q^n = q\prod_{n\geq 1}(1-q^{2n})^4(1-q^{4n})^4$ l'unique forme parabolique normalisée de poids 4 pour $\Gamma_0(8)$. Pour $r \geq 1$, $m \geq 1$ impair et p premier impair, on a la congruence suivante (qui ressemble à celles d'Atkin - Swinnerton-Dyer, voir [H7] §VI.33) :

$$u_{\frac{1}{2}(mp^r-1)} - \gamma_p u_{\frac{1}{2}(mp^{r-1}-1)} + p^3 u_{\frac{1}{2}(mp^{r-2}-1)} \equiv 0 \mod p^r \tag{1.18}$$

avec la convention $u_t = 0$ si $t \notin \mathbb{Z}$. Beukers la démontre [B10] en utilisant la construction modulaire du paragraphe 1.10. On en déduit $u_{\frac{p-1}{2}} \equiv \gamma_p \mod p$, congruence dont Beukers a conjecturé [B10] qu'elle est vraie modulo p^2. Ceci a été prouvé par Ishikawa [I3] si p ne divise pas $u_{\frac{p-1}{2}}$, puis par Ahlgren et Ono [A2] dans le cas général. Ahlgren et Ono utilisent des séries hypergéométriques sur $\mathbb{F}_p$ et la modularité de la variété d'équation $x+\frac{1}{x}+y+\frac{1}{y}+z+\frac{1}{z}+w+\frac{1}{w}=0$ (dont la famille de surfaces K3 considérée par Beukers-Peters est un quotient : voir [P1], Théorème 4).

Pour $r, m \geq 1$ et $p \geq 5$ premier, Beukers a démontré [B8], de manière élémentaire, qu'on a $u_{mp^r-1} \equiv u_{mp^{r-1}-1} \mod p^{3r}$. La même congruence, mais seulement modulo p^r, s'interprète en disant que $\int_0^T U(t)\mathrm{d}t$ est (vue comme série formelle en T) le logarithme d'une loi de groupe formel sur $\mathbb{Z}$ qui est isomorphe à $\mathbb{G}_m$ sur $\mathbb{Z}$ ([B8] ; voir aussi l'appendice de [S8] ou [H7], §VI.33).

2 Irrationalité d'une infinité de $\zeta(2k+1)$

2.1 Énoncé des résultats

Dans cette partie, on démontre les résultats suivants, dont le premier implique le théorème 0.3 :

Théorème 2.1 ([R5], [B2]) *Pour $\ell \geq 3$ impair, notons δ_ℓ la dimension du $\mathbb{Q}$-espace vectoriel engendré par 1, $\zeta(3)$, $\zeta(5)$, …, $\zeta(\ell)$. Pour tout $\varepsilon > 0$ il existe un entier ℓ_0 tel que pour* *tout $\ell \geq \ell_0$ impair on ait :*

$$\delta_\ell \geq \frac{1-\varepsilon}{1+\log(2)} \log(\ell).$$

Remarque 2.2 *Si dans le théorème 2.1 on remplace $\frac{1-\varepsilon}{1+\log(2)}$ par $\frac{1}{3}$ alors [B2] on peut prendre $\ell_0 = 3$.*

Théorème 2.3 ([B2]) *Il existe un entier impair ℓ, avec $\ell \leq 169$, tel que 1, $\zeta(3)$ et $\zeta(\ell)$ soient linéairement indépendants sur $\mathbb{Q}$.*

Ce théorème a été amélioré par Zudilin [Z7], qui remplace 169 par 145, grâce à un raffinement du lemme 2.12 ci-dessous.

Les deux ingrédients essentiels de la démonstration du théorème 2.1 sont l'absence de $\zeta(2)$, $\zeta(4)$, ..., $\zeta(\ell-1)$ d'une part, et la minoration en $\log(\ell)$ de la dimension d'autre part. Seule cette deuxième idée est utile pour démontrer le théorème suivant.

Théorème 2.4 ([R6]) *Soient $z \in \mathbb{Q}$, $|z| > 1$, et $\varepsilon > 0$. Il existe un entier ℓ_0 (qui dépend de z et ε) tel que, pour tout $\ell \geq \ell_0$, la dimension du $\mathbb{Q}$-espace vectoriel engendré par $1, \mathrm{Li}_1(1/z), \mathrm{Li}_2(1/z), \ldots, \mathrm{Li}_\ell(1/z)$ soit minorée par $\frac{1-\varepsilon}{1+\log(2)} \log(\ell)$.*

En conséquence, pour tout nombre rationnel z de valeur absolue supérieure à 1 il existe une infinité d'entiers j tels que $\mathrm{Li}_j(1/z)$ soit irrationnel. Par ailleurs, quand z est un entier négatif tel que $|z| > (4\ell)^{\ell(\ell-1)}$, Nikishin a démontré [N4] que les nombres $1, \mathrm{Li}_1(1/z), \mathrm{Li}_2(1/z), \ldots, \mathrm{Li}_\ell(1/z)$ sont linéairement indépendants sur $\mathbb{Q}$; sa méthode a inspiré en partie la construction exposée au paragraphe suivant. Hata a raffiné ([H4], [H5]) le résultat de Nikishin : par exemple 1, $\mathrm{Li}_1(1/z)$ et $\mathrm{Li}_2(1/z)$ sont linéairement indépendants sur $\mathbb{Q}$ pour $z \leq -5$ ou $z \geq 7$.

2.2 Structure de la preuve

Soient a et r deux entiers, avec $a \geq 3$ et $1 \leq r < \frac{a}{2}$. Soit $n \geq 1$. Définissons $\mathbf{R}_n$ et $\mathbf{S}_n$ (qui dépendent aussi de a et r) par :

$$\begin{aligned}
\mathbf{R}_n(k) &= 2n!^{a-2r}(k+\frac{n}{2})\frac{(k-rn)_{rn}(k+n+1)_{rn}}{(k)_{n+1}^a} \\
&= 2n!^{a-2r}(k+\frac{n}{2})\frac{(k-1)(k-2)\ldots(k-rn)(k+n+1)(k+n+2)\ldots(k+(r+1)n)}{k^a(k+1)^a\ldots(k+n)^a}
\end{aligned}$$

et

$$\mathbf{S}_n(z) = \sum_{k\geq 1} \mathbf{R}_n(k) z^{-k}. \tag{1.19}$$

Cette série converge absolument pour tout nombre complexe z tel que $|z| \geq 1$, car $\mathbf{R}_n(k) = \mathrm{O}(k^{-2})$ quand k tend vers l'infini.

Les propriétés de cette série étudiées au paragraphe 2.3 permettent de démontrer les théorèmes 2.1 (en prenant $z = 1$ et a pair), 2.3 (avec $z = 1$, $a = 169$, $r = 10$ et n impair ; on utilise le théorème d'Apéry) et 2.4 (avec $z \in \mathbb{Q}$, $z > 1$; pour $z < -1$ il suffirait de modifier le lemme 2.9). Les trois preuves sont parallèles ; on détaille dans ce paragraphe la structure de celle du théorème 2.1.

On suppose a pair ; on construit des formes linéaires en 1, $\zeta(3)$, $\zeta(5)$, ..., $\zeta(a-1)$ grâce à la proposition suivante :

Proposition 2.5 *Supposons a pair. Notons d_n le p.p.c.m des entiers de 1 à n. Alors il existe des nombres rationnels κ_0, κ_3, κ_5, ..., κ_{a-1} tels que :*

1. *On a $\mathbf{S}_n(1) = \kappa_0 + \kappa_3\,\zeta(3) + \kappa_5\,\zeta(5) + \kappa_7\,\zeta(7) + \ldots + \kappa_{a-1}\,\zeta(a-1)$.*
2. *Pour tout $j \in \{0, 3, 5, \ldots, a-1\}$ on a $\limsup_{n\to+\infty} |\kappa_j|^{1/n} \leq 2^{a-2r}(2r+1)^{2r+1}$.*
3. *Pour tout $j \in \{0, 3, 5, \ldots, a-1\}$, le nombre rationnel $d_n^a \kappa_j$ est un entier.*
4. *Il existe un réel $\psi_{r,a} > 0$ tel que $\lim_{n\to+\infty} |\mathbf{S}_n(1)|^{1/n} = \psi_{r,a} \leq \frac{2^{r+1}}{r^{a-2r}}$.*

En fait on conjecture que l'amélioration suivante est possible :

Conjecture 2.6 ([R6]) *Dans l'assertion* (3) *de la proposition 2.5, on peut remplacer d_n^a par d_n^{a-1}.*

Remarque 2.7 *En prenant $a = 4$ (et $r = 1$), on obtient les formes linéaires en 1 et $\zeta(3)$ du paragraphe 1.9, donc la conjecture 2.6 est vraie quand $a = 4$. Elle est démontrée aussi quand $a = 6$ et $r = 1$ (voir la fin du paragraphe 2.4). On ne connaît pas de conséquence directe de cette conjecture, mais une version forte de celle-ci pourrait éventuellement permettre de démontrer que parmi $\zeta(5)$, $\zeta(7)$ et $\zeta(9)$, l'un au moins est irrationnel (voir la remarque 3.4). En tout cas, il serait intéressant d'obtenir une preuve de la conjecture 2.6 grâce à une interprétation (par exemple géométrique, comme au paragraphe 1.10) de κ_0, ..., κ_{a-1}.*

La proposition 2.5 fournit des formes linéaires en 1, $\zeta(3)$, ..., $\zeta(a-1)$ (si a est pair). Si cette suite de formes linéaires tend vers 0, sans être nulle à partir d'un certain rang, alors l'un au moins des nombres $\zeta(3)$, ..., $\zeta(a-1)$ est irrationnel. Cette remarque sera utilisée pour démontrer le théorème 0.4. Ici on veut obtenir les théorèmes 2.1 à 2.4, donc on a besoin d'un critère d'indépendance linéaire, qui donne une minoration plus fine de la dimension du $\mathbb{Q}$-espace vectoriel engendré par 1, $\zeta(3)$, ..., $\zeta(a-1)$. On va utiliser à cet effet le théorème 2.8 ci-dessous.

La meilleure minoration qu'on puisse espérer est donnée par le principe des tiroirs, de la manière suivante. Soient α et β des réels, avec $0 < \alpha < 1$ et $\beta > 1$. Soient $\theta_1, \ldots, \theta_s$ des réels qui engendrent un $\mathbb{Q}$-espace vectoriel de dimension au moins $1 - \frac{\log(\alpha)}{\log(\beta)}$. Alors il existe une suite (ℓ_n) de formes linéaires en θ_1, ..., θ_s dont les coefficients entiers $p_{j,n}$ vérifient $\limsup_{n\to+\infty} |p_{j,n}|^{1/n} \leq \beta$ pour tout j et telle que $\limsup_{n\to+\infty} |\ell_n(\theta_1, \ldots, \theta_s)|^{1/n} \leq \alpha$. Essentiellement, plus la dimension du $\mathbb{Q}$-espace vectoriel engendré est grande, plus les formes linéaires qu'on peut construire sont petites. On cherche une réciproque à cette assertion. Une contrainte supplémentaire est nécessaire : si $\frac{\theta_2}{\theta_1}$ est un nombre de Liouville, on peut construire des formes linéaires extrêmement petites même si la dimension du $\mathbb{Q}$-espace vectoriel engendré est seulement 2. Ce contre-exemple ne tient plus si on demande que les formes linéaires en $\theta_1, \ldots, \theta_s$ ne soient pas trop petites. On a alors la réciproque suivante (pour une preuve, voir [N1] ou [C8], §II.1) :

Théorème 2.8 ([N1]) *Soient $\theta_1, \ldots, \theta_s$ des réels. Pour tout $n \geq 1$, soit $\ell_n = p_{1,n}X_1 + \ldots + p_{s,n}X_s$ une forme linéaire à coefficients entiers. Soient α et β des réels, avec $0 < \alpha < 1$ et $\beta > 1$.*

Supposons qu'on ait $\limsup_{n\to+\infty} |p_{j,n}|^{1/n} \leq \beta$ pour tout j compris entre 1 et s, et

$$\lim_{n\to+\infty} |\ell_n(\theta_1, \ldots, \theta_s)|^{1/n} = \alpha.$$

Alors le $\mathbb{Q}$-espace vectoriel engendré par $\theta_1, \ldots, \theta_s$ est de dimension au moins $1 - \frac{\log(\alpha)}{\log(\beta)}$.

Pour déduire le théorème 2.1 de la proposition 2.5 et de ce critère d'indépendance linéaire, il suffit de considérer $d_n^a \mathbf{S}_n(1)$, qui est une forme linéaire à coefficients entiers en 1, $\zeta(3)$, $\zeta(5)$, …, $\zeta(a-1)$. On choisit a suffisamment grand, et r égal à la partie entière de $\frac{a}{(\log(a))^2}$. Alors r^r est négligeable devant c^a (pour toute constante c), et on peut prendre β essentiellement égal à $(2e)^a = e^{a(1+\log(2))}$ et α essentiellement majoré par r^{-a}, qui est de l'ordre de $e^{-a\log(a)}$. Cela démontre le théorème 2.1.

2.3 Quelques détails sur la preuve

Soit z un nombre complexe de module supérieur ou égal à 1. La série $\mathbf{S}_n(z)$ peut s'écrire comme une série hypergéométrique très bien équilibrée, de la manière suivante :

$$\mathbf{S}_n(z) = z^{-rn-1} n!^{a-2r} \frac{(rn)!((r+1)n+2)_{rn+1}}{(rn+1)_{n+1}^a} \times$$
$$ {}_{a+3}F_{a+2}\left(\begin{matrix} (2r+1)n+2, & (r+\frac{1}{2})n+2, & rn+1, & \ldots, & rn+1 \\ & (r+\frac{1}{2})n+1, & (r+1)n+2, & \ldots, & (r+1)n+2 \end{matrix} \,\middle|\, z^{-1} \right).$$

Cette identité provient de simplifications dans les symboles de Pochhammer.

Représentation intégrale et estimation analytique

On a la représentation intégrale suivante, pour $|z| \geq 1$:

$$\mathbf{S}_n(z) = \frac{((2r+1)n+2)!}{n!^{2r+1}} z^{(r+1)n+1} \int_{[0,1]^{a+1}} \left(\frac{\prod_{j=1}^{a+1} t_j^r (1-t_j)}{(z - t_1 t_2 \ldots t_{a+1})^{2r+1}} \right)^n \frac{z + t_1 \ldots t_{a+1}}{(z - t_1 \ldots t_{a+1})^3} \mathrm{d}t_1 \ldots \mathrm{d}t_{a+1}.$$

Cette formule (voir par exemple [R9], Lemme 1) se déduit de l'écriture de $\mathbf{S}_n(z)$ comme série hypergéométrique : pour $|z| > 1$ on applique les relations (4.1.2) et (1.5.21) de [S4], puis on prolonge à $|z| = 1$ par continuité (voir la preuve du lemme 2 de [B2]). On peut aussi obtenir une preuve directe en développant en série le dénominateur de l'intégrande ([C8], [H1]).

En calculant le maximum sur $[0,1]^{a+1}$ de la fonction dont on intègre la puissance n-ième, on déduit de cette représentation intégrale l'estimation analytique suivante :

Lemme 2.9 *On suppose $z \in \mathbb{R}$, $z \geq 1$. Le polynôme*

$$Q_{r,a,z}(s) = rs^{a+2} - (r+1)s^{a+1} + (r+1)zs - rz$$

admet une racine unique $s_0 \in [0,1]$, et elle vérifie $s_0 > \frac{r}{r+1}$. De plus, si

$$\phi_{r,a,z} = z^{-r}((r+1)s_0 - r)^r (r+1-rs_0)^{r+1}(1-s_0)^{a-2r},$$

alors

$$\lim_{n\to\infty} |\mathbf{S}_n(z)|^{1/n} = \phi_{r,a,z} \leq \frac{2^{r+1}}{z^r r^{a-2r}}.$$

Pour démontrer ce lemme, il suffit d'adapter les preuves du lemme 2.2 de [R6] et du lemme 3 de [B2]. On pourrait aussi donner une démonstration élémentaire de ce comportement asymptotique, sans utiliser la représentation intégrale (comme la deuxième preuve du lemme 3 de [B2]). Enfin, une troisième possibilité serait d'écrire $\mathbf{S}_n(z)$ comme intégrale complexe et d'appliquer la méthode du col ; mais cette méthode est très difficile à mettre en œuvre quand r, a et z sont des paramètres.

Remarque 2.10 *Pour démontrer les théorèmes 2.1 et 2.4, il suffit de connaître l'existence de la limite de $|\mathbf{S}_n(z)|^{1/n}$, et sa majoration par $\frac{2^{r+1}}{z^r r^{a-2r}}$. La valeur exacte de $\phi_{r,a,z}$ n'est utile que pour obtenir des estimations numériques précises (par exemple pour le théorème 2.3).*

Décomposition en polylogarithmes

Pour démontrer que $\mathbf{S}_n(z)$ est une combinaison linéaire (à coefficients rationnels) de 1, $\mathrm{Li}_1(1/z)$, …, $\mathrm{Li}_a(1/z)$ quand $|z|>1$, il suffit de décomposer la fraction rationnelle $\mathbf{R}_n$ en éléments simples, sous la forme suivante :

$$\mathbf{R}_n(k) = \sum_{i=0}^{n}\sum_{j=1}^{a} \frac{c_{i,j}}{(k+i)^j} \tag{1.20}$$

où les coefficients $c_{i,j}$ sont des rationnels, donnés par

$$c_{i,j} = \frac{1}{(a-j)!}\left(\frac{\mathrm{d}}{\mathrm{d}X}\right)^{a-j} (\mathbf{R}_n(X)(X+i)^a)_{|X=-i}. \tag{1.21}$$

On a pour $|z|>1$:

$$\begin{aligned}
\mathbf{S}_n(z) &= \sum_{i=0}^{n}\sum_{j=1}^{a} c_{i,j} \sum_{k\geq 1} \frac{z^{-k}}{(k+i)^j} \\
&= \sum_{i=0}^{n}\sum_{j=1}^{a} c_{i,j} z^i \mathrm{Li}_j(1/z) - \sum_{i=0}^{n}\sum_{j=1}^{a} c_{i,j} \sum_{q=1}^{i} \frac{z^{i-q}}{q^j},
\end{aligned}$$

d'où

$$\mathbf{S}_n(z) = P_0(z) + \sum_{j=1}^{a} P_j(z)\mathrm{Li}_j(1/z) \tag{1.22}$$

en posant

$$P_0(z) = -\sum_{\ell=0}^{n-1}\left(\sum_{i=\ell+1}^{n}\sum_{j=1}^{a} \frac{c_{i,j}}{(i-\ell)^j}\right) z^\ell \tag{1.23}$$

et

$$P_j(z) = \sum_{i=0}^{n} c_{i,j} z^i \text{ pour } j \in \{1,\dots,a\}. \tag{1.24}$$

Bien sûr, les P_j et les $c_{i,j}$ dépendent aussi de n, a et r.

Propriété de symétrie

La fonction $\mathbf{R}_n$ vérifie la propriété de symétrie suivante :

$$\mathbf{R}_n(-k-n) = (-1)^{a(n+1)+1}\mathbf{R}_n(k).$$

Cette symétrie est rendue possible par la présence des deux facteurs de Pochhammer au numérateur de $\mathbf{R}_n(k)$: quand k est changé en $-k-n$, ils sont permutés (on applique la formule $(-\alpha)_p = (-1)^p(\alpha - p + 1)_p$).

L'unicité du développement en éléments simples montre que $c_{i,j} = (-1)^{j+a(n+1)+1} c_{n-i,j}$ pour tous $i \in \{0, \ldots, n\}$ et $j \in \{1, \ldots, a\}$, ce qui donne pour tout $j \in \{1, \ldots, a\}$:

$$P_j(z) = (-1)^{j+a(n+1)+1} z^n P_j(1/z). \tag{1.25}$$

En particulier, si $j + a(n+1)$ est pair alors $P_j(1) = 0$. De plus on a $P_1(1) = 0$, car $P_1(1) = \sum_{i=0}^{n} c_{i,1}$ est l'opposé du résidu à l'infini de $\mathbf{R}_n$ (on peut aussi faire tendre z vers 1 dans (1.22) et constater que le seul terme qui puisse tendre vers l'infini est $P_1(z)\mathrm{Li}_1(1/z)$). Quand a est pair, on obtient donc :

$$\mathbf{S}_n(1) = P_0(1) + P_3(1)\zeta(3) + P_5(1)\zeta(5) + \ldots + P_{a-1}(1)\zeta(a-1).$$

Quand a est impair et n pair, on obtient de même une forme linéaire en 1, $\zeta(2)$, $\zeta(4)$, ..., $\zeta(a-1)$ dont on peut se servir pour montrer qu'une infinité de puissances de π sont linéairement indépendantes sur $\mathbb{Q}$, i.e. que π est transcendant. On peut aussi en déduire une mesure de transcendance de π, à la manière de Reyssat [R2].

Enfin, quand a et n sont impairs, on obtient une forme linéaire en 1, $\zeta(3)$, $\zeta(5)$, ..., $\zeta(a)$; c'est ce qu'on utilise pour démontrer le théorème 2.3.

Majoration des coefficients de la forme linéaire

Lemme 2.11 *Pour tout $j \in \{0, \ldots, a\}$ on a :*

$$\limsup_{n \to +\infty} |P_j(z)|^{1/n} \leq 2^{a-2r}(2r+1)^{2r+1}|z|.$$

Démonstration : On peut suivre la démonstration du lemme 4 de [B2] en écrivant la formule de Cauchy sur le cercle C de centre $-i$ et de rayon $1/2$:

$$c_{i,j} = \frac{1}{2i\pi} \int_C \mathbf{R}_n(t)(t+i)^{j-1} \mathrm{d}t.$$

On majore ensuite le module de l'intégrande, et le lemme en découle. Une autre preuve, qui conduit à une majoration légèrement moins précise, est donnée dans [C8] et [H1].

Estimation arithmétique

Les polynômes $P_0, \ldots, P_a$ sont à coefficients rationnels ; on a besoin d'un dénominateur commun pour leurs coefficients.

Lemme 2.12 *Pour tout $j \in \{0, \ldots, a\}$, le polynôme $d_n^{a-j} P_j(z)$ est à coefficients entiers.*

Remarque 2.13 *On peut ([Z7], §4) raffiner ce lemme, ce qui permet de remplacer 169 par 145 dans l'énoncé du théorème 2.3. Cependant, des exemples montrent qu'on ne peut pas espérer remplacer d_n^{a-j} par d_n^{a-1-j}. La conjecture 2.6 signifie que pour $z = 1$ on a des compensations particulières qui font chuter le dénominateur.*

DÉMONSTRATION : Posons $F_s(X) = \frac{(X-sn)_n}{(X)_{n+1}}$ et $G_s(X) = \frac{(X+sn+1)_n}{(X)_{n+1}}$ pour tout $s \in \{1, \ldots, r\}$, ainsi que $H(X) = \frac{n!}{(X)_{n+1}}$ et $I(X) = 2X + n$. Alors on a $F_s(X) = \sum_{p=0}^n \frac{f_{p,s}}{X+p}$ avec $f_{p,s} = (-1)^{n-p}\binom{n}{p}\binom{p+sn}{n} \in \mathbb{Z}$, et de même (avec des notations évidentes) $g_{p,s} \in \mathbb{Z}$ et $h_p \in \mathbb{Z}$ pour tous p, s. On obtient alors le développement en éléments simples de $\mathbf{R}_n(X) = (\prod_{s=1}^r F_s(X)) \cdot (\prod_{s=1}^r G_s(X)) \cdot H(X)^{a-2r} \cdot I(X)$ en faisant le produit des développements des facteurs. On utilise les formules $\frac{2X+n}{X+p} = 2 + \frac{n-2p}{X+p}$ et $\frac{1}{(X+p)(X+p')} = \frac{1}{(p'-p)(X+p)} + \frac{1}{(p-p')(X+p')}$ pour $p \neq p'$; les dénominateurs n'apparaissent que par application de la seconde. Ce calcul montre que $d_n^{a-j} c_{i,j}$ est entier pour tous i, j, ce qui achève la preuve (suivant [C8] et [H1]) du lemme.

2.4 Quelques remarques

Soit Q_n un polynôme à coefficients rationnels, de degré inférieur ou égal à $a(n+1)-1$. On peut toujours considérer $\mathbf{R}_n(k) = \frac{Q_n(k)}{(k)_{n+1}^a}$ et $\mathbf{S}_n(z) = \sum_{k\geq 1} \mathbf{R}_n(k) z^{-k}$, qui converge quand $|z| > 1$. Une difficulté majeure consiste à bien choisir le polynôme Q_n.

Quel que soit ce choix, on peut décomposer $\mathbf{R}_n$ en éléments simples, définir P_0, ..., P_a et obtenir une décomposition de $\mathbf{S}_n(z)$ en polylogarithmes : les formules (1.20) à (1.24) restent valables. Pour obtenir une forme linéaire en valeurs de ζ, il faut[1] faire tendre z vers 1. Tous les termes de la décomposition en polylogarithmes ont une limite finie, sauf peut-être $P_1(z)\mathrm{Li}_1(1/z)$. C'est pourquoi on suppose $P_1(1) = 0$, ce qui signifie que $\mathbf{R}_n$ n'a pas de résidu à l'infini, i.e. $\deg(Q_n) \leq a(n+1) - 2$; alors la série qui définit $\mathbf{S}_n(z)$ converge absolument dès que $|z| \geq 1$.

En outre on souhaite[2] obtenir une forme linéaire en les $\zeta(2k+1)$ seulement, c'est-à-dire avoir $P_j(1) = 0$ pour tout $j \geq 2$ pair. Pour assurer cela il est suffisant d'avoir une propriété de symétrie du polynôme Q_n, en l'occurrence $Q_n(-k-n) = (-1)^{a(n+1)+1} Q_n(k)$. C'est cette remarque qui constitue le cœur des progrès récents ([R5], [B2]). On ne sait pas du tout la généraliser, par exemple pour construire des formes linéaires en $\zeta(s)$ dans lesquelles les s appartenant à une certaine progression arithmétique n'apparaissent pas.

La forme linéaire $\mathbf{S}_n(1)$ ne sera intéressante que si elle tend suffisamment vite vers 0 quand n tend vers l'infini. Intuitivement, ce sera le cas si les premiers termes de la série qui définit $\mathbf{S}_n(1)$ sont nuls. C'est pourquoi on cherche un polynôme $Q_n(k)$ qui s'annule aux premiers entiers, en l'occurrence entre 1 et rn ; ceci signifie que $Q_n(k)$ est multiple de $(k - rn)_{rn}$. Il s'agit en fait d'un problème de type Padé : on demande aux polynômes P_0, ..., P_a d'être tels que

$$\mathbf{S}_n(z) = P_0(z) + \sum_{j=1}^{a} P_j(z)\mathrm{Li}_j(1/z) = \mathrm{O}(z^{-rn-1}) \text{ quand } z \to \infty.$$

Parmi tous les polynômes symétriques $Q_n(k)$ multiples de $(k-rn)_{rn}$ (donc nécessairement aussi multiples de $(k+n+1)_{rn}$), on a intérêt à en prendre un de degré minimal, pour que $\mathbf{S}_n(1)$ soit aussi petit que possible. Si $a(n+1)$ est impair, le polynôme $(k-rn)_{rn}(k+n+1)_{rn}$ a la bonne parité, et on peut considérer $Q_n(k) = n!^{a-2r}(k-rn)_{rn}(k+n+1)_{rn}$: on obtient la série hypergéométrique bien équilibrée de [R5] et [B2]. Si $a(n+1)$ est pair, pour obtenir le bon signe dans la propriété de symétrie de Q_n on est amené à introduire un facteur $k + \frac{n}{2}$,

[1] Voir cependant la remarque 2.14.

[2] Sauf pour démontrer le théorème 2.4 ; pour ce dernier, le polynôme $Q_n(k) = (k - rn)_{rn}$ convient aussi. C'est celui qui est utilisé dans le Chapitre 2 de [R6].

ce qui donne la série très bien équilibrée du paragraphe 2.2. Dans les deux cas, $\mathbf{S}_n(z)$ est la solution unique d'un problème de Padé (voir [H12] et [F5]).

Plus a est grand (en prenant, pour chaque a, la valeur optimale de r), plus la forme linéaire à coefficients entiers $d_n^a \mathbf{S}_n(1)$ est petite (et la présence, ou l'absence, du facteur $k + \frac{n}{2}$ a une influence négligeable sur ce comportement). Donc si on cherche des formes linéaires en 1, $\zeta(3)$, $\zeta(5)$, ..., $\zeta(2\ell+1)$, celles obtenues avec la série très bien équilibrée pour $a = 2\ell + 2$ seront meilleures que celles obtenues avec la série bien équilibrée pour $a = 2\ell + 1$ et n pair. Ceci n'a aucune influence quand ℓ tend vers l'infini, mais peut s'avérer crucial si ℓ est fixé (comme dans le théorème 0.4). En outre, si la conjecture 2.6 (qui n'a aucun équivalent pour des séries seulement bien équilibrées) est vraie alors il suffit de multiplier $\mathbf{S}_n(1)$ par d_n^{a-1}, ce qui donne une forme linéaire encore plus petite. Pour $a = 4$, on retrouve ainsi les formes linéaires d'Apéry en 1 et $\zeta(3)$ (ce qui n'est pas le cas avec la série bien équilibrée quand $a = 3$).

Remarque 2.14 *Pour démontrer le théorème 2.1 on pourrait évaluer les formes linéaires en polylogarithmes en $z = -1$ plutôt qu'en $z = 1$. Ceci induit peu de changements. Le plus notable est que* $\log(2) = -\mathrm{Li}_1(-1)$ *remplace le divergent* $\mathrm{Li}_1(1)$ *; pour $\ell \geq 2$ on a* $\mathrm{Li}_\ell(-1) = -(1 - 2^{1-\ell})\zeta(\ell)$. *Pour $a = 3$ et $z = -1$ les formes linéaires construites au paragraphe 2.3 sont [K4] celles utilisées par Apéry ([A10], [P4]) pour prouver que $\zeta(2)$ est irrationnel. En particulier d_n^2 suffit comme dénominateur des coefficients de cette forme linéaire. Plus généralement, la conjecture 2.6 devrait être valable aussi quand a est impair et $z = -1$.*

Considérons l'opérateur différentiel hypergéométrique suivant, où $\delta = z\frac{\mathrm{d}}{\mathrm{d}z}$:

$$\mathbf{L} = \delta^{a+1}(\delta - \frac{n}{2} - 1)(\delta - (r+1)n - 1) - z(\delta - n)^{a+1}(\delta - \frac{n}{2} + 1)(\delta + rn + 1).$$

L'écriture de $\mathbf{S}_n(z)$ comme série hypergéométrique très bien équilibrée montre que $\mathbf{S}_n(z)$ est une solution de l'équation différentielle $\mathbf{L}y = 0$. Par monodromie on voit, grâce à (1.22), que pour tout $b \in \{1, \ldots, a\}$ la fonction $\sum_{j=b}^{a}(-1)^{j-1}P_j(z)\frac{\log^{j-b}(z)}{(j-b)!}$ est aussi une solution de $\mathbf{L}y = 0$. En particulier pour $b = a$ on obtient le polynôme P_a, qu'on peut écrire comme polynôme hypergéométrique très bien équilibré (avec un petit abus de langage : ici les paramètres inférieurs $-\frac{n}{2}$ et $-(r+1)n$ sont négatifs, mais la série ${}_{a+3}F_{a+2}$ est quand même bien définie) :

$$\begin{aligned} P_a(z) &= (-1)^{rn} n (rn)!((r+1)n)! n!^{-2r-1} \times \\ & {}_{a+3}F_{a+2}\left(\left. \begin{matrix} -n, & -\frac{n}{2}+1, & rn+1, & -n, & \ldots, & -n \\ & -\frac{n}{2}, & -(r+1)n, & 1, & \ldots, & 1 \end{matrix} \right| z \right). \end{aligned}$$

L'aspect bien équilibré de ce polynôme hypergéométrique lui confère (voir [A8] ou [A9], §3.5) la propriété de réciprocité (1.25). En effet, si $y(z)$ est une solution de l'équation différentielle $\mathbf{L}y = 0$ alors $z^n y(1/z)$ est aussi une solution de cette même équation. Quant aux autres polynômes P_{a-1}, ..., P_1, ils s'obtiennent par la méthode de Frobenius (voir [I1]) et vérifient, eux aussi, (1.25). Toutes ces considérations valent aussi pour la série bien équilibrée de [R5] et [B2], et permettent [H12] d'écrire celle-ci comme solution unique d'un problème de Padé.

Un autre intérêt des définitions utilisées dans ce texte est que $\mathbf{S}_n(1)$ possède (pour a pair) plusieurs représentations intégrales assez simples. Tout d'abord, on a ([Z8], Théorème 5) l'intégrale suivante, qui généralise $I_{\mathbb{R},n}(1)$ et les intégrales introduites par Vasilenko [V1] et Vasilyev ([V2], [V3]) :

$$\mathbf{S}_n(1) = \frac{(rn)!^2}{n!^{2r}} \int_{[0,1]^{a-1}} \frac{\prod_{j=1}^{a-1} x_j^{rn}(1-x_j)^n}{(Q_{a-1}(x_1, \ldots, x_{a-1}))^{rn+1}} \mathrm{d}x_1 \ldots \mathrm{d}x_{a-1}, \tag{1.26}$$

en posant $Q_{a-1}(x_1, \ldots, x_{a-1}) = 1 - x_1(1 - x_2(\ldots(1 - x_{a-1})\ldots))$. Vasilyev a démontré [V3] que si $a = 6$ et $r = 1$ alors cette intégrale s'écrit $\kappa'_0 + \kappa'_3\zeta(3) + \kappa'_5\zeta(5)$ avec $d_n^5\kappa'_0$, $d_n^5\kappa'_3$ et $d_n^5\kappa'_5$ entiers. Ceci prouve la conjecture 2.6 dans ce cas. Il n'est pas évident que κ'_0, κ'_3 et κ'_5 soient les $P_0(1)$, $P_3(1)$ et $P_5(1)$ du paragraphe 2.3, mais cela découle de l'indépendance linéaire conjecturale de 1, $\zeta(3)$ et $\zeta(5)$.

D'autre part, en appliquant à (1.26) un théorème de Zlobin [Z5] ou le changement de variables qui figure dans [F2] (§2) on obtient l'intégrale suivante, qui ressemble à celles utilisées par Sorokin ([S6], [S7]) :

$$\mathbf{S}_n(1) = \frac{(rn)!^2}{n!^{2r}} \int_{[0,1]^{a-1}} \frac{\prod_{j=1}^{a-1} x_j^{rn}(1-x_j)^n \mathrm{d}x_j}{(1-x_1x_2)^{n+1}(1-x_1x_2x_3x_4)^{n+1}\ldots(1-x_1\ldots x_{a-2})^{n+1}(1-x_1\ldots x_{a-1})^{rn+1}}.$$

Il serait intéressant d'arriver à démontrer le théorème 2.1 en utilisant seulement des intégrales multiples comme celle-ci (ou celle de (1.26)). Le problème est qu'a priori on s'attend à ce qu'une telle intégrale $(a-1)$-uple soit une forme linéaire, à coefficients rationnels, en les polyzêtas de poids au plus $(a-1)$ (voir [W7] et [Z5], Théorème 3). Or le théorème 5 de [Z8] montre que ces intégrales sont égales à $\mathbf{S}_n(1)$, donc seuls 1 et les valeurs de ζ aux entiers impairs apparaissent.

3 Résultats quantitatifs

3.1 Exposant d'irrationalité de $\zeta(3)$

On appelle *exposant d'irrationalité* d'un nombre réel irrationnel α, et on note $\mu(\alpha)$, la borne inférieure de l'ensemble des réels ν pour lesquels il n'existe qu'un nombre fini de nombres rationnels p/q tels que $|\alpha - \frac{p}{q}| < \frac{1}{q^\nu}$. La théorie des fractions continues ([H2], §11.1), ou le principe des tiroirs de Dirichlet ([H2], §11.3), montre qu'un exposant d'irrationalité est toujours supérieur ou égal à 2. Si α est algébrique, Liouville a démontré ([L5] ; voir aussi [H2], §11.7) que $\mu(\alpha)$ est inférieur ou égal au degré de α. Ce résultat a été amélioré par Roth en 1955 : on a $\mu(\alpha) = 2$ pour tout nombre algébrique irrationnel α (voir [F1], Chapitre 1, §7). On a aussi $\mu(\alpha) = 2$ pour presque tout réel α, au sens de la mesure de Lebesgue ([H2], §11.11). À l'opposé, un nombre de Liouville est un nombre dont l'exposant d'irrationalité est infini : il est extrêmement bien approché par des nombres rationnels (un exemple de tel nombre est $\sum_{k\geq 1} \frac{1}{10^{k!}}$).

Les formes linéaires d'Apéry montrent que l'exposant d'irrationalité de $\zeta(3)$ est majoré par 13,4179 (voir [F1], Chapitre 2, §5.6) ; en particulier $\zeta(3)$ n'est pas un nombre de Liouville. Ce résultat a été amélioré notamment par Hata [H6] puis Rhin-Viola, qui ont démontré la meilleure majoration de $\mu(\zeta(3))$ connue à ce jour :

Théorème 3.1 ([R4]) *L'exposant d'irrationalité de $\zeta(3)$ est majoré par 5,5139, c'est-à-dire qu'il n'existe qu'un nombre fini de nombres rationnels p/q tels que*

$$|\zeta(3) - \frac{p}{q}| < \frac{1}{q^{5,5139}}.$$

Pour obtenir ce résultat, Rhin et Viola considèrent les intégrales suivantes :

$$J_n = \int_0^1\int_0^1\int_0^1 \frac{u^{hn}(1-u)^{ln}v^{kn}(1-v)^{sn}w^{jn}(1-w)^{qn}}{(1-w(1-uv))^{(q+h-r)n+1}} \mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w, \tag{1.27}$$

où $h, \dots, s$ sont des paramètres dont on fixe les valeurs de la manière suivante : $h = 16$, $j = 17$, $k = 19$, $l = 15$, $q = 11$, $r = 9$, $s = 13$. Si on prenait tous ces paramètres égaux à un même entier, on obtiendrait les intégrales du paragraphe 1.3, donc la suite des formes linéaires d'Apéry (ou, plus précisément, une suite extraite), conduisant à la même mesure d'irrationalité. L'intérêt réside donc dans le fait de ne pas prendre tous les paramètres égaux ; l'asymptotique obtenue pour $J_n^{1/n}$ est un peu moins bonne, mais on gagne beaucoup sur les dénominateurs par lesquels il faut multiplier J_n pour obtenir une forme linéaire en 1 et $\zeta(3)$ à coefficients entiers. Ce gain provient de l'action sur des intégrales de la forme (1.27) d'un groupe isomorphe au produit semi-direct $H \rtimes \mathfrak{S}_5$, où H est l'hyperplan d'équation $\varepsilon_1 + \dots + \varepsilon_5 = 0$ dans $(\mathbb{Z}/2\mathbb{Z})^5$. D'autres interprétations de cette action de groupe se trouvent dans [Z9] et [F3].

Remarque 3.2 *Les majorations de* $\mu(\zeta(3))$ *mentionnées ci-dessus sont effectives : on peut donner une majoration explicite de la hauteur* $\max(|p|, |q|)$ *des approximations rationnelles* p/q *"exceptionnellement bonnes". Ceci contraste avec le théorème de Roth, dans lequel on sait seulement majorer le nombre d'exceptions* p/q*, mais pas leur hauteur.*

3.2 Irrationalité d'un nombre parmi $\zeta(5), \dots, \zeta(21)$

Soit a un entier pair, avec $a \geq 6$. Dans ce paragraphe, on construit (en suivant [R7]) des formes linéaires à coefficients rationnels en 1, $\zeta(5)$, $\zeta(7)$, ..., $\zeta(a+1)$. Si, après multiplication par un dénominateur commun des coefficients, elles tendent vers zéro sans être nulles à partir d'un certain rang, alors l'un au moins des nombres $\zeta(5)$, $\zeta(7)$, ..., $\zeta(a+1)$ est irrationnel; c'est ce qui va se produire avec $a = 20$. On pose :

$$\overline{\mathbf{R}}_n(k) = n!^{a-6}(k + \frac{n}{2})\frac{(k-n)_n^3(k+n+1)_n^3}{(k)_{n+1}^a}$$

et

$$\overline{\mathbf{S}}_n(z) = \frac{1}{2}\sum_{k=1}^{\infty} \overline{\mathbf{R}}''_n(k) z^{-k}.$$

On développe $\overline{\mathbf{R}}_n$ en éléments simples, ce qui définit des coefficients $\overline{c}_{i,j}$ (les formules (1.20) et (1.21) restant valables). On définit $\overline{P}_1, \dots, \overline{P}_a$ à partir des $\overline{c}_{i,j}$ par la relation (1.24) ; seul $\overline{P}_0$ est défini par une formule légèrement différente :

$$\overline{P}_0(z) = -\sum_{\ell=0}^{n-1}\left(\sum_{i=\ell+1}^{n}\sum_{j=1}^{a}\frac{j(j+1)\overline{c}_{i,j}}{2(i-\ell)^{j+2}}\right)z^{\ell}.$$

On obtient la décomposition suivante exactement comme au paragraphe 2.3, mais un décalage se produit car on dérive $\overline{\mathbf{R}}_n$ (voir le paragraphe 1.4) :

$$\overline{\mathbf{S}}_n(z) = \overline{P}_0(z) + \sum_{j=1}^{a}\frac{j(j+1)}{2}\,\overline{P}_j(z)\mathrm{Li}_{j+2}(1/z).$$

Les arguments de symétrie du paragraphe 2.3 restent valables, et montrent (car a est pair) que $\overline{\mathbf{S}}_n(1)$ est une forme linéaire à coefficients rationnels en 1, $\zeta(5)$, $\zeta(7)$, ..., $\zeta(a+1)$. De plus un dénominateur commun pour ces coefficients est $2d_n^{a+2}$; on conjecture ([R6], §5.1) que $2d_n^{a+1}$

convient aussi. La majoration de ces coefficients (qui est effectuée au paragraphe 2.3) est inutile ici : elle servait à appliquer le critère de Nesterenko, dont on n'a pas besoin puisqu'on applique seulement la remarque évidente qu'une forme linéaire, à coefficients entiers, en des rationnels fixés ne peut pas être arbitrairement petite sans être nulle.

Le point délicat de la preuve est l'estimation asymptotique de $\overline{\mathbf{S}}_n(1)$. En effet, on ne connaît pas d'écriture de $\overline{\mathbf{S}}_n(1)$ comme intégrale multiple réelle. On utilise donc la méthode du col. Posons

$$K_n(u) = \frac{-1}{2i\pi}\int_{c-i\infty}^{c+i\infty} \overline{\mathbf{R}}_n(s)\left(\frac{\pi}{\sin(\pi s)}\right)^3 e^{us}\mathrm{d}s,$$

où c est un réel avec $0 < c < n+1$, et u un nombre complexe tel que $\mathrm{Re}(u) \leq 0$ et $|\mathrm{Im}(u)| < 3\pi$. Cette intégrale est à rapprocher de celle notée $I_{\mathbb{C},n}(z)$ au paragraphe 1.5. On peut appliquer le théorème des résidus, pour faire apparaître les pôles de l'intégrande qui sont situés aux entiers $n+1$, $n+2$, ...Au voisinage d'un tel entier k, on a $(\frac{\pi}{\sin(\pi s)})^3 = \frac{(-1)^k}{(s-k)^3} + \frac{(-1)^k\pi^2}{2(s-k)} + \mathrm{O}(s-k)$. On obtient donc (voir [H8] et [Z7] pour des résultats analogues) :

$$K_n(u) = \frac{\pi^2+u^2}{2}\sum_{k=n+1}^{\infty}\overline{\mathbf{R}}_n(k)(-e^u)^k + u\sum_{k=n+1}^{\infty}\overline{\mathbf{R}}'_n(k)(-e^u)^k + \frac{1}{2}\sum_{k=n+1}^{\infty}\overline{\mathbf{R}}''_n(k)(-e^u)^k.$$

En choisissant $u = i\pi$, le premier terme disparaît, et on obtient $\overline{\mathbf{S}}_n(1) = \mathrm{Re}(K_n(i\pi))$.

La méthode du col donne ([R7], Lemme 5) deux nombres complexes non nuls c_0 et α, qu'on peut calculer, tels que $K_n(i\pi) \sim c_0 n^{-8} e^{\alpha n}$ quand n tend vers l'infini. Comme la partie imaginaire de α n'est pas un multiple entier de π, il existe une suite strictement croissante $\varphi(n)$ d'entiers tels que l'argument de $c_0 e^{\alpha\varphi(n)}$, vu modulo 2π, ait une limite autre que $\pm\pi/2$. On a alors :

$$\lim_{n\to\infty} |\overline{\mathbf{S}}_{\varphi(n)}(1)|^{1/\varphi(n)} = e^{\mathrm{Re}(\alpha)}.$$

Le choix $a = 20$ donne $\mathrm{Re}(\alpha) = -22{,}02\ldots$ d'où $\mathrm{Re}(\alpha) + a + 2 < 0$. Donc la forme linéaire $d_{\varphi(n)}^{22}\overline{\mathbf{S}}_{\varphi(n)}(1)$ en 1, $\zeta(5)$, $\zeta(7)$, ..., $\zeta(21)$, à coefficients entiers, tend vers 0 quand n tend vers l'infini et est non nulle pour n assez grand. Cela montre que l'un au moins parmi $\zeta(5)$, $\zeta(7)$, ..., $\zeta(21)$ est irrationnel.

Remarque 3.3 *Si on savait démontrer la conjecture mentionnée ci-dessus (i.e. $2d_n^{a+1}\overline{P}_j(1)$ est un entier pour tout j), on pourrait ([R6], §5.1) appliquer la même méthode avec $a = 18$, et démontrer ainsi que l'un au moins des nombres $\zeta(5)$, $\zeta(7)$, ..., $\zeta(19)$, est irrationnel.*

3.3 Irrationalité d'un nombre parmi $\zeta(5)$, $\zeta(7)$, $\zeta(9)$ et $\zeta(11)$

La structure de la preuve est la même que dans le paragraphe précédent. La différence principale vient de dénominateurs nettement plus petits, grâce à une étude fine de leurs valuations p-adiques et à l'utilisation d'une fraction rationnelle modifiée :

$$\widetilde{\mathbf{R}}_n(k) = \frac{\prod_{u=1}^{10}((13+2u)n)!}{(27n)!^6}(37n+2k)\frac{(k-27n)_{27n}^3(k+37n+1)_{27n}^3}{\prod_{u=1}^{10}(k+(12-u)n)_{(13+2u)n+1}}.$$

Pour $|z| \geq 1$ on pose $\widetilde{\mathbf{S}}_n(z) = \frac{1}{2}\sum_{k=1}^{\infty}\widetilde{\mathbf{R}}''_n(k)z^{-k}$. La décomposition en éléments simples $\widetilde{\mathbf{R}}_n(k) = \sum_{j=1}^{10}\sum_{i=(j+1)n}^{(36-j)n}\frac{\tilde{c}_{i,j}}{(k+i)^j}$ définit les $\tilde{c}_{i,j}$ à partir desquels on construit les polynômes

$\widetilde{P}_j(z) = \sum_{i=(j+1)n}^{(36-j)n} \tilde{c}_{i,j} z^i$ pour $j \in \{1, 2, \dots, 10\}$ et

$$\widetilde{P}_0(z) = -\sum_{\ell=0}^{35n-1} \left(\sum_{j=1}^{10} \sum_{i=\max((j+1)n, \ell+1)}^{(36-j)n} \frac{j(j+1)\tilde{c}_{i,j}}{2(i-\ell)^{j+2}} \right) z^\ell.$$

On a alors $\widetilde{\mathbf{S}}_n(z) = \widetilde{P}_0(z) + \sum_{j=1}^{10} \frac{j(j+1)}{2} \widetilde{P}_j(z) \mathrm{Li}_{j+2}(1/z)$.

Le problème est de majorer de façon très précise le dénominateur des rationnels $\tilde{c}_{i,j}$. En suivant la méthode utilisée pour démontrer le lemme 2.12, on obtiendrait $d_{33n}^{10-j} \tilde{c}_{i,j} \in \mathbb{Z}$ pour tous i et j. Une étude fine de la valuation p-adique des coefficients binomiaux permet d'obtenir un dénominateur nettement plus petit : on trouve un entier Φ_n "assez grand" tel que $d_{33n}^{10-j} \Phi_n^{-1} \tilde{c}_{i,j} \in \mathbb{Z}$. On en déduit directement que $2 d_{35n}^3 d_{34n} d_{33n}^8 \Phi_n^{-1} \widetilde{P}_j(z)$ est à coefficients entiers pour tout $j \in \{0, 1, \dots, 10\}$.

La symétrie $\widetilde{\mathbf{R}}_n(-37n-k) = -\widetilde{\mathbf{R}}_n(k)$ donne $z^{37n} \widetilde{P}_j(1/z) = (-1)^{j+1} \widetilde{P}_j(z)$, d'où $\widetilde{P}_j(1) = 0$ pour $j = 2, 4, \dots, 10$. En outre on a $\widetilde{P}_1(1) = 0$ car $\widetilde{\mathbf{R}}_n(k) = \mathrm{O}(k^{-2})$ quand k tend vers l'infini. Donc $\widetilde{\mathbf{S}}_n(1)$ est une forme linéaire en 1, $\zeta(5)$, $\zeta(7)$, $\zeta(9)$ et $\zeta(11)$. Pour l'estimer, et démontrer qu'elle est non nulle pour une infinité de n, on transforme $\widetilde{\mathbf{S}}_n(1)$ en une intégrale complexe, à laquelle on applique la méthode du col (voir [Z7], §2). On obtient les comportements asymptotiques suivants quand n tend vers l'infini : $\limsup |\widetilde{\mathbf{S}}_n(1)|^{1/n} \leq e^{-227,58\dots}$, $\limsup |\Phi_n^{-1}|^{1/n} \leq e^{-176,75\dots}$ et $(d_{35n}^3 d_{34n} d_{33n}^8)^{1/n} \to e^{403}$. Comme $403 < 227,58 + 176,75$ on obtient la conclusion cherchée.

Remarque 3.4 *Zudilin conjecture ([Z9], §9) que des compensations ont lieu quand $z = 1$, ce qui permettrait de trouver un dénominateur plus petit pour les $P_j(1)$. Peut-être pourrait-on alors démontrer que parmi $\zeta(5)$, $\zeta(7)$ et $\zeta(9)$ l'un au moins est irrationnel.*

Remarque 3.5 *En utilisant des méthodes similaires, on peut démontrer [Z7] que pour tout $\ell \geq 1$ impair l'un au moins des nombres $\zeta(\ell+2)$, $\zeta(\ell+4)$, $\dots$, $\zeta(8\ell-1)$, est irrationnel.*

Publications Mathématiques de l'UFR Sciences et Techniques de Besançon
Théorie des Nombres, Colloque Jeunes Chercheurs
Mars 2002, Besançon

Chapitre 2

Mesures d'irrationalité de $\zeta(2)$ et $\zeta(3)$

1 Résultats qualitatifs

Cet exposé est consacré à l'étude des valeurs aux points entiers $s \geq 2$ de la fonction ζ de Riemann, définie par $\zeta(s) = \sum_{n \geq 1} n^{-s}$. Pour s pair, on a $\zeta(s) = c_s \pi^s$, où les c_s sont des nombres rationnels liés aux nombres de Bernoulli. Cette formule était connue d'Euler. Or Lindemann a démontré en 1882 que π est transcendant (voir l'appendice de [L2] pour une démonstration). Donc $\zeta(s)$ est transcendant pour tout entier s pair.

En ce qui concerne les entiers s impairs, on ne connaît pas d'analogue de la formule d'Euler, et on conjecture qu'il n'y en pas (pour un énoncé précis, voir la conjecture 1.4 ci-dessous). La nature arithmétique des valeurs de ζ aux entiers impairs est beaucoup plus difficile à déterminer. Le premier résultat dans cette direction date de 1978 :

Théorème 1.1 (Apéry, 1978) *$\zeta(3)$ est irrationnel.*

Apéry a annoncé ce résultat [A10] lors des Journées Arithmétiques de Luminy. Les détails de la preuve (qui sont loin d'être triviaux) ont été publiés par Van Der Poorten [P4], grâce à des contributions de Cohen et Zagier. Par la suite, plusieurs autres démonstrations du théorème d'Apéry sont parues, en particulier celle de Beukers [B4]. Des versions quantitatives ont également été obtenues ; c'est l'objet des paragraphes suivants.

La grande percée suivante est due à Rivoal ([B2], [R5]) :

Théorème 1.2 (Rivoal, 2000) *Il existe une infinité d'entiers s impairs tels que $\zeta(s)$ soit irrationnel.*

Le théorème de Rivoal est en fait un peu plus précis : le $\mathbb{Q}$-espace vectoriel engendré par 1 et les $\zeta(s)$ pour s impair est de dimension infinie. Cela signifie qu'on peut trouver une suite infinie strictement croissante $s_1, s_2, \ldots$ d'entiers impairs tels que 1, $\zeta(s_1)$, $\zeta(s_2)$, $\ldots$soient linéairement indépendants sur $\mathbb{Q}$. On peut donner des versions quantitatives de ce théorème :

Rivoal a montré [R7] que parmi les neuf nombres $\zeta(5)$, $\zeta(7)$, ..., $\zeta(21)$ l'un au moins est irrationnel. Ce résultat a été amélioré ensuite [Z6] :

Théorème 1.3 (Zudilin, 2001) *L'un au moins des quatre nombres $\zeta(5)$, $\zeta(7)$, $\zeta(9)$, $\zeta(11)$ est irrationnel.*

Malgré ces développements récents, il n'existe aucun entier $s \geq 5$ impair pour lequel on sache si $\zeta(s)$ est rationnel ou non. On est encore bien loin de ce qu'on espère :

Conjecture 1.4 *Les nombres π, $\zeta(3)$, $\zeta(5)$, $\zeta(7)$, ...sont algébriquement indépendants sur $\mathbb{Q}$. En particulier ils sont tous transcendants, donc irrationnels.*

Cette conjecture est un cas particulier de la conjecture diophantienne énoncée dans [W7] en termes de polyzêtas.

2 Exposant d'irrationalité

Dans tout ce paragraphe, on désigne par α un réel irrationnel. On s'intéresse à la précision à laquelle α peut être approché par des rationnels.

Définition 2.1 *On dit qu'un réel positif ν est* un exposant d'irrationalité *de α si, pour tout $\varepsilon > 0$, il n'y a qu'un nombre fini de rationnels p/q tels que*

$$|\alpha - \frac{p}{q}| < \frac{1}{q^{\nu+\varepsilon}}.$$

La théorie des fractions continues ([H2], §11.1), ou le principe des tiroirs de Dirichlet ([H2], §11.3), montre qu'il existe une infinité de rationnels p/q tels que $|\alpha - \frac{p}{q}| < \frac{1}{q^2}$. Donc un exposant d'irrationalité est toujours supérieur ou égal à 2.

Il est clair que si ν est un exposant d'irrationalité de α alors tout $\nu' \geq \nu$ l'est aussi. L'ensemble des exposants d'irrationalité de α est donc un intervalle fermé de $\mathbb{R}$, qui est vide ou bien de la forme $[\mu, +\infty[$.

Définition 2.2 L'exposant d'irrationalité *de α est le plus petit élément de l'ensemble des exposants d'irrationalité de α ; on le note $\mu(\alpha)$. Si cet ensemble est vide, on pose $\mu(\alpha) = +\infty$.*

On a toujours $\mu(\alpha) \geq 2$. Si α est un nombre algébrique (toujours supposé irrationnel), Liouville a démontré ([L5] ; voir aussi [H2], §11.7) qu'il existe une constante $c(\alpha)$ telle que, pour tout rationnel p/q, on ait $|\alpha - \frac{p}{q}| > \frac{c(\alpha)}{q^{\deg(\alpha)}}$. Cela implique $\mu(\alpha) \leq \deg(\alpha)$. Ce résultat a été amélioré par Roth en 1955 : il a démontré qu'on a $\mu(\alpha) = 2$ pour tout nombre algébrique (irrationnel) α.

En ce qui concerne les nombres transcendants, la situation est moins uniforme. Posons $\lambda = \sum_{n\geq 1} \frac{1}{10^{n!}}$. Ce nombre, considéré par Liouville [L5], possède des approximations rationnelles extrêmement bonnes (obtenues en tronquant la série) ; on en déduit facilement qu'il vérifie $\mu(\lambda) = +\infty$. C'est d'ailleurs de cette manière que Liouville a montré que λ est transcendant. Une variante de cette construction permet, pour tout réel $\mu_0 > 2$, de construire un réel λ_{μ_0} (nécessairement transcendant, d'après le théorème de Roth) tel que $\mu(\lambda_{\mu_0}) = \mu_0$.

Il faut noter que les approximations rationnelles qui interviennent dans la définition de $\mu(\alpha)$ sont toujours des convergents du développement en fraction continue (voir [H2], §10.15, théorème 184). Quand on connaît explicitement le développement de α en fraction continue, on peut espérer mesurer avec précision la qualité des meilleures approximations rationnelles de α. C'est le cas pour $\alpha = e$: Davis a démontré (voir [F1], §5.1) que l'équation $|e - \frac{p}{q}| < c\frac{\log(\log(q))}{q^2 \log(q)}$ a une infinité de solutions si $c > \frac{1}{2}$, mais seulement un nombre fini si $c < \frac{1}{2}$. On en déduit immédiatement que $\mu(e) = 2$.

Les exemples donnés jusqu'à présent montrent que l'exposant d'irrationalité mesure très grossièrement la qualité des approximations rationnelles de α. Presque tout nombre α (au sens de la mesure de Lebesgue) vérifie $\mu(\alpha) = 2$ (voir [K1], théorème 32). On conjecture que toute période α (au sens de [K3]) vérifie $\mu(\alpha) = 2$; ceci s'applique en particulier à tous les $\zeta(s)$ pour $s \geq 2$ entier. Les seuls nombres α explicites pour lesquels on sait que $\mu(\alpha) > 2$ sont des nombres construits à cet effet.

Très souvent, lorsqu'on démontre *directement* qu'un réel α est irrationnel, on peut rendre cette preuve quantitative et obtenir une majoration de $\mu(\alpha)$. Par exemple, la preuve d'Apéry donne $\mu(\zeta(3)) \leq 13,41782$. En revanche, bien qu'on sache montrer que e^π est transcendant, on ne sait pas si son exposant d'irrationalité est fini. Pour plus de détails, on pourra consulter [W6].

La méthode d'Apéry s'applique aussi pour démontrer l'irrationalité de $\zeta(2)$. Ce résultat n'est pas nouveau (puisque $\zeta(2) = \frac{\pi^2}{6}$), mais il permet d'obtenir la majoration $\mu(\zeta(2)) \leq 11,85078$. Cette majoration, ainsi que celle pour $\zeta(3)$, a été améliorée successivement par plusieurs auteurs ; le meilleur résultat connu à ce jour est dû à Rhin-Viola, en 1996 [R3] pour $\zeta(2)$ et en 2001 [R4] pour $\zeta(3)$:

Théorème 2.3 (Rhin-Viola) *On a* $\mu(\zeta(2)) \leq 5,441243$ *et* $\mu(\zeta(3)) \leq 5,513891$.

3 Groupe de Rhin-Viola

Dans ce paragraphe, on définit la structure de groupe qui est centrale dans la preuve du théorème 2.3. Tous les détails sur cette structure, et sur l'ensemble de la méthode mise en jeu, se trouvent dans [R3] pour le cas de $\zeta(2)$. Le cas de $\zeta(3)$ n'est pas étudié dans la suite de ce texte, car il est tout à fait analogue à celui de $\zeta(2)$, mais plus long à présenter. Le lecteur intéressé pourra consulter [R4].

Soient h, i, j, k, l cinq entiers naturels. Rhin et Viola considèrent l'intégrale suivante :

$$I(h,i,j,k,l) = \int_0^1 \int_0^1 \frac{x^h(1-x)^i y^k (1-y)^j}{(1-xy)^{i+j-l}} \frac{\mathrm{d}x\,\mathrm{d}y}{1-xy}.$$

C'est une intégrale convergente. Quand tous les paramètres sont égaux à un même entier N, on retrouve les intégrales introduites par Beukers [B4] pour donner une nouvelle preuve de l'irrationalité de $\zeta(2)$. Beukers a démontré que dans ce cas, l'intégrale s'écrit $u_N\zeta(2)+v_N$, avec u_N entier et v_N rationnel. Rhin et Viola montrent que l'intégrale $I(h,i,j,k,l)$ s'écrit toujours sous la forme $u\zeta(2)+v$, avec u entier et v rationnel qui dépendent des paramètres h,i,j,k,l. Ils obtiennent (par des méthodes classiques) une certaine majoration du dénominateur de v, et utilisent un groupe de transformations pour améliorer cette majoration. C'est cette structure de groupe, et le gain arithmétique qu'ils en déduisent, qui constitue le cœur de leur méthode.

Pour définir cette action de groupe, on travaille formellement (c'est-à-dire qu'on considère h, i, j, k, l comme des indéterminées) et on définit $\alpha_1, \ldots, \alpha_5$ par les formules suivantes :

$$\begin{aligned}
\alpha_1 &= h+i \\
\alpha_2 &= i+j \\
\alpha_3 &= j+k \\
\alpha_4 &= k+l \\
\alpha_5 &= l+h
\end{aligned}$$

Le changement de variables réciproque est donné par :

$$\begin{aligned}
h &= \frac{1}{2}(\alpha_5+\alpha_1+\alpha_3-\alpha_2-\alpha_4) \\
i &= \frac{1}{2}(\alpha_1+\alpha_2+\alpha_4-\alpha_3-\alpha_5) \\
j &= \frac{1}{2}(\alpha_2+\alpha_3+\alpha_5-\alpha_4-\alpha_1) \\
k &= \frac{1}{2}(\alpha_3+\alpha_4+\alpha_1-\alpha_5-\alpha_2) \\
l &= \frac{1}{2}(\alpha_4+\alpha_5+\alpha_2-\alpha_1-\alpha_3)
\end{aligned}$$

On considère l'action naturelle de $\mathfrak{S}_5$ sur $\{\alpha_1, \ldots, \alpha_5\}$ (par permutation des indices). On l'étend par linéarité : par exemple, pour $\gamma \in \mathfrak{S}_5$ on a $\gamma(h) = \frac{1}{2}(\alpha_{\gamma(5)}+\alpha_{\gamma(1)}+\alpha_{\gamma(3)}-\alpha_{\gamma(2)}-\alpha_{\gamma(4)})$. Si, dans le pentagone de sommets $1, 2, 3, 4, 5$, les sommets $\gamma(2)$ et $\gamma(4)$ ne sont pas consécutifs, alors $\gamma(h)$ est l'un parmi h, i, j, k, l. Sinon, on voit facilement qu'on obtient l'une des cinq valeurs suivantes :

$$\begin{aligned}
h+i-k &= \frac{1}{2}(\alpha_5+\alpha_1+\alpha_2-\alpha_3-\alpha_4) \\
i+j-l &= \frac{1}{2}(\alpha_1+\alpha_2+\alpha_3-\alpha_4-\alpha_5) \\
j+k-h &= \frac{1}{2}(\alpha_2+\alpha_3+\alpha_4-\alpha_5-\alpha_1) \\
k+l-i &= \frac{1}{2}(\alpha_3+\alpha_4+\alpha_5-\alpha_1-\alpha_2) \\
l+h-j &= \frac{1}{2}(\alpha_4+\alpha_5+\alpha_1-\alpha_2-\alpha_3)
\end{aligned}$$

Posons $S = \{h, i, j, k, l, h+i-k, i+j-l, j+k-h, k+l-i, l+h-j\}$. Le groupe $\mathfrak{S}_5$ agit transitivement sur S. Sous l'action du groupe diédral $\mathcal{D}_5$, l'ensemble S se décompose en deux orbites : d'une part $\{h, i, j, k, l\}$, qui correspondent (via la position des signes moins dans les formules ci-dessus) aux paires de sommets non consécutifs du pentagone (c'est-à-dire aux diagonales), et d'autre part $\{h+i-k, i+j-l, j+k-h, k+l-i, l+h-j\}$, qui correspondent aux paires de sommets consécutifs, i.e. aux côtés du pentagone. L'action de $\mathcal{D}_5$ sur $\{h, i, j, k, l\}$ est la même que sur un pentagone de sommets h, i, j, k, l ; mais l'action de $\mathfrak{S}_5$ ne stabilise pas $\{h, i, j, k, l\}$.

Un point crucial de la méthode de la Rhin-Viola pour $\zeta(2)$ est le théorème suivant, qu'ils démontrent à l'aide de deux changements de variables et d'une identité intégrale liée à la fonction hypergéométrique de Gauss :

Théorème 3.1 *Pour l'action de $\mathfrak{S}_5$ sur S définie ci-dessus :*
– *L'intégrale $I(h,i,j,k,l)$ est invariante par $\mathcal{D}_5$.*
– *Le quotient $\frac{I(h,i,j,k,l)}{h!i!j!k!l!}$ est invariant par $\mathfrak{S}_5$.*

C'est en exploitant les valuations p-adiques du produit de factorielles $h!i!j!k!l!$, et de ses images sous l'action de $\mathfrak{S}_5$, que Rhin et Viola améliorent sensiblement leur majoration du dénominateur de v.

4 Interprétation géométrique

Dans ce paragraphe, on donne une interprétation du théorème 3.1 en termes d'automorphismes. Davantage de détails, ainsi que l'interprétation analogue pour $\zeta(3)$, se trouvent dans [F3].

Notons $\mathcal{V}$ la sous-variété algébrique affine de $\mathbb{R}^5$ définie par les équations suivantes :

$$\begin{aligned} x_1x_2(1+x_5x_3) &= 1\\ x_2x_3(1+x_1x_4) &= 1\\ x_3x_4(1+x_2x_5) &= 1\\ x_4x_5(1+x_3x_1) &= 1\\ x_5x_1(1+x_4x_2) &= 1 \end{aligned}$$

On note Ω l'ensemble des points de $\mathcal{V}$ à coordonnées strictement positives, et ω la 2-forme différentielle $\mathrm{d}^\times(x_{i-1}x_i)\wedge\mathrm{d}^\times(x_ix_{i+1})$ sur $\mathcal{V}$, avec $\mathrm{d}^\times f = \frac{\mathrm{d}f}{f}$ pour toute fonction f. On montre facilement que $\mathcal{V}$ est de dimension 2, et que ω ne dépend pas du choix de l'indice $i\in\{1,\dots,5\}$ (en posant $x_0 = x_5$ et $x_6 = x_1$). On peut paramétrer bijectivement Ω par $]0,1[^2$, ce qui permet de faire le lien avec les intégrales du paragraphe 3 (dont on conserve les notations) :

$$I(h,i,j,k,l) = \int_\Omega x_1^{\alpha_1}x_2^{\alpha_2}x_3^{\alpha_3}x_4^{\alpha_4}x_5^{\alpha_5}\omega.$$

On peut alors interpréter comme suit la première partie du théorème 3.1 :

Théorème 4.1 *Le groupe $\mathcal{D}_5$, qui agit naturellement sur $\mathbb{R}^5$, stabilise $\mathcal{V}$, Ω et ω (au signe près).*

Ce théorème n'est pas difficile à démontrer (par exemple, on voit immédiatement que les équations de $\mathcal{V}$ sont globalement stables par $\mathcal{D}_5$). Il suffit alors, pour chaque $\gamma\in\mathcal{D}_5$, d'effectuer un changement de variables pour en déduire la première assertion du théorème 3.1.

Notons U l'ensemble des $(t_1,\dots,t_5)\in]0,1[^5$ tels que $t_1+\ldots+t_5 = 1$, et η la 4-forme différentielle $\mathrm{d}t_1\wedge\mathrm{d}t_2\wedge\mathrm{d}t_3\wedge\mathrm{d}t_4$ sur U. Pour tout $i\in\{1,\dots,5\}$ on définit une fonction z_i de $\Omega\times U$ dans $\mathbb{R}$ par la formule suivante (dans laquelle les indices sont pris modulo 5) :

$$z_i(x_1,\dots,x_5,t_1,\dots,t_5) = x_i\left(\frac{t_{i-1}t_it_{i+1}}{t_{i-2}t_{i+2}}\right)^{\frac{1}{2}}.$$

On montre, à l'aide de la représentation intégrale de la fonction Γ d'Euler, la formule suivante qui fait le lien avec le paragraphe 3 :

$$\frac{I(h,i,j,k,l)}{h!i!j!k!l!} = \frac{\prod_{n\in S} n!}{(\alpha_1+\ldots+\alpha_5+4)!}\int_{\Omega\times U} z_1^{\alpha_1}z_2^{\alpha_2}z_3^{\alpha_3}z_4^{\alpha_4}z_5^{\alpha_5}\omega\wedge\eta. \qquad (2.1)$$

Sous l'action de $\mathfrak{S}_5$, le produit $\prod_{n \in S} n!$ est invariant car $\mathfrak{S}_5$ agit sur S. Le dénominateur $(\alpha_1 + \ldots + \alpha_5 + 4)!$ étant lui aussi invariant, pour démontrer la deuxième assertion du théorème 3.1 il suffit de montrer que l'intégrale qui figure au membre de droite de (2.1) est invariante par $\mathfrak{S}_5$. Ceci provient du théorème suivant (dans lequel les automorphismes sont des fonctions algébriques) :

Théorème 4.2 *Il existe un groupe H d'automorphismes de $\Omega \times U$ (qui fixent, au signe près, $\omega \wedge \eta$) et un homomorphisme de groupes surjectif $\pi : H \to \mathfrak{S}_5$ tels que, pour tout $f \in H$ et tout $i \in \{1, \ldots, 5\}$, on ait :*

$$z_i \circ f = z_{\pi(f)^{-1}(i)}.$$

C. R. Acad. Sci. Paris, Ser. I 335 (2002), 1-4

Chapitre 3

Formes linéaires en polyzêtas et intégrales multiples

Résumé. Le problème considéré ici est de définir des familles d'intégrales n-uples, munies d'une action de groupe comme dans les travaux de Rhin-Viola ([R3], [R4]), dont les valeurs soient des formes linéaires, sur le corps des rationnels, en les polyzêtas de poids au plus n. On généralise pour cela les approches de Vasilyev [V3] et Sorokin [S6], en les reliant par un changement de variables. On décrit aussi une structure de groupe pour une intégrale n-uple qui donne, pour $n = 2$ et $n = 3$, celles obtenues par Rhin et Viola.

Abstract. *The problem we consider is to define families of n-dimensional integrals, endowed with group actions as in Rhin-Viola's work ([R3], [R4]), the values of which are linear forms, over the rationals, in multiple zeta values of weight at most n. We generalize Vasilyev's [V3] and Sorokin's [S6] approaches, and give a change of variables that connects them to each other. We describe a group structure for a n-dimensional integral that specializes, for $n = 2$ and $n = 3$, to the ones obtained by Rhin and Viola.*

1 Introduction

Après la démonstration de l'irrationalité de $\zeta(3)$ par Apéry [A10], plusieurs variantes ont été proposées, parmi lesquelles celles de Beukers [B4] et Sorokin [S7]. Dans ces deux preuves, les formes linéaires en 1 et $\zeta(3)$ sont écrites comme des intégrales triples :

$$B(N) = \int_{[0,1]^3} \frac{x^N(1-x)^N y^N(1-y)^N z^N(1-z)^N}{(1-z(1-y(1-x)))^{N+1}} \mathrm{d}x\mathrm{d}y\mathrm{d}z$$

pour Beukers, et

$$S(N) = \int_{[0,1]^3} \frac{x^N(1-x)^N y^N(1-y)^N z^N(1-z)^N}{(1-xy)^{N+1}(1-xyz)^{N+1}} \mathrm{d}x\mathrm{d}y\mathrm{d}z$$

pour Sorokin. Or on constate que ces formes linéaires coïncident. Si on croit à la philosophie des périodes [K3], l'égalité de ces intégrales doit pouvoir se démontrer par une suite de changements de variables et d'applications des règles d'additivité (par rapport à l'intégrande ou au domaine) et du théorème de Stokes. En l'occurrence, un seul changement de variables suffit (voir le corollaire 2.2).

Les intégrales $B(N)$ ont été généralisées par Vasilyev [V2], qui pose $\delta_k(x_1, \ldots, x_n) = 1 - x_k\delta_{k-1}(x_1, \ldots, x_n)$ pour $n \geq 2$ et $k \in \{1, \ldots, n\}$ avec $\delta_0 = 1$, et considère

$$\int_{[0,1]^n} \frac{\prod_{k=1}^n x_k^N (1-x_k)^N}{\delta_n(x_1, \ldots, x_n)^{N+1}} \mathrm{d}x_1 \ldots \mathrm{d}x_n. \tag{3.1}$$

Il démontre que pour $N = 0$ cette intégrale vaut $2(-1)^{n-1}\mathrm{Li}_n((-1)^{n-1})$, donc est un multiple rationnel de $\zeta(n)$, et [V3] que pour $n = 5$ (respectivement $n = 4$) et N quelconque c'est une forme linéaire en 1, $\zeta(3)$ et $\zeta(5)$ (respectivement 1, $\zeta(2)$ et $\zeta(4)$). Il conjecture que pour tous n et N on obtient une forme linéaire en 1 et les valeurs de ζ aux entiers compris entre 2 et n ayant la même parité que n.

D'autre part, les intégrales $S(N)$ sont à rapprocher de celles que Sorokin introduit [S6] en relation avec $\zeta(2, 2, \ldots, 2)$, quand $n = 2r$ est pair :

$$\int_{[0,1]^n} \frac{(y_1y_2)^{rN+(r-1)}(y_3y_4)^{(r-1)N+(r-2)} \ldots (y_{n-1}y_n)^N \prod_{k=1}^n (1-y_k)^N}{\prod_{k\in\{2,\ldots,n\}\mathrm{pair}} (1 - y_1y_2 \ldots y_k)^{N+1}} \mathrm{d}y_1 \ldots \mathrm{d}y_n. \tag{3.2}$$

Au paragraphe 2 ci-dessous, on définit deux familles d'intégrales, qui généralisent (3.1), donc $B(N)$, d'une part, (3.2) et $S(N)$ d'autre part, et on montre que ces deux familles se correspondent par un changement de variables. On montre en outre qu'un groupe agit sur ces intégrales, de manière analogue à ce que Rhin et Viola considèrent ([R3], [R4]) pour obtenir les meilleures mesures d'irrationalité connues pour $\zeta(2)$ et $\zeta(3)$. Au paragraphe 3, on définit une autre famille d'intégrales n-uples, qui généralise (3.1) et sur laquelle agit aussi un groupe ; on retrouve alors dans les cas particuliers $n = 2$ et $n = 3$ les groupes obtenus par Rhin et Viola.

On adopte la définition suivante :

Définition 1.1 *On dit qu'une famille de nombres $I(\underline{p}) \in \mathbb{R}_+^* \cup \{\infty\}$, paramétrés par $\underline{p} \in \mathbb{Z}^s$, admet pour groupe de Rhin-Viola un sous-groupe G de $\mathrm{GL}_s(\mathbb{Z})$ si $I(g\underline{p})/I(\underline{p})$ est un rationnel (ou ∞) pour tous $g \in G$ et $\underline{p} \in \mathbb{Z}^s$ (avec $\infty/\infty = 1$).*

Dans ce texte, on considère des familles d'intégrales n-uples dont les valeurs (lorsqu'elles sont finies) sont conjecturalement des formes linéaires sur $\mathbb{Q}$ en les polyzêtas de poids au plus n. On peut espérer que, pour de bons choix des exposants $\underline{p}$, ces intégrales soient assez petites et qu'on ait un contrôle sur le dénominateur des coefficients de la forme linéaire. Alors l'étude de la valuation p-adique des nombres $I(g\underline{p})/I(\underline{p})$, quand g parcourt G et p un certain ensemble de nombres premiers, peut permettre d'améliorer ce contrôle du dénominateur, donc de raffiner des mesures d'irrationalité ou d'indépendance linéaire de certains polyzêtas.

Davantage de détails, en particulier sur l'action des groupes de Rhin-Viola, seront donnés dans [F3].

2 Une généralisation commune des intégrales de Vasilyev et de Sorokin

Dans tout ce texte, n désigne un entier supérieur ou égal à 2.

Pour $k \in \{0, \ldots, n\}$ et $\underline{x} = (x_1, \ldots, x_n) \in [0,1]^n$ on pose

$$D_k(\underline{x}) = \sum_{j=0}^{k} (-1)^j x_n x_{n-1} \ldots x_{n-j+1}.$$

On a alors $D_0(\underline{x}) = 1$, $D_1(\underline{x}) = 1 - x_n$, $D_2(\underline{x}) = 1 - x_n(1 - x_{n-1})$ et $D_n(\underline{x}) = \delta_n(\underline{x})$. À tout $\underline{p} = (a_1, \ldots, a_n, b_1, \ldots, b_n, c_2, \ldots, c_n) \in \mathbb{Z}^{3n-1}$ on associe l'intégrale (éventuellement infinie)

$$J(\underline{p}) = \int_{[0,1]^n} \frac{\prod_{k=1}^n x_k^{a_k}(1-x_k)^{b_k}}{\prod_{k=2}^n D_k(\underline{x})^{c_k}} \frac{\prod_{k \in \{2,\ldots,n-2\} \text{pair}} D_k(\underline{x})}{\prod_{k \in \{3,\ldots,n-1\} \text{impair}} D_k(\underline{x})} \frac{\mathrm{d}x_1 \ldots \mathrm{d}x_n}{D_n(\underline{x})}.$$

Par ailleurs, à tout $\underline{P} = (A_1, \ldots, A_n, B_1, \ldots, B_n, C_2, \ldots, C_n) \in \mathbb{Z}^{3n-1}$ on associe

$$K(\underline{P}) = \int_{[0,1]^n} \frac{\prod_{k=1}^n y_k^{A_k}(1-y_k)^{B_k}}{\prod_{k=2}^n (1 - y_1 y_2 \ldots y_k)^{C_k+1}} \mathrm{d}y_1 \ldots \mathrm{d}y_n.$$

Cette intégrale est finie si, et seulement si[1], on a $A_k \geq 0$, $B_k \geq 0$ et $\sum_{j=2}^k C_j \leq \sum_{j=1}^k B_j$ pour tout $k \in \{1, \ldots, n\}$. L'intérêt des intégrales $K(\underline{P})$ (donc du théorème suivant) est qu'elles se développent "naturellement" en séries multiples, dont on peut espérer démontrer que ce sont des formes linéaires sur $\mathbb{Q}$ en les polyzêtas de poids au plus n.

Théorème 2.1 *Pour tout $\underline{p} \in \mathbb{Z}^{3n-1}$ on a $J(\underline{p}) = K(\underline{P})$, où $\underline{P}$ est donné en fonction de $\underline{p}$ par :*

$A_k = a_{n+1-k}$ *pour* $1 \leq k \leq n$,
$B_k = b_{n+1-k}$ *pour* $2 \leq k \leq n$ *et* $B_1 = a_{n-1} + b_n - c_2 - c_3 - \ldots - c_n$,
$C_k = a_{n+1-k} + b_{n+1-k} - c_k - c_{k+1} - \ldots - c_n$ *pour tout* $k \in \{2, \ldots, n\}$ *pair,*
$C_k = c_k + c_{k+1} + \ldots + c_n - a_{n-k}$ *pour tout* $k \in \{3, \ldots, n\}$ *impair, avec la convention* $a_0 = 0$.

Ce résultat provient du changement de variables défini par $x_k = y_{n+1-k}$ pour $k \equiv n \mod 2$ et $x_k = \frac{(1-y_1\ldots y_{n-k})y_{n+1-k}}{1-y_1\ldots y_{n+1-k}}$ pour $k \not\equiv n \mod 2$. On peut bien sûr l'inverser, donc exprimer $\underline{p}$ en fonction de $\underline{P}$.

Corollaire 2.2 *Avec les notations de l'introduction, on a $B(N) = S(N)$ pour tout $N \geq 0$, et l'intégrale* (3.1) *est égale à* $\int_{[0,1]^n} \frac{\prod_{k=1}^n y_k^N(1-y_k)^N}{\prod_{k\in\{2,\ldots,n\} pair}(1-y_1\ldots y_k)^{N+1}} \mathrm{d}y_1 \ldots \mathrm{d}y_n$ *si n est pair, et à* $\int_{[0,1]^n} \frac{\prod_{k=1}^n y_k^N(1-y_k)^N}{(1-y_1\ldots y_n)^{N+1}\prod_{k\in\{2,\ldots,n\} pair}(1-y_1\ldots y_k)^{N+1}} \mathrm{d}y_1 \ldots \mathrm{d}y_n$ *si n est impair.*

Pour $N = 0$, l'intégrale (3.1) vaut donc

$$\sum_{l_1 \geq l_2 \geq \ldots \geq l_{n/2} \geq 1} \frac{1}{l_1^2 l_2^2 \ldots l_{n/2}^2}$$

si n est pair, et

$$\sum_{l_1 \geq l_2 \geq \ldots \geq l_{(n-1)/2} \geq l_{(n+1)/2} \geq 1} \frac{1}{l_1^2 l_2^2 \ldots l_{(n-1)/2}^2 l_{(n+1)/2}}$$

si n est impair. Le résultat de Vasilyev selon lequel cette intégrale égale $2(-1)^{n-1}\mathrm{Li}_n((-1)^{n-1})$ se ramène ainsi à une identité linéaire entre polyzêtas, qu'on doit pouvoir démontrer de manière combinatoire (voir [W7]).

[1]Noter ici l'erreur dans la version parue aux Comptes-Rendus de l'Académie des Sciences.

Proposition 2.3 *Si $n \geq 3$, la famille $(K(\underline{P}))$ admet un groupe de Rhin-Viola d'ordre 32, isomorphe à $(V \times V) \rtimes \mathbb{Z}/2\mathbb{Z}$, où $V = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ est le groupe de Klein et $\mathbb{Z}/2\mathbb{Z}$ agit en permutant les facteurs de $V \times V$. Cela provient d'analogues des transformations σ et χ de [R4], et du changement de variables défini par $y'_k = \frac{1-y_1 \dots y_{n-k+1}}{1-y_1 \dots y_{n-k+2}}$ pour tout $k \in \{1, \dots, n\}$ (avec la convention $y_{n+1} = 0$).*

3 Une généralisation du groupe de Rhin-Viola

À tout $\underline{p} = (a_1, \dots, a_n, b_1, \dots, b_n, c_2, \dots, c_n) \in \mathbb{Z}^{3n-1}$ on associe (avec la notation δ_k utilisée dans l'introduction) :

$$L(\underline{p}) = \int_{[0,1]^n} \frac{\prod_{k=1}^n x_k^{a_k}(1-x_k)^{b_k}}{\prod_{k=2}^n \delta_k(\underline{x})^{c_k}} \frac{\mathrm{d}x_1 \dots \mathrm{d}x_n}{\delta_n(\underline{x})}.$$

On pose $\varrho_n = c_n - b_n$ et $\varrho_{n-1} = c_{n-1} - 1 - b_{n-1}$, puis $\varrho_k = \varrho_{k+2}^+ + c_k - 1 - b_k$ pour tout $k \in \{1, \dots, n-2\}$ avec la convention $c_1 = 1$ (en notant $\alpha^+ = \max(\alpha, 0)$). Alors l'intégrale $L(\underline{p})$ est finie si, et seulement si, on a $a_k \geq 0$, $b_k \geq 0$ et $\varrho_k \leq a_{k-1}$ pour tout $k \in \{1, \dots, n\}$, avec la convention $a_0 = 0$.

Notons σ l'automorphisme de $\mathbb{Z}^{3n-1}$ qui échange a_1 et b_2, ainsi que a_2 et b_1, en fixant les autres coordonnées. Notons ψ celui qui à $\underline{p}$ associe $\underline{p}' = (a'_1, \dots, a'_n, b'_1, \dots, b'_n, c'_2, \dots, c'_n)$ défini par :

$$\begin{aligned}
&a'_k = a_{n+1-k} \text{ pour } 1 \leq k \leq n \text{ , } b'_k = b_{n+2-k} \text{ pour } 2 \leq k \leq n \text{ et } b'_1 = a_{n-1} + b_n - c_n,\\
&c'_k = a_{n+2-k} + b_{n+2-k} + c_{n+1-k} - b_{n+1-k} - a_{n-k} \text{ pour } 2 \leq k \leq n-1,\\
&c'_n = a_2 + b_2 - b_1.
\end{aligned}$$

Alors des changements de variables montrent qu'on a $L(\underline{p}) = L(\sigma(\underline{p})) = L(\psi(\underline{p}))$ pour tout $\underline{p}$. En outre, notons χ l'automorphisme de $\mathbb{Z}^{3n-1}$ qui fixe toutes les composantes, sauf a_n et c_n qu'il échange et b_n qu'il remplace par $a_n + b_n - c_n$; il vérifie $L(\underline{p}) = \frac{a_n! b_n!}{c_n!(a_n+b_n-c_n)!} L(\chi(\underline{p}))$ pour tout $\underline{p} \in \mathbb{Z}^{3n-1}$ tel que a_n, b_n, c_n et $a_n + b_n - c_n$ soient positifs.

Proposition 3.1 *Si $n \geq 3$, la famille $(L(\underline{p}))$ paramétrée par $\underline{p} \in \mathbb{Z}^{3n-1}$ admet un groupe de Rhin-Viola d'ordre 32, isomorphe à $(V \times V) \rtimes \mathbb{Z}/2\mathbb{Z}$, qui est engendré par σ, ψ et χ.*

Remarque 1 Peut-être y a-t-il un lien entre les intégrales $L(\underline{p})$ et $K(\underline{P})$ qui permette d'expliquer le parallélisme entre les propositions 2.3 et 3.1 ?

Pour obtenir une structure de groupe plus riche que celle de la proposition 3.1, on peut se restreindre aux intégrales $L(\underline{p})$ telles que $c_2 = \dots = c_{n-1} = 0$. Pour conserver l'action de σ et ψ, on doit imposer en outre $a_1 + b_2 = a_3 + b_3$ si $n = 3$, et les relations suivantes si $n \geq 4$:

$$a_2 = b_1 \text{ et } b_n = c_n \text{ et } a_k + b_{k+1} = a_{k+2} + b_{k+2} \text{ pour tout } k \in \{1, \dots, n-2\}.$$

On note $\mathcal{E}$ l'ensemble des $\underline{p}$ vérifiant ces relations ; il est stable par σ et ψ. Notons φ l'automorphisme de $\mathcal{E}$ qui stabilise toutes les coordonnées, sauf a_{n-1} et c_n qu'il échange, b_{n-1} qu'il remplace par $a_{n-1} + b_{n-1} - c_n$ et b_n qu'il remplace par $a_{n-1} + b_n - c_n$. On a alors $L(\underline{p}) = \frac{a_{n-1}! b_{n-1}!}{c_n!(a_{n-1}+b_{n-1}-c_n)!} L(\varphi(\underline{p}))$ pour tout $\underline{p} \in \mathcal{E}$ tel que a_{n-1}, b_{n-1}, c_n et $a_{n-1} + b_{n-1} - c_n$ soient positifs.

Théorème 3.2 *La famille des $L(\underline{p})$, pour $\underline{p} \in \mathcal{E}$, admet un groupe de Rhin-Viola G engendré par σ, ψ et φ. Plus précisément (en se restreignant aux $\underline{p} \in \mathcal{E}$ pour lesquels les quotients ont un sens) :*

- *Pour $n \geq 4$, le groupe G est isomorphe à $(\mathfrak{S}_3 \times \mathfrak{S}_3) \rtimes \mathbb{Z}/2\mathbb{Z}$, donc d'ordre 72 ; il laisse stable $\frac{L(\underline{p})}{a_{n-1}!b_{n-1}!a_2!b_3!}$ si $n \geq 5$, et $\frac{L(\underline{p})}{a_3!b_3!a_2!}$ si $n = 4$.*
- *Pour $n = 3$, le groupe G est isomorphe à $H \rtimes \mathfrak{S}_5$, où H est l'hyperplan $\varepsilon_1 + \ldots + \varepsilon_5 = 0$ de $(\mathbb{Z}/2\mathbb{Z})^5$; il laisse stable $\frac{L(\underline{p})}{a_1!a_2!a_3!b_1!b_2!b_3!(a_2+b_3-c_3)!(b_1+b_3-c_3)!}$.*
- *Pour $n = 2$, le groupe G est isomorphe à $\mathfrak{S}_5$, et laisse stable $\frac{L(\underline{p})}{a_1!a_2!b_1!b_2!(a_1+b_2-c_2)!}$.*

Pour $n \in \{2, 3\}$, on retrouve exactement les situations considérées par Rhin et Viola. Pour $n = 2$, le groupe diédral $\mathcal{D}_5$ de [R3] est exactement celui engendré par ψ et σ. Pour $n = 3$, on pose $x = x_2$, $y = 1 - x_1$ et $z = x_3$; la transformation ϑ^2 de [R4] est alors $\psi \circ \sigma$, et $\vartheta \in \mathrm{Aut}(\mathcal{E})$ est donné par $(\varphi\sigma\psi\sigma\varphi\sigma)^2\varphi$.

Pour $n = 3$, un phénomène mystérieux se produit dans [R4] : lorsqu'on impose la relation $a_1 + b_2 = a_3 + b_3$, ce qui permet d'avoir une action de groupe, les intégrales obtenues sont des formes linéaires en 1 et $\zeta(3)$ seulement : $\zeta(2)$ n'apparaît plus. On peut se demander si un phénomène analogue survient pour $n \geq 4$.

Remarque 2 Dans tout ce texte, les propriétés de χ et φ proviennent, comme dans [R3] et [R4], de la formule $\int_0^1 \frac{x^a(1-x)^b}{(1+\beta x)^{c+1}} \mathrm{d}x = \frac{a!b!}{c!(a+b-c)!} \int_0^1 \frac{x^c(1-x)^{a+b-c}}{(1+\beta x)^{a+1}} \mathrm{d}x$. Un analogue de cette formule, dans lequel le dénominateur serait de la forme $(1+\beta x)^{c+1}(1+\beta' x)^{c'+1}$, permettrait d'enrichir les structures de groupe obtenues ici.

Remerciements : Je remercie Pierre Cartier et Tanguy Rivoal, dont les questions sont à l'origine de ce travail, ainsi que Jacky Cresson et Michel Waldschmidt.

Note ajoutée aux épreuves : Le corollaire 2.2 a été obtenu, indépendamment et par une méthode différente, par Zlobin [Z5].

À paraître au J. Théor. Nombres Bordeaux

Chapitre 4

Groupes de Rhin-Viola et intégrales multiples

Résumé. Ce texte donne une nouvelle présentation, et une généralisation, des groupes qui apparaissent dans les travaux de Rhin-Viola ([R3], [R4]) sur les mesures d'irrationalité de $\zeta(2)$ et $\zeta(3)$. D'une part, on interprète ces groupes comme des groupes d'automorphismes, ce qui permet de déduire chacune des relations entre intégrales utilisées par Rhin-Viola d'un changement de variables. D'autre part, on considère plusieurs familles d'intégrales n-uples, et on montre que chacune d'elles est munie d'une action de groupe comme dans les travaux de Rhin-Viola. De plus, les valeurs de ces intégrales sont (conjecturalement, pour certaines) des formes linéaires, sur le corps des rationnels, en les polyzêtas de poids au plus n. Ces familles englobent beaucoup d'intégrales qui sont apparues dans l'étude des valeurs de ζ aux entiers. On exhibe un changement de variables entre deux de ces familles, qui permet de relier les approches de Beukers, Rhin-Viola, Vasilenko, Vasilyev d'une part, Sorokin et Rivoal d'autre part.

Mots clés. Mesure d'irrationalité, mesure d'indépendance linéaire, groupe d'automorphismes, changement de variables, fonction ζ de Riemann, fonction hypergéométrique de Gauss.

Abstract. This paper gives a new presentation, and a generalization, of the group structures in Rhin-Viola's work ([R3], [R4]) on irrationality measures of $\zeta(2)$ and $\zeta(3)$. On the one hand, these groups are seen as automorphism groups, which makes it possible to prove all relations betweeen Rhin-Viola integrals using changes of variables. On the other hand, several families of n-dimensional integrals are considered, and each of them is shown to be equipped with a group action in the same fashion as in Rhin-Viola's work. Moreover, the values of these integrals are (sometimes conjecturally) linear forms, over the rationals, in multiple zeta values of weight at most n. Among these families lie many integrals that have appeared in the study of the values of ζ at integer points. A change of variables is given between two of these families, which connects the approach of Beukers, Rhin-Viola, Vasilenko, Vasilyev to that of Sorokin and Rivoal.

Key words. Irrationality measure, linear independence measure, automorphism group, change of variables, Riemann ζ function, Gauss hypergeometric function.

Les meilleures mesures d'irrationalité de $\zeta(2)$ et $\zeta(3)$ connues à ce jour sont dues à Rhin et Viola ([R3], [R4]) :

$$\mu(\zeta(2)) < 5.441243 \text{ et } \mu(\zeta(3)) < 5.513891.$$

Pour les obtenir, Rhin et Viola utilisent une famille d'intégrales doubles pour $\zeta(2)$:

$$\int_0^1 \int_0^1 \frac{x^h(1-x)^i y^k (1-y)^j}{(1-xy)^{i+j-l}} \frac{\mathrm{d}x\,\mathrm{d}y}{1-xy}, \tag{4.1}$$

et triples pour $\zeta(3)$:

$$\int_0^1 \int_0^1 \int_0^1 \frac{x^h(1-x)^l y^k(1-y)^s z^j(1-z)^q}{(1-(1-xy)z)^{q+h-r}} \frac{\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}{1-(1-xy)z} \text{ avec } j+q=l+s. \tag{4.2}$$

Le point clef de leur méthode est la présence d'un groupe de transformations sur les paramètres entiers h, i, j, k, l (respectivement h, j, k, l, m, q, r, s, avec $m = k+r-h$) qui multiplient l'intégrale par un facteur rationnel. L'étude de la valuation p-adique de ce facteur rationnel (qui est un quotient de produits de factorielles) permet d'obtenir un meilleur dénominateur pour les coeffici[ents de la forme linéaire sur $\mathbb{Q}$ en 1 et $\zeta(2)$ (respectivement $\zeta(3)$) qu'est cette intégrale.

Dans les deux premières parties de ce texte, on construit une sous-variété $\mathcal{V}$ de $\mathbb{R}^p$ (avec $p=5$ pour $\zeta(2)$ et $p=6$ pour $\zeta(3)$), un ouvert Ω de $\mathcal{V}$ (pour la topologie réelle), et une forme différentielle ω tels que les intégrales (4.1) et (4.2) s'écrivent sous la forme :

$$J(\alpha_1, \ldots, \alpha_p) = \int_\Omega x_1^{\alpha_1} \ldots x_p^{\alpha_p} \omega.$$

L'invariance de (4.1) et (4.2) par le groupe diédral $\mathcal{D}_k$ (avec $k=5$ pour $\zeta(2)$ et $k=8$ pour $\zeta(3)$) correspond alors à une action sur $\mathcal{V}$ de ce groupe. Ces constructions sont analogues à celle que Pierre Cartier avait mentionnée, dans le cas de $\zeta(2)$, à la fin de son exposé aux douzièmes Rencontres Arithmétiques de Caen (en juin 2001). La nouveauté (en plus du cas de $\zeta(3)$) est que le groupe de Rhin-Viola tout entier s'interprète comme groupe d'automorphismes de $\Omega \times U$, en posant

$$U = \{(t_1, \ldots, t_k) \in]0,1[^k, t_1 + \ldots + t_k = 1\}.$$

On démontre ainsi toutes les égalités entre intégrales à partir d'automorphismes, dans l'esprit de la "philosophie des périodes" [K3].

Un autre objectif de ce texte est d'étudier, avec autant de généralité que possible, des familles d'intégrales dont certains cas particuliers apparaissent dans les travaux récents sur l'irrationalité ou l'indépendance linéaire de valeurs de la fonction ζ aux entiers. Le point central est la recherche de changements de variables non triviaux ; ceci fournit des analogues, pour des intégrales n-uples, des situations considérées par Rhin et Viola pour $n=2$ et $n=3$. Toutes les intégrales n-uples qui apparaissent dans ce texte sont (conjecturalement, pour certaines) des formes linéaires sur $\mathbb{Q}$ en les polyzêtas de poids au plus n.

Une première direction consiste à généraliser les intégrales de Beukers [B4] pour $\zeta(3)$:

$$\int_{[0,1]^3} \frac{x^N(1-x)^N y^N(1-y)^N z^N(1-z)^N}{(1-z(1-y(1-x)))^{N+1}} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \tag{4.3}$$

Ceci a été entrepris par Vasilenko [V1] puis Vasilyev [V2] qui posent

$$\delta_n(x_1,\dots,x_n) = 1 - x_n(1 - x_{n-1}(1 - \dots (1-x_1)))$$

et considèrent, pour N entier naturel :

$$\int_{[0,1]^n} \frac{\prod_{k=1}^n x_k^N(1-x_k)^N}{\delta_n(x_1,\dots,x_n)^N} \frac{\mathrm{d}x_1\dots\mathrm{d}x_n}{\delta_n(x_1,\dots,x_n)}. \tag{4.4}$$

Vasilyev démontre que pour $N=0$ cette intégrale vaut $2(-1)^{n-1}\mathrm{Li}_n((-1)^{n-1})$, donc est un multiple rationnel de $\zeta(n)$, et [V3] que pour $n=5$ (respectivement $n=4$) et N quelconque c'est une forme linéaire en 1, $\zeta(3)$ et $\zeta(5)$ (respectivement 1, $\zeta(2)$ et $\zeta(4)$). Il conjecture que pour tous n et N on obtient une forme linéaire en 1 et les valeurs de ζ aux entiers compris entre 2 et n ayant la même parité que n ; cette conjecture vient d'être démontrée par Zudilin [Z8]. Dans la troisième partie de ce texte, on considère une famille d'intégrales qui contient à la fois (4.1), (4.2) et (4.4) :

$$I(a_1,\dots,a_n,b_1,\dots,b_n,c) = \int_{[0,1]^n} \frac{\prod_{k=1}^n x_k^{a_k}(1-x_k)^{b_k}}{\delta_n(x_1,\dots,x_n)^c} \frac{\mathrm{d}x_1\dots\mathrm{d}x_n}{\delta_n(x_1,\dots,x_n)}. \tag{4.5}$$

En supposant que les $2n+1$ exposants vérifient certaines relations linéaires, on démontre qu'un groupe de Rhin-Viola agit sur cette famille, au sens suivant.

Définition : On dit qu'une famille de nombres $I(\underline{p}) \in \mathbb{R}_+^* \cup \{\infty\}$, paramétrés par $\underline{p} \in \mathbb{Z}^s$, admet pour groupe de Rhin-Viola un sous-groupe G de $\mathrm{GL}_s(\mathbb{Z})$ si $I(g\underline{p})/I(\underline{p})$ est un rationnel (ou ∞) pour tous $g \in G$ et $\underline{p} \in \mathbb{Z}^s$ (avec $\infty/\infty = 1$).

Dans ce texte, les nombres $I(\underline{p})$ seront toujours donnés par des intégrales, et les facteurs rationnels $I(g\underline{p})/I(\underline{p})$ seront des quotients de produits de factorielles. De tels groupes sont connus depuis longtemps : ils apparaissent, par exemple, dans les travaux de Bailey, Dixon et Whipple.

Concernant la famille (4.5) (restreinte à des paramètres satisfaisant certaines relations linéaires), on exhibe deux changements de variables σ et ψ qui transforment une intégrale de cette famille en une de la même famille. On construit aussi un analogue de la transformation hypergéométrique φ de [R4], ce qui donne (pour chaque $n \geq 2$) un groupe de Rhin-Viola G, qui dans les cas particuliers $n=2$ et $n=3$ se spécialise en ceux de [R3] et [R4]. Cela permet aussi (voir le paragraphe 3.4) d'obtenir les formules (assez mystérieuses) qui définissent le changement de variables ϑ de [R4] à partir d'autres formules de changements de variables ou de transformation hypergéométrique, moins mystérieuses car généralisables à tout $n \geq 2$.

Toujours dans un souci de généralité, on considère dans la quatrième partie de ce texte la famille suivante, qui contient (4.5) et dont on montre qu'elle admet aussi un groupe de Rhin-Viola :

$$L(a_1,\dots,a_n,b_1,\dots,b_n,c_2,\dots,c_n) = \int_{[0,1]^n} \frac{\prod_{k=1}^n x_k^{a_k}(1-x_k)^{b_k}}{\prod_{k=2}^n \delta_k(x_1,\dots,x_k)^{c_k}} \frac{\mathrm{d}x_1\dots\mathrm{d}x_n}{\delta_n(x_1,\dots,x_n)}.$$

Une deuxième direction pour étudier les valeurs de ζ aux entiers (et, plus généralement, les polyzêtas) consiste à partir de la preuve de Sorokin [S7] de l'irrationalité de $\zeta(3)$, qui utilise les intégrales suivantes :

$$\int_{[0,1]^3} \frac{x^N(1-x)^N y^N(1-y)^N z^N(1-z)^N}{(1-xy)^{N+1}(1-xyz)^{N+1}} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \tag{4.6}$$

Ces intégrales sont à rapprocher de celles que Sorokin introduit [S6] en relation avec $\zeta(2,2,\ldots,2)$, quand $n=2p$ est pair :

$$\int_{[0,1]^n} \frac{(y_1y_2)^{p(N+1)-1}(y_3y_4)^{(p-1)(N+1)-1}\ldots(y_{n-1}y_n)^N\prod_{k=1}^n(1-y_k)^N}{\prod_{k\in\{2,\ldots,n\}\text{ pair}}(1-y_1y_2\ldots y_k)^{N+1}} \mathrm{d}y_1\ldots\mathrm{d}y_n, \tag{4.7}$$

et de celles qui apparaissent dans les travaux de Rivoal ([B2], [R5]) pour $r<\frac{n-1}{2}$:

$$\int_{[0,1]^n} \frac{\prod_{k=1}^n y_k^{rN}(1-y_k)^N}{(1-y_1\ldots y_n)^{(2r+1)N+2}} \mathrm{d}y_1\ldots\mathrm{d}y_n. \tag{4.8}$$

Dans la cinquième partie de ce texte, on considère la famille suivante, qui contient ces trois cas particuliers :

$$K(A_1,\ldots,A_n,B_1,\ldots,B_n,C_2,\ldots,C_n) = \int_{[0,1]^n} \frac{\prod_{k=1}^n y_k^{A_k}(1-y_k)^{B_k}}{\prod_{k=2}^n(1-y_1\ldots y_k)^{C_k+1}} \mathrm{d}y_1\ldots\mathrm{d}y_n. \tag{4.9}$$

On démontre que ces intégrales admettent un groupe de Rhin-Viola, et on exhibe (en réponse à une question de Tanguy Rivoal) un changement de variables qui les transforme en des intégrales de la forme :

$$\int_{[0,1]^n} \frac{\prod_{k=1}^n x_k^{\tilde{a}_k}(1-x_k)^{\tilde{b}_k}}{\prod_{k=2}^n \delta_k(x_{n+1-k},\ldots,x_n)^{\tilde{c}_k}} \mathrm{d}x_1\ldots\mathrm{d}x_n.$$

Dans l'autre sens, ce changement de variables permet d'écrire toute intégrale de la forme (4.5) sous la forme (4.9). L'intérêt de cette manipulation est que les intégrales (4.9) se développent naturellement en séries multiples. Par exemple, on transforme ainsi l'intégrale (4.3) en (4.6) : Beukers et Sorokin construisent exactement les mêmes formes linéaires en 1 et $\zeta(3)$.

Dans ce texte, on suppose toujours que les exposants dans les intégrales n-uples sont entiers ; cependant tous les résultats démontrés ici sont de nature analytique et se généralisent aisément à des exposants complexes.

Les résultats obtenus dans les trois dernières parties ont été annoncés dans [F2].

Remerciements : Ce texte a beaucoup bénéficié de discussions avec Pierre Cartier et Tanguy Rivoal. Je tiens également à remercier Francesco Amoroso, Jacky Cresson, Gilles Damamme, Pierre Grinspan, Federico Pellarin, Serge Perrine, Georges Racinet, Eric Reyssat, Georges Rhin et Michel Waldschmidt pour l'intérêt qu'ils ont porté à ce travail, ainsi que le referee pour sa lecture attentive.

1 Le cas de $\zeta(2)$

Dans cette partie, on donne une présentation “géométrique” du groupe de Rhin-Viola pour $\zeta(2)$ [R3]. Au paragraphe 1.1, on écrit l'intégrale notée (4.1) dans l'introduction grâce à une sous-variété $\mathcal{V}_2$ de $\mathbb{R}^5$, dont on note Ω_2 l'ensemble des points à coordonnées strictement positives. L'action de $\mathcal{D}_5$ sur $\mathcal{V}_2$ (et sur Ω_2) traduit l'invariance de cette intégrale sous une action de $\mathcal{D}_5$. Au paragraphe 1.2, on interprète la transformation hypergéométrique de Rhin-Viola par un automorphisme non trivial d'une variété liée à $\mathcal{V}_2$. Le paragraphe suivant décrit le groupe de Rhin-Viola tout entier comme groupe d'automorphismes de $\Omega_2 \times U_2$, où U_2 est l'ensemble des $(t_1, \dots, t_5) \in]0,1[^5$ tels que $t_1 + \ldots + t_5 = 1$. Enfin, le paragraphe 1.4 contient des rappels, et quelques compléments, sur l'action du groupe de Rhin-Viola.

Dans cette partie et dans la suivante, pour toute fonction f on note $\mathrm{d}^\times f = \frac{\mathrm{d}f}{f}$.

1.1 La variété $\mathcal{V}_2$ et l'invariance par le groupe diédral

Notons $\mathcal{V}_2$ la sous-variété algébrique affine de $\mathbb{R}^5$ définie par les équations suivantes :

$$\begin{aligned}
x_1x_2(1+x_5x_3) &= 1\\
x_2x_3(1+x_1x_4) &= 1\\
x_3x_4(1+x_2x_5) &= 1\\
x_4x_5(1+x_3x_1) &= 1\\
x_5x_1(1+x_4x_2) &= 1
\end{aligned}$$

On note Ω_2 l'ensemble des points de $\mathcal{V}_2$ à coordonnées strictement positives, et ω_2 la 2-forme différentielle $\mathrm{d}^\times(x_{i-1}x_i) \wedge \mathrm{d}^\times(x_ix_{i+1})$ sur $\mathcal{V}_2$, qui ne dépend pas du choix de $i \in \{1, \dots, 5\}$ (voir la preuve de la proposition ci-dessous) ; on pose $x_0 = x_5$, $x_6 = x_1$ et ainsi de suite.

Proposition 1.1 *La variété $\mathcal{V}_2$ est de dimension deux, et le groupe diédral $\mathcal{D}_5$ agit dessus par permutation des coordonnées. Cette action laisse stable Ω_2, ainsi que la forme différentielle ω_2 (au signe près).*

Démonstration : Les équations de $\mathcal{V}_2$ sont visiblement stables sous l'action naturelle de $\mathcal{D}_5$ sur les cinq coordonnées. Quant à la forme différentielle, elle est changée en son opposé par les symétries, et invariante par les rotations (ce qui montre que sa définition ne dépend pas du choix de $i \in \{1, \dots, 5\}$). Ceci provient, en faisant le produit extérieur par $\mathrm{d}^\times(x_ix_{i+1})$, de la relation suivante, elle-même obtenue par différentiation des équations de $\mathcal{V}_2$:

$$\mathrm{d}^\times(x_{i+1}x_{i+2}) = -\mathrm{d}^\times(x_{i-1}x_i) - \frac{\mathrm{d}^\times(x_ix_{i+1})}{x_{i-1}x_{i+2}}. \tag{4.10}$$

Pour montrer que $\mathcal{V}_2$ est de dimension deux, on montre que la première, la troisième et la quatrième équations qui la définissent sont indépendantes et engendrent les deux autres. Par exemple, on peut en déduire la deuxième comme suit, en posant $u_i = x_ix_{i+1}$:

$$\begin{aligned}
u_1u_3 &= (1-u_5u_2)(1-u_2u_4)\\
&= 1-u_2(u_5+u_4(1-u_2u_5))\\
&= 1-u_2(u_5+u_4u_1)\\
&= 1-u_2.
\end{aligned}$$

Proposition 1.2 *Les formules suivantes définissent une paramétrisation bijective de Ω_2 par $]0,1[^2$, dans laquelle ω_2 correspond à $\frac{\mathrm{d}x\wedge\mathrm{d}y}{1-xy}$:*

$$\begin{aligned}
x_1 &= \left(\frac{xy(1-x)}{(1-y)(1-xy)}\right)^{\frac{1}{2}} \\
x_2 &= \left(\frac{(1-x)(1-y)}{xy(1-xy)}\right)^{\frac{1}{2}} \\
x_3 &= \left(\frac{xy(1-y)}{(1-x)(1-xy)}\right)^{\frac{1}{2}} \\
x_4 &= \left(\frac{(1-x)y(1-xy)}{x(1-y)}\right)^{\frac{1}{2}} \\
x_5 &= \left(\frac{x(1-y)(1-xy)}{(1-x)y}\right)^{\frac{1}{2}}
\end{aligned}$$

Démonstration : Il s'agit de vérifications sans difficulté ; la réciproque est donnée par $x = x_1x_5$ et $y = x_3x_4$. On a alors : $\frac{\mathrm{d}x\wedge\mathrm{d}y}{1-xy} = \frac{\mathrm{d}(xy)\wedge\mathrm{d}y}{y(1-xy)} = \mathrm{d}^\times(x_3x_4)\wedge\mathrm{d}^\times(x_4x_5) = \omega_2$.

N.B. On pourrait faire disparaître les racines carrées dans les formules de la proposition 1.2, en remplaçant partout les x_i par leurs carrés (ce qui changerait les équations de $\mathcal{V}_2$).

Étant donnés cinq nombres réels $\alpha_1,\ldots,\alpha_5$, on leur associe (pour faire le lien avec les notations de [R3]) :

$$\begin{aligned}
h &= \frac{1}{2}(\alpha_5+\alpha_1+\alpha_3-\alpha_2-\alpha_4) \\
i &= \frac{1}{2}(\alpha_1+\alpha_2+\alpha_4-\alpha_3-\alpha_5) \\
j &= \frac{1}{2}(\alpha_2+\alpha_3+\alpha_5-\alpha_4-\alpha_1) \\
k &= \frac{1}{2}(\alpha_3+\alpha_4+\alpha_1-\alpha_5-\alpha_2) \\
l &= \frac{1}{2}(\alpha_4+\alpha_5+\alpha_2-\alpha_1-\alpha_3)
\end{aligned}$$

Dans toute la suite de cette partie, on suppose que h,i,j,k,l sont entiers. Les variables auxiliaires de [R3] s'écrivent :

$$\begin{aligned}
h+i-k &= \frac{1}{2}(\alpha_5+\alpha_1+\alpha_2-\alpha_3-\alpha_4) \\
i+j-l &= \frac{1}{2}(\alpha_1+\alpha_2+\alpha_3-\alpha_4-\alpha_5) \\
j+k-h &= \frac{1}{2}(\alpha_2+\alpha_3+\alpha_4-\alpha_5-\alpha_1) \\
k+l-i &= \frac{1}{2}(\alpha_3+\alpha_4+\alpha_5-\alpha_1-\alpha_2) \\
l+h-j &= \frac{1}{2}(\alpha_4+\alpha_5+\alpha_1-\alpha_2-\alpha_3)
\end{aligned}$$

A l'inverse, on peut retrouver $\alpha_1, \ldots, \alpha_5$ à partir de h, i, j, k, l grâce aux formules suivantes :

$$\begin{aligned} h+i &= \alpha_1 \\ i+j &= \alpha_2 \\ j+k &= \alpha_3 \\ k+l &= \alpha_4 \\ l+h &= \alpha_5 \end{aligned}$$

On oriente Ω_2 par la paramétrisation de la proposition 1.2. Considérons l'intégrale suivante (qui est positive, éventuellement égale à $+\infty$) :

$$J(\alpha_1, \ldots, \alpha_5) = \int_{\Omega_2} x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} x_4^{\alpha_4} x_5^{\alpha_5} \omega_2.$$

D'après la proposition 1.2, on retrouve exactement l'intégrale notée $I(h, i, j, k, l)$ dans [R3], à savoir :

$$\int_0^1 \int_0^1 \frac{x^h(1-x)^i y^k (1-y)^j}{(1-xy)^{i+j-l}} \frac{\mathrm{d}x\,\mathrm{d}y}{1-xy}.$$

En particulier, on voit facilement (ou on déduit de la proposition 5.1) que cette intégrale est finie si, et seulement si, les entiers h, i, j, k, l associés à $\alpha_1, \ldots, \alpha_5$ sont positifs. De la proposition 1.1 et de la formule de changement de variables (voir par exemple [C1]) découle immédiatement le résultat suivant :

Corollaire 1.3 *L'action naturelle du groupe diédral $\mathcal{D}_5$ sur $(\alpha_1, \ldots, \alpha_5)$ laisse invariant $J(\alpha_1, \ldots, \alpha_5)$.*

N.B. Cette invariance par $\mathcal{D}_5$ correspond aux changements de variables τ et σ de [R3].

1.2 La variété $\tilde{\mathcal{V}}_2$ et la transformation hypergéométrique

Notons $\tilde{\mathcal{V}}_2$ la sous-variété algébrique de $\mathbb{R}^5$ définie par les équations suivantes :

$$\begin{aligned} y_3 y_4 (1 + y_2 y_5) &= 1 \\ y_5 y_1 (1 + y_4 y_2) + y_1 y_2 (1 + y_5 y_3) &= 1 \end{aligned} \tag{4.11}$$

On note $\tilde{\Omega}_2$ l'ensemble des points de $\tilde{\mathcal{V}}_2$ à coordonnées strictement positives, et $\tilde{\omega}_2$ la forme différentielle $y_1 y_2 y_3 y_5 \mathrm{d}^\times(y_1 y_2) \wedge \mathrm{d}^\times(y_2 y_3) \wedge \mathrm{d}^\times(y_3 y_5)$ sur $\tilde{\Omega}_2$.

Proposition 1.4 *La variété $\tilde{\mathcal{V}}_2$ est de dimension trois. Quand on échange y_2 et y_5, $\tilde{\mathcal{V}}_2$ et $\tilde{\Omega}_2$ sont invariantes, et $\tilde{\omega}_2$ est changée en son opposé.*

N.B. Quand on échange y_3 et y_4, la variété $\tilde{\mathcal{V}}_2$ est aussi invariante (voir le paragraphe 1.4).

Démonstration de la proposition 1.4 : Elle est immédiate en ce qui concerne $\tilde{\mathcal{V}}_2$ et $\tilde{\Omega}_2$ (il suffit de développer la seconde équation de $\tilde{\mathcal{V}}_2$) ; pour la forme différentielle, on prend le produit extérieur par $\mathrm{d}^\times(y_2 y_3) \wedge \mathrm{d}^\times(y_3 y_5)$ de la relation $\mathrm{d}^\times(y_1 y_2) - \mathrm{d}^\times(y_1 y_5) = \mathrm{d}^\times(y_2 / y_5) = \mathrm{d}^\times(y_2 y_3) - \mathrm{d}^\times(y_3 y_5)$.

Proposition 1.5 *Les formules suivantes définissent une bijection de* $\Omega_2 \times]0,1[$ *dans* $\tilde{\Omega}_2$ *:*

$$\begin{aligned}
y_1 &= x_1 \left(t(1-t)\right)^{\frac{1}{2}} \\
y_2 &= x_2 \left(\frac{t}{1-t}\right)^{\frac{1}{2}} \\
y_3 &= x_3 \left(\frac{t}{1-t}\right)^{\frac{1}{2}} \\
y_4 &= x_4 \left(\frac{1-t}{t}\right)^{\frac{1}{2}} \\
y_5 &= x_5 \left(\frac{1-t}{t}\right)^{\frac{1}{2}}
\end{aligned}$$

Dans cette bijection, la 3-forme $\tilde{\omega}_2$ *correspond à* $\omega_2 \wedge \mathrm{d}t$.

Démonstration : Soit $(x_1, \ldots, x_5, t) \in \Omega_2 \times]0,1[$. Le point $(y_1, \ldots, y_5)$ qui lui est associé vérifie $y_3y_4 = x_3x_4$, $y_5y_1 = (1-t)x_5x_1$, $y_2y_5 = x_2x_5$, $y_4y_2 = x_4x_2$, $y_1y_2 = tx_1x_2$ et $y_5y_3 = x_5x_3$. On en déduit immédiatement que les deux équations de $\tilde{\mathcal{V}}_2$ sont satisfaites pour le point $(y_1, \ldots, y_5)$: la seconde traduit l'identité $(1-t)+t = 1$.

Exhibons la bijection réciproque en posant, pour $(y_1, \ldots, y_5) \in \tilde{\mathcal{V}}_2$:

$$t = y_1y_2(1 + y_3y_5), \tag{4.12}$$

et en définissant $(x_1, \ldots, x_5)$ à l'aide des formules de la proposition 1.5. On a visiblement $t > 0$, et aussi $t < 1$ d'après (4.11). On vérifie aisément que $(x_1, \ldots, x_5, t) \in \Omega_2 \times]0,1[$, et que les deux applications sont réciproques l'une de l'autre. En ce qui concerne la forme différentielle, on écrit :

$$\begin{aligned}
\omega_2 \wedge \mathrm{d}t &= \mathrm{d}^\times(\frac{1}{t}y_1y_2) \wedge \mathrm{d}^\times(\frac{1-t}{t}y_2y_3) \wedge \mathrm{d}t \\
&= \mathrm{d}^\times(y_1y_2) \wedge \mathrm{d}^\times(y_2y_3) \wedge ((1+y_3y_5)\mathrm{d}(y_1y_2) + y_1y_2\mathrm{d}(y_3y_5)) \\
&= y_1y_2y_3y_5\mathrm{d}^\times(y_1y_2) \wedge \mathrm{d}^\times(y_2y_3) \wedge \mathrm{d}^\times(y_3y_5).
\end{aligned}$$

Ceci termine la preuve de la proposition 1.5.

Corollaire 1.6 *On définit un automorphisme de* $\Omega_2 \times]0,1[$ *en associant à* $(x_1, x_2, x_3, x_4, x_5, t)$

le point $(x'_1, x'_2, x'_3, x'_4, x'_5, t')$ défini par les équations suivantes :

$$\begin{aligned}
t' &= (1-x_1x_2)t + x_1x_5(1-t) \\
x'_1 &= x_1 \left(\frac{t(1-t)}{t'(1-t')}\right)^{\frac{1}{2}} \\
x'_2 &= x_5 \left(\frac{(1-t)(1-t')}{tt'}\right)^{\frac{1}{2}} \\
x'_3 &= x_3 \left(\frac{t(1-t')}{(1-t)t'}\right)^{\frac{1}{2}} \\
x'_4 &= x_4 \left(\frac{(1-t)t'}{t(1-t')}\right)^{\frac{1}{2}} \\
x'_5 &= x_2 \left(\frac{tt'}{(1-t)(1-t')}\right)^{\frac{1}{2}}
\end{aligned}$$

De plus, cet automorphisme change la 3-forme $\omega_2 \wedge \mathrm{d}t$ sur $\Omega_2 \times]0,1[$ en son opposé.

DÉMONSTRATION : Il s'agit du conjugué, par la bijection de la proposition 1.5, de l'automorphisme de $\tilde{\mathcal{V}}_2$ qui échange y_2 et y_5. Bien sûr, on peut aussi démontrer ce corollaire par un calcul direct, sans utiliser la variété $\tilde{\mathcal{V}}_2$.

Théorème 1.7 *Supposons que les entiers h, i, j, k, l associés à $\alpha_1, \ldots, \alpha_5$ sont positifs, ainsi que $i+j-l$ et $l+h-j$. Alors on a :*

$$(i+j-l)!(l+h-j)!J(\alpha_1, \ldots, \alpha_5) = h!i!J(\alpha_1, \alpha_5, \alpha_3, \alpha_4, \alpha_2).$$

Ce théorème est crucial dans la méthode de Rhin-Viola. Il s'agit de la formule notée (3.3) dans [R3] ; elle y est déduite de la représentation intégrale d'Euler de la fonction hypergéométrique de Gauss :

$$\frac{1}{h!i!} I(h,i,j,k,l) = \frac{1}{(i+j-l)!(l+h-j)!} I(i+j-l, l+h-j, j, k, l).$$

DÉMONSTRATION du théorème 1.7 : Le théorème de Fubini et l'expression classique $\frac{a!b!}{(a+b+1)!} = \int_0^1 t^a(1-t)^b \mathrm{d}t$ de la fonction Bêta d'Euler montrent que le membre de gauche dans l'énoncé du théorème s'écrit :

$$\Gamma(\alpha_1) \int_{\Omega_2 \times]0,1[} t^{\frac{1}{2}(\alpha_1+\alpha_2+\alpha_3-\alpha_4-\alpha_5)}(1-t)^{\frac{1}{2}(\alpha_4+\alpha_5+\alpha_1-\alpha_2-\alpha_3)} x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} x_4^{\alpha_4} x_5^{\alpha_5} \omega_2 \wedge \mathrm{d}t. \quad (4.13)$$

La proposition 1.5 définit une orientation de $\tilde{\Omega}_2$ et montre que l'intégrale précédente vaut :

$$\Gamma(\alpha_1) \int_{\tilde{\Omega}_2} y_1^{\alpha_1} y_2^{\alpha_2} y_3^{\alpha_3} y_4^{\alpha_4} y_5^{\alpha_5} \tilde{\omega}_2.$$

La proposition 1.4 montre qu'on peut appliquer à cette intégrale le changement de variables qui échange y_2 et y_5 (voir [C1] pour la formule de changement de variables quand la forme différentielle est changée en son opposé). Donc le membre de gauche du théorème 1.7 est invariant quand on échange α_2 et α_5 : il est égal au membre de droite.

N.B. On aurait pu appliquer le corollaire 1.6 pour démontrer directement que l'intégrale (4.13) est invariante quand on échange α_2 et α_5.

N.B. Le groupe symétrique $\mathfrak{S}_5$ agit sur $\mathbb{R}^5$ par permutation des coordonnées, donc aussi sur l'ensemble des sous-variétés algébriques de $\mathbb{R}^5$. Le stabilisateur de $\mathcal{V}_2$ pour cette action est le groupe diédral $\mathcal{D}_5$; l'orbite de $\mathcal{V}_2$ est donc formée par douze variétés. Les $J(\alpha_{\gamma(1)}, \dots, \alpha_{\gamma(5)})$, pour $\gamma \in \mathfrak{S}_5$, sont les intégrales sur ces variétés du monôme $x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} x_4^{\alpha_4} x_5^{\alpha_5}$.

1.3 Le groupe de Rhin-Viola comme groupe d'automorphismes

Notons U_2 le simplexe de dimension 4 défini par :

$$U_2 = \{(t_1, \dots, t_5) \in]0, 1[^5,\ t_1 + \dots + t_5 = 1\}.$$

On considère sur U_2 la 4-forme différentielle η_2 définie, pour $i \in \{1, \dots, 5\}$, par :

$$\eta_2 = \mathrm{d}t_i \wedge \mathrm{d}t_{i+1} \wedge \mathrm{d}t_{i+2} \wedge \mathrm{d}t_{i+3},$$

où on note $t_6 = t_1$, et ainsi de suite. Il est immédiat que η_2 ne dépend pas du choix de i.

Sur $\Omega_2 \times U_2$ on définit des fonctions $z_1, \dots, z_5$ qui à tout point $(\underline{x}, \underline{t}) = (x_1, \dots, x_5, t_1, \dots, t_5)$ associent :

$$\begin{aligned}
z_1(\underline{x}, \underline{t}) &= x_1 \left(\frac{t_5 t_1 t_2}{t_4 t_3} \right)^{\frac{1}{2}} \\
z_2(\underline{x}, \underline{t}) &= x_2 \left(\frac{t_1 t_2 t_3}{t_5 t_4} \right)^{\frac{1}{2}} \\
z_3(\underline{x}, \underline{t}) &= x_3 \left(\frac{t_2 t_3 t_4}{t_1 t_5} \right)^{\frac{1}{2}} \\
z_4(\underline{x}, \underline{t}) &= x_4 \left(\frac{t_3 t_4 t_5}{t_2 t_1} \right)^{\frac{1}{2}} \\
z_5(\underline{x}, \underline{t}) &= x_5 \left(\frac{t_4 t_5 t_1}{t_3 t_2} \right)^{\frac{1}{2}}
\end{aligned}$$

On peut résumer ces équations sous la forme générale suivante, pour $i \in \{1, \dots, 5\}$:

$$z_i(\underline{x}, \underline{t}) = x_i \left(\frac{t_{i-1} t_i t_{i+1}}{t_{i-2} t_{i+2}} \right)^{\frac{1}{2}}.$$

Il convient de préciser les conventions qu'on utilise pour les différentes actions de $\mathfrak{S}_5$. Le groupe $\mathfrak{S}_5$ agit naturellement à droite sur l'ensemble des points $(x_1, \dots, x_5)$ de $\mathbb{R}^5$, par $\gamma \cdot (x_1, \dots, x_5) = (x_{\gamma(1)}, \dots, x_{\gamma(5)})$ (et de même pour l'action sur U_2). On en déduit une action à gauche de $\mathfrak{S}_5$ sur l'ensemble des fonctions de $\mathbb{R}^5$ dans $\mathbb{R}$, par composition. Ainsi, $\mathfrak{S}_5$ agit à gauche sur l'ensemble des cinq fonctions coordonnées $z_1, \dots, z_5$ (qui vont de $\mathbb{R}^5$ dans $\mathbb{R}$), par $(\gamma \cdot z_i)(\underline{x}) = z_i(\gamma \cdot \underline{x}) = z_{\gamma(i)}(\underline{x})$.

Concernant les exposants $\alpha_1, \dots, \alpha_5$, il y a deux manières de les considérer. Ou bien on note $(\alpha_1, \dots, \alpha_5)$ un point de $\mathbb{R}^5$, et $\mathfrak{S}_5$ agit à droite sur l'ensemble de ces points : c'est ce qu'on fait dans ce paragraphe. Ou bien on note $\alpha_1, \dots, \alpha_5$ les formes linéaires coordonnées sur $\mathbb{R}^5$, et dans ce cas $\mathfrak{S}_5$ agit à gauche sur $\{\alpha_1, \dots, \alpha_5\}$. Dans les deux cas, on étend cette action par linéarité, et $\mathfrak{S}_5$ agit sur $\{h, \dots, l, h+i-k, \dots, l+h-j\}$. On reviendra sur cette ambiguïté au paragraphe 3.2.

Théorème 1.8 *Il existe un groupe H d'automorphismes de $\Omega_2 \times U_2$ (qui fixent, au signe près, $\omega_2 \wedge \eta_2$) et un homomorphisme de groupes surjectif $\pi : H \to \mathfrak{S}_5$ tels que, pour tout $f \in H$ et tout $i \in \{1, \dots, 5\}$, on ait :*

$$z_i \circ f = z_{\pi(f)^{-1}(i)}.$$

N.B. Les éléments de H sont des fonctions algébriques (au sens où elles sont définies par composition de fractions rationnelles et d'extractions de racines).

Corollaire 1.9 *La quantité suivante ne dépend pas du choix de $\gamma \in \mathfrak{S}_5$:*

$$\frac{J(\gamma(\alpha_1), \dots, \gamma(\alpha_5))}{\gamma(h)!\gamma(i)!\gamma(j)!\gamma(k)!\gamma(l)!}.$$

Ce corollaire contient toutes les identités entre intégrales utilisées par Rhin-Viola. On peut le déduire directement du corollaire 1.3 (qui donne l'invariance par $\mathcal{D}_5$) et du théorème 1.7 (qui donne l'invariance par φ) ; la démonstration présentée ici le relie aux automorphismes de $\Omega_2 \times U_2$.

Démonstration du corollaire 1.9 : Le théorème 1.8 montre que la quantité suivante est invariante sous l'action de $\mathfrak{S}_5$:

$$\begin{aligned} & (\alpha_1 + \ldots + \alpha_5 + 4)! \int_{\Omega_2 \times U_2} z_1^{\alpha_1} z_2^{\alpha_2} z_3^{\alpha_3} z_4^{\alpha_4} z_5^{\alpha_5} \omega_2 \wedge \eta_2 \\ = \quad & J(\alpha_1, \dots, \alpha_5) \int_{]0,+\infty[\times U_2} e^{-u} u^{\alpha_1+\ldots+\alpha_5} t_1^{h+i-k} t_2^{i+j-l} t_3^{j+k-h} t_4^{k+l-i} t_5^{l+h-j} u^4 \mathrm{d}u \wedge \eta_2. \end{aligned}$$

En utilisant la relation

$$\alpha_1 + \ldots + \alpha_5 = (h+i-k) + (i+j-l) + (j+k-h) + (k+l-i) + (l+h-j)$$

et le changement de variables donné par $t'_i = u t_i$ pour tout $i \in \{1, \dots, 5\}$ on obtient :

$$\begin{aligned} & J(\alpha_1, \dots, \alpha_5) \int_{]0,+\infty[^5} e^{-(t'_1+\ldots+t'_5)} {t'_1}^{h+i-k} {t'_2}^{i+j-l} {t'_3}^{j+k-h} {t'_4}^{k+l-i} {t'_5}^{l+h-j} \mathrm{d}t'_1 \wedge \ldots \wedge \mathrm{d}t'_5 \\ = \quad & (h+i-k)!(i+j-l)!(j+k-h)!(k+l-i)!(l+h-j)! J(\alpha_1, \dots, \alpha_5). \end{aligned}$$

En divisant par la quantité

$$(h+i-k)!(i+j-l)!(j+k-h)!(k+l-i)!(l+h-j)!h!i!j!k!l!$$

qui est elle aussi invariante par $\mathfrak{S}_5$ on conclut la démonstration du corollaire.

Démonstration du théorème 1.8 : Pour $\gamma \in \mathcal{D}_5$, notons f_γ l'automorphisme donné par l'action diagonale de $\gamma^{-1} \in \mathcal{D}_5$ sur $\Omega_2 \times U_2$, c'est-à-dire :

$$f_\gamma(x_1, \dots, x_5, t_1, \dots, t_5) = (x_{\gamma^{-1}(1)}, \dots, x_{\gamma^{-1}(5)}, t_{\gamma^{-1}(1)}, \dots, t_{\gamma^{-1}(5)}).$$

Pour φ, on définit $(X_1, \dots, X_5, T_1, \dots, T_5) = f_\varphi(x_1, \dots, x_5, t_1, \dots, t_5)$ par les formules suivantes, directement inspirées de celles du corollaire 1.6 et de l'action de φ sur $h+i-k, \dots, l+$

$h-j$:

$$\begin{aligned}
T_1 &= t_1\\
T_2 &= (1-x_1x_2)t_2 + x_1x_5t_5\\
T_3 &= t_4\\
T_4 &= t_3\\
T_5 &= x_1x_2t_2 + (1-x_1x_5)t_5\\
X_1 &= x_1\left(\frac{t_2t_5}{T_2T_5}\right)^{\frac{1}{2}}\\
X_2 &= x_5\left(\frac{t_5T_5}{t_2T_2}\right)^{\frac{1}{2}}\\
X_3 &= x_3\left(\frac{t_2T_5}{t_5T_2}\right)^{\frac{1}{2}}\\
X_4 &= x_4\left(\frac{t_5T_2}{t_2T_5}\right)^{\frac{1}{2}}\\
X_5 &= x_2\left(\frac{t_2T_2}{t_5T_5}\right)^{\frac{1}{2}}
\end{aligned}$$

En fait les cinq dernières formules s'écrivent, pour $i \in \{1,\dots,5\}$:

$$X_i = x_{\varphi(i)}\left(\frac{T_{i-1}T_{i+2}}{T_{i-1}T_iT_{i+1}}\frac{t_{\varphi(i)-1}t_{\varphi(i)}t_{\varphi(i)+1}}{t_{\varphi(i)-1}t_{\varphi(i)+2}}\right)^{\frac{1}{2}}.$$

Alors f_φ (respectivement chacun des f_γ pour $\gamma \in \mathcal{D}_5$) est un automorphisme de $\Omega_2 \times U_2$ qui fixe au signe près la forme différentielle $\omega_2 \wedge \eta_2$, et induit sur les fonctions $z_1,\dots,z_5$ la permutation φ^{-1} (resp. γ^{-1}), c'est-à-dire qu'on a $z_i \circ f_\varphi = z_{\varphi^{-1}(i)}$ (resp. $z_i \circ f_\gamma = z_{\gamma^{-1}(i)}$) pour tout $i \in \{1,\dots,5\}$. Notons H le groupe d'automorphismes engendré par f_φ et les f_γ. Chaque élément f de H induit une permutation, notée $\pi(f)^{-1}$, des fonctions $z_1,\dots,z_5$. Cela termine la preuve du théorème 1.8.

N.B. Le fait qu'on ait $T_1 = t_1$ dans la définition de f_φ traduit l'invariance par φ de $h+i-k$. De même, les relations $T_3 = t_4$ et $T_4 = t_3$ traduisent que φ échange $j+k-h$ et $k+l-i$. Le niveau de φ s'interprète comme le nombre d'indices i tels que T_i ne soit pas l'un des t_j (voir le paragraphe suivant).

Question : Le morphisme π construit ici est-il injectif ?

1.4 Lien avec la présentation de Rhin-Viola

Les dix variables $h, i, j, k, l, h+i-k, i+j-l, j+k-h, k+l-i, l+h-j$ correspondent (via la position des signes moins dans les expressions qui les définissent ci-dessus) aux dix parties à deux éléments de l'ensemble $\{1,2,3,4,5\}$. Si on considère un pentagone régulier dont les sommets sont numérotés de 1 à 5, les variables auxiliaires correspondent aux paires de sommets consécutifs (donc aux côtés du pentagone), alors que h, i, j, k, l correspondent aux paires de sommets non consécutifs (donc aux diagonales). Le groupe $\mathfrak{S}_5$ agit transitivement sur ces dix variables ; pour l'action du sous-groupe $\mathcal{D}_5$, il y a deux orbites : les variables auxiliaires d'une part, et $\{h, i, j, k, l\}$ d'autre part. Dans ce paragraphe, on considère $\alpha_1, \ldots, \alpha_5$ comme les formes linéaires coordonnées sur $\mathbb{R}^5$, donc $\mathfrak{S}_5$ agit à gauche sur les formes linéaires h, ..., $l+h-j$.

On remplace maintenant ces variables $\alpha_1, \ldots, \alpha_5$ par des réels, et on suppose que les dix combinaisons linéaires $h, i, j, k, l, h+i-k, i+j-l, j+k-h, k+l-i, l+h-j$ sont des entiers positifs. On conserve la même action de $\mathfrak{S}_5$. Notons, en accord avec [R3], φ la transposition qui échange α_2 et α_5, τ la permutation circulaire qui envoie chaque α_i sur α_{i+1} et σ la double tranposition qui échange α_1 et α_3 ainsi que α_4 et α_5.

D'après le corollaire 1.9, en posant

$$\xi(\gamma; \alpha_1, \ldots, \alpha_5) = \frac{h!i!j!k!l!}{\gamma(h)!\gamma(i)!\gamma(j)!\gamma(k)!\gamma(l)}$$

on a pour tout $\gamma \in \mathfrak{S}_5$:

$$J(\alpha_{\gamma(1)}, \ldots, \alpha_{\gamma(5)}) = \xi(\gamma; \alpha_1, \ldots, \alpha_5) J(\alpha_1, \ldots, \alpha_5).$$

Dans le quotient de factorielles qu'est $\xi(\gamma; \alpha_1, \ldots, \alpha_5)$, des simplifications peuvent apparaître (même si $\alpha_1, \ldots, \alpha_5$ sont génériques, ce qu'on suppose ici). Rhin et Viola appellent *niveau* de $\gamma \in \mathfrak{S}_5$ le nombre de facteurs qu'il reste au numérateur une fois celles-ci effectuées ; c'est aussi le nombre de facteurs au dénominateur. Posons $\mathcal{A}_1 = \{h, i, j, k, l\}$ et $\mathcal{A}_2 = \{h+i-k, i+j-l, j+k-h, k+l-i, l+h-j\}$; ce sont deux $\mathcal{D}_5$-orbites dont la réunion est une $\mathfrak{S}_5$-orbite. Le niveau de γ est le nombre d'éléments de $\mathcal{A}_1$ que γ envoie dans $\mathcal{A}_2$: le numérateur est formé par les factorielles des éléments de $\mathcal{A}_1 \setminus \gamma \cdot \mathcal{A}_1$, et le dénominateur par les factorielles des éléments de $\mathcal{A}_2 \cap \gamma \cdot \mathcal{A}_1$. Comme $\mathcal{A}_1$ et $\mathcal{A}_2$ sont stables sous l'action de $\mathcal{D}_5$, le niveau de γ ne dépend que de la double classe $\mathcal{D}_5 \gamma \mathcal{D}_5$. En particulier, les éléments de $\mathcal{D}_5$ sont de niveau 0, et ceux de $\mathcal{D}_5 \varphi \mathcal{D}_5$ de niveau 2. Plus généralement, tout élément γ s'écrit comme un mot en les trois lettres φ, τ et σ ; le niveau de γ est alors inférieur ou égal au double du nombre minimal de φ qui apparaissent dans une telle écriture.

Un intérêt de la méthode présentée ici est qu'elle ne privilégie pas les générateurs σ, τ et φ. Ainsi, le théorème 1.8 traduit l'invariance de $\frac{J(\alpha_1, \ldots, \alpha_5)}{h!i!j!k!l!}$ par l'existence d'un groupe d'automorphismes de $\Omega_2 \times U_2$. Par ailleurs, pour chaque $\gamma \in \mathfrak{S}_5$ de niveau 2 on peut construire une variété $\tilde{\mathcal{V}}_2(\gamma)$ de manière analogue à $\tilde{\mathcal{V}}_2 = \tilde{\mathcal{V}}_2(\varphi)$ au paragraphe 1.2. On démontre la relation $J(\alpha_{\gamma(1)}, \ldots, \alpha_{\gamma(5)}) = \xi(\gamma) J(\alpha_1, \ldots, \alpha_5)$ de même qu'on démontre le théorème 1.7, en remplaçant $\tilde{\mathcal{V}}_2$ par $\tilde{\mathcal{V}}_2(\gamma)$. D'ailleurs si on note $\sigma_1 = \tau^{-1} \sigma \tau$ l'élément de $\mathcal{D}_5$ qui échange α_2 et α_5, ainsi que α_3 et α_4, alors on a $\tilde{\mathcal{V}}_2 = \tilde{\mathcal{V}}_2(\varphi) = \tilde{\mathcal{V}}_2(\sigma_1 \varphi)$. Ceci traduit que $\tilde{\mathcal{V}}_2$ est invariante par l'action de $\sigma_1 \varphi$, qui échange y_3 et y_4.

En termes d'automorphismes, ce processus revient à composer à droite et/ou à gauche l'automorphisme de $\Omega_2 \times]0, 1[$ donné par le corollaire 1.6 par des automorphismes de $\Omega_2 \times]0, 1[$ provenant de ceux de Ω_2 fournis par l'action de $\mathcal{D}_5$. Ce point de vue conduit à ne pas privilégier φ, mais seulement la double classe $\mathcal{D}_5 \varphi \mathcal{D}_5$: le groupe $\mathfrak{S}_5$ est engendré par $\mathcal{D}_5$ (i.e. l'ensemble des éléments de niveau zéro) et un élément, quelconque, de $\mathcal{D}_5 \varphi \mathcal{D}_5$ (i.e. de niveau 2).

Enfin, le choix de paramètres effectué dans [R3] pour obtenir la meilleure mesure d'irrationalité connue (qui est $h = i = 12$, $j = k = 14$, $l = 13$) correspond, avec les notations introduites ici, à $\alpha_1 = 24$, $\alpha_2 = 26$, $\alpha_3 = 28$, $\alpha_4 = 27$, $\alpha_5 = 25$.

2 Le cas de $\zeta(3)$

Cette partie concerne le groupe de Rhin-Viola pour $\zeta(3)$ [R4] ; elle est parallèle à la précédente.

2.1 La variété $\mathcal{V}_3$ et l'invariance par le groupe diédral

Notons $\mathcal{V}_3$ la sous-variété algébrique quasiprojective de $\mathbb{R}^4$ définie par l'équation et les inéquations suivantes :

$$\begin{aligned}
x_1 - x_2 + x_3 - x_4 + \frac{1}{x_1} - \frac{1}{x_2} + \frac{1}{x_3} - \frac{1}{x_4} &= 0 && (4.14)\\
x_1x_2x_3x_4 &\neq 0 &&\\
(x_1-x_2)(x_2-x_3)(x_3-x_4)(x_4-x_1) &\neq 0 && (4.15)\\
(x_1x_2-1)(x_2x_3-1)(x_3x_4-1)(x_4x_1-1) &\neq 0 && (4.16)
\end{aligned}$$

La relation (4.14) peut être remplacée par chacune des deux équations suivantes :

$$\begin{aligned}
x_1x_2(x_3x_4-1)(x_3-x_4) &= x_3x_4(x_1x_2-1)(x_2-x_1) && (4.17)\\
x_2x_3(x_4-x_1)(x_4x_1-1) &= x_1x_4(x_2x_3-1)(x_3-x_2), && (4.18)
\end{aligned}$$

et elle implique[1] :

$$x_1(x_3-x_4)(x_2x_3-1)(x_3x_4-x_1x_2) = x_3(x_4-x_1)(x_1x_2-1)(x_2x_3-x_1x_4). \qquad (4.19)$$

Considérons les trois fonctions $\mathfrak{a}$, $\mathfrak{b}$ et $\mathfrak{c}$ qui à tout point $\underline{x} = (x_1, \ldots, x_4) \in \mathcal{V}_3$ associent respectivement (en utilisant les relations (4.17) et (4.18)) :

$$\begin{aligned}
\mathfrak{a}(\underline{x}) &= \frac{x_3-x_4}{x_1x_2-1} = \frac{x_3x_4}{x_1x_2}\frac{x_2-x_1}{x_3x_4-1}\\
\mathfrak{b}(\underline{x}) &= \frac{x_1x_4-1}{x_2x_3-1} = \frac{x_1x_4}{x_2x_3}\frac{x_3-x_2}{x_4-x_1}\\
\mathfrak{c}(\underline{x}) &= \frac{x_2x_3-x_1x_4}{x_2x_3-1} = 1-\mathfrak{b}(\underline{x})
\end{aligned}$$

On définit des fonctions ε_5 et ε_6, de $\mathcal{V}_3$ dans $\mathbb{R}$, en posant :

$$\varepsilon_5(\underline{x}) = \frac{x_1}{x_3}\frac{\mathfrak{a}(\underline{x})}{\mathfrak{b}(\underline{x})} \text{ et } \varepsilon_6(\underline{x}) = \frac{x_2}{x_4}\mathfrak{a}(\underline{x})\mathfrak{b}(\underline{x}). \qquad (4.20)$$

On note Ω_3 l'ensemble des points $\underline{x} \in \mathcal{V}_3$ dont les quatre coordonnées sont strictement positives, et tels que les réels $\mathfrak{a}(\underline{x})$, $\mathfrak{b}(\underline{x})$ et $\mathfrak{c}(\underline{x})$ associés soient eux aussi strictement positifs. On introduit la forme différentielle suivante :

$$\omega_3 = -2\frac{x_2\mathfrak{a}(\underline{x})\mathfrak{b}(\underline{x})}{x_4^2-1}\mathrm{d}^\times x_1 \wedge \mathrm{d}^\times x_2 \wedge \mathrm{d}^\times x_3.$$

Considérons l'action naturelle du groupe diédral $\mathcal{D}_8$ à seize éléments sur l'octogone régulier dont les sommets sont, dans le sens direct, $x_1, x_2, x_3, x_4, \frac{1}{x_1}, \frac{1}{x_2}, \frac{1}{x_3}, \frac{1}{x_4}$. Elle définit une action à

[1] Avec équivalence si $x_3 \neq x_1$.

droite de $\mathcal{D}_8$ sur $\mathbb{R}^{*4}$ (par exemple, la rotation d'un huitième de tour dans le sens direct envoie (x_1,x_2,x_3,x_4) sur $(x_2,x_3,x_4,\frac{1}{x_1})$). On en déduit une action à gauche de $\mathcal{D}_8$ sur l'ensemble des fonctions de $\mathbb{R}^{*4}$ (ou d'une partie de $\mathbb{R}^{*4}$) dans $\mathbb{R}$: si f est une telle fonction et $\gamma \in \mathcal{D}_8$, on pose $\gamma \cdot f = f \circ \gamma$, c'est-à-dire :

$$(\gamma \cdot f)(x_1,\ldots,x_4) = f(\gamma \cdot (x_1,\ldots,x_4)).$$

Proposition 2.1 *La variété $\mathcal{V}_3$ est de dimension trois. L'action du groupe diédral $\mathcal{D}_8$ sur $\mathbb{R}^{*4}$ laisse stable $\mathcal{V}_3$, Ω_3, et la forme différentielle ω_3. De plus tout $\gamma \in \mathcal{D}_8$ de signature η_γ transforme ε_5 en $\varepsilon_5^{\eta_\gamma}$ et fixe ε_6.*

Démonstration : Il est clair que $\mathcal{V}_3$ est de dimension trois, et qu'elle est stable par $\mathcal{D}_8$. Notons $\vartheta \in \mathcal{D}_8$ la rotation d'un huitième de tour dans le sens direct : on a $\vartheta(x_1,x_2,x_3,x_4) = (x_2,x_3,x_4,\frac{1}{x_1})$. Par ailleurs, notons σ la symétrie définie par $\sigma(x_1,x_2,x_3,x_4) = (\frac{1}{x_4},\frac{1}{x_3},\frac{1}{x_2},\frac{1}{x_1})$. Alors ϑ échange les équations (4.17) et (4.18), alors que σ fixe (au signe près) la relation (4.14). Par ailleurs, on a :

$$\mathfrak{a}(\vartheta(\underline{x})) = \frac{1}{x_1}\mathfrak{b}(\underline{x}) \qquad \mathfrak{b}(\vartheta(\underline{x})) = \frac{x_2}{x_3x_4}\mathfrak{a}(\underline{x})$$

$$\mathfrak{a}(\sigma(\underline{x})) = \mathfrak{a}(\underline{x}) \qquad \mathfrak{b}(\sigma(\underline{x})) = \frac{x_2x_3}{x_1x_4}\mathfrak{b}(\underline{x})$$

Cela donne $\vartheta \cdot \varepsilon_5 = \frac{1}{\varepsilon_5}$, $\sigma \cdot \varepsilon_5 = \varepsilon_5$, et $\vartheta \cdot \varepsilon_6 = \sigma \cdot \varepsilon_6 = \varepsilon_6$, ce qui prouve l'assertion de la proposition concernant ε_5 et ε_6. Pour obtenir la stabilité de Ω_3, il suffit d'utiliser la relation suivante, qu'on déduit de (4.19) :

$$\mathfrak{c}(\vartheta(\underline{x})) = 1 - \frac{x_2\mathfrak{a}(\underline{x})}{x_3x_4} = 1 - (\frac{x_1x_2\mathfrak{a}(\underline{x})}{x_3\mathfrak{c}(\underline{x})})(\frac{\mathfrak{c}(\underline{x})}{x_1x_4}) = 1 - (1 - \frac{x_2\mathfrak{a}(\underline{x})\mathfrak{b}(\underline{x})}{\mathfrak{c}(\underline{x})})(1 - \frac{x_2x_3\mathfrak{b}(\underline{x})}{x_1x_4}).$$

Quant à la forme différentielle, son invariance sous l'action de $\mathcal{D}_8$ provient de la formule suivante, qu'on déduit de (4.14) :

$$((x_4 - \frac{1}{x_4})\mathrm{d}^\times x_4 - (x_1 - \frac{1}{x_1})\mathrm{d}^\times x_1) \wedge \mathrm{d}^\times x_2 \wedge \mathrm{d}^\times x_3 = 0.$$

Ceci termine la preuve de la proposition 2.1.

Proposition 2.2 *Les formules suivantes définissent une paramétrisation bijective de Ω_3 par $]0,1[^3\backslash Y$, où Y est une partie de $]0,1[^3$ dont la mesure de Lebesgue est nulle :*

$$\begin{aligned}
x_1 &= \left(\frac{xyz(1-x)}{(1-y)(1-z)(1-(1-xy)z)}\right)^{\frac{1}{2}} \\
x_2 &= \left(\frac{yz(1-x)(1-(1-xy)z)}{x(1-y)(1-z)}\right)^{\frac{1}{2}} \\
x_3 &= \left(\frac{y(1-x)(1-z)}{xz(1-y)(1-(1-xy)z)}\right)^{\frac{1}{2}} \\
x_4 &= \left(\frac{(1-x)(1-z)(1-(1-xy)z)}{xyz(1-y)}\right)^{\frac{1}{2}}
\end{aligned}$$

De plus, dans cette paramétrisation, ω_3 *correspond à* $\frac{\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z}{1-(1-xy)z}$, *et on a :*

$$\begin{aligned}
\varepsilon_5(x_1,\ldots,x_4) &= \left(\frac{z(1-x)(1-y)(1-z)}{xy(1-(1-xy)z)}\right)^{\frac{1}{2}} \\
\varepsilon_6(x_1,\ldots,x_4) &= \left(\frac{xyz(1-x)(1-y)(1-z)}{1-(1-xy)z}\right)^{\frac{1}{2}}
\end{aligned}$$

Démonstration : La partie Y de $]0,1[^3$ correspond aux inégalités qui apparaissent dans la définition de la variété quasiprojective $\mathcal{V}_3$ (voir ci-dessous). Pour $(x,y,z)\in]0,1[^3\setminus Y$, on vérifie que le point $(x_1,\ldots,x_4)$ associé appartient bien à Ω_3, et que les formules concernant $\varepsilon_5(\underline{x})$ et $\varepsilon_6(\underline{x})$ sont correctes. Pour obtenir la bijection réciproque, on pose :

$$x=\mathfrak{b}(\underline{x}), \qquad y=\frac{x_2x_3}{x_1x_4}\mathfrak{b}(\underline{x}), \qquad z=\frac{x_1x_2}{x_3}\frac{\mathfrak{a}(\underline{x})}{\mathfrak{c}(\underline{x})}.$$

On a alors, d'après (4.19) et (4.14) :

$$1-x=\mathfrak{c}(\underline{x}), \qquad 1-y=\frac{\mathfrak{c}(\underline{x})}{x_1x_4}, \qquad 1-z=x_2\frac{\mathfrak{a}(\underline{x})\mathfrak{b}(\underline{x})}{\mathfrak{c}(\underline{x})},$$

$$1-(1-xy)z=\frac{x_2}{x_1}\mathfrak{b}(\underline{x}).$$

Concernant la forme différentielle, il suffit de calculer $\mathrm{d}^\times x_1\wedge\mathrm{d}^\times x_2\wedge\mathrm{d}^\times x_3$, par exemple en l'écrivant sous la forme $\mathrm{d}^\times\frac{x_1}{x_3}\wedge\mathrm{d}^\times(x_2x_3)\wedge\mathrm{d}^\times x_3$, et de conclure grâce à l'égalité suivante :

$$\frac{x_2\mathfrak{a}(\underline{x})\mathfrak{b}(\underline{x})}{x_4^2-1}=\frac{xyz(1-x)(1-y)(1-z)}{(1-x)(1-z)(1-z(1-xy))-xyz(1-y)}.$$

N.B. La partie Y est donnée par les cas particuliers suivants :

$$\begin{array}{rcl}
x_1=x_2 & \text{si, et seulement si, } x_3x_4=1 & \text{si, et seulement si, } x=1-z(1-xy); \\
x_1x_2=1 & \text{si, et seulement si, } x_3=x_4 & \text{si, et seulement si, } y=1-z(1-xy); \\
x_2=x_3 & \text{si, et seulement si, } x_1=x_4 & \text{si, et seulement si, } 1-z=z(1-z(1-xy)); \\
x_2x_3=1 & \text{si, et seulement si, } x_1x_4=1 & \text{si, et seulement si, } x=y.
\end{array}$$

N.B. Les équations de $\mathcal{V}_3$ montrent qu'on a aussi :

$$x=\mathfrak{b}(\underline{x})=\sqrt{\frac{x_1x_4\varepsilon_6(\underline{x})}{x_2x_3\varepsilon_5(\underline{x})}} \text{ et } y=\sqrt{\frac{x_2x_3\varepsilon_6(\underline{x})}{x_1x_4\varepsilon_5(\underline{x})}}. \tag{4.21}$$

Soient $\alpha_1, \dots, \alpha_6$ six nombres réels. Pour faire le lien avec [R4], on leur associe :

$$\begin{aligned}
h &= \frac{1}{2}(\alpha_1 - \alpha_2 - \alpha_3 - \alpha_4 - \alpha_5 + \alpha_6)\\
j &= \frac{1}{2}(\alpha_1 + \alpha_2 - \alpha_3 - \alpha_4 + \alpha_5 + \alpha_6)\\
k &= \frac{1}{2}(\alpha_1 + \alpha_2 + \alpha_3 - \alpha_4 - \alpha_5 + \alpha_6)\\
l &= \frac{1}{2}(\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6)\\
m &= \frac{1}{2}(-\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \alpha_5 + \alpha_6)\\
q &= \frac{1}{2}(-\alpha_1 - \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6)\\
r &= \frac{1}{2}(-\alpha_1 - \alpha_2 - \alpha_3 + \alpha_4 - \alpha_5 + \alpha_6)\\
s &= \frac{1}{2}(-\alpha_1 - \alpha_2 - \alpha_3 - \alpha_4 + \alpha_5 + \alpha_6)
\end{aligned}$$

Dans toute la suite, on suppose que h, j, k, l, m, q, r, s sont entiers. Dans [R4], ces huit entiers sont soumis aux contraintes $j + q = l + s$ et $h + m = k + r$; ici, ces contraintes sont des conséquences des expressions de $h, \dots, s$ en fonction de $\alpha_1, \dots, \alpha_6$. Les six variables $\alpha_1, \dots, \alpha_6$ ne sont soumises a priori à aucune relation linéaire.

Les “entiers auxiliaires” de [R4] (formules (2.8) et (4.7)) sont alors :

$$\begin{array}{ccccccc}
h' &=& h+l-j &=& h+q-s &=& \frac{1}{2}(\alpha_1 - \alpha_2 + \alpha_3 + \alpha_4 - \alpha_5 + \alpha_6)\\
j' &=& j+m-k &=& j+r-h &=& \frac{1}{2}(-\alpha_1 + \alpha_2 - \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6)\\
k' &=& k+q-l &=& k+s-j &=& \frac{1}{2}(-\alpha_1 - \alpha_2 + \alpha_3 - \alpha_4 - \alpha_5 + \alpha_6)\\
l' &=& l+r-m &=& l+h-k &=& \frac{1}{2}(\alpha_1 - \alpha_2 - \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6)\\
m' &=& m+s-q &=& m+j-l &=& \frac{1}{2}(-\alpha_1 + \alpha_2 - \alpha_3 - \alpha_4 - \alpha_5 + \alpha_6)\\
q' &=& q+h-r &=& q+k-m &=& \frac{1}{2}(\alpha_1 - \alpha_2 + \alpha_3 - \alpha_4 + \alpha_5 + \alpha_6)\\
r' &=& r+j-s &=& r+l-q &=& \frac{1}{2}(\alpha_1 + \alpha_2 - \alpha_3 + \alpha_4 - \alpha_5 + \alpha_6)\\
s' &=& s+k-h &=& s+m-r &=& \frac{1}{2}(-\alpha_1 + \alpha_2 + \alpha_3 - \alpha_4 + \alpha_5 + \alpha_6)
\end{array}$$

On oriente Ω_3 par la paramétrisation de la proposition 2.2. On considère l'intégrale suivante (qui est positive, et éventuellement égale à $+\infty$) :

$$J(\alpha_1, \dots, \alpha_6) = \int_{\Omega_3} x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} x_4^{\alpha_4} \varepsilon_5(\underline{x})^{\alpha_5} \varepsilon_6(\underline{x})^{\alpha_6} \omega_3.$$

D'après la proposition 2.2, on retrouve exactement l'intégrale notée $I(h, j, k, l, m, q, r, s)$ dans [R4], à savoir :

$$\int_0^1 \int_0^1 \int_0^1 \frac{x^h (1-x)^l y^k (1-y)^s z^j (1-z)^q}{(1-(1-xy)z)^{q+h-r}} \frac{\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}{1-(1-xy)z}.$$

D'après la proposition 3.1 ci-dessous (ou d'après la deuxième partie de [R4]), cette intégrale est finie si, et seulement si, les huit entiers h, j, k, l, m, q, r, s associés à $\alpha_1, \dots, \alpha_6$ sont positifs ou nuls.

Considérons l'action naturelle (à gauche) de $\mathcal{D}_8$ sur un octogone régulier dont les sommets sont $\alpha_1, \alpha_2, \alpha_3, \alpha_4, -\alpha_1, -\alpha_2, -\alpha_3, -\alpha_4$. Par exemple, la rotation d'un huitième de tour envoie α_i sur α_{i+1} pour $i \in \{1,2,3\}$ et α_4 sur $-\alpha_1$. En outre, si $\gamma \in \mathcal{D}_8 \subset \mathfrak{S}_8$ a pour signature η_γ on pose $\gamma \cdot \alpha_5 = \eta_\gamma \alpha_5$ et $\gamma \cdot \alpha_6 = \alpha_6$. Par linéarité, on définit une action de $\mathcal{D}_8$ sur l'ensemble des combinaisons linéaires de $\alpha_1, \ldots, \alpha_5$. On voit que $\{h, \ldots, s\}$ est alors une $\mathcal{D}_8$-orbite, sur laquelle l'action est la même que sur un octogone dont les sommets seraient indexés par ces entiers (dans leur ordre alphabétique).

De la proposition 2.1 découle immédiatement le résultat suivant :

Corollaire 2.3 *L'action de $\mathcal{D}_8$ sur $(\alpha_1, \ldots, \alpha_6)$ laisse invariant $J(\alpha_1, \ldots, \alpha_6)$.*

N.B. Cette invariance par $\mathcal{D}_8$ correspond aux changements de variables ϑ et σ de [R4].

2.2 La variété $\tilde{\mathcal{V}}_3$ et la transformation hypergéométrique

Notons $\tilde{\mathcal{V}}_3$ la sous-variété algébrique quasiprojective de $\mathbb{R}^6$ définie par les équations et l'inéquation suivantes :

$$\begin{aligned}
(y_2 - y_1 - y_1y_2y_4)(y_3 + y_5) &= (y_1 - y_6)(y_4 - y_5 - y_2y_3y_4 + y_3y_4y_5) && (4.22)\\
y_1y_2y_6(y_3 + y_5)^2 &= y_3y_4y_5(y_1 - y_6)^2 && (4.23)\\
y_3y_5(y_3 + y_5)(y_1 - y_6) &\neq 0
\end{aligned}$$

Notons $\tilde{\Omega}_3$ l'ensemble des points de $\tilde{\mathcal{V}}_3$ dont les six coordonnées sont strictement positives et vérifient :

$$\begin{aligned}
y_1y_4y_5 &> y_2y_3y_6 && (4.24)\\
\frac{y_3}{y_5} &> (y_1y_2 - 1)\sqrt{\frac{y_3y_4y_6}{y_1y_2y_5}} && (4.25)
\end{aligned}$$

Notons $\tilde{\omega}_3$ la 4-forme différentielle suivante :

$$\tilde{\omega}_3 = \frac{-y_2y_4y_6 \mathrm{d}^\times y_1 \wedge \mathrm{d}^\times y_2 \wedge \mathrm{d}^\times y_3 \wedge ((y_3y_4 + 1)\mathrm{d}^\times y_4 + (y_3y_4 - 1)(\mathrm{d}^\times y_5 + \mathrm{d}^\times y_6))}{y_2(y_4^2 + 1)(y_3y_4 - 1) + (y_1(y_4^2 + 1) - 1)\sqrt{\frac{y_2y_3y_4}{y_1y_5y_6}}}$$

Proposition 2.4 *La variété $\tilde{\mathcal{V}}_3$ est de dimension quatre. Quand on échange y_3 et $\frac{1}{y_5}$ (donc aussi y_5 et $\frac{1}{y_3}$) en fixant les autres coordonnées, $\tilde{\mathcal{V}}_3$ et $\tilde{\Omega}_3$ sont invariantes, et $\tilde{\omega}_3$ aussi au signe près.*

Démonstration : Les vérifications sont immédiates, sauf pour la forme différentielle (pour laquelle elles sont plus laborieuses).

N.B. On peut démontrer que $\tilde{\omega}_3$ est invariante (au signe près) de la manière suivante. Le théorème 2.7 ci-dessous (qui est démontré dans [R4]) signifie que l'intégrale

$$\int_{\tilde{\Omega}_3} y_1^{\alpha_1} y_2^{\alpha_2} y_3^{\alpha_3} y_4^{\alpha_4} y_5^{\alpha_5} y_6^{\alpha_6} \tilde{\omega}_3$$

est invariante par le changement de variables qui échange y_3 et $\frac{1}{y_5}$ (donc y_5 et $\frac{1}{y_3}$) en fixant y_1, y_2, y_4 et y_6 (voir la preuve du théorème 2.7 pour les détails). On a donc, pour tous $\alpha_1, \ldots, \alpha_6$ tels que les dix entiers associés $h, j, k, l, m, q, r, s, q + h - r, r + j - s$ soient positifs :

$$\int_{\tilde{\Omega}_3} y_1^{\alpha_1} y_2^{\alpha_2} y_3^{\alpha_3} y_4^{\alpha_4} y_5^{\alpha_5} y_6^{\alpha_6} \tilde{\omega}_3 = \varepsilon \int_{\tilde{\Omega}_3} y_1^{\alpha_1} y_2^{\alpha_2} y_3^{\alpha_3} y_4^{\alpha_4} y_5^{\alpha_5} y_6^{\alpha_6} \tilde{\omega}'_3,$$

où $\tilde{\omega}_3'$ est l'image par ce changement de variables de $\tilde{\omega}_3$, et $\varepsilon = 1$ si ce changement de variables conserve l'orientation de $\tilde{\Omega}_3$, et $\varepsilon = -1$ sinon. Donc l'intégrale de tous ces monômes en $y_1, \dots, y_6$ contre $\tilde{\omega}_3 - \varepsilon\tilde{\omega}_3'$ est nulle : il en découle $\tilde{\omega}_3' = \varepsilon\tilde{\omega}_3$.

Proposition 2.5 *Les formules suivantes définissent une bijection de $\Omega_3\times]0,1[$ dans $\tilde{\Omega}_3 \setminus Y'$, où Y' est une partie de $\tilde{\Omega}_3$ dont la mesure de Lebesgue est nulle :*

$$\begin{aligned} y_1 &= x_1\sqrt{t(1-t)} \\ y_2 &= x_2\sqrt{\frac{1-t}{t}} \\ y_3 &= x_3\sqrt{\frac{t}{1-t}} \\ y_4 &= x_4\sqrt{\frac{1-t}{t}} \\ y_5 &= \varepsilon_5(\underline{x})\sqrt{\frac{t}{1-t}} \\ y_6 &= \varepsilon_6(\underline{x})\sqrt{t(1-t)} \end{aligned}$$

Dans cette bijection, la 4-forme $\tilde{\omega}_3$ correspond à $\omega_3 \wedge \mathrm{d}t$.

Démonstration : On note Y' l'ensemble des $(y_1, \dots, y_6)$ qui proviennent, par cette application, de points $(x_1, \dots, x_4, t)$ pour lesquels l'une au moins des inéquations (4.15), (4.16) n'est pas vérifiée. Tout d'abord, montrons que pour tout $(x_1, \dots, x_4, t)$ le point $(y_1, \dots, y_6)$ défini par la proposition appartient à $\tilde{\Omega}_3 \setminus Y'$. On déduit de (4.14) et des définitions de $\mathfrak{a}$ et $\mathfrak{b}$ la relation $x_2(\mathfrak{b}(\underline{x})x_3^2 + x_1\mathfrak{a}(\underline{x})) = x_3(x_1x_4 - x_2\mathfrak{a}(\underline{x})\mathfrak{b}(\underline{x}))$, qui s'écrit aussi, en utilisant (4.21) :

$$\frac{x_1 - \varepsilon_6(\underline{x})}{x_3 + \varepsilon_5(\underline{x})} = \frac{x_2\mathfrak{b}(\underline{x})}{x_4} = \sqrt{\frac{x_1x_2\varepsilon_6(\underline{x})}{x_3x_4\varepsilon_5(\underline{x})}}. \tag{4.26}$$

La relation (4.23) découle immédiatement de (4.26). Par ailleurs, on peut traduire (4.22) par la nullité d'une fonction affine de t. Il suffit alors de montrer que les deux coefficients de cette fonction sont nuls, ce qui ne pose pas de difficulté à partir de (4.26) et des équations de $\mathcal{V}_3$.

Pour vérifier (4.24), il suffit de voir que $\frac{y_2y_3y_6}{y_1y_4y_5} = \frac{x_2x_3\varepsilon_6(\underline{x})}{x_1x_4\varepsilon_5(\underline{x})}$ est compris entre 0 et 1 d'après (4.21). Quant à (4.25), on peut la déduire de la relation $\frac{x_3}{\varepsilon_5(\underline{x})} > (x_1x_2 - 1)\sqrt{\frac{x_3x_4\varepsilon_6(\underline{x})}{x_1x_2\varepsilon_5(\underline{x})}}$, qui elle-même provient de (4.21).

On a donc montré que tout point $(y_1, \dots, y_6)$ défini par les formules de la proposition 2.5 appartient à $\tilde{\Omega}_3 \setminus Y'$. Réciproquement, à tout $(y_1, \dots, y_6) \in \tilde{\Omega}_3 \setminus Y'$ on associe

$$t = 1 - y_1y_4 + (y_2y_3 - 1)\sqrt{\frac{y_1y_4y_6}{y_2y_3y_5}}, \tag{4.27}$$

puis $x_1, \dots, x_4$ donnés par les quatre premières formules de la proposition 2.5. D'après la relation (4.24), on a $y_2y_3\sqrt{\frac{y_1y_4y_6}{y_2y_3y_5}} - y_1y_4 < 0$ donc $t < 1$. Pour montrer que t est strictement positif, il suffit d'après (4.25) de prouver la relation suivante :

$$t = \frac{y_3 - (y_1y_2 - 1)\sqrt{\frac{y_3y_4y_5y_6}{y_1y_2}}}{y_3 + y_4 + \sqrt{\frac{y_3y_4y_5y_6}{y_1y_2}}}. \tag{4.28}$$

Or déduire (4.28) de (4.22) revient à démontrer :

$$(1 - y_2(y_3 + y_4 + y_5) + y_4(\frac{1}{y_3} + y_5))\sqrt{\frac{y_1y_3y_6}{y_2y_4y_5}} = 1 - y_1y_4 - \frac{y_6}{y_2} + y_3(y_6 - y_1), \tag{4.29}$$

et le membre de droite de cette équation s'écrit, en remplaçant y_6 par sa valeur compte tenu de (4.23) :

$$(1 - y_1y_4 - \frac{y_1}{y_2}) + (1 + \frac{y_5}{y_3} - y_2(y_3 + y_5))\sqrt{\frac{y_1y_3y_6}{y_2y_4y_5}}.$$

Cette expression est égale au membre de gauche de (4.29), puisque la combinaison de (4.22) et (4.23) donne :

$$1 - y_1y_4 - \frac{y_1}{y_2} = (\frac{y_4 - y_5}{y_3} - y_2y_4 + y_4y_5)\sqrt{\frac{y_1y_3y_6}{y_2y_4y_5}}.$$

On a donc montré que t est strictement compris entre 0 et 1. Il reste à vérifier que le point $(x_1, \ldots, x_4)$ appartient à Ω_3, et à démontrer les deux dernières relations de la proposition. On commence par exprimer t de la manière suivante, en utilisant les équations (4.22) et (4.27) :

$$t = \frac{y_1}{y_2} + (y_3y_4 - 1)\sqrt{\frac{y_1y_5y_6}{y_2y_3y_4}}. \tag{4.30}$$

Les équations (4.27), (4.28) et (4.30) se traduisent respectivement par les égalités suivantes, dans lesquelles les dénominateurs sont non nuls car $(y_1, \ldots, y_6) \notin Y'$:

$$\begin{aligned} \frac{y_2y_3 - 1}{y_1y_4 - (1-t)} &= \sqrt{\frac{y_2y_3y_5}{y_1y_4y_6}} \\ \frac{(1-t)y_3 - ty_4}{y_1y_2 - (1-t)} &= \sqrt{\frac{y_3y_4y_5y_6}{y_1y_2}} \\ \frac{y_3y_4 - 1}{ty_2 - y_1} &= \sqrt{\frac{y_3y_4}{y_1y_2y_5y_6}} \end{aligned}$$

Les deux premières égalités montrent que $\mathfrak{b}(\underline{x})$ et $\mathfrak{a}(\underline{x})$ sont strictement positifs. De plus, le produit et le quotient des deux premières égalités, ainsi que le produit de la deuxième et de la troisième, donnent :

$$\begin{aligned} y_5 &= \frac{y_1((1-t)y_3 - ty_4)(y_2y_3 - 1)}{y_3(y_1y_2 - (1-t))(y_1y_4 - (1-t))} \\ y_6 &= \frac{y_2((1-t)y_3 - ty_4)(y_1y_4 - (1-t))}{y_4(y_1y_2 - (1-t))(y_2y_3 - 1)} \\ y_1y_2(y_3y_4 - 1)((1-t)y_3 - ty_4) &= y_3y_4(y_1y_2 - (1-t))(ty_2 - y_1) \end{aligned}$$

De ces trois formules on déduit immédiatement d'une part les deux dernières relations de la proposition, et d'autre part l'égalité (4.17). Enfin, $\frac{x_2x_3 - x_1x_4}{x_2x_3 - 1}$ est égal à $\frac{y_1y_4 - y_2y_3\sqrt{\frac{y_1y_4y_6}{y_2y_3y_5}}}{1-t}$ d'après (4.27), donc strictement positif par l'hypothèse (4.24) : on a bien $\mathfrak{c}(\underline{x}) > 0$. Il en découle que le point $(x_1, \ldots, x_4, t)$ appartient à $\Omega_3 \times]0, 1[$.

Quant à la forme différentielle, son expression se déduit immédiatement de la formule (4.30) pour t, et de la différentiation de cette même formule qui donne :

$$\mathrm{d}^{\times}y_1\wedge\mathrm{d}^{\times}y_2\wedge\mathrm{d}^{\times}y_3\wedge\mathrm{d}t=\frac{1}{2}\sqrt{\frac{y_1y_5y_6}{y_2y_3y_4}}\mathrm{d}^{\times}y_1\wedge\mathrm{d}^{\times}y_2\wedge\mathrm{d}^{\times}y_3\wedge((y_3y_4+1)\mathrm{d}^{\times}y_4+(y_3y_4-1)(\mathrm{d}^{\times}y_5+\mathrm{d}^{\times}y_6)).$$

Ceci termine la démonstration de la proposition 2.5.

Corollaire 2.6 *On définit un automorphisme de $\Omega_3\times]0,1[$ en associant à (x_1,x_2,x_3,x_4,t) le point (x_1',x_2',x_3',x_4',t') défini par les équations suivantes :*

$$\begin{aligned}
t' &= (1-x_1\mathfrak{a}(\underline{x}))t+\mathfrak{b}(\underline{x})(1-t)\\
x_1' &= x_1\left(\frac{t(1-t)}{t'(1-t')}\right)^{\frac{1}{2}}\\
x_2' &= x_2\left(\frac{(1-t)t'}{t(1-t')}\right)^{\frac{1}{2}}\\
x_3' &= \frac{1}{\varepsilon_5(x_1,\ldots,x_4)}\left(\frac{(1-t)(1-t')}{tt'}\right)^{\frac{1}{2}}\\
x_4' &= x_4\left(\frac{(1-t)t'}{t(1-t')}\right)^{\frac{1}{2}}
\end{aligned}$$

De plus, cet automorphisme change la 3-forme $\omega_3\wedge\mathrm{d}t$ sur $\Omega_3\times]0,1[$ en son opposé, on a :

$$\begin{aligned}
\varepsilon_5(x_1',\ldots,x_4') &= \frac{1}{x_3}\left(\frac{(1-t)(1-t')}{tt'}\right)^{\frac{1}{2}}\\
\varepsilon_6(x_1',\ldots,x_4') &= \varepsilon_6(x_1,\ldots,x_4)\left(\frac{t(1-t)}{t'(1-t')}\right)^{\frac{1}{2}}
\end{aligned}$$

Démonstration : Il s'agit du conjugué, par la bijection de la proposition 2.5, de l'automorphisme de $\tilde{\mathcal{V}}_3$ qui échange y_3 et $\frac{1}{y_5}$. On peut aussi démontrer ce corollaire par un calcul direct, sans utiliser la variété $\tilde{\mathcal{V}}_3$.

Théorème 2.7 *Supposons que les dix entiers $h,j,k,l,m,q,r,s,q+h-r,r+j-s$ associés à $\alpha_1,\ldots,\alpha_6$ sont positifs. Alors on a :*

$$(q+h-r)!\,(r+l-q)!\,J(\alpha_1,\ldots,\alpha_6)=h!\,l!\,J(\alpha_1,\alpha_2,-\alpha_5,\alpha_4,-\alpha_3,\alpha_6).$$

Ce théorème est crucial dans la méthode de Rhin-Viola. Il s'agit de la formule notée (4.1) dans [R4] ; elle y est déduite de la représentation intégrale d'Euler de la fonction hypergéométrique de Gauss :

$$\frac{1}{h!l!}I(h,j,k,l,m,q,r,s)=\frac{1}{(q+h-r)!(r+l-q)!}I(q+h-r,j,k,r+l-q,m,r,q,s).$$

Démonstration du théorème 2.7 : Le membre de gauche dans l'énoncé du théorème s'écrit, d'après le théorème de Fubini et la relation $\frac{a!b!}{(a+b+1)!}=\int_0^1 t^a(1-t)^b\mathrm{d}t$:

$$\Gamma(\alpha_1+\alpha_6)\int_{\Omega_3\times]0,1[}t^{\frac{1}{2}(\alpha_1-\alpha_2+\alpha_3-\alpha_4+\alpha_5+\alpha_6)}(1-t)^{\frac{1}{2}(\alpha_1+\alpha_2-\alpha_3+\alpha_4-\alpha_5+\alpha_6)}x_1^{\alpha_1}x_2^{\alpha_2}x_3^{\alpha_3}x_4^{\alpha_4}\varepsilon_5(\underline{x})^{\alpha_5}\varepsilon_6(\underline{x})^{\alpha_6}\omega_3\wedge\mathrm{d}t.$$

La proposition 2.5 permet de transformer cette intégrale en l'intégrale suivante (dans laquelle l'orientation de $\tilde{\Omega}_3$ vient de sa paramétrisation par $]0,1[^3$) :

$$\Gamma(\alpha_1+\alpha_6)\int_{\tilde{\Omega}_3} y_1^{\alpha_1}y_2^{\alpha_2}y_3^{\alpha_3}y_4^{\alpha_4}y_5^{\alpha_5}y_6^{\alpha_6}\tilde{\omega}_3.$$

La proposition 2.4 et le changement de variables qui échange y_3 et $\frac{1}{y_5}$ (donc y_5 et $\frac{1}{y_3}$) en fixant y_1, y_2, y_4 et y_6 prouvent que cette intégrale est invariante quand on échange α_3 et $-\alpha_5$ en fixant α_1, α_2, α_4 et α_6. Donc $(q+h-r)!(r+l-q)!J(\alpha_1,\ldots,\alpha_6)$ l'est aussi, ce qui termine la preuve du théorème 2.7.

2.3 Le groupe de Rhin-Viola comme groupe d'automorphismes

Notons U_3 le simplexe de dimension 7 défini par :

$$U_3=\{(t_1,\ldots,t_8)\in]0,1[^8,\ t_1+\ldots+t_8=1\}.$$

On considère sur U_3 la 7-forme différentielle η_3 définie par :

$$\eta_3=\mathrm{d}t_1\wedge\ldots\wedge\mathrm{d}t_7,$$

dont on voit qu'elle est invariante au signe près sous l'action naturelle de $\mathcal{D}_8$ sur U_3.

Sur $\Omega_3\times U_3$ on définit des fonctions $z_1,\ldots,z_6$ qui à tout point $(\underline{x},\underline{t})=(x_1,\ldots,x_4,t_1,\ldots,t_8)$ associent :

$$\begin{aligned}
z_1(\underline{x},\underline{t}) &= x_1\left(\frac{t_1t_4t_6t_7}{t_5t_8t_2t_3}\right)^{\frac{1}{2}}\\
z_2(\underline{x},\underline{t}) &= x_2\left(\frac{t_2t_5t_7t_8}{t_6t_1t_3t_4}\right)^{\frac{1}{2}}\\
z_3(\underline{x},\underline{t}) &= x_3\left(\frac{t_3t_6t_8t_1}{t_7t_2t_4t_5}\right)^{\frac{1}{2}}\\
z_4(\underline{x},\underline{t}) &= x_4\left(\frac{t_4t_7t_1t_2}{t_8t_3t_5t_6}\right)^{\frac{1}{2}}\\
z_5(\underline{x},\underline{t}) &= \varepsilon_5(x_1,\ldots,x_4)\left(\frac{t_2t_4t_6t_8}{t_1t_3t_5t_7}\right)^{\frac{1}{2}}\\
z_6(\underline{x},\underline{t}) &= \varepsilon_6(x_1,\ldots,x_4)(t_1t_2\ldots t_8)^{\frac{1}{2}}
\end{aligned}$$

On peut résumer les quatre premières équations sous la forme suivante, pour $i\in\{1,\ldots,4\}$:

$$z_i(\underline{x},\underline{t})=x_i\left(\frac{t_it_{i+3}t_{i+5}t_{i+6}}{t_{i+4}t_{i+7}t_{i+1}t_{i+2}}\right)^{\frac{1}{2}}.$$

Notons Φ le groupe de Rhin-Viola, qui est isomorphe à $K\rtimes\mathfrak{S}_5$ où K est l'hyperplan d'équation $\varepsilon_1+\ldots+\varepsilon_5=0$ dans $(\mathbb{Z}/2\mathbb{Z})^5$ (voir le paragraphe 2.4). On le fait agir à gauche sur $\{z_1,\ldots,z_5,\frac{1}{z_1},\ldots,\frac{1}{z_5}\}$ de la manière suivante. Si $\gamma\in\Phi$ envoie α_i sur $\varepsilon_i\alpha_{\gamma^*(i)}$, avec $i\in\{1,\ldots,5\}$ et $\varepsilon_i\in\{-1,1\}$, alors on pose $\gamma\cdot z_i=(z_{\gamma^*(i)})^{\varepsilon_i}$. Par ailleurs on pose $\gamma\cdot z_6=z_6$.

Théorème 2.8 *Il existe un groupe H d'automorphismes de $\Omega_3 \times U_3$ (qui fixent, au signe près, $\omega_3 \wedge \eta_3$) et un homomorphisme de groupes surjectif $\pi : H \to \Phi$ tels que, pour tout $f \in H$ et tout $i \in \{1, \dots, 6\}$, on ait :*

$$z_i \circ f = \pi(f)^{-1} \cdot z_i.$$

N.B. Les éléments de H sont des fonctions algébriques (au sens où elles sont définies par composition de fractions rationnelles et d'extractions de racines).

Corollaire 2.9 *La quantité suivante ne dépend pas du choix de $\gamma \in \Phi$:*

$$\frac{J(\gamma(\alpha_1), \dots, \gamma(\alpha_6))}{\gamma(h)!\gamma(j)!\gamma(k)!\gamma(l)!\gamma(m)!\gamma(q)!\gamma(r)!\gamma(s)!}.$$

N.B. Comme Φ est engendré par $\mathcal{D}_8$ et φ (voir le paragraphe 2.4), on peut déduire directement ce corollaire du corollaire 2.3 (qui donne l'invariance par $\mathcal{D}_8$) et du théorème 2.7 (qui donne l'invariance par φ). La démonstration qui suit fait le lien avec les automorphismes de $\Omega_3 \times U_3$.

Démonstration du corollaire 2.9 : Le théorème 2.8 montre que la quantité suivante est invariante sous l'action de Φ :

$$\begin{aligned} & (4\alpha_6 + 7)! \int_{\Omega_3 \times U_3} z_1^{\alpha_1} z_2^{\alpha_2} z_3^{\alpha_3} z_4^{\alpha_4} z_5^{\alpha_5} z_6^{\alpha_6} \omega_3 \wedge \eta_3 \\ = \; & J(\alpha_1, \dots, \alpha_6) \int_{]0,+\infty[\times U_3} e^{-u} u^{4\alpha_6} t_1^{h'} t_2^{j'} t_3^{k'} t_4^{l'} t_5^{m'} t_6^{q'} t_7^{r'} t_8^{s'} u^7 \mathrm{d}u \wedge \eta_3. \end{aligned}$$

En utilisant la relation $h' + \ldots + s' = 4\alpha_6$ et le changement de variables donné par $t_i' = ut_i$ pour tout $i \in \{1, \dots, 8\}$ on obtient :

$$\begin{aligned} & J(\alpha_1, \dots, \alpha_6) \int_{]0,+\infty[^8} e^{-(t_1' + \ldots + t_8')} {t_1'}^{h'} {t_2'}^{j'} {t_3'}^{k'} {t_4'}^{l'} {t_5'}^{m'} {t_6'}^{q'} {t_7'}^{r'} {t_8'}^{s'} \mathrm{d}t_1' \wedge \ldots \wedge \mathrm{d}t_8' \\ = \; & h'!j'!k'!l'!m'!q'!r'!s'!J(\alpha_1, \dots, \alpha_6). \end{aligned}$$

En divisant par la quantité $h'!j'!k'!l'!m'!q'!r'!s'!h!j!k!l!m!q!r!s!$, qui est elle aussi invariante par Φ, on conclut la démonstration du corollaire.

Démonstration du théorème 2.8 : Pour $\gamma \in \mathcal{D}_8$, notons f_γ l'automorphisme donné par l'action diagonale de $\gamma^{-1} \in \mathcal{D}_8$ sur $\Omega_3 \times U_3$. Pour φ, on définit $(X_1, \dots, X_4, T_1, \dots, T_8) = f_\varphi(x_1, \dots, x_4, t_1, \dots, t_8)$ par les formules suivantes, directement inspirées de celles du corollaire 2.6 et de l'action de φ sur $h', \ldots, s'$:

$$\begin{aligned}
T_i &= t_i \text{ pour } i \in \{1, \ldots, 4\} \\
T_5 &= t_8 \\
T_6 &= (1 - x_1 \mathfrak{a}(\underline{x}))t_6 + \mathfrak{b}(\underline{x})t_7 \\
T_7 &= x_1 \mathfrak{a}(\underline{x})t_6 + (1 - \mathfrak{b}(\underline{x}))t_7 \\
T_8 &= t_5 \\
X_1 &= x_1 \left(\frac{t_6 t_7}{T_6 T_7}\right)^{\frac{1}{2}} \\
X_2 &= x_2 \left(\frac{t_7 T_6}{t_6 T_7}\right)^{\frac{1}{2}} \\
X_3 &= \frac{1}{\varepsilon_5(x_1, \ldots, x_4)} \left(\frac{t_7 T_7}{t_6 T_6}\right)^{\frac{1}{2}} \\
X_4 &= x_4 \left(\frac{t_7 T_6}{t_6 T_7}\right)^{\frac{1}{2}}
\end{aligned}$$

On en déduit :

$$\begin{aligned}
\varepsilon_5(X_1, \ldots, X_4) &= \frac{1}{x_3} \left(\frac{t_7 T_7}{t_6 T_6}\right)^{\frac{1}{2}} \\
\varepsilon_6(X_1, \ldots, X_4) &= \varepsilon_6(x_1, \ldots, x_4) \left(\frac{t_6 t_7}{T_6 T_7}\right)^{\frac{1}{2}}
\end{aligned}$$

Alors f_φ (respectivement chacun des f_γ pour $\gamma \in \mathcal{D}_8$) est un automorphisme de $\Omega_3 \times U_3$ qui fixe au signe près la forme différentielle $\omega_3 \wedge \eta_3$, et agit sur les fonctions $z_1, \ldots, z_5, \frac{1}{z_1}, \ldots, \frac{1}{z_5}, z_6$ comme φ^{-1} (resp. γ^{-1}), c'est-à-dire qu'on a $z_i \circ f_\varphi = \varphi^{-1} \cdot z_i$ (resp. $z_i \circ f_\gamma = \gamma^{-1} \cdot z_i$) pour tout $i \in \{1, \ldots, 6\}$. Notons H le groupe d'automorphismes engendré par f_φ et les f_γ. Chaque élément f de H agit sur les fonctions $z_1, \ldots, z_5, \frac{1}{z_1}, \ldots, \frac{1}{z_5}, z_6$ comme un unique élément de Φ, noté $\pi(f)^{-1}$. Cela termine la preuve du théorème 2.8.

Question : L'homomorphisme π ainsi construit est-il injectif ?

2.4 Lien avec la présentation de Rhin-Viola

Dans ce paragraphe, on considère alternativement $\alpha_1, \ldots, \alpha_6$ (ainsi que leurs combinaisons linéaires) comme des formes linéaires sur $\mathbb{R}^6$ ou comme des nombres réels. Dans ce deuxième cas, on suppose que les seize réels $h, j, k, l, m, q, r, s, h+l-j, j+m-k, k+q-l, l+r-m, m+s-q, q+h-r, r+j-s, s+k-h$ sont des entiers positifs. Toutes les actions considérées sont des actions à gauche.

Notons Φ le groupe formé par les permutations paires γ de l'ensemble à dix éléments $(\alpha_1, \ldots, \alpha_5, -\alpha_1, \ldots, -\alpha_5)$ qui vérifient $\gamma \cdot (-\alpha_i) = -\gamma \cdot \alpha_i$ pour tout $i \in \{1, \ldots, 5\}$. Tout tel élément induit une permutation γ^* de l'ensemble formé par les cinq paires $\{\alpha_1, -\alpha_1\}$, ..., $\{\alpha_5, -\alpha_5\}$. On définit ainsi un morphisme de groupes surjectif de Φ dans $\mathfrak{S}_5$, dont le noyau K est isomorphe à l'hyperplan d'équation $\varepsilon_1 + \ldots + \varepsilon_5 = 0$ dans $(\mathbb{Z}/2\mathbb{Z})^5$: les éléments de K envoient chaque α_i sur $(-1)^{\varepsilon_i}\alpha_i$, où $\varepsilon_1, \ldots, \varepsilon_5 \in \{0, 1\}$ sont de somme paire.

On a une section de $\mathfrak{S}_5$ dans Φ, qui permet d'identifier Φ au produit semi-direct $K \rtimes \mathfrak{S}_5$ relatif à l'action naturelle de $\mathfrak{S}_5$ sur K. Tout élément $\gamma \in \Phi$ agit sur $(\alpha_1, \dots, \alpha_5, -\alpha_1, \dots, -\alpha_5)$ par une permutation des indices $1, 2, 3, 4, 5$, composée avec un changement de signe sur un nombre pair d'indices $i \in \{1, \dots, 5\}$.

Cette action induit, par linéarité, une action de Φ sur l'ensemble $\mathcal{A}_1$ des sommes de la forme $\varepsilon_1\alpha_1 + \dots + \varepsilon_5\alpha_5$ avec $\varepsilon_1, \dots, \varepsilon_5 \in \{-1, 1\}$. Sous cette action, l'ensemble $\mathcal{A}_1$ se décompose en deux orbites, chacune de cardinal 16 ; celle qui servira dans la suite est l'ensemble $\mathcal{A}_2$ des sommes $\varepsilon_1\alpha_1 + \dots + \varepsilon_5\alpha_5$ où le produit des ε_i vaut 1. Cette Φ-orbite se scinde en deux $\mathcal{D}_8$-orbites[2] $\mathcal{A}_3$ et $\mathcal{A}_4$, chacune de cardinal 8.

Pour décrire ces $\mathcal{D}_8$-orbites, on remarque qu'un élément $\varepsilon_1\alpha_1 + \dots + \varepsilon_5\alpha_5$ de $\mathcal{A}_2$ est déterminé de manière unique par $\varepsilon_1, \dots, \varepsilon_4$, qui sont quelconques dans $\{-1, 1\}$. Un élément de $\mathcal{A}_2$ peut donc être vu comme une somme de 4 parmi les 8 termes $\alpha_1, \alpha_2, \alpha_3, \alpha_4, -\alpha_1, -\alpha_2, -\alpha_3, -\alpha_4$, de telle sorte que pour chaque indice i l'un exactement parmi α_i et $-\alpha_i$ fasse partie des 4 termes retenus. Cette interprétation est compatible avec l'action de $\mathcal{D}_8$ sur l'octogone régulier dont les sommets sont les huit termes ci-dessus. Les éléments de $\mathcal{A}_3$ correspondent alors aux choix de 4 sommets consécutifs de cet octogone.

Les huit éléments de $\mathcal{A}_3$, quand on leur ajoute α_6 et qu'on les divise par 2 (ce qui ne change rien à l'action de Φ, qui laisse α_6 invariant), sont exactement les huit nombres h, j, k, l, m, q, r, s définis ci-dessus. L'action de $\mathcal{D}_8$ induite sur $h, \dots, s$ est celle dans laquelle la rotation d'un huitième de tour envoie h sur j, j sur k, et ainsi de suite. Quant aux huit éléments de $\mathcal{A}_4$, ils correspondent aux "entiers auxiliaires" $h', j', k', l', m', q', r', s'$ (voir le paragraphe 2.1).

A l'inverse, on peut retrouver les α_i à partir de $\beta_1 = h, \dots, \beta_8 = s$ par les formules suivantes (où on pose $\beta_9 = \beta_1$, $\beta_{10} = \beta_2$ et ainsi de suite) :

$$\begin{aligned}
\alpha_i &= \frac{1}{2}(\beta_i + \beta_{i+3} - \beta_{i+4} - \beta_{i+7}) \text{ pour } 1 \le i \le 4;\\
\alpha_5 &= \frac{(-1)^i}{2}(\beta_i + \beta_{i+4} - \beta_{i+1} - \beta_{i+5}) \text{ est indépendant de } i \in \{1, \dots, 8\};\\
\alpha_6 &= \frac{1}{2}(\beta_i + \beta_{i+4} + \beta_{i+1} + \beta_{i+5}) \text{ est indépendant de } i \in \{1, \dots, 8\}.
\end{aligned}$$

L'action de Φ sur $\alpha_1, \dots, \alpha_5, -\alpha_1, \dots, -\alpha_5$ correspond au plongement de Φ dans $\mathfrak{A}_{10}$ donné dans [R4] (quatrième partie), puisque l'action sur α_6 est triviale et qu'on a les relations suivantes :

$$\begin{aligned}
\alpha_1 + \alpha_6 &= h + l, \text{ noté } u_1 \text{ dans [R4]}\\
\alpha_2 + \alpha_6 &= j + m, \text{ noté } u_2\\
\alpha_3 + \alpha_6 &= k + q, \text{ noté } u_3\\
\alpha_4 + \alpha_6 &= l + r, \text{ noté } u_4\\
-\alpha_1 + \alpha_6 &= m + s, \text{ noté } u_5\\
-\alpha_2 + \alpha_6 &= q + h, \text{ noté } u_6\\
-\alpha_3 + \alpha_6 &= r + j, \text{ noté } u_7\\
-\alpha_4 + \alpha_6 &= s + k, \text{ noté } u_8\\
\alpha_5 + \alpha_6 &= j + q = l + s, \text{ noté } u_9\\
-\alpha_5 + \alpha_6 &= k + r = h + m, \text{ noté } u_{10}
\end{aligned}$$

[2]Le groupe diédral $\mathcal{D}_8$ se plonge dans Φ grâce à l'action de $\mathcal{D}_8$ sur $\{\alpha_1, \dots, \alpha_5, -\alpha_1, \dots, -\alpha_5\}$ définie au paragraphe 2.1.

Adoptons les notations suivantes, cohérentes avec celles de [R4] et celles introduites cidessus. On note φ l'élément de Φ qui fixe α_1, α_2 et α_4, et échange α_3 et $-\alpha_5$ (donc aussi $-\alpha_3$ et α_5). On note ϑ l'élément de $\mathcal{D}_8$ qui correspond à la rotation d'un huitième de tour dans le sens direct : ϑ envoie α_i sur α_{i+1} pour $i \in \{1,2,3\}$, α_4 sur $-\alpha_1$ et α_5 sur $-\alpha_5$. Enfin, on note σ l'élément de $\mathcal{D}_8$ qui échange α_1 et $-\alpha_4$, ainsi que α_2 et $-\alpha_3$; on a $\sigma \cdot \alpha_5 = \alpha_5$.

Il est crucial de remarquer que Φ est engendré par $\mathcal{D}_8$ et φ. Une démonstration de ce fait se trouve dans [R4] ; on la reproduit brièvement ici. Tout d'abord, on remarque que les éléments ϑ^4, $\vartheta\varphi\vartheta^4\varphi\vartheta^7$, $\varphi\vartheta^4\varphi$ et $\vartheta^3\varphi\vartheta^4\varphi\vartheta^5$ appartiennent à K, et agissent chacun en changeant le signe de quatre parmi les cinq α_i. On en déduit immédiatement que K est inclus dans le sous-groupe de Φ engendré par $\mathcal{D}_8$ et φ. En outre, ce sous-groupe est envoyé surjectivement sur $\mathfrak{S}_5$, car $\vartheta^*\varphi^*$ est un 5-cycle et φ^* est une transposition. Cela montre que ce sous-groupe est Φ tout entier.

À $\alpha_1, \dots, \alpha_6$ et à tout $\gamma \in \Phi$ on associe le quotient de factorielles suivant :

$$\xi(\gamma; \alpha_1, \dots, \alpha_6) = \frac{h!j!k!l!m!q!r!s!}{\gamma(h)!\gamma(j)!\gamma(k)!\gamma(l)!\gamma(m)!\gamma(q)!\gamma(r)!\gamma(s)!}.$$

On a alors, pour tout $\gamma \in \Phi$:

$$J(\alpha_1, \dots, \alpha_6) = \xi(\gamma; \alpha_1, \dots, \alpha_6) J(\gamma(\alpha_1), \dots, \gamma(\alpha_6))$$

Pour $\gamma \in \Phi$, il peut y avoir des simplifications dans le quotient $\xi(\gamma; \alpha_1, \dots, \alpha_6)$. Rhin et Viola appellent *niveau* de γ le nombre de factorielles présentes au numérateur (ou au dénominateur) après ces simplifications (sans tenir compte des simplifications dues à un choix particulier des paramètres $\alpha_1, \dots, \alpha_6$). C'est le nombre d'éléments de $\mathcal{A}_3$ que γ envoie dans $\mathcal{A}_4$: le numérateur est formé par les factorielles des éléments de $\mathcal{A}_3 \setminus \gamma \cdot \mathcal{A}_3$, et le dénominateur par les factorielles des éléments de $\mathcal{A}_4 \cap \gamma \cdot \mathcal{A}_3$. Comme $\mathcal{A}_3$ et $\mathcal{A}_4$ sont stables sous l'action de $\mathcal{D}_8$, le niveau de γ ne dépend que de la double classe $\mathcal{D}_8 \gamma \mathcal{D}_8$. En particulier, les éléments de $\mathcal{D}_8$ sont de niveau 0, et ceux de $\mathcal{D}_8 \varphi \mathcal{D}_8$ sont de niveau 2. Plus généralement, tout élément γ s'écrit comme un mot en les trois lettres φ, ϑ et σ ; le niveau de γ est alors inférieur ou égal au double du nombre minimal de φ qui apparaissent dans une telle écriture.

Le choix de paramètres utilisé dans [R4] pour obtenir la meilleure mesure d'irrationalité connue est $h = 16$, $j = 17$, $k = 19$, $l = 15$, $m = 12$, $q = 11$, $r = 9$, $s = 13$. Il correspond à $\alpha_1 = 3$, $\alpha_2 = 1$, $\alpha_3 = 2$, $\alpha_4 = -4$, $\alpha_5 = 0$, $\alpha_6 = 28$. Au vu de ce choix, on peut essayer de suivre la méthode de Rhin-Viola en utilisant seulement le sous-groupe de Φ formé par les γ qui fixent $\{\alpha_5, -\alpha_5\}$, de manière à ne faire apparaître dans la preuve que des familles $(\alpha_1, \dots, \alpha_6)$ vérifiant $\alpha_5 = 0$. Toutefois, cela donne une mesure d'irrationalité de $\zeta(3)$ strictement moins précise que celle obtenue par Rhin et Viola.

Ce choix de paramètres semble complètement mystérieux. Zudilin a fait de nombreux tests numériques [Z9] et n'a pas trouvé de meilleur choix ; en ce qui concerne l'exposant d'irrationalité de $\log(2)$ (qui peut se traiter parallèlement à $\zeta(2)$ et $\zeta(3)$), Nicolas Brisebarre a essayé aussi [B16] sans plus de succès. Quand on considère les paramètres $\alpha_1, \dots, \alpha_6$, les valeurs correspondant au choix de Rhin-Viola sont plus petites (sauf α_6, qui joue un rôle à part et n'intervient pas dans l'action du groupe), mais pas assez pour rendre “naturel” le choix de ces valeurs. Un autre changement de paramètres figure dans [Z9], ainsi qu'un lien entre les valeurs utilisées pour $\zeta(2)$ et pour $\zeta(3)$.

3 Une généralisation du groupe de Rhin-Viola à une famille d'intégrales n-uples

3.1 Notations

Dans toute la suite de ce texte, on désigne par n un entier supérieur ou égal à deux. Soient $\underline{a} = (a_1, \dots, a_n)$ et $\underline{b} = (b_1, \dots, b_n)$ deux n-uplets d'entiers relatifs, et soit c un entier relatif. On note $\underline{x} = (x_1, \dots, x_n)$ les coordonnées d'un point de $[0,1]^n$, et par convention on pose $x_{n+1} = 0$. Suivant Vasilenko et Vasilyev, on définit des réels $\delta_0(\underline{x}), \dots, \delta_n(\underline{x})$ compris entre 0 et 1 en posant $\delta_0(\underline{x}) = 1$ et $\delta_{k+1}(\underline{x}) = 1 - x_{k+1}\delta_k(\underline{x})$ pour $k \in \{0, \dots, n-1\}$. Enfin, on pose $\delta_{n+1}(\underline{x}) = 1$. On peut aussi définir les $\delta_k(\underline{x})$ par la formule suivante, pour $k \in \{1, \dots, n\}$:

$$\delta_k(\underline{x}) = \sum_{j=0}^{k} (-1)^{k-j} x_{j+1} \cdot \ldots \cdot x_k,$$

ou bien comme dans l'introduction par :

$$\delta_k(x_1, \dots, x_n) = \delta_k(x_1, \dots, x_k) = 1 - x_k(1 - x_{k-1}(1 - \ldots (1 - x_1))).$$

On a alors la propriété suivante :

$$\delta_{k+2} - x_{k+1}\delta_k(\underline{x}) = (1 - x_{k+2})\delta_{k+1}(\underline{x}) \text{ pour tout } k \in \{0, \dots, n-2\}. \tag{4.31}$$

On note ω la n-forme différentielle sur $[0,1]^n$ définie par $\omega = \frac{\mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n}{\delta_n(\underline{x})}$.

On considère, pour $n \geq 2$, des intégrales (éventuellement infinies) :

$$I(\underline{a}, \underline{b}, c) = \int_{[0,1]^n} \delta_n(\underline{x})^{-c} \prod_{k=1}^{n} x_k^{a_k} (1 - x_k)^{b_k} \omega. \tag{4.32}$$

Proposition 3.1 *Posons, pour tout $k \in \{1, \dots, n\}$ qui a la même parité que n, avec la convention $a_0 = -1$:*

$$\tilde{a}_k = a_{k-1} + \frac{n-k}{2} + \sum_{\substack{j \in \{k, \dots, n\} \\ j \equiv n \mod 2}} b_j.$$

Alors l'intégrale $I(\underline{a}, \underline{b}, c)$ est finie si, et seulement si, les conditions suivantes sont réalisées :

$$c \leq \min_{\substack{k \in \{1, \dots, n\} \\ k \equiv n \mod 2}} \tilde{a}_k \text{ et, pour tout } k \in \{1, \dots, n\},\ a_k \geq 0 \text{ et } b_k \geq 0. \tag{4.33}$$

Cette proposition est un cas particulier de la proposition 4.1 qui est démontrée dans la quatrième partie de ce texte.

Pour retrouver la formulation classique de ces intégrales (par exemple pour $n = 2$ et $n = 3$), il convient de poser $x = 1 - x_1$, $y = x_2$, $z = x_3$ et ainsi de suite (on peut aussi inverser x et y). On a alors $\delta_2(\underline{x}) = 1 - xy$ et $\delta_3(\underline{x}) = 1 - z(1 - xy)$; l'utilisation de $1 - x$ comme première variable permet d'avoir des formules plus symétriques dans la définition des $\delta_k(\underline{x})$ donnée ci-dessus et dans celle de la transformation ψ ci-dessous.

Comme dans l'ensemble de ce texte, on suppose dans cette partie que les exposants (ici $a_1, \dots, a_n, b_1, \dots, b_n, c$) sont entiers, mais tous les résultats (sauf bien sûr ceux du paragraphe 3.5) se généralisent au cas où ils sont complexes.

3.2 Définition des transformations

Pour tout automorphisme g de $\mathbb{Z}^{2n+1}$, on note :

$$g(a_1,\dots,a_n,b_1,\dots,b_n,c)=(\bar{g}(a_1),\dots,\bar{g}(a_n),\bar{g}(b_1),\dots,\bar{g}(b_n),\bar{g}(c)).$$

Alors $\bar{g}$ est essentiellement la transposée de g ; l'anti-isomorphisme qui à g associe $\bar{g}$ est le même que celui qui consiste à mettre en caractères gras dans les articles de Rhin-Viola. La notation $\bar{g}$ signifie que l'on pense à g comme à une permutation des lettres $a_1,\dots,a_n,b_1,\dots,b_n,c$ (ou, plus généralement, à une application qui envoie chacune de ces lettres sur une combinaison linéaire de lettres).

Tout sous-groupe G de $\mathrm{GL}_{2n+1}(\mathbb{Z})$ agit naturellement à gauche sur les $(2n+1)$-uplets $(\underline{a},\underline{b},c)$ vus comme points de $\mathbb{Z}^{2n+1}$, et à droite sur les mêmes $(2n+1)$-uplets vus comme familles de formes linéaires sur $\mathbb{Z}^{2n+1}$. On note $\bar{g}$ quand on considère cette deuxième action ; on a $\overline{gg'}=\overline{g'}\,\overline{g}$, si bien que $\bar{g}$ agit à gauche.

N.B. La même distinction entre g et $\bar{g}$ est également utilisée dans les quatrième et cinquième parties de ce texte. Dans les deux premières parties, on notait simplement g les éléments du groupe de Rhin-Viola, dont on considérait des actions à droite ou à gauche.

Proposition 3.2 *Notons σ l'automorphisme de $\mathbb{Z}^{2n+1}$ tel que $\bar{\sigma}$ échange a_1 et b_2, ainsi que a_2 et b_1, et fixe les autres coordonnées. Alors on a, pour tout $(\underline{a},\underline{b},c)\in\mathbb{Z}^{2n+1}$:*

$$I(\underline{a},\underline{b},c)=I(\sigma(\underline{a},\underline{b},c)).$$

N.B. En particulier, si $(\underline{a},\underline{b},c)$ vérifie les conditions (4.33) alors $\sigma(\underline{a},\underline{b},c)$ aussi.

Démonstration de la proposition 3.2 : Considérons le changement de variables (encore noté σ) de $[0,1]^n$ dans lui-même qui envoie x_1 sur $1-x_2$, x_2 sur $1-x_1$ et qui fixe x_k pour tout $k\in\{3,\dots,n\}$. Alors σ change la forme différentielle ω en son opposé ; on en déduit immédiatement le résultat, grâce au théorème de changement de variables dans une intégrale multiple (voir [C1], chapitre 1, paragraphe 4.10).

N.B. Le changement de variables σ est involutif, préserve le facteur $\frac{\prod_{k=1}^n x_k(1-x_k)}{\delta_n(\underline{x})}$, et vérifie $\delta_1(\sigma(\underline{x}))=x_2$ et $\delta_k(\sigma(\underline{x}))=\delta_k(\underline{x})$ pour tout $k\in\{0,\dots,n+1\}\setminus\{1\}$. En fonction des variables $x,y,\dots$, utilisées couramment, il s'agit de la transformation qui échange x et y en fixant les autres variables.

Proposition 3.3 *Notons ψ l'automorphisme de $\mathbb{Z}^{2n+1}$ défini par :*

$$\begin{aligned}\bar{\psi}(a_k) &= a_{n+1-k} \text{ pour tout } k\in\{1,\dots,n\}\\ \bar{\psi}(b_1) &= a_{n-1}+b_n-c\\ \bar{\psi}(b_k) &= b_{n+2-k} \text{ pour tout } k\in\{2,\dots,n\}\\ \bar{\psi}(c) &= a_2+b_2-b_1\end{aligned}$$

Soit $(\underline{a},\underline{b},c)\in\mathbb{Z}^{2n+1}$ un triplet tel que :

$$a_k+b_{k+1}=a_{k+2}+b_{k+2} \text{ pour tout } k\in\{1,\dots,n-2\}, \tag{4.34}$$

cette condition étant supposée remplie si $n=2$. Alors cette condition est aussi vérifiée par $\psi(\underline{a},\underline{b},c)$, et on a :

$$I(\underline{a},\underline{b},c)=I(\psi(\underline{a},\underline{b},c)).$$

N.B. En particulier, si $(\underline{a},\underline{b},c)$ est un tel triplet qui vérifie les conditions (4.33) alors $\psi(\underline{a},\underline{b},c)$ aussi.

N.B. L'hypothèse (4.34) signifie que $a_k + a_{k+1} + b_{k+1}$ ne dépend pas de $k \in \{1,\dots,n-1\}$.

Démonstration : Notons ψ l'application de $[0,1]^n$ dans lui-même qui à $\underline{x} = (x_1,\dots,x_n)$ associe $\psi(\underline{x}) = (X_1,\dots,X_n)$ défini par :

$$X_k = \frac{x_{n+1-k}\delta_{n-k}(\underline{x})}{\delta_{n+2-k}(\underline{x})} \text{ pour tout } k \in \{1,\dots,n\}.$$

La relation (4.31) et les conventions $\delta_{n+1}(\underline{x}) = 1$ et $x_{n+1} = 0$ montrent qu'on a :

$$1 - X_k = \frac{(1-x_{n+2-k})\delta_{n+1-k}(\underline{x})}{\delta_{n+2-k}(\underline{x})} \text{ pour tout } k \in \{1,\dots,n\}.$$

En outre, on voit que

$$\delta_k(\psi(\underline{x})) = \delta_{n+1-k}(\underline{x}) \text{ pour tout } k \in \{0,\dots,n+1\}.$$

Il est alors immédiat de vérifier que ψ transforme l'intégrande de $I(\underline{a},\underline{b},c)$ en celle de $I(\psi(\underline{a},\underline{b},c))$. La condition (4.34) assure que les facteurs $\delta_2(\underline{x}),\dots,\delta_{n-1}(\underline{x})$ apparaissent avec un exposant nul. Pour conclure la démonstration de la proposition 3.3, il suffit de montrer que ψ laisse stable la forme différentielle ω, au signe près. Cela découle immédiatement de la relation

$$\mathrm{d}\delta_1 \wedge \dots \wedge \mathrm{d}\delta_k = (-1)^k \delta_1(\underline{x}) \cdot \dots \cdot \delta_{k-1}(\underline{x}) \mathrm{d}x_1 \wedge \dots \wedge \mathrm{d}x_k, \tag{4.35}$$

valable pour tout $k \in \{1,\dots,n\}$, et du lemme suivant (utilisé avec $k = n-1$), dans lequel on considère les δ_j, ainsi que les coordonnées X_1, …, X_n de $\psi(\underline{x})$, comme des fonctions de $\underline{x}$:

Lemme 3.4 *Pour tout $k \in \{0,\dots,n-1\}$ on a :*

$$\delta_n(\underline{x})\delta_{n-1}(\underline{x})\dots\delta_{n-k+1}(\underline{x})\mathrm{d}X_n \wedge \dots \wedge \mathrm{d}X_1 = (-1)^{k+1}\mathrm{d}X_n \wedge \dots \wedge \mathrm{d}X_{k+2} \wedge \mathrm{d}\delta_{n-k} \wedge \dots \wedge \mathrm{d}\delta_n.$$

Démonstration du lemme : Pour $k = 0$, cette formule provient de la relation $1 - X_1 = \delta_n(\underline{x})$. Supposons que cette formule est vraie pour un indice $k \in \{0,\dots,n-2\}$, et démontrons-la pour $k+1$. Les $n-k$ fonctions $X_n,\dots,X_{k+3},\delta_{n-k}(\underline{x})X_{k+2},\delta_{n-k-1}(\underline{x})$ ne dépendent que des $n-k-1$ variables $x_1,\dots,x_{n-k-1}$, donc on a :

$$\delta_{n-k}(\underline{x})\mathrm{d}X_n \wedge \dots \wedge \mathrm{d}X_{k+2} \wedge \mathrm{d}\delta_{n-k-1} = -X_{k+2}\mathrm{d}X_n \wedge \dots \wedge \mathrm{d}X_{k+3} \wedge \mathrm{d}\delta_{n-k} \wedge \mathrm{d}\delta_{n-k-1}.$$

En écrivant $\delta_{n-k-1}(\underline{x}) = 1 - X_{k+2}\delta_{n-k}(\underline{x})$, on en déduit :

$$\delta_{n-k}(\underline{x})\mathrm{d}X_n \wedge \dots \wedge \mathrm{d}X_{k+2} \wedge \mathrm{d}\delta_{n-k} = -\mathrm{d}X_n \wedge \dots \wedge \mathrm{d}X_{k+3} \wedge \mathrm{d}\delta_{n-k-1} \wedge \mathrm{d}\delta_{n-k}.$$

Cette relation permet de terminer la démonstration du lemme par récurrence.

N.B. Ce changement de variables ψ est involutif, et préserve le facteur $\frac{\prod_{k=1}^n x_k(1-x_k)}{\delta_n(\underline{x})}$.

N.B. L'hypothèse (4.34) est nécessaire et suffisante pour que le changement de variables ψ transforme $I(\underline{a},\underline{b},c)$ en une intégrale de la même forme. Sans cette hypothèse, des facteurs $\delta_2(\underline{x}),\dots,\delta_{n-1}(\underline{x})$ apparaissent au dénominateur ; c'est ce qui conduit à introduire la famille plus générale étudiée dans la quatrième partie de ce texte.

Il est facile de calculer l'ordre de la transformation $\sigma \circ \psi$; on voit que c'est cinq si $n = 2$, et quatre si $n \geq 3$. Comme σ et ψ sont involutives, elles engendrent un groupe diédral d'ordre 10 ou 8 (respectivement). De plus, $\psi\sigma\psi$ est le changement de variables qui fixe les $n-1$ premières coordonnées et remplace x_n par $\frac{1-x_n}{\delta_n(\underline{x})}$. On retrouve ainsi, pour $n = 3$, le changement de variables utilisé par Beukers [B4].

Proposition 3.5 *Notons φ l'automorphisme de $\mathbb{Z}^{2n+1}$ défini par :*

$$\begin{aligned}
\bar{\varphi}(a_k) &= a_k \text{ pour tout } k \in \{1, \dots, n-2\} \cup \{n\} \\
\bar{\varphi}(b_k) &= b_k \text{ pour tout } k \in \{1, \dots, n-2\} \\
\bar{\varphi}(a_{n-1}) &= c \\
\bar{\varphi}(b_{n-1}) &= b_{n-1} + a_{n-1} - c \\
\bar{\varphi}(b_n) &= b_n + a_{n-1} - c \\
\bar{\varphi}(c) &= a_{n-1}
\end{aligned}$$

Alors on a, pour tout $(\underline{a}, \underline{b}, c) \in \mathbb{Z}^{2n+1}$ qui vérifie les conditions (4.33) et tel que c et $a_{n-1} + b_{n-1} - c$ soient positifs :

$$I(\underline{a}, \underline{b}, c) = \frac{a_{n-1}!b_{n-1}!}{c!(a_{n-1} + b_{n-1} - c)!} I(\varphi(\underline{a}, \underline{b}, c)).$$

DÉMONSTRATION : Supposons que les conditions (4.33) sont réalisées, et que c et $a_{n-1}+b_{n-1}-c$ sont positifs. On a pour tout $\beta \in \mathbb{C} \setminus]-\infty, -1]$:

$$\int_0^1 \frac{x_{n-1}^{a_{n-1}}(1-x_{n-1})^{b_{n-1}}}{(1+\beta x_{n-1})^{c+1}} \mathrm{d}x_{n-1} = \frac{a_{n-1}!b_{n-1}!}{c!(a_{n-1}+b_{n-1}-c)!} \int_0^1 \frac{x_{n-1}^{c}(1-x_{n-1})^{a_{n-1}+b_{n-1}-c}}{(1+\beta x_{n-1})^{a_{n-1}+1}} \mathrm{d}x_{n-1}.$$

Rhin et Viola déduisent ([R3], troisième partie) cette formule de la représentation sous forme d'intégrale eulérienne de la fonction hypergéométrique de Gauss (ils l'utilisent si $|\beta| < 1$, mais elle est vraie pour toute valeur de $\beta \in \mathbb{C} \setminus]-\infty, -1]$ par prolongement analytique). On peut aussi la démontrer par un changement de variables qui figure dans [D3]. On utilise cette formule à $x_1, \dots, x_{n-2}, x_n$ fixés, avec $\beta = \frac{x_n \delta_{n-2}(\underline{x})}{1-x_n}$. On multiplie les deux membres par $x_n^{a_n}(1-x_n)^{b_n-c-1} \prod_{k=1}^{n-2} x_k^{a_k}(1-x_k)^{b_k} \mathrm{d}x_1 \dots \mathrm{d}x_{n-2}\mathrm{d}x_n$ et on intègre de 0 à 1 par rapport à $x_1, \dots, x_{n-2}, x_n$, en remarquant qu'on a $(1-x_n)(1+\beta x_{n-1}) = \delta_n(\underline{x})$. Cela donne le résultat.

3.3 Action du groupe

On souhaite faire agir sur une famille d'intégrales le sous-groupe de $\mathrm{GL}_{2n+1}(\mathbb{Z})$ engendré par σ, ψ et φ. Pour cela, on suppose que $(\underline{a}, \underline{b}, c)$, ainsi que tous les éléments de son orbite sous l'action de ce sous-groupe, vérifient les conditions (4.34) de la proposition 3.3. Ceci revient à faire les hypothèses suivantes sur $(\underline{a}, \underline{b}, c)$:

$$\begin{aligned}
&\text{aucune hypothèse si } n = 2, \\
&a_1 + b_2 = a_3 + b_3 \text{ si } n = 3, \qquad\qquad (4.36)\\
&a_2 = b_1 \text{ et } b_n = c \text{ et } a_k + b_{k+1} = a_{k+2} + b_{k+2} \text{ pour tout } k \in \{1, \dots, n-2\} \text{ si } n \geq 4.
\end{aligned}$$

Notons $\mathcal{E}$ le sous-$\mathbb{Z}$-module de $\mathbb{Z}^{2n+1}$ formé par les $(2n+1)$-uplets vérifiant ces conditions. Il est de rang cinq si $n = 2$, six si $n = 3$, et $n+1$ si $n \geq 4$ (car les n relations qu'on impose si $n \geq 4$ sont indépendantes ; on peut le vérifier directement, mais cela sera démontré au paragraphe 3.4). Pour $(\underline{a}, \underline{b}, c) \in \mathcal{E}$, les conditions (4.33) de convergence se réduisent, si $n \geq 4$, aux conditions suivantes :

$$a_k \geq 0 \text{ et } b_k \geq 0 \text{ pour tout } k \in \{1, \ldots, n\}. \tag{4.37}$$

Comme σ, ψ et φ laissent $\mathcal{E}$ stable, ils induisent des automorphismes de $\mathcal{E}$, encore notés σ, ψ et φ. On note G le sous-groupe de $\mathrm{Aut}(\mathcal{E})$ engendré par ces trois automorphismes. On montre au paragraphe 3.4 que G est isomorphe à $(\mathfrak{S}_3 \times \mathfrak{S}_3) \rtimes \mathbb{Z}/2\mathbb{Z}$ si $n \geq 4$, et qu'on retrouve les situations considérées par Rhin et Viola si $n = 2$ ou $n = 3$. C'est pourquoi on suppose, dans la suite de ce paragraphe, qu'on a $n \geq 4$.

Notons $\mathcal{E}^+$ l'ensemble des $(\underline{a}, \underline{b}, c) \in \mathcal{E}$ dont tous les éléments de la G-orbite vérifient les conditions (4.37) de convergence. La discussion qui suit le théorème 3.6 montre que $(\underline{a}, \underline{b}, c)$ appartient à $\mathcal{E}^+$ si, et seulement si, l'ensemble suivant est formé de nombres positifs :

$$\{a_1, \ldots, a_n, b_1, \ldots, b_n, a_{n-1} + b_{n-1} - c, a_{n-1} + b_{n-1} - a_n, a_2 + b_3 - b_2, a_2 + b_3 - a_1\} \text{ si } n \geq 5,$$
$$\{a_1, \ldots, a_4, b_1, \ldots, b_4, a_3 + b_3 - c, a_3 + b_3 - a_4, a_2 + b_3 - b_2, a_2 + b_3 - a_1, a_1 + a_2 - c,$$
$$a_2 + b_2 - c, a_3 + c - b_2, a_3 + c - a_1\} \text{ si } n = 4.$$

À tout $(\underline{a}, \underline{b}, c) \in \mathcal{E}^+$ et à tout $g \in G$ on associe, si $n \geq 5$:

$$\xi(g; \underline{a}, \underline{b}, c) = \frac{a_{n-1}! a_2! b_{n-1}! b_3!}{(\bar{g}a_{n-1})!(\bar{g}a_2)!(\bar{g}b_{n-1})!(\bar{g}b_3)!}.$$

Si $n = 4$, on ne répète pas le facteur b_3, c'est-à-dire qu'on pose :

$$\xi(g; \underline{a}, \underline{b}, c) = \frac{a_3! a_2! b_3!}{(\bar{g}a_3)!(\bar{g}a_2)!(\bar{g}b_3)!}.$$

Théorème 3.6 *Pour tout $(\underline{a}, \underline{b}, c) \in \mathcal{E}^+$ et tout $g \in G$ on a (avec $n \geq 4$) :*

$$I(\underline{a}, \underline{b}, c) = \xi(g; \underline{a}, \underline{b}, c)\, I(g(\underline{a}, \underline{b}, c)).$$

DÉMONSTRATION : Il suffit de constater que d'après les propositions 3.2, 3.3 et 3.5, le quotient $\frac{I(\underline{a}, \underline{b}, c)}{a_{n-1}! a_2! b_{n-1}! b_3!}$ est invariant (si $n \geq 5$) sous l'action de σ, ψ et φ. Si $n = 4$, il en est de même avec le quotient $\frac{I(\underline{a}, \underline{b}, c)}{a_3! a_2! b_3!}$.

Dans la suite de ce paragraphe, on omet les arguments $\underline{a}, \underline{b}, c$ de ξ.

Pour tout $\gamma \in \mathcal{D}_4$ et tout $g \in G$ on a $\xi(\gamma g) = \xi(g)$ car $\bar{\gamma}$ permute les facteurs du numérateur (par exemple au dénominateur de $\xi(\gamma g)$ on a $\overline{\gamma g}(a_{n-1}) = \overline{g}(\overline{\gamma}(a_{n-1}))$ qui est $\overline{g}(a_{n-1})$ ou $\overline{g}(a_2)$). Donc ξ est constant sur chaque classe de G modulo $\mathcal{D}_4$. On montre au paragraphe 3.4 que G est d'ordre 72 ; il y a donc 9 classes modulo $\mathcal{D}_4$. Les valeurs de ξ en ces 9 classes sont deux à deux distinctes (comme le prouvent les expressions ci-dessous), quand $\underline{a}, \underline{b}, c$ sont suffisamment généraux. Ceci montre que $\mathcal{D}_4$ est exactement l'ensemble des $\gamma \in G$ tels que $\xi(\gamma) = 1$ (pour $\underline{a}, \underline{b}, c$ génériques).

Quand $n \geq 5$, il y a une classe (modulo $\mathcal{D}_4$) de niveau 1, quatre de niveau 2 et quatre de niveau 4 :

$$\begin{aligned}
\xi(\mathrm{Id}) &= 1\\
\xi(\varphi) &= \frac{a_{n-1}!b_{n-1}!}{c!(a_{n-1}+b_{n-1}-c)!}\\
\xi(\varphi\psi) &= \frac{a_2!b_3!}{b_2!(a_2+b_3-b_2)!}\\
\xi(\varphi\psi\sigma) &= \frac{a_2!b_3!}{a_1!(a_2+b_3-a_1)!}\\
\xi(\varphi\psi\sigma\psi) &= \frac{a_{n-1}!b_{n-1}!}{a_n!(a_{n-1}+b_{n-1}-a_n)!}\\
\xi(\varphi\psi\varphi) &= \frac{a_{n-1}!a_2!b_{n-1}!b_3!}{b_2!c!(a_2+b_3-b_2)!(a_{n-1}+b_{n-1}-c)!}\\
\xi(\varphi\psi\varphi\sigma) &= \frac{a_{n-1}!a_2!b_{n-1}!b_3!}{a_1!c!(a_2+b_3-a_1)!(a_{n-1}+b_{n-1}-c)!}\\
\xi(\varphi\psi\varphi\sigma\psi) &= \frac{a_{n-1}!a_2!b_{n-1}!b_3!}{a_n!b_2!(a_{n-1}+b_{n-1}-a_n)!(a_2+b_3-b_2)!}\\
\xi(\varphi\psi\varphi\sigma\psi\sigma) &= \frac{a_{n-1}!a_2!b_{n-1}!b_3!}{a_n!a_1!(a_{n-1}+b_{n-1}-a_n)!(a_2+b_3-a_1)!}
\end{aligned}$$

Quand $n = 4$, il y a une classe (modulo $\mathcal{D}_4$) de niveau 1, quatre de niveau 2 et quatre de niveau 3 :

$$\begin{aligned}
\xi(\mathrm{Id}) &= 1\\
\xi(\varphi) &= \frac{a_3!b_3!}{c!(a_3+b_3-c)!}\\
\xi(\varphi\psi) &= \frac{a_2!b_3!}{b_2!(a_2+b_3-b_2)!}\\
\xi(\varphi\psi\sigma) &= \frac{a_2!b_3!}{a_1!(a_2+b_3-a_1)!}\\
\xi(\varphi\psi\sigma\psi) &= \frac{a_3!b_3!}{a_4!(a_3+b_3-a_4)!}\\
\xi(\varphi\psi\varphi) &= \frac{a_3!a_2!b_3!}{b_2!c!(a_1+a_2-c)!}\\
\xi(\varphi\psi\varphi\sigma) &= \frac{a_3!a_2!b_3!}{a_1!c!(a_2+b_2-c)!}\\
\xi(\varphi\psi\varphi\sigma\psi) &= \frac{a_3!a_2!b_3!}{a_4!b_2!(a_3+c-b_2)!}\\
\xi(\varphi\psi\varphi\sigma\psi\sigma) &= \frac{a_3!a_2!b_3!}{a_4!a_1!(a_3+c-a_1)!}
\end{aligned}$$

Les propriétés de ξ sont les mêmes que dans [R3] et [R4] (on suppose que les paramètres $\underline{a}, \underline{b}, c$ sont génériques) :

– Pour chaque $g \in G$, le numérateur et le dénominateur de $\xi(g)$ sont le produit d'un même nombre de factorielles, appelé *niveau* de g.

- La somme des entiers dont les factorielles apparaissent au numérateur est égale à celle des entiers dont les factorielles apparaissent au dénominateur.
- Les entiers dont les factorielles apparaissent au numérateur appartiennent aux $\mathcal{D}_4$-orbites de a_{n-1} et b_{n-1}. Ceux dont les factorielles apparaissent au dénominateur appartiennent aux G-orbites de a_{n-1} et b_{n-1}, mais pas à leurs $\mathcal{D}_4$-orbites.

Une différence par rapport aux cas $n = 2$ et $n = 3$ est que pour $n \geq 4$ les G-orbites de a_{n-1} et b_{n-1} sont distinctes. En outre, pour $n = 2$ et $n = 3$ la G-orbite de a_{n-1} est la réunion de deux orbites de même cardinal sous l'action du groupe diédral associé, alors qu'ici la situation est plus compliquée. Pour $n \geq 4$, la G-orbite de a_{n-1} est de cardinal 6, et se scinde en deux $\mathcal{D}_4$-orbites : l'une de cardinal 2 (formée de a_{n-1} et a_2), l'autre de cardinal 4 (formée de c, b_2, a_1, a_n). Si $n \geq 5$, la G-orbite de b_{n-1} est aussi de cardinal 6, et se scinde également en deux $\mathcal{D}_4$-orbites : l'une de cardinal 2 (formée de b_{n-1} et b_3), l'autre de cardinal 4 (formée de $a_{n-1} + b_{n-1} - c, a_2 + b_3 - b_2, a_2 + b_3 - a_1, a_{n-1} + b_{n-1} - a_n$).

Quand $n = 4$, la G-orbite de b_3 est de cardinal 9, et se scinde en trois $\mathcal{D}_4$-orbites : l'une est réduite à b_3, et les deux autres sont de cardinal 4, formées respectivement de $a_3 + b_3 - c, a_2 + b_3 - b_2, a_2 + b_3 - a_1, a_3 + b_3 - a_4$ et de $a_1 + a_2 - c, a_2 + b_2 - c, a_3 + c - b_2, a_3 + c - a_1$.

3.4 Structure du groupe G

Dans ce paragraphe, on étudie la structure du groupe G défini au paragraphe précédent. Pour $n = 2$ et $n = 3$, on retrouve exactement les situations considérées par Rhin et Viola. Pour $n \geq 4$, on détermine la structure de G en le faisant agir par permutation des coordonnées sur une variété (dans le même esprit que dans les deux premières parties de ce texte).

Théorème 3.7 *Le groupe G est un groupe fini, isomorphe :*
- *à $\mathfrak{S}_5$ si $n = 2$, donc d'ordre 120 ;*
- *à $H \rtimes \mathfrak{S}_5$ si $n = 3$, donc d'ordre 1920 (on note H l'hyperplan d'équation $\varepsilon_1 + \ldots + \varepsilon_5 = 0$ dans $(\mathbb{Z}/2\mathbb{Z})^5$) ;*
- *à $(\mathfrak{S}_3 \times \mathfrak{S}_3) \rtimes \mathbb{Z}/2\mathbb{Z}$ si $n \geq 4$, donc d'ordre 72 (on fait agir $\mathbb{Z}/2\mathbb{Z}$ sur $\mathfrak{S}_3 \times \mathfrak{S}_3$ par permutation des deux facteurs).*

N.B. Pour les applications arithmétiques, ce qui compte est l'indice dans G du sous-groupe formé par les $g \in G$ tels que $\xi(g; \underline{a}, \underline{b}, c) = 1$ pour tout $(\underline{a}, \underline{b}, c) \in \mathcal{E}^+$. Cet indice vaut 12 si $n = 2$, 120 si $n = 3$ et 9 si $n \geq 4$ (d'après le paragraphe précédent).

La suite de ce paragraphe est consacrée à la démonstration du théorème 3.7, et à une étude un peu plus détaillée de chacun de ces trois cas.

Dans le cas $n = 2$, posons $x = x_2$ et $y = 1 - x_1$. Alors la transformation σ est exactement celle de [R3], et ψ correspond, avec les notations de [R3], à $\tau \circ \sigma$ (i.e. $\bar{\psi} = \bar{\sigma}\bar{\tau}$, où $\bar{\tau}$ est le τ en caractères gras de [R3]). Donc le groupe engendré par ψ et σ est le même que celui engendré par τ et σ ; il est isomorphe au groupe diédral $\mathcal{D}_5$. Quant à la transformation hypergéométrique, celle notée φ dans [R3] correspond au cas particulier $n = 2$ de la transformation χ_L définie dans la proposition 4.5 (voir le paragraphe 4.2). Alors que cette transformation $\bar{\varphi}$ échange $\alpha_2 = i + j$ et $\alpha_5 = l + h$ en fixant α_1, α_3 et α_4, le cas particulier $n = 2$ de la transformation $\bar{\varphi}$ de la proposition 3.5 échange $\alpha_1 = h + i$ et α_5 en fixant α_2, α_3 et α_4. Mais le groupe engendré par $\mathcal{D}_5$ et φ reste le même ; il est isomorphe à $\mathfrak{S}_5$. Plus précisément $\mathcal{D}_5 \varphi \mathcal{D}_5$ est le même, car ces deux transformations φ sont de niveau 2. Par ailleurs, les conditions (4.36) sont vides quand $n = 2$: on retrouve exactement la situation considérée par Rhin et Viola.

Dans le cas $n = 3$, posons $x = x_2$, $y = 1 - x_1$ et $z = x_3$. Alors les transformations σ et φ sont exactement celles de [R4], et ψ correspond, avec les notations de [R4], à $\vartheta^2 \circ \sigma$ (ce qui signifie que $\bar{\psi} = \bar{\sigma}\bar{\vartheta}^2$). Le groupe engendré par ψ et σ est donc le même que celui engendré par σ et ϑ^2 ; il est isomorphe au groupe du carré $\mathcal{D}_4$. Le groupe $G = <\sigma, \psi, \varphi>$ est donc égal au sous-groupe de $\Phi = <\sigma, \vartheta, \varphi>$ engendré par σ, ϑ^2 et φ. Or on remarque que ce sous-groupe est Φ tout entier, car $\vartheta \in \mathrm{Aut}(\mathcal{E})$ est donné par

$$\vartheta = (\varphi\sigma\psi\sigma\varphi\sigma)^2\varphi, \text{ c'est-à-dire } \bar{\vartheta} = \bar{\varphi}(\bar{\sigma}\bar{\varphi}\bar{\sigma}\bar{\psi}\bar{\sigma}\bar{\varphi})^2, \tag{4.38}$$

où $\bar{\vartheta}$ est l'application notée ϑ en caractères gras dans [R4]. Ainsi, G est isomorphe à Φ, donc à $H \rtimes \mathfrak{S}_5$ (voir le paragraphe 2.4 ci-dessus).

De plus, quand $n = 3$, la contrainte (4.36) est exactement l'hypothèse $j + q = l + s$ de [R4]. On obtient donc exactement la situation considérée par Rhin et Viola, à une exception près : le groupe diédral engendré par σ et ψ est d'ordre 8, alors que celui engendré par σ et ϑ est d'ordre 16. Avec la présentation adoptée ici, il n'est pas clair a priori que l'élément $\vartheta \in G$ défini par (4.38) provienne d'un changement de variables. On peut cependant s'en douter en calculant $\xi(\vartheta; \underline{a}, \underline{b}, c)$ et en voyant que c'est 1 quels que soient $\underline{a}, \underline{b}, c$; il est alors facile par identification de trouver le changement de variables en question.

Cette ambiguïté sur le groupe diédral n'apparaît ni pour $n = 2$, ni pour $n \geq 4$. En effet, dans ce dernier cas, on a montré au paragraphe 3.3 que $\mathcal{D}_4 = <\sigma, \psi>$ est exactement l'ensemble des $g \in G$ tels que $\xi(g; \underline{a}, \underline{b}, c) = 1$ pour tous $\underline{a}, \underline{b}, c$.

On aimerait obtenir un groupe dont la structure soit de plus en plus riche quand n augmente ; ici, le théorème 3.7 montre qu'on n'y parvient pas. Pour enrichir cette structure, il faudrait probablement chercher du côté de la transformation hypergéométrique. En effet, on utilise seulement une formule en une variable x_k, et cette formule n'est valable que si un seul facteur du dénominateur dépend de x_k :

$$\int_0^1 \frac{x_k^a(1-x_k)^b}{(1+\beta x_k)^c}\frac{\mathrm{d}x_k}{(1+\beta x_k)} = \frac{a!b!}{c!(a+b-c)!}\int_0^1 \frac{x_k^c(1-x_k)^{a+b-c}}{(1+\beta x_k)^a}\frac{\mathrm{d}x_k}{(1+\beta x_k)}.$$

Une formule analogue, dans laquelle le dénominateur serait un produit de plusieurs facteurs distincts, pourrait être appliquée à x_k pour d'autres valeurs de k que $k = n - 1$ (qui donne φ à la proposition 3.5) et $k = n$ (qui donne χ_L et χ_K, voir les propositions 4.5 et 5.5). Elle permettrait aussi d'obtenir un analogue de φ, c'est-à-dire une transformation hypergéométrique par rapport à la variable x_{n-1}, pour les intégrales considérées dans les quatrième et cinquième parties de ce texte.

Une autre direction est de chercher quelles contraintes on met sur les paramètres : plus on met de contraintes et plus on a de chances que le groupe de transformations obtenu soit gros, mais moins l'étude de la valuation p-adique des $\xi(g; \underline{a}, \underline{b}, c)$ fournira d'informations. Par exemple, si on considère uniquement le cas où les $2n + 1$ exposants sont égaux, alors les transformations σ, ψ et φ sont triviales (voir aussi la remarque qui suit la proposition 4.5).

Il y a peut-être une direction dans laquelle le présent travail a une chance d'être exhaustif : c'est la recherche de changements de variables qui préservent la forme de l'intégrale. Précisément, on peut espérer une réponse positive aux questions suivantes :

Question : Soit $g \in \mathrm{GL}(\mathbb{Z}^{2n+1})$ tel que

$$I(\underline{a}, \underline{b}, c) = I(g(\underline{a}, \underline{b}, c))$$

pour tout $(\underline{a}, \underline{b}, c) \in \mathbb{Z}^{2n+1}$. A-t-on nécessairement $g = \mathrm{Id}$ ou $g = \sigma$, lorsque $n \geq 3$? Quand $n = 2$, g appartient-il nécessairement au sous-groupe (d'ordre huit) engendré par ψ et σ ?

Question : Soit g un élément de $\mathrm{GL}(\mathbb{Z}^{2n+1})$ qui vérifie

$$I(\underline{a}, \underline{b}, c) = I(g(\underline{a}, \underline{b}, c))$$

pour tout $(\underline{a}, \underline{b}, c) \in \mathbb{Z}^{2n+1}$ tel que

$$a_k + b_{k+1} = a_{k+2} + b_{k+2} \text{ pour tout } k \in \{1, \dots, n-2\}.$$

Lorsque $n \neq 3$, g appartient-il nécessairement au sous-groupe (d'ordre huit) engendré par ψ et σ ? Quand $n = 3$, appartient-il nécessairement au sous-groupe (d'ordre seize) engendré par ϑ et σ ?

N.B. Si on restreint ces questions aux éléments $g \in G$, la réponse est positive puisque la contrainte sur g implique $\xi(g; \underline{a}, \underline{b}, c) = 1$ pour tout $(\underline{a}, \underline{b}, c) \in \mathcal{E}^+$.

Considérons, dans toute la suite de ce paragraphe, le cas $n \geq 4$.

Soit $(\underline{a}, \underline{b}, c)$ un $(2n+1)$-uplet vérifiant les conditions (4.36). On pose $\nu_\infty = \frac{2}{3}(a_k + a_{k+1} + b_{k+1})$ pour $k \in \{1, \dots, n-1\}$; cette valeur ne dépend pas du choix de k. On pose aussi :

$$\begin{array}{rcl}
\nu_0 & = & a_1 + a_2 - \nu_\infty = a_1 + b_1 - \nu_\infty \\
\nu_1 & = & a_2 + b_2 - \nu_\infty = b_1 + b_2 - \nu_\infty \\
\nu_k & = & a_{k-1} + b_k - \nu_\infty = a_{k+1} + b_{k+1} - \nu_\infty \text{ pour tout } k \in \{2, \dots, n-1\} \\
\nu_n & = & a_{n-1} + b_n - \nu_\infty \\
\nu_{n+1} & = & a_{n-1} + a_n - \nu_\infty
\end{array}$$

On a alors :

$$\nu_0 + \nu_1 + \nu_2 = \nu_{n-1} + \nu_n + \nu_{n+1} = 0. \tag{4.39}$$

Réciproquement, tout $(n+3)$-uplet $(\nu_0, \dots, \nu_{n+1}, \nu_\infty)$ qui vérifie les relations (4.39) provient d'un $(2n+1)$-uplet $(\underline{a}, \underline{b}, c)$ qui vérifie les conditions (4.36), puisqu'on a :

$$\begin{array}{rcl}
a_k & = & \frac{1}{2}\nu_\infty - \nu_k \text{ pour tout } k \in \{1, \dots, n\} \\
b_k & = & \nu_{k-1} + \nu_k + \frac{1}{2}\nu_\infty \text{ pour tout } k \in \{1, \dots, n\} \\
c & = & b_n = \nu_{n-1} + \nu_n + \frac{1}{2}\nu_\infty = -\nu_{n+1} + \frac{1}{2}\nu_\infty
\end{array}$$

En outre, les conditions (4.37) s'écrivent :

$$\begin{array}{rrcl}
 & \nu_\infty & \geq & 2\nu_k \text{ pour tout } k \in \{1, \dots, n\}, \\
\text{et} & \nu_\infty + 2\nu_k + 2\nu_{k-1} & \geq & 0 \text{ pour tout } k \in \{1, \dots, n\}.
\end{array}$$

L'isomorphisme ci-dessus entre le sous-$\mathbb{Z}$-module $\mathcal{E}$ de $\mathbb{Z}^{2n+1}$ défini par les équations (4.36) et le sous-$\mathbb{Z}$-module $\mathcal{F}$ de $\mathbb{Z}^{n+3}$ défini par les équations (4.39) montre que $\mathcal{E}$ est de rang $n+1$. De plus, il permet d'identifier $\mathrm{GL}(\mathcal{E})$ et $\mathrm{GL}(\mathcal{F})$. On peut expliciter l'action de G sur $\mathcal{F}$ ainsi obtenue : $\bar{\sigma}$ fixe chaque composante, sauf ν_0 et ν_1 qu'elle échange ; $\bar{\psi}$ échange ν_k et ν_{n+1-k} pour tout $k \in \{0, \dots, n+1\}$, et fixe ν_∞ ; $\bar{\varphi}$ fixe ν_k pour tout $k \in \{0, \dots, n-2\} \cup \{n, \infty\}$ et échange ν_{n-1} et ν_{n+1}. On peut résumer cela par le diagramme suivant, dans lequel $\bar{\psi}$ est la symétrie par rapport à la droite verticale :

$$\begin{array}{ccccc|ccccc}
\nu_2 & \cdots & \cdots & \nu_k & \cdots & \cdots & \nu_{n+1-k} & \cdots & \cdots & \nu_{n-1} \\
 & & & & & & & & & \updownarrow \bar{\varphi} \\
\nu_0 & \overset{\bar{\sigma}}{\longleftrightarrow} & \nu_1 & & & & & & \nu_n & \nu_{n+1}
\end{array}$$

Tout élément $\gamma \in G$ induit une permutation de l'ensemble formé par les deux triangles $\{\nu_0, \nu_1, \nu_2\}$ et $\{\nu_{n-1}, \nu_n, \nu_{n+1}\}$. On définit ainsi un homomorphisme de groupes surjectif de G dans $\mathbb{Z}/2\mathbb{Z}$. Son noyau contient entre autres $\bar{\sigma}$, $\bar{\varphi}$, $\bar{\psi}\bar{\sigma}\bar{\psi}$ et $\bar{\psi}\bar{\varphi}\bar{\psi}$: il est isomorphe au groupe $\mathfrak{S}_3 \times \mathfrak{S}_3$ des permutations de l'ensemble à six éléments $\{\nu_0, \nu_1, \nu_2, \nu_{n-1}, \nu_n, \nu_{n+1}\}$ qui laissent stables les deux parties $\{\nu_0, \nu_1, \nu_2\}$ et $\{\nu_{n-1}, \nu_n, \nu_{n+1}\}$. Le groupe G se plonge ainsi dans $\mathfrak{S}_6$, et on voit qu'il est isomorphe à $(\mathfrak{S}_3 \times \mathfrak{S}_3) \rtimes \mathbb{Z}/2\mathbb{Z}$. Le sous-groupe de G engendré par σ et ψ est isomorphe au groupe diédral d'ordre 8 ; il agit sur le carré dont les sommets sont $\nu_0, \nu_n, \nu_1, \nu_{n+1}$ (dans cet ordre). On voit ainsi que $\bar{\sigma}\bar{\psi} = \overline{\psi\sigma}$ est d'ordre 4, comme le changement de variables dont il provient.

Les exposants ν_k introduits ici sont à rapprocher des paramètres $h_0, \dots, h_{n+2}$ utilisés par Zudilin [Z8]. Le lien est le suivant :

$$\nu_\infty = \frac{2}{3}(h_0 - 2) \text{ et } \nu_k = \frac{h_0 + 1}{3} - h_{n+2-k} \text{ pour tout } k \in \{0, \dots, n+1\}.$$

Il y a toutefois une différence importante : ici, on considère seulement des triplets $(\underline{a}, \underline{b}, c)$ qui appartiennent à $\mathcal{E}$. Cela signifie qu'à la condition

$$a_k + b_{k+1} = a_{k+2} + b_{k+2} \text{ pour tout } k \in \{1, \dots, n-2\} \tag{4.40}$$

qui apparaît chez Zudilin on adjoint (si $n \geq 4$) les hypothèses $a_2 = b_1$ et $b_n = c$.

Par une méthode assez différente de celle utilisée ici, Zudilin obtient [Z8] un groupe de Rhin-Viola, isomorphe à $\mathfrak{S}_{n+2}$, qui agit sur l'ensemble des paramètres $(\underline{a}, \underline{b}, c)$ vérifiant (4.40). Il constate que pour $n \geq 4$ ce groupe ne contient pas d'analogue de la transformation σ qui échange x et y. Ici, les restrictions supplémentaires $a_2 = b_1$ et $b_n = c$ (qui se traduisent par la relation (4.39)) permettent d'obtenir cette transformation σ. En utilisant les résultats de Zudilin, on voit que la famille $I(\underline{a}, \underline{b}, c)$, paramétrée par $(\underline{a}, \underline{b}, c) \in \mathcal{E}$, admet (si $n \geq 4$) un groupe de Rhin-Viola G' isomorphe à $((\mathfrak{S}_3 \times \mathfrak{S}_3) \rtimes \mathbb{Z}/2\mathbb{Z}) \times \mathfrak{S}_{n-4}$. Ceci améliore notre théorème 3.7 si $n \geq 6$. Ce groupe G' fixe ν_∞, agit sur $\{\nu_0, \nu_1, \nu_2, \nu_{n-1}, \nu_n, \nu_{n+1}\}$ comme G et sur $\{\nu_3, \dots, \nu_{n-2}\}$ par toutes les permutations de $\nu_3, \dots, \nu_{n-2}$. Les éléments $g' \in G'$ tels que $I(\underline{a}, \underline{b}, c) = I(g'(\underline{a}, \underline{b}, c))$ pour tout $(\underline{a}, \underline{b}, c) \in \mathcal{E}$ appartiennent à G : ce sont les éléments du sous-groupe engendré par ψ et σ.

Question : Peut-on démontrer que G' est un groupe de Rhin-Viola pour la famille $I(\underline{a}, \underline{b}, c)$ paramétrée par $\mathcal{E}$ en s'inspirant de la méthode utilisée dans les deux premières parties ?

Pour répondre à cette question, la première étape est de considérer les coordonnées suivantes, qui sont permutées par G' :

$$
\begin{array}{rrcl}
& y_0 & = & \varepsilon \dfrac{x_1(1-x_1)x_2}{1-x_2} \\
& y_1 & = & \varepsilon \dfrac{(1-x_1)x_2(1-x_2)}{x_1} \\
& y_2 & = & \varepsilon \dfrac{x_1(1-x_2)(1-x_3)^2}{(1-x_1)x_2} \\
& y_k & = & \dfrac{(1-x_k)(1-x_{k+1})}{x_k} \text{ pour tout } k \in \{3, \dots, n-2\} \\
& y_{n-1} & = & \eta \dfrac{x_n(1-x_n)(1-x_{n-1})^2}{x_{n-1}} \dfrac{\delta_{n-2}(\underline{x})}{\delta_n(\underline{x})} \\
& y_n & = & \eta \dfrac{x_{n-1}(1-x_n)}{x_n} \dfrac{\delta_{n-2}(\underline{x})}{\delta_n(\underline{x})} \\
& y_{n+1} & = & \eta \dfrac{x_{n-1}x_n}{1-x_n} \delta_{n-2}(\underline{x}) \delta_n(\underline{x}) \\
& y_\infty & = & \delta_n(\underline{x})^{-1} \displaystyle\prod_{k=1}^{n} x_k(1-x_k) \\
\text{avec} & \varepsilon & = & (x_1(1-x_1)x_2(1-x_2)(1-x_3)^2)^{-1/3} \\
\text{et} & \eta & = & (\dfrac{\delta_{n-2}(\underline{x})^3}{\delta_n(\underline{x})} x_{n-1}(1-x_{n-1})^2 x_n(1-x_n))^{-1/3}.
\end{array}
$$

Notons Ω l'ensemble des points $(y_0, \dots, y_{n+1}, y_\infty) \in \mathbb{R}^{n+3}$ ainsi obtenus quand $(x_1, \dots, x_n)$ décrit $]0,1[^n$, et $\mathcal{V}$ l'adhérence de Zariski de Ω. Alors $\mathcal{V}$ est incluse dans les hypersurfaces d'équations $y_0y_1y_2 = 1$ et $y_{n-1}y_ny_{n+1} = 1$ (c'est le rôle des facteurs ε et η dans les formules ci-dessus). Le groupe du carré (engendré par $\bar{\sigma}$ et $\bar{\psi}$) agit sur Ω (donc sur $\mathcal{V}$) en permutant les coordonnées : $\bar{\sigma}$ échange y_0 et y_1, alors que $\bar{\psi}$ échange y_k et y_{n+1-k} pour tout $k \in \{0, \dots, n+1\}$. En outre, l'intégrande de $I(\underline{a}, \underline{b}, c)$ apparaît comme un monôme en les $n+3$ coordonnées en posant $\alpha_k = \frac{1}{2}\nu_k$ pour $k \in \{0, 1, 2, n-1, n, n+1, \infty\}$ et $\alpha_k = \nu_k$ pour $k \in \{3, \dots, n-2\}$, car on a :

$$\delta_n(\underline{x})^{-c} \prod_{k=1}^{n} x_k^{a_k}(1-x_k)^{b_k} = y_\infty^{\alpha_\infty} \prod_{k=0}^{n+1} y_k^{\alpha_k}.$$

La variable y_∞ joue ici le même rôle que ε_6 dans la partie 2 : elle correspond au facteur $\frac{\prod_{k=1}^n x_k(1-x_k)}{\delta_n(\underline{x})}$, sur lequel le groupe agit trivialement.

3.5 Aspects arithmétiques

Dans le cas $n = 3$, la contrainte $j + q = l + s$ qui est imposée dans [R4] pour obtenir une action intéressante de groupe permet d'obtenir des formes linéaires en 1 et $\zeta(3)$, alors que sans cette hypothèse $\zeta(2)$ peut aussi apparaître (voir [R4], page 276). Le théorème suivant [Z8] montre qu'il s'agit d'un fait général ; ceci démontre partiellement une conjecture de [V3], et répond à une question posée dans [F2].

Théorème 3.8 *Soit $(\underline{a}, \underline{b}, c)$ un $(2n+1)$-uplet d'entiers relatifs vérifiant les conditions (4.33) de convergence, et les conditions (4.34) de la proposition 3.3. Alors $I(\underline{a}, \underline{b}, c)$ est une combinaison linéaire, à coefficients rationnels, de 1 et des $\zeta(j)$ pour j entier, compris entre 2 et n, ayant la même parité que n.*

En particulier ce théorème s'applique aux éléments de $\mathcal{E}$, sous réserve que les conditions de convergence soient vérifiées.

4 Une famille plus générale d'intégrales n-uples

Soient $\underline{a} = (a_1, \ldots, a_n)$ et $\underline{b} = (b_1, \ldots, b_n)$ deux n-uplets d'entiers relatifs, et soit $\underline{c} = (c_2, \ldots, c_n)$ un $(n-1)$-uplet d'entiers relatifs.

Comme dans la partie précédente, on note $\delta_k(\underline{x}) = 1 - x_k(1 - x_{k-1}(\ldots(1 - x_1)))$ pour $\underline{x} = (x_1, \ldots, x_n) \in [0,1]^n$ et $k \in \{1, \ldots, n\}$. On considère les intégrales suivantes, qui sont positives et éventuellement infinies :

$$L(\underline{a}, \underline{b}, \underline{c}) = \int_{[0,1]^n} \frac{\prod_{k=1}^n x_k^{a_k}(1-x_k)^{b_k}}{\prod_{k=2}^n \delta_k(\underline{x})^{c_k}} \frac{\mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n}{\delta_n(\underline{x})}.$$

4.1 Convergence

Aux paramètres $\underline{a}, \underline{b}, \underline{c}$ on associe des entiers $\varrho_1, \ldots, \varrho_{n+2}$ de la manière suivante. On pose $\varrho_{n+2} = \varrho_{n+1} = 0$, $\varrho_n = c_n - b_n$ et pour tout $k \in \{1, \ldots, n-1\}$:

$$\varrho_k = \varrho_{k+2}^+ + c_k - 1 - b_k,$$

en notant $\alpha^+ = \max(\alpha, 0)$ et avec la convention $c_1 = 1$. On note aussi $\tilde{c}_i = c_i$ pour $i \in \{1, \ldots, n-1\}$ et $\tilde{c}_n = c_n + 1$.

Dans ce paragraphe uniquement, on autorise l'entier n à valoir 1. Dans ce cas, on pose $\varrho_1 = -b_1$ et $\varrho_2 = \varrho_3 = 0$.

Proposition 4.1 *L'intégrale $L(\underline{a}, \underline{b}, \underline{c})$ converge si, et seulement si, les a_k et les b_k sont tous positifs ou nuls, et qu'on a pour tout $k \in \{1, \ldots, n\}$, avec la convention $a_0 = 0$:*

$$\varrho_k \leq a_{k-1}.$$

N.B. On n'utilisera donc l'intégrale $L(\underline{a}, \underline{b}, \underline{c})$ que quand les a_k et les b_k sont tous positifs ou nuls ; en revanche, on a intérêt à autoriser les c_k à être éventuellement négatifs, c'est-à-dire à ne pas développer une puissance de $\delta_k(\underline{x})$ qui figurerait au numérateur. En effet, dans certains cas, cela reviendrait à écrire une intégrale convergente comme somme d'intégrales divergentes.

Démonstration : La convergence de l'intégrale résulte des deux lemmes ci-dessous. Montrons la réciproque. Tout d'abord, s'il existe k tel que $a_k \leq -1$ (respectivement $b_k \leq -1$) alors pour $\varepsilon > 0$ assez petit on restreint l'intégrale au domaine défini par $x_k \in [0, \varepsilon]$ (resp. $x_k \in [1-\varepsilon, 1]$) et $x_j \in [\frac{1}{3}, \frac{2}{3}]$ pour tout $j \neq k$. Sur ce domaine, on a $\delta_p(\underline{x}) \in [\frac{1}{9}, 1]$ pour tout $p \geq 2$, donc l'intégrale est déjà infinie. Supposons à partir de maintenant que les a_k et les b_k sont tous positifs ou nuls, et qu'il existe $k \in \{1, \ldots, n\}$ tel que $\varrho_k \geq 1 + a_{k-1}$. Dans un premier temps, supposons $k \geq 2$. Notons p le plus petit entier naturel pair tel que $\varrho_{k+p} \leq 0$. Soit $\varepsilon > 0$ assez petit. Restreignons-nous au domaine défini par les contraintes $x_{k-1} \in [\varepsilon, 2\varepsilon]$, $x_k, x_{k+2}, x_{k+4}, \ldots, x_{k+p-2} \in [1-2\varepsilon, 1-\varepsilon]$, et $x_i \in [\frac{1}{3}, \frac{2}{3}]$ pour les autres indices i. Dans ce domaine, on a $\delta_{k-1}(\underline{x}) \geq 1-2\varepsilon$ d'où $\delta_i(\underline{x}) \leq c\varepsilon$ pour tout $i \in \{k, k+2, k+4, \ldots, k+p-2\}$ avec une certaine constante $c > 0$ indépendante de ε. Par ailleurs, on a $\frac{1}{3} \leq \delta_i(\underline{x}) \leq 1$ pour tout $i \notin \{k, k+2, k+4, \ldots, k+p-2\}$. L'intégrale sur ce domaine est donc minorée, à une constante multiplicative près (indépendante de ε), par ε à la puissance $a_{k-1}+1+\sum_{j=0}^{(p-2)/2}(b_{k+2j}+1-\tilde{c}_{k+2j})$, qui vaut $a_{k-1}+1-\varrho_k \leq 0$ (par définition de k et p). Comme ce minorant ne tend pas vers zéro quand ε tend vers 0, l'intégrale $L(\underline{a}, \underline{b}, \underline{c})$ diverge. Dans le cas où $k = 1$, on procède de même, en omettant la contrainte sur x_{k-1} et en utilisant la convention $c_1 = 1$ qui apparaît dans la définition de ϱ_1.

Lemme 4.2 *Supposons que les a_k et les b_k sont tous positifs ou nuls, et qu'on a $\varrho_k \leq a_{k-1}$ pour tout $k \in \{1, \ldots, n\}$, avec la convention $a_0 = 0$. Alors la fonction $\xi_n(\underline{a}, \underline{b}, \underline{c})$ définie par*

$$\xi_n(\underline{a}, \underline{b}, \underline{c})(\underline{x}) = \frac{\prod_{k=1}^n x_k^{a_k}(1-x_k)^{b_k}}{\prod_{k=2}^n \delta_k(\underline{x})^{\tilde{c}_k-1}}$$

est bornée sur $[0,1]^n$.

Démonstration : Ce résultat est vrai pour $n = 1$. Soit $n \geq 2$ tel que le lemme soit vrai pour $n-1$. On peut supposer $\tilde{c}_n \geq 2$. On a $\delta_n(\underline{x}) \geq 1-x_n$ et $\delta_n(\underline{x}) = 1-x_n+x_n x_{n-1}\delta_{n-2}(\underline{x}) \geq x_{n-1}\delta_{n-2}(\underline{x})$, d'où :

$$\delta_n(\underline{x})^{\tilde{c}_n-1} \geq (1-x_n)^{\min(b_n,c_n)}(x_{n-1}\delta_{n-2}(\underline{x}))^{(c_n-b_n)^+}.$$

On en déduit :

$$\xi_n(\underline{a}, \underline{b}, \underline{c})(\underline{x}) \leq x_n^{a_n}(1-x_n)^{b_n-\min(b_n,c_n)}\xi_{n-1}(\underline{a'}, \underline{b'}, \underline{c'})(x_1, \ldots, x_{n-1}),$$

en posant :

$$\begin{aligned}
a'_k &= a_k \text{ pour tout } k \in \{1, \ldots, n-2\} \\
a'_{n-1} &= a_{n-1} - (c_n - b_n)^+ \\
b'_k &= b_k \text{ pour tout } k \in \{1, \ldots, n-1\} \\
c'_k &= c_k \text{ pour tout } k \in \{2, \ldots, n-3\} \\
c'_{n-2} &= c_{n-2} + (c_n - b_n)^+ \text{ si } n \geq 4 \\
c'_{n-1} &= c_{n-1} - 1 \text{ si } n \geq 3
\end{aligned}$$

On a alors $\varrho'_k = \varrho_k$ pour tout $k \in \{1, \ldots, n-1\}$: par récurrence, $\xi_{n-1}(\underline{a'}, \underline{b'}, \underline{c'})$ est donc bornée. Comme on a $\xi_n(\underline{a}, \underline{b}, \underline{c}) \leq \xi_{n-1}(\underline{a'}, \underline{b'}, \underline{c'})$, cela termine la démonstration du lemme.

Lemme 4.3 *Pour tout entier $p \geq 0$, l'intégrale $\int_{[0,1]^n} \frac{(-\log(\delta_n(\underline{x})))^p \mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n}{\prod_{k=2}^n \delta_k(\underline{x})}$ est convergente.*

Démonstration : On procède par récurrence sur n ; le cas $n = 1$ est clair. Supposons le résultat vrai jusqu'à $n-1$, avec $n \geq 2$. Alors il suffit de montrer que $\int_{[0,1]^n} \frac{(-\log(\delta_n(\underline{x})))^p \mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n}{\prod_{k \equiv n \mod 2} \delta_k(\underline{x})}$ converge, car là où $\delta_n(\underline{x})$ est proche de 0 les $\delta_k(\underline{x})$, pour $k \not\equiv n \mod 2$, sont minorés par une constante strictement positive. On découpe le domaine d'intégration en deux. Là où $1-x_n \leq x_{n-1}\delta_{n-2}(\underline{x})$, on écrit $\delta_n(\underline{x}) = (1-x_n)+x_n x_{n-1}\delta_{n-2}(\underline{x}) \geq x_n x_{n-1}\delta_{n-2}(\underline{x})$ ce qui permet de majorer $(-\log(\delta_n(\underline{x})))^p$ par une combinaison linéaire de termes de la forme $(-\log(\delta_{n-2}(\underline{x})))^{p_1}(-\log(x_{n-1}))^{p_2}(-\log(x_n))^{p_3}$. Comme le problème de convergence se pose quand x_n est proche de 1, on peut supposer que $\frac{(-\log(x_n))^{p_3}}{x_n}$ est majoré par une constante ; on en déduit immédiatement le résultat. Dans le reste de l'hypercube, on a $1-x_n > x_{n-1}\delta_{n-2}(\underline{x})$, et on écrit $\delta_n(\underline{x}) \geq 1-x_n$. Cette fois, on majore $\int_0^{1-x_{n-1}\delta_{n-2}(\underline{x})} \frac{(-\log(1-x_n))^p}{1-x_n}\mathrm{d}x_n$ par $(-\log(x_{n-1}\delta_{n-2}(\underline{x})))^{p+1}$, et on conclut de manière analogue.

4.2 Définition des transformations

Notons σ_L et ψ_L les automorphismes de $\mathbb{Z}^{3n-1}$ définis comme suit : $\overline{\sigma_L}$ échange a_1 et b_2, ainsi que a_2 et b_1, et fixe les autres coordonnées. Quant à $\overline{\psi_L}$, elle est donnée par les formules suivantes :

$$
\begin{aligned}
\overline{\psi_L}(a_k) &= a_{n+1-k} \text{ pour tout } k \in \{1, \dots, n\} \\
\overline{\psi_L}(b_k) &= b_{n+2-k} \text{ pour tout } k \in \{2, \dots, n\} \\
\overline{\psi_L}(b_1) &= a_{n-1} - (c_n - b_n) \\
\overline{\psi_L}(c_k) &= a_{n+2-k} + b_{n+2-k} + c_{n+1-k} - b_{n+1-k} - a_{n-k} \text{ pour tout } k \in \{2, \dots, n-1\} \\
\overline{\psi_L}(c_n) &= a_2 + b_2 - b_1
\end{aligned}
$$

Proposition 4.4 *Pour tout triplet* $(\underline{a}, \underline{b}, \underline{c}) \in \mathbb{Z}^{3n-1}$ *on a :*

$$L(\underline{a}, \underline{b}, \underline{c}) = L(\psi_L(\underline{a}, \underline{b}, \underline{c})) = L(\sigma_L(\underline{a}, \underline{b}, \underline{c})).$$

DÉMONSTRATION : On utilise les mêmes changements de variables que dans les preuves des propositions 3.2 et 3.3.

Proposition 4.5 *Notons* χ_L *l'automorphisme de* $\mathbb{Z}^{3n-1}$ *défini par :*

$$
\begin{aligned}
\overline{\chi_L}(a_k) &= a_k \textit{ pour tout } k \in \{1, \dots, n-1\} \\
\overline{\chi_L}(b_k) &= b_k \textit{ pour tout } k \in \{1, \dots, n-1\} \\
\overline{\chi_L}(c_k) &= c_k \textit{ pour tout } k \in \{2, \dots, n-1\} \\
\overline{\chi_L}(a_n) &= c_n \\
\overline{\chi_L}(b_n) &= b_n + a_n - c_n \\
\overline{\chi_L}(c_n) &= a_n
\end{aligned}
$$

Alors on a, pour tout $(\underline{a}, \underline{b}, \underline{c}) \in \mathbb{Z}^{3n-1}$ *qui vérifie les conditions de convergence de la proposition 4.1 et tel que* c_n *et* $a_n + b_n - c_n$ *soient positifs :*

$$L(\underline{a}, \underline{b}, \underline{c}) = \frac{a_n! b_n!}{c_n!(a_n + b_n - c_n)!} L(\chi_L(\underline{a}, \underline{b}, \underline{c})).$$

DÉMONSTRATION : Cela provient de la formule

$$\int_0^1 \frac{x_n^{a_n}(1-x_n)^{b_n}}{(1-\delta_{n-1}(\underline{x})x_n)^{c_n+1}} \mathrm{d}x_n = \frac{a_n! b_n!}{c_n!(a_n + b_n - c_n)!} \int_0^1 \frac{x_n^{c_n}(1-x_n)^{a_n+b_n-c_n}}{(1-\delta_{n-1}(\underline{x})x_n)^{a_n+1}} \mathrm{d}x_n$$

en raisonnant comme dans la preuve de la proposition 3.5.

N.B. En se restreignant aux $(\underline{a}, \underline{b}, \underline{c})$ tels que $c_2 = \ldots = c_{n-1} = 0$, on voit que χ_L agit sur les intégrales $I(\underline{a}, \underline{b}, c)$ de la partie 3. En fait χ_L appartient au groupe G défini au paragraphe 3.3. Quand $n \geq 4$, l'hypothèse $b_n = c$ qui apparaît dans la définition de $\mathcal{E}$ fait que le quotient de factorielles se simplifie, et χ_L agit sur $\mathcal{E}$ comme le changement de variables $\psi\sigma\psi$.

4.3 Action et structure du groupe

Dans ce paragraphe, on suppose $n \geq 3$. Notons G_L le groupe engendré par σ_L, ψ_L et χ_L. Il agit sur l'ensemble à huit éléments $\{a_n, b_n, c_n, a_n+b_n-c_n, a_1, b_2, a_1+b_1-a_2, a_2+b_2-b_1\}$ comme indiqué sur le diagramme suivant, où $\overline{\psi_L}$ est la symétrie par rapport à la droite verticale :

$$\begin{array}{ccc|cc} c_n & \xleftrightarrow{\overline{\chi_L}} & a_n & a_1 & a_2+b_2-b_1 \\ & & & \updownarrow \overline{\sigma_L} & \updownarrow \overline{\sigma_L} \\ a_n+b_n-c_n & \xleftrightarrow{\overline{\chi_L}} & b_n & b_2 & a_1+b_1-a_2 \end{array}$$

On a deux carrés : $\{a_n, b_n, c_n, a_n+b_n-c_n\}$ et $\{a_1, b_2, a_1+b_1-a_2, a_2+b_2-b_1\}$. On a alors un homomorphisme surjectif de G_L dans $\mathbb{Z}/2\mathbb{Z}$, qui à tout $g \in G_L$ associe 0 si g fixe globalement chacun des deux carrés, et 1 si g les échange. Le noyau de cet homomorphisme est isomorphe à $V \times V$, où $V = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ est le groupe de Klein (engendré par les symétries par rapport aux médiatrices d'un carré). L'image de G_L dans $\mathfrak{S}_8$ est donc d'ordre 32, isomorphe à $(V \times V) \rtimes \mathbb{Z}/2\mathbb{Z}$. Or l'action de G_L sur cet ensemble à huit éléments est fidèle ; donc G_L est d'ordre 32.

On peut résumer les constructions de ce paragraphe dans l'énoncé suivant :

Théorème 4.6 *La famille $L(\underline{a}, \underline{b}, \underline{c})$, paramétrée par $(\underline{a}, \underline{b}, \underline{c}) \in \mathbb{Z}^{3n-1}$, admet un groupe de Rhin-Viola d'ordre 32, isomorphe à $(V \times V) \rtimes \mathbb{Z}/2\mathbb{Z}$, qui est engendré par deux changements de variables σ_L et ψ_L et une transformation hypergéométrique χ_L.*

5 Des intégrales adaptées au développement en série entière

Les intégrales considérées dans cette partie font apparaître au dénominateur des facteurs $1 - y_1 \dots y_k$, et non des δ_k. C'est pourquoi il est plus naturel de les développer en séries multiples. On montre qu'elles admettent un groupe de Rhin-Viola isomorphe à celui du paragraphe 4.3. En outre, au paragraphe 5.5 on exhibe un changement de variables qui relie ces intégrales à celles considérées dans la partie 3 (et même à une famille plus générale, définie au paragraphe 5.5). Comme application, on transforme les intégrales considérées par Beukers, Vasilenko et Vasilyev en intégrales qui se développent naturellement en séries.

Dans cette partie comme dans le reste du texte, on suppose que les exposants sont entiers, même si les résultats se généralisent.

5.1 Définition

À tout point $\underline{y} = (y_1, \dots, y_n) \in [0,1]^n$, et à tout entier $k \in \{0, \dots, n\}$, on associe le réel $\alpha_k(\underline{y}) = y_1 y_2 \dots y_k$. De plus, on pose par convention $y_{n+1} = 0$ et $\alpha_{n+1}(\underline{y}) = 0$.

Soient $\underline{A} = (A_1, \dots, A_n)$, $\underline{B} = (B_1, \dots, B_n)$ et $\underline{C} = (C_2, \dots, C_n)$ trois familles d'entiers relatifs.

On note $\bar{\omega}$ la n-forme différentielle sur $[0,1]^n$ définie par :

$$\bar{\omega} = \frac{\mathrm{d}y_1 \wedge \dots \wedge \mathrm{d}y_n}{(1-\alpha_2(\underline{y})) \cdot (1-\alpha_3(\underline{y})) \cdot \ldots \cdot (1-\alpha_n(\underline{y}))}.$$

On considère les intégrales suivantes, qui sont positives et éventuellement infinies :

$$K(\underline{A}, \underline{B}, \underline{C}) = \int_{[0,1]^n} \frac{\prod_{k=1}^n y_k^{A_k}(1-y_k)^{B_k}}{\prod_{k=2}^n (1-\alpha_k(\underline{y}))^{C_k+1}} \mathrm{d}y_1 \dots \mathrm{d}y_n.$$

5.2 Convergence

Dans ce paragraphe, on autorise l'entier n à valoir 1. La condition nécessaire et suffisante de convergence de l'intégrale $K(\underline{A}, \underline{B}, \underline{C})$, dans laquelle on suppose que les exposants sont entiers, est donnée par la proposition suivante :

Proposition 5.1 *L'intégrale $K(\underline{A}, \underline{B}, \underline{C})$ est convergente si, et seulement si, les A_k et les B_k sont tous positifs ou nuls, et qu'on a :*

$$\sum_{k=2}^{p} C_k \leq \sum_{j=1}^{p} B_j \text{ pour tout } p \in \{2, \dots, n\}.$$

DÉMONSTRATION : La convergence de l'intégrale résulte des deux lemmes ci-dessous. Montrons la réciproque, en supposant que les A_k et les B_k sont tous positifs ou nuls, et qu'il existe p tel que $\sum_{k=2}^{p} C_k \geq 1 + \sum_{j=1}^{p} B_j$. Soit $\varepsilon > 0$ assez petit. Restreignons-nous au domaine défini par les inégalités $1 - 2\varepsilon < y_p < \ldots < y_2 < y_1 < 1 - \varepsilon$ et $y_{p+1}, \dots, y_n \in [\frac{1}{3}, \frac{2}{3}]$. Sur ce domaine, on a $\varepsilon \leq 1 - \alpha_k(\underline{y}) \leq 2n\varepsilon$ et $\frac{1-y_k}{1-\alpha_k(\underline{y})} \geq \frac{1}{k}$ pour tout $k \in \{2, \dots, p\}$, donc $\frac{(1-y_k)^{B_k}}{(1-\alpha_k(\underline{y}))^{C_k+1}}$ est minoré par $\varepsilon^{-(C_k+1-B_k)}$, à une constante multiplicative près (qui est indépendante de ε). L'intégrale sur ce domaine est donc minorée par un multiple constant de $\varepsilon^{B_1+p-\sum_{k=2}^{p}(C_k+1-B_k)}$, qui ne tend pas vers 0 avec ε. Ceci démontre que l'intégrale $K(\underline{A}, \underline{B}, \underline{C})$ diverge.

Lemme 5.2 *Supposons que les A_k et les B_k sont tous positifs ou nuls, et qu'on a $\sum_{k=2}^{p} C_k \leq \sum_{j=1}^{p} B_j$ pour tout $p \in \{2, \dots, n\}$. Notons $\xi_n(\underline{A}, \underline{B}, \underline{C})$ la fonction dont $K(\underline{A}, \underline{B}, \underline{C})$ est l'intégrale contre $\bar{\omega}$:*

$$\xi_n(\underline{A}, \underline{B}, \underline{C})(\underline{y}) = \frac{\prod_{k=1}^{n} y_k^{A_k}(1-y_k)^{B_k}}{\prod_{k=2}^{n}(1-\alpha_k(\underline{y}))^{C_k}}.$$

Alors $\xi_n(\underline{A}, \underline{B}, \underline{C})$ est une fonction bornée sur $[0,1]^n$.

DÉMONSTRATION : Quitte à remplacer chaque C_k par $C_k - B_k$, pour $k \in \{2, \dots, n\}$, on peut supposer $B_2 = \ldots = B_n = 0$ (puisque $1 - y_k \leq 1 - \alpha_k(\underline{y})$ sur $[0,1]^n$). De même, en remplaçant C_2 par $C_2 - B_1$, on se ramène à $B_1 = \ldots = B_n = 0$. On a alors $C_2 \leq 0$, donc l'inégalité $1 - \alpha_2(\underline{y}) \leq 1 - \alpha_3(\underline{y})$ permet de remplacer C_2 par 0 et C_3 par $C_2 + C_3$, qui est négatif par hypothèse. On peut donc itérer ce processus, pour se ramener au cas où $B_1 = \ldots = B_n = 0$ et $C_2 = \ldots = C_{n-1} = 0$, avec $C_n \leq 0$. Dans ce cas, la fonction ξ_n est évidemment bornée.

Lemme 5.3 *L'intégrale*

$$K(\underline{0}, \underline{0}, \underline{0}) = \int_{[0,1]^n} \bar{\omega} = \int_{[0,1]^n} \frac{\mathrm{d}y_1 \wedge \ldots \wedge \mathrm{d}y_n}{(1-\alpha_2(\underline{y})) \cdot (1-\alpha_3(\underline{y})) \cdot \ldots \cdot (1-\alpha_n(\underline{y}))}$$

est convergente, et est égale à la série

$$\sum_{\ell_1 \geq \ell_2 \geq \ldots \geq \ell_{n-1} \geq 1} \frac{1}{\ell_1^2 \ell_2 \ell_3 \ldots \ell_{n-1}}.$$

DÉMONSTRATION : Immédiat par développement en série entière.

5.3 Définition des transformations

Définissons deux automorphismes σ_K et λ de $\mathbb{Z}^{3n-1}$ de la manière suivante. La permutation $\overline{\sigma_K}$ associée à σ_K (voir le début du paragraphe 3.2) échange A_1 et A_2, ainsi que B_1 et B_2, en fixant les autres composantes. Quant à $\bar{\lambda}$, elle est donnée par les formules suivantes (avec la convention $C_1 = 0$) :

$$\begin{aligned}
\bar{\lambda}(A_k) &= \sum_{j=1}^{n+1-k} B_j - C_j \text{ pour tout } k \in \{1, \ldots, n\} \\
\bar{\lambda}(B_1) &= A_n \\
\bar{\lambda}(B_k) &= B_{n+2-k} \text{ pour tout } k \in \{2, \ldots, n\} \\
\bar{\lambda}(C_k) &= A_{n+2-k} + B_{n+2-k} - A_{n+1-k} \text{ pour tout } k \in \{2, \ldots, n\}
\end{aligned}$$

On a alors la proposition suivante :

Proposition 5.4 *Pour tout* $(\underline{A}, \underline{B}, \underline{C}) \in \mathbb{Z}^{3n-1}$ *on a :*

$$K(\underline{A}, \underline{B}, \underline{C}) = K(\sigma_K(\underline{A}, \underline{B}, \underline{C})) = K(\lambda(\underline{A}, \underline{B}, \underline{C})).$$

Démonstration : Pour σ_K il suffit d'appliquer le changement de variables qui échange y_1 et y_2 en fixant les y_k pour $k \geq 3$. Pour λ on applique le changement de variables noté également λ, qui à $\underline{y} = (y_1, \ldots, y_n)$ associe $\underline{y'} = (y_1, \ldots, y_n)$ défini (avec la convention $\alpha_{n+1}(\underline{y}) = 0$) par :

$$y_k = \frac{1 - \alpha_{n+1-k}(\underline{y})}{1 - \alpha_{n+2-k}(\underline{y})}$$

pour tout $k \in \{1, \ldots, n\}$. Il vérifie $1 - y_k = \frac{\alpha_{n+1-k}(\underline{y})(1-y_{n+2-k})}{1-\alpha_{n+2-k}(\underline{y})}$ et $1 - \alpha_k(\underline{y'}) = \alpha_{n+1-k}(\underline{y})$ pour tout $k \in \{1, \ldots, n\}$. Pour voir que λ stabilise (au signe près) la forme différentielle $\bar{\omega}$, il suffit de remarquer que $\mathrm{d}\alpha_1 \wedge \ldots \wedge \mathrm{d}\alpha_n = \alpha_1(\underline{y})\alpha_2(\underline{y}) \ldots \alpha_n(\underline{y})\mathrm{d}y_1 \wedge \ldots \wedge \mathrm{d}y_n$ est fixée, au signe près, par λ. Ceci termine la démonstration de la proposition 5.4.

N.B. On voit facilement que λ et σ_K sont d'ordre 2, et que leur composée $\lambda\sigma_K$ est d'ordre 4 : le groupe engendré par λ et σ_K est isomorphe au groupe diédral $\mathcal{D}_4$, qui est d'ordre 8. De plus, $\lambda\sigma_K\lambda$ est le changement de variables qui fixe les $n-1$ premières coordonnées et remplace y_n par $\frac{1-y_n}{1-\alpha_n(\underline{y})}$.

Notons χ_K l'automorphisme de $\mathbb{Z}^{3n-1}$ défini par :

$$\begin{aligned}
\overline{\chi_K}(A_k) &= A_k \text{ pour tout } k \in \{1, \ldots, n-1\} \\
\overline{\chi_K}(B_k) &= B_k \text{ pour tout } k \in \{1, \ldots, n-1\} \\
\overline{\chi_K}(C_k) &= C_k \text{ pour tout } k \in \{2, \ldots, n-1\} \\
\overline{\chi_K}(A_n) &= C_n \\
\overline{\chi_K}(B_n) &= A_n + B_n - C_n \\
\overline{\chi_K}(C_n) &= A_n
\end{aligned}$$

On a alors la proposition suivante, qu'on démontre de manière analogue à la proposition 4.5 :

Proposition 5.5 *Pour tout* $(\underline{A}, \underline{B}, \underline{C}) \in \mathbb{Z}^{3n-1}$ *qui vérifie les conditions de convergence de la proposition 5.1 et tel que* C_n *et* $A_n + B_n - C_n$ *soient positifs on a :*

$$K(\underline{A}, \underline{B}, \underline{C}) = \frac{A_n! B_n!}{C_n!(A_n + B_n - C_n)!} K(\chi_K(\underline{A}, \underline{B}, \underline{C})).$$

5.4 Action et structure du groupe

Dans ce paragraphe, on suppose $n \geq 3$. Notons G_K le groupe engendré par σ_K, λ et χ_K. Il agit sur l'ensemble à huit éléments $\{A_n, B_n, C_n, A_n + B_n - C_n, B_1, B_2, A_1 + B_1 - A_2, A_2 + B_2 - A_1\}$ comme indiqué dans le diagramme suivant, où λ est la symétrie par rapport à la droite verticale :

$$\begin{array}{ccc|cc} C_n & \overset{\chi_K}{\longleftrightarrow} & A_n & B_1 & A_2 + B_2 - A_1 \\ & & & \updownarrow \sigma_K & \updownarrow \sigma_K \\ A_n + B_n - C_n & \overset{\chi_K}{\longleftrightarrow} & B_n & B_2 & A_1 + B_1 - A_2 \end{array}$$

On a deux carrés : $\{A_n, B_n, C_n, A_n + B_n - C_n\}$ et $\{B_1, B_2, A_1 + B_1 - A_2, A_2 + B_2 - A_1\}$. Considérons l'homomorphisme surjectif de G_K dans $\mathbb{Z}/2\mathbb{Z}$ qui à tout $g \in G_K$ associe 0 si g fixe globalement chacun des deux carrés, et 1 si g les échange. Le noyau de cet homomorphisme est isomorphe à $V \times V$, où V est le groupe de Klein. Comme l'action est fidèle, on voit que G_K est d'ordre 32, isomorphe à $(V \times V) \rtimes \mathbb{Z}/2\mathbb{Z}$.

On peut résumer les constructions de ce paragraphe dans l'énoncé suivant :

Théorème 5.6 *La famille $K(\underline{A}, \underline{B}, \underline{C})$, paramétrée par $(\underline{A}, \underline{B}, \underline{C}) \in \mathbb{Z}^{3n-1}$, admet un groupe de Rhin-Viola d'ordre 32, isomorphe à $(V \times V) \rtimes \mathbb{Z}/2\mathbb{Z}$, qui est engendré par deux changements de variables σ_K et λ et une transformation hypergéométrique χ_K.*

Question : Ce paragraphe ressemble beaucoup au paragraphe 4.3. Peut-on trouver une explication à cette similitude ?

5.5 Lien avec les intégrales de Vasilyev

Définitions

Dans ce paragraphe, on définit une famille d'intégrales qui contient les $I(\underline{a}, \underline{b}, c)$ de la partie 3 et qui correspond, par un changement de variables, aux intégrales $K(\underline{A}, \underline{B}, \underline{C})$.

À tout point $\underline{x} = (x_1, \ldots, x_n) \in [0,1]^n$ et à tout entier $k \in \{0, \ldots, n\}$ on associe le réel

$$D_k(\underline{x}) = \sum_{j=0}^{k} (-1)^j x_n x_{n-1} \cdot \ldots \cdot x_{n-j+1}.$$

On a alors $D_0(\underline{x}) = 1$, $D_1(\underline{x}) = 1 - x_n$, $D_2(\underline{x}) = 1 - x_n(1 - x_{n-1})$, $D_n(\underline{x}) = \delta_n(\underline{x})$ et, pour tout $k \in \{0, \ldots, n\}$:

$$\begin{aligned} D_k(\underline{x}) &= \delta_k(x_{n+1-k}, \ldots, x_{n-1}, x_n) \\ &= 1 - x_n(1 - x_{n-1}(1 - \ldots (1 - x_{n+1-k}))). \end{aligned}$$

En outre, on dispose de la relation de récurrence suivante :

$$D_{k+1}(\underline{x}) = D_k(\underline{x}) + (-1)^{k+1} x_n x_{n-1} \cdot \ldots \cdot x_{n-k}.$$

Soient $\underline{a} = (a_1, \ldots, a_n)$, $\underline{b} = (b_1, \ldots, b_n)$ et $\underline{c} = (c_2, \ldots, c_n)$ trois familles d'entiers relatifs. On note $\tilde{\omega}$ la n-forme différentielle sur $[0,1]^n$ définie par :

$$\tilde{\omega} = \frac{\prod_{k \in \{2, \ldots, n-2\} \text{ pair}} D_k(\underline{x})}{\prod_{k \in \{3, \ldots, n-1\} \text{ impair}} D_k(\underline{x})} \frac{\mathrm{d}x_1 \wedge \ldots \wedge \mathrm{d}x_n}{D_n(\underline{x})}.$$

On considère des intégrales de la forme suivante :

$$J(\underline{a},\underline{b},\underline{c}) = \int_{[0,1]^n} \frac{\prod_{k=1}^n x_k^{a_k}(1-x_k)^{b_k}}{\prod_{k=2}^n D_k(\underline{x})^{c_k}} \tilde{\omega}.$$

Les intégrales $J(\underline{a},\underline{b},\underline{c})$ sont une généralisation des intégrales $I(\underline{a},\underline{b},c)$ considérées dans la partie 3, puisqu'on a $D_n(\underline{x}) = \delta_n(\underline{x})$.

N.B. Le changement de variables défini par $u_i = \frac{1}{x_{n+1-i}}$ pour tout $i \in \{1,\dots,n\}$ permet de transformer toute intégrale $J(\underline{a},\underline{b},\underline{c})$ en une intégrale de la forme

$$\int_{]1,+\infty[^n} \frac{\prod_{k=1}^n u_k^{\mathfrak{a}_k}(u_k-1)^{\mathfrak{b}_k}}{\prod_{k=2}^n \delta_k(\underline{u})^{\mathfrak{c}_k}} \mathrm{d}u_1 \wedge \ldots \wedge \mathrm{d}u_n,$$

grâce à la relation $D_k(\underline{x}) = (-1)^k x_n x_{n-1} \cdot \ldots \cdot x_{n+1-k} \delta_k(\underline{u})$ valable pour tout $k \in \{1,\dots,n\}$. Ces intégrales ressemblent à celles de la quatrième partie de ce texte, car elles font intervenir les $\delta_k(\underline{u})$. Mais le domaine d'intégration est différent ; cette remarque ne constitue donc pas une réponse à la question posée à la fin du paragraphe 5.4.

Énoncé et démonstration du théorème principal

Ce théorème fait le lien entre les intégrales $J(\underline{a},\underline{b},\underline{c})$ définies ci-dessus et les intégrales $K(\underline{A},\underline{B},\underline{C})$ introduites au paragraphe 5.1. On le formule ici avec des exposants entiers, mais il se généralise au cas où les exposants sont complexes.

Théorème 5.7 *Soit $(\underline{a},\underline{b},\underline{c}) \in \mathbb{Z}^{3n-1}$. Définissons $(\underline{A},\underline{B},\underline{C})$ par les formules suivantes (avec la convention $a_0 = 0$) :*

$$\begin{array}{rcl}
A_k &=& a_{n+1-k} \text{ pour tout } k \in \{1,\dots,n\} \\
B_1 &=& a_{n-1} + b_n - c_2 - c_3 - \ldots - c_n \\
B_k &=& b_{n+1-k} \text{ pour tout } k \in \{2,\dots,n\} \\
C_k &=& a_{n+1-k} + b_{n+1-k} - c_k - c_{k+1} - \ldots - c_n \text{ pour tout } k \in \{2,\dots,n\} \text{ pair} \\
C_k &=& c_k + c_{k+1} + \ldots + c_n - a_{n-k} \text{ pour tout } k \in \{3,\dots,n\} \text{ impair}
\end{array}$$

Alors on a $J(\underline{a},\underline{b},\underline{c}) = K(\underline{A},\underline{B},\underline{C})$, c'est-à-dire :

$$\begin{aligned}
& \int_{[0,1]^n} \frac{\prod_{k=1}^n x_k^{a_k}(1-x_k)^{b_k}}{\prod_{k=2}^n D_k(\underline{x})^{c_k}} \frac{\prod_{k\in\{2,\dots,n-2\} \text{ pair}} D_k(\underline{x})}{\prod_{k\in\{3,\dots,n-1\} \text{ impair}} D_k(\underline{x})} \frac{\mathrm{d}x_1 \ldots \mathrm{d}x_n}{D_n(\underline{x})} \\
= & \int_{[0,1]^n} \frac{\prod_{k=1}^n y_k^{A_k}(1-y_k)^{B_k}}{\prod_{k=2}^n (1-\alpha_k(\underline{y}))^{C_k+1}} \mathrm{d}y_1 \ldots \mathrm{d}y_n.
\end{aligned}$$

Comme les fonctions qu'on intègre sont positives, l'une des intégrales du théorème est finie (i.e. convergente) si, et seulement si, l'autre l'est. En particulier, ce théorème permet (grâce à la proposition 5.1) d'obtenir les conditions nécessaires et suffisantes pour qu'une intégrale de la forme $J(\underline{a},\underline{b},\underline{c})$ soit convergente.

Ce théorème permet de traduire les hypothèses du théorème de Zudilin cité au paragraphe 3.5, donnant ainsi des conditions suffisantes sur $(\underline{A},\underline{B},\underline{C})$ pour que $K(\underline{A},\underline{B},\underline{C})$ soit une combinaison linéaire, à coefficients rationnels, de 1 et des $\zeta(j)$ pour j entier, compris entre 2 et n, ayant la même parité que n.

L'application linéaire $(\underline{a}, \underline{b}, \underline{c}) \mapsto (\underline{A}, \underline{B}, \underline{C})$ définie dans ce théorème est bijective, et sa réciproque est définie (avec la convention $A_j = B_j = C_j = 0$ pour $j > n$) par :

$$\begin{aligned}
a_k &= A_{n+1-k} \text{ pour tout } k \in \{1, \ldots, n\} \\
b_k &= B_{n+1-k} \text{ pour tout } k \in \{1, \ldots, n-1\} \\
b_n &= B_1 + B_2 - C_2 \\
c_k &= A_k + B_k - A_{k+2} - C_k - C_{k+1} \text{ pour tout } k \in \{2, \ldots, n\} \text{ pair} \\
c_k &= C_k + C_{k+1} - B_{k+1} \text{ pour tout } k \in \{3, \ldots, n\} \text{ impair}
\end{aligned}$$

Démonstration du théorème 5.7 : Il suffit d'appliquer le changement de variables défini par les formules suivantes :

$$\begin{aligned}
x_p &= y_{n+1-p} \text{ si } p \equiv n \mod 2, \\
\text{et } x_p &= \frac{(1 - \alpha_{n-p}(\underline{y}))y_{n+1-p}}{1 - \alpha_{n+1-p}(\underline{y})} \text{ si } p \not\equiv n \mod 2.
\end{aligned}$$

On a alors :

$$\begin{aligned}
1 - x_p &= \frac{1 - y_{n+1-p}}{1 - \alpha_{n+1-p}(\underline{y})} \text{ si } p \not\equiv n \mod 2, \\
\text{et } D_p(\underline{x}) &= \prod_{j=1}^{p} (1 - \alpha_j(\underline{y}))^{(-1)^{j+1}} \text{ pour tout } p \in \{1, \ldots, n\}.
\end{aligned}$$

La réciproque est donnée par :

$$\begin{aligned}
y_k &= \frac{D_{k-2}(\underline{x}) x_{n+1-k}}{D_k(\underline{x})} \text{ pour tout } k \text{ pair}, \\
\text{et } y_k &= x_{n+1-k} \text{ pour tout } k \text{ impair}.
\end{aligned}$$

On a alors :

$$\begin{aligned}
1 - y_k &= \frac{D_{k-1}(\underline{x})(1 - x_{n+1-k})}{D_k(\underline{x})} \text{ pour tout } k \text{ pair}, \\
\text{et } 1 - \alpha_k(\underline{y}) &= \left(\frac{D_{k-1}(\underline{x})}{D_k(\underline{x})} \right)^{(-1)^k} \text{ pour tout } k \in \{1, \ldots, n\}.
\end{aligned}$$

En outre, ce changement de variables transforme $\tilde{\omega}$ en $\bar{\omega}$, au signe près. En effet, on peut montrer par récurrence sur $k \in \{1, \ldots, n\}$ la relation suivante :

$$\mathrm{d}x_n \wedge \ldots \wedge \mathrm{d}x_{n+1-k} = \frac{\prod_{j \in \{1, \ldots, k-1\} \text{ impair}} (1 - \alpha_j(\underline{y}))}{\prod_{j \in \{2, \ldots, k\} \text{ pair}} (1 - \alpha_j(\underline{y}))^2} \mathrm{d}y_1 \wedge \ldots \wedge \mathrm{d}y_k.$$

Cela termine la démonstration du théorème 5.7.

Corollaires

On obtient comme cas particulier du théorème 5.7 le résultat suivant :

Corollaire 5.8 *L'intégrale* $\int_{[0,1]^n} \frac{\prod_{k=1}^n x_k^N(1-x_k)^N}{\delta_n(x_1,\ldots,x_n)^{N+1}} \mathrm{d}x_1 \ldots \mathrm{d}x_n$ *est égale, pour tout* $N \geq 0$*, à*

$$\int_{[0,1]^n} \frac{\prod_{k=1}^n y_k^N(1-y_k)^N}{\prod_{k\in\{2,\ldots,n\} \ \mathit{pair}} (1-y_1\ldots y_k)^{N+1}} \mathrm{d}y_1 \ldots \mathrm{d}y_n$$

si n *est pair, et à*

$$\int_{[0,1]^n} \frac{\prod_{k=1}^n y_k^N(1-y_k)^N}{(1-y_1\ldots y_n)^{N+1}\prod_{k\in\{2,\ldots,n\} \ \mathit{pair}} (1-y_1\ldots y_k)^{N+1}} \mathrm{d}y_1 \ldots \mathrm{d}y_n$$

si n *est impair.*

En particulier, pour $n = 3$, on obtient que les intégrales (4.3) et (4.6) de l'introduction coïncident pour tout $N \geq 0$: les preuves de Beukers et de Sorokin de l'irrationalité de $\zeta(3)$ utilisent exactement les mêmes formes linéaires en 1 et $\zeta(3)$.

Quand $N = 0$, le corollaire 5.8 montre que l'intégrale $\int_{[0,1]^n} \frac{\mathrm{d}x_1\ldots\mathrm{d}x_n}{\delta_n(x_1,\ldots,x_n)}$ est égale à

$$\int_{[0,1]^n} \frac{\mathrm{d}y_1 \wedge \ldots \wedge \mathrm{d}y_n}{(1-\alpha_2(\underline{y}))\cdot(1-\alpha_4(\underline{y}))\cdot\ldots\cdot(1-\alpha_n(\underline{y}))} \text{ si } n \text{ est pair,}$$

$$\text{et à} \quad \int_{[0,1]^n} \frac{\mathrm{d}y_1 \wedge \ldots \wedge \mathrm{d}y_n}{(1-\alpha_2(\underline{y}))\cdot(1-\alpha_4(\underline{y}))\cdot\ldots\cdot(1-\alpha_{n-1}(\underline{y}))\cdot(1-\alpha_n(\underline{y}))} \text{ si } n \text{ est impair.}$$

Sous cette forme, il est facile d'obtenir un développement en série :

$$\sum_{l_1\geq l_2\geq\ldots\geq l_{n/2}\geq 1} \frac{1}{l_1^2 l_2^2\cdot\ldots\cdot l_{n/2}^2} \text{ si } n \text{ est pair,}$$

$$\text{et} \quad \sum_{l_1\geq l_2\geq\ldots\geq l_{(n-1)/2}\geq l_{(n+1)/2}\geq 1} \frac{1}{l_1^2 l_2^2\cdot\ldots\cdot l_{(n-1)/2}^2 l_{(n+1)/2}} \text{ si } n \text{ est impair.}$$

Il est alors immédiat d'écrire l'intégrale $\int_{[0,1]^n} \frac{\mathrm{d}x_1\ldots\mathrm{d}x_n}{\delta_n(x_1,\ldots,x_n)}$ comme combinaison linéaire, sur $\mathbb{Q}$, de polyzêtas de poids au plus n. Or Vasilyev [V2] a démontré que cette intégrale vaut $2(-1)^{n-1}\mathrm{Li}_n((-1)^{n-1})$. Pour obtenir une nouvelle preuve de cette égalité, il suffit donc de démontrer une certaine relation linéaire sur $\mathbb{Q}$ entre des polyzêtas ; il existe des outils combinatoires à cet effet (voir par exemple [W7]).

En appliquant la proposition 5.4 au corollaire 5.8, on obtient des intégrales qui ressemblent à celles utilisées par Sorokin [S6] et notées (4.7) dans l'introduction :

Corollaire 5.9 *L'intégrale* $\int_{[0,1]^n} \frac{\prod_{k=1}^n x_k^N(1-x_k)^N}{\delta_n(x_1,\ldots,x_n)^{N+1}} \mathrm{d}x_1 \ldots \mathrm{d}x_n$ *est aussi égale à*

$$\int_{[0,1]^n} \frac{(y_1y_2)^{r(N+1)-1}(y_3y_4)^{(r-1)(N+1)-1}\cdots(y_{n-1}y_n)^N\prod_{k=1}^n(1-y_k)^N}{\prod_{k=2}^n(1-y_1\ldots y_k)^{N+1}} \mathrm{d}y_1 \ldots \mathrm{d}y_n$$

si $n = 2r$ *est pair, et à*

$$\int_{[0,1]^n} \frac{y_1^{r(N+1)-1}(y_2y_3)^{r(N+1)-1}(y_4y_5)^{(r-1)(N+1)-1}\cdots(y_{n-1}y_n)^N\prod_{k=1}^n(1-y_k)^N}{\prod_{k=2}^n(1-y_1\ldots y_k)^{N+1}} \mathrm{d}y_1 \ldots \mathrm{d}y_n$$

si $n = 2r+1$ *est impair.*

6 Conclusion

Un prolongement possible à ce texte serait de chercher d'autres équations des variétés qui sont introduites dans les deux premières parties, pour mieux faire ressortir le parallélisme entre ces deux situations et arriver à une compréhension aussi géométrique que possible du groupe de Rhin-Viola. Concrètement, on peut considérer un autre plongement de ces variétés (par exemple en prenant les $x_i - x_{i+1}$ et les $x_i x_{i+1}$ comme coordonnées pour $\mathcal{V}_3$), ou bien remplacer partout les x_i par leurs carrés. En outre, les automorphismes qui apparaissent dans ce texte sont toujours d'une forme assez particulière (qui permet de faire apparaître des facteurs Γ).

Le corollaire 5.8 a été obtenu, indépendamment, par Zlobin [Z5] comme cas particulier d'un résultat général qui écrit n'importe quelle intégrale de la forme $I(\underline{a}, \underline{b}, c)$ comme produit d'un nombre rationnel (qui est un quotient de produits de factorielles) par une intégrale de la forme $K(\underline{A}, \underline{B}, \underline{C})$. En comparant ce résultat général avec notre théorème 5.7, on devrait pouvoir enrichir la structure de groupe sur les intégrales $I(\underline{a}, \underline{b}, c)$.

En commun avec Tanguy Rivoal
Rapport de Recherche 2002-27
LMNO, Université de Caen
À paraître au J. Math. Pures Appl.

Chapitre 5

Approximants de Padé et séries hypergéométriques équilibrées

Résumé. Dans cet article, nous énonçons et résolvons des problèmes d'approximation de Padé nouveaux et très généraux dont les solutions s'expriment à l'aide de séries hypergéométriques : par spécialisation, ces séries permettent de retrouver l'irrationalité de $\zeta(3)$, d'une infinité de $\zeta(2n+1)$, n entier ≥ 1, et essentiellement tous les résultats de cette nature déjà présents dans la littérature. Nous présentons également deux nouvelles applications diophantiennes de notre méthode.

Mots clés. Approximants de Padé, séries hypergéométriques, approximation diophantienne, fonction ζ de Riemann.

Abstract. In this article, we present and solve some very general new Padé approximant problems, whose solutions can be expressed with hypergeometric series. These series appear in the proofs of the irrationality of $\zeta(3)$, of infinitely many $\zeta(2n+1)$, and in essentially all results of this kind in the literature. We also prove two new diophantine results with this method.

Key words. Padé approximants, hypergeometric series, diophantine approximation, Riemann ζ function.

1 Introduction

Les démonstrations données par Apéry en 1978 [A10] de l'irrationalité de $\zeta(2)$ et $\zeta(3)$ sont apparues initialement comme très mystérieuses. Néanmoins, dans [B5] et [B6], Beukers est parvenu à les replacer dans le cadre plus connu des approximants de Padé des polylogarithmes, définis (pour $s \geq 1$ et $|z| < 1$) par le développement en série entière

$$\mathrm{Li}_s(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^s}\ .$$

De façon précise, il considère les deux problèmes suivants : déterminer pour tout entier $n \geq 0$ des polynômes a, b, c de degré au plus n tels que

$$\begin{cases} S(z) = a(z)\mathrm{Li}_2(1/z) + b(z)\mathrm{Li}_1(1/z) + c(z) = \mathrm{O}(z^{-n-1}) \\ R(z) = a(z)\log(z) - b(z) = \mathrm{O}\big((1-z)^{n+1}\big) \end{cases} \tag{5.1}$$

et des polynômes A, B, C et D de degré au plus n tels que[1]

$$\begin{cases} U(z) = A(z)\mathrm{Li}_2(1/z) + B(z)\mathrm{Li}_1(1/z) + C(z) = \mathrm{O}(z^{-n-1}) \\ V(z) = 2A(z)\mathrm{Li}_3(1/z) + B(z)\mathrm{Li}_2(1/z) + D(z) = \mathrm{O}(z^{-n-1}) \\ W(z) = A(z)\log(z) - B(z) = \mathrm{O}(1-z)\ . \end{cases} \tag{5.2}$$

Remarque : étant donnée une fonction $F(w)$ développable en série de Laurent $F(w) = \sum_{n=-m}^{+\infty} a_n w^n$ au voisinage de $w = 0$ (avec $w = z$, $1/z$ ou $1 - z$ dans la suite), on note $F(w) = \mathrm{O}(w^{N+1})$ si $a_{-m} = a_{-m+1} = \cdots = a_N = 0$. Il s'agit d'une majoration quand z tend vers 0, l'infini ou 1 (suivant la valeur de w).

Les solutions de ces deux problèmes et de ceux qui suivent font intervenir les séries hypergéométriques ${}_{q+1}F_q$ définies (pour $q \geq 1$) par :

$${}_{q+1}F_q\left(\begin{matrix} \alpha_0, & \alpha_1, & \ldots, & \alpha_q \\ & \beta_1, & \ldots, & \beta_q \end{matrix} \,\middle|\, z\right) = \sum_{k=0}^{\infty} \frac{(\alpha_0)_k(\alpha_1)_k \cdots (\alpha_q)_k}{(1)_k(\beta_1)_k \cdots (\beta_q)_k} z^k\ ,$$

où les α_j, β_j et z sont des complexes convenables, et $(\alpha)_k = \alpha(\alpha+1)\cdots(\alpha+k-1)$ est le symbole de Pochhammer. Dans les ouvrages traitant de ces fonctions (par exemple [A9]), on trouve les définitions suivantes :

- ${}_{q+1}F_q$ est *quasi équilibrée* si $\alpha_1 + \beta_1 = \cdots = \alpha_q + \beta_q$
- ${}_{q+1}F_q$ est *bien équilibrée* si $\alpha_0 + 1 = \alpha_1 + \beta_1 = \cdots = \alpha_q + \beta_q$
- ${}_{q+1}F_q$ est *très bien équilibrée* si elle est bien équilibrée et $\alpha_1 = \frac{1}{2}\alpha_0 + 1$.

Dans [B6], Beukers indique que la solution de (5.1) est unique (à une constante multiplicative près), donnée par une certaine intégrale que l'on transforme facilement en la série quasi équilibrée suivante[2] :

$$\begin{aligned} S(z) &= n! \sum_{k=1}^{\infty} \frac{(k-n)_n}{(k)_{n+1}^2} z^{-k} & (5.3) \\ &= z^{-n-1} \frac{n!^4}{(2n+1)!^2}\ {}_3F_2\left(\begin{matrix} n+1, & n+1 & , & n+1 \\ & 2n+2 & , & 2n+2 \end{matrix} \,\middle|\, z^{-1}\right)\ . & (5.4) \end{aligned}$$

Remarque : la série (5.3) n'est pas sous forme hypergéométrique. On passe à la forme (5.4) en appliquant des formules telles que $(\alpha)_k = \frac{k!}{(\alpha-1)!}(k+1)_{\alpha-1}$ (pour $\alpha \geq 1$ et $k \geq 0$). Cette remarque s'applique à toutes les séries ci-dessous.

---

[1] Beukers n'énonce pas la condition sur $W(z)$ mais la condition équivalente $B(1) = 0$.

[2] Bien que cela soit un cas particulier du Théorème 1.1 ci-dessous, nous donnons une démonstration alternative de (5.4) au § 5, en adaptant une technique due à Sorokin [S7].

Beukers montre aussi [B5] que (5.2) a une solution unique (à une constante multiplicative près) et son argument donne immédiatement :

$$U(z) = \sum_{k=1}^{\infty} \frac{(k-n)_n^2}{(k)_{n+1}^2} z^{-k} \tag{5.5}$$

$$= z^{-n-1} \frac{n!^4}{(2n+1)!^2} \; {}_4F_3\left(\begin{matrix} n+1, & n+1 & , & n+1 & , & n+1 \\ & 1 & , & 2n+2 & , & 2n+2 \end{matrix} \middle| z^{-1} \right) \tag{5.6}$$

et

$$V(z) = -\sum_{k=1}^{\infty} \frac{\mathrm{d}}{\mathrm{d}k} \left(\frac{(k-n)_n^2}{(k)_{n+1}^2} \right) z^{-k} . \tag{5.7}$$

Armé de ces solutions explicites, on peut déduire les théorèmes d'Apéry (voir par exemple [B4] et [N2]). Plus récemment, en cherchant une démonstration élémentaire de l'irrationalité de $\zeta(3)$, K. Ball a introduit la série

$$\begin{aligned} B_n &= n!^2 \sum_{k=1}^{\infty} \left(k + \frac{n}{2}\right) \frac{(k-n)_n (k+n+1)_n}{(k)_{n+1}^4} \\ &= \frac{n!^7 (3n+2)!}{2(2n+1)!^5} \; {}_7F_6\left(\begin{matrix} 3n+2, & \frac{3}{2}n+2, & n+1 & ,\ldots, & n+1 \\ & \frac{3}{2}n+1, & 2n+2 & ,\ldots, & 2n+2 \end{matrix} \middle| 1 \right) \end{aligned}$$

qui vérifie $B_n = \alpha_n \zeta(3) + \beta_n$ pour certains rationnels α_n et β_n, alors que l'on s'attend aussi à voir apparaître $\zeta(2)$ et $\zeta(4)$. Ce remarquable phénomène est dû à la nature très[3] bien équilibrée de B_n.

Généralisant cette construction, le second auteur de cet article a introduit ([R6], chapitre 3) la série bien équilibrée suivante, pour $1 \le r \le a/2$:

$$\begin{aligned} S_{n,a,r}(z) &= n!^{a-2r} \sum_{k=1}^{\infty} \frac{(k-rn)_{rn}(k+n+1)_{rn}}{(k)_{n+1}^a} z^{-k} \\ &= z^{-rn-1} n!^{a-2r} \frac{(rn)!^{a+1}((2r+1)n+1)!}{((r+1)n+1)!^{a+1}} \\ &\quad \times {}_{a+2}F_{a+1}\left(\begin{matrix} (2r+1)n+2, & rn+1 & ,\ldots, & rn+1 \\ & (r+1)n+2 & ,\ldots, & (r+1)n+2 \end{matrix} \middle| z^{-1} \right) \end{aligned}$$

qui se décompose elle aussi en polylogarithmes et permet de prouver qu'une infinité des nombres $\zeta(2j+1)$ sont irrationnels (voir également [R5] et [B2]).

Pour montrer des résultats d'indépendance linéaire de valeurs de polylogarithmes aux rationnels, la série quasi équilibrée suivante (avec $1 \le r \le a$), qui généralise à la fois (5.4) et une construction de Nikishin [N4], est également introduite dans [R6], chapitre 2 :

$$\begin{aligned} N_{n,a,r}(z) = n!^{a-r} \sum_{k=1}^{\infty} \frac{(k-rn)_{rn}}{(k)_{n+1}^a} z^{-k} &= z^{-rn-1} n!^{a-r} \frac{(rn)!^{a+1}}{((r+1)n+1)!^a} \\ &\times {}_{a+1}F_a\left(\begin{matrix} rn+1, & rn+1 & ,\ldots, & rn+1 \\ & (r+1)n+2 & ,\ldots, & (r+1)n+2 \end{matrix} \middle| z^{-1} \right) . \end{aligned}$$

[3]La disparition de la moitié des valeurs de ζ attendues est en fait liée seulement à l'aspect bien équilibré de la série.

À la page 53 de [R6], est posée la question d'expliciter si possible des problèmes de Padé tels que (5.1) et (5.2) dont la solution ferait intervenir ces trois types de séries hypergéométriques. Dans cet article, nous répondons positivement à cette question en construisant et résolvant des problèmes de Padé très généraux, sur lesquels se lisent en outre les propriétés de réciprocité des solutions dans le cas bien équilibré. Commençons par décrire le cas des séries quasi et bien équilibrées.

Considérons des entiers $n \geq 0$, $a \geq 1$ et $\rho, \sigma \geq 0$ vérifiant $\rho + \sigma \leq a(n+1) - 1$ et que nous supposons fixés. Nous voulons déterminer des polynômes P_0, $\overline{P}_0$ et P_j (pour $1 \leq j \leq a$) de degré au plus n et des fonctions S, $\overline{S}$, R (qui dépendent de n, a, ρ, σ) tels que

$$\begin{cases} S(z) = P_0(z) + \displaystyle\sum_{j=1}^{a} P_j(z)\mathrm{Li}_j(1/z) = \mathrm{O}(z^{-\rho-1}) \\ \overline{S}(z) = \overline{P}_0(z) + \displaystyle\sum_{j=1}^{a} (-1)^j P_j(z)\mathrm{Li}_j(z) = \mathrm{O}(z^{n+\sigma+1}) \\ R(z) = \displaystyle\sum_{j=1}^{a} (-1)^{j-1} P_j(z) \frac{\log^{j-1}(z)}{(j-1)!} = \mathrm{O}\big((1-z)^{a(n+1)-\rho-\sigma-1}\big) . \end{cases} \tag{5.8}$$

(Ici et dans toute la suite, la fonction $\log(z)$ est définie avec sa détermination principale.)

En fait le problème est de trouver les polynômes $P_1, \ldots, P_a$. En effet, quand ceux-ci sont fixés, il existe au plus un choix pour P_0, $\overline{P}_0$, S, $\overline{S}$ et R. De plus, comme $\mathrm{Li}_j(0) = 0$ pour tout $j \geq 1$, on aura automatiquement $\deg(P_0) \leq n-1$ et $\overline{P}_0(0) = 0$.

Remarque : tous les problèmes de Padé de cet article se traduisent par un système d'équations linéaires dont les inconnues sont les coefficients des polynômes. Il y aura toujours une inconnue de plus que d'équations, ce qui implique l'existence d'au moins une solution non identiquement nulle. Nous montrerons que la matrice du système est de rang maximal, c'est-à-dire que la solution du problème est unique à une constante multiplicative près, ce qui sera le sens de l'expression "unicité" dans ce texte ; un tel problème est aussi dit *normal.*

Théorème 1.1 *À constante multiplicative près, le problème de Padé* (5.8) *a une unique solution, et elle vérifie pour tout $z \in \mathbb{C}$ tel que $|z| \geq 1$,*

$$S(z) = \sum_{k=1}^{\infty} \frac{(k-\rho)_\rho (k+n+1)_\sigma}{(k)_{n+1}^a} z^{-k} , \tag{5.9}$$

$$P_a(z) = \sum_{k=0}^{n} (-1)^{ka} \frac{(-k-\rho)_\rho (n-k+1)_\sigma}{k!^a (n-k)!^a} z^k \tag{5.10}$$

et, si $z \notin]-\infty, 0]$,

$$R(z) = \frac{1}{2i\pi} \int_{\mathcal{C}} \frac{(s-\rho)_\rho (s+n+1)_\sigma}{(s)_{n+1}^a} z^{-s} \mathrm{d}s \tag{5.11}$$

où $\mathcal{C}$ est un lacet entourant les points $-n, -n+1, \ldots, 0$ dans le sens direct.

Ce résultat concerne des séries quasi équilibrées ; dans le cas particulier $\rho = \sigma$ elles sont bien équilibrées. En effet, on peut écrire (5.9) sous la forme :

$$S(z) = z^{-\rho-1} \frac{\rho!^{a+1}(\rho+\sigma+n+1)!}{(\rho+n+1)!^{a+1}} \times {}_{a+2}F_{a+1}\left(\begin{matrix} \rho+\sigma+n+2, & \rho+1 & ,\ldots, & \rho+1 \\ & \rho+n+2 & ,\ldots, & \rho+n+2 \end{matrix} \middle| z^{-1} \right).$$

Remarque : On peut démontrer [H12] ce théorème en calculant les exposants de l'équation différentielle associée. Mais dans la suite on utilise une méthode complètement différente.

Proposition 1.2 *Pour chaque couple (ρ, σ), notons $(P_{j,\rho,\sigma}(z))_{1\le j\le a}$ la solution du problème (5.8) associé à (n, a, ρ, σ) qui vérifie la normalisation (5.10), et $S_{\rho,\sigma}(z)$, $\overline{S}_{\rho,\sigma}(z)$, $R_{\rho,\sigma}(z)$ les quantités correspondantes. On a alors :*

$$\begin{aligned} z^n P_{j,\sigma,\rho}(1/z) &= (-1)^{a(n+1)+\rho+\sigma+j} P_{j,\rho,\sigma}(z) \text{ pour } j \in \{1,\ldots,a\} \\ z^n \overline{S}_{\sigma,\rho}(1/z) &= (-1)^{a(n+1)+\rho+\sigma} S_{\rho,\sigma}(z) \\ z^n R_{\sigma,\rho}(1/z) &= (-1)^{a(n+1)+\rho+\sigma+1} R_{\rho,\sigma}(z) \end{aligned}$$

En particulier quand $\sigma = \rho$, on a pour $j \in \{1,\ldots,a\}$:

$$z^n P_{j,\rho,\rho}(1/z) = (-1)^{a(n+1)+j} P_{j,\rho,\rho}(z) \text{ et } z^n \overline{S}_{\rho,\rho}(1/z) = (-1)^{a(n+1)} S_{\rho,\rho}(z).$$

Remarque : Cette proposition affirme que si $a(n+1)+j$ est impair alors $P_{j,\rho,\rho}(1)$ est nul. Donc quand $a(n+1)$ est impair, resp. pair, $S_{\rho,\rho}(1)$ est une forme linéaire à coefficients rationnels en 1, $\zeta(3)$, $\zeta(5)$, …, $\zeta(a)$, resp. 1, $\zeta(2)$, $\zeta(4)$, …, $\zeta(2[a/2])$: c'est ce qui apparaît dans [R5] et [B2]. Le cas le plus simple présentant une telle dichotomie correspond à $\rho = 0$, et nous n'en avons pas trouvé mention dans la littérature :

$$S_{0,0}(1) = \sum_{k=1}^{\infty} \frac{1}{(k)_{n+1}^a}.$$

Démonstration de la Proposition 1.2. En changeant z en $1/z$ on voit que la famille $((-1)^j z^n P_{j,\sigma,\rho}(1/z))_{1\le j\le a}$ est solution du problème (5.8) associé à (n, a, ρ, σ), et les fonctions associées sont $z^n\overline{S}_{\sigma,\rho}(1/z)$, $z^n S_{\sigma,\rho}(1/z)$ et $-z^n R_{\sigma,\rho}(1/z)$. D'après le Théorème 1.1, cette solution est proportionnelle à $(P_{j,\rho,\sigma}(z))_{1\le j\le a}$. Le coefficient de proportionnalité s'obtient à l'aide de (5.10) et de l'identité triviale $(\alpha)_k = (-1)^k(-\alpha-k+1)_k$.

Une conséquence de ces considérations est donnée par les deux théorèmes suivants.

Théorème 1.3 *Soient λ, μ deux réels et α un rationnel tels que $(\lambda,\mu) \neq (0,0)$ et $0 < \alpha < 1$. L'ensemble*

$$\left\{ \lambda \mathrm{Li}_s(\alpha) + \mu \frac{\log^s(\alpha)}{(s-1)!}, s \in \mathbb{N}^* \right\}$$

contient une infinité de nombres linéairement indépendants sur $\mathbb{Q}$.

Quand $\lambda \neq 0$, on peut supposer $\lambda = 1$ dans l'énoncé de ce théorème. Lorsque $\mu = 0$, on retrouve le théorème démontré au chapitre 2 de [R6]. Pour $\lambda = 0$, on en déduit la transcendance de $\log(\alpha)$ et dans ce cas, la méthode est essentiellement celle utilisée par Reyssat [R2] pour obtenir une mesure de transcendance de $\log(\alpha)$. Notre résultat n'apporte malheureusement aucun éclairage sur l'éventuelle indépendance linéaire de $\mathrm{Li}_s(\alpha)$ et $\log^s(\alpha)$ pour tout entier $s \geq 2$. En effet, la conclusion reste vraie si pour tout entier $s \geq 2$ et tout rationnel $0 < \alpha < 1$, il existe des rationnels $c_s(\alpha)$ et $d_s(\alpha)$ tels que $\mathrm{Li}_s(\alpha) = c_s(\alpha)\log^s(\alpha) + d_s(\alpha)$.

Remarque : Le Théorème 1.3 reste valable pour $-1 \leq \alpha < 0$, en choisissant une détermination convenable du logarithme (voir la remarque qui suit la Proposition 2.1). En utilisant des séries bien équilibrées, et $\alpha = -1$, on obtient un résultat analogue pour l'ensemble des $\lambda \mathrm{Li}_s(-1) + i\mu \dfrac{(i\pi)^s}{(s-1)!}$, s entier impair.

Théorème 1.4 *Au moins l'un des trois nombres*

$$\mathrm{Li}_2(1/2) + \log^2(2)\,, \quad \mathrm{Li}_3(1/2) - \frac{1}{2}\log^3(2)\,, \quad \mathrm{Li}_4(1/2) + \frac{1}{6}\log^4(2)$$

est irrationnel.

Remarques : En appliquant la méthode de Hata [H4], on pourrait peut-être éliminer $\mathrm{Li}_4(1/2)+\frac{1}{6}\log^4(2)$ du résultat précédent. Remarquons par ailleurs que $\mathrm{Li}_2(1/2) = \frac{1}{2}\zeta(2) - \frac{1}{2}\log^2(2)$ et $\mathrm{Li}_3(1/2) = \frac{7}{8}\zeta(3) - \frac{1}{2}\zeta(2)\log(2) + \frac{1}{6}\log^3(2)$ (voir [L4]).

La démonstration du Théorème 1.1 est donnée au § 2. Elle se décompose en trois lemmes, chacun traduisant l'une des équations du problème (5.8). Au § 3, nous énonçons et résolvons (Théorème 3.1) un autre problème de Padé permettant de traiter le cas de séries très bien équilibrées. Au § 4, nous généralisons les Théorèmes 1.1 et 3.1 de manière à englober les problèmes du type (5.2) : on obtient ainsi les Théorèmes 4.1 et 4.2. Au § 5, nous donnons une solution alternative au problème de Padé (5.1), sans passer par le Théorème 1.1. Enfin, aux § 6 et § 7, nous démontrons les Théorèmes 1.3 et 1.4.

2 Séries quasi et bien équilibrées

Démonstration du Théorème 1.1. Il est plus commode de chercher à résoudre le problème équivalent

$$\begin{cases} S(z) = P_0(z) + \displaystyle\sum_{j=1}^{a} P_j(z)\mathrm{Li}_j(1/z) = \mathrm{O}(z^{-\rho-1}) \\ \tilde{S}(z) = z^n\overline{P_0}(1/z) + \displaystyle\sum_{j=1}^{a} (-1)^j z^n P_j(1/z)\mathrm{Li}_j(1/z) = \mathrm{O}(z^{-\sigma-1}) \\ R(z) = \displaystyle\sum_{j=1}^{a} (-1)^{j-1} P_j(z)\frac{\log^{j-1}(z)}{(j-1)!} = \mathrm{O}\big((1-z)^{a(n+1)-\rho-\sigma-1}\big) \end{cases} \tag{5.12}$$

avec $\tilde{S}(z) = z^n\overline{S}(1/z)$.

Fixons les notations suivantes, dans lesquelles on ne présage pas du fait que les P_j forment une solution. Pour $j \in \{1, \ldots, a\}$ on écrit :

$$P_j(z) = \sum_{t=0}^{n} p_{j,t} z^t.$$

On a :

$$\begin{aligned} \sum_{j=1}^{a} P_j(z) \mathrm{Li}_j(1/z) &= \sum_{j=1}^{a} \sum_{t=0}^{n} p_{j,t} \sum_{m \geq 1} \frac{z^{-(m-t)}}{m^j} \\ &= \sum_{k=1-n}^{+\infty} z^{-k} \sum_{j=1}^{a} \sum_{t=\max(0,1-k)}^{n} \frac{p_{j,t}}{(t+k)^j} \,. \end{aligned} \tag{5.13}$$

Pour $k \geq 1$, le coefficient de z^{-k} dans cette série est donné par la fraction rationnelle

$$A(k) = \sum_{j=1}^{a} \sum_{t=0}^{n} \frac{p_{j,t}}{(t+k)^j} \tag{5.14}$$

en la variable k, qu'on peut aussi écrire sous la forme $\dfrac{Q(k)}{(k)_{n+1}^a}$ où le polynôme $Q(k)$ est de degré strictement inférieur à $a(n+1)$, car il n'y a pas de partie principale dans l'équation (5.14). D'après l'unicité de la décomposition en éléments simples, la donnée de $P_1, \ldots, P_a$ est équivalente à la donnée du polynôme Q. Les Lemmes 2.1 à 2.3 ci-dessous montrent que le seul polynôme Q qui donne une solution de (5.12) est (à constante multiplicative près)

$$Q(k) = (k-\rho)_\rho (k+n+1)_\rho.$$

Cela démontre "l'unicité" de la solution et la relation (5.9). On en déduit facilement (5.10) en calculant les coefficients $p_{a,t}$ grâce à (5.14). L'expression de $R(z)$ découle du théorème des résidus appliqué à l'intégrale de (5.11).

Lemme 2.1 *Les polynômes $P_1, \ldots, P_a$ satisfont à la première condition du système* (5.12) *(avec P_0 bien choisi) si, et seulement si, le polynôme*

$$\prod_{i=1}^{\rho} (k-i) = (k-\rho)_\rho \quad \text{divise} \quad Q(k)\,.$$

Démonstration. Le coefficient de z^{-k} dans l'équation (5.13) est nul pour tout $k \in \{1, \ldots, \rho\}$ si, et seulement si, A s'annule aux points $1, \ldots, \rho$.

Lemme 2.2 *Les polynômes $P_1, \ldots, P_a$ satisfont à la deuxième condition du système* (5.12) *(avec $\overline{P_0}$ bien choisi) si, et seulement si, le polynôme*

$$\prod_{i=n+1}^{n+\sigma} (k+i) = (k+n+1)_\sigma \quad \text{divise} \quad Q(k)\,.$$

Démonstration. On a :

$$
\begin{aligned}
&\sum_{j=1}^{a}(-1)^j z^n P_j(1/z)\mathrm{Li}_j(1/z)\\
&= \sum_{j=1}^{a}(-1)^j \sum_{t=0}^{n} p_{j,n-t} \sum_{m\geq 1} \frac{z^{-(m-t)}}{m^j}\\
&= \sum_{k=1-n}^{+\infty} z^{-k} \sum_{j=1}^{a}(-1)^j \sum_{t=\max(0,1-k)}^{n} \frac{p_{j,n-t}}{(t+k)^j}\\
&= \sum_{k=1-n}^{+\infty} z^{-k} \sum_{j=1}^{a} \sum_{\ell=0}^{\min(n,n+k-1)} (-1)^j \frac{p_{j,\ell}}{(n+k-\ell)^j}\ .
\end{aligned}
\tag{5.15}
$$

Pour $k \geq 1$, le coefficient de z^{-k} dans la série $\tilde{S}(z)$ est donné par l'équation (5.15), et vaut $A(-k-n)$. Donc l'existence de $\overline{P_0}$ vérifiant la deuxième condition du système (5.12) équivaut à l'annulation de A en $-n-1, \ldots, -n-\sigma$, ce qui termine la preuve du Lemme 2.2.

Lemme 2.3 *Soit D un entier inférieur ou égal à $a(n+1)-1$. Alors $\deg(Q) \leq D$ si, et seulement si, quand z tend vers 1, on a :*

$$R(z) = \mathrm{O}\big((1-z)^{a(n+1)-D-1}\big)$$

En particulier si $D = a(n+1)-1$ alors les deux assertions sont trivialement vraies. Si $D = a(n+1)-2$ on retrouve une observation qui apparaît dans [B5] : A est sans résidu à l'infini si, et seulement si, $P_1(1) = 0$. Pour démontrer le Théorème 1.1 on utilise ce Lemme avec $D = \rho + \sigma$.

Démonstration du Lemme 2.3. Dans un premier temps, on traduit la condition $\deg(Q) \leq D$. On a :

$$\frac{1}{(k+i)^j} = \sum_{\ell\geq 0} \binom{\ell+j-1}{\ell} \frac{(-i)^\ell}{k^{\ell+j}}$$

d'où, en posant $d = \ell + j$, le développement asymptotique de $A(k)$ à l'infini :

$$A(k) = \sum_{d\geq 1} \frac{\mathfrak{A}_d}{k^d} \tag{5.16}$$

en notant

$$\mathfrak{A}_d = (-1)^d \sum_{i=0}^{n} \sum_{j=1}^{\min(a,d)} (-1)^j \binom{d-1}{d-j} i^{d-j} p_{j,i}\ . \tag{5.17}$$

La condition $\deg(Q) \leq D$ se traduit par :

$$\mathfrak{A}_d = 0 \text{ pour tout } d \in \{1, \ldots, a(n+1)-D-1\}\ .$$

On déduit alors immédiatement le Lemme 2.3 de la proposition suivante :

Proposition 2.1 *On a, pour tout $z \in \mathbb{C}\backslash]-\infty, 0]$:*

$$R(z) = \sum_{d\geq 1} (-1)^{d-1} \mathfrak{A}_d \frac{\log^{d-1}(z)}{(d-1)!}.$$

Remarque : la série du membre de droite converge uniformément en z, sur tout compact de $\mathbb{C}\backslash]-\infty, 0]$. En outre, on peut remplacer $\mathbb{C}\backslash]-\infty, 0]$ par n'importe quel ouvert simplement connexe qui contient 1 et pas 0, en prenant sur cet ouvert la détermination du logarithme qui s'annule en 1. Bien entendu, il faut effectuer ce changement à la fois dans (5.12) (qui définit $R(z)$) et dans la Proposition 2.1.

Première démonstration de la Proposition 2.1. Posons $x = \log(z)$. On a :

$$\begin{aligned}
R(e^x) &= \sum_{j=1}^{a} P_j(e^x)(-1)^{j-1} \frac{x^{j-1}}{(j-1)!} \\
&= \sum_{j=1}^{a} \sum_{t=0}^{n} p_{j,t} e^{tx} (-1)^{j-1} \frac{x^{j-1}}{(j-1)!} \\
&= \sum_{j=1}^{a} \sum_{t=0}^{n} \sum_{\ell\geq 0} p_{j,t} (-1)^{j-1} \frac{t^\ell x^{\ell+j-1}}{(j-1)!\ell!} \\
&= \sum_{d\geq 1} \frac{x^{d-1}}{(d-1)!} \Big(\sum_{t=0}^{n} \sum_{j=1}^{\min(a,d)} \binom{d-1}{j-1} t^{d-j} p_{j,t} (-1)^{j-1} \Big) \\
&= \sum_{d\geq 1} (-1)^{d-1} \mathfrak{A}_d \frac{x^{d-1}}{(d-1)!},
\end{aligned}$$

en utilisant la formule (5.17).

Deuxième démonstration de la Proposition 2.1. Un calcul de résidus montre qu'on a :

$$R(z) = \frac{1}{2i\pi} \int_C A(s) z^{-s} \mathrm{d}s, \tag{5.18}$$

où C désigne le cercle de centre 0 et de rayon $2n$, parcouru dans le sens positif. En posant $u = 1/s$ on obtient :

$$R(z) = \frac{1}{2i\pi} \int_{C'} A(1/u) \exp(-\frac{\log(z)}{u}) \frac{\mathrm{d}u}{u^2}$$

en notant C' le cercle de centre 0 et de rayon $\frac{1}{2n}$, parcouru dans le sens direct. Grâce au développement $A(1/u) = \sum_{d\geq 1} \mathfrak{A}_d u^d$, valable quand u parcourt C', on a :

$$R(z) = \frac{1}{2i\pi} \sum_{d\geq 1} \sum_{\ell\geq 0} (-1)^\ell \frac{\log^\ell(z)}{\ell!} \mathfrak{A}_d \int_{C'} u^{d-\ell-2} \mathrm{d}u.$$

Or l'intégrale est nulle sauf quand $d = \ell + 1$, et dans ce cas elle vaut $2i\pi$. On a donc démontré la Proposition 2.1.

Remarque : de la relation (5.18) on déduit immédiatement que l'ordre d'annulation de $R(z)$ en $z = 1$ est égal à l'ordre d'annulation de $A(k)$ en l'infini, diminué de 1. Ceci démontre le Lemme 2.3, sans passer par la proposition 2.1.

3 Séries très bien équilibrées

Considérons des entiers $n \geq 0$, $a \geq 1$ et $\rho \geq 0$ vérifiant $2\rho \leq a(n+1)-2$ et que nous supposons fixés. Nous désirons maintenant résoudre le problème d'approximation de Padé suivant : déterminer des polynômes P_0, $\overline{P}_0$ et P_j (pour $1 \leq j \leq a$) de degré au plus n (dépendant aussi de a et ρ) et des fonctions S, $\overline{S}$, R tels que

$$\begin{cases} S(z) = P_0(z) + \sum_{j=1}^{a} P_j(z)\mathrm{Li}_j(1/z) = \mathrm{O}(z^{-\rho-1}) \\ \overline{S}(z) = \overline{P}_0(z) + \sum_{j=1}^{a} (-1)^j P_j(z)\mathrm{Li}_j(z) = \mathrm{O}(z^{n+\rho+1}) \\ R(z) = \sum_{j=1}^{a} (-1)^{j-1} P_j(z) \dfrac{\log^{j-1}(z)}{(j-1)!} = \mathrm{O}\big((1-z)^{a(n+1)-2\rho-2}\big) \\ P_a((-1)^a) = 0\ . \end{cases} \tag{5.19}$$

Théorème 3.1 *À une constante multiplicative près, le problème de Padé* (5.19) *a une solution unique, et elle vérifie*

$$S(z) = \sum_{k=1}^{\infty} \left(k + \frac{n}{2}\right) \frac{(k-\rho)_\rho (k+n+1)_\rho}{(k)_{n+1}^a} z^{-k} \tag{5.20}$$

et

$$P_a(z) = \sum_{k=0}^{n} (-1)^{ka} \left(\frac{n}{2} - k\right) \frac{(-k-\rho)_\rho (n-k+1)_\rho}{k!^a (n-k)!^a} z^k\ . \tag{5.21}$$

De plus, pour tout $j = 1, \ldots, a$, on a

$$z^n P_j(1/z) = (-1)^{a(n+1)+j+1} P_j(z) \quad \text{et} \quad z^n \overline{S}(1/z) = (-1)^{a(n+1)+1} S(z)\ . \tag{5.22}$$

Remarque : on peut écrire (5.20) comme une série hypergéométrique très bien équilibrée :

$$S(z) = \frac{1}{2} z^{-\rho-1} \frac{\rho!^{a+1}(2\rho+n+2)!}{(\rho+n+1)!^{a+1}} \times {}_{a+3}F_{a+2}\left(\begin{matrix} 2\rho+n+2, & \frac{1}{2}n+\rho+2, & \rho+1 & , \ldots , & \rho+1 \\ & \frac{1}{2}n+\rho+1, & \rho+n+2 & , \ldots , & \rho+n+2 \end{matrix} \middle| z^{-1} \right)\ .$$

Démonstration. Reprenons les notations du § 2. Les Lemmes 2.1, 2.2 et 2.3 montrent que les solutions sont données par

$$S(z) = \sum_{k=1}^{\infty} \frac{Q(k)}{(k)_{n+1}^a} z^{-k},$$

avec un polynôme $Q(k)$ de degré inférieur ou égal à $2\rho+1$ et divisible par $(k-\rho)_\rho(k+n+1)_\rho$. Pour toute solution il existe donc un polynôme $\pi(k)$ de degré au plus 1 tel que

$$Q(k) = \pi(k)(k-\rho)_\rho(k+n+1)_\rho\ .$$

En développant en éléments simples la fraction rationnelle

$$\pi(k)\frac{(k-\rho)_\rho(k+n+1)_\rho}{(k)_{n+1}^a} ,$$

on voit que

$$\begin{aligned} P_a(z) &= \sum_{k=0}^{n}(-1)^{ka}\pi(-k)\frac{(-k-\rho)_\rho(n-k+1)_\rho}{k!^a(n-k)!^a}z^k \\ &= (-1)^\rho\sum_{k=0}^{n}(-1)^{ka}\pi(-k)\frac{(k+1)_\rho(n-k+1)_\rho}{k!^a(n-k)!^a}z^k . \end{aligned}$$

Sur cette formule, il est clair que si $\pi(k)=1$ pour tout k alors $P_a((-1)^a)\neq 0$ (car c'est une somme de termes non nuls du même signe), et que si $\pi(k)=k+\frac{n}{2}$ alors $P_a((-1)^a)=0$ (car le changement de k en $n-k$ change $P_a((-1)^a)$ en son opposé). Dans le cas général, on peut écrire $\pi(k)=\alpha+\beta(k+\frac{n}{2})$. La condition $P_a((-1)^a)=0$ se traduit alors par $\alpha=0$: on a démontré que le problème (5.19) a une solution unique (à constante multiplicative près), et qu'elle vérifie (5.20) et (5.21). On en déduit (5.22) comme dans la preuve de la Proposition 1.2.

4 Généralisations

Nous considérons maintenant deux problèmes de Padé qui englobent le problème (5.2) et font intervenir des séries "dérivées" telles que (5.7). Considérons des entiers $n\geq 0$, $a\geq 1$, $L,M\geq 0$ et $\rho,\sigma\geq 0$ vérifiant $L\rho+M\sigma\leq a(n+1)-1$ et que nous supposons fixés. Nous voulons déterminer des polynômes $P_{0,\ell}$, $\overline{P}_{0,m}$, P_j (pour $0\leq\ell\leq L-1$, $0\leq m\leq M-1$ et $1\leq j\leq a$) de degré au plus n et des fonctions S_ℓ, $\overline{S}_m$ et R (qui dépendent aussi de ρ, σ, L, M, a, n) tels que l'on ait les $M+L+1$ conditions simultanées pour $\ell=0,\dots,L-1$ et $m=0,\dots,M-1$:

$$\begin{cases} S_\ell(z):=P_{0,\ell}(z)+\displaystyle\sum_{j=1}^{a}\binom{\ell+j-1}{j-1}P_j(z)\mathrm{Li}_{\ell+j}(1/z)=\mathrm{O}(z^{-\rho-1}) \\ \overline{S}_m(z):=\overline{P}_{0,m}(z)+\displaystyle\sum_{j=1}^{a}(-1)^j\binom{m+j-1}{j-1}P_j(z)\mathrm{Li}_{m+j}(z)=\mathrm{O}(z^{n+\sigma+1}) \\ R(z):=\displaystyle\sum_{j=1}^{a}(-1)^{j-1}P_j(z)\frac{\log^{j-1}(z)}{(j-1)!}=\mathrm{O}\big((1-z)^{a(n+1)-L\rho-M\sigma-1}\big) . \end{cases} \tag{5.23}$$

Les entiers a et n étant fixés, on note $S_\ell(z)=S_\ell\binom{L,M}{\rho,\sigma}(z)$, etc.

Théorème 4.1 *À une constante multiplicative près, le problème de Padé* (5.23) *a une unique solution, et elle vérifie pour tout* $\ell=0,\dots,L-1$ *:*

$$S_\ell\binom{L,M}{\rho,\sigma}(z)=\frac{(-1)^\ell}{\ell!}\sum_{k=1}^{\infty}\frac{\mathrm{d}^\ell}{\mathrm{d}k^\ell}\left(\frac{(k-\rho)_\rho^L(k+n+1)_\sigma^M}{(k)_{n+1}^a}\right)z^{-k}$$

et

$$P_a\binom{L,M}{\rho,\sigma}(z)=\sum_{k=0}^{n}(-1)^{ka}\frac{(-k-\rho)_\rho^L(n-k+1)_\sigma^M}{k!^a(n-k)!^a}z^k .$$

De plus, pour tout $j = 1, \dots, a$, on a

$$z^n P_j\begin{pmatrix} L, M \\ \rho, \sigma \end{pmatrix}(1/z) = (-1)^{a(n+1)+L\rho+M\sigma+j} P_j\begin{pmatrix} M, L \\ \sigma, \rho \end{pmatrix}(z)$$
$$\text{et} \quad z^n \overline{S}_\ell\begin{pmatrix} L, M \\ \rho, \sigma \end{pmatrix}(1/z) = (-1)^{a(n+1)+L\rho+M\sigma} S_\ell\begin{pmatrix} M, L \\ \sigma, \rho \end{pmatrix}(z) .$$

On peut également généraliser le Théorème 3.1. On garde les mêmes notations que précédemment, et on suppose $M = L$, $\sigma = \rho$ et $2L\rho \leq a(n+1) - 2$. On cherche des polynômes $P_{0,\ell}$, $\overline{P}_{0,\ell}$, P_j (pour $1 \leq j \leq a$ et $0 \leq \ell \leq L-1$) de degré au plus n et des fonctions S_ℓ, $\overline{S}_\ell$ et R (qui dépendent aussi de ρ, L, a et n) tels que l'on ait simultanément les $2L+2$ conditions suivantes (pour $\ell = 0, \dots, L-1$) :

$$\begin{cases} S_\ell(z) := P_{0,\ell}(z) + \displaystyle\sum_{j=1}^{a} \binom{\ell+j-1}{j-1} P_j(z) \mathrm{Li}_{\ell+j}(1/z) = \mathrm{O}(z^{-\rho-1}) \\ \overline{S}_\ell(z) := \overline{P}_{0,\ell}(z) + \displaystyle\sum_{j=1}^{a} (-1)^j \binom{\ell+j-1}{j-1} P_j(z) \mathrm{Li}_{\ell+j}(z) = \mathrm{O}(z^{n+\sigma+1}) \\ R(z) := \displaystyle\sum_{j=1}^{a} (-1)^{j-1} P_j(z) \frac{\log^{j-1}(z)}{(j-1)!} = \mathrm{O}\big((1-z)^{a(n+1)-2L\rho-2}\big) \\ P_a((-1)^a) = 0 . \end{cases} \tag{5.24}$$

Théorème 4.2 *À une constante multiplicative près, le problème de Padé* (5.24) *a une unique solution, et elle vérifie pour tout $\ell = 0, \dots, L-1$:*

$$S_\ell(z) = \frac{(-1)^\ell}{\ell!} \sum_{k=1}^{\infty} \frac{\mathrm{d}^\ell}{\mathrm{d}k^\ell} \left(\left(k + \frac{n}{2}\right) \frac{(k-\rho)_\rho^L (k+n+1)_\rho^L}{(k)_{n+1}^a} \right) z^{-k}$$

et

$$P_a(z) = \sum_{k=0}^{n} (-1)^{ka} \left(\frac{n}{2} - k\right) \frac{(-k-\rho)_\rho^L (n-k+1)_\rho^L}{k!^a (n-k)!^a} z^k .$$

De plus, pour tous $j = 1, \dots, a$ et $\ell = 0, \dots, L-1$, on a

$$z^n P_j(1/z) = (-1)^{a(n+1)+j+1} P_j(z) \ \text{ et } \ z^n \overline{S}_\ell(1/z) = (-1)^{a(n+1)+1} S_\ell(z) .$$

Nous omettons les démonstrations des Théorèmes 4.1 et 4.2 car elles sont tout à fait similaires à celles des Théorèmes 1.1 et 3.1 : les conditions à l'infini, resp. en 0, se traduisent par l'existence d'une fraction rationnelle $A(k) = \dfrac{Q(k)}{(k)_{n+1}^a}$ telle que $(k-\rho)_\rho^L$, resp. $(k+n+1)_\sigma^M$, divise $Q(k)$, et la condition en 1 implique

- dans le cas du Théorème 4.1, que $\deg(Q) \leq L\rho + M\sigma$, ce qui suffit.
- dans le cas du Théorème 4.2, que $\deg(Q) \leq 2L\rho + 1$; on adapte alors la démonstration du Théorème 3.1 .

Remarque : Si dans le problème (5.24), on remplace la condition $P_a((-1)^a) = 0$ par $P_a((-1)^{a+1}) = 0$, on trouve que ce nouveau problème a également une solution "unique", qui est la même que celle du problème (5.23) (avec $M = L$ et $\sigma = \rho$). C'est un exemple où plusieurs problèmes de Padé admettent une série donnée comme solution unique.

Avec $L = 3$, $a = 20$, $\rho = n$ et $\ell = 2$, le Théorème 4.2 donne la série utilisée dans [R7] pour prouver l'irrationalité d'au moins un des neuf nombres $\zeta(5)$, $\zeta(7)$, ..., $\zeta(21)$:

$$n!^{14} \sum_{k=1}^{\infty} \frac{\mathrm{d}^2}{\mathrm{d}k^2} \left(\left(k + \frac{n}{2} \right) \frac{(k-n)_n^3 (k+n+1)_n^3}{(k)_{n+1}^{20}} \right)$$
$$= \alpha_{0,n} + \alpha_{5,n}\zeta(5) + \alpha_{7,n}\zeta(7) + \cdots + \alpha_{21,n}\zeta(21) ,$$

pour certains rationnels $\alpha_{j,n}$.

Remarque : On pourrait aussi considérer des problèmes de Padé "non diagonaux", c'est-à-dire dans lesquels on demande $\deg(P_j) \leq n_j$, sans avoir forcément $n_1 = \cdots = n_a = n$. Ceci permettrait peut-être d'englober les séries considérées par Zudilin [Z6] pour démontrer qu'il y a au moins un irrationnel parmi les nombres $\zeta(5)$, $\zeta(7)$, $\zeta(9)$ et $\zeta(11)$ (voir aussi [F4], §3.3).

5 Résolution du problème de Padé pour $\zeta(2)$

Nous donnons ici une solution alternative au problème de Padé (5.1) : déterminer pour tout entier $n \geq 0$ des polynômes a, b et c de degré au plus n et des fonctions S et R tels que

$$\begin{cases} S(z) = a(z)\mathrm{Li}_2(1/z) + b(z)\mathrm{Li}_1(1/z) + c(z) = \mathrm{O}(z^{-n-1}) \\ R(z) = a(z)\log(z) - b(z) = \mathrm{O}\big((1-z)^{n+1}\big) . \end{cases}$$

En adaptant une méthode utilisée par Sorokin [S7] (et qui se rapproche d'une méthode exposée par Siegel [S3] pour déterminer les approximants de Padé usuels de la fonction exponentielle), nous allons montrer la

Proposition 5.1 *À une constante multiplicative près, le problème* (5.1) *admet une solution unique qui vérifie :*

$$S(z) = \int_0^1 \int_0^1 \frac{x^n(1-x)^n y^n (1-y)^n}{(z-xy)^{n+1}} \mathrm{d}x\mathrm{d}y . \tag{5.25}$$

En développant l'intégrale comme série entière en $1/z$, on retrouve la série (5.4). Pour $z = 1$, on obtient l'intégrale introduite par Beukers dans [B4].

Nous aurons besoin des deux lemmes suivants, dont nous omettons les démonstrations car ils sont classiques (voir par exemple [A3] pour le Lemme 5.1 et [S2] (page 60) pour le Lemme 5.2).

Lemme 5.1 *Le problème de Padé consistant à déterminer des polynômes P et Q de degré au plus n tels que $V(z) = Q(z)\mathrm{Li}_1(1/z) + P(z) = \mathrm{O}(z^{-n-1})$ admet une solution unique (à une constante multiplicative près) et on a*

$$V(z) = \int_0^1 \frac{x^n(1-x)^n}{(z-x)^{n+1}} \mathrm{d}x .$$

Lemme 5.2 *Soit P_{n+1} l'opérateur intégral défini, pour toute fonction f analytique dans un voisinage de l'infini, par*

$$P_{n+1}(f)(z) = \frac{(-1)^{n+1}}{n!} \int_z^{\infty} (x-z)^n f(x) \mathrm{d}x \; .$$

Soit g une fonction analytique dans un voisinage de l'infini et telle que $\lim_{z\to\infty} g(z) = 0$. Alors

$$g^{(n+1)} = f \quad \text{équivaut à} \quad g = P_{n+1}(f) \; . \tag{5.26}$$

Démonstration de la Proposition 5.1. Il existe au moins une solution non triviale, qui donne des fonctions $S(z)$ et $R(z)$. Nous allons montrer que $S(z)$ est un multiple de l'intégrale (5.25). Remarquons tout d'abord que[4]

$$\mathrm{Li}_2(1/z) + \log(z)\mathrm{Li}_1(1/z) = \pi^2/6 - \mathrm{Li}_2(1-1/z) \; .$$

On en déduit que la fonction $T(z)$ définie par

$$T(z) = S(z) + \mathrm{Li}_1(1/z)R(z) = a(z)(\mathrm{Li}_2(1/z) + \log(z)\mathrm{Li}_1(1/z)) + c(z)$$

est analytique au voisinage de 1. De plus, la condition en 1 sur $R(z)$ implique que $(\mathrm{Li}_1(1/z)R(z))^{(n+1)}$ a, au plus, une singularité logarithmique en 1. Donc $S^{(n+1)}(z) = T^{(n+1)}(z) - (\mathrm{Li}_1(1/z)R(z))^{(n+1)}$ a, au plus, une singularité logarithmique en 1. Comme par ailleurs,

$$S^{(n+1)}(z) = \frac{\tilde{Q}(z)}{z^{n+1}(1-z)^{n+1}} \mathrm{Li}_1(1/z) + \frac{\tilde{P}(z)}{z^{n+1}(1-z)^{n+1}}$$

où $\tilde{P}$ et $\tilde{Q}$ sont des polynômes de degré au plus $2n+1$, on en déduit que $(1-z)^{n+1}$ divise $\tilde{P}$ et $\tilde{Q}$.

En résumé, on a montré l'existence de polynômes P et Q de degré au plus n tels que $z^{n+1}S^{(n+1)}(z) = Q(z)\mathrm{Li}_1(1/z) + P(z) = \mathrm{O}(z^{-n-1})$: le Lemme 5.1 assure qu'il existe une constante c telle que :

$$S^{(n+1)}(z) = (-1)^{n+1} c z^{-n-1} \int_0^1 \frac{x^n(1-x)^n}{(z-x)^{n+1}} \mathrm{d}x \; .$$

Il suffit alors d'appliquer le Lemme 5.2 à $g(z) = S(z)$ (qui tend vers 0 à l'infini) et $f(z) = S^{(n+1)}(z)$:

$$\begin{aligned}
S(z) &= P_{n+1}(S^{(n+1)}(z)) \\
&= \frac{c}{n!} \int_z^{\infty} \left(1 - \frac{z}{u}\right)^n \left(\int_0^1 \frac{x^n(1-x)^n}{(u-x)^{n+1}} \mathrm{d}x \right) \frac{\mathrm{d}u}{u} \\
&= \frac{c}{n!} \int_0^1 (1-y)^n \left(\int_0^1 \frac{x^n(1-x)^n}{(z/y-x)^{n+1}} \mathrm{d}x \right) \frac{\mathrm{d}y}{y} \\
&= \frac{c}{n!} \int_0^1 \int_0^1 \frac{x^n(1-x)^n y^n(1-y)^n}{(z-xy)^{n+1}} \mathrm{d}x \, \mathrm{d}y \; .
\end{aligned}$$

[4]Cela se vérifie en dérivant les deux membres, avec $z\frac{d}{dz}\mathrm{Li}_2(z) = \mathrm{Li}_1(z) = -\log(1-z)$.

6 Démonstration du Théorème 1.3

6.1 Structure de la preuve

Considérons les fonctions $S_n(z)$ et $R_n(z)$ de la variable réelle $z \geq 1$ définies par :

$$S_n(z) = n!^{a-r} \sum_{k=1}^{\infty} \frac{(k-rn)_{rn}}{(k)_{n+1}^a} z^{-k} \text{ et } R_n(z) = \frac{n!^{a-r}}{2i\pi} \int_{\mathcal{C}} \frac{(s-rn)_{rn}}{(s)_{n+1}^a} z^{-s} \mathrm{d}s\,,$$

où $\mathcal{C}$ est un lacet direct entourant les points $-n, -n+1, \cdots, 0$, et r un entier tel que $1 \leq r \leq a$. On déduit du Théorème 1.1 l'existence de polynômes $P_{j,n}$ de degré au plus n tels que :

$$\begin{cases} S_n(z) = P_{0,n}(z) + \displaystyle\sum_{j=1}^{a} P_{j,n}(z)\mathrm{Li}_j(1/z) = \mathrm{O}(z^{-rn-1}) \\ R_n(z) = \displaystyle\sum_{j=1}^{a} P_{j,n}(z) \frac{\log^{j-1}(1/z)}{(j-1)!} = \mathrm{O}\big((1-z)^{(a-r)n+a-1}\big)\,. \end{cases}$$

Soient λ et μ deux réels et z un rationnel tels que $(\lambda, \mu) \neq (0,0)$ et $z > 1$. La démontration du Théorème 1.3 consiste à appliquer le critère d'indépendance linéaire ci-dessous à l'expression

$$\lambda S_n(z) + \mu \log(1/z) R_n(z) = \lambda P_{0,n}(z) + \sum_{j=1}^{a} P_{j,n}(z) \left(\lambda \mathrm{Li}_j(1/z) + \mu \frac{\log^j(1/z)}{(j-1)!} \right).$$

Proposition 6.1 (Critère de Nesterenko) *Soit $\vartheta_1, \vartheta_2, \ldots, \vartheta_N$ des nombres réels. Supposons qu'il existe N suites d'entiers $(p_{l,n})_{n \geq 0}$ et des réels $\alpha, \beta > 0$ tels que :*

i) $\left| \sum_{l=1}^{N} p_{l,n} \vartheta_l \right| = \alpha^{n+o(n)}$;

ii) $\forall l = 1, \ldots, N$, $|p_{l,n}| \leq \beta^{n+o(n)}$.

Dans ces conditions, on a

$$\dim_{\mathbb{Q}}(\mathbb{Q}\,\vartheta_1 + \mathbb{Q}\,\vartheta_2 + \cdots + \mathbb{Q}\,\vartheta_N) \geq \frac{\log(\beta) - \log(\alpha)}{\log(\beta)}\,.$$

Nous devrons donc déterminer un dénominateur commun aux $P_{j,n}(z)$, ainsi que le comportement asymptotique de $|\lambda S_n(z) + \mu \log(1/z) R_n(z)|^{1/n}$ et de $|P_{j,n}(z)|^{1/n}$. Pour cela, nous estimerons le comportement asymptotique de $|R_n(z)|^{1/n}$ en utilisant la Proposition 2.1.
On note d_n le p.p.c.m. des entiers 1, 2, …, n. Le théorème des nombres premiers implique que

$$d_n = e^{n+o(n)}\,.$$

Nous ne donnons pas les démonstrations des Lemmes 6.1 et 6.2 ci-dessous. En effet, la série $S_n(z)$ coïncide avec la série $N_{n,a,r}(z)$ mentionnée dans l'introduction et ces lemmes sont démontrés au chapitre 2 de [R6].

6.2 Estimations arithmétiques et asymptotiques

Lemme 6.1 *Pour tout entier $j \in \{0, 1, \ldots, a\}$, on a $d_n^{a-j} P_{j,n}(z) \in \mathbb{Z}[z]$. De plus, les coefficients $p_{j,i,n}$ des polynômes $P_{j,n}(z)$ vérifient, pour tout $j \in \{0, \ldots, a\}$ et tout $i \in \{0, \ldots, n\}$:*

$$\limsup_{n \to +\infty} |p_{j,i,n}|^{1/n} \leq 2^{a+r+1} r^r\,.$$

En particulier, pour tout réel $z > 1$ et pour tout j on a :

$$\limsup_{n\to+\infty} |P_{j,n}(z)|^{1/n} \leq 2^{a+r+1} r^r z \,.$$

Lemme 6.2 *Pour tout réel $z > 1$, on a*

$$S_n(z) = \frac{(rn)!}{n!^r} \int_{[0,1]^a} \left(\frac{\prod_{j=1}^a x_j^r (1-x_j)}{(z - x_1 x_2 \cdots x_a)^r} \right)^n \frac{\mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_a}{z - x_1 x_2 \cdots x_a} \,. \tag{5.27}$$

On en déduit que pour tout réel $z > 1$, la limite $\varphi_{a,r,z}$ de la suite $|S_n(z)|^{1/n}$ existe et vérifie

$$\frac{e^{-a}}{z^r (r+1)^a} \leq \varphi_{a,r,z} \leq \frac{1}{z^r r^{a-r}} \,. \tag{5.28}$$

Remarque. Montrons la minoration de $\varphi_{a,r,z}$, puisque le reste est démontré dans [R6]. La formule de Stirling donne $\lim_{n\to\infty}(\frac{(rn)!}{n!^r})^{1/n} = r^r$ et $\varphi_{a,r,z}$ est égal au maximum sur $[0,1]^a$ de la fonction $\underline{x} \mapsto r^r \left| \frac{\prod_{j=1}^a x_j^r (1-x_j)}{(z - x_1 x_2 \cdots x_a)^r} \right|$. Comme $z > 1$, on peut majorer son dénominateur par z^r. En considérant le point $x_1 = \cdots = x_a = \frac{r}{r+1}$, on obtient

$$\varphi_{a,r,z} \geq \frac{r^r}{z^r} \left(\frac{r^r}{(r+1)^{r+1}} \right)^a \geq \frac{e^{-a}}{z^r (r+1)^a} \,,$$

en remarquant que $(r/(r+1))^r \geq e^{-1}$.

Lemme 6.3 *Pour tout réel $z > 1$, si $a > r \geq 1$, alors*

$$\psi_{a,r,z} = \limsup_{n\to+\infty} |R_n(z)|^{1/n} \leq \frac{e^{2a+1} r^r}{(a-r)^{a-r}} \cdot z \log^{a-r}(z) \,. \tag{5.29}$$

Démonstration. Les diverses suites $c(n)$ qui apparaissent ci-dessous sont explicitables ; elles dépendent éventuellement de a et de r, mais pas de z, et vérifient $\lim_{n\to+\infty} c(n)^{1/n} = 1$. D'après la Proposition 2.1, dont on reprend les notations, on a

$$R_n(z) = \sum_{d=(a-r)n+a}^{\infty} (-1)^{d-1} \mathfrak{A}_d \frac{\log^{d-1}(z)}{(d-1)!}$$

avec $\mathfrak{A}_d = \displaystyle\sum_{i=0}^{n} \sum_{j=1}^{a} (-1)^{j+d} \binom{d-1}{d-j} i^{d-j} p_{j,i,n}$. Comme $|p_{j,i,n}| \leq c_1(n) 2^{(a+r+1)n} r^{rn}$ d'après le Lemme 6.1, on en déduit pour $d \geq 2a-1$:

$$\begin{aligned} |\mathfrak{A}_d| &\leq c_1(n) 2^{(a+r+1)n} r^{rn} \sum_{i=0}^{n} \sum_{j=1}^{a} \binom{d-1}{d-j} i^{d-j} \\ &\leq c_2(n) 2^{(a+r+1)n} r^{rn} \binom{d-1}{d-a} n^{d-1} \,. \end{aligned}$$

Donc

$$\begin{aligned} |R_n(z)| &\leq c_2(n) 2^{(a+r+1)n} r^{rn} \sum_{d=(a-r)n+a-1}^{\infty} \frac{\binom{d}{d-a+1}}{d!} n^d \log^d(z) \\ &\leq c_3(n) 2^{(a+r+1)n} r^{rn} \log(z)^{a-1} \sum_{d=(a-r)n}^{\infty} \frac{(n \log(z))^d}{d!} \,. \end{aligned}$$

En appliquant l'identité $\displaystyle\sum_{d=\ell+1}^{\infty}\frac{t^d}{d!}=\frac{1}{\ell!}\int_0^t(t-x)^\ell e^x\mathrm{d}x$, on obtient

$$\sum_{d=(a-r)n}^{\infty}\frac{(n\log(z))^d}{d!}\leq z^n\cdot\frac{(n\log(z))^{(a-r)n}}{((a-r)n-1)!}\ .$$

Par la formule de Stirling, on en déduit que

$$|R_n(z)|\leq c_4(n)2^{(a+r+1)n}r^{rn}\cdot\frac{e^{(a-r)n}}{(a-r)^{(a-r)n}}\cdot z^n\log(z)^{(a-r)n+a-1}\ .$$

Finalement

$$\limsup_{n\to\infty}|R_n(z)|^{1/n}\quad\leq\quad\frac{e^{2a+1}r^r}{(a-r)^{a-r}}\cdot z\log^{a-r}(z)\ .$$

6.3 Fin de la démonstration du Théorème 1.3

Fixons les réels λ et μ et le rationnel $z=q/p>1$. Si $\lambda=0$, le Théorème 1.3 est démontré (essentiellement par cette méthode, avec $r=0$) dans [R2]. On peut donc supposer que $\lambda=1$. Définissons les entiers $\pi_{0,n}(z)=d_n^ap^nP_{0,n}(z)$ et $\pi_{j,n}(z)=d_n^ap^nP_{j,n}(z)$ pour $j\in\{1,\dots,a\}$, ainsi que la combinaison linéaire à coefficients entiers

$$\begin{aligned}\ell_n(z)&=d_n^ap^n\big(S_n(z)+\mu\log(1/z)R_n(z)\big)\\&=\pi_{0,n}(z)+\sum_{j=1}^{a}\pi_{j,n}(z)\big(\mathrm{Li}_j(1/z)+\mu\frac{\log^j(1/z)}{(j-1)!}\big)\ .\end{aligned}$$

Notons $\delta_{z,\mu}(a)$ la dimension de l'espace vectoriel engendré sur $\mathbb{Q}$ par 1 et les $\mathrm{Li}_j(1/z)+\mu\frac{\log^j(1/z)}{(j-1)!}$ pour $j\in\{1,\dots,a\}$. Le Théorème 1.3 découle immédiatement de la proposition plus précise suivante.

Proposition 6.2 *Si $a>r=[a/\log^2(a)]\gg_z 0$, la limite de $|\ell_n(z)|^{1/n}$ existe et vérifie*

$$\lim_{n\to+\infty}|\ell_n(z)|^{1/n}=e^a\varphi_{a,r,z}p\ .\tag{5.30}$$

Pour tout $j\in\{0,1,\dots,a\}$, on a

$$\limsup_{n\to+\infty}|\pi_{j,n}(z)|^{1/n}\leq e^a2^{a+r+1}r^rq\ .\tag{5.31}$$

Enfin, pour tout $\varepsilon>0$ il existe un entier $a_0(\varepsilon,z)$, qui dépend de ε et de z, tel que pour tout $a\geq a_0(\varepsilon,z)$ on ait :

$$\delta_{z,\mu}(a)\geq\frac{1-\varepsilon}{1+\log(2)}\log(a)\ .$$

Démonstration de la Proposition 6.2. Montrons tout d'abord que si $a > r = [a/\log^2(a)] \gg_z 0$, alors
$$\psi_{a,r,z} = \limsup_{n\to+\infty} |R_n(z)|^{1/n} < \lim_{n\to+\infty} |S_n(z)|^{1/n} = \varphi_{a,r,z} .$$
On vérifie en effet que, avec $r = [a/\log^2(a)]$:
$$\frac{\varphi_{a,r,z}}{\psi_{a,r,z}} \geq \frac{1}{e^{3a+1} z^{r+1} \log(z)^{a-r}} \cdot \frac{(a-r)^{a-r}}{r^r (r+1)^a} \to +\infty \quad \text{quand} \quad a \to +\infty .$$
On a donc
$$\lim_{n\to+\infty} |\ell_n(z)|^{1/n} = \lim_{n\to+\infty} |d_n^a p^n S_n(z)|^{1/n} ,$$
ce qui prouve (5.30). L'assertion (5.31) se déduit immédiatement du Lemme 6.1. On peut appliquer le critère de Nesterenko avec $\alpha = e^a \varphi_{a,r,z} p \leq e^a z^{-r} r^{r-a} p$ et $\beta = e^a 2^{a+r+1} r^r q$: il vient alors
$$\delta_{z,\mu}(a) \geq \frac{(a+r+1)\log(2) + (r+1)\log(z) + a\log(r)}{a + (a+r+1)\log(2) + \log(q) + r\log(r)} . \tag{5.32}$$
Un développement limité du membre de droite de (5.32) lorsque $a \to +\infty$, avec $r = [a/\log^2(a)]$, conclut la démonstration de la Proposition 6.2.

Remarque : lorsque a et r sont fixés, on montre que $\lim_{z\to+\infty} \varphi_{a,r,z} = 0$ alors que $\lim_{z\to+\infty} \psi_{a,r,z} = +\infty$. Il semble donc difficile de montrer avec nos méthodes que les nombres $\lambda \mathrm{Li}_a(1/z) + \mu \frac{\log^a(1/z)}{(a-1)!}$ sont irrationnels pour $z \gg_a 0$ quand λ et μ sont non nuls, bien que cela soit vrai quand $\lambda = 0$ ou $\mu = 0$. C'est pourquoi nous avons adopté le point de vue inverse en fixant z et en faisant varier a et r.

7 Démonstration du Théorème 1.4

On utilise maintenant les deux fonctions
$$S_n(z) = \frac{(2n)!^4}{n!} \sum_{k=1}^{\infty} \frac{(k-n)_n}{(k)_{2n+1}^4} z^{-k} \text{ et } R_n(z) = \frac{(2n)!^4}{n!\, 2i\pi} \int_{\mathcal{C}} \frac{(s-n)_n}{(s)_{2n+1}^4} z^{-s} \mathrm{d}s ,$$
où $\mathcal{C}$ est un lacet direct entourant les points $-2n, -2n+1, \cdots, 0$. Le Théorème 1.1 (en changeant n en $2n$) montre qu'il existe des polynômes $P_{j,n}$ de degré au plus $2n$ tels que :
$$\begin{cases} S_n(z) = P_{0,n}(z) + \displaystyle\sum_{j=1}^{4} P_{j,n}(z) \mathrm{Li}_j(1/z) = \mathrm{O}(z^{-n-1}) \\ R_n(z) = \displaystyle\sum_{j=1}^{4} P_{j,n}(z) \frac{\log^{j-1}(1/z)}{(j-1)!} = \mathrm{O}\big((1-z)^{7n+3}\big) . \end{cases}$$
Posons $\ell_n = S_n(2) - \log(2) R_n(2)$. Puisque $\mathrm{Li}_1(1/2) = \log(2)$, on a :
$$\ell_n = P_{0,n}(2) + P_{2,n}(2)\lambda_2 + P_{3,n}(2)\lambda_3 + P_{4,n}(2)\lambda_4$$
avec $\lambda_j = \mathrm{Li}_j(1/2) + \frac{\log^j(1/2)}{(j-1)!}$: cette combinaison linéaire ne fait pas intervenir $\log(2)$ puisque $\lambda_1 = 0$. Pour conclure, il suffit donc de déterminer un dénominateur commun D_n aux rationnels $P_{j,n}(2)$ et de montrer qu'on a $0 < \liminf_{n\to+\infty} |D_n \ell_n|^{1/n} < 1$.

Proposition 7.1 *i) Pour tout $j \in \{0, \ldots, 4\}$, on a $d_{2n}^4 \, P_{j,n}(z) \in \mathbb{Z}[z]$.*
ii) On a $\lim\limits_{n\to+\infty} |S_n(2)|^{1/n} \approx e^{-8,325073064}$.
iii) On a $\lim\limits_{n\to+\infty} |R_n(2)|^{1/n} \approx e^{-9,799017432}$.

On déduit de cette proposition que d_{2n}^4 est le dénominateur D_n recherché et que

$$\lim_{n\to+\infty} |d_{2n}^4 \, \ell_n|^{1/n} = \lim_{n\to+\infty} |d_{2n}^4 \, S_n(2)|^{1/n} \approx e^{-0,325073064} \in]0,1[\,,$$

ce qui prouve le Théorème 1.4.

Démonstration de la Proposition 7.1. i) On écrit

$$\frac{(2n)!^4}{n!} \cdot \frac{(k-n)_n}{(k)_{2n+1}^4} = \left(\frac{(2n)!(k-n)_n}{n!(k)_{2n+1}} \right) \cdot \left(\frac{(2n)!}{(k)_{2n+1}} \right)^3$$

et on adapte ensuite la technique utilisée au Lemme 5 de [B2], par exemple.
ii) On pourrait utiliser la représentation intégrale de $S_n(z)$ mais nous donnons ici une démonstration alternative basée sur la seconde démonstration du Lemme 3 de [B2]. En appliquant la formule de Stirling à $\dfrac{(2n)!^4}{n!} \cdot \dfrac{(k-n)_n}{2^k (k)_{2n+1}^4}$, avec $k = xn$, $x \geq 1$, on montre que

$$\begin{aligned} \lim_{n\to+\infty} |S_n(2)|^{1/n} &= 2^8 \max_{x\in[1,+\infty[} \left(\frac{2^{-x} x^{5x}}{(x-1)^{x-1} (x+2)^{4(x+2)}} \right) \\ &= 2^8 \frac{\alpha - 1}{(\alpha+2)^8} \approx e^{-8,325073064} \,, \end{aligned}$$

où $\alpha \approx 1,006316912$ est l'unique solution > 1 de $s^5 - 2(s-1)(s+2)^4 = 0$.
iii) Par le changement de variables $s \mapsto -s$, on voit que

$$R_n(2) = (-1)^n \frac{(2n)!^4}{n! \, 2i\pi} \int_{\mathcal{C}} \frac{(s+1)_n}{(s-2n)_{2n+1}^4} 2^s \mathrm{d}s$$

où $\mathcal{C}$ est un lacet direct entourant les points $0, 1, \ldots, 2n$. En suivant [R2], on peut déformer ce contour de telle sorte que pour tout réel $c > 2$, on ait

$$\begin{aligned} R_n(2) &= (-1)^n \frac{(2n)!^4}{n! \, 2i\pi} \int_{cn-i\infty}^{cn+i\infty} \frac{(s+1)_n}{(s-2n)_{2n+1}^4} 2^s \mathrm{d}s \\ &= (-1)^n n \frac{(2n)!^4}{n! \, 2i\pi} \int_{c-i\infty}^{c+i\infty} \frac{\Gamma\big((s+1)n+1\big) \Gamma\big((s-2)n\big)^4}{\Gamma(sn+1)^5} 2^{ns} \mathrm{d}s \,. \end{aligned}$$

La formule de Stirling s'étend aux nombres complexes :

$$\Gamma(z) = \big(z/e\big)^z \sqrt{2\pi/z} \, \big(1 + \mathrm{O}(1/|z|)\big)$$

lorsque $|\arg(z)| < \pi$ et on en déduit que

$$R_n(2) = 2^8 \, \alpha(n) \int_{c-i\infty}^{c+i\infty} g(s) e^{nw(s)} \mathrm{d}s \cdot \big(1 + \mathrm{O}(1/n)\big)$$

avec $c > 2$ quelconque, $g(s) = \sqrt{\frac{s+1}{s(s-2)^4}}$, $w(s) = (s+1)\log(s+1) + 4(s-2)\log(s-2) + s\log(2) - 5s\log(s)$ et $\alpha(n)$ telle que $\lim_{n\to+\infty} |\alpha(n)|^{1/n} = 1$. Nous sommes maintenant en position d'appliquer la méthode du col, dont nous rappelons le principe ci-dessous. Comme $w'(s) = \log(s+1) + 4\log(s-2) + \log(2) - 5\log(s)$, les solutions de l'équation $w'(s) = 0$ sont parmi les solutions de $s^5 - 2(s+1)(s-2)^4 = 0$. Cette dernière équation admet une unique solution réelle $\beta > 2$, qui vérifie *de facto* $w'(\beta) = 0$. On a $\beta \approx 11,31757856$ et le réel $w''(\beta)$ est non nul. L'étude des variations de la fonction $y \to \mathrm{Re}(w)(\beta + iy)$ (à la manière de [R7], page 165) montre que la droite $\mathrm{Re}(s) = c$ avec $c = \beta$ est admissible. On peut donc appliquer la Proposition 7.2 ci-dessous et en déduire que

$$\lim_{n\to+\infty} |R_n(2)|^{1/n} = 2^8 \frac{\beta+1}{(\beta-2)^8} \approx e^{-9,799017432} .$$

Pour conclure ce paragraphe, nous donnons maintenant un des énoncés possibles de la méthode du col : on se réfèrera à [C9], pp. 91-94 ou [D2], pp. 279-285 pour plus de détails. Soit w une fonction analytique au voisinage d'un point z_0 et telle que $w'(z_0) = 0$ et $w''(z_0) = |w''(z_0)|e^{i\alpha_0} \neq 0$. Au voisinage de z_0, on a $w(z) = w(z_0) + \frac{1}{2}w''(z_0)(z-z_0)^2 + O((z-z_0)^3)$. Soit L un chemin passant par z_0, et admettant une tangente Δ en ce point. Notons θ l'argument de $z - z_0$, défini modulo π, pour $z \in \Delta$. Supposons que $\cos(\alpha_0 + 2\theta) < 0$. Alors $\mathrm{Re}(\frac{1}{2}w''(z_0)(z - z_0)^2) < 0$ quand z parcourt L, au voisinage de z_0. La fonction $\mathrm{Re}(w(z))$ admet donc un maximum local en z_0 le long de L. On dit qu'un chemin L est admissible en z_0 si les conditions précédentes sont remplies, et que $\mathrm{Re}(w(z_0))$ est le maximum global de $\mathrm{Re}(w(z))$ le long de L.

Proposition 7.2 *Soient g et w deux fonctions analytiques dans un ouvert simplement connexe $\mathcal{D}$ du plan. Supposons qu'il existe $z_0 \in \mathcal{D}$ tel que $w'(z_0) = 0$ et $w''(z_0) = |w''(z_0)|e^{i\alpha_0} \neq 0$. Si L est un chemin inclus dans $\mathcal{D}$ et admissible en z_0, alors*

$$\int_L g(z)e^{nw(z)}\mathrm{d}z = \pm i\, g(z_0)\sqrt{\frac{2\pi}{n|w''(z_0)|}}\, e^{-i\alpha_0/2}e^{nw(z_0)} \cdot \big(1 + \mathrm{O}(1/n)\big) ,$$

où le signe $\pm$ dépend de l'orientation de L. De plus, cette estimation est encore valable si L est un chemin que l'on peut déformer en un chemin admissible en z_0.

8 Représentation intégrale réelle de $R_n(z)$

Lorsqu'on prend $\rho = \sigma = 0$ dans le Théorème 1.1, on a pour tous réels c et z tels que $c > n$ et $z > 0$:

$$R_n(z) = \frac{(-1)^{a(n+1)}}{2i\pi} \int_{c-i\infty}^{c+i\infty} \frac{n!^a}{(s-n)_n^a} z^s \mathrm{d}s .$$

Aux notations près, cette représentation complexe apparaît dans [R2] où Reyssat montre que pour tout réel $w \geq 0$, $R_n(e^w)$ se représente comme la puissance de convolution $R_n(e^w) = q_n^{*a}(w)$ avec $q_n(w) = (e^w - 1)^n$. Dans [A5], Amoroso remarque qu'au moyen de changements de variables simples on peut transformer ce produit de convolution en l'intégrale suivante :

$$R_n(1-z) = z^{a(n+1)-1} \int_{[0,1]^{a-1}} \prod_{j=1}^{a-1} \frac{x_j^{j(n+1)-1}(1-x_j)^n}{(1 - zx_j \cdots x_{a-1})^{n+1}} \mathrm{d}x_j .$$

En particulier, on en déduit que pour tout $z < 1$,

$$\lim_{n\to+\infty} |R_n(1-z)|^{1/n} = |z|^a \max_{\underline{x}\in[0,1]^{a-1}} \left(\prod_{j=1}^{a-1} \frac{x_j^j(1-x_j)}{1-zx_j\cdots x_{a-1}} \right) \leq \frac{|z|^a}{a^a}\,.$$

Il serait intéressant de généraliser cette représentation intégrale au cas où ρ et σ sont quelconques.

Remerciements : Nous tenons à remercier F. Amoroso, D. Bertrand, P. Grinspan, S. Khemira et M. Waldschmidt pour leurs remarques qui ont permis d'améliorer ce texte.

Deuxième partie

Transcendance dans les groupes algébriques

Acta Arithmetica 100.2 (2001), 167-187

Chapitre 6

Orbits under algebraic groups and logarithms of algebraic numbers

1 Introduction

[1]

Let K be either the field $\mathbb{C}$ of complex numbers, or its p-adic analog $\mathbb{C}_p$, which is the completion of an algebraic closure of $\mathbb{Q}_p$. Then K is an algebraically closed field of characteristic zero. We fix an algebraic closure $\bar{\mathbb{Q}}$ of $\mathbb{Q}$ in K.

In the complex case, let $L = \{\lambda \in \mathbb{C};\ \exp(\lambda) \in \bar{\mathbb{Q}}^*\}$ be the $\mathbb{Q}$-vector space of logarithms of algebraic numbers. In the p-adic case, we define $\log(1+x)$ by its power series expansion if $x \in \mathbb{C}_p$ has positive valuation, and we extend the logarithm function into a morphism of $\mathbb{C}_p^*$ to $\mathbb{C}_p$ such that $\log(p) = 0$. Then we denote by L the set of logarithms of nonzero elements of $\bar{\mathbb{Q}} \subset \mathbb{C}_p$; this set is a $\mathbb{Q}$-vector subspace of $\mathbb{C}_p$.

In the sequel, we consider simultaneously (unless otherwise specified) both the complex and the p-adic cases. We denote by $\mathcal{L}$ the $\bar{\mathbb{Q}}$-vector subspace of K spanned by 1 and L. The algebraic independence Conjecture (see [L1], Historical Notes of Chapter III) asserts that elements $\lambda_1, \dots, \lambda_q$ of $\mathcal{L}$ are algebraically independent over $\bar{\mathbb{Q}}$ if, and only if, $1, \lambda_1, \dots, \lambda_q$ are linearly independent over $\bar{\mathbb{Q}}$.

A new approach to the algebraic independence Conjecture has been suggested by Damien Roy. Instead of fixing elements $\lambda_1, \dots, \lambda_q$ of $\mathcal{L}$ and studying polynomial relations between them, he proceeds in the opposite direction : first fix a system of polynomial relations with coefficients in $\bar{\mathbb{Q}}$, in q variables, then focus on the q-tuples of elements of $\mathcal{L}$ which satisfy these relations. In geometric terms, this means that we are given a closed algebraic subset X of K^q, defined over $\bar{\mathbb{Q}}$ (that is, defined as the zero locus of a collection of polynomials with coefficients in $\bar{\mathbb{Q}}$), and we study the set $X(\mathcal{L}) = X \cap \mathcal{L}^q$ of points of X with coordinates in $\mathcal{L}$.

In this situation, Damien Roy states the following conjecture, which is equivalent [R13] to that of algebraic independence :

[1] 2000 Mathematics Subject Classification : Primary 11J85 ; Secondary 11J81, 33B10, 14R20, 14L35.

Conjecture 1 *Let q be a positive integer, and X be a closed algebraic subset of K^q, defined over $\bar{\mathbb{Q}}$.*

Then the following holds :

$$X(\mathcal{L}) = \bigcup_{E \subset X} E(\mathcal{L})$$

where E runs through the linear affine subspaces of K^q contained in X and defined over $\bar{\mathbb{Q}}$.

Conjecture 1 is a very precise description of $X(\mathcal{L})$. The following conjecture is less precise, therefore easier to prove for a given closed algebraic subset X ; nevertheless, it is equivalent to Conjecture 1, since both are equivalent to the algebraic independence Conjecture :

Conjecture 2 *Let q be a positive integer, and X be a closed algebraic subset of K^q, different from K^q, defined over $\bar{\mathbb{Q}}$.*

Let $x \in X(\mathcal{L}) = X \cap \mathcal{L}^q$.

Then x belongs to some affine hyperplane of K^q defined over $\bar{\mathbb{Q}}$.

These conjectures are very far from being proved ; actually, there are only a few closed algebraic subsets X for which anything is known about $X(\mathcal{L})$:

• Let X be a linear affine subspace of K^q, defined over $\bar{\mathbb{Q}}$. Then Conjectures 1 and 2 hold trivially for X.

• Let d, l and r be positive integers such that $r < \frac{dl}{d+l}$. Identify K^{dl} with the space of matrices with d rows and l columns, with entries in K. Let X be the subset of K^{dl} consisting in those matrices of rank at most r. Then Conjecture 2 holds for X : this is a straightforward consequence of Theorem 2.1 stated below, which is due to Damien Roy [R11].

• Let k and m be integers such that $2 \leq k \leq m-2$, and denote by $G(k,m)$ the affine cone over the Grassmannian which parametrizes the subspaces of dimension k of K^m. Assume $(k,m) \neq (2,4)$. Then Conjecture 2 holds for $G(k,m)$: this follows both from Theorem 2.1 of [R13] and from the results in the present paper (see Section 5.1). Further, if Conjecture 1 holds for $G(2,4)$ then it holds for any $G(k,m)$ ([R13], Proposition 2.5).

• Finally, if $K = \mathbb{C}$ and X is an algebraic curve, there are a few results ([R15], Theorem 0.2 and Corollary 7.2), but they concern only the points $x \in X(L)$.

In this paper, we prove new results about these sets $X(\mathcal{L})$. On the one hand, we prove Conjecture 1 for some algebraic subsets X, including the affine cone $\mathcal{V}(k,m) \subset \mathrm{Sym}^k(K^m)$ over the Veronese variety if $k \geq 3$ (see Section 6). This yields the following result :

Theorem 1.1 *Let k and m be integers, with $k \geq 3$.*

Let P be a homogeneous polynomial of degree k, in m variables, with coefficients in $\mathcal{L}$.

Assume that P is the k-th power of a linear form Λ with coefficients in K.

Then there exist a linear form ϕ, with coefficients in $\bar{\mathbb{Q}}$, and an element a of $\mathcal{L}$, such that $P = a\phi^k$.

On the other hand, we prove Conjecture 2 for some orbits under algebraic groups. In more precise terms, we consider an affine algebraic group G defined over $\bar{\mathbb{Q}}$ acting on a vector space W equipped with a $\bar{\mathbb{Q}}$-structure (see Section 2.1). We assume the representation $\varrho : G \to GL(W)$ to be a morphism of algebraic varieties defined over $\bar{\mathbb{Q}}$. Let X be an orbit under this action ; then X is a locally closed subset of W ([H10], Proposition 8.3). The following conjecture may be stated :

Conjecture 3 *Assume X is not of maximal dimension among ϱ-orbits. Then every $x \in X(\mathcal{L})$ belongs to some affine hyperplane of W defined over $\bar{\mathbb{Q}}$.*

Conjecture 3 follows from Conjecture 2, for the union of orbits of dimension less than or equal to $\dim(X)$ is a closed algebraic subset of W defined over $\bar{\mathbb{Q}}$.

In this paper, we prove Conjecture 3 for the orbits X which satisfy some additional assumptions (see Section 4). But the result we obtain is slightly more precise : for these orbits X, every $x \in X(\mathcal{L})$ belongs to some vector hyperplane of W defined over $\bar{\mathbb{Q}}$. The proof of the results stated in Section 4 involves a transcendence theorem due to Damien Roy [R11], which yields a lower bound for the rank of a matrix with entries in $\mathcal{L}$ by taking into account the linear relations, with coefficients in $\bar{\mathbb{Q}}$, between its entries. This result is stated in Section 2, together with definitions and notation. Section 3 is devoted to applying this transcendence theorem. In Section 5, we provide examples to which the previous results apply ; for instance, the following statement is proved :

Theorem 1.2 *Let n be an integer, and M be a square matrix of size n, with entries in $\mathcal{L}$. Assume the n^2 entries of M are linearly independent over $\bar{\mathbb{Q}}$. Then the centralizer of M has dimension less than or equal to $\frac{1}{2}(n^2+1)$.*

Finally, notice that all the results obtained in this paper concerning the points of $X(\mathcal{L})$ apply, in particular, to the points of $X(L)$. Let $x \in X(L)$; then x belongs to some affine hyperplane of K^q defined over $\bar{\mathbb{Q}}$ if, and only if, x belongs to some vector hyperplane of K^q defined over $\mathbb{Q}$: this follows from the Theorem of Baker-Brumer ([B1], Chapters 1 and 2). Consequently, it is possible, in every statement in this paper, to replace $(\bar{\mathbb{Q}}, \mathcal{L})$ by $(\mathbb{Q}, L)$.

2 Preliminaries

In the sequel, all the subspaces we deal with are vector subspaces (unless otherwise specified).

2.1 Some Elementary Facts about $\bar{\mathbb{Q}}$-Structures

This Section is devoted to $\bar{\mathbb{Q}}$-structures ; a detailed account of this can be found in [B14], §8.

Let q be an integer. An element of K^q is said to be *defined over* $\bar{\mathbb{Q}}$ if it belongs to $\bar{\mathbb{Q}}^q$. A basis $(e_1, \dots, e_q)$ of K^q is said to be *defined over* $\bar{\mathbb{Q}}$ if the vectors e_j are defined over $\bar{\mathbb{Q}}$. A K-subspace of K^q is said to be *defined over* $\bar{\mathbb{Q}}$ if it is spanned by vectors defined over $\bar{\mathbb{Q}}$. It is equivalent to being defined by linear equations with coefficients in $\bar{\mathbb{Q}}$. More generally, a closed algebraic subset of K^q is said to be *defined over* $\bar{\mathbb{Q}}$ if it is the zero locus of a collection of polynomials with coefficients in $\bar{\mathbb{Q}}$.

Let W be a vector space of dimension q over K. We call $\bar{\mathbb{Q}}$-*structure* on W any $\bar{\mathbb{Q}}$-subspace of W that is spanned, over $\bar{\mathbb{Q}}$, by a K-basis of W.

Let W be a vector space of dimension q over K, equipped with a $\bar{\mathbb{Q}}$-structure denoted by $W(\bar{\mathbb{Q}})$. There is a bijective K-linear map f, from W to K^q, that sends $W(\bar{\mathbb{Q}})$ onto $\bar{\mathbb{Q}}^q$. Thanks to f, it makes sense for a vector, basis, subspace or closed algebraic subset of W to be *defined over* $\bar{\mathbb{Q}}$; and this does not depend on the choice of f, but only on the $\bar{\mathbb{Q}}$-structure $W(\bar{\mathbb{Q}})$.

A linear map $f : W \to W'$, where W and W' are equipped with $\bar{\mathbb{Q}}$-structures, is said to be *defined over* $\bar{\mathbb{Q}}$ if $f(W(\bar{\mathbb{Q}})) \subset W'(\bar{\mathbb{Q}})$.

We denote by $W(\mathcal{L})$ the subset of W consisting of those vectors whose coordinates, in a basis of W defined over $\bar{\mathbb{Q}}$, belong to $\mathcal{L}$ (since $\mathcal{L}$ is a vector space over $\bar{\mathbb{Q}}$, the set $W(\mathcal{L})$ does not depend on the basis we choose). Moreover, if X is any subset of W, we let $X(\bar{\mathbb{Q}}) = X \cap W(\bar{\mathbb{Q}})$ and $X(\mathcal{L}) = X \cap W(\mathcal{L})$. If W is the vector space K^q, equipped with the $\bar{\mathbb{Q}}$-structure $\bar{\mathbb{Q}}^q$, this agrees with the notation $X(\mathcal{L}) = X \cap \mathcal{L}^q$ used in the introduction.

Let W be a vector space equipped with a $\bar{\mathbb{Q}}$-structure $W(\bar{\mathbb{Q}})$, and let k be an integer. Then the symmetric power $\mathrm{Sym}^k(W)$ is equipped with an induced $\bar{\mathbb{Q}}$-structure $\mathrm{Sym}^k(W(\bar{\mathbb{Q}}))$. In more concrete terms, if $(e_1, \dots, e_q)$ is a basis of W defined over $\bar{\mathbb{Q}}$, then the corresponding basis (e_I) of $\mathrm{Sym}^k(W)$ is defined over $\bar{\mathbb{Q}}$ (where $I = (i_1, \dots, i_q)$ runs through the q-tuples of integers such that $i_1 + \ldots + i_q = k$, and e_I is the symmetric product in which each e_j is repeated i_j times). An analogous argument applies to $W^{\otimes k}$ and $\Lambda^k(W)$.

2.2 Statement of the Transcendence Theorems

The transcendence theorem we will use in Section 3 is the following statement. It is a generalization of Theorem 2.1 of [W2] ; it results from Theorem 4 of [R11] (see the proof of Corollary 1 of the same theorem) :

Theorem 2.1 *Let E and F be nonzero vector spaces, equipped with $\bar{\mathbb{Q}}$-structures, of respective dimensions l and d.*

Let u be a linear map from E to F, of rank r.

We assume the following :

1. *The matrix of u, in bases of E and F defined over $\bar{\mathbb{Q}}$, has entries in $\mathcal{L}$.*
2. *The kernel of u contains no nonzero element of E defined over $\bar{\mathbb{Q}}$.*
3. *The image of u is contained in no hyperplane of F defined over $\bar{\mathbb{Q}}$.*
4. *We have : $r < \frac{dl}{d+l}$.*

Then there exist vector subspaces $S \subset E$ and $T \subset F$, of respective codimensions $l_1 \geq 1$ and $d_1 \geq 1$, defined over $\bar{\mathbb{Q}}$, with the following properties : $u(S) \subset T$ and there is an integer $r_1 \geq 1$ such that

$$\frac{r_1}{d_1} < \frac{r}{d} \text{ and } r_1 \geq \frac{l_1 d_1}{l_1 + d_1} \tag{6.1}$$

In this theorem, the assumptions 2 and 3 are not essential : if they are not met, we can consider the induced map $\bar{u} : E/E' \to F'$, where E' is the maximal subspace of E, defined over $\bar{\mathbb{Q}}$, that is contained in $\ker(u)$, and F' is the minimal subspace of F, defined over $\bar{\mathbb{Q}}$, that contains $\mathrm{Im}(u)$. This yields the following corollary, which is Corollary 1 of [R11] :

Corollary 2.2 *Let E and F be nonzero vector spaces, equipped with $\bar{\mathbb{Q}}$-structures, of respective dimensions l and d.*

Let u be a linear map from E to F, of rank r, whose matrix in bases of E and F defined over $\bar{\mathbb{Q}}$ has entries in $\mathcal{L}$.

Then there exist vector subspaces $S \subset E$ and $T \subset F$, of respective codimensions $l_1 \geq 0$ and $d_1 \geq 1$, defined over $\bar{\mathbb{Q}}$, such that $u(S) \subset T$ and $(d - r)l_1 \leq rd_1$.

In Section 6, we will use the following corollary of Theorem 2.1 ([R11], Corollary 2), which is a generalization of the six exponential theorem ([L1], Chapter II, Theorem 1) :

Theorem 2.3 *Let d and l be integers, with $d \geq 2$ and $l \geq 3$.*

Let M be a matrix, with d rows and l columns, of rank 1, with entries in $\mathcal{L}$.

Then either the rows or the columns of M are linearly dependent over $\bar{\mathbb{Q}}$.

2.3 Notation

In the sequel, we consider the following situation.

Let W be a finite dimensional vector space over K, equipped with a $\bar{\mathbb{Q}}$-structure.

Let G be an affine algebraic group defined over $\bar{\mathbb{Q}}$, that is an affine algebraic variety defined over $\bar{\mathbb{Q}}$, equipped with a group structure such that the map $(x, y) \mapsto xy^{-1}$ is a morphism of varieties defined over $\bar{\mathbb{Q}}$. Then $G(K)$ is a Lie group; the notation $G(K)$ underlines the fact that it is equipped with complex or p-adic topology.

Let $\varrho : G \to GL(W)$ be a linear representation of G, assumed to be at the same time a morphism defined over $\bar{\mathbb{Q}}$ between the algebraic varieties G and $GL(W)$.

Let X be an orbit under this action. Then X is a smooth locally closed subset of W ([H10], Proposition 8.3).

Notation: Let $d = \dim(W)$, $l = \dim(G)$ and $r = \dim(X)$.

We denote by G^{op} the opposite group to G, that is the group with underlying set that of G and law $*^{op}$ defined by $a *^{op} b = b * a$ for all $a, b \in G$ (where $*$ is the law of G). We denote by W^* the dual space of W, and by $\varrho^{op} : G^{op} \to GL(W^*)$ the contragredient representation associated with ϱ, defined by $\varrho^{op}(g) = {}^t(\varrho(g))$ for all $g \in G$.

The tangent map of ϱ at the unit element (denoted by Id) of G is the map $\mathrm{Lie}(\varrho) : \mathfrak{g} \to \mathfrak{gl}(W)$, where $\mathfrak{g}$ is the Lie algebra of G and $\mathfrak{gl}(W) = \mathrm{End}(W)$ that of $GL(W)$. Both Lie algebras are equipped with $\bar{\mathbb{Q}}$-structures, and $\mathrm{Lie}(\varrho)$ is defined over $\bar{\mathbb{Q}}$ ([H10], §34.2).

For $A \in \mathfrak{g}$ and $\alpha \in W$, we let :

$$(\mathrm{Lie}(\varrho)(A))(\alpha) = f_A(\alpha) = M_\alpha(A)$$

We define in this way linear maps (for $A \in \mathfrak{g}$ and $\alpha \in W$)

$$f_A : W \to W \text{ and } M_\alpha : \mathfrak{g} \to W$$

The following lemma ([B15], Chapter III, §1.7, Proposition 14) implies that the union of orbits whose dimension is less than some given integer is a closed algebraic subset of W defined over $\bar{\mathbb{Q}}$:

Lemma 2.4 *For $x \in X$, the image of M_x is the tangent space to X at x; accordingly, we have :*

$$\mathrm{rk}(M_x) = \dim(X)$$

This paper originates in the following remark, due to Damien Roy : if X is an orbit of dimension less than $\frac{dl}{d+l}$, then for $x \in X(\mathcal{L})$ the map M_x has rank less than $\frac{dl}{d+l}$, therefore Theorem 2.1 may apply to M_x. This idea is developed in a more precise way in the next Section.

3 Applying a Transcendence Theorem

In this Section, we state and prove Proposition 3.1, the only arithmetical step in the proof of the results mentioned in Sections 4 and 5. This Proposition follows from Theorem 2.1, the assumptions of which lead to the following definition :

Definition: The pair (ϱ, X) is said to be *suitable* if the following holds :

1. The map $\mathrm{Lie}(\varrho)$ is injective.
2. There is no pair (V, ϕ), consisting in an open subgroup V of the Lie group $G^{op}(K)$ and a nonzero element $\phi \in W^*$, such that ϕ is invariant under $\varrho^{op}(V)$.
3. We have $r = \dim(X) < \frac{dl}{d+l} = \frac{\dim(W)\dim(G)}{\dim(W)+\dim(G)}$.

For the sake of brevity, we shall sometimes say that X, rather than (ϱ, X), is suitable.

N.B. If ϱ is faithful then the first condition is met. On the other hand, the second condition is satisfied as soon as for every open subgroup V of $G^{op}(K)$ there exists $g \in V$ such that 1 is not an eigenvalue of $\varrho(g)$. A sufficient condition for this to hold is the existence of $A \in \mathfrak{g}$ such that $\mathrm{Lie}(\varrho)(A) = \mathrm{Id}$.

Now we can state, and prove, the following :

Proposition 3.1 *Assume (ϱ, X) is suitable.*

Let $x \in X(\mathcal{L})$ be an element that belongs to no hyperplane of W defined over $\bar{\mathbb{Q}}$.

Then there exist vector subspaces $S \subset \mathfrak{g}$ and $T \subset W$, of respective codimensions $l_1 \geq 1$ and $d_1 \geq 1$, with the following property : $\mathrm{Im}(f_A) \subset T$ for all $A \in S$, and there is an integer $r_1 \geq 1$ such that

$$\frac{r_1}{d_1} < \frac{r}{d} \text{ and } r_1 \geq \frac{l_1 d_1}{l_1 + d_1} \tag{6.2}$$

N.B. In particular, the conclusion yields $r_1 < r$ and the following relation :

$$l_1 < \frac{r}{d-r} d_1 \tag{6.3}$$

Proof of Proposition 3.1 : Let us check that Theorem 2.1 applies to M_x. First of all, the matrix of M_x in bases of $\mathfrak{g}$ and W defined over $\bar{\mathbb{Q}}$ has entries in $\mathcal{L}$ since $x \in W(\mathcal{L})$. Further, we have $\mathrm{rk}(M_x) = \dim(X) < \frac{dl}{d+l}$ by Lemma 2.4, and because (ϱ, X) is suitable.

Furthermore, let $A \in \ker(M_x)(\bar{\mathbb{Q}})$. Then x belongs to the kernel of f_A, which is a subspace of W defined over $\bar{\mathbb{Q}}$. Accordingly, this subspace is equal to W itself, that is $\mathrm{Lie}(\varrho)(A) = 0$; by assumption, this implies $A = 0$.

Finally, let H be a hyperplane defined over $\bar{\mathbb{Q}}$ which contains the image of M_x. Then $x \in f_A^{-1}(H)$ for all $A \in \mathfrak{g}$. This yields $\mathrm{Im}(f_A) \subset H$ for all $A \in \mathfrak{g}(\bar{\mathbb{Q}})$, hence for all $A \in \mathfrak{g}$ by linearity. In the complex case, this implies (thanks to [B15], Chapter III, §6.5, Proposition 13) $\varrho(g)(y) - y \in H$ for every $y \in W$ and every g in the neutral component of $G(\mathbb{C})$. In the p-adic case, we have $\mathrm{Im}(\exp(f_A) - \mathrm{Id}) \subset H$ whenever A is close enough to the origin. But there exist open subgroups (in the p-adic topology) $U \subset \mathfrak{g}(K)$ and $V \subset G(K)$, and a bijective exponential map (denoted by $\exp_U$) from U to V. Restricting U and V if necessary, we can assume $\mathrm{Im}(\exp(f_A) - \mathrm{Id}) \subset H$ for all $A \in U$, and $\varrho \circ \exp_U = \exp \circ \mathrm{Lie}(\varrho)$ on U. Then for every $g \in V$ there exists $A \in U$ such that $\exp_U(A) = g$, hence $\mathrm{Im}(\varrho(g) - \mathrm{Id}) = \mathrm{Im}(\exp(f_A) - \mathrm{Id}) \subset H$. Therefore, in the p-adic as well as in the complex case, there is an open subgroup V of $G(K)$ such that $\mathrm{Im}(\varrho(g) - \mathrm{Id}) \subset H$ for all $g \in V$. Let ϕ be a linear form whose kernel is H; the previous relation means $\phi \in \ker(\varrho^{op}(g) - \mathrm{Id}_{W^*})$ for all $g \in V$: this contradicts the assumption that (ϱ, X) is suitable. Consequently, the image of M_x is contained in no hyperplane of W defined over $\bar{\mathbb{Q}}$.

Thus Theorem 2.1 applies, and produces some subspaces S and T defined over $\bar{\mathbb{Q}}$ such that relations (6.1) hold and with the property $M_x(S) \subset T$. For $A \in S(\bar{\mathbb{Q}})$, this yields $f_A^{-1}(T) = W$, i.e. $\mathrm{Im}(f_A) \subset T$. As A runs through a basis of S defined over $\bar{\mathbb{Q}}$, this concludes the proof of Proposition 3.1.

Proposition 3.1 will allow us to prove Conjecture 3 for some orbits X. However, applying this statement prevents us from proving more precise results on the points $x \in X(\mathcal{L})$. This is the reason why the following proposition is useful :

Proposition 3.2 *Let $x \in X(\mathcal{L})$ and denote by $\mathcal{E}$ the minimal subspace of W, defined over $\bar{\mathbb{Q}}$, that contains x. Then there exist vector subspaces $S \subset \mathfrak{g}$ and $T \subset W$, of respective codimensions $l_1 \geq 0$ and $d_1 \geq 1$, such that $(d-r)l_1 \leq rd_1$ and $f_A(\alpha) \in T$ for every $A \in S$ and every $\alpha \in \mathcal{E}$.*

Proof of Proposition 3.2 : Corollary 2.2, applied to M_x, produces subspaces S and T, defined over $\bar{\mathbb{Q}}$, such that $M_x(S) \subset T$. For $A \in S(\bar{\mathbb{Q}})$, this implies $f_A(\mathcal{E}) \subset T$; as A runs through a basis of S defined over $\bar{\mathbb{Q}}$, this ends the proof of Proposition 3.2.

4 General Results

In this Section, we try to prove that every point $x \in X(\mathcal{L})$ belongs to some hyperplane of W defined over $\bar{\mathbb{Q}}$ (except in Theorem 4.2, where a stronger statement is obtained). With this aim in view, we let $x \in X(\mathcal{L})$ be a point that does not belong to any such hyperplane, and we proceed in the following way :

• We assume that (ϱ, X) is suitable, so that Proposition 3.1 applies and produces subspaces S and T.

• We make a geometric assumption on the orbits under ϱ (in Section 4.1) or under the contragredient representation ϱ^{op} (in Sections 4.3 and 4.4). This assumption allows us to derive a relation between $l_1 = \mathrm{codim}(S)$ and $d_1 = \mathrm{codim}(T)$ from the property $\mathrm{Im}(f_A) \subset T$ for all $A \in S$.

• We assume the dimension r of X to be "small enough", and sometimes we add a technical assumption, in order to derive a contradiction from the relations between l_1 and d_1.

These assumptions are of a different kind : the first one appears to be necessary to apply Proposition 3.1. The second one has an important drawback : the property $\mathrm{Im}(f_A) \subset T$ for all $A \in S$ is much stronger than the relation between l_1 and d_1 that we derive from it. Therefore it could be interesting to imagine other (geometric) assumptions than those made in this paper. Finally, the last assumption is fitted in such a way that it is possible to derive a contradiction.

4.1 A Large Dimensional ϱ-Orbit

Theorem 4.1 *Let k be an integer greater than 1. Assume :*

- *that (ϱ, X) is suitable.*
- *that there exists a ϱ-orbit of codimension less than k in W.*
- *that X has dimension $r \leq \frac{d}{k}$.*

Then every element $x \in X(\mathcal{L})$ belongs to some hyperplane of W defined over $\bar{\mathbb{Q}}$.

N.B. In particular, if ϱ has a Zariski-dense orbit then Theorem 4.1 applies with $k = 2$. But in this case, Theorem 4.2 stated below provides a more precise result (except if $r = \frac{d}{2}$).

Proof of Theorem 4.1 : Assume there exists $x \in X(\mathcal{L})$ that belongs to no hyperplane of W defined over $\bar{\mathbb{Q}}$. Denote by S and T the subspaces produced by Proposition 3.1, and let α be an element of W whose orbit has codimension $c \leq k-1$. Then $M_\alpha(S) \subset T$, hence $\dim(T) \geq \dim(S) - \dim(\ker(M_\alpha))$. As $\mathrm{rk}(M_\alpha) = d - c \geq d - k + 1$ by Lemma 2.4, we obtain : $d_1 \leq l_1 + k - 1$.

This inequality, together with assumption $r \leq \frac{d}{k}$ and inequality (6.3), yields : $d_1 - k + 1 \leq l_1 < \frac{r}{d-r} d_1 \leq \frac{d_1}{k-1}$, hence $\frac{k-2}{k-1} d_1 < k - 1$. But inequalities (6.2), with $r_1 \geq 1$ and $r \leq \frac{d}{k}$, imply $d_1 > \frac{d}{r} r_1 \geq k$, i.e. $d_1 \geq k+1$. Therefore the following holds : $\frac{(k-2)(k+1)}{k-1} < k - 1$. If $k \geq 3$, this yields a contradiction.

Assume now $k = 2$. Then $d_1 \leq l_1 + 1$, and $l_1 < d_1$ because of assumption $r \leq \frac{d}{2}$ and inequality (6.3) ; therefore $l_1 = d_1 - 1$. But inequalities (6.2) imply $d_1 > 2r_1$, that is $d_1 \geq 2r_1 + 1$. Hence $l_1 \geq 2r_1$ and $r_1 \geq \frac{(2r_1)(2r_1+1)}{4r_1+1}$, which is impossible because $r_1 \geq 1$.

In conclusion, Theorem 4.1 is proved for any k greater than or equal to 2.

Actually it is possible (under stronger assumptions) to prove the following more precise description of the points $x \in X(\mathcal{L})$:

Theorem 4.2 *Assume :*
- *that ϱ has a Zariski-dense orbit Y.*
- *that X has dimension $r < \frac{d}{2}$.*

Then for every $x \in X(\mathcal{L})$ there is a subspace $\mathcal{E}$ of W, defined over $\bar{\mathbb{Q}}$, that contains x and is disjoint from Y.

Proof of Theorem 4.2 : Let $\mathcal{E}$ be the minimal subspace of W, defined over $\bar{\mathbb{Q}}$, that contains x. Assume there is an element $\alpha \in \mathcal{E}$ that belongs to Y. Then Proposition 3.2 produces subspaces S and T such that $M_\alpha(S) \subset T$, hence $d_1 \leq l_1 + d - \mathrm{rk}(M_\alpha)$. Now Lemma 2.4 implies $\mathrm{rk}(M_\alpha) = d$, thereby proving $d_1 \leq l_1$. But this contradicts the relation $(d-r)l_1 \leq r d_1$, with $r < \frac{d}{2}$ and $d_1 \geq 1$. This ends the proof of Theorem 4.2.

4.2 Additional Notation

We denote by $\mathfrak{g}^{op}$ the Lie algebra of G^{op}. This is the Lie algebra with underlying vector space that of $\mathfrak{g}$ and bracket the opposite of the bracket of $\mathfrak{g}$.

In the same way as in Section 2.3, we use the following notation for $A \in \mathfrak{g}^{op}$ and $\phi \in W^*$: $(\mathrm{Lie}(\varrho^{op})(A))(\phi) = g_A(\phi) = N_\phi(A)$. For every $A \in \mathfrak{g}^{op}$ and every $\phi \in W^*$, this defines linear maps $g_A : W^* \to W^*$ and $N_\phi : \mathfrak{g}^{op} \to W^*$.

For all $A \in \mathfrak{g}$, we have ${}^t(\mathrm{Lie}(\varrho)(A)) = \mathrm{Lie}(\varrho^{op})(A)$, which reads ${}^t f_A = g_A$. A straightforward consequence of this relation is the following :

Lemma 4.3 *Let $S \subset \mathfrak{g}$ and $T \subset W$ be subspaces such that* $\mathrm{Im}(f_A) \subset T$ *for all $A \in S$.*

Let $T^* \subset W^*$ be the orthogonal subspace to T.

Then $g_A(\phi) = N_\phi(A) = 0$ for every $A \in S$ and every $\phi \in T^*$.

4.3 Finite Number of ϱ^{op}-Orbits

Theorem 4.4 *Assume :*

- *that* (ϱ, X) *is suitable.*
- *that the contragredient representation* ϱ^{op} *associated with* ϱ *has only finitely many orbits.*
- *that* X *has dimension* $r \leq \frac{d}{2}$.

Then every element $x \in X(\mathcal{L})$ *belongs to some hyperplane of* W *defined over* $\bar{\mathbb{Q}}$.

Proof of Theorem 4.4 : Assume, on the contrary, that some $x \in X(\mathcal{L})$ belongs to no hyperplane of W defined over $\bar{\mathbb{Q}}$. Then Proposition 3.1 applies, and produces subspaces S and T.

Let q be the maximal dimension of the ϱ^{op}-orbits that intersect T^*. Let Y_0 be an orbit realizing this maximum, and let $\phi \in T^* \cap Y_0$. We have $S \subset \ker(N_\phi)$ by Lemma 4.3, hence $l_1 \geq q$ thanks to Lemma 2.4 applied to the representation ϱ^{op} and the orbit Y_0 of ϕ. On the other hand, T^* is contained in the (finite) union of the ϱ^{op}-orbits of its elements, therefore $d_1 \leq q$. Let us compare the relation $l_1 \geq q \geq d_1$ obtained here with formula (6.3) and assumption $r \leq \frac{d}{2}$; the following contradiction appears : $d_1 \leq l_1 < \frac{r}{d-r} d_1 \leq d_1$. Therefore Theorem 4.4 is proved.

4.4 Assumptions on the Small ϱ^{op}-Orbits

Throughout this Section, we make some assumptions on the "small" ϱ^{op}-orbits, precisely those which have dimension less than $r = \dim(X)$. Assuming that there is no such orbit (except the trivial one), the following statement is obtained :

Theorem 4.5 *Assume :*

- *that* (ϱ, X) *is suitable.*
- *that all nonzero* ϱ^{op}*-orbits have dimension at least* r.
- *that* X *has dimension* r *with* $d \geq \frac{r(r+1)}{2}$.
- *that one, at least, of the following holds :*
    1. *The union of* $\{0\}$ *and of* ϱ^{op}*-orbits of dimension* r *contains no vector subspace of* W^* *of codimension* $r-1$.
    2. *If* ϕ *and* ϕ' *are elements of* W^* *the* ϱ^{op}*-orbits of which have dimension* r, *and if* $\ker(N_\phi) = \ker(N_{\phi'})$, *then* ϕ *and* ϕ' *are collinear.*

Then every element $x \in X(\mathcal{L})$ *belongs to some hyperplane of* W *defined over* $\bar{\mathbb{Q}}$.

N.B. Assume $r \leq \frac{d}{2}$ and ϱ^{op} has only finitely many orbits of dimension r. Then the union of $\{0\}$ and of ϱ^{op}-orbits of dimension r has dimension at most r, with $r < d-(r-1)$; therefore assumption 1 holds.

Proof of Theorem 4.5 : Assume there exists $x \in X(\mathcal{L})$ that does not belong to any hyperplane of W defined over $\bar{\mathbb{Q}}$. Then applying Proposition 3.1 yields subspaces S and T.

As $d_1 \geq 1$, there exists $\phi \in T^*$ such that $\phi \neq 0$, therefore the ϱ^{op}-orbit of ϕ has dimension at least r. We have $S \subset \ker(N_\phi)$ by Lemma 4.3, therefore Lemma 2.4 yields : $l_1 \geq \mathrm{rk}(N_\phi) \geq r$.

Let us prove that equality $l_1 = r$ does not hold. In fact, otherwise, we would have $\mathrm{rk}(N_\phi) = r$ and $S = \ker(N_\phi)$ for all nonzero $\phi \in T^*$; in particular, by Lemma 2.4, each nonzero $\phi \in T^*$ would belong to some ϱ^{op}-orbit of dimension r. Moreover, relation (6.3) would yield $d_1 > d-r$. Assumption 1 could not hold. Neither could assumption 2, for it would imply $d_1 \leq 1$, hence $d = r < \frac{dl}{d+l}$, which is impossible.

Therefore we have $l_1 \geq r+1$. Thanks to assumption $d \geq \frac{r(r+1)}{2}$ and relations (6.2), we obtain :

$$\frac{(r+1)(d-r)}{d} \geq r-1 \geq r_1 \geq \frac{l_1 d_1}{l_1+d_1} \geq \frac{(r+1)d_1}{(r+1)+d_1}$$

This yields $d(r+1) \geq r(d_1+r+1)$, hence the following contradiction : $r+1 \leq l_1 < \frac{r}{d-r}d_1 \leq r+1$. This concludes the proof of Theorem 4.5.

Another result is the following, which applies when it is possible to control the union of "small" ϱ^{op}-orbits:

Theorem 4.6 *Assume :*

- *that (ϱ, X) is suitable.*
- *that the union of all ϱ^{op}-orbits of dimension less than r contains no vector subspace of W^* of dimension greater than $\frac{d}{r}$.*
- *that X has dimension $r \leq \sqrt{d}$.*

Then every element $x \in X(\mathcal{L})$ belongs to some hyperplane of W defined over $\bar{\mathbb{Q}}$.

Proof of Theorem 4.6 : Assume some $x \in X(\mathcal{L})$ belongs to no hyperplane of W defined over $\bar{\mathbb{Q}}$. Then Proposition 3.1 produces some subspaces S and T.

As $r_1 \geq 1$, relation (6.2) yields $d_1 > \frac{d}{r}$, therefore the union of ϱ^{op}-orbits of dimension less than r contains no vector subspace of W^* of dimension d_1. In particular, T^* is not contained in this union : there exists $\phi \in T^*$ such that $\mathrm{rk}(N_\phi) \geq r$ (thanks to Lemma 2.4). As $S \subset \ker(N_\phi)$ by Lemma 4.3, we obtain $l_1 \geq r$, hence $d_1 > d-r$ and the following contradiction :

$$r-1 \geq r_1 \geq \frac{l_1 d_1}{l_1+d_1} > \frac{r(d-r)}{d} = r - \frac{r^2}{d} \geq r-1$$

This proves Theorem 4.6.

5 Special Cases

5.1 Symmetric, Tensor and Exterior Powers

Let V be a vector space of dimension $m \geq 2$ over K, equipped with a $\bar{\mathbb{Q}}$-structure, and let k be an integer greater than or equal to 2. We fix a basis $(e_1, \ldots, e_m)$ of V.

In this Section, we consider the natural action ϱ of $G = GL(V)$, first on $W = \mathrm{Sym}^k(V)$, then on $W = V^{\otimes k}$, and finally on $W = \Lambda^k(V)$. We sketch the proofs of Theorems 5.1, 5.3 and 5.4 as consequences of the results obtained in Section 4. Detailed proofs are omitted, because stronger statements can be proved by other methods.

First of all, denote by W the vector space $\mathrm{Sym}^k(V)$; it is equipped with an induced $\bar{\mathbb{Q}}$-structure (see the end of Section 2.1). The action ϱ is given by $\varrho(g)(\alpha_1 \cdot \ldots \cdot \alpha_k) = g(\alpha_1) \cdot \ldots \cdot g(\alpha_k)$ for $g \in G$ and $\alpha_1, \ldots, \alpha_k \in V$. For $A \in \mathfrak{g} = \mathrm{End}(V)$ and $\alpha_1, \ldots, \alpha_k \in V$, we have :

$$(\mathrm{Lie}(\varrho)(A))(\alpha_1 \cdot \ldots \cdot \alpha_k) = \sum_{i=1}^{k} \alpha_1 \cdot \ldots \cdot \alpha_{i-1} \cdot A(\alpha_i) \cdot \alpha_{i+1} \cdot \ldots \cdot \alpha_k$$

Denote by $X = \mathcal{V}(k, V)$ the set of elements of the shape $v \cdot \ldots \cdot v$, with $v \in V$. Then $X \setminus \{0\}$ is a ϱ-orbit of dimension m, known as the affine cone (without the origin) over the Veronese variety ([H3], Lecture 2).

Let α be a nonzero element of $\mathrm{Sym}^k(V)$. Denote by g_n the linear automorphism of V which sends e_j to $t_j e_1 + u_n e_j$ for every $j \in \{1, \dots, m\}$, where $t_1, \dots, t_m$ are elements of K and (u_n) is a sequence of elements of $K \setminus \{0, -t_1\}$ which tends to zero. Then $g_n(\alpha)$ tends to $\lambda e_1 \cdot \ldots \cdot e_1$ for the norm topology on $\mathrm{Sym}^k(V)$, where $\lambda \in K$ is nonzero if $t_1, \dots, t_m$ are chosen in a proper way : the closure, for the norm topology, of the orbit Y_α of α intersects X. Therefore the Zariski closure of Y_α, which is a union of ϱ-orbits ([H10], Proposition 8.3), contains X. This proves that X lies in the Zariski closure of any nonzero ϱ-orbit. Now, the contragredient representation ϱ^{op} is isomorphic to the representation ϱ^* of $GL(V^*)$ on $\mathrm{Sym}^k(V^*)$. Therefore every nonzero ϱ^{op}-orbit has dimension at least m, with equality only for the affine cone $X^* \setminus \{0\}$ over the Veronese variety in $\mathrm{Sym}^k(V^*) \simeq (\mathrm{Sym}^k(V))^*$. Assume $(k,m) \neq (2,2)$ and $(k,m) \neq (3,2)$; then (ϱ, X) is suitable. Accordingly the remark following Theorem 4.5 applies (with $r = m$ and $d = \binom{k+m-1}{m-1}$), thereby proving the following :

Theorem 5.1 *Assume $(k,m) \neq (2,2)$ and $(k,m) \neq (3,2)$.*

Then Conjecture 2 holds for the affine cone $\mathcal{V}(k,V)$ over the Veronese variety.

Actually it is possible, as soon as $(k,m) \neq (2,2)$, to prove a more precise result : see Theorem 6.2 below.

Let us now move to $W = V^{\otimes k}$ and $X = \mathcal{T}(k,V) = \{v \otimes \ldots \otimes v;\ v \in V\}$. The situation is quite similar to the previous one, except that there may be nonzero ϱ-orbits of dimension less than $m = \dim(X \setminus \{0\})$. For instance, if $k = m$, the orbit of $\alpha = \sum_{\sigma \in \mathfrak{S}_k} \varepsilon_\sigma e_{\sigma(1)} \otimes \ldots \otimes e_{\sigma(k)}$ has dimension 1 (in this formula, ε_σ is the sign of σ).

We are going to apply Theorem 4.6 (with $r = m$ and $d = m^k$) ; to do this, we need to control the orbits of dimension less than m. With this aim in view, we consider the basis (e_I) of $V^{\otimes k}$, where $I = (i_1, \dots, i_k)$ runs through $\{1, \dots, m\}^k$, and its dual basis (e_I^*).

Two families $I = (i_1, \dots, i_k)$ and $I' = (i'_1, \dots, i'_k)$ are said to be *anagrams* if there exists a permutation σ of $\{1, \dots, k\}$ such that the following holds : $i'_1 = i_{\sigma(1)}, \dots, i'_k = i_{\sigma(k)}$.

For any family $I = (i_1, \dots, i_k)$ and any integer $s \in \{1, \dots, m\}$, we denote by $\mathcal{N}_s(I)$ the number of indices $j \in \{1, \dots, k\}$ such that $i_j = s$. Two families I and I' are anagrams if, and only if, the equality $\mathcal{N}_s(I) = \mathcal{N}_s(I')$ holds for all s.

These definitions allow us to prove the following lemma :

Lemma 5.2 *Let $\alpha = \sum_I \alpha_I e_I$ be a nonzero element of $V^{\otimes k}$, the orbit of which has dimension less than m.*

Then k is a multiple of m, and the families I such that $\alpha_I \neq 0$ are anagrams of the family $(1, \dots, 1, 2, \dots, 2, \dots, m)$ in which every integer from 1 to m is repeated $\frac{k}{m}$ times.

Sketch of proof of Lemma 5.2 : Assume the conclusion does not hold. Considering the linear automorphism of V which maps e_j to $\mu_j e_j$, for suitable values of $\mu_1, \dots, \mu_m \in K$, makes it possible to find a family I_0 and a nonzero element β of $V^{\otimes k}$ with the following properties :

- Either k is not a multiple of m, or I_0 is not an anagram of the family $(1, \dots, 1, 2, \dots, 2, \dots, m)$ in which every integer from 1 to m is appears $\frac{k}{m}$ times.
- The element β lies in the closure (for the norm topology, and therefore for the Zariski topology) of the orbit of α.
- We have $\beta = \sum_I \beta_I e_I$ where $\beta_I = \alpha_I$ if I is an anagram of I_0, and $\beta_I = 0$ otherwise ; moreover, $\beta_{I_0} = \alpha_{I_0} \neq 0$.

Then the orbit of β is contained in the Zariski closure of that of α, hence has dimension less than m. Thanks to Lemma 2.4, this implies $\mathrm{rk}(M_\beta) < m$. We shall now construct m families

$J_1, \dots, J_m$ such that $(e^*_{J_1} \circ M_\beta, \dots, e^*_{J_m} \circ M_\beta)$ are m linearly independent linear forms ; this will give the desired contradiction.

Let H be the set of elements $s \in \{1, \dots, m\}$ such that $\mathcal{N}_s(I_0) = u$, where u is chosen in such a way that H is neither empty nor equal to $\{1, \dots, m\}$. Denote by M the complement of H in $\{1, \dots, m\}$, and let $(s', s'') \in H \times M$. Let $\mathcal{J}_{s',s''}$ be the set of families that can be deduced from I_0 by replacing exactly one occurrence of s' by s''. For $J \in \mathcal{J}_{s',s''}$, let $\lambda_J = \sum_I \alpha_I$, where the sum is taken over all anagrams I of I_0 that can be deduced from J by replacing exactly one occurrence of s'' by s'. Then the relation $e^*_J(M_\beta(A)) = \lambda_J a_{s'',s'}$ holds for any $A \in \mathrm{End}(V)$ identified with its matrix $(a_{i,j})$ in the basis $(e_1, \dots, e_m)$. Define $\mathcal{J}_{s'',s'}$ in the same way ; then for any $J \in \mathcal{J}_{s'',s'}$ there is $\lambda_J \in K$ such that $e^*_J(M_\beta(A)) = \lambda_J a_{s',s''}$ for every A. Now there exists a family $J(s', s'')$, which belongs either to $\mathcal{J}_{s',s''}$ or to $\mathcal{J}_{s'',s'}$, such that $\lambda_{J(s',s'')} \neq 0$. As (s', s'') ranges through $H \times M$, we obtain in this way $(\#H)(\#M) \geq m - 1$ families $J(s', s'')$ among which we select $J_1, \dots, J_{m-1}$. We let $J_m = I_0$; then an easy computation shows that $(e^*_{J_1} \circ M_\beta, \dots, e^*_{J_m} \circ M_\beta)$ are linearly independent linear forms. This proves Lemma 5.2.

Theorem 5.3 *Assume* $(k, m) \neq (2, 2)$.
Then Conjecture 2 holds for the subset $\mathcal{T}(k, V)$ *of* $V^{\otimes k}$.

Again, a more precise result can be obtained : see Theorem 6.4 in Section 6.

Proof of Theorem 5.3 : If k is a multiple of m, let I_1 be the family $(1, \dots, 1, 2, \dots, 2, \dots, m)$ in which every integer from 1 to m is repeated $\frac{k}{m}$ times, and denote by F the subspace of $V^{\otimes k}$ spanned by those vectors e_I such that I is an anagram of I_1. If k is not a multiple of m, let $F = \{0\}$. Then, in both cases, F has dimension less than or equal to m^{k-1}, and contains every ϱ-orbit of dimension less than m thanks to Lemma 5.2. Further, $(\varrho, X \setminus \{0\})$ is suitable as soon as $(k, m) \neq (2, 2)$, and ϱ^{op} is isomorphic to ϱ^* (in the same way as in the proof of Theorem 5.1). Therefore Theorem 4.6 applies.

Finally, let us turn to $W = \Lambda^k(V)$. Let $X = G(k, V) = \{v_1 \wedge \ldots \wedge v_k \in \Lambda^k(V);\ v_1, \dots, v_k \in V\}$ be the affine cone over the Grassmannian whose points are the k-dimensional subspaces of V. Then $X \setminus \{0\}$ is a ϱ-orbit of dimension $r = k(m-k) + 1$ ([H3], Lecture 6).

Theorem 5.4 *Assume that* $2 \leq k \leq m - 2$ *and* $(k, m) \neq (2, 4)$.
Then Conjecture 2 holds for $G(k, V)$.

Theorem 5.4 is a weaker statement than Theorem 2.1 of [R13]. It can be proved as a consequence of Theorems 4.1 and 4.5, except for a few pairs (k, m). Indeed, changing k and V into $m - k$ and V^* if necessary, we may assume $k \leq \frac{m}{2}$. If $k = 2$, each element α of $\Lambda^2(V)$ has an even rank $p \leq m$, which is the only nonnegative integer such that α can be written $v_1 \wedge v_2 + \ldots + v_{p-1} \wedge v_p$ with linearly independent vectors $v_1, \dots, v_p$ ([M5], pages 177 and 192). There are $[\frac{m}{2}] + 1$ ϱ-orbits, each of them corresponding to a value p of the rank. One of them (with $p = m$ or $p = m - 1$) is dense, therefore Theorem 4.1 applies if $m \geq 8$. On the other hand, if $3 \leq k \leq m - 3$ and $m \geq 25$ then $d = \binom{m}{k} \geq \frac{r(r+1)}{2}$. Moreover, ϱ^{op} is isomorphic to the natural representation of $GL(V^*)$ on $\Lambda^k(V^*)$, hence every ϱ^{op}-orbit has dimension greater than r, except $\{0\}$ and $G(k, V^*) \setminus \{0\}$. This allows us to apply Theorem 4.5. To conclude the proof of Theorem 5.4, that is to deal with the pairs (k, m) such that $k \in \{2, m-2\}$ and $5 \leq m \leq 7$, or $3 \leq k \leq m - 3$ and $m \leq 24$, we apply Proposition 3.1, Lemma 4.3, and we use arguments that are specific to $X = G(k, V)$ (for instance, we bound from below the rank of

f_A as soon as $A \in \mathfrak{gl}(V)$ is nonzero). As far as the remaining pair $(k,m) = (2,4)$ is concerned, nothing can be deduced from Proposition 3.1 because $(\varrho, X \setminus \{0\})$ is not suitable. It may also be noted that Theorem 2.1 of [R13] is trivial when $(k,m) = (2,4)$: nothing is known about the points of $G(2, K^4)(\mathcal{L})$.

5.2 Centralizers of Matrices

In this Section, we consider the action of $\mathrm{GL}_n(K)$ on $\mathrm{Mat}_n(K)$ by conjugation. For $M \in \mathrm{Mat}_n(K)$, let $\mathcal{C}(M)$ be the centralizer of M, that is the space of all matrices A such that $[A,M] = 0$ (where $[A,M] = AM - MA$). Then the orbit of M under the action of $\mathrm{GL}_n(K)$ has dimension equal to the codimension, in $\mathrm{Mat}_n(K)$, of $\mathcal{C}(M)$. This enables us to prove Theorem 1.2 stated in the Introduction, as a corollary of the following statement :

Proposition 5.5 *Let M be a square matrix of size n, with entries in $\mathcal{L}$, whose centralizer $\mathcal{C}(M)$ has dimension greater than $\frac{1}{2}(n^2+1)$. Then there exist vector subspaces U and V of $\mathrm{Mat}_n(K)$, defined over $\bar{\mathbb{Q}}$, with the following properties :*

- *We have* $\dim(U) + \dim(V) \geq n^2 + 2$.
- *For every $A \in U$ and every $B \in V$, we have* $\mathrm{Trace}(M\,[A,B]) = 0$.

To deduce Theorem 1.2 from Proposition 5.5, it suffices to exclude the case where $[A,B] = 0$ for any $A \in U$ and any $B \in V$. This is done by the following lemma, the proof of which was provided to me by Gaël Rémond :

Lemma 5.6 *Let U and V be vector subspaces of* $\mathrm{Mat}_n(K)$ *such that each element of U commutes with each element of V. Then the following holds :*

$$\dim(U) + \dim(V) \leq n^2 + 1$$

Proof of Lemma 5.6 : This statement is obvious for $n = 1$; let us prove it by induction on n. First of all, it is possible to replace U and V by $U \cap V$ and $U + V$, so we can assume $U \subset V$. Now, if $U \subset K\mathrm{Id}$ then the conclusion holds trivially, therefore we can assume there is a matrix $M \in U$ such that $M \notin K\mathrm{Id}$. Let λ be an eigenvalue of M, and $F = \ker(M - \lambda\mathrm{Id})$. Then F is stable under every matrix that commutes with M, in particular under every matrix of V. Choose a basis of K^n whose $\dim(F)$ first vectors belong to F. In this basis, the elements of V are of the shape $\left(\begin{smallmatrix} A & B \\ 0 & C \end{smallmatrix}\right)$.

Let U' (respectively V') be the set of those matrices A for which there exist matrices B and C such that $\left(\begin{smallmatrix} A & B \\ 0 & C \end{smallmatrix}\right) \in U$ (respectively $\left(\begin{smallmatrix} A & B \\ 0 & C \end{smallmatrix}\right) \in V$). In the same way, define U'' (respectively V'') to be the set of those matrices C for which there exist matrices A and B such that $\left(\begin{smallmatrix} A & B \\ 0 & C \end{smallmatrix}\right) \in U$ (respectively $\left(\begin{smallmatrix} A & B \\ 0 & C \end{smallmatrix}\right) \in V$).

Then U is contained in the set of all matrices $\left(\begin{smallmatrix} A & B \\ 0 & C \end{smallmatrix}\right)$ such that $A \in U'$ and $C \in U''$. This implies :

$$\dim(U) \leq \dim(U') + \dim(U'') + (\dim(F))(\mathrm{codim}(F))$$

An analogous inequality holds for V ; by summing up and applying induction to (U', V') and (U'', V''), we obtain :

$$\dim(U) + \dim(V) \leq 1 + (\dim(F))^2 + 1 + (\mathrm{codim}(F))^2 + 2(\dim(F))(\mathrm{codim}(F))$$

This inequality means $\dim(U) + \dim(V) \leq n^2 + 2$; in order to conclude the proof, it suffices to check that equality does not hold. Assume, on the contrary, that equality holds. Then U is equal to the set of all matrices $\left(\begin{smallmatrix} A & B \\ 0 & C\end{smallmatrix}\right)$ such that $A \in U'$ and $C \in U''$, and V is equal to the set of matrices $\left(\begin{smallmatrix} A & B \\ 0 & C\end{smallmatrix}\right)$ such that $A \in V'$ and $C \in V''$. Moreover, we then have $\dim(U') + \dim(V') = 1 + (\dim(F))^2$, therefore U' contains a nonzero matrix A. There is a matrix B, with $\dim(F)$ rows and $\operatorname{codim}(F)$ columns, such that $AB \neq 0$. But $\left(\begin{smallmatrix} A & B \\ 0 & 0\end{smallmatrix}\right) \in U$ commutes with $\left(\begin{smallmatrix} 0 & B \\ 0 & 0\end{smallmatrix}\right) \in V$, that is $AB = 0$. This contradiction concludes the proof of Lemma 5.6.

N.B. Using the same kind of methods, it is possible to prove that under the assumptions of Lemma 5.6, if $\dim(U) + \dim(V) = n^2 + 1$ then $U = K\mathrm{Id}$ or $V = K\mathrm{Id}$.

Proof of Proposition 5.5 : Denote by H the space of those matrices A such that $\mathrm{Trace}(A) = 0$; it is equipped with a $\bar{\mathbb{Q}}$-structure $H(\bar{\mathbb{Q}})$ consisting in those matrices with entries in $\bar{\mathbb{Q}}$. Let M be a square matrix of size n, whose n^2 entries belong to $\mathcal{L}$ and whose centralizer $\mathcal{C}(M)$ has codimension r, with $r < \frac{1}{2}(n^2 - 1)$.

Let $u = \mathrm{adM}$ be the endomorphism of H which sends any matrix N to $[M, N]$. Then u has rank r, and Corollary 2.2 (applied with E and F equal to H) produces subspaces S and T of H, defined over $\bar{\mathbb{Q}}$, such that $l_1 < d_1$ and $[M, A] \in T$ for all $A \in S$. Denote by T^* the orthogonal subspace to T, for the non-degenerate symmetric bilinear form $(X, Y) \mapsto \mathrm{Trace}(XY)$ on H. Then the following holds for every $A \in S$ and every $B \in T^*$:

$$0 = \mathrm{Trace}([M, A]\, B) = \mathrm{Trace}(M\, [A, B])$$

Let $U = S \oplus K\mathrm{Id}$ and $V = T^* \oplus K\mathrm{Id}$; then $\dim(U) + \dim(V) \geq n^2 + 2$, thereby proving Proposition 5.5.

N.B. We actually proved that it is enough, in Theorem 1.2, to assume that there is no hyperplane of $\mathrm{Mat}_n(K)$, defined over $\bar{\mathbb{Q}}$, which contains both M and Id. A weaker assumption might be sufficient ; such an improvement of Theorem 1.2 could be derived from an answer to the following question. Let U and V be vector subspaces of $\mathrm{Mat}_n(K)$ such that $\dim(U) + \dim(V) \geq n^2 + 2$. How small can the vector space spanned by $[U, V]$ be ?

In another direction, if the algebraic independence Conjecture holds, then a result stronger than Theorem 1.2 follows :

Proposition 5.7 *Take for granted the algebraic independence Conjecture.*

Let M be a square matrix of size n, whose n^2 entries belong to $\mathcal{L}$ and are linearly independent over $\bar{\mathbb{Q}}$.

Then the centralizer of M has dimension n.

Proof of Proposition 5.7 : Assume, on the contrary, that $\mathcal{C}(M)$ has dimension greater than n. Then not all eigenvalues of M are simple, and the characteristic polynomial χ_M of M has zero discriminant. Apply Conjecture 1 stated in the Introduction : M belongs to a linear affine subspace E of $\mathrm{Mat}_n(K)$, defined over $\bar{\mathbb{Q}}$ and contained in the set of matrices N such that χ_N has zero discriminant. We can assume that E is an affine hyperplane of $\mathrm{Mat}_n(K)$, otherwise E would be contained in a vector hyperplane and the n^2 entries of M would not be linearly independent over $\bar{\mathbb{Q}}$. Moreover, we can assume $n \geq 3$, otherwise Proposition 5.7 holds trivially. Consider now the linear affine subspace consisting in those matrices which are upper triangular with diagonal entries equal to $(1, 2, \ldots, n)$. This subspace has dimension at least 2, therefore it intersects E ; this is impossible because for every $N \in E$ the polynomial χ_N has zero discriminant. This ends the proof of Proposition 5.7.

N.B. Following the same lines as in this Section, it is possible to study the space of symmetric (or skew-symmetric, or triangular) matrices commuting with a given symmetric (or skew-symmetric, or triangular) matrix M with entries in $\mathcal{L}$.

6 Proof of Conjecture 1 in Special Cases

Let k and m be positive integers. The following notation is analogous to that used in Section 5.1 :

$$\begin{aligned}\mathcal{V}(k,m) &= \{v\cdot\ldots\cdot v;\ v\in K^m\}\subset \mathrm{Sym}^k(K^m)\\ \mathcal{T}(k,m) &= \{v\otimes\ldots\otimes v;\ v\in K^m\}\subset (K^m)^{\otimes k}\end{aligned}$$

When $k=1$ or $m=1$, we have $\mathcal{V}(k,m)=\mathrm{Sym}^k(K^m)$ and $\mathcal{T}(k,m)=(K^m)^{\otimes k}$: Conjecture 1 holds trivially for these subsets. This is the reason why we now assume $k\geq 2$ and $m\geq 2$.

In order to prove Theorems 6.2 and 6.4 stated below, we shall use the following lemma :

Lemma 6.1 *Let $\lambda_1,\ldots,\lambda_q$ be nonzero elements of $\mathcal{L}$ in geometric progression with transcendental ratio. Then $q\leq 3$.*

Proof of Lemma 6.1 : Let t be the ratio ; then $\lambda_j=\lambda_1 t^{j-1}$ for all $j\in\{1,\ldots,q\}$. The matrix $M=\begin{bmatrix}\lambda_1 & \lambda_1 t & \ldots & \lambda_1 t^{q-2}\\ \lambda_1 t & \lambda_1 t^2 & \ldots & \lambda_1 t^{q-1}\end{bmatrix}$ has rank 1 and entries in $\mathcal{L}$; if $q\geq 4$, Theorem 2.3 shows that either the rows or the columns of M are linearly dependent over $\bar{\mathbb{Q}}$. In the former case, t would be algebraic, which is impossible. In the latter case, t would be a root of a nonzero polynomial with coefficients in $\bar{\mathbb{Q}}$, which is impossible too. This ends the proof of Lemma 6.1.

Let us now state our result concerning the sets $\mathcal{V}(k,m)$. It is helpful to see $\mathrm{Sym}^k(K^m)$ as the space of homogeneous polynomials of degree k, in m variables, with coefficients in K ; then $\mathcal{V}(k,m)$ consists in those polynomials which are the k-th power of a linear form.

Theorem 6.2 *Let k and m be integers greater than or equal to 2.*

Let P be a homogeneous polynomial of degree k, in m variables, with coefficients in $\mathcal{L}$. Assume that P is the k-th power of a linear form Λ with coefficients in K.

1. *If $k\geq 3$ then there exist a linear form ϕ, with coefficients in $\bar{\mathbb{Q}}$, and an element a of $\mathcal{L}$, such that $P=a\phi^k$.*
2. *If $k=2$ then there exist two linear forms ϕ_1 and ϕ_2, with coefficients in $\bar{\mathbb{Q}}$, and two elements a and b of K, such that $P=(a\phi_1+b\phi_2)^k$.*

Corollary 6.3 1. *Conjecture 1 holds for $\mathcal{V}(k,m)$ as soon as $k\geq 3$.*

2. *If Conjecture 1 holds for $\mathcal{V}(2,2)$ then it holds for $\mathcal{V}(k,m)$, for any pair (k,m).*

Proof of Theorem 6.2 : Let n be the minimal integer such that there exist linear forms $\phi_1,\ldots,\phi_n$, with coefficients in $\bar{\mathbb{Q}}$, and elements $c_1,\ldots,c_n$ of K such that $\Lambda=\sum_{i=1}^n c_i\phi_i$. Then $c_1,\ldots,c_n$ are linearly independent over $\bar{\mathbb{Q}}$; on the other hand, $\phi_1,\ldots,\phi_n$ are algebraically independent over K : we denote them by $Y_1,\ldots,Y_n$ and write $\Lambda=\sum_{i=1}^n c_iY_i$. Furthermore, we let $P=\Lambda^k=\sum p_{i_1,\ldots,i_n}Y_1^{i_1}\ldots Y_n^{i_n}$; note that the coefficients $p_{i_1,\ldots,i_n}$ belong to $\mathcal{L}$ since $\phi_1,\ldots,\phi_n$ are defined over $\bar{\mathbb{Q}}$ and linearly independent over K. To begin with, we shall prove that n is at most 2.

Assume $n \geq 3$. Let $Q = (\frac{\partial}{\partial Y_1})^{k-2}P$. Then Q is a nonzero homogeneous polynomial of degree 2 in $Y_1, \ldots, Y_n$, with coefficients in $\mathcal{L}$, such that $Q = \frac{k!}{2}c_1^{k-2}\Lambda^2$. Let δ be a square root of $\frac{k!}{2}c_1^{k-2}$, and $\Lambda' = \delta\Lambda$; then $Q = \Lambda'^2$. Now, associate with each homogeneous polynomial $R = \sum_{i,j} r_{i,j}Y_iY_j$ of degree 2 (written in such a way that $r_{i,j} = r_{j,i}$ for all $i, j \in \{1, \ldots, n\}$) the symmetric matrix $\tilde{R} = (r_{i,j})$ of size n. Denoting by $u \in K^n$ the coordinate vector $(\delta c_1, \ldots, \delta c_n)$ of Λ', the relation $Q = \Lambda'^2$ yields $\tilde{Q} = u\,{}^t u$. Consequently, $\tilde{Q}$ is a square matrix, of size $n \geq 3$, of rank 1, with entries in $\mathcal{L}$: Theorem 2.3 applies to $\tilde{Q}$. Therefore the n coordinates of Λ' are linearly dependent over $\bar{\mathbb{Q}}$, and so are $c_1, \ldots, c_n$, in contradiction with the definition of n.

Therefore $n \leq 2$: Theorem 6.2 is proved if $k = 2$. Assume now $k \geq 3$ and $n = 2$. Then $P = (c_1Y_1 + c_2Y_2)^k = \sum_{j=0}^k a_j Y_1^j Y_2^{k-j}$, with $a_0, \ldots, a_k \in \mathcal{L}$. Let $t = \frac{c_1}{c_2} \in K$ and $a_j = \binom{k}{j}a'_j$ for all $j \in \{0, \ldots, k\}$; then $a'_j = t^j a'_0$. But t is transcendental, because c_1 and c_2 are linearly independent over $\bar{\mathbb{Q}}$. This contradicts Lemma 6.1, thereby proving Theorem 6.2.

Proof of Corollary 6.3 : Let P be as in Theorem 6.2. If $k \geq 3$ then $P = a\phi^k$ belongs to the subspace $K\phi^k$, which is contained in $\mathcal{V}(k,m)$ and defined over $\bar{\mathbb{Q}}$. Assume now that $k = 2$, and that Conjecture 1 holds for $\mathcal{V}(2,2)$. Let $P = (a\phi_1 + b\phi_2)^2$ be a polynomial with coefficients in $\mathcal{L}$, and consider (in the same way as in the proof of Theorem 6.2) $P_1 = (aY_1 + bY_2)^2$. Then P_1 belongs to a linear affine subspace E_1 contained in $\mathcal{V}(2,2)$ and defined over $\bar{\mathbb{Q}}$. Let E be the linear affine subspace of $\mathrm{Sym}^2(K^m)$ consisting in those polynomials Q such that there is $Q_1 \in E_1$ with $Q(X_1, \ldots, X_m) = Q_1(\phi_1(X_1, \ldots, X_m), \phi_2(X_1, \ldots, X_m))$. Then $P \in E$, E is defined over $\bar{\mathbb{Q}}$ and $E \subset \mathcal{V}(2,m)$, thereby proving Corollary 6.3.

The proof of Theorem 6.2 given above can be easily translated in terms of symmetric powers ; dealing with tensor powers, the following result is obtained in a similar way :

Theorem 6.4 *Let k and m be integers greater than or equal to 2.*
Let $v \in K^m$ be such that $x = v \otimes \ldots \otimes v$ belongs to $(K^m)^{\otimes k}(\mathcal{L})$. Then :

1. *If $k \geq 3$, there exist $a \in \mathcal{L}$ and $v' \in \bar{\mathbb{Q}}^m$ such that $x = av' \otimes \ldots \otimes v'$.*
2. *If $k = 2$, there exists a vector subspace F of K^m, defined over $\bar{\mathbb{Q}}$, of dimension 2, which contains v.*

Corollary 6.5

1. *Conjecture 1 holds for $\mathcal{T}(k,m)$ as soon as $k \geq 3$.*
2. *If Conjecture 1 holds for $\mathcal{T}(2,2)$ then it holds for $\mathcal{T}(k,m)$, for any pair (k,m).*

Proof of Theorem 6.4 : Let F be the smallest vector subspace of K^m, defined over $\bar{\mathbb{Q}}$, that contains v. Let $n = \dim(F)$.

First of all, assume $n \geq 3$. Let ξ be a nonzero linear form on F, defined over $\bar{\mathbb{Q}}$; then $< \xi, v > \neq 0$. Denote by D the linear endomorphism of the tensor algebra $T(F)$ which maps any $\alpha = \alpha_1 \otimes \ldots \otimes \alpha_q$ to

$$D(\alpha) = \sum_{j=1}^q < \xi, \alpha_j > \alpha_1 \otimes \ldots \otimes \alpha_{j-1} \otimes \alpha_{j+1} \otimes \ldots \otimes \alpha_q$$

Let $Q = D^{k-2}(x) = \frac{k!}{2} < \xi, v >^{k-2} v \otimes v$ and $v' = \delta v$ where δ is a square root of $\frac{k!}{2} < \xi, v >^{k-2}$. Choose a basis $(f_1, \ldots, f_n)$ of F, defined over $\bar{\mathbb{Q}}$, and associate with any $R = \sum_{i,j} r_{i,j} f_i \otimes f_j$ the matrix $\tilde{R} = (r_{i,j})$. Then $\tilde{Q} = u\,{}^t u$ where u is the coordinate vector of v' in the basis

$(f_1, \dots, f_n)$. As $n \geq 3$, Theorem 2.3 shows that v' belongs to some vector hyperplane of F defined over $\bar{\mathbb{Q}}$; then so does v, in contradiction with the definition of F.

Therefore $n \leq 2$, and Theorem 6.4 is proved if $k = 2$. Let us assume $k \geq 3$ and $n = 2$. Let (f_1, f_2) be a basis of F defined over $\bar{\mathbb{Q}}$; since $n = 2$, we can write $v = a(f_1 + tf_2)$ with $a, t \in K^\star$. Then, for every $j \in \{0, \dots, k\}$, $a^k t^j$ belongs to $\mathcal{L}$: Lemma 6.1 implies $t \in \bar{\mathbb{Q}}$, in contradiction with the definition of F. This ends the proof of Theorem 6.4; Corollary 6.5 immediately follows.

N.B. As Damien Roy pointed out to me, Theorem 6.2 and Theorem 6.4 can be easily deduced from each other by considering the linear embedding of $\mathrm{Sym}^k(K^m)$ into $(K^m)^{\otimes k}$ which sends $v_1 \cdot \ldots \cdot v_k$ to $\frac{1}{k!} \sum v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(k)}$, where the sum is over all permutations σ of $\{1, \dots, k\}$. Indeed, this embedding maps $\mathcal{V}(k,m)$ onto $\mathcal{T}(k,m)$.

N.B. The following conjectures are equivalent :

1. Conjecture 1 holds for $\mathcal{V}(2,2)$.
2. Conjecture 1 holds for $\mathcal{T}(2,2)$.
3. It is possible to replace $q \leq 3$ by $q \leq 2$ in the conclusion of Lemma 6.1.

These conjectures are consequences of the four exponential Conjecture (i.e. the assertion that Theorem 2.3 holds when $d = l = 2$). There seems to be a gap between these conjectures and the theorems proved up to now; actually, it is impossible ([R10], Proposition 2) to derive "algebraically" any of these conjectures from Theorem 2.1.

Acknowledgements: I am grateful to Michel Waldschmidt and Damien Roy for the time they spent with me, in discussions I always found exciting.

Chapitre 7

Lemme d'interpolation dans un groupe algébrique commutatif

1 Introduction

Dans ce texte, on démontre un lemme d'interpolation avec multiplicités (énoncé au paragraphe 1.1). L'un des ingrédients essentiels de la preuve est la traduction en termes de fonctionnelles (au paragraphe 1.2) de cet énoncé. On étudie au paragraphe 2 des opérations sur les fonctionnelles (passage au quotient, translation, dérivation). La partie 3 est dédiée à une étude générale des sous-groupes obstructeurs à l'interpolation et au lemme de zéros. Enfin, dans la partie 4 on utilise les outils introduits précédemment pour démontrer le lemme d'interpolation.

1.1 Énoncé du théorème

On considère la situation de [W11]. Soit G un groupe algébrique commutatif connexe, plongé de manière quasiprojective dans $\mathbb{P}^N$ (via le choix d'un diviseur très ample). Le corps de base est toujours $\mathbb{C}$ (on pourrait aussi considérer son analogue p-adique $\mathbb{C}_p$) ; on identifie G à l'ensemble $G(\mathbb{C})$ de ses points complexes. Notons n la dimension de G, et TG l'espace tangent à G en l'identité. On identifie TG à l'espace des champs de vecteurs sur G invariants par translation à gauche.

Notons $\mathbb{C}[G] = \oplus_{D \geq 0} \mathbb{C}[G]_D^h$ l'algèbre graduée des polynômes sur G (plongé dans $\mathbb{P}^N$), où $\mathbb{C}[G]_D^h = H^0(G, \mathcal{O}(D))$ est l'espace des polynômes homogènes sur G de degré D.

Soit W un sous-espace vectoriel de TG, de dimension d. Soit $(\partial_1, \dots, \partial_d)$ une base de W. Pour tout $i \in \{1, \dots, d\}$ il existe [W11] une dérivation Δ_i de $\mathbb{C}[G]$, qui envoie tout polynôme P homogène de degré D sur un polynôme homogène de degré $D + \kappa$ tel que

$$\partial_i \left(\frac{P}{X_0^D} \right) = \frac{\Delta_i P}{X_0^{D+\kappa}}$$

comme fonctions analytiques sur l'ouvert de G où X_0 ne s'annule pas.

Les dérivations $\Delta_1, \dots, \Delta_d$ commutent. Pour $\sigma \in \mathbb{N}^d$ on pose $\underline{\Delta}^\sigma = \Delta_1^{\sigma_1} \dots \Delta_d^{\sigma_d}$. On utilisera la formule de Leibniz : pour $P, R \in \mathbb{C}[G]$ et $\sigma \in \mathbb{N}^d$ on a

$$\underline{\Delta}^\sigma(PR) = \sum_{\tilde{\sigma} \leq \sigma} \binom{\sigma}{\tilde{\sigma}} (\underline{\Delta}^{\tilde{\sigma}} P)(\underline{\Delta}^{\sigma - \tilde{\sigma}} R)$$

Soit Γ un sous-groupe de G de type fini, engendré par $(\gamma_1, \dots, \gamma_l)$. Pour $S \in \mathbb{R}_+$ on note $\Gamma(S)$ l'ensemble des combinaisons linéaires $n_1\gamma_1 + \dots + n_l\gamma_l$, avec n_j entier et $0 \leq n_j < S$ pour tout $j \in \{1, \dots, l\}$.

Supposons (sans perte de généralité) :

$$\Gamma \subset \{X_0 \neq 0\} \subset \mathbb{P}^N \tag{7.1}$$

Théorème 1.1 *Il existe une constante c_1 strictement positive, qui dépend de toutes les données précédentes, ayant la propriété suivante. Soient D, S, T trois entiers strictement positifs tels que, pour tout sous-groupe algébrique H connexe, non nul, de G on ait*

$$S^{\mathrm{rg}(\Gamma\cap H)} T^{\dim(W\cap TH)} < c_1 D^{\dim(H)}$$

Pour tous $\gamma \in \Gamma(S)$ et $\sigma \in \mathbb{N}^d$, $|\sigma| < T$, soit $a_{\gamma,\sigma}$ un nombre complexe. Alors il existe $P \in \mathbb{C}[G]$, homogène de degré D, tel que $\frac{\Delta^\sigma P}{X_0^{D+|\sigma|\kappa}}(\gamma) = a_{\gamma,\sigma}$ pour tous $\gamma \in \Gamma(S)$ et $\sigma \in \mathbb{N}^d$, $|\sigma| < T$.

Cet énoncé est optimal, à la valeur près de la constante c_1. En effet, soit H un sous-groupe algébrique connexe, non nul, de G tel que $S^{\mathrm{rg}(\Gamma\cap H)} T^{\dim(W\cap TH)} \geq c_1' D^{\dim(H)}$. Comme H n'apparaît dans cette condition qu'à travers $\mathrm{rg}(\Gamma \cap H)$, $\dim(W \cap TH)$ et $\dim(H)$, on peut supposer que H appartient à un ensemble fini de sous-groupes algébriques de G, fixé à l'avance (en fonction de G, Γ et W seulement). Notons Φ l'application linéaire qui à $P \in \mathbb{C}[G]_D^h$ associe la famille des valeurs, aux points de $\Gamma(S) \cap H$, de ses dérivées le long de $W \cap TH$ jusqu'à l'ordre $T - 1$. Alors Φ se factorise à travers $\mathbb{C}[H]_D^h$ (qui est un quotient de $\mathbb{C}[G]_D^h$). Si la constante c_1' est assez grande (en fonction de G, Γ et W), l'inégalité $S^{\mathrm{rg}(\Gamma\cap H)} T^{\dim(W\cap TH)} \geq c_1' D^{\dim(H)}$ implique que l'espace d'arrivée de Φ est de dimension strictement plus grande que $\mathbb{C}[H]_D^h$; alors Φ ne peut pas être surjective.

Dans le cas où il n'y a pas de multiplicités (i.e. $W = \{0\}$ ou $T = 1$), ce théorème est dû à Masser [M2]. La démonstration donnée au paragraphe 4 ci-dessous suit les mêmes lignes que celle de [M2] : on traite un cas particulier "non dégénéré" (§ 4.1), en utilisant un lemme de zéros[1]. Puis on en déduit le cas général (§ 4.2 à 4.4), en utilisant une récurrence sur la dimension du groupe algébrique G.

N.B. Si le groupe G est affine, le théorème 1.1 est un cas particulier du théorème 2.3 de [F7] (voir le lemme 1.2 ci-dessous).

1.2 Formulation en termes de fonctionnelles

On note $\mathbb{C}\Gamma$ l'algèbre du groupe Γ : les éléments de $\mathbb{C}\Gamma$ sont les combinaisons linéaires finies (formelles, à coefficients dans $\mathbb{C}$) d'éléments de Γ. On note $\mathrm{Sym}(W) = \oplus_{k\geq 0}\mathrm{Sym}^k W$ l'algèbre symétrique de W. On pose :

$$\mathcal{F} = \mathbb{C}\Gamma \otimes \mathrm{Sym}(W)$$

On appelle *fonctionnelle* tout élément de $\mathcal{F}$, i.e. toute combinaison linéaire finie d'éléments de la forme $\gamma \otimes (\partial^{(1)} \cdot \ldots \cdot \partial^{(k)})$ avec $\gamma \in \Gamma$ et $\partial^{(1)}, \dots, \partial^{(k)} \in W$. Si une confusion est possible, on écrira $\mathcal{F}^{\Gamma,W}$ au lieu de $\mathcal{F}$.

[1]On parle ici de *lemme de zéros* même quand il y a des multiplicités (certains auteurs préfèrent utiliser le terme *lemme de multiplicités* dans ce cas).

On peut évaluer une fonctionnelle $\eta \in \mathcal{F}$ sur un polynôme $P \in \mathbb{C}[G]_D^h$: par linéarité, il suffit de définir cette évaluation quand $\eta = \gamma \otimes (\partial^{(1)} \cdot \ldots \cdot \partial^{(k)})$, et dans ce cas on pose :

$$\eta(P) = \partial^{(1)} \ldots \partial^{(k)} \left(\frac{P(\underline{X})}{X_0^D} \right) (\gamma)$$

ce qui a bien un sens d'après (7.1). On définit ainsi une forme bilinéaire de $\mathcal{F} \times \mathbb{C}[G]_D^h$ dans $\mathbb{C}$.

On définit une bi-filtration sur $\mathcal{F}$ en posant, pour $S, T \in \mathbb{N}$:

$$\mathcal{F}_{S,T} = (\mathrm{Vect}_{\mathbb{C}} \Gamma(S)) \otimes \left(\oplus_{0 \leq k < T} \mathrm{Sym}^k W \right) \tag{7.2}$$

La définition de $\mathcal{F}$ et de son action sur $\mathbb{C}[G]$ ne dépend ni du choix d'une base de W, ni du choix d'une famille génératrice de Γ. La définition de $\mathcal{F}_{S,T}$, quant à elle, dépend du choix de γ_1, …, γ_l (qui est implicite dans la notation $\Gamma(S)$). Pour toute partie finie Σ de $\mathbb{P}^N$, contenue dans l'ouvert $\{X_0 \neq 0\}$, on pose $\mathcal{F}^{\Sigma,W} = (\mathrm{Vect}_{\mathbb{C}} \Sigma) \otimes \mathrm{Sym}(W)$, où les combinaisons linéaires d'éléments de Σ sont formelles. L'espace $\mathcal{F}^{\Sigma,W}$ est filtré par les $\mathcal{F}_T^{\Sigma,W}$, définis par $\mathcal{F}_T^{\Sigma,W} = (\mathrm{Vect}_{\mathbb{C}} \Sigma) \otimes \left(\oplus_{0 \leq k < T} \mathrm{Sym}^k(W) \right)$.

Fixons une base $(\partial_1, \ldots, \partial_d)$ de W. Alors une base de $\mathrm{Sym}(W)$ est donnée par les $\partial^\sigma = \partial_1^{\sigma_1} \cdot \ldots \cdot \partial_d^{\sigma_d}$ pour $\sigma \in \mathbb{N}^d$. Donc une base de $\mathcal{F} = \mathcal{F}^{\Gamma,W}$ est donnée par les fonctionnelles $\mathrm{ev}_{\gamma,\sigma} = \gamma \otimes \partial^\sigma$ pour $\gamma \in \Gamma$ et $\sigma \in \mathbb{N}^d$. En particulier, $\mathcal{F}$ est isomorphe à la $\mathbb{C}$-algèbre du semi-groupe $\Gamma \times \mathbb{N}^d$ (vue comme espace vectoriel).

On peut alors exprimer l'action de $\mathcal{F}$ sur $\mathbb{C}[G]$ à l'aide de Δ_1, …, Δ_d. En effet, pour tout $\gamma \in \Gamma$ il existe (d'après l'hypothèse (7.1)) un unique $\widehat{\gamma} = (1, \widehat{\gamma}_1, \ldots, \widehat{\gamma}_N) \in \mathbb{C}^{N+1}$ tel que $[1 : \widehat{\gamma}_1 : \ldots : \widehat{\gamma}_N] = \gamma$. Pour $\eta = \sum_{\gamma,\sigma} \lambda_{\gamma,\sigma} \mathrm{ev}_{\gamma,\sigma} \in \mathcal{F}$ et $P \in \mathbb{C}[G]_D^h$ on a :

$$\eta(P) = \sum_{\gamma,\sigma} \lambda_{\gamma,\sigma} (\underline{\Delta}^\sigma P)(\widehat{\gamma}) = \sum_{\gamma,\sigma} \lambda_{\gamma,\sigma} \frac{\underline{\Delta}^\sigma P}{X_0^{D+|\sigma|\kappa}}(\gamma) \tag{7.3}$$

Dans toute la suite, D, S et T représentent des nombres réels (pas forcément entiers) avec $D \geq 0$ et $S, T \geq 1$. On note $\mathbb{C}[G]_D^h$ l'espace des polynômes homogènes sur G, de degré $[D]$ (la partie entière de D), et on pose $\mathcal{F}_{S,T} = \mathcal{F}_{\lceil S \rceil, \lceil T \rceil}$, de telle sorte que l'équation (7.2) soit vérifiée même quand S et T ne sont pas entiers (ici $\lceil x \rceil$ est le plus petit entier supérieur ou égal à x).

L'intérêt des fonctionnelles dans ce contexte est mis en lumière par le lemme suivant :

Lemme 1.2 *Soient $D \geq 0$ et $S, T \geq 1$ des nombres réels. Les deux assertions suivantes sont équivalentes :*

- *Quels que soient les nombres complexes $a_{\gamma,\sigma}$, indexés par $\gamma \in \Gamma(S)$ et $\sigma \in \mathbb{N}^d$, $|\sigma| < T$, il existe $P \in \mathbb{C}[G]_D^h$ tel que $\frac{\underline{\Delta}^\sigma P}{X_0^{D+|\sigma|\kappa}}(\gamma) = a_{\gamma,\sigma}$ pour tous ces γ, σ.*
- *Il n'existe pas de fonctionnelle $\eta \in \mathcal{F}_{S,T}$ non nulle telle que $\eta(P) = 0$ pour tout $P \in \mathbb{C}[G]_D^h$.*

Démonstration : Considérons l'application linéaire de $\mathbb{C}[G]_D^h$ dans $\mathbb{C}^{\Gamma(S) \times \{\sigma \in \mathbb{N}^d, |\sigma| < T\}}$ qui à tout polynôme P associe la famille des $\frac{\underline{\Delta}^\sigma P}{X_0^{D+|\sigma|\kappa}}(\gamma)$. La première assertion est que cette application est surjective, la seconde est que son image n'est contenue dans aucun hyperplan.

On utilise ce lemme dans toute la suite : pour démontrer le théorème 1.1, on suppose qu'il existe une fonctionnelle $\eta \in \mathcal{F}_{S,T}$ non nulle qui s'annule (identiquement[2]) sur $\mathbb{C}[G]^h_D$. On en déduit l'existence d'un sous-groupe obstructeur H.

N.B. Pour $P \in \mathbb{C}[G]^h_D$ et $k \in \mathbb{N}$, on a $\eta(P) = \eta(X_0^k P)$ pour tout $\eta \in \mathcal{F}$. En particulier, si $\eta \in \mathcal{F}_{S,T}$ s'annule sur $\mathbb{C}[G]^h_D$ alors pour tous $S' \geq S$, $T' \geq T$ et $D' \leq D$ on a $\eta \in \mathcal{F}_{S',T'}$ et η s'annule sur $\mathbb{C}[G]^h_{D'}$. Par ailleurs la seconde assertion du lemme 1.2 ne dépend pas du choix d'une base de W, donc on peut choisir n'importe quelle base pour énoncer le théorème 1.1.

Notons a et b des entiers tels qu'il existe un recouvrement ouvert de G tel que, sur chacun des ouverts, l'addition de G (plongé dans $\mathbb{P}^N$) soit donnée par des polynômes bihomogènes de bidegré (a, b). Notons $E_0(\underline{X}, \underline{Y}), \dots, E_N(\underline{X}, \underline{Y})$ de tels polynômes bihomogènes, qui définissent l'addition sur un ouvert contenant Γ. On note $\underline{X} = (X_0, \dots, X_N)$, $\underline{Y} = (Y_0, \dots, Y_N)$ et $\underline{E} = (E_0, \dots, E_N)$. Par hypothèse, pour $\gamma, \delta \in \Gamma$ le point $[E_0(\widehat{\gamma}, \widehat{\delta}) : \dots : E_N(\widehat{\gamma}, \widehat{\delta})]$ est $\gamma +_G \delta$, donc l'hypothèse (7.1) donne :

$$E_0(\widehat{\gamma}, \widehat{\delta}) \neq 0 \text{ pour tous } \gamma, \delta \in \Gamma \subset G \subset \mathbb{P}^N \tag{7.4}$$

N.B. En fait il suffit de faire l'hypothèse (7.1) pour un nombre fini de γ, ceux qui interviennent dans les démonstrations ci-dessous. De même, il suffit de supposer que $\underline{E}$ définit l'addition sur un ouvert qui contient un certain nombre (fini) de $\gamma \in \Gamma$.

2 Opérations sur les fonctionnelles

Dans cette partie, on montre comment projeter une fonctionnelle sur un quotient (§ 2.1), et comment translater une fonctionnelle ou la dériver (§ 2.2). Enfin, au paragraphe 2.3 on relie les constructions effectuées aux paragraphes 2.1 et 2.2, et on fait quelques remarques complémentaires (qui ne seront pas utilisées dans la preuve du lemme d'interpolation).

La difficulté que posent les fonctionnelles dans ce contexte est qu'elles s'appliquent à des polynômes homogènes, et pas à de "vraies" fonctions sur G. L'action de $\mathcal{F}$ sur $\mathbb{C}[G]$ dépend du choix de la forme linéaire X_0, puisque l'évaluation de $P \in \mathbb{C}[G]^h_D$ en $\gamma \in \Gamma$ est $\frac{P}{X_0^D}(\gamma)$.

Considérons le cas où G est un groupe algébrique commutatif affine, et notons $\mathbb{C}[G]$ l'algèbre des fonctions régulières sur G. Soit H un sous-groupe algébrique de G. La projection $\pi : G \to G/H$ envoie Γ dans $\overline{\Gamma} = \Gamma/\Gamma \cap H$ et W dans $\overline{W} = W/W \cap TH$, donc elle induit une application linéaire $\pi_{\mathcal{F}}$ de $\mathcal{F} = \mathcal{F}^{\Gamma, W}$ dans $\overline{\mathcal{F}} = \mathcal{F}^{\overline{\Gamma}, \overline{W}}$. Elle induit aussi une application injective $i : \mathbb{C}[G/H] \to \mathbb{C}[G]$, qui est la composition par π. Par construction, on a $\eta(i(P)) = \pi_{\mathcal{F}}(\eta)(P)$ pour tous $P \in \mathbb{C}[G/H]$ et $\eta \in \mathcal{F}$: les applications i et $\pi_{\mathcal{F}}$ sont adjointes, relativement aux produits bilinéaires $\mathcal{F} \times \mathbb{C}[G] \to \mathbb{C}$ et $\overline{\mathcal{F}} \times \mathbb{C}[G/H] \to \mathbb{C}$.

Quand le groupe algébrique G n'est plus supposé affine, il faut travailler avec des polynômes homogènes et la situation est plus compliquée. L'application $\iota_{D'} : \mathbb{C}[G/H]^h_{D'} \to \mathbb{C}[G]^h_{\delta_H D'}$ dépend du choix d'une famille $p = (p_0, \dots, p_M)$ de polynômes homogènes, de même degré δ_H, qui représente π sur un ouvert contenant Γ. On voit que $\iota_{D'}$ et $\pi_{\mathcal{F}}$ ne sont pas adjoints en général. Par exemple, pour $P \in \mathbb{C}[G/H]^h_{D'}$ et $\eta = \gamma \otimes 1 \in \mathcal{F}$ on a $\pi_{\mathcal{F}}(\eta)(P) = \frac{P(\underline{Y})}{Y_0^{D'}} \circ \pi(\gamma) = \frac{P(p(\underline{X}))}{p_0^{D'}(\underline{X})}(\gamma) = \frac{X_0^{\delta_H}}{p_0(\underline{X})}(\gamma)^{D'} \, \eta(\iota_{D'}(P))$. Au paragraphe 2.1, on construit un adjoint pour $\iota_{D'}$. Au paragraphe 2.2, on construit un adjoint pour l'application qui consiste à translater un polynôme homogène de degré D' par un élément de Γ, et à le dériver.

[2] Dans la suite, on dira simplement que η s'annule sur $\mathbb{C}[G]^h_D$.

2.1 Fonctionnelle quotient

Soit H un sous-groupe algébrique (pas nécessairement connexe) de G. Le quotient G/H est un groupe algébrique, donc admet un plongement dans un espace projectif $\mathbb{P}^M$. Relativement à ce plongement, la projection $\pi : G \to G/H$ est donnée (sur un ouvert qui contient Γ) par des polynômes homogènes p_0, ..., p_M de même degré, qu'on note δ_H. Ceci signifie que pour $[x_0 : \ldots : x_N]$ dans un ouvert de G qui contient Γ, les polynômes p_0, ..., p_M ne s'annulent pas tous en $(x_0 : \ldots : x_N)$, et que le point $[p_0(x_0, \ldots, x_N) : \ldots : p_M(x_0, \ldots, x_N)]$ de G/H est la classe, modulo H, de $[x_0 : \ldots : x_N]$.

Notons $p(\underline{X}) = (p_0(\underline{X}), \ldots, p_M(\underline{X}))$, et $\underline{Y} = (Y_0, \ldots, Y_M)$ les coordonnées sur $\mathbb{P}^M$. Supposons que $\pi(\Gamma)$ est contenu dans l'ouvert $\{Y_0 \neq 0\}$, c'est-à-dire qu'on a :

$$p_0(\widehat{\gamma}) \neq 0 \text{ pour tout } \gamma \in \Gamma \tag{7.5}$$

Soit $P(\underline{Y})$ un polynôme homogène sur G/H. On note $\iota P = P \circ p$ le polynôme sur G induit par P :

$$\iota P(\underline{X}) = P(p_0(\underline{X}), \ldots, p_M(\underline{X}))$$

Alors ιP est un polynôme homogène sur G (en les variables $\underline{X} = X_0, \ldots, X_N$) de degré $\delta_H \deg(P)$.

Notons $\overline{W}$ l'image de W dans TG/TH. Alors $\overline{W}$ est canoniquement isomorphe à $W/(W \cap TH)$ et à $(W + TH)/TH$. Pour $i \in \{1, \ldots, d\}$, notons $\overline{\partial}_i$ la projection de ∂_i. Les dérivations $\overline{\partial}_i$, pour $i \in \{1, \ldots, d\}$, forment une famille génératrice de $\overline{W}$, mais en général cette famille n'est pas libre. On a aussi des dérivations $\overline{\Delta}_i$ qui agissent sur $\mathbb{C}[G/H]^h$, de manière analogue à Δ_i. On pose $\overline{\underline{\Delta}}^\sigma = \overline{\Delta}_1^{\sigma_1} \ldots \overline{\Delta}_d^{\sigma_d}$ pour $\sigma \in \mathbb{N}^d$.

Lemme 2.1 *Pour $\sigma \in \mathbb{N}^d$ et $\gamma \in \Gamma$ on a :*

$$(\underline{\Delta}^\sigma \iota P)(\widehat{\gamma}) = \sum_{\tilde{\sigma} \leq \sigma} \binom{\sigma}{\tilde{\sigma}} \left(\underline{\Delta}^{\sigma - \tilde{\sigma}} p_0^{\deg(P)}\right)(\widehat{\gamma}) \left(\overline{\underline{\Delta}}^{\tilde{\sigma}} P\right)(\widehat{\overline{\gamma}})$$

où $\overline{\gamma}$ est le projeté de γ dans G/H, et $\widehat{\overline{\gamma}} \in \mathbb{C}^{M+1}$ est une famille de coordonnées de $\overline{\gamma}$, la première valant 1 (ce qui est possible d'après (7.5)*).*

N.B. La conclusion de ce lemme s'écrit aussi, la constante $\overline{\kappa}$ étant l'analogue, pour G/H et les $\overline{\Delta}_i$, de la constante κ :

$$\frac{\underline{\Delta}^\sigma \iota P(\underline{X})}{X_0^{\delta_H \deg(P) + |\sigma|\kappa}}(\gamma) = \sum_{\tilde{\sigma} \leq \sigma} \binom{\sigma}{\tilde{\sigma}} \frac{\underline{\Delta}^{\sigma - \tilde{\sigma}} p_0^{\deg(P)}(\underline{X})}{X_0^{\delta_H \deg(P) + |\sigma - \tilde{\sigma}|\kappa}}(\gamma) \times \frac{\overline{\underline{\Delta}}^{\tilde{\sigma}} P(\underline{Y})}{Y_0^{\deg(P) + |\tilde{\sigma}|\overline{\kappa}}}(\overline{\gamma})$$

Démonstration : On a :

$$\frac{\iota P(\underline{X})}{X_0^{\delta_H \deg(P)}} = \frac{P(p(\underline{X}))}{p_0(\underline{X})^{\deg(P)}} \times \frac{p_0(\underline{X})^{\deg(P)}}{X_0^{\delta_H \deg(P)}}$$

comme fonctions analytiques sur l'ouvert de G où $X_0 \neq 0$ et $p_0(\underline{X}) \neq 0$. On applique $\partial_1^{\sigma_1} \ldots \partial_d^{\sigma_d}$ et on évalue en γ (en utilisant la relation de Leibniz). La fonction $\frac{P(p(\underline{X}))}{p_0(\underline{X})^{\deg(P)}}$ s'écrit $\frac{P(\underline{Y})}{Y_0^{\deg(P)}} \circ \pi$, donc pour $i \in \{1, \ldots, d\}$ on a

$$\partial_i \left(\frac{P(p(\underline{X}))}{p_0(\underline{X})^{\deg(P)}}\right) = \left(\overline{\partial}_i \frac{P(\underline{Y})}{Y_0^{\deg(P)}}\right) \circ \pi$$

Ceci termine la démonstration du lemme.

On note $\overline{\Gamma}$ l'image de Γ dans G/H, et $\overline{\gamma}_1, \ldots, \overline{\gamma}_l$ les images des générateurs $\gamma_1, \ldots, \gamma_l$. Alors pour tout $S \geq 0$ l'ensemble $\overline{\Gamma}(S)$ (relatif au sous-groupe $\overline{\Gamma}$ de G/H et à ses générateurs $\overline{\gamma}_1, \ldots, \overline{\gamma}_l$) est l'image dans G/H de $\Gamma(S)$.

On note $\overline{\mathcal{F}} = \mathbb{C}\overline{\Gamma} \otimes \mathrm{Sym}(\overline{W})$ l'espace vectoriel bi-filtré des fonctionnelles sur G/H obtenues par combinaisons linéaires d'évaluations de dérivées (le long de $\overline{W}$) en des points de $\overline{\Gamma}$. On a une projection naturelle $\pi_{\mathcal{F}} : \mathcal{F} \to \overline{\mathcal{F}}$, qui envoie $\mathcal{F}_{S,T}$ dans $\overline{\mathcal{F}}_{S,T}$.

Dans la suite, on aura besoin, étant donné un polynôme homogène $R \in \mathbb{C}[G]$, de considérer (pour $\eta \in \mathcal{F}$) l'application qui à $P \in \mathbb{C}[G/H]^h_{D'}$ associe $\eta(R \cdot (P \circ p))$. C'est ce qui motive la définition suivante (voir la proposition 2.2).

Définition : Soient $R \in \mathbb{C}[G]$ un polynôme homogène et D' un entier. Notons $\mathcal{P}_{R,D'}$ l'application linéaire de $\mathcal{F}$ dans $\overline{\mathcal{F}}$ définie, pour tous $\gamma \in \Gamma$ et $\sigma \in \mathbb{N}^d$, par :

$$\mathcal{P}_{R,D'}\mathrm{ev}_{\gamma,\sigma} = \pi_{\mathcal{F}}\left(\sum_{\nu+\mu+\tilde{\sigma}=\sigma} \begin{pmatrix} & \sigma & \\ \tilde{\sigma} & \nu & \mu \end{pmatrix} \left(\underline{\Delta}^{\mu} p_0^{D'}\right)(\widehat{\gamma})\ (\underline{\Delta}^{\nu} R)(\widehat{\gamma})\ \mathrm{ev}_{\gamma,\tilde{\sigma}}\right)$$

Avec cette définition, on a visiblement $\mathcal{P}_{R,D'}\mathcal{F}_{S,T} \subset \overline{\mathcal{F}}_{S,T}$ pour tous S, T.

Proposition 2.2 *Soient $P \in \mathbb{C}[G/H]^h$ un polynôme homogène de degré D', $R \in \mathbb{C}[G]$ un polynôme homogène et $\eta \in \mathcal{F}$. On pose $\iota P(\underline{X}) = P(p_0(\underline{X}), \ldots, p_M(\underline{X}))$. Alors :*

$$\eta(R \cdot \iota P) = \mathcal{P}_{R,D'}\eta(P)$$

Démonstration : Par linéarité on peut supposer $\eta = \mathrm{ev}_{\gamma,\sigma}$. On a alors :

$$\eta(R \cdot \iota P) = \underline{\Delta}^{\sigma}(R \cdot \iota P)(\widehat{\gamma}) = \sum_{\nu \leq \sigma} \binom{\sigma}{\nu} (\underline{\Delta}^{\sigma-\nu}\iota P)(\widehat{\gamma})(\underline{\Delta}^{\nu} R)(\widehat{\gamma})$$

d'où le résultat en utilisant le lemme 2.1.

N.B. En prenant $R = 1$, on obtient $\eta(\iota P) = \mathcal{P}_{1,D'}\eta(P)$ donc $\mathcal{P}_{1,D'} : \mathcal{F} \to \mathcal{F}$ et $\iota : \mathbb{C}[G/H]^h_{D'} \to \mathbb{C}[G]^h_{\delta_H D'}$ sont adjoints.

Supposons maintenant (sans perte de généralité) que la base de W soit choisie de telle sorte que $\partial_{\bar{d}+1}, \ldots, \partial_d$ forment une base de $W \cap TH$. Alors $\overline{\partial}_1, \ldots, \overline{\partial}_{\bar{d}}$ forment une base de $\overline{W}$, et $\bar{d} = \dim(\overline{W})$; on identifie $\mathbb{N}^d$ à $\mathbb{N}^{\bar{d}} \times \mathbb{N}^{d-\bar{d}}$. En outre $\pi_{\mathcal{F}}(\mathrm{ev}_{\gamma,\sigma})$ est nul, sauf si $\sigma \in \mathbb{N}^{\bar{d}} \times \{0\}^{d-\bar{d}}$, et dans ce cas c'est exactement l'élément $\overline{\mathrm{ev}}_{\pi(\gamma),\sigma}$ de la base de $\overline{\mathcal{F}}$ obtenue à partir de $\overline{\Gamma}$ et de $\overline{\partial}_1, \ldots, \overline{\partial}_{\bar{d}}$.

Lemme 2.3 *Soit $\eta = \sum_{\gamma,\sigma} \lambda_{\gamma,\sigma}\mathrm{ev}_{\gamma,\sigma} \in \mathcal{F}$ une fonctionnelle non nulle. Soit $\sigma_0 \in \mathbb{N}^{\bar{d}}$ de longueur maximale tel qu'il existe $\gamma_0 \in \Gamma$ et $\sigma_1 \in \mathbb{N}^{d-\bar{d}}$ avec $\lambda_{\gamma_0,(\sigma_0,\sigma_1)} \neq 0$. Soit D' un entier. Considérons la fonctionnelle non nulle :*

$$\eta^{\gamma_0,\sigma_0,D'} = \sum_{\substack{\gamma \in \Gamma \\ \gamma \equiv \gamma_0 \mod H}} \sum_{\nu,\mu \in \mathbb{N}^{d-\bar{d}}} \binom{\nu+\mu}{\nu} \underline{\Delta}^{(0,\mu)} p_0^{D'}(\widehat{\gamma})\ \lambda_{\gamma,(\sigma_0,\nu+\mu)}\ \mathrm{ev}_{\gamma,(0,\nu)}$$

Soit R un polynôme homogène sur G tel que $\mathcal{P}_{R,D'}\eta$ soit la fonctionnelle nulle. Alors on a $\eta^{\gamma_0,\sigma_0,D'}(R) = 0$.

Démonstration : Supposons $\mathcal{P}_{R,D'}\eta = 0$. Comme $\pi_{\mathcal{F}}(\mathrm{ev}_{\gamma,(\sigma_0,\sigma_1)}) = 0$ dès que $\sigma_1 \neq 0$, la composante sur $\overline{\mathrm{ev}}_{\pi(\gamma_0),\sigma_0}$ de $\mathcal{P}_{R,D'}\eta$ donne :

$$\sum_{\substack{\gamma\,\in\,\Gamma\\ \gamma\,\equiv\,\gamma_0 \mod H}} \sum_{\nu,\mu\in\mathbb{N}^d} \begin{pmatrix} (\sigma_0,0)+\nu+\mu \\ (\sigma_0,0) \quad \nu \quad \mu \end{pmatrix} \underline{\Delta}^{\mu} p_0^{D'}(\widehat{\gamma})\ \underline{\Delta}^{\nu} R(\widehat{\gamma})\ \lambda_{\gamma,(\sigma_0,0)+\nu+\mu} = 0$$

Or $\lambda_{\gamma,(\sigma_0,0)+\nu+\mu}$ est nul, sauf si $\nu,\mu \in \{0\}\times\mathbb{N}^{d-\bar{d}}$, d'après le choix de σ_0. Ceci signifie exactement $\eta^{\gamma_0,\sigma_0,D'}(R) = 0$. Il reste à montrer que la fonctionnelle $\eta^{\gamma_0,\sigma_0,D'}$ est non nulle. Soit $\sigma_1 \in \mathbb{N}^{d-\bar{d}}$ de longueur maximale tel que $\lambda_{\gamma_0,(\sigma_0,\sigma_1)} \neq 0$. Alors la coordonnée de $\eta^{\gamma_0,\sigma_0,D'}$ sur $\mathrm{ev}_{\gamma_0,(0,\sigma_1)}$ est $p_0^{D'}(\widehat{\gamma})\lambda_{\gamma,(\sigma_0,\sigma_1)}$, qui est non nul d'après (7.5).

Dans la suite, on utilisera le lemme 2.3 seulement à travers le corollaire suivant :

Corollaire 2.4 *Soient D, D', D'' trois entiers tels que $D = \delta_H D' + D''$. Soient $S,T \geq 1$, et $\eta \in \mathcal{F}_{S,T}\setminus\{0\}$ qui s'annule identiquement sur $\mathbb{C}[G]^h_D$. Alors :*
- *Ou bien il existe $\eta_1 \in \overline{\mathcal{F}}_{S,T}\setminus\{0\}$ qui s'annule identiquement sur $\mathbb{C}[G/H]^h_{D'}$.*
- *Ou bien il existe $\overline{\gamma} \in \overline{\Gamma}(S)$ et $\eta_2 \in \mathcal{F}_T^{\Gamma(S)\cap\pi^{-1}(\overline{\gamma}),W\cap TH}\setminus\{0\}$ qui s'annule identiquement sur $\mathbb{C}[G]^h_{D''}$.*

Démonstration : Par hypothèse, il existe $\eta \in \mathcal{F}_{S,T}$ non nulle telle que $\eta(R\cdot\iota P) = \mathcal{P}_{R,D'}\eta(P) = 0$ pour tous $R \in \mathbb{C}[G]^h_{D''}$ et $P \in \mathbb{C}[G/H]^h_{D'}$. Le premier cas est celui où $\mathcal{P}_{R,D'}\eta$ est non nulle pour un certain R ; le second provient du lemme 2.3.

On peut donner une deuxième démonstration de ce corollaire, en traduisant toutes les assertions grâce au lemme 1.2. On suppose toujours que $\partial_{\bar{d}+1},\dots,\partial_d$ forment une base de $W\cap TH$. On suppose que η_1 et η_2 n'existent pas, et on veut démontrer que η ne peut pas exister (en étant non nulle). Une variante du lemme 1.2 montre qu'on aura atteint cet objectif si, pour tous $\gamma \in \Gamma(S)$ et $\sigma \in \mathbb{N}^d$, $|\sigma| < T$, on construit $P_{\gamma,\sigma} \in \mathbb{C}[G]^h_D$ qui vérifie les propriétés suivantes (où on note $\overline{\gamma} = \pi(\gamma)$ et $\sigma = (\sigma_0,\sigma_1) \in \mathbb{N}^d = \mathbb{N}^{\bar{d}}\times\mathbb{N}^{d-\bar{d}}$, avec la convention analogue pour $\tilde{\sigma}$) :
- $P_{\gamma,\sigma}$ s'annule à l'ordre T le long de W en tout point de $\Gamma(S)\setminus\pi^{-1}(\overline{\gamma})$.
- Pour tout $\gamma' \in \Gamma(S)$ congru à γ modulo H, mais distinct de γ, on a $\underline{\Delta}^{\tilde{\sigma}}P_{\gamma,\sigma}(\widehat{\gamma'}) = 0$ pour tout $\tilde{\sigma} \in \mathbb{N}^d$ tel que $|\tilde{\sigma}| < T$ et $|\tilde{\sigma}_0| \leq |\sigma_0|$.
- On a $\underline{\Delta}^{\sigma}P_{\gamma,\sigma}(\widehat{\gamma}) \neq 0$ et $\underline{\Delta}^{\tilde{\sigma}}P_{\gamma,\sigma}(\widehat{\gamma}) = 0$ pour tout $\tilde{\sigma} \in \mathbb{N}^d$ tel que $|\tilde{\sigma}_0| \leq |\sigma_0|$, $|\tilde{\sigma}_1| \leq |\sigma_1|$ et $\tilde{\sigma} \neq \sigma$.

Soient $\gamma \in \Gamma(S)$ et $\sigma \in \mathbb{N}^d$, $|\sigma| < T$. Par hypothèse, il n'existe aucune fonctionnelle $\eta_1 \in \overline{\mathcal{F}}_{S,T}$ non nulle qui s'annule identiquement sur $\mathbb{C}[G/H]^h_{D'}$. On en déduit l'existence de $A_{\gamma,\sigma} \in \mathbb{C}[G/H]^h_{D'}$ qui s'annule à l'ordre T le long de $\overline{W}$ en tout point de $\overline{\Gamma}(S)\setminus\{\overline{\gamma}\}$, et vérifie $\overline{\underline{\Delta}}^{\sigma_0}A_{\gamma,\sigma}(\widehat{\overline{\gamma}}) \neq 0$ et $\overline{\underline{\Delta}}^{\tilde{\sigma}_0}A_{\gamma,\sigma}(\widehat{\overline{\gamma}}) = 0$ pour tout $\tilde{\sigma}_0 \in \mathbb{N}^{\bar{d}}$ tel que $|\tilde{\sigma}_0| \leq |\sigma_0|$ et $\tilde{\sigma}_0 \neq \sigma_0$. Alors $B_{\gamma,\sigma} = A_{\gamma,\sigma}\circ p \in \mathbb{C}[G]^h_{\delta_H D'}$ s'annule à l'ordre T le long de W en tout point de $\Gamma(S)\setminus\pi^{-1}(\overline{\gamma})$, et vérifie, pour tout $\gamma' \in \Gamma(S)\cap\pi^{-1}(\overline{\gamma})$, les relations $\underline{\Delta}^{(\sigma_0,0)}B_{\gamma,\sigma}(\widehat{\gamma'}) \neq 0$ et $\underline{\Delta}^{\tilde{\sigma}}B_{\gamma,\sigma}(\widehat{\gamma'}) = 0$ pour tout $\tilde{\sigma} \in \mathbb{N}^d$ tel que $|\tilde{\sigma}_0| \leq |\sigma_0|$ et $\tilde{\sigma}_0 \neq \sigma_0$. Or, par hypothèse, il n'existe aucune fonctionnelle $\eta_2 \in \mathcal{F}_T^{\Gamma(S)\cap\pi^{-1}(\overline{\gamma}),W\cap TH}$ non nulle qui s'annule identiquement sur $\mathbb{C}[G]^h_{D''}$. On en déduit l'existence de $C_{\gamma,\sigma} \in \mathbb{C}[G]^h_{D''}$ qui s'annule à l'ordre T le long de $W\cap TH$ en tout point de $\Gamma(S)\cap\pi^{-1}(\overline{\gamma})$ distinct de γ, et vérifie $\underline{\Delta}^{(0,\sigma_1)}C_{\gamma,\sigma}(\widehat{\gamma}) \neq 0$ et $\underline{\Delta}^{(0,\tilde{\sigma}_1)}C_{\gamma,\sigma}(\widehat{\gamma}) = 0$ dès que $|\tilde{\sigma}_1| < T$ et $\tilde{\sigma}_1 \neq \sigma_1$. Alors le polynôme $P_{\gamma,\sigma} = B_{\gamma,\sigma}C_{\gamma,\sigma}$ a les propriétés voulues.

2.2 Translation et dérivation d'une fonctionnelle

Considérons le groupe algébrique $G \times G$. On a une injection de $W \times W$ dans $TG \times TG$ (qui s'identifie canoniquement à l'espace tangent en l'identité à $G \times G$). En outre $\Gamma \times \Gamma$ est un sous-groupe de type fini de $G \times G$. La base $(\partial_1, \dots, \partial_d)$ de W fournit $2d$ dérivations linéairement indépendantes, le long de $W \times W$: on note $\partial_{\underline{X},i}$ et $\Delta_{\underline{X},i}$ celles le long de $W \times \{0\}$, et $\partial_{\underline{Y},i}$ et $\Delta_{\underline{Y},i}$ celles le long de $\{0\} \times W$, pour $i \in \{1, \dots, d\}$. Avec des coordonnées $(\underline{X}, \underline{Y}) = (X_0, \dots, X_N, Y_0, \dots, Y_N)$ sur $G \times G$, $\Delta_{\underline{X},i}$ agit seulement sur les variables $\underline{X}$.

Soit $P \in \mathbb{C}[G]$ un polynôme homogène de degré D'. Soient $\tau \in \mathbb{N}^d$ et $\delta \in \Gamma$. On considère :

$$t_{\delta,\tau} P(\underline{X}) = \underline{\Delta}_{\underline{Y}}^{\tau}(P(\underline{E}(\underline{X}, \underline{Y})))_{|\underline{Y}=\widehat{\delta}} = \frac{\underline{\Delta}_{\underline{Y}}^{\tau}(P(\underline{E}(\underline{X}, \underline{Y})))}{Y_0^{bD'+|\tau|\kappa}}(\delta) \tag{7.6}$$

où $\underline{\Delta}_{\underline{Y}}^{\tau} = \Delta_{\underline{Y},1}^{\tau_1} \dots \Delta_{\underline{Y},d}^{\tau_d}$.

Le polynôme $P(\underline{E}(\underline{X}, \underline{Y}))$ est bihomogène, de bidegré (aD', bD'). L'action de $\underline{\Delta}_{\underline{Y}}^{\tau}$ conserve le degré en $\underline{X}$, donc $t_{\delta,\tau} P(\underline{X})$ est homogène de degré aD'.

Le polynôme $t_{\delta,\tau} P(\underline{X})$ est une variante de la dérivée $\underline{\Delta}^{\tau} P$ de P, translatée par δ. L'intérêt de la définition (7.6) est que $\deg(t_{\delta,\tau} P) = aD'$ alors que le degré de $\underline{\Delta}^{\tau} P$ croît avec $|\tau|$ (c'est l'astuce dite de Baker-Coates-Anderson). Le lemme suivant montre que $\underline{\Delta}^{\sigma} t_{\delta,\tau} P(\widehat{\gamma})$ est essentiellement $\underline{\Delta}^{\sigma+\tau} P(\widehat{\gamma +_G \delta})$, à un facteur multiplicatif non nul et à des dérivées d'ordre inférieur près.

Lemme 2.5 *Pour $\sigma \in \mathbb{N}^d$ et $\gamma \in \Gamma$ on a :*

$$(\underline{\Delta}^{\sigma} t_{\delta,\tau} P)(\widehat{\gamma}) = \sum_{\substack{\tilde{\sigma} \le \sigma \\ \tilde{\tau} \le \tau}} \binom{\sigma}{\tilde{\sigma}} \binom{\tau}{\tilde{\tau}} \left(\underline{\Delta}_{\underline{X}}^{\sigma-\tilde{\sigma}} \underline{\Delta}_{\underline{Y}}^{\tau-\tilde{\tau}} E_0^{D'} \right)(\widehat{\gamma}, \widehat{\delta}) \left(\underline{\Delta}^{\tilde{\sigma}+\tilde{\tau}} P \right)(\widehat{\gamma +_G \delta})$$

N.B. Cette équation s'écrit aussi :

$$\frac{\underline{\Delta}^{\sigma} t_{\delta,\tau} P}{X_0^{aD'+|\sigma|\kappa}}(\gamma) = \sum_{\substack{\tilde{\sigma} \le \sigma \\ \tilde{\tau} \le \tau}} \binom{\sigma}{\tilde{\sigma}} \binom{\tau}{\tilde{\tau}} \frac{\underline{\Delta}_{\underline{X}}^{\sigma-\tilde{\sigma}} \underline{\Delta}_{\underline{Y}}^{\tau-\tilde{\tau}} E_0^{D'}}{X_0^{aD'+|\sigma-\tilde{\sigma}|\kappa} Y_0^{bD'+|\tau-\tilde{\tau}|\kappa}}(\gamma, \delta) \frac{\underline{\Delta}^{\tilde{\sigma}+\tilde{\tau}} P}{X_0^{D'+|\tilde{\sigma}+\tilde{\tau}|\kappa}}(\gamma +_G \delta)$$

Démonstration : On a

$$\frac{P(\underline{E}(\underline{X}, \underline{Y}))}{X_0^{aD'} Y_0^{bD'}}(g, d) = \frac{P(Z)}{Z_0^{D'}}(g +_G d) \frac{E_0(\underline{X}, \underline{Y})^{D'}}{X_0^{aD'} Y_0^{bD'}}(g, d)$$

comme fonctions de (g, d), analytiques sur l'ouvert de $G \times G$ défini par $X_0 \neq 0$, $Y_0 \neq 0$ et $E_0(\underline{X}, \underline{Y}) \neq 0$. On applique $\partial_{\underline{X},1}^{\sigma_1} \dots \partial_{\underline{X},d}^{\sigma_d} \partial_{\underline{Y},1}^{\tau_1} \dots \partial_{\underline{Y},d}^{\tau_d}$ aux deux membres, et on utilise la relation de Leibniz. On obtient le résultat, car les dérivations ∂_i commutent aux translations.

Ce lemme incite à poser la définition suivante.

Définition : Soient $\delta \in \Gamma$, $\tau \in \mathbb{N}^d$ et D' un entier. On note $\mathcal{T}_{\delta,\tau,D'}$ l'application linéaire de $\mathcal{F}$ dans lui-même définie, pour tous $\gamma \in \Gamma$ et $\sigma \in \mathbb{N}^d$, par :

$$\mathcal{T}_{\delta,\tau,D'} \mathrm{ev}_{\gamma,\sigma} = \sum_{\sigma_1 \le \sigma+\tau} \nu_{\sigma_1} \, \mathrm{ev}_{\gamma+\delta,\sigma_1}$$

avec

$$\nu_{\sigma_1} = \sum_{\substack{0 \leq \tilde{\sigma} \leq \sigma \\ 0 \leq \tilde{\tau} \leq \tau \\ \tilde{\sigma} + \tilde{\tau} = \sigma_1}} \binom{\sigma}{\tilde{\sigma}} \binom{\tau}{\tilde{\tau}} \left(\underline{\Delta}_{\underline{X}}^{\sigma - \tilde{\sigma}} \underline{\Delta}_{\underline{Y}}^{\tau - \tilde{\tau}} E_0^{D'} \right) (\widehat{\gamma}, \widehat{\delta})$$

N.B. En particulier la coordonnée de $\mathcal{T}_{\delta,\tau,D'}\mathrm{ev}_{\gamma,\sigma}$ sur $\mathrm{ev}_{\gamma+\delta,\sigma+\tau}$ est $E_0^{D'}(\widehat{\gamma}, \widehat{\delta})$, qui est non nul d'après (7.4).

On a alors la proposition suivante, qui démontre que $\mathcal{T}_{\delta,\tau,D'} : \mathcal{F} \to \mathcal{F}$ et $t_{\delta,\tau} : \mathbb{C}[G]_{D'}^h \to \mathbb{C}[G]_{aD'}^h$ sont adjoints.

Proposition 2.6 *L'opérateur $\mathcal{T}_{\delta,\tau,D'}$ est injectif, et envoie $\mathcal{F}_T^{\Gamma(S),W}$ dans $\mathcal{F}_{T+|\tau|}^{\delta+\Gamma(S),W}$; en particulier si $\delta \in \Gamma(S')$ alors $\mathcal{T}_{\delta,\tau,D'}\mathcal{F}_{S,T}^{\Gamma,W} \subset \mathcal{F}_{S+S',T+|\tau|}^{\Gamma,W}$. En outre, pour tout $P \in \mathbb{C}[G]_{D'}^h$ on a*

$$\eta(t_{\delta,\tau}P) = (\mathcal{T}_{\delta,\tau,D'}\eta)(P)$$

où $t_{\delta,\tau}P(\underline{X})$ est défini par l'équation (7.6).

Démonstration : Soit $\eta \neq 0$. Notons T' le plus petit entier tel que $\eta \in \mathcal{F}_{S,T'}$ pour un certain S. Soient $\gamma \in \Gamma(S)$ et $\sigma \in \mathbb{N}^d$ tels que $|\sigma| = T' - 1$ et $\lambda_{\gamma,\sigma} \neq 0$. La coordonnée de $\mathcal{T}_{\delta,\tau,D'}\eta$ sur $\mathrm{ev}_{\gamma+\delta,\sigma+\tau}$ est $\lambda_{\gamma,\sigma}E_0(\widehat{\gamma}, \widehat{\delta})^{D'}$ donc non nulle d'après (7.4) : la fonctionnelle $\mathcal{T}_{\delta,\tau,D'}\eta$ est non nulle. En outre, la relation $\mathcal{T}_{\delta,\tau,D'}\mathcal{F}_T^{\Gamma(S),W} \subset \mathcal{F}_{T+|\tau|}^{\delta+\Gamma(S),W}$ est immédiate, et $\eta(t_{\delta,\tau}P) = \mathcal{T}_{\delta,\tau,D'}\eta(P)$ découle du lemme 2.5 et de la définition de $\mathcal{T}_{\delta,\tau,D'}$. Ceci termine la démonstration de la proposition 2.6.

Lemme 2.7 *Supposons Γ sans torsion. Soient $\eta \in \mathcal{F}$ une fonctionnelle non nulle, et D' un entier. Alors les fonctionnelles $\mathcal{T}_{\delta,\tau,D'}\eta$, pour $\delta \in \Gamma$ et $\tau \in \mathbb{N}^d$, sont linéairement indépendantes.*

Démonstration : Il s'agit de montrer que toute combinaison linéaire non triviale des opérateurs $\mathcal{T}_{\delta,\tau,D'}$ (avec D' fixé) est injective. Soient $\eta = \sum_{\gamma,\sigma} \lambda_{\gamma,\sigma}\mathrm{ev}_{\gamma,\sigma} \in \mathcal{F} \setminus \{0\}$ et $\mathcal{T} = \sum_{\delta,\tau} \mu_{\delta,\tau}\mathcal{T}_{\delta,\tau,D'}$, avec les $\mu_{\delta,\tau}$ non tous nuls.

Comme Γ est un $\mathbb{Z}$-module de type fini sans torsion, il est libre, donc il existe un ordre total sur $\Gamma \times \mathbb{N}^d$ compatible à l'addition. Notons (γ_0, σ_0) le plus grand élément, pour cet ordre, tel que $\lambda_{\gamma_0,\sigma_0} \neq 0$, et (δ_0, τ_0) le plus grand tel que $\mu_{\delta_0,\tau_0} \neq 0$. Alors la coordonnée de $\mathcal{T}\eta$ sur $\mathrm{ev}_{\gamma_0+\delta_0,\sigma_0+\tau_0}$ est $\lambda_{\gamma_0,\sigma_0}\mu_{\delta_0,\tau_0}E_0^{D'}(\widehat{\gamma_0}, \widehat{\delta_0}) \neq 0$, donc $\mathcal{T}\eta \neq 0$. Ceci termine la démonstration du lemme.

2.3 Quelques remarques sur les fonctionnelles

Soient G_1, Γ_1 et W_1 comme au paragraphe 1.1. Soit $\pi : G_1 \to G_2$ un morphisme de groupes algébriques, où G_2 est un groupe algébrique commutatif plongé dans $\mathbb{P}^{N_2}$. Posons $\Gamma_2 = \pi(\Gamma_1)$ et $W_2 = T\pi(W_1)$, où $T\pi$ est l'application linéaire tangente à π en l'identité. Notons $p = (p_0, \ldots, p_{N_2})$ une famille de polynômes homogènes de même degré δ_π qui définit π sur un ouvert contenant Γ_1. La composition par p définit $\iota_{D'} : \mathbb{C}[G_2]_{D'}^h \to \mathbb{C}[G_1]_{\delta_\pi D'}^h$ pour tout entier D'. De même qu'au paragraphe 2.1, on peut définir $\mathcal{P}_{D'} : \mathcal{F}^{\Gamma_1,W_1} \to \mathcal{F}^{\Gamma_2,W_2}$ qui soit adjoint à $\iota_{D'}$, c'est-à-dire tel que $\eta(P \circ p) = \mathcal{P}_{D'}\eta(P)$ pour tous $P \in \mathbb{C}[G_2]_{D'}^h$ et $\eta \in \mathcal{F}^{\Gamma_1,W_1}$.

Considérons le cas particulier où $G_1 = G \times G$, $G_2 = G$ et π est l'addition de G. L'argument ci-dessus se généralise au cas où G_1 est plongé dans un produit d'espaces projectifs, en l'occurrence $\mathbb{P}^N \times \mathbb{P}^N$. Avec $\Gamma_1 = \Gamma \times \Gamma$ et $W_1 = W \times W$, on a $\mathcal{F}^{\Gamma_1, W_1} = \mathcal{F}^{\Gamma, W} \otimes \mathcal{F}^{\Gamma, W}$, et l'application $\mathcal{P}_{D'}$ (liée au choix de $(E_0, \dots, E_N)$) va de $\mathcal{F} \otimes \mathcal{F}$ dans $\mathcal{F}$. Elle définit un produit bilinéaire sur $\mathcal{F}$, noté $\star_{D'}$ (qui dépend de l'entier D'). Ce produit est une “perturbation” du produit naturel sur l'algèbre associative $\mathcal{F} = \mathbb{C}\Gamma \otimes \mathrm{Sym}(W)$. A priori le produit $\star_{D'}$ n'est pas associatif ; il est commutatif si, et seulement si, le polynôme E_0 est symétrique en $\underline{X}$ et $\underline{Y}$. Quand G est affine, contenu dans l'ouvert $\{X_0 \neq 0\}$ et qu'on a $E_0(\underline{X}, \underline{Y}) = 1$, le produit $\star_{D'}$ ne dépend pas de D', et c'est exactement le produit de l'algèbre associative $\mathbb{C}\Gamma \otimes \mathrm{Sym}(W)$, i.e.

$$\Big(\gamma \otimes (\delta^{(1)} \cdot \dots \cdot \delta^{(k)})\Big) \star_{D'} \Big(\gamma' \otimes ({\delta'}^{(1)} \cdot \dots \cdot {\delta'}^{(k')})\Big) = (\gamma +_G \gamma') \otimes (\delta^{(1)} \cdot \dots \cdot \delta^{(k)} \cdot {\delta'}^{(1)} \cdot \dots \cdot {\delta'}^{(k')})$$

Ce produit $\star_{D'}$ est relié aux opérateurs $\mathcal{T}_{\delta, \tau, D'}$ du paragraphe 2.2 de la manière suivante : pour tout $\eta \in \mathcal{F}$ on a

$$(\delta \otimes \partial^\tau) \star_{D'} \eta = \mathcal{T}_{\delta, \tau, D'} \eta$$

Ainsi, $\mathcal{T}_{\delta, \tau, D'}$ est la multiplication par $\delta \otimes \partial^\tau$ pour le produit $\star_{D'}$. Le lemme 2.7 affirme que si Γ est sans torsion alors pour $\eta, \eta' \in \mathcal{F} \setminus \{0\}$ on a $\eta \star_{D'} \eta' \neq 0$. Dans ce cas, $\mathcal{F}$ est isomorphe (comme espace vectoriel) à une algèbre de polynômes $\mathbb{C}[z_1, \dots, z_r]$. Le produit $\star_{D'}$ transporté sur $\mathbb{C}[z_1, \dots, z_r]$ est de la forme suivante. Si $P = \sum_{|\mu| \leq D_P} p_\mu z^\mu$ et $Q = \sum_{|\nu| \leq D_Q} q_\nu z^\nu$ alors $P \star_{D'} Q = \sum_{|\chi| \leq D_P + D_Q} s_\chi z^\chi$, et pour $|\mu| = D_P$ et $|\nu| = D_Q$ tels que $p_\mu \neq 0$ et $q_\nu \neq 0$ on a $s_{\mu+\nu} \neq 0$. C'est ce qui permet de démontrer le lemme 2.7.

3 Répartition de (Γ, W)

Soient G, Γ et W comme au paragraphe 1.1. On note toujours $d = \dim(W)$ et $n = \dim(G)$; dans cette partie on pose en outre $r = \mathrm{rg}(\Gamma)$.

Sauf mention explicite du contraire, on n'a pas besoin de choisir une famille génératrice de Γ ou une base de W. En outre les constructions et les résultats présentés dans cette partie sont inchangés quand on quotiente G par un sous-groupe algébrique fini (en prenant l'image de Γ dans le quotient), ou quand on remplace Γ par un sous-groupe d'indice fini de Γ.

3.1 Obstructions à l'interpolation et au lemme de zéros

On appelle *obstruction à l'interpolation* (relative aux paramètres S, T, D) tout sous-groupe algébrique connexe H non nul de G tel que

$$S^{\mathrm{rg}(\Gamma \cap H)} T^{\dim(W \cap TH)} > D^{\dim(H)}$$

Il existe une constante $c > 0$, qui dépend seulement de G, Γ, W et du choix d'une famille génératrice de Γ, ayant la propriété suivante. Si H est une obstruction à l'interpolation relative à (S, T, D), alors il existe $\eta \in \mathcal{F}_{S,T}$ non nulle qui s'annule identiquement sur $\mathbb{C}[G]^h_{cD}$. En effet, on peut supposer que H appartient à un ensemble fini de sous-groupes de G, qui dépend seulement de G, Γ et W.

Réciproquement, le théorème 1.1 donne une constante $\overline{c}$ telle que, si il n'y a aucune obstruction à l'interpolation relative à (S, T, D), alors aucune fonctionnelle $\eta \in \mathcal{F}_{S,T}$ non nulle ne s'annule identiquement sur $\mathbb{C}[G]^h_{\overline{c}D}$.

De même, une *obstruction au lemme de zéros* (relative à S, T, D) est un sous-groupe algébrique connexe H de G, distinct de G, tel que

$$S^{\mathrm{rg}(\Gamma/\Gamma\cap H)} T^{\dim(W/W\cap TH)} < D^{\dim(G/H)}$$

Il existe une constante $c' > 0$, qui dépend seulement de G, Γ, W et du choix d'une famille génératrice de Γ, telle que si H est une obstruction au lemme de zéros relative à (S, T, D) alors il existe $P \in \mathbb{C}[G]^h_{c'D}$ non nul qui s'annule en tout point de $\Gamma(S)$ à l'ordre T le long de W. Réciproquement, le lemme de zéros fournit une constante $\overline{c'}$ ayant la propriété suivante : si il n'y a aucune obstruction au lemme de zéros relative à (S, T, D) alors aucun polynôme $P \in \mathbb{C}[G]^{h}_{\underline{c'}D}$ non nul ne s'annule à l'ordre T le long de W en tout point de $\Gamma(S)$.

3.2 Exposants de répartition

L'exposant de répartition ([M3], [W1]) qui intervient dans le lemme de zéros sans multiplicités est :

$$\mu(\Gamma, G) = \min_{H \neq G} \frac{\mathrm{rg}(\Gamma/\Gamma \cap H)}{\dim(G/H)}$$

où le minimum porte sur les sous-groupes algébriques H de G, distincts de G, qu'on peut supposer connexes.

Masser définit [M2] un autre exposant de répartition, qui apparaît dans son lemme d'interpolation :

$$\mu^*(\Gamma, G) = \max_{H \neq 0} \frac{\mathrm{rg}(\Gamma \cap H)}{\dim(H)}$$

où le maximum porte sur les sous-groupes algébriques connexes H, non nuls, de G. Dans la suite de cette partie, tous les sous-groupes H considérés sont supposés connexes.

On a toujours $\mu(\Gamma, G) \leq \frac{r}{n} \leq \mu^*(\Gamma, G)$ (pour le voir, il suffit de prendre $H = 0$ dans la définition de μ, et $H = G$ dans celle de μ^*). En outre, pour tout H distinct de 0 et de G on a $\frac{\mathrm{rg}(\Gamma/\Gamma\cap H)}{\dim(G/H)} < \frac{r}{n}$ si, et seulement si, $\frac{\mathrm{rg}(\Gamma\cap H)}{\dim(H)} > \frac{r}{n}$. Donc $\mu(\Gamma, G) = \frac{r}{n}$ équivaut à $\mu^*(\Gamma, G) = \frac{r}{n}$. Quand c'est le cas, on dit que Γ est *bien réparti* dans G.

On peut reprendre les définitions précédentes en remplaçant Γ par W. Précisément, on note

$$\mu(W, G) = \min_{H \neq G} \frac{\dim(W/W \cap TH)}{\dim(G/H)}$$

et

$$\mu^*(W, G) = \max_{H \neq 0} \frac{\dim(W \cap TH)}{\dim(H)}$$

où le minimum et le maximum portent toujours sur des sous-groupes algébriques connexes de G, dont seul l'espace tangent en l'identité intervient.

On a de même $\mu(W, G) \leq \frac{d}{n} \leq \mu^*(W, G)$, et si l'une des inégalités est une égalité alors l'autre aussi. Dans ce cas, on dit que W est *bien réparti* dans G (ou dans TG).

L'intérêt de l'exposant de répartition $\mu^*(\Gamma, G)$ est le suivant. On considère une situation sans[3] multiplicités, i.e. $T = 1$ ou $W = \{0\}$. Soit H_0 un sous-groupe de G qui réalise le

[3] On peut traiter de la même manière le cas où T est strictement plus grand que 1, mais majoré par une constante qui dépend seulement de Γ, W, γ_1, ..., γ_l.

maximum dans la définition de $\mu^*(\Gamma, G)$. Alors H_0 a la propriété suivante : si S et D sont des paramètres, et H une obstruction à l'interpolation relative à (S, D), alors H_0 est aussi une obstruction à l'interpolation relative à (S, D). Ceci signifie qu'on peut conclure à la présence ou à l'absence d'obstruction en considérant seulement H_0. On a bien sûr l'analogue pour l'obstruction au lemme de zéros avec $\mu(\Gamma, G)$, et on peut faire de même avec $\mu(W, G)$ et $\mu^*(W, G)$ quand il n'y a qu'un nombre fixé de points (i.e. $\mathrm{rg}(\Gamma) = 0$ ou S majoré par une constante).

Dans la suite de cette partie, on généralise la notion d'exposant de répartition pour tenir compte des cas où on considère à la fois plusieurs points et des dérivées. Au paragraphe 3.3, on construit des sous-groupes algébriques $H_1, \ldots, H_{p-1}$ de G tels que si S, T, D sont des paramètres pour lesquels il existe une obstruction à l'interpolation, alors l'un des H_i est une telle obstruction (voir la proposition 3.1). Au paragraphe 3.4, on construit des sous-groupes analogues $\widetilde{H}_j$ pour le lemme de zéros, et on étudie leur lien avec $H_1, \ldots, H_{p-1}$.

3.3 Construction et propriétés du polygone d'interpolation

Dans ce paragraphe, on suppose $\mathrm{rg}(\Gamma) \geq 1$ et $\dim(W) \geq 1$ (sauf à la fin, où le cas particulier $\mathrm{rg}(\Gamma) = 0$ ou $\dim(W) = 0$ est traité séparément).

Définitions

Le fait qu'un sous-groupe H soit une obstruction à l'interpolation dépend de S, T, D de manière homogène au sens suivant : si S, T et D sont remplacés par S^α, T^α et D^α pour un même réel $\alpha > 0$, cela ne change pas la nature de H (i.e. le fait qu'il soit ou non une obstruction à l'interpolation).

Pour exploiter cette homogénéité, on considère $x = \frac{\log(S)}{\log(D)}$ et $y = \frac{\log(T)}{\log(D)}$, et on se place dans le plan affine avec des coordonnées (x, y) (en supposant $D > 1$). On pourrait aussi se placer dans l'espace projectif $\mathbb{P}(\mathbb{R}^3)$, ce qui permettrait de garder une situation homogène (avec $\log(S)$, $\log(T)$ et $\log(D)$).

Dans ce paragraphe, H désigne toujours un sous-groupe algébrique connexe de G. À tout H on associe l'ensemble Δ_H d'équation $\varphi_H(x, y) = 0$, où

$$\varphi_H(x, y) = x \,\mathrm{rg}(\Gamma \cap H) + y \dim(W \cap T H) - \dim(H)$$

On a $\Delta_0 = \mathbb{R}^2$, et Δ_H est vide si, et seulement si, $\mathrm{rg}(\Gamma \cap H) = \dim(W \cap T H) = 0$ et $H \neq 0$. Dans les autres cas, Δ_H est une droite.

Pour tout H, on a $\varphi_H(x, y) > 0$ si, et seulement si, H est une obstruction à l'interpolation pour (S, T, D), avec $x = \frac{\log(S)}{\log(D)}$ et $y = \frac{\log(T)}{\log(D)}$. Ceci ne peut arriver que si Δ_H est une droite (sinon φ_H est négative ou nulle sur $\mathbb{R}^2$). Pour tout H on pose

$$x_H = \frac{\dim(H)}{\mathrm{rg}(\Gamma \cap H)} \text{ et } y_H = \frac{\dim(H)}{\dim(W \cap T H)}$$

à condition que les dénominateurs soient non nuls. L'ensemble des points (x, y) de $\mathbb{R}_+^2$ pour lesquels H n'est pas une obstruction à l'interpolation, i.e. tels que $\varphi_H(x, y) \leq 0$, est :

1. $\mathbb{R}_+^2$ tout entier si $\mathrm{rg}(\Gamma \cap H) = \dim(W \cap T H) = 0$.
2. La bande horizontale $\mathbb{R}_+ \times [0, y_H]$ si $\mathrm{rg}(\Gamma \cap H) = 0$ et $\dim(W \cap T H) \neq 0$.

3. La bande verticale $[0, x_H] \times \mathbb{R}_+$ si $\mathrm{rg}(\Gamma \cap H) \neq 0$ et $\dim(W \cap TH) = 0$.
4. Le triangle de sommets $O = (0,0)$, $(x_H, 0)$ et $(0, y_H)$ si $\mathrm{rg}(\Gamma \cap H) \neq 0$ et $\dim(W \cap TH) \neq 0$.

En outre, pour $H = G$ on est dans le quatrième cas (car on a supposé $\mathrm{rg}(\Gamma) > 0$ et $\dim(W) > 0$). Donc l'intersection de tous ces domaines est un polygone, qu'on appelle le *polygone d'interpolation* de (Γ, W). C'est l'ensemble des (x, y) qui correspondent à des triplets (S, T, D) pour lesquels il n'y aucune obstruction à l'interpolation.

Le polygone d'interpolation est convexe, comme intersection de convexes. Il s'agit d'ailleurs d'une intersection finie, car les fonctions φ_H sont en nombre fini (même si il peut y avoir une infinité de sous-groupes algébriques H).

Notons $OA_1A_2 \dots A_p$ le polygone d'interpolation, où O est l'origine, les sommets sont parcourus dans le sens trigonométrique, et trois sommets consécutifs ne sont jamais alignés (voir la figure 7.1, page 160). L'entier p est supérieur ou égal à deux.

Pour tout $i \in \{1, \dots, p-1\}$, soit H_i un sous-groupe de G tel que la droite (A_iA_{i+1}) soit Δ_{H_i}. Il peut exister plusieurs tels sous-groupes H_i ; dans ce cas, on en choisit un de dimension maximale. Par définition même, Δ_{H_i} est une droite donc H_i est non nul.

Notons χ_i la pente de Δ_{H_i}. On a $\chi_i = \frac{-\mathrm{rg}(\Gamma \cap H_i)}{\dim(W \cap TH_i)} \in [-\infty, 0]$. En outre on a $\chi_1 < \chi_2 < \dots < \chi_{p-1}$ par convexité du polygone d'interpolation ; en particulier seul χ_1 peut être infini, seul χ_{p-1} peut être nul, et les sous-groupes H_1, …, H_{p-1} sont deux à deux distincts. En outre on a $A_1 = (x_{H_1}, 0)$ et $A_{p-1} = (0, y_{H_{p-1}})$.

Pour $i \in \{1, \dots, p-1\}$, notons $\mathcal{C}_i$ le secteur angulaire fermé délimité dans $\mathbb{R}_+^2$ par les demi-droites (OA_i) et (OA_{i+1}). Alors les cônes $\mathcal{C}_1$, …, $\mathcal{C}_{p-1}$ forment une partition de $\mathbb{R}_+^2$ (à des parties de mesure nulle près).

Caractère obstructeur des sous-groupes $H_1, \dots, H_{p-1}$

L'intérêt des sous-groupes $H_1, \dots, H_{p-1}$ et des cônes $\mathcal{C}_1$, …, $\mathcal{C}_{p-1}$ est mis en évidence par la proposition suivante :

Proposition 3.1 *Soient $S, T \geq 1$ et $D > 1$ des réels tels que le point $(\frac{\log(S)}{\log(D)}, \frac{\log(T)}{\log(D)})$ appartienne au cône $\mathcal{C}_i$. Soient $c > 0$ et H un sous-groupe algébrique non nul de G tels que*

$$S^{\mathrm{rg}(\Gamma \cap H)} T^{\dim(W \cap TH)} > (cD)^{\dim(H)}$$

Alors on a :

$$S^{\mathrm{rg}(\Gamma \cap H_i)} T^{\dim(W \cap TH_i)} > (cD)^{\dim(H_i)}$$

En prenant $c = 1$, on voit que si il existe une obstruction à l'interpolation pour S, T, D alors H_i est une telle obstruction. En outre, le fait de pouvoir prendre c quelconque montre que dans l'énoncé du théorème 1.1 il suffit de supposer $S^{\mathrm{rg}(\Gamma \cap H)} T^{\dim(W \cap TH)} < c_1 D^{\dim(H)}$ quand $H = H_i$ (quitte à modifier la valeur de c_1).

Démonstration de la proposition 3.1 : Si $cD < 1$, il n'y a rien à démontrer. Si $cD \geq 1$, en augmentant légèrement cD si nécessaire on peut supposer $cD > 1$. Alors le point $(\frac{\log(S)}{\log(cD)}, \frac{\log(T)}{\log(cD)})$ appartient à $\mathcal{C}_i$ (par homothétie). Par hypothèse, φ_H y est strictement positive, donc ce point n'appartient pas au polygone d'interpolation. Ce polygone étant délimité, dans $\mathcal{C}_i$, par la droite Δ_{H_i}, on en déduit que φ_{H_i} est strictement positive en ce point : c'est exactement la conclusion souhaitée.

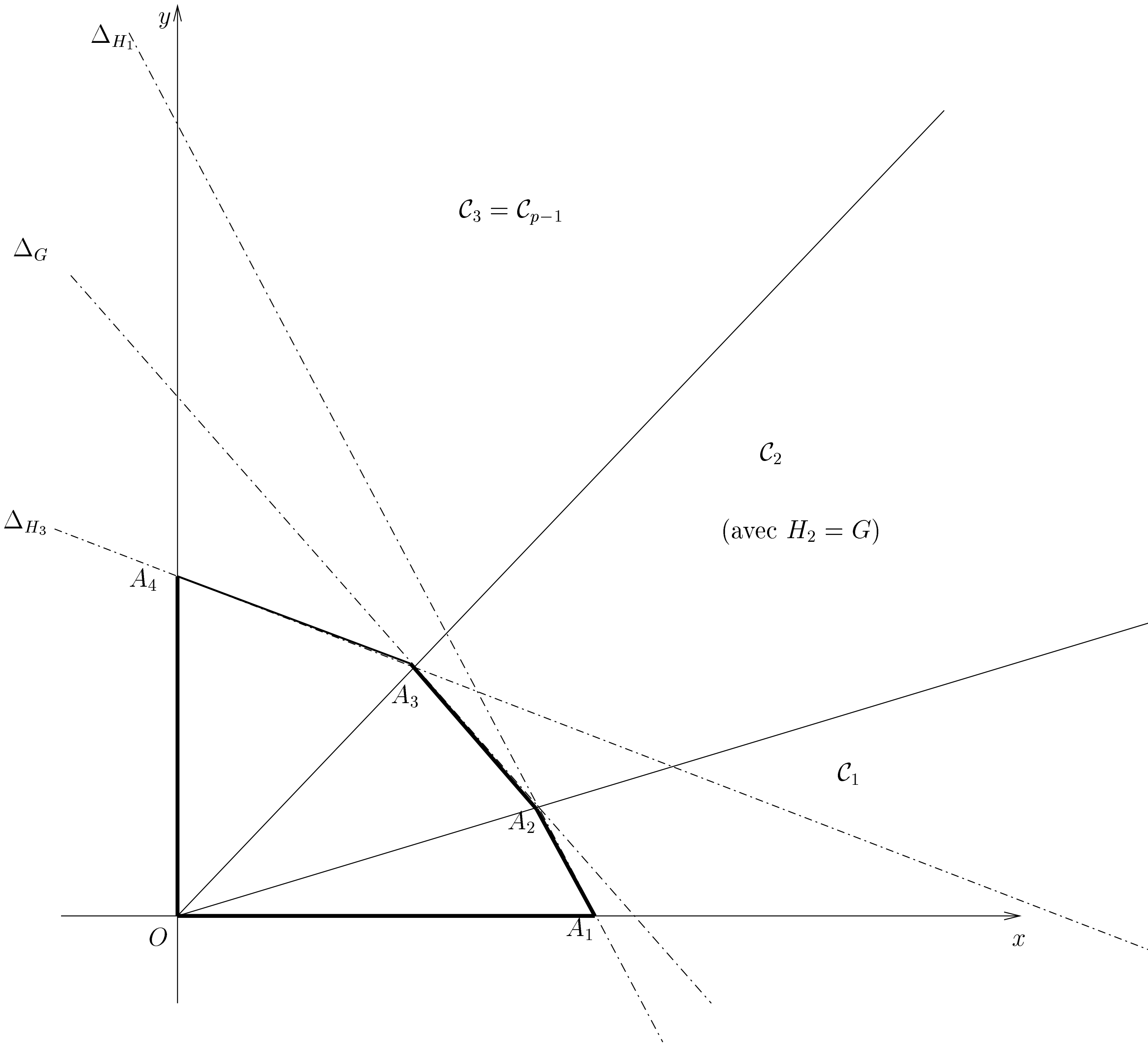

FIG. 7.1: Décomposition $\mathbb{R}_+^2 = \mathcal{C}_1 \cup \ldots \cup \mathcal{C}_{p-1}$ qui vient de l'interpolation

On aura aussi besoin du lemme suivant, à travers la proposition 3.3 ci-dessous qui en découle :

Lemme 3.2 *Soit (x,y) un point de $\mathbb{R}_+^2$, qui appartient au cône $\mathcal{C}_i$. Pour tout sous-groupe H non nul de G on a :*

$$\frac{\varphi_H(x,y)}{\dim(H)} \leq \frac{\varphi_{H_i}(x,y)}{\dim(H_i)}$$

En outre, si H contient strictement H_i on a

$$\frac{(\varphi_H - \varphi_{H_i})(x,y)}{\dim(H/H_i)} \leq \frac{\varphi_H(x,y)}{\dim(H)} \leq \frac{\varphi_{H_i}(x,y)}{\dim(H_i)}$$

et si H est contenu strictement dans H_i on a

$$\frac{\varphi_H(x,y)}{\dim(H)} \leq \frac{\varphi_{H_i}(x,y)}{\dim(H_i)} \leq \frac{(\varphi_{H_i} - \varphi_H)(x,y)}{\dim(H_i/H)}$$

Démonstration : Si $x = y = 0$ le lemme est évident. Sinon on paramètre la demi-droite joignant l'origine au point (x,y) en posant $x(t) = tx$ et $y(t) = ty$ pour $t \in \mathbb{R}_+$. Alors la fonction $\dim(H_i)\varphi_H(x(t),y(t)) - \dim(H)\varphi_{H_i}(x(t),y(t))$ est nulle pour $t = 0$ (qui correspond à l'origine $(0,0)$), et négative (ou nulle) pour le paramètre $t > 0$ qui correspond à l'intersection de cette demi-droite avec Δ_{H_i} (car ce point appartient au polygone d'interpolation). Comme c'est une fonction affine, elle prend des valeurs négatives en tous les réels $t \geq 0$, en particulier en $t = 1$: la première inégalité du lemme est démontrée. Le reste en découle (en utilisant le fait 3.10 ci-dessous).

N.B. L'hypothèse que H contient strictement H_i pourrait être remplacée par $\dim(H) > \dim(H_i)$, en notant $\dim(H) - \dim(H_i)$ au lieu de $\dim(H/H_i)$.

En traduisant la première assertion du lemme 3.2, on retrouve la proposition 3.1 ; en traduisant la deuxième, on obtient l'énoncé suivant (qui sera utilisé dans la preuve du théorème 1.1) :

Proposition 3.3 *Soient $S, T \geq 1$ et $D > 1$ des réels tels que le point $(\frac{\log(S)}{\log(D)}, \frac{\log(T)}{\log(D)})$ appartienne au cône $\mathcal{C}_i$. Soient H un sous-groupe algébrique de G qui contient strictement H_i, et c un réel strictement positif, tels que*

$$S^{\mathrm{rg}(\Gamma\cap H/\Gamma\cap H_i)} T^{\dim(W\cap TH/W\cap TH_i)} > (cD)^{\dim(H/H_i)}$$

Alors on a :

$$S^{\mathrm{rg}(\Gamma\cap H_i)} T^{\dim(W\cap TH_i)} > (cD)^{\dim(H_i)}$$

Démonstration : De même que dans la preuve de la proposition 3.1, on peut supposer $cD > 1$. Par hypothèse, la fonction $\varphi_H - \varphi_{H_i}$ est strictement positive au point $(\frac{\log(S)}{\log(cD)}, \frac{\log(T)}{\log(cD)})$ de $\mathcal{C}_i$. Le lemme 3.2 montre que φ_{H_i} l'est aussi.

Coordonnées des sommets du polygone d'interpolation

Les points $A_1 = (x_{H_1}, 0)$ et $A_{p-1} = (0, y_{H_{p-1}})$ sont reliés aux exposants de répartition de la manière suivante.

Proposition 3.4 *On a $\mu^*(\Gamma, G) = \frac{1}{x_{H_1}} = \frac{\mathrm{rg}(\Gamma\cap H_1)}{\dim(H_1)}$ et $\mu^*(W, G) = \frac{1}{y_{H_{p-1}}} = \frac{\dim(W\cap TH_{p-1})}{\dim(H_{p-1})}$.*

Démonstration : Le maximum que représente $\mu^*(\Gamma, G)$ est l'inverse du minimum des x_H, qui est atteint pour $H = H_1$. De même pour $\mu^*(W, G)$.

Concernant les autres sommets du polygone d'interpolation, on aura besoin des encadrements suivants :

Proposition 3.5 *Notons (x_i, y_i) les coordonnées du point A_i. Pour tout $i \in \{2, \dots, p-1\}$ on a :*

$$\frac{1}{dr} \leq x_i \leq \frac{n}{r} \text{ et } \frac{1}{dr} \leq y_i \leq \frac{n}{d}$$

Démonstration : Les majorations proviennent de la proposition 3.4 et des minorations évidentes des exposants μ^*, puisque le polygone d'interpolation est inclus dans le rectangle dont trois sommets consécutifs sont $A_1 O A_p$ (ceci provient du fait que les pentes des droites $A_i A_{i+1}$ sont négatives). Pour les minorations, on constate que pour $i \in \{2, \dots, p-1\}$ le point A_i est l'intersection des droites $\Delta_{H_{i-1}}$ et Δ_{H_i}. Leurs équations donnent :

$$\begin{aligned} x_i &= \frac{\dim(W \cap TH_i)\dim(H_{i-1}) - \dim(W \cap TH_{i-1})\dim(H_i)}{\dim(W \cap TH_i)\mathrm{rg}(\Gamma \cap H_{i-1}) - \dim(W \cap TH_{i-1})\mathrm{rg}(\Gamma \cap H_i)} \\ \text{et } y_i &= \frac{\mathrm{rg}(\Gamma \cap H_{i-1})\dim(H_i) - \mathrm{rg}(\Gamma \cap H_i)\dim(H_{i-1})}{\dim(W \cap TH_i)\mathrm{rg}(\Gamma \cap H_{i-1}) - \dim(W \cap TH_{i-1})\mathrm{rg}(\Gamma \cap H_i)} \end{aligned}$$

Par hypothèse, ces droites s'intersectent en un point, A_i, de ${\mathbb{R}_+^*}^2$. Les numérateurs de ces fractions sont des entiers non nuls, donc supérieurs à un en valeur absolue ; ceci termine la démonstration de la proposition.

Bonne répartition de (Γ, W)

Quand $(\frac{\log(S)}{\log(D)}, \frac{\log(T)}{\log(D)}) \in \mathcal{C}_{i_0}$ avec $H_{i_0} = G$, on dit que (Γ, W) est *bien réparti relativement à* (S, T, D). En effet, dans ce cas, si il existe une obstruction à l'interpolation (pour les paramètres (S, T, D)) alors G lui-même est une telle obstruction. Le cas où (Γ, W) est bien réparti relativement à tous les choix de paramètres (S, T, D) est le suivant :

Définition : On dit que (Γ, W) est *bien réparti* si $p = 2$ et $H_1 = G$.

Ceci signifie que le polygone d'interpolation est un triangle dont l'arête oblique est portée par Δ_G.

Proposition 3.6 *Le couple (Γ, W) est bien réparti si, et seulement si, Γ et W sont bien répartis.*

Démonstration : D'après la proposition 3.4, si (Γ, W) est bien réparti alors Γ et W le sont. Réciproquement, supposons Γ et W bien répartis. Alors le polygone d'interpolation admet pour sommets $A_1 = (x_G, 0)$ et $A_{p-1} = (0, y_G)$. Par convexité, ce polygone contient le triangle OA_1A_{p-1} ; il est aussi contenu dedans car ce triangle a pour équation $\varphi_G(x, y) \leq 0$ dans $\mathbb{R}_+^2$. Donc le polygone d'interpolation est exactement ce triangle : on a $p = 2$, et H_1 est un sous-groupe tel que $\Delta_{H_1} = \Delta_G$. Il peut y en avoir plusieurs, mais H_1 est de dimension maximale donc $H_1 = G$.

Cas particulier $\mathrm{rg}(\Gamma) = 0$ ou $\dim(W) = 0$

Considérons maintenant le cas particulier où W est réduit à $\{0\}$. Pour chaque H, l'ensemble des points de $\mathbb{R}_+^2$ où φ_H est négative est ou bien $\mathbb{R}_+^2$, ou bien une bande verticale. L'intersection de ces domaines, quand H varie, est la bande $[0, \mu^*(\Gamma, G)] \times \mathbb{R}_+$; cette bande joue le rôle du polygone d'interpolation. On a $p = 2$, et l'unique cône $\mathcal{C}_1 = \mathbb{R}_+^2$ correspond à un sous-groupe H_1 qui réalise le maximum dans la définition de $\mu^*(\Gamma, G)$. En outre, si ce maximum est réalisé par plusieurs sous-groupes, H_1 est choisi de dimension maximale parmi ceux-ci. Ainsi, on a $H_1 = G$ si, et seulement si, Γ est bien réparti dans G ; dans ce cas, on dit que (Γ, W) est bien réparti.

Quand $\mathrm{rg}(\Gamma) = 0$ et $\dim(W) \neq 0$, la situation est analogue.

3.4 Construction et propriétés analogues pour le lemme de zéros

On suppose que si H contient un sous-groupe d'indice fini de Γ et que TH contient W alors $H = G$. En effet, si ce n'est pas le cas, alors H est une obstruction systématique au lemme de zéros, et le lieu d'injectivité défini ci-dessous est vide. Par ailleurs, cette hypothèse n'est pas restrictive, puisque si un tel sous-groupe H existe on peut travailler dans H plutôt que dans G tout entier.

En outre, on suppose $\mathrm{rg}(\Gamma) \geq 1$ et $\dim(W) \geq 1$ (sauf à la fin de ce paragraphe, où le cas particulier $\mathrm{rg}(\Gamma) = 0$ ou $\dim(W) = 0$ est traité séparément).

Définitions

Dans ce paragraphe, H désigne toujours un sous-groupe algébrique connexe de G. Un tel H est une obstruction au lemme de zéros pour (S, T, D) si, et seulement si, on a $(\varphi_G - \varphi_H)(x, y) < 0$, avec $x = \frac{\log(S)}{\log(D)}$ et $y = \frac{\log(T)}{\log(D)}$. En effet, on a $(\varphi_G - \varphi_H)(x, y) = x\,\mathrm{rg}(\Gamma/\Gamma \cap H) + y \dim(W/W \cap TH) - \dim(G/H)$. La fonction $\varphi_G - \varphi_H$ n'est pas constante, sauf si H contient un sous-groupe d'indice fini de Γ et que TH contient W ; par hypothèse, ceci ne se produit que si $H = G$. L'ensemble des points (x, y) de $\mathbb{R}_+^2$ pour lesquels H n'est pas une obstruction au lemme de zéros, i.e. tels que $(\varphi_G - \varphi_H)(x, y) \geq 0$, est (puisque $\mathrm{rg}(\Gamma/\Gamma \cap H) = \dim(W/W \cap TH) = 0$ implique $H = G$) :

1. $\mathbb{R}_+^2$ tout entier si $H = G$.
2. Le quart de plan $\mathbb{R}_+ \times [\frac{\dim(G/H)}{\dim(W/W \cap TH)}, +\infty[$ si $\mathrm{rg}(\Gamma/\Gamma \cap H) = 0$ et $H \neq G$.
3. Le quart de plan $[\frac{\dim(G/H)}{\mathrm{rg}(\Gamma/\Gamma \cap H)}, +\infty[\times \mathbb{R}_+$ si $\dim(W/W \cap TH) = 0$ et $H \neq G$.
4. Le complémentaire, dans $\mathbb{R}_+^2$, du triangle de sommets $O = (0, 0)$, $(\frac{\dim(G/H)}{\mathrm{rg}(\Gamma/\Gamma \cap H)}, 0)$ et $(0, \frac{\dim(G/H)}{\dim(W/W \cap TH)})$ si $\mathrm{rg}(\Gamma/\Gamma \cap H) \neq 0$ et $\dim(W/W \cap TH) \neq 0$.

En outre, pour $H = 0$ on est dans le quatrième cas (car on a supposé $\mathrm{rg}(\Gamma) > 0$ et $\dim(W) > 0$).

On appelle *lieu d'injectivité* l'intersection de ces domaines ; c'est l'ensemble des points (x, y) qui correspondent à des triplets (S, T, D) pour lesquels il n'y a aucune obstruction au lemme de zéros.

Le lieu d'injectivité est un "polygone" infini, de sommets $\infty_x B_1 B_2 \dots B_q \infty_y$, où ∞_x (respectivement ∞_y) est le point à l'infini de la demi-droite horizontale issue de B_1 et contenue dans $\mathbb{R}_+^2$ (respectivement de la droite verticale issue de B_q). Il est convexe (comme intersection de convexes). On a $q \geq 1$; voir la figure 7.2, page 165.

Pour $j \in \{1, \dots, q-1\}$, on choisit $\widetilde{H}_j$ de dimension minimale tel qu'une équation de la droite $(B_j B_{j+1})$ soit $(\varphi_G - \varphi_{\widetilde{H}_j})(x,y) = 0$. La pente $\tilde{\chi}_j$ de cette droite est $\frac{-\mathrm{rg}(\Gamma/\Gamma\cap\widetilde{H}_j)}{\dim(W/W\cap T\widetilde{H}_j)} \in [-\infty, 0]$. Par convexité on a $\tilde{\chi}_1 > \tilde{\chi}_2 > \dots > \tilde{\chi}_{q-1}$, donc les sous-groupes $\widetilde{H}_1, \dots, \widetilde{H}_{q-1}$ sont deux à deux distincts (et distincts de G). En outre toutes ces pentes sont finies et non nulles, car le polygone d'injectivité a déjà un côté vertical et un côté horizontal.

Pour $j \in \{1, \dots, q-1\}$ notons $\widetilde{\mathcal{C}}_j$ le secteur angulaire fermé délimité dans $\mathbb{R}_+^2$ par les demi-droites (OB_j) et (OB_{j+1}).

Deux cas peuvent se présenter concernant le point B_1 :

1. Le point B_1 est sur l'axe (Ox). Ceci équivaut à $\mu(\Gamma, G) \neq 0$, i.e. aucun sous-groupe d'indice fini de Γ n'est contenu dans aucun sous-groupe algébrique H distinct de G. Dans ce cas, on pose $\widetilde{\mathcal{C}}_0 = \emptyset$, et on a $\mu(\Gamma, G) = \frac{\mathrm{rg}(\Gamma/\Gamma\cap\widetilde{H}_1)}{\dim(G/\widetilde{H}_1)}$; l'abscisse de B_1 est l'inverse de ce nombre (de même que dans la proposition 3.4).
2. Le point B_1 n'est pas sur l'axe (Ox), i.e. $\mu(\Gamma, G) = 0$. On note $\widetilde{\mathcal{C}}_0$ le secteur angulaire fermé délimité dans $\mathbb{R}_+^2$ par les demi-droites (Ox) et (OB_1), et $\widetilde{H}_0$ un sous-groupe (de dimension minimale) tel qu'une équation de la droite $(B_1 \infty_x)$ soit $\varphi_G - \varphi_{\widetilde{H}_0} = 0$. Alors $\widetilde{H}_0$ contient un sous-groupe d'indice fini de Γ ; en particulier $\widetilde{H}_0 \neq 0$.

On a une alternative similaire pour B_q :

1. Le point B_q est sur (Oy), c'est-à-dire $\mu(W, G) \neq 0$ (i.e. W n'est inclus dans TH pour aucun sous-groupe algébrique H distinct de G). On pose $\widetilde{\mathcal{C}}_q = \emptyset$, et on a $\mu(W, G) = \frac{\dim(W/W\cap T\widetilde{H}_{q-1})}{\dim(G/\widetilde{H}_{q-1})}$; l'ordonnée de B_q est l'inverse de ce nombre.
2. Le point B_q n'est pas sur (Oy), i.e. $\mu(W, G) = 0$. On note $\widetilde{\mathcal{C}}_q$ le secteur angulaire fermé délimité par (Oy) et (OB_q), et $\widetilde{H}_q$ un sous-groupe (de dimension minimale) tel qu'une équation de $(B_q \infty_y)$ soit $\varphi_G - \varphi_{\widetilde{H}_q} = 0$. Alors $T\widetilde{H}_q$ contient W ; en particulier $\widetilde{H}_q \neq 0$.

Dans tous les cas, les secteurs $\widetilde{\mathcal{C}}_0, \dots, \widetilde{\mathcal{C}}_q$ forment une partition de $\mathbb{R}_+^2$ (à des parties de mesure nulle près). En outre chaque cône $\widetilde{\mathcal{C}}_j$ non vide est fermé, convexe, d'intérieur non vide (pour la topologie habituelle de $\mathbb{R}_+^2$), et correspond à un sous-groupe $\widetilde{H}_j$. Les seuls cônes qui peuvent éventuellement être vides sont $\widetilde{\mathcal{C}}_0$ et $\widetilde{\mathcal{C}}_q$; dans ce cas, on ne leur associe aucun sous-groupe de G.

Caractère obstructeur des sous-groupes $\widetilde{H}_j$

Soit $j \in \{0, \dots, q\}$ tel que $\widetilde{\mathcal{C}}_j$ soit non vide ; le sous-groupe $\widetilde{H}_j$ vérifie la propriété suivante :

Proposition 3.7 *Soient $S, T \geq 1$ et $D > 1$ des réels tels que le point $(\frac{\log(S)}{\log(D)}, \frac{\log(T)}{\log(D)})$ appartienne au cône $\widetilde{\mathcal{C}}_j$. Soient $c > 0$ et H un sous-groupe algébrique de G, distinct de G, tels que*

$$S^{\mathrm{rg}(\Gamma/\Gamma\cap H)} T^{\dim(W/W\cap TH)} < (cD)^{\dim(G/H)}$$

Alors on a :

$$S^{\mathrm{rg}(\Gamma/\Gamma\cap\widetilde{H}_j)} T^{\dim(W/W\cap T\widetilde{H}_j)} < (cD)^{\dim(G/\widetilde{H}_j)}$$

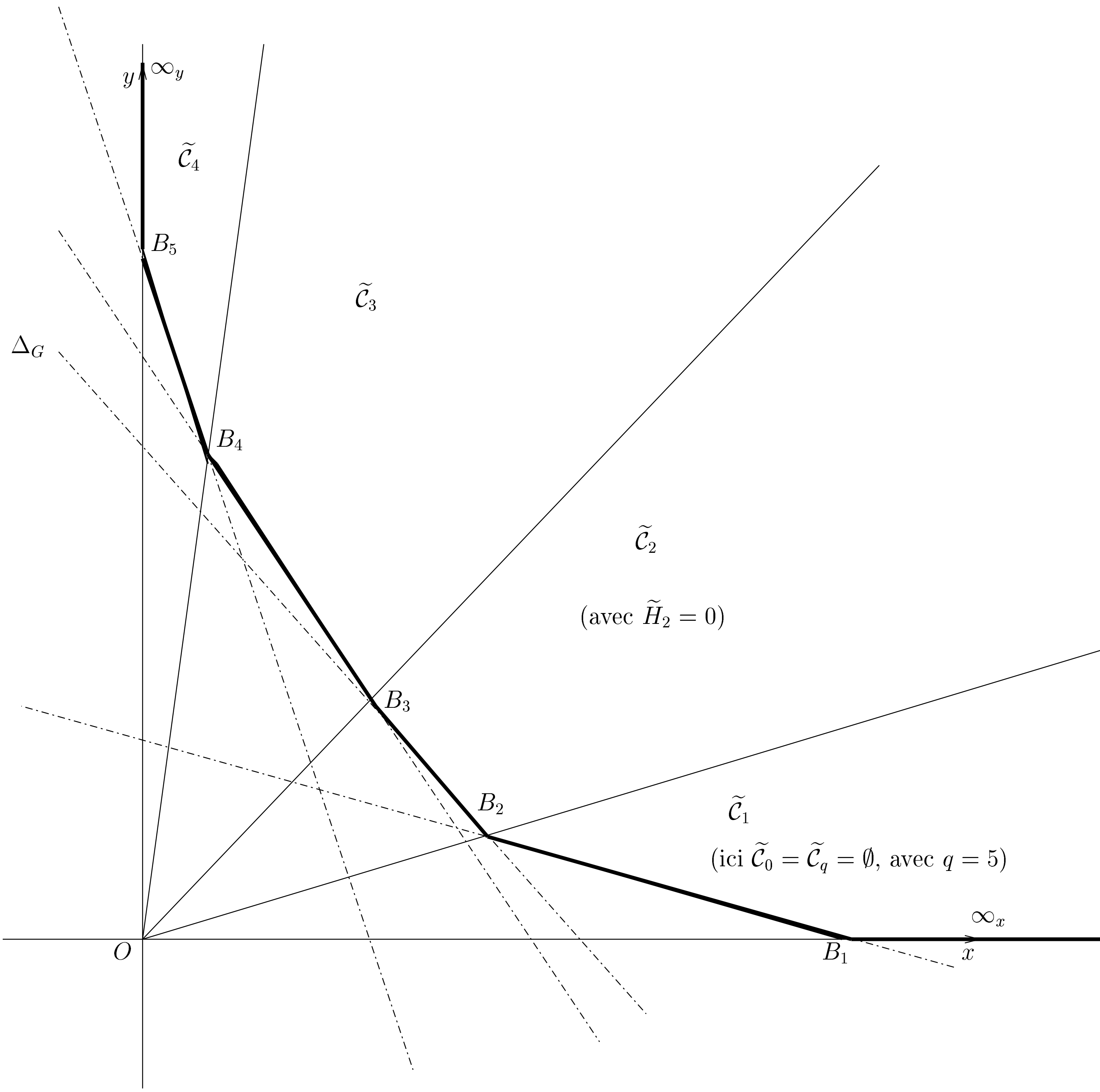

FIG. 7.2: Décomposition $\mathbb{R}^2_+ = \widetilde{\mathcal{C}}_0 \cup \ldots \cup \widetilde{\mathcal{C}}_q$ qui vient du lemme de zéros

Démonstration : Par hypothèse, on a $cD > 1$, et le point $(\frac{\log(S)}{\log(cD)}, \frac{\log(T)}{\log(cD)})$ appartient à $\widetilde{\mathcal{C}}_j$ mais pas au lieu d'injectivité. Donc $\varphi_G - \varphi_{\widetilde{H}_j}$ est strictement négative en ce point, ce qui est la conclusion cherchée.

Ainsi, si il existe une obstruction au lemme de zéros pour S, T, D alors $\widetilde{H}_j$ est une telle obstruction. En outre on peut, dans l'énoncé d'un lemme de zéros, supposer que le sous-groupe obstructeur est $\widetilde{H}_j$ (quitte à changer la constante multiplicative qui y apparaît). C'est ce qu'on fait au paragraphe suivant, dans le cas où $\widetilde{H}_j = 0$.

Un lemme de zéros sous hypothèse de bonne répartition

On dit que (Γ, W) est *bien réparti* pour le lemme de zéros, relativement aux paramètres (S, T, D), si $(\frac{\log(S)}{\log(D)}, \frac{\log(T)}{\log(D)}) \in \widetilde{\mathcal{C}}_{j_0}$ avec $\widetilde{H}_{j_0} = 0$ (d'où nécessairement $j_0 \in \{1, \ldots, q-1\}$). On verra plus loin que c'est le cas si, et seulement si, (Γ, W) est bien réparti pour l'interpolation relativement à (S, T, D) (i.e. $(\frac{\log(S)}{\log(D)}, \frac{\log(T)}{\log(D)}) \in \mathcal{C}_{i_0}$ avec $H_{i_0} = G$). Le lemme de zéros suivant concerne cette situation. On rappelle que, par hypothèse, sur tout ouvert de G suffisamment petit l'addition est donnée par des polynômes bihomogènes de bidegré (a, b).

Théorème 3.8 *Soient G, Γ, W, $\gamma_1, \ldots, \gamma_l$ comme au paragraphe 1.1. Soient $D > 1$ et $S, T \geq 1$ des nombres réels. Supposons qu'il existe $j_0 \in \{1, \ldots, q-1\}$ tel que $\widetilde{H}_{j_0} = 0$, et que le point $(\frac{\log(S)}{\log(D)}, \frac{\log(T)}{\log(D)})$ appartient au cône $\widetilde{\mathcal{C}}_{j_0}$.*

Soit $P \in \mathbb{C}[G]^h_D$ un polynôme non nul, homogène de degré D, qui s'annule à l'ordre T le long de W en tout point de $\Gamma(S)$. Alors on a :

$$S^{\mathrm{rg}(\Gamma)} T^{\dim(W)} \leq c_2 D^{\dim(G)}$$

avec (en notant $r = \mathrm{rg}(\Gamma)$ et $d = \dim(W)$)

$$c_2 = (d!(2n)^{d+r} \deg(G) a^n)^n$$

N.B. On note $\deg(G)$ le degré dans $\mathbb{P}^N$ de l'adhérence de Zariski de G.

Démonstration : On applique le lemme de zéros de Philippon (voir [P3] ou [R14]), qui donne un sous-groupe algébrique connexe H de G, distinct de G, tel que

$$\binom{[\frac{T-1}{n}] + s}{s} \mathrm{Card}\left(\frac{\Gamma(S/n) + H}{H}\right) \deg(H) \leq \deg(G) a^n D^{\dim(G/H)}$$

en posant $s = \dim(W/W \cap TH)$. Or $\binom{[\frac{T-1}{n}]+s}{s} \geq \frac{(T-1)^s}{s! n^s} \geq \frac{T^s}{d!(2n)^d}$ si $T \geq 2$, l'inégalité extrémale étant vraie pour tout T. En outre, quitte à permuter les générateurs de Γ on peut supposer que $\gamma_1, \ldots, \gamma_{\mathrm{rg}(\Gamma/\Gamma\cap H)}$ ont des images linéairement indépendantes dans le groupe quotient $\Gamma/\Gamma \cap H$. Comme $\frac{\Gamma(S/n)+H}{H}$ contient les projetés des combinaisons linéaires de $\gamma_1, \ldots, \gamma_{\mathrm{rg}(\Gamma/\Gamma\cap H)}$ à coefficients compris entre 0 et S/n, on obtient $\mathrm{Card}(\frac{\Gamma(S/n)+H}{H}) \geq (\frac{S}{n} - 1)^{\mathrm{rg}(\Gamma/\Gamma\cap H)} \geq \frac{(S-n)^{\mathrm{rg}(\Gamma/\Gamma\cap H)}}{n^r} \geq \frac{S^{\mathrm{rg}(\Gamma/\Gamma\cap H)}}{(2n)^r}$ si $S \geq 2n$, l'inégalité extrémale étant vraie pour tout S. Comme $\deg(H) \geq 1$ il vient :

$$S^{\mathrm{rg}(\Gamma/\Gamma\cap H)} T^{\dim(W/W\cap TH)} \leq d!(2n)^{d+r} \deg(G) a^n D^{\dim(G/H)}$$

La proposition 3.7 permet alors de conclure.

Lien avec le polygone d'interpolation

On a un lien entre la décomposition en cônes qui provient du lemme de zéros et celle qui vient du lemme d'interpolation. On aura besoin, pour le démontrer, du lemme suivant.

Lemme 3.9 *Soit Δ une demi-droite issue de l'origine, incluse dans $\mathbb{R}_+^2$ et distincte des axes (Ox) et (Oy). Soit H un sous-groupe de G, distinct de G, tel que Δ_H soit une droite. Notons P_H, P_G et $P_{G/H}$ les points d'intersection respectifs de Δ avec Δ_H, Δ_G et la droite d'équation $(\varphi_G - \varphi_H)(x,y) = 0$. Alors P_G est entre P_H et $P_{G/H}$. De plus si P_G est confondu avec l'un de ces deux points alors il est aussi confondu avec l'autre.*

Démonstration : Soit $u = (u_1, u_2) \in {\mathbb{R}_+^*}^2$ un vecteur directeur de Δ. On paramètre Δ en posant $x = u_1 t$ et $y = u_2 t$ pour $t \in \mathbb{R}_+$. On a alors $\varphi_G(x,y) = at - \dim(G)$ et $\varphi_H(x,y) = a_H t - \dim(H)$ où les réels a et a_H sont donnés par $a = u_1 \mathrm{rg}(\Gamma) + u_2 \dim(W)$ et $a_H = u_1 \mathrm{rg}(\Gamma \cap H) + u_2 \dim(W \cap TH)$. On a $a > a_H > 0$ (en utilisant l'hypothèse que si H contient un sous-groupe d'indice fini de Γ et que TH contient W alors $H = G$). Les points P_H, P_G et $P_{G/H}$ correspondent respectivement aux paramètres $\frac{\dim(H)}{a_H}$, $\frac{\dim(G)}{a}$ et $\frac{\dim(G/H)}{a - a_H}$. Le lemme découle alors du fait standard suivant :

Fait 3.10 *Soient a', a'', b', b'' quatre nombres réels strictement positifs. Alors $\frac{b'+b''}{a'+a''}$ est entre $\frac{b'}{a'}$ et $\frac{b''}{a''}$.*

Démonstration : L'inégalité $\frac{b'}{a'} \leq \frac{b'+b''}{a'+a''}$ équivaut à $b'a'' \leq b''a'$ donc à $\frac{b'+b''}{a'+a''} \leq \frac{b''}{a''}$.

Proposition 3.11 *L'existence d'un entier $i_0 \in \{1, \dots, p-1\}$ tel que $H_{i_0} = G$ équivaut à l'existence d'un entier $j_0 \in \{1, \dots, q-1\}$ tel que $\widetilde{H}_{j_0} = 0$. Quand i_0 et j_0 existent, ils sont uniques et on a $\mathcal{C}_{i_0} = \widetilde{\mathcal{C}}_{j_0}$.*

Démonstration : Soit Δ une demi-droite dans $\mathbb{R}_+^2$, issue de l'origine. Soit $i_0 \in \{1, \dots, p-1\}$ tel que $H_{i_0} = G$. Alors Δ appartient à l'intérieur du cône $\mathcal{C}_{i_0}$ si, et seulement si, pour tout sous-groupe H tel que Δ_H soit une droite le point P_H n'est pas entre O et P_G (avec les notations du lemme 3.9). D'après le lemme, ceci équivaut à ce que le point $P_{G/H}$ soit entre O et P_G. Or cette assertion traduit qu'il existe j_0 tel que $\widetilde{H}_{j_0} = 0$, et que Δ appartient à l'intérieur du cône $\widetilde{\mathcal{C}}_{j_0}$. La réciproque se démontre de la même manière. L'unicité de i_0 et j_0 est évidente, puisque $H_1, \dots, H_{p-1}$ sont deux à deux distincts, ainsi que $\widetilde{H}_1, \dots, \widetilde{H}_{q-1}$.

N.B. Le polygone d'interpolation et le lieu d'injectivité sont de part et d'autre de la droite Δ_G (voir les figures 7.1 et 7.2, qui correspondent aux mêmes G, Γ, W). D'après la proposition 3.11, leur intersection est égale à l'intersection avec Δ_G de chacun d'eux : c'est l'ensemble vide ou bien un segment. Dans ce deuxième cas, si le segment est d'intérieur non vide on a $A_{i_0} = B_{j_0}$ et $A_{i_0+1} = B_{j_0+1}$.

Corollaire 3.12 *Le couple (Γ, W) est bien réparti si, et seulement si, on a $q = 2$, $\widetilde{H}_1 = 0$, et que les points B_1 et B_2 sont respectivement sur (Ox) et (Oy).*

Démonstration : C'est immédiat à partir de la proposition précédente.

On a donc équivalence entre la bonne répartition de (Γ, W) relative à l'interpolation (i.e. $\mathcal{C}_1 = \mathbb{R}_+^2$ et $H_1 = G$) et celle relative au lemme de zéros (i.e. $\widetilde{\mathcal{C}}_1 = \mathbb{R}_+^2$ et $\widetilde{H}_1 = 0$). Il est à noter que l'analogue sans multiplicités de ce résultat (i.e. l'équivalence $\mu(\Gamma, G) = \frac{r}{n}$ si, et seulement si, $\mu^*(\Gamma, G) = \frac{r}{n}$) se démontre essentiellement en utilisant le fait 3.10.

Cas particuliers rg(Γ) = 0 et dim(W) = 0

Considérons maintenant le cas particulier où W est réduit à $\{0\}$. On suppose que Γ est de rang strictement positif, et qu'aucun sous-groupe d'indice fini de Γ n'est inclus dans aucun sous-groupe algébrique H distinct de G. Pour chaque H, l'ensemble des points de $\mathbb{R}_+^2$ où $\varphi_G - \varphi_H$ est positive est $\mathbb{R}_+^2$ si $H = G$, et le quart de plan $[\frac{\dim(G/H)}{\mathrm{rg}(\Gamma/\Gamma\cap H)}, +\infty[\times\mathbb{R}_+$ sinon. Ce deuxième cas se produit quand $H = 0$, donc l'intersection de ces domaines (qui joue le rôle de lieu d'injectivité) est un quart de plan $[\frac{\dim(G/\widetilde{H}_1)}{\mathrm{rg}(\Gamma/\Gamma\cap\widetilde{H}_1)}, +\infty[\times\mathbb{R}_+$, où $\widetilde{H}_1$ est choisi de dimension minimale. Alors le sous-groupe $\widetilde{H}_1$ réalise le minimum dans la définition de $\mu(\Gamma, G)$. On a deux cônes $\widetilde{\mathcal{C}}_0 = \emptyset$ et $\widetilde{\mathcal{C}}_1 = \mathbb{R}_+^2$, avec $q = 1$. On a $\widetilde{H}_1 = 0$ si, et seulement si, $H_1 = G$, i.e. (Γ, W) est bien réparti.

Quand $\mathrm{rg}(\Gamma) = 0$ et $\dim(W) \neq 0$, la situation est analogue.

4 Démonstration du lemme d'interpolation

Dans ce paragraphe, on démontre le théorème 1.1 en utilisant les outils introduits dans les paragraphes précédents.

Soient G, Γ, W, γ_1, …, γ_l comme au paragraphe 1.1. Les constantes qui apparaissent dans ce paragraphe dépendent de $G, \Gamma, W, \gamma_1, \ldots, \gamma_l$. La dépendance en G est aussi présente à travers le choix du plongement de G dans $\mathbb{P}^N$, des polynômes E_0, …, E_N qui représentent l'addition de G, et ainsi de suite.

Pour démontrer le théorème 1.1, on peut supposer qu'il n'existe aucun sous-groupe H de G, distinct de G, qui contienne un sous-groupe d'indice fini de Γ et tel que TH contienne W.

Cette partie suit la preuve de Masser [M2]. Au paragraphe 4.1, on traite un cas particulier "non dégénéré", dans lequel on suppose que (Γ, W) est bien réparti relativement aux paramètres considérés, et que $(\gamma_1, \ldots, \gamma_l)$ est une base de Γ. On s'affranchit de cette seconde hypothèse aux paragraphes 4.2 et 4.3. Puis on s'affranchit de l'hypothèse de bonne répartition au paragraphe 4.4, en quotientant par un sous-groupe H_i qui représente une obstruction à l'interpolation (on applique le théorème 1.1 dans H_i et dans G/H_i, en raisonnant par récurrence sur la dimension du groupe algébrique G considéré).

4.1 Un cas non dégénéré

Dans ce paragraphe, on suppose qu'il existe $j_0 \in \{1, \ldots, q-1\}$ tel que $\widetilde{H}_{j_0} = 0$, et que $(\frac{\log(S)}{\log(D)}, \frac{\log(T)}{\log(D)})$ est dans le cône correspondant, c'est-à-dire que pour ce choix de paramètres Γ et W sont bien répartis. On suppose en outre que les générateurs γ_1, …, γ_l sont linéairement indépendants sur $\mathbb{Z}$ (ce qui signifie que Γ est un $\mathbb{Z}$-module de type fini sans torsion, donc libre, et que $(\gamma_1, \ldots, \gamma_l)$ en est une base). Cette hypothèse équivaut à ce que Γ soit sans torsion, et de rang égal à son nombre l de générateurs.

Proposition 4.1 *Supposons γ_1, …, γ_l linéairement indépendants. Soient $S, T \geq 1$ et $D > 1$ des réels tels que le point $(\frac{\log(S)}{\log(D)}, \frac{\log(T)}{\log(D)})$ appartienne au cône $\widetilde{\mathcal{C}}_{j_0}$. Soit $\eta \in \mathcal{F}_{S,T}$ non nulle qui s'annule sur $\mathbb{C}[G]_D^h$. Alors :*

$$S^{\mathrm{rg}(\Gamma)} T^{\dim(W)} \geq c_3 D^{\dim(G)}$$

où c_3 est une constante explicite, qu'on peut définir (si $d \geq 1$ et $l \geq 1$) par

$$c_3 = \left((d!(2n)^{d+l+1} \deg(G) a^n)^n 2^{d+l} \right)^{-2nld}$$

N.B. Si (Γ, W) est bien réparti et $d, l \geq 1$, la preuve donne une meilleure constante c_3 : l'exposant $2nld$ peut être remplacé par $d + l$. La formule obtenue est alors valable même si $d = 0$ ou $l = 0$.

Démonstration : Supposons $d \neq 0$ et $l \neq 0$. On choisit, de la manière suivante, une demi-droite contenue dans le cône $\widetilde{\mathcal{C}}_{j_0}$ et dont on va noter χ la pente. Si (Γ, W) est bien réparti, on choisit la première bissectrice, i.e. $\chi = 1$. Sinon, le polygone d'interpolation a au moins quatre côtés, i.e. $p \geq 3$, et le cône $\widetilde{\mathcal{C}}_{j_0} = \mathcal{C}_{i_0}$ est limité par les demi-droites (OA_{i_0}) et (OA_{i_0+1}) pour un certain $i_0 \in \{1, \ldots, p-1\}$ (d'après la proposition 3.11). Si $i_0 \geq 2$, on note χ la pente de la droite (OA_{i_0}) ; si $i_0 = 1$, on note χ la pente de la droite (OA_2). Dans tous les cas, on a (d'après la proposition 3.5) :

$$\frac{1}{nd} \leq \chi \leq nl \tag{7.7}$$

Définissons un nombre réel $\alpha > 0$ en posant :

$$\alpha + 1 = \left((d!(2n)^{d+l} \deg(G) a^n)^n n! 2^{d+l+1} \right)^{\max(1, \frac{1}{\chi})}$$

Définissons aussi $\beta > 0$ par :

$$\beta + 1 = (\alpha + 1)^\chi$$

Notons $f(s, t)$ le cardinal de $\Gamma(s) \times \{\tau \in \mathbb{N}^d, |\tau| < t\}$, c'est-à-dire la dimension de $\mathcal{F}_{s,t}$. Posons $k = f((1+\alpha)S, (1+\beta)T) - f(\alpha S, \beta T)$. On a $f(s, t) = \lceil s \rceil^l \binom{\lceil t \rceil + d - 1}{d}$ donc

$$k \leq ((1+\alpha)S + 1)^l \binom{(1+\beta)T + d}{d} - (\alpha S)^l \binom{\beta T + d - 1}{d}$$

On utilise le fait suivant :

Fait 4.2 *Soient u un réel supérieur ou égal à un, et j un entier positif ou nul. Alors, pour $\varepsilon \geq 0$:*

$$((\varepsilon + 1)u + 1)^j - (\varepsilon u)^j \leq 2^j (\varepsilon + 1)^{j-1} u^j$$

Démonstration de ce fait : On peut supposer $j \geq 1$, sinon le résultat est évident. On développe alors le membre de gauche, comme polynôme en u qu'on note $\Phi(u)$. Le degré de Φ est j et ses coefficients sont tous positifs, donc $\Phi(u)$ (pour $u \geq 1$) est majoré par $u^j \Phi(1)$. Or $\Phi(1) = (\varepsilon' + 1)^j - (\varepsilon' - 1)^j$ est un polynôme à coefficients positifs en $\varepsilon' = \varepsilon + 1 \geq 1$, d'où (de même) $\Phi(1) \leq 2^j \varepsilon'^{j-1}$. Ceci conclut la preuve du fait.

Démonstration de la proposition (suite) : Le fait, appliqué à S, l et α, donne :

$$k \leq (\alpha S)^l \left(\binom{(1+\beta)T + d}{d} - \binom{\beta T + d - 1}{d} \right) + 2^l (\alpha + 1)^{l-1} S^l \binom{(1+\beta)T + d}{d} \tag{7.8}$$

Considérons le polynôme $R(X) = \binom{X+d-1}{d}$. Alors on peut écrire $R(X) = \sum_{j=0}^{d} r_j X^j$, où les coefficients r_j sont positifs ou nuls (c'est évident en développant le coefficient binomial). On a, d'après le fait appliqué à T, j et β :

$$\begin{aligned} R((1+\beta)T+1) - R(\beta T) &= \sum_{j=0}^{d} r_j \left(((1+\beta)T+1)^j - (\beta T)^j\right) \\ &\leq \sum_{j=0}^{d} r_j 2^j (\beta+1)^{j-1} T^j \\ &\leq 2^d(\beta+1)^{d-1}T^d = 2^d(\alpha+1)^{(d-1)\chi}T^d \end{aligned}$$

puisque $R(1) = 1$ est la somme des r_j. Comme on a $R(X+1) \leq (d+1)X^d$ pour tout $X \geq 1$ (par exemple car $(X+1)+i \leq (i+2)X$ pour tout $i \in \{0, \ldots, d-1\}$), l'équation (7.8) donne :

$$k \leq \left[2^d(\alpha+1)^{l+(d-1)\chi} + (d+1)2^l(\alpha+1)^{l-1+d\chi}\right] S^l T^d$$

En posant $c_k = 2^d + 2^l(d+1)$ on obtient :

$$k \leq c_k(\alpha+1)^{l+d\chi-\min(1,\chi)} S^l T^d \tag{7.9}$$

Posons

$$D_0 = [(n!k)^{1/n}]$$

Si $D \leq aD_0$ alors $D^n \leq a^n n! k$ d'où la conclusion du théorème, avec $c_3^{-1} = c_k a^n n!(\alpha+1)^{l+d\chi-\min(1,\chi)}$, c'est-à-dire

$$c_3^{-1} = (2^d + (d+1)2^l)a^n n! \left((d!(2n)^{d+l} \deg(G) a^n)^n n! 2^{d+l+1}\right)^{l\max(1,\frac{1}{\chi})+d\max(1,\chi)-1}$$

La valeur de c_3 annoncée dans l'énoncé du théorème s'en déduit en appliquant (7.7).

Supposons $D > aD_0$; il reste à en déduire une contradiction. Par définition de D_0, on a (pour la première minoration, voir la preuve du lemme 2 de [M2]) :

$$\dim \mathbb{C}[G]^h_{D_0} \geq \binom{D_0+n}{n} \geq \frac{(D_0+1)^n}{n!} > k$$

Or, par hypothèse, la fonctionnelle $\eta \in \mathcal{F}_{S,T}$ s'annule sur tous les polynômes homogènes de degré D. Comme $D \geq aD_0$, η s'annule aussi sur $\mathbb{C}[G]^h_{aD_0}$ (d'après la remarque qui suit le lemme 1.2). Pour $\delta \in \Gamma(\alpha S)$ et $\tau \in \mathbb{N}^d$, $|\tau| < \beta T$, considérons la fonctionnelle $\mathcal{T}_{\delta,\tau,D_0}\eta$ (avec les notations du paragraphe 2.2). On définit ainsi $f(\alpha S, \beta T)$ fonctionnelles, qui appartiennent à $\mathcal{F}_{(\alpha+1)S,(\beta+1)T}$, sont linéairement indépendantes (d'après le lemme 2.7) et s'annulent sur tous les polynômes homogènes de degré D_0 (d'après la proposition 2.6).

Donc l'application

$$\Psi : \mathcal{F}_{(\alpha+1)S,(\beta+1)T} \to (\mathbb{C}[G]^h_{D_0})^*$$

qui à une fonctionnelle associe la forme linéaire sur $\mathbb{C}[G]^h_{D_0}$ définie par la formule (7.3) a un noyau de dimension au moins $f(\alpha S, \beta T)$. Son rang est donc majoré par $f((\alpha+1)S, (\beta+1)T) - f(\alpha S, \beta T) = k < \dim(\mathbb{C}[G]^h_{D_0})^*$. Donc Ψ n'est pas surjective : l'image de Ψ est contenue dans un hyperplan, qui est l'orthogonal d'un poynôme $P \in \mathbb{C}[G]^h_{D_0}$ non nul. Ce polynôme s'annule à l'ordre $(\beta+1)T$ le long de W en tout point de $\Gamma((\alpha+1)S)$. On peut lui appliquer le lemme de zéros démontré ci-dessus (théorème 3.8). En effet, on a supposé que le point $(\frac{\log(S)}{\log(D)}, \frac{\log(T)}{\log(D)})$ appartient au cône $\widetilde{\mathcal{C}}_{j_0}$, donc le point $(\frac{\log(S)}{\log(D_0)}, \frac{\log(T)}{\log(D_0)})$ appartient aussi à ce cône. Grâce au choix de χ, la translation de vecteur $(\frac{\log(\alpha+1)}{\log(D_0)}, \frac{\log(\beta+1)}{\log(D_0)})$ envoie $\widetilde{\mathcal{C}}_{j_0}$ dans lui-même. Donc le point $(\frac{\log((1+\alpha)S)}{\log(D_0)}, \frac{\log((1+\beta)T)}{\log(D_0)})$ appartient aussi à ce cône. On obtient ainsi :

$$\begin{aligned}(\alpha+1)^{d\chi+l}S^lT^d &\leq c_2D_0^n \\ &\leq c_2 n!k \\ &\leq c_2 n! c_k(\alpha+1)^{d\chi+l-\min(1,\chi)}S^lT^d \\ &< (d!(2n)^{d+l}\deg(G)a^n)^n n! 2^{d+l+1}(\alpha+1)^{d\chi+l-\min(1,\chi)}S^lT^d\end{aligned}$$

Ceci contredit la définition de α. La proposition 4.1 est donc démontrée, dans le cas où $d \neq 0$ et $l \neq 0$.

Supposons $W = \{0\}$. On reprend la preuve ci-dessus, en omettant les multiplicités. On choisit α tel que $\alpha+1 = ((2n)^l \deg(G)a^n)^n n! 2^{l+1}$, et on montre que $k \leq ((1+\alpha)S+1)^l - (\alpha S)^l \leq 2^l(1+\alpha)^{l-1}S^l$. Le reste de la preuve est identique.

Dans le cas où $\mathrm{rg}(\Gamma) = 0$ (i.e. $\Gamma = \{0\}$), on procède de manière analogue, en posant $\beta+1 = (d!(2n)^d \deg(G)a^n)^n n! 2^{d+1}$ et en majorant k par $2^d(1+\beta)^{d-1}T^d$.

4.2 Générateurs superflus

Comme au paragraphe précédent, on suppose qu'il existe j_0 tel que $\widetilde{H}_{j_0} = 0$, et que Γ est sans torsion. Mais on ne fait plus d'hypothèse sur la famille génératrice de Γ.

Proposition 4.3 *Supposons Γ sans torsion. Il existe une constante $c_4 > 0$, qui dépend seulement de G, Γ, W, $\gamma_1, \ldots, \gamma_l$, ayant la propriété suivante. Soient $S, T \geq 1$ et $D > 1$ des réels tels que le point $(\frac{\log(S)}{\log(D)}, \frac{\log(T)}{\log(D)})$ appartienne au cône $\widetilde{\mathcal{C}}_{j_0}$. Soit $\eta \in \mathcal{F}_{S,T}$ non nulle qui s'annule sur $\mathbb{C}[G]^h_D$. Alors :*

$$S^{\mathrm{rg}(\Gamma)}T^{\dim(W)} \geq c_4 D^{\dim(G)}$$

On utilise le lemme suivant ([M2], Lemme 3) :

Lemme 4.4 *Soient Γ' un sous-groupe de G qui contient Γ, et $\gamma'_1, \ldots, \gamma'_{l'}$ des générateurs de Γ'. Il existe une constante c_5, qui dépend seulement de $\gamma_1, \ldots, \gamma_l$ et de $\gamma'_1, \ldots, \gamma'_{l'}$, telle que pour tout $S \geq 0$ il existe $\gamma'_S \in \Gamma'$ avec :*

$$\Gamma(S) \subset \gamma'_S + \Gamma'(c_5 S)$$

Démonstration de la proposition 4.3 : Si $\mathrm{rg}(\Gamma) = 0$ (i.e. $\Gamma = \{0\}$), la proposition 4.3 n'est autre que la proposition 4.1. Supposons $\mathrm{rg}(\Gamma) \geq 1$. Notons χ la pente d'une demi-droite contenue dans le cône $\widetilde{\mathcal{C}}_{j_0}$, choisie comme dans la preuve de la proposition 4.1 si $\dim(W) \geq 1$; on pose $\chi = 0$ si $\dim(W) = 0$. Le groupe abélien Γ est de type fini sans torsion, donc libre : notons $(\gamma'_1, \ldots, \gamma'_r)$ une base du $\mathbb{Z}$-module Γ. Posons $\Gamma' = \Gamma$, en munissant Γ' des générateurs $\gamma'_1, \ldots, \gamma'_r$ et Γ des générateurs $\gamma_1, \ldots, \gamma_l$ (donc $\Gamma(S)$ n'est, en général, pas égal à $\Gamma'(S)$). Le lemme 4.4 donne $\gamma'_S \in \Gamma'$ et une constante c_5 (indépendante de S) telle que $\Gamma(S) \subset \gamma'_S + \Gamma'(c_5 S)$; on peut supposer $c_5 \geq 1$. La proposition 2.6 fournit $\mathcal{T}_{-\gamma'_S,0,D/a}\eta \in \mathcal{F}^{\Gamma',W}_{c_5 S,T} \subset \mathcal{F}^{\Gamma',W}_{c_5 S, c_5^{\chi} T}$ non nulle qui s'annule sur $\mathbb{C}[G]^h_{D/a}$. La proposition 4.1 s'applique ; en effet, les générateurs de Γ' sont linéairement indépendants, le cône $\widetilde{\mathcal{C}}_{j_0}$ relatif à Γ' est le même que celui relatif à Γ, et le point $(\frac{\log(c_5 S)}{\log(D/a)}, \frac{\log(c_5^\chi T)}{\log(D/a)})$ appartient à $\widetilde{\mathcal{C}}_{j_0}$. On obtient donc le résultat voulu, avec $c_4 = c_3 c_5^{-\mathrm{rg}(\Gamma)-\chi \dim(W)} a^{-\dim(G)}$.

4.3 Présence éventuelle de torsion

Dans ce paragraphe, on ne fait plus aucune hypothèse restrictive sur Γ ; on garde seulement l'hypothèse de bonne répartition.

Proposition 4.5 *Il existe une constante $c_6 > 0$, qui dépend seulement de G, Γ, W, γ_1, …, γ_l, ayant la propriété suivante. Soient $S, T \geq 1$ et $D > 1$ des réels tels que le point $(\frac{\log(S)}{\log(D)}, \frac{\log(T)}{\log(D)})$ appartienne au cône $\widetilde{\mathcal{C}}_{j_0}$. Soit $\eta \in \mathcal{F}_{S,T}$ non nulle qui s'annule sur $\mathbb{C}[G]_D^h$. Alors :*

$$S^{\mathrm{rg}(\Gamma)} T^{\dim(W)} \geq c_6 D^{\dim(G)}$$

On utilise le lemme suivant, qui est une forme "triviale" de lemme d'interpolation :

Lemme 4.6 *Soit Σ une partie finie de $\mathbb{P}^N$, contenue dans $\{X_0 \neq 0\}$. Soient T et D des entiers naturels, et $\eta \in \mathcal{F}_T^{\Sigma,W}$ non nulle qui s'annule sur $\mathbb{C}[G]_D^h$. Alors $T^{\dim(W)}\mathrm{Card}(\Sigma) > D$.*

Quand $W = \{0\}$ ou $T = 1$, il s'agit du Lemme 6 de [M2]. Dès que $\dim(W) \geq 1$, on démontre en fait une conclusion plus forte : $T\,\mathrm{Card}(\Sigma) > D$.

N.B. Pour démontrer la proposition 4.5, on utilise ce lemme sans multiplicités, donc le lemme 6 de [M2] suffit. En revanche, le lemme 4.6 est utilisé (avec multiplicités) pour démontrer le théorème 4.7 dans le cas $n = 1$, ce qui permet d'initialiser la preuve (par récurrence) de ce théorème.

Démonstration du lemme 4.6 : On utilise les fonctions suivantes. Tout d'abord, pour tous $\gamma, \delta \in \Sigma$ distincts il existe une forme linéaire $L_{\gamma,\delta}$ en X_0, …, X_N qui s'annule en δ et pas en γ. En effet, si $\delta = [\xi_0 : \ldots : \xi_N]$, pas toutes les formes linéaires $\xi_i X_j - \xi_j X_i$ (pour $i, j \in \{0, \ldots, N\}$) ne s'annulent au point γ. Si $W = \{0\}$ il suffit de considérer, pour $\gamma \in \Sigma$, les polynômes $\prod_{\delta \in \Sigma \setminus \{\gamma\}} L_{\gamma,\delta}$ pour conclure la démonstration du lemme.

Supposons que $d = \dim(W)$ est strictement positif. Pour $\gamma \in \Sigma$, on note $\widehat{\gamma}$ un $(N+1)$-uplet de coordonnées de γ, dont la première vaut 1. Pour tous $\gamma \in \Sigma$ et $i \in \{1, \ldots, d\}$ il existe une forme linéaire $M_{\gamma,i}$ en X_0, …, X_N qui s'annule en γ et telle que $\Delta_j M_{\gamma,i}(\widehat{\gamma}) = \delta_{i,j}$ pour tout $j \in \{1, \ldots, d\}$, où $\delta_{i,j}$ est le symbole de Kronecker. En effet, la matrice $[\Delta_j X_k(\widehat{\gamma})]_{1 \leq k \leq N, 1 \leq j \leq d}$ est de rang d (c'est, par exemple, une conséquence de l'égalité $\rho_n = \tau_n$ démontrée dans [W11], § 7). En ajoutant à $M_{\gamma,i}$ un multiple convenable de X_0 (ce qui ne change pas la propriété $\Delta_j M_{\gamma,i}(\widehat{\gamma}) = \delta_{i,j}$ puisque $\Delta_j X_0 = 0$ pour tout j), on peut supposer qu'elle s'annule en γ (grâce à l'hypothèse (7.1)).

Soient $\gamma \in \Sigma$ et $|\sigma| < T$. On construit un polynôme $P_{\gamma,\sigma}$, homogène de degré $T\,\mathrm{Card}(\Sigma)$, qui s'annule en tout point de $\Sigma \setminus \{\gamma\}$ à l'ordre T le long de W, et vérifie $\underline{\Delta}^\sigma P_{\gamma,\sigma}(\widehat{\gamma}) \neq 0$ et $\underline{\Delta}^{\tilde{\sigma}} P_{\gamma,\sigma}(\widehat{\gamma}) = 0$ pour tout $\tilde{\sigma} \in \mathbb{N}^d$ tel que $\tilde{\sigma} \neq \sigma$ et $|\tilde{\sigma}| \leq |\sigma|$. Pour ce faire, on pose (en suivant, par exemple, la preuve de la proposition 3.8 de [R14]) :

$$P_{\gamma,\sigma} = X_0^{T-|\sigma|} M_{\gamma,1}^{\sigma_1} \ldots M_{\gamma,d}^{\sigma_d} \prod_{\delta \in \Sigma \setminus \{\gamma\}} L_{\gamma,\delta}^T$$

Ceci termine la démonstration du lemme 4.6.

Démonstration de la proposition 4.5 : Notons H la partie de torsion de Γ (i.e. l'ensemble des éléments de Γ qui sont de torsion dans G). C'est un sous-groupe algébrique de dimension zéro de G (non connexe, sauf si $H = \{0\}$) ; son ordre est fini. On peut supposer que D est un multiple entier de $2\delta_H$. On applique alors le corollaire 2.4 avec $D' = \frac{D}{2\delta_H}$ et $D'' = D/2$. La seconde alternative implique, d'après le lemme 4.6, qu'on a $\mathrm{Card}(H) > D/2$; dans ce cas la proposition 4.5 est démontrée. Supposons maintenant que $\eta_1 \in \overline{\mathcal{F}}_{S,T}$ s'annule sur $\mathbb{C}[G/H]^h_{D/2\delta_H}$, avec $\eta_1 \neq 0$. Le polygone d'interpolation associé à $(\overline{\Gamma}, W)$, où $\overline{\Gamma} = \Gamma/H$, est le même que celui associé à (Γ, W) (par la bijection entre les sous-groupes algébriques connexes de G et ceux de G/H). Le point $(\frac{\log(S)}{\log(\frac{D}{2\delta_H})}, \frac{\log(T)}{\log(\frac{D}{2\delta_H})})$ appartient au cône $\widetilde{\mathcal{C}}_{j_0}$, et $\overline{\Gamma}$ est sans torsion, donc la proposition 4.3 donne la conclusion voulue, avec $c_6 = c_4(2\delta_H)^{-\dim(G)}$.

4.4 Démonstration du théorème dans le cas général

On peut maintenant démontrer le théorème suivant, qui est une reformulation du théorème 1.1 (compte tenu du lemme 1.2).

Théorème 4.7 *Il existe une constante c_7 strictement positive, qui dépend seulement de G, Γ, W, γ_1, …, γ_l, ayant la propriété suivante. Soient $S, T \geq 1$ et $D > 0$ des réels, et $\eta \in \mathcal{F}_{S,T}$ une fonctionnelle non nulle qui s'annule identiquement sur $\mathbb{C}[G]^h_D$. Alors il existe un sous-groupe algébrique H connexe de G, non nul, tel que*

$$S^{\mathrm{rg}(\Gamma\cap H)} T^{\dim(W\cap TH)} \geq c_7 D^{\dim(H)}$$

On va utiliser le lemme suivant :

Lemme 4.8 *Soient γ_1, …, γ_l des générateurs de Γ, et Δ un sous-groupe de Γ. Soient δ_1, …, δ_p des générateurs de Δ. Il existe une constante c_8, qui dépend seulement de γ_1, …, γ_l et de δ_1, …, δ_p, telle que pour tous $S \geq 0$ et $\alpha \in \Gamma$ il existe $\beta_{\alpha,S} \in \Gamma$ avec :*

$$\Gamma(S) \cap (\alpha + \Delta) \subset \beta_{\alpha,S} + \Delta(c_8 S)$$

où $\Gamma(S)$ est relatif aux générateurs γ_1, …, γ_l et $\Delta(c_8S)$ aux générateurs δ_1, …, δ_p.

Démonstration du lemme : Notons Λ le sous-groupe de $\mathbb{Z}^l$ formé par les $(n_1, \ldots, n_l)$ tels que $n_1\gamma_1 + \ldots + n_l\gamma_l \in \Delta$. Il existe une base $(e_1, \ldots, e_l)$ de $\mathbb{Z}^l$ telle que $(k_1e_1, \ldots, k_re_r)$ soit une base de Λ, pour des entiers $k_1, \ldots, k_r \geq 1$ et $r \leq l$ convenables. Notons $(\varepsilon_1, \ldots, \varepsilon_l)$ la base canonique de $\mathbb{Z}^l$; on a $e_i = \sum_{j=1}^l a_{i,j}\varepsilon_j$ où la matrice des $a_{i,j}$ appartient à $\mathrm{GL}_l(\mathbb{Z})$. Posons $\tilde{\gamma}_i = \sum_{j=1}^l a_{i,j}\gamma_i$ pour $i \in \{1, \ldots, l\}$ et $\tilde{\delta}_i = k_i\tilde{\gamma}_i$ pour $i \in \{1, \ldots, r\}$. On note $\tilde{\Gamma} = \Gamma$ muni de $(\tilde{\gamma}_1, \ldots, \tilde{\gamma}_l)$, et $\tilde{\Delta} = \Delta$ muni de $(\tilde{\delta}_1, \ldots, \tilde{\delta}_r)$; dans les deux cas il s'agit de familles génératrices. Le lemme 4.4 donne une constante c_9 (indépendante de S) et $\gamma'_S \in \Gamma$ tels que $\Gamma(S) \cap (\alpha + \Delta) \subset \left(\gamma'_S + \tilde{\Gamma}(c_9S)\right) \cap (\alpha + \Delta)$. Notons $\gamma'_S = \kappa_1\tilde{\gamma}_1 + \ldots + \kappa_l\tilde{\gamma}_l$ et $\alpha = \alpha_1\tilde{\gamma}_1 + \ldots + \alpha_l\tilde{\gamma}_l$; alors $\left(\gamma'_S + \tilde{\Gamma}(c_9S)\right) \cap (\alpha + \Delta)$ est inclus dans l'ensemble des $(\kappa_1 + x_1)\tilde{\gamma}_1 + \ldots + (\kappa_r + x_r)\tilde{\gamma}_r + \alpha_{r+1}\tilde{\gamma}_{r+1} + \ldots + \alpha_l\tilde{\gamma}_l$ avec $0 \leq x_j < c_9S$ et $\kappa_j + x_j - \alpha_j \in k_j\mathbb{Z}$ pour tout $j \in \{1, \ldots, r\}$. Pour chaque j, notons χ_j l'entier, compris entre 0 et $k_j - 1$, tel que $\kappa_j - \alpha_j + \chi_j$ soit un multiple de k_j. Posons $\tilde{\beta}_{\alpha,S} = (\kappa_1 + \chi_1)\tilde{\gamma}_1 + \ldots + (\kappa_r + \chi_r)\tilde{\gamma}_r + \alpha_{r+1}\tilde{\gamma}_{r+1} + \ldots + \alpha_l\tilde{\gamma}_l \in \Gamma$. Alors on a $\left(\gamma'_S + \tilde{\Gamma}(c_9S)\right) \cap (\alpha + \Delta) \subset \tilde{\beta}_{\alpha,S} + \tilde{\Delta}(c_9S)$. Le lemme 4.4, appliqué à Δ et $\tilde{\Delta}$, permet alors de conclure.

Démonstration du théorème 4.7 : On démontre le théorème 4.7 par récurrence sur la dimension de G ; pour $n = 1$ il se déduit du lemme 4.6. Supposons-le vrai pour tous les groupes algébriques de dimension au plus $n-1$. Pour tout $i \in \{1, \dots, p-1\}$, on fixe un plongement projectif de G/H_i, et des polynômes homogènes de degré δ_{H_i} qui représentent la projection de G sur G/H_i, comme au paragraphe 2.1. On peut supposer $D > 1$; soit $i \in \{1, \dots, p-1\}$ tel que le point $(\frac{\log(S)}{\log(D)}, \frac{\log(T)}{\log(D)})$ appartienne au cône $\mathcal{C}_i$. Si $H_i = G$, il suffit d'appliquer les propositions 3.11 et 4.5 pour obtenir la conclusion voulue, avec $H = G$ et $c_7 = c_6$. On peut donc supposer que $\dim(H_i)$ est compris entre 1 et $n-1$; en outre on peut supposer que D est un multiple entier de $2\delta_{H_i}$. Appliquons le corollaire 2.4 avec $D' = \frac{D}{2\delta_{H_i}}$ et $D'' = D/2$. Si on est dans la première alternative, on peut (par récurrence) appliquer le théorème 4.7 à G/H_i, $\overline{\Gamma} = \Gamma/\Gamma \cap H_i$ et $\overline{W} = W/W \cap TH_i$; on obtient un sous-groupe algébrique connexe de G/H_i, qui s'écrit H'/H_i où H' est un sous-groupe algébrique connexe de G qui contient strictement H_i et vérifie

$$S^{\mathrm{rg}(\Gamma\cap H'/\Gamma\cap H_i)} T^{\dim(W\cap TH'/W\cap TH_i)} \geq c_{10} \left(\frac{D}{2\delta_{H_i}}\right)^{\dim(H'/H_i)}$$

La proposition 3.3 donne exactement la conclusion cherchée, avec $H = H_i$. Si on est dans la seconde alternative du corollaire 2.4, on a $\alpha \in \Gamma$ et $\eta' \in \mathcal{F}_T^{\Gamma(S)\cap(\alpha+H_i), W\cap TH_i}$ non nulle qui s'annule sur $\mathbb{C}[G]^h_{D/2}$. Choisissons des générateurs $\delta_1, \dots, \delta_p$ de $\Delta = \Gamma \cap H_i$. Le lemme 4.8 et la proposition 2.6 donnent $\mathcal{T}_{-\beta_{\alpha,S},0,D/2a}\eta' \in \mathcal{F}_{c_8 S,T}^{\Delta, W\cap TH_i}$ non nulle qui s'annule sur $\mathbb{C}[G]^h_{D/2a}$. Le théorème 4.7, appliqué à H_i, Δ et $W \cap TH_i$, fournit un sous-groupe algébrique H' non nul de H_i tel que

$$S^{\mathrm{rg}(\Gamma\cap H')} T^{\dim(W\cap TH')} \geq c_{11} \left(\frac{D}{2a}\right)^{\dim(H')}$$

Cette minoration est exactement la conclusion cherchée, avec $H = H'$.

5 Prolongements possibles

Pourrait-on considérer des multi-indices σ tels que $0 \leq \sigma_j < T_j$ pour tout $j \in \{1, \dots, d\}$, plutôt que d'imposer $|\sigma| < T$? De même, pourrait-on considérer $\Gamma(S_1, \dots, S_l)$ au lieu de $\Gamma(S)$?

Pourrait-on obtenir une autre démonstration du lemme d'interpolation, qui soit similaire aux preuves de lemmes de zéros ? Pour cela, il serait probablement utile de donner une définition plus “naturelle” de l'action d'une fonctionnelle sur un polynôme homogène : ici le fait de privilégier X_0 rompt la symétrie. Considérons pour simplifier une situation sans multiplicités. Un autre choix de dénominateur correspond, quand on considère l'action de $\eta = \sum_\gamma \lambda_\gamma \mathrm{ev}_\gamma$, à multiplier chaque λ_γ par une certaine constante non nulle (qui dépend de γ). Ce qui est naturel (avec la définition donnée ici) n'est donc pas la valeur de λ_γ, mais seulement sa nullité ou sa non-nullité. D'ailleurs on n'utilise pas vraiment autre chose dans ce texte (voir notamment la preuve du lemme 2.7).

Chapitre 8

Lemmes de zéros et algèbres de Hopf

Dans ce texte, on formule les lemmes de zéros et d'interpolation pour un groupe algébrique affine commutatif dans le langage des algèbres de Hopf (voir par exemple [A1] ou [H9]). La traduction met en jeu une équivalence de catégories, énoncée et démontrée dans la première partie. Une référence pour le langage des catégories est [M1]. Dans tout ce texte, le corps de base est $\mathbb{C}$ (tout autre corps algébriquement clos de caractéristique zéro convient aussi).

On considère la catégorie des groupes algébriques affines commutatifs. Elle est anti-équivalente (d'après [A1], Chapitre 4, §2.1) à celle des algèbres de Hopf de type fini, commutatives et co-commutatives[1] : tout tel groupe G correspond à l'algèbre $\mathbb{C}[G]$ des fonctions régulières (polynomiales) sur G. En outre G est connexe si, et seulement si, $\mathbb{C}[G]$ est intègre (comme anneau).

Le théorème de structure des groupes algébriques affines commutatifs (voir par exemple [D1]) affirme que tout tel groupe G est un produit $G_0 \times G_1$, où G_0 (respectivement G_1) est la partie additive (respectivement multiplicative) de G ; le groupe G_0 est isomorphe à une puissance de $\mathbb{G}_a$, et la composante neutre du groupe G_1 est isomorphe à une puissance de $\mathbb{G}_m$. Ce résultat se traduit en termes d'algèbres de Hopf. Soit H une algèbre de Hopf de type fini, commutative et co-commutative. Notons Γ le groupe des éléments group-like de H, et W l'espace vectoriel des éléments primitifs de H. Alors Γ est de type fini, et W de dimension finie. Le sous-espace vectoriel de H engendré par Γ est une sous-algèbre de Hopf, isomorphe à l'algèbre $\mathbb{C}\Gamma$ du groupe Γ. La sous-algèbre (associative) de H engendrée par W est une sous-algèbre de Hopf, isomorphe à $\mathrm{Sym}(W)$. Si $H = \mathbb{C}[G]$ alors $\mathbb{C}[G_0]$ est isomorphe à $\mathrm{Sym}(W)$, $\mathbb{C}[G_1]$ à $\mathbb{C}\Gamma$, et H à $\mathbb{C}\Gamma \otimes \mathrm{Sym}(W)$, comme algèbres de Hopf. Le choix d'une base $(\partial_1, \dots, \partial_d)$ de W induit un isomorphisme linéaire entre H et l'algèbre $\mathbb{C}(\Gamma \times \mathbb{N}^d)$ du semi-groupe $\Gamma \times \mathbb{N}^d$. Pour $\sigma = (\sigma_1, \dots, \sigma_d) \in \mathbb{N}^d$ et $\gamma \in \Gamma$, on note $\mathrm{ev}_{\gamma,\sigma}$ l'élément de $\mathbb{C}(\Gamma \times \mathbb{N}^d)$ qui correspond, par cet isomorphisme, à $\gamma \otimes (\partial_1^{\sigma_1} \cdot \ldots \cdot \partial_d^{\sigma_d})$.

[1]Sur un corps de caractéristique zéro, toute algèbre de Hopf commutative est réduite : voir le corollaire 2.5.4 de [A1].

1 Une équivalence de catégories

1.1 Définitions, énoncé du théorème principal

Notons $\mathcal{C}_1$ la catégorie définie comme suit. Les objets de $\mathcal{C}_1$ sont les triplets (G,Γ,W) où G est un groupe algébrique affine commutatif, Γ un sous-groupe de $G(\mathbb{C})$ de type fini, et W un sous-espace vectoriel de l'algèbre de Lie TG de G, tels qu'aucun sous-groupe algébrique K de G, distinct de G, ne vérifie $\Gamma \subset K(\mathbb{C})$ et $W \subset TK$. Les morphismes de (G,Γ,W) dans (G',Γ',W'), dans la catégorie $\mathcal{C}_1$, sont les morphismes de groupes algébriques de G dans G' qui induisent des morphismes de Γ dans Γ' et de W dans W'.

Notons $\mathcal{C}_2$ la catégorie suivante. Les objets de $\mathcal{C}_2$ sont les triplets $(H_1,H_2,<\cdot,\cdot>)$ où H_1 et H_2 sont des algèbres de Hopf de type fini, commutatives et co-commutatives, et $<\cdot,\cdot>$ est un produit bilinéaire non dégénéré, de $H_1 \times H_2$ dans $\mathbb{C}$, qui vérifie les propriétés suivantes :

$$< x, yy' >=< \Delta x, y \otimes y' > \text{ pour tous } x \in H_1 \text{ et } y, y' \in H_2 \tag{8.1}$$

$$< xx', y >=< x \otimes x', \Delta y > \text{ pour tous } x, x' \in H_1 \text{ et } y \in H_2 \tag{8.2}$$

Dans ces équations, on note aussi $<\cdot,\cdot>$ le produit bilinéaire de $H_1^{\otimes 2} \times H_2^{\otimes 2}$ dans $\mathbb{C}$, défini pour $x,x' \in H_1$ et $y,y' \in H_2$ par :

$$< x \otimes x', y \otimes y' >=< x, y >< x', y' >$$

Les morphismes de $(H_1,H_2,<\cdot,\cdot>)$ dans $(H_1',H_2',<\cdot,\cdot>')$ sont les couples (f,g), où f est un morphisme d'algèbres de Hopf de H_1 dans H_1' et g est un morphisme d'algèbres de Hopf de H_2' dans H_2, tels que $< x, g(y') >=< f(x), y' >$ pour tous $x \in H_1$ et $y' \in H_2'$. La composition de deux morphismes est donnée par $(f,g) \circ (f',g') = (f \circ f', g' \circ g)$.

Notons ε_1 et ε_2 les co-unités respectives de H_1 et H_2, et 1 leurs unités (si H_1 est l'algèbre des polynômes sur un groupe G_1, la co-unité ε_1 est l'évaluation en l'élément neutre de G_1). Alors les équations (8.1) et (8.2) impliquent :

$$< \cdot, 1 >= \varepsilon_1 \text{ et } < 1, \cdot >= \varepsilon_2 \tag{8.3}$$

En effet, comme ε_1 et $< \cdot, 1 >$ sont des morphismes d'algèbres (d'après (8.2)) et que la réunion des éléments group-like et des éléments primitifs engendre l'algèbre H_1, il suffit de vérifier que $< \gamma, 1 >= \varepsilon_1(\gamma)$ et $< \partial, 1 >= \varepsilon_1(\partial)$ pour γ group-like et ∂ primitif. Or (8.1) donne $< \gamma, 1 >^2=< \gamma, 1 >$, d'où $< \gamma, 1 >= 1$ car $<\cdot,\cdot>$ est non dégénéré. En particulier $< 1, 1 >= 1$; l'équation (8.1) donne alors $< \partial, 1 >= 0$, ce qui termine la démonstration de (8.3).

Le résultat principal de cette partie est le suivant :

Théorème 1.1 *Les catégories $\mathcal{C}_1$ et $\mathcal{C}_2$ sont équivalentes.*

Ce théorème est démontré au paragraphe 1.3. Auparavant, au paragraphe 1.2, on formule en termes d'algèbres de Hopf duales les axiomes de la catégorie $\mathcal{C}_2$.

N.B. On peut démontrer, comme variante du théorème 1.1, que les deux catégories $\mathcal{C}_1'$ et $\mathcal{C}_2'$ ci-après sont équivalentes. Les objets de $\mathcal{C}_1'$ sont les triplets (G,Γ,W), où G est un groupe algébrique affine commutatif, Γ un groupe abélien de type fini et W un espace vectoriel de dimension finie, munis d'un homomorphisme de groupes de Γ dans $G(\mathbb{C})$ et d'une application linéaire de W dans TG. Quant à $\mathcal{C}_2'$, elle est définie comme $\mathcal{C}_2$ à ceci près qu'on ne suppose pas le produit bilinéaire $<\cdot,\cdot>$ non dégénéré.

1.2 Rappels sur la dualité entre algèbres de Hopf

Soit H une algèbre de Hopf. Notons H^* l'espace vectoriel dual de H, c'est-à-dire l'ensemble des formes linéaires de H dans $\mathbb{C}$. La structure de co-algèbre sur H permet ([A1], p. 55) de munir H^* d'une structure d'algèbre, pour laquelle le produit $\varphi_1\varphi_2$ de $\varphi_1, \varphi_2 \in H^*$ envoie $h \in H$ sur $\sum_i \varphi_1(h'_i)\varphi_2(h''_i)$, en notant $\Delta h = \sum_i h'_i \otimes h''_i$. Cependant, le produit sur H ne permet pas, en général, de munir H^* d'une structure de co-algèbre.

On note H° l'ensemble des $\varphi \in H^*$ tels que $\ker(\varphi)$ contienne un idéal de l'algèbre H de codimension finie. Alors ([A1], p. 87) H° est une sous-algèbre de H^*, et on peut ([A1], Théorème 2.2.12) munir H° d'une structure de co-algèbre en utilisant le produit de H. En effet, pour tout $f \in H^\circ$ il existe $\varphi_1, \dots, \varphi_l, \psi_1, \dots, \psi_l \in H^\circ$ tels que $f(xy) = \sum_{i=1}^{l} \varphi_i(x)\psi_i(y)$ pour tous $x, y \in H$. On a alors $\Delta f = \sum_i \varphi_i \otimes \psi_i$.

Finalement, H° est une algèbre de Hopf, appelée *algèbre de Hopf duale* de H. Si $H = \mathbb{C}[G]$ est l'algèbre des polynômes sur un groupe algébrique affine commutatif G, alors ([A1], Théorème 4.3.13) une base de l'espace vectoriel H° est formée par les $R \mapsto \partial_1^{a_1} \dots \partial_n^{a_n} R(g)$, pour $a_1, \dots, a_n \in \mathbb{N}$ et $g \in G(\mathbb{C})$, où $(\partial_1, \dots, \partial_n)$ désigne une base de TG. On voit sur cet exemple que même en supposant H de type fini, l'algèbre de Hopf duale H° n'est pas de type fini en général.

Soit F un espace vectoriel. On peut munir ([A1], p. 67) l'espace vectoriel dual F^* d'une topologie en prenant pour base de voisinages de $\varphi \in F^*$ les intersections finies d'hyperplans $\varphi + x^\perp$, pour $x \in F$, en notant $x^\perp$ l'ensemble des formes linéaires qui s'annulent en x. Alors un sous-espace vectoriel de F^* est dense si, et seulement si, aucun $x \in F$ non nul n'annule tous ses éléments.

Soient E et F deux espaces vectoriels sur $\mathbb{C}$, munis d'un produit bilinéaire non dégénéré de $E \times F$ dans $\mathbb{C}$. Alors on peut identifier E à un sous-espace vectoriel dense de F^*. Réciproquement, tout isomorphisme entre E et un sous-espace vectoriel dense E' de F^* donne un produit bilinéaire non dégénéré de $E \times F$ dans $\mathbb{C}$.

Définissons une catégorie $\mathcal{C}_3$ comme suit. Les objets de $\mathcal{C}_3$ sont les couples (H_1, H_2), où H_2 est une algèbre de Hopf de type fini, commutative et co-commutative, et H_1 est une sous-algèbre de Hopf de type fini, dense, de H_2°. Les morphismes de (H_1, H_2) dans (H'_1, H'_2) sont les morphismes d'algèbres de Hopf g de H'_2 dans H_2 dont la transposée envoie H_1 dans H'_1. On les compose "à l'envers", c'est-à-dire que $g \circ g'$, vu dans $\mathcal{C}_3$, est $g' \circ g$ vu dans la catégorie des algèbres de Hopf.

Alors on a le lemme suivant :

Lemme 1.2 *Les catégories $\mathcal{C}_2$ et $\mathcal{C}_3$ sont équivalentes.*

N.B. Dans la suite, on utilisera seulement les catégories $\mathcal{C}_1$ et $\mathcal{C}_2$, excepté au paragraphe 2.3.

Démonstration : Soit $(H_1, H_2, < \cdot, \cdot >)$ un objet de $\mathcal{C}_2$. Grâce à $< \cdot, \cdot >$, on peut identifier H_1 à un sous-espace vectoriel dense de H_2^*. Montrons que ce sous-espace est contenu dans H_2°. Soit $x \in H_1$; notons $\Delta x = \sum_i x'_i \otimes x''_i$, et E le sous-espace vectoriel de H_2 formé par les y tels que $< x'_i, y >= 0$ pour tout i. Alors E est de codimension finie, et pour tous $y \in E$ et $y' \in H_2$ on a :

$$< x, yy' >= \sum_i < x'_i, y >< x''_i, y' >= 0$$

Donc l'idéal de H_2 engendré par E est de codimension finie, et il est inclus dans le noyau de la forme linéaire $< x, \cdot >$. Donc H_1 se plonge dans H_2°. En outre les axiomes (8.1) et (8.2) traduisent que ce plongement est un homomorphisme d'algèbres de Hopf.

1.3 Démonstration du théorème principal

On construit une équivalence de catégories comme suit. Soit (G, Γ, W) un objet de la catégorie $\mathcal{C}_1$. Posons $H_2 = \mathbb{C}[G]$ et $H_1 = \mathbb{C}\Gamma \otimes \mathrm{Sym}(W)$. On définit un produit bilinéaire $< \cdot, \cdot >$ de $H_1 \times H_2$ dans $\mathbb{C}$ en posant, pour $R \in H_2$, $\gamma \in \Gamma$ et $\partial^{(1)}, \ldots, \partial^{(k)} \in W$:

$$< \gamma \otimes \left(\partial^{(1)} \cdot \ldots \cdot \partial^{(k)}\right), R >= \left(\partial^{(1)} \ldots \partial^{(k)} R\right)(\gamma) \tag{8.4}$$

où on voit, dans le membre de droite, les $\partial^{(i)}$ comme des dérivations de l'algèbre H_2.

Montrons que $(H_1, H_2, < \cdot, \cdot >)$ est un objet de $\mathcal{C}_2$. On commence par remarquer que pour $\partial \in W$, vu comme une dérivation de H_2, on a

$$\Delta \circ \partial = (\partial, 0) \circ \Delta = (0, \partial) \circ \Delta$$

comme applications linéaires de H_2 dans $H_2^{\otimes 2}$, en voyant $(\partial, 0)$ et $(0, \partial) \in W \times W$ comme des dérivations sur $\mathbb{C}[G \times G] = H_2^{\otimes 2}$ (par exemple $(\partial, 0)(R_1 \otimes R_2) = (\partial R_1) \otimes R_2$ pour $R_1, R_2 \in H_2$).

En outre, la définition de $< \cdot, \cdot >$ donne immédiatement

$$< \partial x, R >=< x, \partial R > \text{ et } < \gamma x, R >=< x, \tau_\gamma R >$$

pour tous $x \in H_1$, $R \in H_2$, $\gamma \in \Gamma$ et $\partial \in W$. Ici les produits ∂x et γx sont des produits dans l'algèbre associative H_1, alors que ∂R est l'image de R par la dérivation invariante ∂. Dans la seconde équation, on note τ_γ la translation par γ, définie par $\tau_\gamma R(g) = R(\gamma + g)$ pour $g \in G$.

Vérifions l'axiome (8.1). Par linéarité, on peut supposer que $x = \gamma \otimes \left(\partial^{(1)} \cdot \ldots \cdot \partial^{(k)}\right)$, avec $k \geq 0$. On raisonne par récurrence sur k. Si $k = 0$, c'est évident car $\Delta(\gamma \otimes 1) = (\gamma \otimes 1) \otimes (\gamma \otimes 1)$. Si le résultat vrai jusqu'à $k - 1$, alors il est vrai pour k grâce au calcul suivant, valable pour $x' \in H_1$, $R_1, R_2 \in H_2$ et $\partial \in W$:

$$\begin{aligned}
< \partial x', R_1 R_2 > &= < x', \partial(R_1 R_2) > \\
&= < x', R_1 \partial R_2 + (\partial R_1) R_2 > \\
&= < \Delta x', R_1 \otimes \partial R_2 + \partial R_1 \otimes R_2 > \\
&= < \Delta x', (1 \otimes \partial + \partial \otimes 1)(R_1 \otimes R_2) > \\
&= < (1 \otimes \partial + \partial \otimes 1) \Delta x', R_1 \otimes R_2 > \\
&= < \Delta(\partial x'), R_1 \otimes R_2 >
\end{aligned}$$

Vérifions maintenant l'axiome (8.2). Par linéarité, on peut supposer $x = \gamma \otimes \left(\partial^{(1)} \cdot \ldots \cdot \partial^{(k)}\right)$ et $x' = \gamma' \otimes \left(\partial'^{(1)} \cdot \ldots \cdot \partial'^{(k')}\right)$. Pour $k = k' = 0$ le résultat est évident. Il suffit alors pour conclure de procéder par récurrence, grâce au calcul suivant :

$$\begin{aligned}
< \partial x x', R > &= < x x', \partial R > \\
&= < x \otimes x', \Delta(\partial R) > \\
&= < x \otimes x', (\partial, 0) \Delta R > \\
&= < (\partial x) \otimes x', \Delta R >
\end{aligned}$$

et de même avec $(0, \partial)$ pour obtenir $x \otimes (\partial x')$.

Montrons que le produit bilinéaire $< \cdot, \cdot >$ est non dégénéré. Soit $R \in H_2$ tel que $< x, R >= 0$ pour tout $x \in H_1$. Alors $R \in \mathbb{C}[G]$ s'annule en tout point de Γ à un ordre infini le long de W, donc $R \circ \exp_G$ s'annule en tout point de $\Sigma = \exp_G^{-1}(\Gamma) + W$. Or $\exp_G(\Sigma)$ n'est contenu (par minimalité de G) dans aucun sous-groupe algébrique de G, distinct de G lui-même. Donc l'adhérence de Zariski de $\exp_G(\Sigma)$, qui est un sous-groupe algébrique de G ([A1], théorème 4.2.2), est G tout entier. Or R est identiquement nulle sur $\exp_G(\Sigma)$, donc $R = 0$. Pour démontrer qu'il n'existe pas de $x \in H_1$ non nul tel que $< x, R >= 0$ pour tout $R \in H_2$, on peut appliquer la dualité de Fourier-Borel au résultat précédent, ou bien utiliser un lemme d'interpolation (une généralisation du lemme 6 de [M2] suffit, par exemple le lemme 4.6 de [F6]).

On a donc montré que $(H_1, H_2, < \cdot, \cdot >)$ est un objet de $\mathcal{C}_2$.

Considérons maintenant un morphisme entre deux objets (G, Γ, W) et (G', Γ', W') de $\mathcal{C}_1$, donné par $\phi : G \to G'$. Notons $(H_1, H_2, < \cdot, \cdot >)$ et $(H_1', H_2', < \cdot, \cdot >')$ les objets de $\mathcal{C}_2$ associés. Alors ϕ induit un morphisme d'algèbres de Hopf g de $H_2' = \mathbb{C}[G']$ dans $H_2 = \mathbb{C}[G]$. En outre les morphismes de Γ dans Γ' et de W dans W' induits par ϕ permettent de définir un morphisme d'algèbres de Hopf f de $H_1 = \mathbb{C}\Gamma \otimes \mathrm{Sym}(W)$ dans $H_1' = \mathbb{C}\Gamma' \otimes \mathrm{Sym}(W')$. On vérifie immédiatement que (f, g) est un morphisme de $(H_1, H_2, < \cdot, \cdot >)$ dans $(H_1', H_2', < \cdot, \cdot >')$, dans la catégorie $\mathcal{C}_2$.

Réciproquement, soit $(H_1, H_2, < \cdot, \cdot >)$ un objet de $\mathcal{C}_2$. Notons Γ le groupe des éléments group-like de H_1, W l'espace vectoriel de ses éléments primitifs, et G un groupe algébrique affine commutatif tel que $H_2 = \mathbb{C}[G]$. La classification rappelée dans l'introduction montre que Γ est de type fini, W de dimension finie, et que H_1 est isomorphe à $\mathbb{C}\Gamma \otimes \mathrm{Sym}(W)$. On va utiliser le produit bilinéaire $< \cdot, \cdot >$ pour plonger Γ dans $G(\mathbb{C})$ et W dans TG.

Les points complexes de G sont exactement les homomorphismes d'algèbres de H_2 dans $\mathbb{C}$. On définit donc une application de Γ dans $G(\mathbb{C})$ en associant à tout $\gamma \in \Gamma$ l'homomorphisme $< \gamma, \cdot >$ (il s'agit bien d'un homomorphisme d'algèbres d'après (8.1), puisque $\Delta\gamma = \gamma \otimes \gamma$). L'axiome (8.2) montre qu'on définit ainsi un homomorphisme de groupes de Γ dans $G(\mathbb{C})$. Cet homomorphisme est injectif car $< \cdot, \cdot >$ est non dégénéré (si $\gamma \neq \gamma' \in \Gamma$ alors il existe $R \in H_2$ tel que $< \gamma, R > \neq < \gamma', R >$). Donc Γ est (isomorphe à) un sous-groupe de $G(\mathbb{C})$.

En ce qui concerne W, on commence par énoncer le fait suivant. Notons $\mathcal{E}_G$ l'ensemble des formes linéaires $u : \mathbb{C}[G] \to \mathbb{C}$ qui vérifient $u(1) = 0$ et $u(R_1 R_2) = R_1(e)u(R_2) + u(R_1)R_2(e)$ pour tous $R_1, R_2 \in \mathbb{C}[G]$, où e est l'élément neutre de G. Alors l'application qui à ∂ associe la forme linéaire $R \mapsto (\partial R)(e)$ définit un isomorphisme entre TG et $\mathcal{E}_G$. En considérant les formes linéaires $< \partial, \cdot >$, on plonge W dans $\mathcal{E}_G$ (grâce à (8.1) et (8.3)), donc dans TG.

Montrons que le produit bilinéaire $< \cdot, \cdot >$ est donné par (8.4). Par construction, on a $< \gamma, R >= R(\gamma)$ et $< \partial, R >= (\partial R)(e) =< 1, \partial R >$ pour tous $\gamma \in \Gamma$, $\partial \in W$ et $R \in H_2$. Pour tout $x \in H_1$, on en déduit :

$$\begin{aligned} < x, \partial R > &= < 1 \otimes x, \Delta(\partial R) > \\ &= < 1 \otimes x, (\partial, 0)\Delta R > \\ &= < \partial \otimes x, \Delta R > \\ &= < \partial x, R > \end{aligned}$$

La relation (8.4) en découle (par récurrence sur k).

Si K est un sous-groupe algébrique de G, distinct de G, qui contient Γ et tel que TK contienne W, alors il existe $R \in H_2$ non nul qui induit une fonction identiquement nulle sur K. Alors $< x, R >= 0$ pour tout $x \in H_1$, ce qui est contradictoire. Ceci démontre que (G, Γ, W) est un objet de $\mathcal{C}_1$, et termine la preuve du théorème 1.1.

2 Formulation de lemmes de zéros et d'interpolation

Pour tout groupe algébrique affine commutatif K, on note K_0 et K_1 les parties additive et multiplicative (respectivement) de K ; on a alors $K = K_0 \times K_1$.

2.1 Graduations

Soit $H = \mathbb{C}[G]$ l'algèbre de Hopf d'un groupe algébrique affine commutatif G. Si on note Γ et W, respectivement, l'ensemble des éléments group-like et primitifs de H, alors H est isomorphe à $\mathbb{C}\Gamma \otimes \mathrm{Sym}(W)$ et G_0 est isomorphe à $\mathbb{G}_a^{\dim(W)}$. Fixons un plongement de G_1 dans $\mathbb{G}_m^{d_1}$, pour un certain entier d_1. Ceci correspond à se donner un morphisme surjectif de $\mathbb{Z}^{d_1}$ dans Γ, c'est-à-dire une famille génératrice $(\gamma_1, \ldots, \gamma_{d_1})$ de Γ. On définit alors une double filtration sur H en posant, pour $S, T \in \mathbb{N}$:

$$H_{S,T} = \mathrm{Vect}_{\mathbb{C}}(\Gamma(S)) \otimes \left(\oplus_{0 \leq k < T} \mathrm{Sym}^k(W) \right)$$

où $\Gamma(S)$ est l'ensemble des combinaisons linéaires $n_1\gamma_1 + \ldots + n_{d_1}\gamma_{d_1}$ avec des entiers n_j tels que $0 \leq n_j < S$. Si on note $X_1, \ldots, X_{d_0}$ des coordonnées sur G_0 (avec $d_0 = \dim(W)$) et $Y_1, \ldots, Y_{d_1}$ les coordonnées sur $\mathbb{G}_m^{d_1}$, alors H est un quotient de $\mathbb{C}[X_1, \ldots, X_{d_0}, Y_1, \ldots, Y_{d_1}, Y_1^{-1}, \ldots, Y_{d_1}^{-1}]$ et $H_{S,T}$ est l'image dans H de l'espace des polynômes de degré strictement inférieur à T en $X_1, \ldots, X_{d_0}$ et strictement inférieur à S en $Y_1, \ldots, Y_{d_1}$.

Soit $H' = \mathbb{C}[G']$ une algèbre de Hopf quotient de H. La projection canonique p de H dans H' correspond à un plongement de G' dans G, qui permet de plonger G' dans $\mathbb{G}_m^{d_1}$. On a donc une double filtration sur H', telle que $p(H_{S,T}) \subset H'_{S,T}$ pour tous $S, T \in \mathbb{N}$.

2.2 Transformée de Fourier-Borel

Dans la catégorie $\mathcal{C}_2$ on a une dualité[2] : à tout objet $(H_1, H_2, < \cdot, \cdot >)$ on associe l'objet $(\widetilde{H}_1, \widetilde{H}_2, < \cdot, \cdot \tilde{>})$ défini par $\widetilde{H}_1 = H_2$, $\widetilde{H}_2 = H_1$ et $< u, v \tilde{>} =< v, u >$ pour tous $v \in H_1$ et $u \in H_2$. L'équivalence de catégories du théorème 1.1 permet de transporter dans la catégorie $\mathcal{C}_1$ cette dualité ; on obtient alors la *dualité de Fourier-Borel* (voir [W5] ou [W8]). Cette dualité est involutive, et contravariante.

Soit (G, Γ, W) un objet de $\mathcal{C}_1$. Son image $(\widetilde{G}, \widetilde{\Gamma}, \widetilde{W})$ par la transformation de Fourier-Borel a les propriétés suivantes. Tout d'abord, on a les isomorphismes suivants entre algèbres de Hopf :

$$\begin{array}{rcl} \mathbb{C}[G_0] & \simeq & \mathrm{Sym}(\widetilde{W}) \\ \mathbb{C}[G_1] & \simeq & \mathbb{C}\widetilde{\Gamma} \\ \mathrm{Sym}(W) & \simeq & \mathbb{C}[\widetilde{G}_0] \\ \mathbb{C}\Gamma & \simeq & \mathbb{C}[\widetilde{G}_1] \end{array}$$

[2]Précisément, il s'agit d'un isomorphisme contravariant de $\mathcal{C}_2$ dans $\mathcal{C}_2$.

En outre, le choix d'un isomorphisme de G_0 avec $\mathbb{G}_a^{d_0}$ donne une base de $\widetilde{W}$. Le choix d'un plongement de G_1 dans $\mathbb{G}_m^{d_1}$ donne une famille génératrice $(\widetilde{\gamma}_1, \ldots, \widetilde{\gamma}_{d_1})$ de $\widetilde{\Gamma}$; le rang de $\widetilde{\Gamma}$ est égal à la dimension de G_1. D'autre part, le choix d'une base de W permet d'identifier $\widetilde{G}_0$ à $\mathbb{G}_a^d$, avec $d = \dim(W)$. Supposons Γ muni d'une famille génératrice $(\gamma_1, \ldots, \gamma_l)$; alors $\widetilde{G}_1$ est plongé dans $\mathbb{G}_m^l$, et on a $\dim(\widetilde{G}_1) = \mathrm{rg}(\Gamma)$.

Notons $X_1, \ldots, X_{d_0}$ les coordonnées sur G_0 (identifié avec $\mathbb{G}_a^{d_0}$) et $Y_1, \ldots, Y_{d_1}$ les coordonnées sur $\mathbb{G}_m^{d_1}$; alors $\mathbb{C}[G]$ est un quotient de $\mathbb{C}[X_1, \ldots, X_{d_0}, Y_1, \ldots, Y_{d_1}, Y_1^{-1}, \ldots, Y_{d_1}^{-1}]$. L'image dans $\mathbb{C}[G]$ du monôme $X_1^{n_1} \ldots X_{d_0}^{n_{d_0}} Y_1^{m_1} \ldots Y_{d_1}^{m_{d_1}}$ correspond à la fonctionnelle $\mathrm{ev}_{\widetilde{\gamma}, \widetilde{\sigma}}$ où $\widetilde{\gamma} = m_1\widetilde{\gamma}_1 + \ldots + m_{d_1}\widetilde{\gamma}_{d_1} \in \widetilde{\Gamma}$ et $\widetilde{\sigma} = (n_1, \ldots, n_{d_0}) \in \mathbb{N}^{d_0}$. De même, la fonctionnelle $\mathrm{ev}_{\gamma,\sigma}$, avec $\gamma = a_1\gamma_1 + \ldots + a_l\gamma_l \in \Gamma$ et $\sigma = (b_1, \ldots, b_d) \in \mathbb{N}^d$, correspond à l'image dans $\mathbb{C}[\widetilde{G}]$ du monôme $\widetilde{X}_1^{b_1} \ldots \widetilde{X}_d^{b_d} \widetilde{Y}_1^{a_1} \ldots \widetilde{Y}_l^{a_l}$, en notant $\widetilde{X}_1, \ldots, \widetilde{X}_d$ les coordonnées sur $\widetilde{G}_0$ identifié à $\mathbb{G}_a^d$ et $\widetilde{Y}_1, \ldots, \widetilde{Y}_l$ les coordonnées sur $\mathbb{G}_m^l$.

N.B. Le groupe algébrique G est connexe si, et seulement si, $\widetilde{\Gamma}$ est sans torsion, car ces deux assertions signifient que l'algèbre $\mathbb{C}[G_1] \simeq \mathbb{C}\widetilde{\Gamma}$ est intègre.

2.3 Sous-algèbre obstructrice

Soit $(H_1, H_2, < \cdot, \cdot >)$ un objet de la catégorie $\mathcal{C}_2$. On a vu au paragraphe 1.2 qu'on peut identifier H_1 à une sous-algèbre de Hopf de type fini, dense, de H_2°.

Soit H_2' une sous-algèbre de Hopf de H_2. La restriction à H_2' fournit un homomorphisme d'algèbres de Hopf α de H_1 dans ${H_2'}^\circ$. Alors $H_1' = H_1/\ker(\alpha)$ est isomorphe à une sous-algèbre de Hopf de type fini, dense, de ${H_2'}^\circ$. On en déduit un produit bilinéaire non dégénéré $< \cdot, \cdot >'$, tel que $(H_1', H_2', < \cdot, \cdot >')$ soit un objet de $\mathcal{C}_2$. En outre, la projection et l'injection canoniques donnent un épimorphisme[3] (p, i) de $(H_1, H_2, < \cdot, \cdot >)$ dans $(H_1', H_2', < \cdot, \cdot >')$.

Réciproquement, il est clair que tout épimorphisme de $(H_1, H_2, < \cdot, \cdot >)$ dans un objet de $\mathcal{C}_2$ s'obtient ainsi.

De même, la donnée d'une algèbre de Hopf H_1' quotient de H_1 suffit à déterminer un épimorphisme de $(H_1, H_2, < \cdot, \cdot >)$ vers un objet $(H_1', H_2', < \cdot, \cdot >')$ de $\mathcal{C}_2$: on prend pour H_2' l'ensemble des éléments de H_2 dont l'image dans H_1° appartient à ${H_1'}^\circ$.

Un épimorphisme (p, i) de $(H_1, H_2, < \cdot, \cdot >)$ vers $(H_1', H_2', < \cdot, \cdot >')$ est susceptible d'être une obstruction au "lemme de zéros" dans $(H_1, H_2, < \cdot, \cdot >)$. En effet, si $R_0' \in H_2'$ vérifie $< x', R_0' >= 0$ pour tout $x' \in H'_{1,S_1,T_1}$ alors $R_0 = i(R_0') \in H_2$ vérifie $< x, R_0 >= 0$ pour tout $x \in H_{1,S_1,T_1}$. Cette obstruction n'est pas intéressante quand H_1' est un espace vectoriel de dimension 1 (i.e. correspond au groupe trivial réduit à un élément). Dans ce cas, on dit que $(H_1, H_2, < \cdot, \cdot >)$ est trivial. Ceci équivaut à $\dim(H_2') = 1$.

Au paragraphe 2.4, on va traduire un énoncé général en un lemme de zéros, et en un lemme d'interpolation. Pour effectuer la traduction au niveau des sous-groupes obstructeurs, on a besoin d'expliciter concrètement ce qu'est un épimorphisme (p, i) de $(H_1, H_2, < \cdot, \cdot >)$ dans $(H_1', H_2', < \cdot, \cdot >')$. Pour cela, notons $H_2 = \mathbb{C}[G]$ et $H_1 = \mathbb{C}\Gamma \otimes \mathrm{Sym}(W)$ comme dans la première partie ; alors H_2' s'écrit $\mathbb{C}[G/G']$ pour un certain sous-groupe algébrique G' de G, et H_1' est isomorphe à un produit tensoriel $\mathbb{C}\Gamma' \otimes \mathrm{Sym}(W')$. La projection de H_1 dans H_1' induit un morphisme de groupes surjectif p_Γ de Γ dans Γ', et une application linéaire

[3] C'est-à-dire l'analogue, dans la catégorie $\mathcal{C}_2$, d'une surjection dans la catégorie des ensembles.

surjective p_W de W dans W'. Le noyau de p_Γ est $\Gamma \cap G'$, et celui de p_W est $W \cap TG'$. Ainsi, la donnée de l'épimorphisme (p, i) correspond à celle du sous-groupe G', avec $H'_2 = \mathbb{C}[G/G']$ et $H'_1 = \mathbb{C}_{\frac{\Gamma}{\Gamma \cap G'}} \otimes \mathrm{Sym}(W/W \cap TG')$. En outre, si $H'_1 = \mathbb{C}[\widetilde{G}']$ alors on a $\mathrm{rg}(\Gamma/\Gamma \cap G') = \dim(\widetilde{G}'_1)$ et $\dim(W/W \cap TG') = \dim(\widetilde{G}'_0)$.

La traduction précédente sera utilisée, au paragraphe 2.4, pour obtenir un lemme de zéros. Pour le lemme d'interpolation, la transformation de Fourier-Borel intervient. C'est pourquoi il faut plutôt considérer la situation suivante, où (p, i) désigne toujours un épimorphisme de $(H_1, H_2, < \cdot, \cdot >)$ dans $(H'_1, H'_2, < \cdot, \cdot >')$. Notons $H_1 = \mathbb{C}[\widetilde{G}]$ et $H_2 = \mathbb{C}\widetilde{\Gamma} \otimes \mathrm{Sym}(\widetilde{W})$ comme dans la première partie. Il existe un sous-groupe algébrique $\widetilde{G}'$ de $\widetilde{G}$ tel que $H'_1 = \mathbb{C}[\widetilde{G}']$. En outre H'_2 s'écrit $\mathbb{C}\widetilde{\Gamma}' \otimes \mathrm{Sym}(\widetilde{W}')$, où $\widetilde{\Gamma}'$ est un sous-groupe de $\widetilde{\Gamma}$ et $\widetilde{W}'$ est un sous-espace vectoriel de $\widetilde{W}$. On voit que $\widetilde{\Gamma}' = \widetilde{\Gamma} \cap \widetilde{G}'$ et $\widetilde{W}' = \widetilde{W} \cap T\widetilde{G}'$. En particulier, si $H'_2 = \mathbb{C}[K]$ pour un certain groupe algébrique affine commutatif K, alors $\mathrm{rg}(\widetilde{\Gamma} \cap \widetilde{G}') = \dim(K_1)$ et $\dim(\widetilde{W} \cap T\widetilde{G}') = \dim(K_0)$.

2.4 Énoncé du lemme de zéros

On a l'énoncé général suivant, qui porte sur la catégorie $\mathcal{C}_2$:

Théorème 2.1 *Soit $(H_1, H_2, < \cdot, \cdot >)$ un objet de $\mathcal{C}_2$, avec $H_1 = \mathbb{C}[\widetilde{G}]$ et $H_2 = \mathbb{C}[G]$. On suppose $\widetilde{G}_1$ et G_1 plongés dans des puissances de $\mathbb{G}_m$. Il existe une constante $c > 0$ ayant la propriété suivante. Soient S_1, S_2, T_1, T_2 des entiers strictement positifs, et $R \in H_{2,S_2,T_2}$ non nul tel que*

$$< x, R >= 0 \text{ pour tout } x \in H_{1,S_1,T_1}$$

Alors il existe un épimorphisme de $(H_1, H_2, < \cdot, \cdot >)$ dans un objet non trivial $(H'_1, H'_2, < \cdot, \cdot >')$, avec $H'_1 = \mathbb{C}[\widetilde{G}']$ et $H'_2 = \mathbb{C}[G/G']$, tel que :

$$T_2^{\dim(G_0/G'_0)} S_2^{\dim(G_1/G'_1)} > c\, T_1^{\dim(\widetilde{G}'_0)} S_1^{\dim(\widetilde{G}'_1)}$$

On peut traduire ce théorème de deux manières : comme un lemme de zéros ou comme un lemme d'interpolation. Ceci fournit une preuve du théorème 2.1, puisqu'on obtient un lemme de zéros connu.

Pour obtenir un lemme de zéros, on applique l'équivalence de catégories ci-dessus : on obtient un objet (G, Γ, W) de $\mathcal{C}_1$ qui correspond à $(H_1, H_2, < \cdot, \cdot >)$. La traduction du théorème 2.1 est alors la suivante (dans laquelle on pose $D_0 = T_2$, $D_1 = S_2$, $S = S_1$ et $T = T_1$ pour obtenir les notations habituelles) :

Théorème 2.2 *Soit (G, Γ, W) un objet de $\mathcal{C}_1$, muni d'une famille génératrice $(\gamma_1, \ldots, \gamma_l)$ de Γ et d'un plongement de G_1 dans une puissance de $\mathbb{G}_m$. Alors il existe une constante $c > 0$ ayant la propriété suivante. Soient D_0, D_1, S, T des entiers strictement positifs, et R un polynôme non nul sur G, de bidegré inférieur à (D_0, D_1) , qui s'annule à l'ordre T le long de W en tout point de $\Gamma(S)$. Alors il existe un sous-groupe algébrique G' de G, distinct de G, tel que :*

$$D_0^{\dim(G_0/G'_0)} D_1^{\dim(G_1/G'_1)} > c\, S^{\mathrm{rg}(\Gamma/\Gamma \cap G')} T^{\dim(W/W \cap TG')}$$

N.B. En général, on suppose G connexe dans un tel lemme de zéros. Ici ce n'est pas nécessaire, car on peut d'y ramener. En effet, en notant G° la composante neutre de G et $\Gamma^\circ = \Gamma \cap$

G° (muni d'une famille génératrice choisie arbitrairement), le lemme 3 de [M2] fournit une constante c' telle que pour tout $S \geq 0$ il existe $\alpha_S \in \Gamma$ tel que $\Gamma^\circ(S) \subset \alpha_S + \Gamma(c'S)$.

Pour obtenir un lemme d'interpolation, on échange les rôles de H_1 et H_2 dans le théorème 2.1 avant d'appliquer l'équivalence de catégories. Ceci revient à appliquer la transformation de Fourier-Borel au théorème 2.2. On obtient l'énoncé suivant :

Théorème 2.3 *Soit $(\widetilde{G}, \widetilde{\Gamma}, \widetilde{W})$ un objet de $\mathcal{C}_1$, muni d'une famille génératrice $(\widetilde{\gamma}_1, \dots, \widetilde{\gamma}_{d_1})$ de $\widetilde{\Gamma}$, d'une base de $\widetilde{W}$ et d'un plongement de $\widetilde{G}_1$ dans une puissance de $\mathbb{G}_m$. Il existe une constante $c > 0$ ayant la propriété suivante. Soient D_0, D_1, S, T des entiers strictement positifs. Soit η une fonctionnelle non nulle, combinaison linéaire des $\mathrm{ev}_{\gamma,\sigma}$ pour $\gamma \in \widetilde{\Gamma}(S)$ et $\sigma \in \mathbb{N}^{\dim(\widetilde{W})}$, $|\sigma| < T$, qui annule tous les polynômes sur $\widetilde{G}$ de bidegré inférieur à (D_0, D_1). Alors il existe un sous-groupe algébrique $\widetilde{G}'$ de $\widetilde{G}$, non nul, tel que :*

$$S^{\mathrm{rg}(\widetilde{\Gamma} \cap \widetilde{G}')} T^{\dim(\widetilde{W} \cap T\widetilde{G}')} > c\, D_0^{\dim(\widetilde{G}'_0)} D_1^{\dim(\widetilde{G}'_1)}$$

N.B. On peut énoncer ce théorème sans choisir de base de W : il suffit de ne pas mentionner les $\mathrm{ev}_{\gamma,\sigma}$ et d'écrire simplement $\eta \in (\mathbb{C}\Gamma \otimes \mathrm{Sym}(W))_{S,T}$.

3 Prolongements possibles

On pourrait étudier ce qui se passe quand l'un des deux groupes algébriques affines (voire les deux) est remplacé par un groupe algébrique quasi-projectif. C'est le cas, par exemple, dans [F6].

On pourrait aussi essayer de considérer des algèbres de Hopf plus générales, sans se restreindre à celles qui sont commutatives, co-commutatives et de type fini.

On pourrait, enfin, essayer de traduire en termes d'algèbres de Hopf des lemmes de zéros plus précis que le théorème 2.2.

Bibliographie

[A1] E. Abe – *Hopf algebras*, Cambridge Tracts in Math., Cambridge Univ. Press, 1977.

[A2] S. Ahlgren et K. Ono – *A Gaussian hypergeometric series evaluation and Apéry number congruences*, J. Reine Angew. Math. 518 (2000), 187-212.

[A3] K. Alladi et M. Robinson – *Legendre polynomials and irrationality*, J. Reine Angew. Math. 318 (1980), 137-155.

[A4] G. Almkvist et A. Granville – *Borwein and Bradley's Apéry-like formulae for* $\zeta(4n+3)$, Experiment. Math. 8.2 (1999), 197-203.

[A5] F. Amoroso – *Osservazioni sull'integrale "di Reyssat"*, manuscrit non publié.

[A6] Y. André – *G-functions and geometry*, Aspects of Math. E13, Vieweg, 1989.

[A7] R. André-Jeannin – *Irrationalité de la somme des inverses de certaines suites récurrentes*, C. R. Acad. Sci. Paris, Ser. I 308 (1989), 539-541.

[A8] G.E. Andrews – *The well-poised thread: an organized chronicle of some amazing summations and their implications*, Ramanujan J. 1.1 (1997), 7-23.

[A9] G.E. Andrews, R. Askey et R. Roy – *Special Functions*, The Encyclopedia of Mathematics and its Applications 71 (G.-C. Rota ed.), Cambridge University Press, Cambridge, 1999.

[A10] R. Apéry – *Irrationalité de* $\zeta(2)$ *et* $\zeta(3)$, in: Journées Arithmétiques (Luminy, 1978), Astérisque 61 (1979), 11-13.

[A11] R. Apéry – *Interpolation de fractions continues et irrationalité de certaines constantes*, in: Comité des Travaux Historiques et Scientifiques (CTHS), Bulletin de la Section des Sciences III (Mathématiques), Bibliothèque Nationale, Paris, 1981, 37-53.

[A12] R. Askey et J.A. Wilson – *A recursive relation generalizing those of Apéry*, J. Austral. Math. Soc. 36 (1984), 267-278.

[A13] W. Van Assche – *Approximation theory and analytic number theory*, in: Special Functions and Differential Equations (Madras, 1997), Allied Publishers, New Delhi, 1998, 336-355.

[B1] A. Baker et D.W. Masser, eds., – *Transcendence theory: advances and applications* (Cambridge, 1976), Academic Press, 1977.

[B2] K.M. Ball et T. Rivoal – *Irrationalité d'une infinité de valeurs de la fonction zêta aux entiers impairs*, Invent. Math. 146.1 (2001), 193-207.

[B3] C. Batut et M. Olivier – *Sur l'accélération de la convergence de certaines fractions continues*, Sém. de Théorie des Nombres de Bordeaux 1979-1980, exp. no. 23 (25 p.).

[B4] F. Beukers – *A note on the irrationality of* $\zeta(2)$ *and* $\zeta(3)$, Bull. London Math. Soc. 11.3 (1979), 268-272.

[B5] F. Beukers – *Padé-approximations in number theory*, in: Padé approximation and its applications (Amsterdam, 1980), Lecture Notes in Math. 888, Springer, 1981, 90-99.

[B6] F. Beukers – *The values of polylogarithms*, in: Topics in classical number theory (Budapest, 1981), Colloq. Math. Soc. János Bolyai 34, 1984, 219-228.

[B7] F. Beukers – *Irrationality of* π^2*, periods of an elliptic curve and* $\Gamma_1(5)$, in: Approximations diophantiennes et nombres transcendants (Luminy, 1982), D. Bertrand et M. Waldschmidt eds., Progress in Math. 31, Birkhäuser, 1983, 47-66.

[B8] F. Beukers – *Some congruences for the Apéry numbers*, J. Number Th. 21 (1985), 141-155.

[B9] F. Beukers – *Irrationality proofs using modular forms*, in: Journées Arithmétiques (Besançon, 1985), Astérisque 147-148 (1987), 271-283.

[B10] F. Beukers – *Another Congruence for the Apéry Numbers*, J. Number Th. 25 (1987), 201-210.

[B11] F. Beukers et C.A.M. Peters – *A family of K3 surfaces and* $\zeta(3)$, J. Reine Angew. Math. 351 (1984), 42-54.

[B12] J. Borwein et D. Bradley – *Empirically determined Apéry-like formulae for* $\zeta(4n+3)$, Experiment. Math. 6 (1997), 181-194.

[B13] P. Borwein et T. Erdélyi – *Polynomials and Polynomial inequalities*, Graduate Texts in Math. 161, Springer, 1995.

[B14] N. Bourbaki – *Algèbre*, Chapitre II, 3ème éd., Hermann, 1962.

[B15] N. Bourbaki – *Groupes et algèbres de Lie*, Chapitres II et III, Hermann, 1972.

[B16] N. Brisebarre – *Irrationality measures of* $\log(2)$ *and* $\pi/\sqrt{3}$, Experimental Math. 10.1 (2001), 35-51.

[B17] P. Bundschuh et K. Väänänen – *Arithmetical investigations of a certain infinite product*, Compositio Math. 91 (1994), 175-199.

[C1] H. Cartan – *Formes différentielles*, Hermann, 1967.

[C2] P. Cartier – *Démonstration automatique d'identités et fonctions hypergéométriques (d'après Zeilberger)*, Sém. Bourbaki 1991-92, exp. no. 746, Astérisque 206 (1992), 41-91.

[C3] P. Cartier – *Fonctions polylogarithmes, nombres polyzêtas et groupes pro-unipotents*, Sém. Bourbaki 2000-01, exp. no. 885, à paraître dans Astérisque.

[C4] S. Chowla, J. Cowles et M. Cowles – *Congruence properties of Apéry numbers*, J. Number Th. 12 (1980), 188-190.

[C5] G.V. Chudnovsky – *Transcendental numbers*, in: Number theory, Proc. Southern Illinois Conf. (Carbondale, 1979), Lecture Notes in Math. 751, Springer, 45-69.

[C6] H. Cohen – *Démonstration de l'irrationalité de* $\zeta(3)$ *(d'après Apéry)*, Sém. de Théorie des Nombres de Grenoble, octobre 1978 (9 p.).

[C7] H. Cohen – *Généralisation d'une construction de R. Apéry*, Bull. Soc. Math. France 109 (1981), 269-281.

[C8] P. Colmez – *Arithmétique de la fonction zêta*, Journées X-UPS 2002, à paraître.

[C9] E.T. Copson – *Asymptotic expansions*, Cambridge University Press, 1967.

[D1] M. Demazure et P. Gabriel – *Groupes algébriques*, Tome I, Masson, 1970.

[D2] J. Dieudonné – *Calcul infinitésimal*, Collection Méthodes, Hermann, 1968.

[D3] A.C. Dixon – *On a certain double integral*, Proc. London Math. Soc. (2) 2 (1905), 8-15.

[D4] B. Dwork – *On Apéry's differential operator*, Groupe d'étude d'analyse ultramétrique 1979-1981, exp. no. 25 (6 p.).

[D5] B. Dwork – *Arithmetic theory of differential equations*, Symposia Math. 24 (INDAM, Rome, 1979), Academic Press, 1981, 225-243.

[D6] B. Dwork, G. Gerotto et F.J. Sullivan – *An introduction to G-functions*, Annals of Math. Studies 133, Princeton Univ. Press, 1994.

[F1] N.I. Fel'dman et Yu.V. Nesterenko – *Number Theory IV, Transcendental Numbers*, A.N. Parshin et I.R. Shafarevich eds., Encyclopaedia of Mathematical Sciences 44, Springer, 1998.

[F2] S. Fischler – *Formes linéaires en polyzêtas et intégrales multiples*, C. R. Acad. Sci. Paris, Ser. I 335.1 (2002), 1-4 ; chapitre 3 de cette thèse.

[F3] S. Fischler – *Groupes de Rhin-Viola et intégrales multiples*, Actes des Rencontres Arithmétiques de Caen (juin 2001), à paraître au J. Théor. Nombres Bordeaux ; chapitre 4 de cette thèse.

[F4] S. Fischler – *Irrationalité de valeurs de zêta (d'après Apéry, Rivoal, ...)*, Séminaire Bourbaki 2002-2003, exposé no. 910 ; chapitre 1 de cette thèse.

[F5] S. Fischler et T. Rivoal – *Approximants de Padé et séries hypergéométriques équilibrées*, Rapport de Recherche 2002-27, LMNO, Université de Caen, à paraître au J. Math. Pures Appl.; chapitre 5 de cette thèse.

[F6] S. Fischler – *Lemme d'interpolation avec multiplicités*, chapitre 7 de cette thèse.

[F7] S. Fischler – *Lemmes de zéros et algèbres de Hopf*, chapitre 8 de cette thèse.

[G1] A.O. Gel'fond – *Calcul des différences finies*, Dunod, 1963.

[G2] I. Gessel – *Some congruences for Apéry numbers*, J. Number Th. 14 (1982), 362-368.

[G3] L.A. Gutnik – *The irrationality of certain quantities involving $\zeta(3)$*, Uspekhi Mat. Nauk [Russian Math. Surveys] 34.3 (1979), 190 [200].

[G4] L.A. Gutnik – *On the irrationality of some quantities containing $\zeta(3)$*, Acta Arith. 42.3 (1983), 255-264 (en russe) ; traduction dans Amer. Math. Soc. Transl. 140 (1988), 45-55.

[H1] L. Habsieger – *Introduction to diophantine approximation*, en préparation.

[H2] G.H. Hardy et E.M. Wright – *An introduction to the theory of numbers*, 3ème édition, Oxford Univ. Press, 1954.

[H3] J. Harris – *Algebraic Geometry, A First Course*, Graduate Texts in Math. 133, Springer-Verlag, 1992.

[H4] M. Hata – *On the linear independence of the values of polylogarithmic functions*, J. Math. Pures Appl. 69.2 (1990), 133-173.

[H5] M. Hata – *Rational approximations to the dilogarithm*, Trans. Amer. Math. Soc. 336.1 (1993), 363-387.

[H6] M. Hata – *A new irrationality measure for* $\zeta(3)$, Acta Arith. 92.1 (2000), 47-57.

[H7] M. Hazewinkel – *Formal groups and applications*, Pure and Applied Mathematics 78, Academic Press, 1978.

[H8] T.G. Hessami Pilehrood – *Linear independence of vectors with polylogarithmic coordinates*, Vestnik Moskov. Univ. Ser. I Mat. Mekh. no. 6 [Moscow Univ. Math. Bull. 54.6] (1999), 54-56 [40-42].

[H9] G. Hochschild – *Introduction to affine algebraic groups*, Holden-Day, 1971.

[H10] J. E. Humphreys – *Linear Algebraic Groups*, Graduate Texts in Math. 21, Springer-Verlag, 1975.

[H11] M. Huttner – *Équations différentielles fuchsiennes. Approximations du dilogarithme, de* $\zeta(2)$ *et de* $\zeta(3)$, Pub. IRMA Lille 43 (1997).

[H12] M. Huttner – *Constructible sets of linear differential equations and effective rational approximations of* G*-functions*, Pub. IRMA Lille 59 (2002).

[I1] E.L. Ince – *Ordinary differential equations*, Dover Publ., 1926.

[I2] A.E. Ingham – *The distribution of prime numbers*, Cambridge Univ. Press, 1932.

[I3] T. Ishikawa – *On Beukers' conjecture*, Kobe J. Math. 6 (1989), 49-52.

[K1] A. Ya. Khintchine – *Continued fractions*, Noordhoff, Groningen, 1963.

[K2] M. Koecher – *Letter*, Math. Intelligencer 2 (1980), 62-64.

[K3] M. Kontsevich et D. Zagier – *Periods*, in: Mathematics Unlimited - 2001 and beyond, Springer, 2001, 771-808.

[K4] C. Krattenthaler – Communication personnelle, 28 Octobre 2002.

[L1] S. Lang – *Introduction to transcendental numbers*, Addison-Wesley, 1966.

[L2] S. Lang – *Algebra*, 3ème édition, Addison-Wesley, 1993.

[L3] D. Leshchiner – *Some new identities for* $\zeta(k)$, J. Number Th. 13 (1981), 355-362.

[L4] L. Lewin – *Polylogarithms and Associated Functions*, North-Holland, New York, 1981.

[L5] J. Liouville – *Sur des classes très étendues de quantités dont la valeur n'est ni algébrique, ni même réductible à des irrationnelles algébriques*, J. Math. Pures Appl. 16 (1851), 133-142.

[L6] Y.L. Luke – *The special functions and their approximations*, Volume I, Mathematics in Science and Engineering 53, Academic Press, 1969.

[M1] S. Mac Lane – *Categories for the working mathematician*, Graduate Texts in Math. 5, Springer, 1971.

[M2] D.W. Masser – *Interpolation on group varieties*, in: Approximations diophantiennes et nombres transcendants (Luminy, 1982), D. Bertrand et M. Waldschmidt eds., Progress in Math. 31, Birkhäuser, 1983, 151-171.

[M3] D.W. Masser et G. Wüstholz – *Zero estimates on group varieties I*, Invent. Math. 64 (1981), 489-516.

[M4] M. Mendès-France – *Roger Apéry et l'irrationnel*, La Recherche 97 (1979), 170-172.

[M5] C. Mutafian – *Algebra Multilineal*, Instituto Cubano del Libro, 1974.

[N1] Yu.V. Nesterenko – *On the linear independence of numbers*, Vestnik Moskov. Univ. Ser. I Mat. Mekh. no. 1 [Moscow Univ. Math. Bull. 40.1] (1985), 46-49 [69-74].

[N2] Yu.V. Nesterenko – *A few remarks on $\zeta(3)$*, Mat. Zametki [Math. Notes] 59.6 (1996), 865-880 [625-636].

[N3] Yu.V. Nesterenko – *Integral identities and constructions of approximations to zeta-values*, Actes des Rencontres Arithmétiques de Caen (juin 2001), soumis au J. Théor. Nombres Bordeaux.

[N4] E.M. Nikishin – *On the irrationality of the values of the functions F(x,s)*, Mat. Sbornik 109.3 [Math. USSR-Sb. 37.3] (1979), 410-417 [381-388].

[N5] E.M. Nikishin et V.N. Sorokin – *Rational approximations and orthogonality*, Translations of Math. Monographs 92, Amer. Math. Soc., 1991.

[N6] I. Niven, H.S. Zuckerman et H.L. Montgomery – *An introduction to the theory of numbers*, 5ème édition, J. Wiley, 2000.

[O1] J. Oesterlé – *Polylogarithmes*, Sém. Bourbaki 1992-93, exp. no. 762, Astérisque 216 (1993), 49-67.

[P1] C. Peters et J. Stienstra – *A pencil of K3-surfaces related to Apéry's recurrence for $\zeta(3)$ and Fermi surfaces for potential zero*, in: Arithmetics of complex manifolds (Erlangen, 1988), W.P. Barth et H. Lange eds., Lecture Notes in Math. 1399, Springer, 110-127.

[P2] M. Petkovšek, H.S. Wilf et D. Zeilberger – *A=B*, A.K. Peters, 1996.

[P3] P. Philippon – *Lemmes de zéros dans les groupes algébriques commutatifs*, Bull. Soc. Math. France 114 (1986), 355-383 ; errata et addenda, 115 (1987), 397-398.

[P4] A. Van Der Poorten – *A proof that Euler missed... Apéry's proof of the irrationality of $\zeta(3)$*, Math. Intelligencer 1.4 (1978/79), 195-203.

[P5] A. Van Der Poorten – *Some wonderful formulae... footnotes to Apéry's proof of the irrationality of $\zeta(3)$*, Sém. Delange-Pisot-Poitou, 20e année, 1978-79, no. 29 (7p.).

[P6] A. Van Der Poorten – *Some wonderful formulas ... an introduction to polylogarithms*, in: Proceedings of the Queen's Number Theory Conference (Kingston, 1979), Queen's Papers in Pure and Applied Mathematics 54, 1980, 269-286.

[P7] M. Prevost – *A new proof of the irrationality of $\zeta(2)$ and $\zeta(3)$ using Padé approximants*, J. Comp. Appl. Math. 67 (1996), 219-235.

[P8] M. Prevost – *On the irrationality of* $\sum \frac{t^n}{A\alpha^n+B\beta^n}$, J. Number Th. 73 (1998), 139-161.

[R1] E. Reyssat – *Irrationalité de $\zeta(3)$ selon Apéry*, Sém. Delange-Pisot-Poitou, 20e année, 1978-79, no. 6 (6 p.).

[R2] E. Reyssat – *Mesures de transcendance pour les logarithmes de nombres rationnels*, in: Approximations diophantiennes et nombres transcendants (Luminy, 1982), D. Bertrand et M. Waldschmidt eds., Progress in Math. 31, Birkhäuser, 1983, 235-245.

[R3] G. Rhin et C. Viola – *On a permutation group related to $\zeta(2)$*, Acta Arith. 77.1 (1996), 23-56.

[R4] G. Rhin et C. Viola – *The group structure for $\zeta(3)$*, Acta Arith. 97.3 (2001), 269-293.

[R5] T. Rivoal – *La fonction zêta de Riemann prend une infinité de valeurs irrationnelles aux entiers impairs*, C. R. Acad. Sci. Paris, Sér. I 331.4 (2000), 267-270.

[R6] T. Rivoal – *Propriétés diophantiennes des valeurs de la fonction zêta de Riemann aux entiers impairs*, thèse de doctorat, Univ. de Caen (2001). Disponible sur `http://theses-EN-ligne.in2p3.fr`.

[R7] T. Rivoal – *Irrationalité d'au moins un des neuf nombres* $\zeta(5)$, $\zeta(7)$, $\ldots$, $\zeta(21)$, Acta Arith. 103.2 (2002), 157-167.

[R8] T. Rivoal – *Séries hypergéométriques et irrationalité des valeurs de la fonction zêta de Riemann*, Actes des Journées Arithmétiques de Lille (juillet 2001), à paraître au J. Théor. Nombres Bordeaux.

[R9] T. Rivoal et W. Zudilin – *Diophantine properties of numbers related to Catalan's constant*, Prépublication 315 de l'Institut de Mathématiques de Jussieu (janvier 2002), Univ. Paris 6, soumis.

[R10] D. Roy – *Sur la conjecture de Schanuel pour les logarithmes de nombres algébriques*, Problèmes Diophantiens 1988-89, Publ. Math. Univ. Paris VI 90.

[R11] D. Roy – *Matrices whose coefficients are linear forms in logarithms*, J. Number Th. 41 (1992), 22-47.

[R12] D. Roy – *Transcendance et questions de répartition dans les groupes algébriques*, in: Approximations diophantiennes et nombres transcendants (Luminy, 1990), de Gruyter, 1992, 249-274.

[R13] D. Roy – *Points whose coordinates are logarithms of algebraic numbers on algebraic varieties*, Acta Math. 175 (1995), 49-73.

[R14] D. Roy – *Zero estimates on commutative algebraic groups*, in: Introduction to algebraic independence theory, Yu. V. Nesterenko et P. Philippon eds., Lecture Notes in Math. 1752, Springer, 2001, 167-186.

[R15] D. Roy et M. Waldschmidt – *Approximation diophantienne et indépendance algébrique de logarithmes*, Ann. Sci. École Norm. Sup., Série 4, 30 (1997), 753-796.

[S1] J.P. Serre – *Cours d'arithmétique*, Presses Univ. de France, 1970.

[S2] A.B. Shidlovskii – *Transcendental numbers*, de Gruyter Studies in Math. 12, Berlin, 1989.

[S3] C.L. Siegel – *Transcendental numbers*, Annals of Math. Studies 16, Princeton Univ. Press, Princeton, 1949.

[S4] L.J. Slater – *Generalized hypergeometric functions*, Cambridge Univ. Press, 1966.

[S5] V.N. Sorokin – *Hermite-Padé approximations for Nikishin systems and the irrationality of* $\zeta(3)$, Uspekhi Mat. Nauk [Russian Math. Surveys] 49.2 (1994), 167-168 [176-177].

[S6] V.N. Sorokin – *A transcendence measure for* π^2, Mat. Sbornik [Sb. Math.] 187.12 (1996), 87-120 [1819-1852].

[S7] V.N. Sorokin – *Apéry's theorem*, Vestnik Moskov. Univ. Ser. I Mat. Mekh. no. 3 [Moscow Univ. Math. Bull. 53.3] (1998), 48-53 [48-52].

[S8] J. Stienstra et F. Beukers – *On the Picard-Fuchs equation and the formal Brauer group of certain elliptic K3-surfaces*, Math. Ann. 271 (1985), 269-304.

[S9] B. Sury – *On a conjecture of Chowla et al.*, J. Number Th. 72 (1998), 137-139.

[V1] O.N. Vasilenko – *Certain formulae for values of the Riemann zeta function at integral points*, in: Number theory and its applications, Proceedings of the science-theoretical conference (Tashkent, 1990), p. 27 (en russe).

[V2] D.V. Vasilyev – *Some formulas for Riemann zeta-function at integer points*, Vestnik Moskov. Univ. Ser. I Mat. Mekh. no. 1 [Moscow Univ. Math. Bull. 51.1] (1996), 81-84 [41-43].

[V3] D.V. Vasilyev – *On small linear forms for the values of the Riemann zeta-function at odd integers* (en russe), Doklady NAN Belarusi (Reports of the Belarus National Academy of Sciences) 45.5 (2001), 36-40.

[W1] M. Waldschmidt – *Nombres transcendants et groupes algébriques*, Astérisque 69-70, Soc. Math. France, 1979.

[W2] M. Waldschmidt – *Transcendance et exponentielles en plusieurs variables*, Invent. Math. 63 (1981), 97-127.

[W3] M. Waldschmidt – *Dépendance de logarithmes dans les groupes algébriques*, in: Approximations diophantiennes et nombres transcendants (Luminy, 1982), D. Bertrand et M. Waldschmidt eds., Progress in Math. 31, Birkhäuser, 1983, 289-328.

[W4] M. Waldschmidt – *La transformation de Fourier-Borel : une dualité en transcendance*, exposé donné à Delphes le 29/9/1989, `http://www.math.jussieu.fr/~miw`

[W5] M. Waldschmidt – *Fonctions auxiliaires et fonctionnelles analytiques I, II*, J. Analyse Math. 56 (1991), 231-254 et 255-279.

[W6] M. Waldschmidt – *Un demi-siècle de transcendance*, in: Development of Mathematics 1950-2000, J.-P. Pier ed., Birkhäuser Verlag, 2000, 1121-1186.

[W7] M. Waldschmidt – *Valeurs zêta multiples : une introduction*, J. Théor. Nombres Bordeaux 12.2 (2000), 581-595.

[W8] M. Waldschmidt – *Diophantine approximation on linear algebraic groups: transcendence properties of the exponential function in several variables*, Grundlehren Math. Wiss. 326, Springer, 2000.

[W9] A. Weil – *Remarks on Hecke's lemma and its use*, in: Oeuvres scientifiques - Collected Papers III, Springer, 1979, 405-412.

[W10] E.T. Whittaker et G.N. Watson – *A course of modern analysis*, 4ème édition, Cambridge Univ. Press, 1927.

[W11] G. Wüstholz – *Multiplicity estimates on group varieties*, Annals of Math. 129 (1989), 471-500.

[Z1] D. Zagier – *Introduction to modular forms*, in: From number theory to physics (Les Houches, 1989), M. Waldschmidt, P. Moussa, J.M. Luck et C. Itzykson eds., Springer, 1992, 238-291.

[Z2] D. Zagier – Cours au Collège de France, mai 2001.

[Z3] D. Zeilberger – *Closed form (pun intended !)*, in: A tribute to Emil Grosswald: Number theory and related analysis, Comtemporary Math. 143 (M. Knopp et M. Sheingorn eds.), Amer. Math. Soc., 1993, 579-607.

[Z4] D. Zeilberger – *Computerized deconstruction*, à paraître dans Adv. Applied Math.

[Z5] S.A. ZLOBIN – *Integrals expressible as linear forms in generalized polylogarithms*, Mat. Zametki [Math. Notes] 71.5 (2002), 782-787 [711-716].

[Z6] W. ZUDILIN – *One of the numbers $\zeta(5)$, $\zeta(7)$, $\zeta(9)$, $\zeta(11)$ is irrational*, Uspekhi Mat. Nauk [Russian Math. Surveys] 56.4 (2001), 149-150 [774-776].

[Z7] W. ZUDILIN – *Irrationality of values of the Riemann zeta function*, Izvestiya RAN Ser. Mat. [Izv. Math.] 66.3 (2002), 49-102 [489-542].

[Z8] W. ZUDILIN – *Well-poised hypergeometric service for diophantine problems of zeta values*, Actes des Rencontres Arithmétiques de Caen (juin 2001), soumis au J. Théor. Nombres Bordeaux.

[Z9] W. ZUDILIN – *Arithmetic of linear forms involving odd zeta values*, preprint, `math.NT/0206176`.

[Z10] W. ZUDILIN – *An elementary proof of Apéry's theorem*, preprint, `math.NT/0202159`.