

GROUPES FORMELS, FONCTIONS AUTOMORPHES ET FONCTIONS ZETA DES COURBES ELLIPTIQUES

par P. CARTIER

A ANDRÉ WEIL et JEAN DIEUDONNÉ,
dont les travaux ont été notre source
d'inspiration constante et féconde

1. Congruences pour les coefficients des fonctions automorphes.

Nous allons rappeler quelques-unes des remarquables congruences satisfaites par les coefficients des formes modulaires, et qui ont été découvertes par Ramanujan, Newman, Atkin, O'Brien et Swinnerton-Dyer (voir Atkin [1] pour les détails). Considérons d'abord la forme modulaire Δ de poids 12 (discriminant) :

$$(1) \quad \Delta(\tau) = e^{2\pi i\tau} \prod_{n=1}^{\infty} (1 - e^{2\pi in\tau})^{24} = \sum_{n=1}^{\infty} \tau(n) \cdot e^{2\pi in\tau}$$

Soit p un nombre premier ; les coefficients $\tau(n)$ de Δ satisfont à la relation de Ramanujan-Mordell⁽¹⁾ :

$$(2) \quad \tau(np) - \tau(p) \cdot \tau(n) + p^{11} \cdot \tau(n/p) = 0$$

pour tout entier $n \geq 1$. Sous l'hypothèse $\tau(p) \not\equiv 0 \pmod{p}$, on déduit de cette égalité des congruences comme suit : définissons par récurrence les nombres rationnels p -entiers B_α par $B_1 = \tau(p)$ et $B_{\alpha+1} = \tau(p) - p^{11}/B_\alpha$; on a alors

$$(3) \quad \tau(np^\alpha) \equiv B_\alpha \cdot \tau(np^{\alpha-1}) \pmod{p^{11\alpha}}$$

pour $n \geq 1$ et $\alpha \geq 1$. On notera qu'il existe une unique unité p -adique B satisfaisant à l'équation $B^2 - \tau(p) \cdot B + p^{11} = 0$ et qu'on a $B_\alpha \equiv B \pmod{p^{11\alpha}}$ pour tout $\alpha \geq 1$; on peut donc remplacer B_α par B dans (3), à condition de se placer dans le domaine des nombres p -adiques.

Considérons par ailleurs les coefficients $c(n)$ définis par $j(\tau) = \sum_{n=-1}^{\infty} c(n) e^{2\pi in\tau}$,

où j est l'invariant modulaire elliptique de poids 0 bien connu. En 1968, Atkin a obtenu le résultat suivant, qui généralise et résume une longue suite de résultats partiels : étant donné un entier $\alpha \geq 1$, on pose $t(n) = c(\mathcal{Q}^\alpha n)/c(\mathcal{Q}^\alpha)$; on a alors les relations

(1) Nous faisons la convention que $\tau(a)$ est nul si a n'est pas entier ; on fera des conventions analogues pour $t(a)$ dans (4), pour $\beta(a)$ dans (9), etc . . .

$$(4) \quad t(np) - t(p) \cdot t(n) + p^{-1} \cdot t(n/p) \equiv 0 \pmod{\ell^\alpha}$$

$$(5) \quad t(n\ell) = t(n) \cdot t(\ell)$$

($n \geq 1$, p premier $\neq \ell$) lorsque $\ell = 13$ et α quelconque ou lorsque $\ell = 17, 19, 23$ et α assez petit. Atkin a formulé une conjecture précise pour le cas des nombres premiers ℓ quelconques [1].

Le troisième exemple que nous considérerons se réfère à des formes modulaires de poids 2, c'est-à-dire à des formes différentielles de première espèce sur des courbes modulaires. D'une manière plus générale (cf. n° 5 pour le rapport entre ces deux points de vue), considérons une cubique plane C d'équation non homogène $Y^2 = X^3 - aX - b$ avec a et b entiers. Choisissons au voisinage du point à l'infini

de C un paramètre local ξ tel que l'on ait $X = \xi^{-2} + \sum_{n=1}^{\infty} \alpha(n) \cdot \xi^n$ avec des coefficients $\alpha(n)$ entiers ; la forme différentielle de première espèce $\omega = -dX/2Y$ sur C se développe sous la forme $\omega = \sum_{n=1}^{\infty} \beta(n) \cdot \xi^{n-1} d\xi$ avec des coefficients $\beta(n)$ entiers,

et $\beta(1) = 1$. Soit p un nombre premier différent de 2 et 3 ; Atkin et Swinnerton-Dyer⁽¹⁾ ont établi les congruences suivantes :

$$(6) \quad \beta(np) \equiv \beta(n) \cdot \beta(p) \pmod{p}$$

$$(7) \quad \beta(p) \equiv \sum_{t \pmod{p}} - \left(\frac{t^3 - at - b}{p} \right) \pmod{p},$$

où $\left(\frac{a}{p} \right)$ est le symbole de Legendre. Supposons qu'on ait $\beta(p) \not\equiv 0 \pmod{p}$, c'est-à-dire que la réduction de C modulo p soit d'invariant de Hasse-Witt non nul ; il existe alors une suite $(k_\alpha)_{\alpha \geq 1}$ de nombres entiers tels que

$$(8) \quad \beta(np^\alpha) \equiv k_\alpha \beta(np^{\alpha-1}) \pmod{p^\alpha} \text{ pour tout } n \geq 1.$$

L'analogie avec la démonstration de (3) à partir de (2) a conduit Atkin et Swinnerton-Dyer à postuler une congruence de la forme

$$(9) \quad \beta(np) - \beta(p) \cdot \beta(n) + p \cdot \beta(n/p) \equiv 0 \pmod{p^\alpha}$$

pour tout entier $n \equiv 0 \pmod{p^{\alpha-1}}$, y compris lorsque $\beta(p) \equiv 0 \pmod{p}$.

Il semble prématuré de faire des conjectures précises contenant tous ces cas particuliers (et d'autres analogues). Le schéma général semble être le suivant : on considère

une certaine forme modulaire de poids $2g$, soit $h(\tau) = \sum_{n=1}^{\infty} r(n) \cdot e^{2\pi i n \tau}$, avec des

coefficients $r(n)$ entiers, normalisée par $r(1) = 1$; on est en droit d'attendre des congruences de la forme

 (1) A notre connaissance, les résultats d'Atkin et Swinnerton-Dyer n'ont pas encore été publiés et sont contenus dans la correspondance échangée entre ces auteurs et Serre. Nous remercions Serre qui, en nous communiquant cette correspondance et en nous obligeant à répondre à ses questions pertinentes, a été à l'origine des résultats exposés ici.

$$(10) \quad r(np) - r(p) \cdot r(n) + p^{2g-1} \cdot r(n/p) \equiv 0 \pmod{p^{(2g-1)\alpha}}$$

lorsque p est premier et $n \equiv 0 \pmod{p^{\alpha-1}}$. Rappelons que la relation (2) de Ramanujan-Mordell signifie que Δ est fonction propre de l'opérateur de Hecke T_p . Par analogie, les résultats sur l'invariant modulaire elliptique j suggèrent la possibilité suivante : soit ℓ premier ; à l'aide des coefficients de Fourier de certaines formes modulaires de poids 0, on pourrait définir une "cohomologie étale ℓ -adique" qui serait un module libre H_ℓ de rang $[\ell/12]$ sur l'anneau \mathbb{Z}_ℓ des entiers ℓ -adiques et un opérateur de Hecke $T_{p,\ell}$ dans H_ℓ pour tout nombre premier $p \neq \ell$. Par contre, les congruences sur les courbes elliptiques suggèrent la possibilité dans certains cas de définir un opérateur de Hecke $T_{p,p}$ dans un module de cohomologie p -adique H'_p analogue à la cohomologie de Washnitzer-Monsky.

2. Groupes p -adiques rigides.

La suite de cet exposé est motivée par les congruences d'Atkin et Swinnerton-Dyer pour les différentielles de première espèce sur les courbes elliptiques. Le cadre naturel semble celui des groupes p -adiques rigides, dont nous empruntons la définition (en la simplifiant pour notre usage) à Tate [6]. Notons p un nombre premier, \mathfrak{o} ou \mathbb{Z}_p l'anneau des entiers p -adiques et K ou \mathbb{Q}_p le corps des fractions de \mathfrak{o} . Pour tout entier $n \geq 0$, on note D^n l'ensemble des vecteurs à n composantes dans \mathfrak{o} divisibles par p , et \mathfrak{A}_n la \mathfrak{o} -algèbre des fonctions sur D^n de la forme

$$f(x) = \sum_{i_1, \dots, i_n} a(i_1, \dots, i_n) \cdot x_1^{i_1} \dots x_n^{i_n}$$

(les coefficients $a(i_1, \dots, i_n)$ étant pris dans \mathfrak{o}). Une *variété rigide de dimension n* est un couple $(X, \mathfrak{A}(X))$ isomorphe à (D^n, \mathfrak{A}_n) ; un système de coordonnées rigide sur X est une suite (ξ_1, \dots, ξ_n) d'éléments de $\mathfrak{A}(X)$ telle que l'application $x \mapsto (\xi_1(x), \dots, \xi_n(x))$ soit un isomorphisme de X sur D^n . Une variété rigide X porte une structure de variété analytique sur le corps K pour laquelle tout système de coordonnées rigide est un système de coordonnées analytique ; les éléments de $\mathfrak{A}(X)$ sont *certaines* fonctions analytiques sur X , qualifiées de *rigides*⁽¹⁾. A partir des fonctions analytiques rigides sur X , on pourra définir les champs de vecteurs (ou les formes différentielles) rigides.

Les variétés rigides forment une catégorie avec produit, et l'on peut par suite définir la notion de groupe p -adique rigide. Deux exemples de tels groupes sont le groupe additif G_a , ayant D^1 pour variété sous-jacente, et l'addition pour opération, et le groupe multiplicatif G_m qui se compose du groupe multiplicatif des $x \equiv 1 \pmod{p}$ dans \mathfrak{o} , avec la coordonnée rigide ξ donnée par $\xi(x) = x - 1$.

Dans la suite, nous désignerons par G un groupe p -adique rigide de dimension 1 (nécessairement commutatif) ; les formes différentielles rigides de degré 1 sur G

 (1) Rappelons qu'une fonction qui est localement égale à une fonction analytique est analytique. Par contre, une fonction qui appartient localement à $\mathfrak{A}(X)$ n'appartient pas nécessairement à $\mathfrak{A}(X)$, d'où la terminologie : "rigide".

invariantes par translation forment un \mathfrak{o} -module libre de rang 1, dont nous choisirons une base ω_0 . Alors ω_0 est la différentielle $d\ell$ d'une fonction analytique ℓ sur G , appelée le *logarithme de G* . Ce logarithme est un isomorphisme de groupes de Lie p -adiques de G sur G_a , mais n'est pas en général une fonction analytique rigide. Pour préciser ce point, introduisons les opérateurs de Lazard Ψ_n ($n \geq 1$) dans $\mathfrak{X}(G)$ par

$$(11) \quad \Psi_n f(x) = \sum_{i=0}^n (-1)^{n-i} \binom{n}{i} f(x^i) \quad ;$$

si ξ est une coordonnée rigide dans G , normalisée par $\omega_0 = d\xi$ à l'origine, on a

$$(12) \quad \ell(x) = \sum_{n=1}^{\infty} (-1)^{n-1} \Psi_n \xi(x)/n.$$

Cette formule de Lazard permet le contrôle des dénominateurs dans ℓ ; lorsque $G = G_m$, $\xi(x) = x - 1$ et $\omega_0 = dx/x$, on a $\Psi_n \xi = \xi^n$ et (12) redonne le développement en série classique du logarithme usuel.

Le lien avec les groupes formels est le suivant. Choisissons une coordonnée rigide ξ sur G ; il existe alors une série formelle $F \in \mathfrak{o}[[X, X']]$ caractérisée par $\xi(xx') = F(\xi(x) ; \xi(x'))$ pour x, x' dans G ("Théorème d'addition"). Cette série satisfait aux identités

$$(13) \quad F(X ; 0) = F(0 ; X) = X \quad , \quad F(X ; Y) = F(Y ; X) \quad , \\ F(F(X ; Y) ; Z) = F(X ; F(Y ; Z)) \quad ;$$

autrement dit, c'est une loi de groupe formel commutatif à coefficients dans \mathfrak{o} .

3. Classification des groupes p -adiques rigides.

Le théorème de classification repose sur deux notions essentielles : la hauteur et le module différentiel. Soit G un groupe p -adique rigide de dimension 1. L'anneau $\mathfrak{X} = \mathfrak{X}(G)$ est local, et son idéal maximal \mathfrak{m} se compose des fonctions analytiques rigides dont les valeurs sont divisibles par p en tout point de G . La *hauteur* de G est la borne supérieure (finie ou non) $ht(G)$ des entiers $h \geq 1$ tels que

$$\Psi_p(\mathfrak{X}) \subset p \cdot \mathfrak{X} + \mathfrak{m}^h.$$

On a $ht(G_m) = 1$ et $ht(G_a) = \infty$; la formule (12) montre facilement que tout groupe de hauteur infinie est isomorphe, comme groupe p -adique rigide, à G_a .

Une *courbe* dans G est un morphisme de variétés rigides $\gamma : D^1 \rightarrow G$, normalisé par $\gamma(0) = e$ (élément neutre de G). Les courbes forment un groupe commutatif $C(G)$ pour l'addition définie par $(\gamma + \gamma')(t) = \gamma(t) \cdot \gamma'(t)$. Pour tout nombre premier ℓ , l'opérateur de décalage dans $C(G)$ est défini par $V_\ell \gamma(t) = \gamma(t^\ell)$, et l'opérateur de Frobenius par $F_\ell \gamma(t) = \prod_{i=1}^{\ell} \gamma(t^i t^{1/\ell})$. Dans cette dernière formule, ζ est une racine ℓ -ième de l'unité, distincte de 1, que l'on adjoint à \mathfrak{o} ainsi que la racine $t^{1/\ell}$ de t , mais le résultat de la multiplication se trouve définit sur 0.

Notons maintenant t la coordonnée naturelle sur D^1 et Ω le \mathfrak{o} -module des formes différentielles rigides sur D^1 ; nous représenterons toujours celles-ci sous la forme

$$(14) \quad \omega = \sum_{n=1}^{\infty} a(n) \cdot t^{n-1} dt \quad (a(n) \in \mathfrak{o} \text{ pour tout } n \geq 1) \quad ;$$

enfin, soit $d\mathfrak{A}_1$ l'ensemble des différentielles des fonctions $f \in \mathfrak{A}_1$. L'application $\gamma \mapsto \gamma^*(\omega_0)$ définit un isomorphisme u du groupe $C(G)$ des courbes de G sur un sous-groupe $\mathfrak{D}(G)$ de Ω . On dit que $\mathfrak{D}(G)$ est le *module différentiel* de G ; il caractérise G à un isomorphisme rigide près. De plus, u transforme V_ℓ et F_ℓ en les opérateurs suivants sur $\mathfrak{D}(G)$:

$$(15) \quad V_\ell \omega = \sum_{n=1}^{\infty} \ell \cdot a(n/\ell) \cdot t^{n-1} dt \quad , \quad F_\ell \omega = \sum_{n=1}^{\infty} a(n\ell) \cdot t^{n-1} dt$$

(pour ω de la forme (13)).

Soit F la loi de groupe formel définie à la fin du n° 2, et soit $F_{(p)}$ la loi de groupe formel à coefficients dans le corps $F_p = \mathfrak{o}/p \cdot \mathfrak{o}$ déduite de F par réduction modulo p . Sa hauteur au sens de Lazard et Dieudonné est égale à la hauteur h de G ; nous la supposons désormais finie⁽¹⁾. Le module de Dieudonné de $F_{(p)}$ est un \mathfrak{o} -module libre $\mathfrak{D}_p(G)$ de rang h muni d'un opérateur linéaire V , donc un module sur l'anneau de polynômes $\mathfrak{o}[V]$. On démontre qu'il existe un unique polynôme d'Eisenstein $P = V^h + b_1 V^{h-1} + \dots + b_{h-1} V + b_h$ dans $\mathfrak{o}[V]$ tel que $\mathfrak{D}_p(G)$ soit isomorphe au $\mathfrak{o}[V]$ -module $\mathfrak{o}[V]/(P)$. De plus, la théorie résumée dans [2] permet d'identifier $\mathfrak{D}_p(G)$ au quotient de $\mathfrak{D}(G)$ par le sous-groupe formé des différentielles de la forme $p \cdot df + \sum_\ell V_\ell \omega_\ell$ avec $f \in \mathfrak{A}_1$ et $\omega_\ell \in \mathfrak{D}(G)$ pour tout nombre premier ℓ , et V provient de V_p par passage au quotient.

Le polynôme d'Eisenstein P , ou ce qui revient au même, les coefficients b_1, \dots, b_h déterminent entièrement le module différentiel $\mathfrak{D}(G)$ qui se compose des formes différentielles ω telles que

$$(16) \quad V_p^h \omega + b_1 \cdot V_p^{h-1} \omega + \dots + b_{h-1} \cdot V_p \omega + b_h \cdot \omega \equiv 0 \quad \text{mod. } p \cdot d\mathfrak{A}_1 .$$

De manière plus explicite, soient $a(1), a(2), \dots, a(n), \dots$ des éléments de \mathfrak{o} ; posons

$$(17) \quad t(n) = a(n) + \frac{pb_{h-1}}{b_h} \cdot a(n/p) + \dots + \frac{p^{h-1}b_1}{b_h} \cdot a(n/p^{h-1}) + \frac{p^h}{b_h} \cdot a(n/p^h) .$$

La forme différentielle $\omega = \sum_{n=1}^{\infty} a(n) \cdot t^{n-1} dt$ appartient à $\mathfrak{D}(G)$ si et seulement si l'on a les congruences $t(n) \equiv 0 \text{ mod. } p^\alpha$ pour tout $\alpha \geq 1$ et tout entier $n \equiv 0 \text{ mod. } p^\alpha$. De plus, tout polynôme d'Eisenstein de degré h provient d'un groupe p -adique rigide de dimension 1 et de hauteur h .

 (1) Lorsque G est de hauteur infinie, il est isomorphe (de manière rigide) à G_a , et l'on a $\mathfrak{D}(G) = d\mathfrak{A}_1$.

En résumé, on peut répartir les groupes p -adiques rigides de hauteur h en familles non vides $F(b_1, \dots, b_h)$ (avec b_1, \dots, b_h dans $p \cdot \mathfrak{o}$ et b_h non divisible par p^2). Supposons que G soit de type $F(b_1, \dots, b_h)$ et soient ξ une coordonnée rigide dans

G , $\omega = \sum_{n=1}^{\infty} a(n) \cdot \xi^{n-1} d\xi$ une forme différentielle rigide invariante par translations

sur G . Alors les coefficients $a(n) \in \mathfrak{o}$ satisfont aux congruences $t(np^\alpha) \equiv 0 \pmod{p^\alpha}$ pour $\alpha \geq 1$ et $n \geq 1$, en définissant $t(n)$ comme plus haut⁽¹⁾.

4. Courbes elliptiques.

On note \mathbf{Z} l'anneau des entiers rationnels, \mathbf{Q} le corps des nombres rationnels et F_p le corps fini à p éléments. Soit $H \in \mathbf{Z}[X, Y, Z]$ un polynôme non nul, homogène de degré 3, irréductible et de discriminant non nul. On suppose que la courbe elliptique d'équation homogène $H = 0$ a un point d'inflexion à coordonnées rationnelles. Quitte à faire un changement linéaire de variables à coefficients entiers, on peut ramener H à la forme

$$(18) \quad H(X, Y, Z) = Y^2Z + (aX + bZ)YZ + (X^3 + uX^2Z + vXZ^2 + wZ^3)$$

et supposer que la réduction $H_{(p)}$ de H modulo p est irréductible dans $F_p[X, Y, Z]$ pour tout nombre premier p . Soit Γ le schéma projectif sur \mathbf{Z} associé à l'algèbre graduée $\mathbf{Z}[X, Y, Z]/(H)$; on pose $C = \mathbf{Q} \otimes_{\mathbf{Z}} \Gamma$ et $C_{(p)} = F_p \otimes_{\mathbf{Z}} \Gamma$, de sorte que C est la courbe elliptique sur \mathbf{Q} d'équation $H = 0$, et que $C_{(p)}$ est la réduction modulo p de C , d'équation $H_{(p)} = 0$. On dit que Γ est le *modèle de Néron* de C (cf. [4]). On considère C (resp. $C_{(p)}$) comme un groupe algébrique sur \mathbf{Q} (resp. F_p), d'élément neutre le point à l'infini e (resp. e_p).

Soit p un nombre premier. Nous associons comme suit un groupe p -adique rigide G_p à Γ : les points de G_p sont les points de Γ dans \mathbf{Z}_p qui se réduisent modulo p en e_p , et les fonctions analytiques rigides sur G_p sont les éléments du complété de l'anneau local du schéma Γ au point $e_p \in \Gamma(F_p)$. De manière plus concrète, G_p se compose des points $g = (x, y, z)$ de C dans \mathbf{Q}_p tels que $x/py \in \mathbf{Z}_p$, et l'on définit une coordonnée rigide ξ par $\xi(g) = x/y$. On note $\omega = \sum_{n=1}^{\infty} \beta(n) \cdot \xi^{n-1} d\xi$ la forme différentielle de première espèce sur C normalisée par $\beta(1) = 1$; c'est une forme différentielle rigide invariante par translations sur G_p .

Supposons d'abord que $C_{(p)}$ soit une courbe elliptique sur F_p , ce qui exclut un nombre fini de valeurs de p . Le nombre des points rationnels de $C_{(p)}$ est de la forme $1 - f_p + p$ avec $|f_p| < 2p^{1/2}$ (inégalité de Hasse-Weil). De plus, la réduction modulo p de ω est une forme de première espèce sur $C_{(p)}$ et "l'opération de Cartier" la multiplie par f_p ; comme cette opération transforme $h^{p^i-1}dh$ en $h^{p^{i-1}-1}dh$, on en déduit les

(1) En particulier, le groupe p -adique rigide G est défini à isomorphisme près par sa réduction modulo p , qui est un groupe formel sur $F_p = \mathfrak{o}/p \cdot \mathfrak{o}$, et il n'y a donc pas de "modules". Cette situation est particulière au cas envisagé $\mathfrak{o} = \mathbf{Z}_p$ (cf. [2]).

congruences $\beta(np) \equiv f_p \cdot \beta(n)$ et en particulier $\beta(p) \equiv f_p \pmod{p}$. Lorsque H est de la forme $Y^2Z - (X^3 - aXZ^2 - bZ^3)$, on a $f_p = \sum_{t \pmod{p}} -\left(\frac{t^3 - at - b}{p}\right)$ et l'on trouve ainsi les congruences (6) et (7) du n° 1 (cette démonstration est due à Serre). Enfin, f_p détermine la structure du groupe p -adique rigide G_p comme suit⁽¹⁾ :

(a) si $f_p \neq 0$, le groupe G_p est de hauteur 1, associé au polynôme d'Eisenstein $V - pu^{-1}$ où l'unité p -adique u satisfait à $u^2 - f_p u + p = 0$;

(b) si $f_p = 0$, le groupe G_p est de hauteur 2, associé au polynôme d'Eisenstein $V^2 + p$.

La congruence (9) du n° 1 se déduit immédiatement de là et des résultats du n° 3.

La fonction zêta de la courbe elliptique C a été définie par A. Weil comme le produit eulérien $\zeta_C(s) = \prod \zeta_p(s)$; lorsque $C_{(p)}$ est une courbe elliptique, on a $\zeta_p(s) = (1 - f_p p^{-s} + p^{1-2s})^{-1}$, et l'on a une recette bien définie [7] lorsque p est un nombre premier exceptionnel pour C . On peut aussi définir le schéma formel $\hat{\Gamma}$ complété de Γ le long de la section neutre ; c'est un groupe formel sur \mathbf{Z} . Le choix du paramètre local ξ permet de représenter $\hat{\Gamma}$ par une loi de groupe formel F à coefficients dans \mathbf{Z} , telle que $\xi(xx') = F(\xi(x) ; \xi(x'))$ pour tout nombre premier p et x, x' dans G_p .

Un de nos résultats fondamentaux (démontré aussi partiellement par Honda [3]) est le suivant : *il existe un paramètre local bien déterminé t dans Γ au voisinage de la section neutre tel que la forme différentielle de première espèce ω s'écrive*

$$\omega = \sum_{n=1}^{\infty} b(n) \cdot t^{n-1} dt$$

et que la fonction zêta de C s'écrive $\zeta_C(s) = \sum_{n=1}^{\infty} b(n) \cdot n^{-s}$ avec les mêmes coefficients entiers $b(n)$. Le choix usuel des facteurs exceptionnels de ζ_C est le seul pour lequel ce résultat soit vrai, et l'on peut donc dire que la fonction zêta de C ne dépend que du groupe formel associé à C .

5. Relation avec les fonctions automorphes.

Les résultats précédents nous semblent jeter une lumière supplémentaire sur les conjectures de Weil [7], [8] (mais non sur leur démonstration !). Notons \mathbf{C} l'ensemble des nombres complexes, \mathbf{P} le demi-plan de Poincaré et, pour tout entier $N > 0$, soit $\Gamma_0(N)$ le groupe des transformations conformes de \mathbf{P} de la forme $z \mapsto \frac{az + b}{cz + d}$ avec

 (1) Lorsque $C_{(p)}$ n'est pas une courbe elliptique, elle est isomorphe comme groupe algébrique sur \mathbf{F}_p , soit à G_a , soit à G_m , soit à la forme non-déployée de G_m qui se déploie sur l'extension quadratique de \mathbf{F}_p .

a, b, c, d entiers, $ad - bc = 1$ et $c \equiv 0 \pmod{N}$. Les coefficients entiers $b(n)$ étant définis comme précédemment, on note φ la forme différentielle holomorphe $\sum_{n=1}^{\infty} b(n) \cdot e^{2\pi i n \tau} d\tau$ sur \mathbf{P} . Enfin, soit N le conducteur de C ; c'est un entier > 0 dont les diviseurs premiers sont les nombres premiers exceptionnels pour C . La conjecture de Weil est que φ est toujours invariante par $\Gamma_0(N)$.

Soient D le disque unité ouvert dans \mathbf{C} , et C_c le tore complexe de dimension 1 formé des points complexes de C . Le groupe commutatif Λ formé des applications holomorphes γ de D dans C_c telles que $\gamma(0) = e$ est l'analogue du groupe $C(G)$ défini au n° 3. On définit pour chaque nombre premier p des opérateurs V_p et F_p par

$$(19) \quad V_p \gamma(q) = \gamma(q^p) \quad , \quad F_p \gamma(q^p) = \sum_{j=1}^p \gamma(\zeta^j q)$$

(avec $\zeta^p = 1, \zeta \neq 1$) ; l'opérateur de Hecke associé à p est $T_p = V_p + F_p$. Le paramètre local t auquel il est fait allusion à la fin du n° 4 définit en fait une coordonnée locale holomorphe au voisinage de e dans C_c et il existe un élément δ de Λ caractérisé par $t(\delta(q)) = q$ pour q assez petit dans D .

Posons $H(\tau) = \delta(e^{2\pi i \tau})$; alors H est une application holomorphe de \mathbf{P} dans C_c , caractérisée par la propriété suivante : l'image réciproque par H de la forme de première espèce ω sur C_c est la forme différentielle holomorphe φ sur \mathbf{P} . Soit p un nombre premier tel que $C_{(p)}$ soit une courbe elliptique; on peut montrer qu'on a $T_p \delta = f_p \delta$, c'est-à-dire la relation

$$(20) \quad H(p\tau) + \sum_{j \text{ mod. } p} H\left(\frac{\tau + j}{p}\right) = f_p \cdot H(\tau) \quad (\tau \text{ dans } \mathbf{P}) \quad .$$

La conjecture de Weil signifie que H se factorise en $\mathbf{P} \rightarrow \mathbf{P}/\Gamma_0(N) \xrightarrow{H'} C_c$. De plus, par adjonction à $\mathbf{P}/\Gamma_0(N)$ des points à l'infini correspondant aux "pointes", on obtient une courbe algébrique complète S_N sur \mathbf{C} . Or Shimura a construit dans [5] un modèle de S_N sur le corps \mathbf{Q} des nombres rationnels, et l'on peut raffiner sa méthode⁽¹⁾ de manière à obtenir un schéma Σ_N sur \mathbf{Z} tel que $S_N = \mathbf{C} \otimes_{\mathbf{Z}} \Sigma_N$. Nos résultats entraînent que, si C satisfait à la conjecture de Weil, H' est un morphisme de schémas de Σ_N dans Γ au-dessus de $\text{Spec}(\mathbf{Z})$.

(1) Pour tout nombre premier p , l'anneau local de Σ_N au point de Σ_N générique au-dessus de p se compose des fonctions méromorphes sur \mathbf{P} , invariantes par $\Gamma_0(N)$ et qui se développent en série de Fourier $\sum_{n=-\infty}^{\infty} c_n \cdot e^{2\pi i n \tau}$ avec des coefficients p -entiers c_n .

BIBLIOGRAPHIE

- [1] ATKIN A.O.L. — Congruences for modular forms, in *Computers in Mathematical Research*, North Holland Publ., 1968.
- [2] CARTIER P. — Relèvement des groupes formels commutatifs, *Séminaire Bourbaki*, 21^e année, 1968-69, exposé n° 359.
- [3] HONDA T. — Formal groups and zeta functions, *Osaka J. Math.*, 5, 1968, p. 199-213.
- [4] NÉRON A. — Modèles minimaux des variétés abéliennes sur les corps locaux et globaux, *Publ. Math. I.H.E.S.*, 21, 1964, p. 1-128.
- [5] SHIMURA G. — Correspondances modulaires et les fonctions ζ des courbes algébriques, *Journ. Math. Soc. Japan*, 10, 1958, p. 1-28.
- [6] TATE J. — Rigid analytic spaces, à paraître dans *Inventiones Mathematicae*.
- [7] WEIL A. — Ueber die Bestimmung Dirichletscher Reihen durch Funktionalgleichungen, *Math. Annalen*, 168, 1967, p. 149-156.
- [8] WEIL A. — Dirichlet series and automorphic functions, *Lecture Notes in Math. 189*, Springer 1971.

Université Louis Pasteur
Dept. de Mathématiques
7, Rue René Descartes,
67 — Strasbourg — France

BOUNDARIES OF LIE GROUPS AND DISCRETE SUBGROUPS

by Harry FURSTENBERG

1. Introduction.

If G is a non-compact semi-simple Lie group, there is a compact homogeneous space $B(G)$ attached to it which plays an important role in the theory of harmonic functions on the symmetric space associated with the group G . $B(G)$ is a boundary component of one of the Satake compactifications of the symmetric space ([8]), but it can be characterized directly in terms of its behavior as a G -space (see § 3). In [4] the space $B(G)$ is shown to play an important role in the theory of spherical functions on G which means that it also has significance for the theory of irreducible unitary representations of G . More recently this space has appeared as a tool in proving "rigidity" theorems. If G is a locally compact topological group, a subgroup Γ is called a *lattice subgroup* if Γ is discrete and G/Γ has finite left-invariant (Haar) measure. Suppose Γ_1 is a lattice in G_1 and Γ_2 is a lattice in G_2 , where G_1 and G_2 are semi-simple Lie groups. One wants to know to what extent an assertion of the following type is valid: an isomorphism of Γ_1 with Γ_2 is induced by an isomorphism of G_1 with G_2 . The case where $G_1 = G_2$ and Γ_2 is obtained from Γ_1 by a continuous deformation had been treated in work by Calabi, Vesentini, Weil, Garland and Ragunathan. In some recent work the space $B(G)$ has played a major role in the arguments. For example, in [5] we considered the question of whether $\Gamma_1 \cong \Gamma_2$ implied $G_1 \cong G_2$, and we treated a fairly special case making use of the notion of Poisson boundary which is closely related to that of the space $B(G)$. Mostow has given a rather conclusive treatment for the case that the Γ_i are uniform (co-compact) subgroups showing that, in general, an isomorphism of Γ_1 with Γ_2 induces a homeomorphism of $B(G_1)$ with $B(G_2)$, and this in turn implies the isomorphism of G_1 with G_2 ([7]).

The latter results suggest the possibility that when Γ is a lattice subgroup of the semi-simple group G , the space $B(G)$ may be attached directly to the group Γ as an abstract group. More precisely, one might expect to be able to define a functor Π on the category of locally compact groups to the category of compact spaces satisfying the following conditions;

- (i) $\Pi(G)$ is a G -space.
- (ii) If $h : G_1 \rightarrow G_2$ is a epimorphism so that every G_2 -space can be viewed as a G_1 -space, then there exists a G_1 -equivariant map $h^* : \Pi(G_2) \rightarrow \Pi(G_1)$.
- (iii) If Γ is a lattice subgroup of G , then there exists a Γ -equivariant map $\Pi(\Gamma) \rightarrow \Pi(G)$ which is an isomorphism.
- (iv) If G is a semi-simple Lie group, then $\Pi(G) = B(G)$.

Without further restriction one cannot expect to attain all of these conditions. One can see this by considering the case of the free group F_n on n generators which is a lattice subgroup of $SL(2, R)$. Since $B(SL(2, R))$ is P^1 , the one-dimensional projective space, and since free groups map readily into any group one would find an abundance of maps of P^1 into every $B(G)$. However there are even automorphisms of F_n which are not compatible with continuous maps of P^1 onto itself.

In what follows we shall discuss several candidates for the functor Π and show to what extent the conditions above are met. While in none of these do we achieve the identity of $\Pi(\Gamma)$ with $\Pi(G)$ for Γ a lattice subgroup, we nonetheless find in one case that these two spaces are sufficiently close to have some implications for rigidity type theorems. In particular we obtain in this way an alternative proof of the result announced in [5] to the effect that if G_1 has R -rank 1 and G_2 is one of the groups $SL(m, R)$, $m \geq 3$, then G_1 and G_2 have no isomorphic lattice subgroups. The same method also seems to show that for $m \neq n$ $SL(m, R)$ and $SL(n, R)$ cannot have isomorphic lattice subgroups. Because of the more precise results of Mostow we haven't pursued this matter to its conclusion. Our interest in the functor Π stems from our expectation that the spaces $\Pi(G)$ will play a role in other problems. Detailed proofs will appear elsewhere.

2. Proximal Minimal G -spaces.

This notion is borrowed from topological dynamics ([2]). Let M be a compact G -space. We say M is *minimal* if M does not contain a non-trivial, closed, G -invariant subset. Equivalently, M is minimal if every G -orbit in M is dense. These spaces are plentiful since every compact G -space must contain a minimal G -space. M is called *proximal* if for every pair $x, y \in M$ there exists a net $\{g_\alpha\}$ in G with

$$\lim g_\alpha x = \lim g_\alpha y .$$

When G is abelian, or more generally, when G is nilpotent, every proximal G -space contains a fixed point for the group, so that the only proximal, minimal G -space is the trivial space. However, when G is semi-simple there exist interesting proximal minimal spaces. Namely one has

THEOREM 2.1. — *If G is semi-simple then $B(G)$ is a proximal G -space.*

Since G is transitive on $B(G)$, the latter is obviously a minimal G -space.

Now products of proximal spaces are proximal, and if M_1 and M_2 are proximal minimal, then a minimal subspace of $M_1 \times M_2$ will be both proximal and minimal and will have M_1 and M_2 as equivariant images. In this way one may prove

THEOREM 2.2. — *For an arbitrary group G , there exists a universal proximal minimal G -space $\Pi_p(G)$ such that if M is any proximal minimal G -space, then exists an equivariant map $\rho : \Pi_p(G) \rightarrow M$.*

Note that ρ is onto inasmuch as M is minimal. Moreover, ρ is unique. For suppose that ρ and σ were two such maps. Let $x \in \Pi_p(G)$ and choose a net with $\lim g_\alpha \rho(x) = \lim g_\alpha \sigma(x)$. If y is a limit of a subnet of $g_\alpha x$, then $\rho(y) = \sigma(y)$,

whence $\rho(gy) = \sigma(gy)$, and since the orbit of y is dense, $\rho = \sigma$. This implies in particular that $\Pi_p(G)$ is unique.

It is easily shown that conditions (i) and (ii) are met for Π_p . The following condition related to (iii) can also be established :

THEOREM 2.3. — *If Γ_1 is a subgroup of finite index in Γ_2 , then $\Pi_p(\Gamma_1) \cong \Pi_p(\Gamma_2)$*

It is an open question whether (iv) is valid in this case. (iv) would be valid if it were true that $\Pi_p(S)$ is trivial for all solvable S . This is also open. In any case, if (iv) is true so that $\Pi_p(S)$ is a manifold when G is semi-simple, then (iii) will certainly not be true since one can prove that for Γ a lattice subgroup of a semi-simple Lie group, $\Pi_p(\Gamma)$ will not be a manifold.

3. Strong proximality.

We will denote by $\mathfrak{R}(M)$ the space of regular probability measures on the compact space M . If M is a G -space, then $\mathfrak{R}(M)$ is a G -space. Moreover we endow $\mathfrak{R}(M)$ with the usual weak topology so that it becomes a compact convex set. The extremals of this set are the point measures and these are in correspondance with the points of M . Now suppose that M is a proximal G -space. One sees easily that for any finite set of points $x_1, x_2, \dots, x_n \in M$, there exists a net g_α in G with $\lim g_\alpha x_1 = \lim g_\alpha x_2 = \dots = \lim g_\alpha x_n$. From this it follows that if ν is any discrete measure in $\mathfrak{R}(M)$ we can find a net with $g_\alpha \nu \rightarrow$ point measure. We shall say that a G -space M is *strongly proximal* if this holds for an arbitrary $\nu \in \mathfrak{R}(M)$. Again one shows :

THEOREM 3.1. — *For an arbitrary group G there exists a universal strongly proximal minimal G -space $\Pi_{sp}(G)$ with the property that if M is any strongly proximal minimal G -space there exists a unique equivariant map ρ of $\Pi_{sp}(G)$ onto M .*

Moreover one has

THEOREM 3.2. — *If G is a connected Lie group and R its (not necessarily connected) radical, then $\Pi_{sp}(G) = B(G/R)$.*

This is an easy consequence of [3].

We will see in the next section that if Γ is a lattice subgroup of the semi-simple group G , then the space $\Pi_{sp}(G)$ is a strongly proximal minimal Γ -space. Hence there is a map of $\Pi_{sp}(\Gamma)$ onto $\Pi_{sp}(G)$. However, in general it will not be one-one. For we can construct a strongly proximal Γ -space M which is a non-trivial extension of $\Pi_{sp}(G)$. Namely we form a Γ -space M which is minimal and of which $\Pi_{sp}(G)$ is an equivariant image such that for some point $x \in \Pi_{sp}(G)$, the inverse image in M consists of a single point. It is easy to see that this implies that M is strongly proximal.

As with Π_p one has

THEOREM 3.3. — *If Γ_1 is of finite index in Γ_2 then*

$$\Pi_{sp}(\Gamma_1) \cong \Pi_{sp}(\Gamma_2) \quad ,$$

Finally we mention another characterization of the space $\Pi_{SP}(G)$.

THEOREM 3.4. — *If G acts by affine transformations on a compact convex set Q leaving no proper compact convex subset invariant, then there exists a unique affine equivariant map of $\mathfrak{E}(\Pi_{SP}(G))$ onto Q .*

4. Mean Proximal G -spaces.

We now introduce a further strengthening of proximality which enables us to say something more about the relation between $\Pi(\Gamma)$ and $\Pi(G)$. Let M be a G -space and let μ be a probability measure on G . Form the sequence of measures

$$\mu_n = \frac{\overbrace{\mu + \mu * \mu + \dots + \mu * \dots * \mu}^n}{n}$$

We shall say that M is μ -proximal if for every neighborhood \mathcal{U} of the diagonal $\Delta(M) \subset M \times M$ we have

$$\mu_n \{g | (gx, gy) \notin \mathcal{U}\} \rightarrow 0$$

as $n \rightarrow \infty$ for each $x, y \in M$. Finally we say that M is mean proximal if it is μ -proximal for any μ whose support generates G .

Once again we have

THEOREM 4.1. — *There exists a universal mean proximal minimal G -space $\Pi_{MP}(G)$ such that if M is any mean proximal minimal G -space then there exists a unique equivariant map of $\Pi_{MP}(G)$ onto M .*

The following theorem implies that mean proximality is in fact stronger than strong proximality.

THEOREM 4.2. — *The following conditions on a metric G -space M are equivalent :*

- (a) M is μ -proximal.
- (b) A measure ν on $M \times M$ satisfying $\mu * \nu = \nu$ is supported by the diagonal .
- (c) If $X_1, X_2, \dots, X_n, \dots$ is a sequence of G -valued independent random variables each having distribution μ , and if ν is a measure on M with $\mu * \nu = \nu$, then with probability one, $X_1, X_2, \dots, X_n \nu$ converges to a point measure.
- (d) If $\theta \in \mathfrak{E}(M)$ and u is an open subset of $\mathfrak{E}(M)$ containing all point measures then $\mu_n \{g | g\theta \notin u\} \rightarrow 0$ as $n \rightarrow \infty$.

The relationship between the various notions of proximality imply

$$\Pi_P(G) \rightarrow \Pi_{SP}(G) \rightarrow \Pi_{MP}(G)$$

As we saw in Theorem 2.2, all equivariant maps between proximal spaces are unique.

In the remainder of the paper we shall sketch a proof of the following result :

THEOREM 4.3. — *If G is a semi-simple Lie group then $\Pi_{MP}(G) = B(G)$. If Γ is a lattice subgroup of G then there is an equivariant map $\rho : \Pi_{MP}(\Gamma) \rightarrow \Pi_{MP}(G)$. The map ρ is a measurable isomorphism in the sense that there exists a map*

$$\sigma : \Pi_{MP}(G) \rightarrow \Pi_{MP}(\Gamma)$$

which is measurable as a map from the manifold $B(G)$ to $\Pi_{MP}(\Gamma)$, and such that $\rho \circ \sigma = \text{identity}$.

A map on a C^∞ -manifold is said to be *measurable* if when composed with a real-valued function, the resulting function is a lebesgue measurable function of the local coordinates.

For the first assertion of the theorem, inasmuch as $B(G)$ is strongly proximal so that there exists a map $B(G)$ onto $\Pi_{SP}(G)$, it would suffice to show that $B(G)$ is mean proximal. For the second assertion we must show that $B(G)$ is mean proximal as a Γ -space. Both of these follow from the next theorem.

THEOREM 4.4. — *Let G be a subgroup of $GL(n, R)$ which together with all its subgroups of finite index acts irreducibly on R^n . Assume moreover that G acts proximally on the projective space P^{n-1} . Then P^{n-1} is a mean proximal space.*

The proof of the theorem depends upon an analysis of the behavior of random products of matrices. The type of argument is similar to that of [6, § 8].

To apply the foregoing theorem to give a proof of Theorem 4.3, we use the fact that the space $B(G)$ occurs as a component of one of the Satake compactifications of the symmetric space G/K . In these compactifications the symmetric space is identified with a subset of the projective space associated with the space of symmetric matrices of a certain dimension. Moreover, the group G acts linearly on this space. It is not hard to show that the action is irreducible and proximal. According to Theorem 4.4, it will therefore be mean proximal.

Now let Γ be a lattice subgroup of G . By [1], if G acts irreducibly on a space so does Γ and so does every subgroup of finite index. Moreover, it is easily seen that if G acts mean proximally on a space, then Γ at least acts proximally. Therefore Theorem 4.4 applies to Γ as well, and this yields the second statement of the theorem.

We now turn to the last assertion of the theorem. Let $\Pi = \Pi_{MP}(\Gamma)$ so that Π is a Γ -space. We construct a G -space by dividing the product $G \times \Pi$ by the relation $(g, x) \sim (g\gamma^{-1}, \gamma x)$ and setting $g_1(g, x) = (g_1 g, x)$. We denote this space $G \times_\Gamma \Pi$. Clearly $G \times_\Gamma \Pi$ has G/Γ as equivariant image. Both of these spaces are locally compact and it follows that the set of probability measures on $G \times_\Gamma \Pi$ which map onto the Haar measure on G/Γ form a compact convex space. Now recall that $B(G)$ can be expressed as G/H where H is a subgroup of G with the fixed point property ([3]). So there exists a measure θ on $G \times_\Gamma \Pi$ which maps onto Haar measure on G/Γ and which is invariant under H . If we lift θ to $G \times \Pi$, we obtain a measure λ on $G \times \Pi$ invariant under the action of both Γ and H . Here Γ acts by sending

(g, x) into $(g\gamma^{-1}, \gamma x)$, and H acts by sending (g, x) into (hg, x) . Since θ maps into Haar measure on G/Γ , λ maps into Haar measure on G . We may therefore decompose λ :

$$\lambda = \int_G \delta_g \times \lambda_g dg$$

with δ_g the Dirac measure and λ_g a Haar-measurable function on G with values in $\mathfrak{R}(\Pi)$.

We now have

$$h\lambda = \int_G \delta_{hg} \times \lambda_g dg = \int_G \delta_g \times \lambda_{h^{-1}g} dg$$

whence $\lambda_{h^{-1}g} = \lambda_g$, so we can define ω_ξ for $\xi \in G/H$ by $\omega_{gH} = \lambda_{g^{-1}}$. Moreover

$$\gamma\lambda = \int_G \delta_{g\gamma^{-1}} \times \gamma\lambda_g dg = \int_G \delta_g \times \gamma\lambda_{g\gamma} dg$$

whence $\lambda_{g\gamma} = \gamma^{-1}\lambda_g$. Therefore $\gamma\omega_{gH} = \gamma\lambda_{g^{-1}} = \lambda_{g^{-1}\gamma^{-1}} = \omega_{\gamma gH}$ and $\gamma\omega_\xi = \omega_{\gamma\xi}$.

We thus find a Γ -equivariant map ω of $B(G)$ into $\mathfrak{R}(\Pi)$. ω need not be continuous, but it is measurable.

We now make use of (the easy part of) Theorem 3 in [5]. According to this theorem, there exists a measure μ on Γ and an absolutely continuous measure ν on $B(G)$ with $\mu * \nu = \nu$. The map ω takes ν into a measure $\tilde{\nu}$ on $\mathfrak{R}(\Pi)$ and by the equivariance of the map Γ it follows that $\mu * \tilde{\nu} = \tilde{\nu}$. By Theorem 4.2, $\tilde{\nu}$ must be concentrated on point measures of Π . In other words, ω defines a measurable Γ -equivariant map from $B(G)$ into Π . This is the σ of our theorem.

BIBLIOGRAPHY

- [1] BOREL A. — Density properties for certain subgroups of semi-simple groups without compact component, *Ann. of Math.*, 72, 1960, p. 179-188.
- [2] ELLIS R. — *Lectures in Topological Dynamics*, W.A. Benjamin Inc., 1969.
- [3] FURSTENBERG H. — A Poisson formula for semi-simple Lie groups, *Ann. of Math.*, 77, 1963, p. 335-386.
- [4] FURSTENBERG H. — Translation invariant cones of functions on semi-simple Lie groups, *Bull. Amer. Math. Soc.*, 71, 1965, p. 271-326.
- [5] FURSTENBERG H. — Poisson boundaries and envelopes of discrete groups, *Bull. Amer. Math. Soc.*, 73, 1967, p. 350-356.
- [6] FURSTENBERG H. — Noncommuting random products, *Trans. Amer. Math. Soc.*, 108, 1963, p. 377-428.
- [7] MOSTOW G.D. — The rigidity of locally symmetric spaces, *Proceedings of this congress*.
- [8] SATAKE I. — On representations and compactifications of symmetric spaces, *Ann. of Math.*, 71, 1960, p. 77-110.

Dept. of Mathematics,
Jérusalem
Israël