

# ANNALES SCIENTIFIQUES DE L'É.N.S.

PIERRE ARNOUX

ARNALDO NOGUEIRA

**Mesures de Gauss pour des algorithmes de fractions  
continues multidimensionnelles**

*Annales scientifiques de l'É.N.S. 4<sup>e</sup> série*, tome 26, n° 6 (1993), p. 645-664

[http://www.numdam.org/item?id=ASENS\\_1993\\_4\\_26\\_6\\_645\\_0](http://www.numdam.org/item?id=ASENS_1993_4_26_6_645_0)

© Gauthier-Villars (Éditions scientifiques et médicales Elsevier), 1993, tous droits réservés.

L'accès aux archives de la revue « Annales scientifiques de l'É.N.S. » (<http://www.elsevier.com/locate/ansens>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

# MESURES DE GAUSS POUR DES ALGORITHMES DE FRACTIONS CONTINUES MULTIDIMENSIONNELLES

PAR PIERRE ARNOUX ET ARNALDO NOGUEIRA

---

**ABSTRACT.** — We give a procedure to compute invariant measures for some multi-dimensional continued fractions algorithms; we use this procedure to build in an elementary way the usual Gauss measure for the map  $x \mapsto \{1/x\}$ , and to build an ergodic invariant measure for the algorithms of Brun and Selmer in any dimension. We show that, in the case of ordinary continued fraction, our construction admits a natural geometric interpretation.

## 1. Introduction

La théorie des fractions continues est directement reliée à l'étude de la transformation de l'intervalle  $[0, 1[$  qui à  $\alpha$  associe  $\{1/\alpha\}$ , où  $\{\alpha\}$  désigne la partie fractionnaire de  $\alpha$ . On sait que cette transformation possède une mesure invariante absolument continue par rapport à la mesure de Lebesgue, la mesure de Gauss, de densité  $(\log 2)^{-1} (1 + \alpha)^{-1}$ .

Une fois la densité donnée, il est facile de vérifier que cette mesure est invariante; il est plus difficile de donner une preuve qui fasse apparaître cette densité de façon naturelle, ce qui est utile pour pouvoir généraliser ce résultat à des algorithmes d'approximation simultanée de plusieurs nombres. De tels algorithmes ont été étudiés en particulier en arithmétique, en essayant de généraliser le théorème de Lagrange sur la périodicité du développement en fraction continue des nombres quadratiques, ou pour trouver les unités fondamentales de corps cubiques en généralisant les résultats classiques sur les corps quadratiques (résolution de l'équation de Pell-Fermat par les fractions continues); ils pourraient aussi jouer un rôle en mécanique, les problèmes d'approximation simultanée de plusieurs réels par des rationnels étant directement reliés aux problèmes de petits diviseurs.

Nous nous proposons dans cet article de fournir une méthode élémentaire qui permet d'obtenir une dérivation naturelle de la mesure de Gauss. Dans la deuxième section, nous expliquons le principe de la construction et nous donnons les notations qui seront

---

Ce travail a été réalisé au cours d'une invitation du premier auteur à l'Instituto de Matemática, Universidade Federal do Rio de Janeiro, Brésil.

utilisées dans la suite. Dans la troisième section, nous montrons comment on peut de cette manière retrouver la mesure de Gauss classique. Dans la quatrième section, nous donnons une interprétation géométrique de cette construction, comme paramétrisation explicite du flot géodésique sur la surface modulaire, et dans la cinquième section, nous retrouvons grâce à cette interprétation la constante de Lévy comme temps de premier retour moyen de ce flot géodésique en une bonne section. Dans la sixième section, nous appliquons cette méthode à des algorithmes de fractions continues généralisées, les versions additives et multiplicatives de l'algorithme de Brun, qui sont définies de la façon suivante (cf. [Br], p. 40) :

On prend deux nombres réels  $0 < \alpha < \beta < 1$ , on remplace 1 par  $1 - \beta$  pour l'algorithme additif, et par  $1 - [1/\beta]\beta$  pour l'algorithme multiplicatif ( $[x]$  désigne la partie entière de  $x$ ), on réordonne les trois nombres dans l'ordre croissant, et on normalise pour que le plus grand soit égal à 1.

Ces algorithmes peuvent donc être vus comme des fonctions sur le triangle  $\{(\alpha, \beta) \in \mathbb{R}^2 \mid 0 < \alpha < \beta < 1\}$ , et nous montrons le théorème suivant :

**THÉORÈME.** — *Il existe une mesure de probabilité, absolument continue par rapport à la mesure de Lebesgue, invariante pour l'algorithme additif de Brun, de densité :*

$$\frac{12}{\pi^2} \frac{1}{\beta(1+\alpha)}$$

*de même, il existe une mesure de probabilité, absolument continue par rapport à la mesure de Lebesgue, invariante pour l'algorithme multiplicatif de Brun, dont la densité est donnée, à une constante près, par :*

$$\frac{1}{(1+\alpha)(1+\alpha+\beta)} + \frac{1}{(1+\beta)(1+\alpha+\beta)}$$

Ce théorème se généralise facilement en dimension  $n$ , nous donnons les formules correspondantes dans la sixième section, et nous étudions un algorithme analogue, celui de Selmer. Dans la septième section, nous montrons que la mesure obtenue est ergodique, et dans la huitième section nous posons quelques problèmes reliés.

L'existence d'une mesure ergodique absolument continue par rapport à la mesure de Lebesgue pour plusieurs algorithmes, en particulier celui de Jacobi (cf. [Sc1]), est connue depuis plusieurs années, et Schweiger a donné, par une méthode différente de la nôtre, des formules explicites pour les algorithmes de Brun et Selmer (cf. [Sc2], [Sc3], [Sc4]) ; il considère comme nous des extensions bijectives de ces algorithmes, mais il se place directement dans le cadre projectif, et trouve des mesures avec une densité non triviale alors que nous obtenons la mesure de Lebesgue. Par exemple, sa méthode appliquée au cas le plus simple donne l'extension naturelle de Ito, Nakada, Tanaka (cf. [INT]) : c'est une application du carré dans lui-même définie par  $(x, y) \mapsto (\{1/x\}, ([1/x] + y)^{-1})$ , avec une mesure invariante de densité  $(1 + xy)^{-2}$ , alors que nous verrons dans la troisième section que nous obtenons l'application  $(x, y) \mapsto (\{1/x\}, x - x^2 y)$  sur l'ensemble  $0 < x < 1, 0 < y < (1 + x)^{-1}$ , qui préserve la mesure de Lebesgue ; ces deux applications sont conjuguées par  $(x, y) \mapsto (x, y/(1 - xy))$ .

Dans le cas des fractions continues usuelles, d'autres auteurs ont déjà eu une approche semblable : on peut citer, en arithmétique, les extensions naturelles des fractions continues (cf. [Na], [Ja]), et du point de vue géométrique le codage de l'espace modulaire (cf. [AF2], [BKS], [Se]), mais ces travaux restent toujours en dimension 1.

Par contre, le procédé d'induction des échanges d'intervalles peut être considéré comme un algorithme complexe de fractions continues généralisées en grande dimension (cf. [Ve1]), et les méthodes employées sont très semblables aux nôtres ; en fait, on peut considérer que nous exposons ici une simplification des méthodes de Veech, en particulier la section 3 est le cas le plus simple pour les échanges d'intervalle.

## 2. Principe de la construction de la mesure

Rappelons d'abord ce qu'est un algorithme vectoriel de fractions continues : il s'agira pour nous d'une application définie sur un sous-cône  $\Lambda$  du cône positif de  $\mathbb{R}^n$ , commutant avec les homothéties, et linéaire par morceaux avec des coefficients entiers (cf. [Br] pour une définition plus élaborée). On peut naturellement passer en projectif, et l'on obtient une application définie sur une partie  $\Delta$  du simplexe standard ; c'est pour ces applications que l'on cherche une mesure invariante. L'algorithme est dit additif s'il est défini sur un nombre fini de morceaux et ne fait intervenir que des constantes, multiplicatif s'il fait intervenir des parties entières  $[\alpha_i/\alpha_j]$ .

Dans toute la suite, nous noterons  $F : \Lambda \rightarrow \Lambda$  l'algorithme vectoriel considéré, et  $f : \Delta \rightarrow \Delta$  l'application projective associée.

L'algorithme additif le plus simple est celui qui, au couple  $(\alpha, \beta)$ ,  $0 < \alpha < \beta$ , associe  $(\alpha, \beta - \alpha)$  si  $\alpha < \beta - \alpha$ , et  $(\beta - \alpha, \alpha)$  dans le cas contraire ; l'algorithme multiplicatif le plus simple associe au couple  $(\alpha, \beta)$ ,  $0 < \alpha < \beta$ , le couple  $(\beta - [\beta/\alpha]\alpha, \alpha)$ . Ces deux procédés, qui ne sont qu'une version sur les nombres réels de l'algorithme d'Euclide, correspondent au développement en fraction continue ; nous les étudions dans la section 3. L'algorithme additif se réalise facilement sous forme géométrique, en prenant deux intervalles de longueur  $\alpha$  et  $\beta$  ; sous cette forme, il était connu dans l'antiquité sous le nom d'anthyphère, et a peut-être servi à la fameuse découverte de l'irrationalité de  $\sqrt{2}$ , puisqu'il permet de prouver géométriquement la périodicité de son développement en fraction continue, et donc l'incommensurabilité de la diagonale du carré par rapport au côté (cf. [Ar]).

*Remarque.* — L'algorithme n'est en général défini que sur le complémentaire d'un nombre fini d'hyperplans ; une extension de la définition à ces hyperplans serait arbitraire et non significative. Comme on veut pouvoir itérer l'algorithme, on se placera en fait sur le complémentaire d'un nombre dénombrable d'hyperplans ; pour chercher des mesures absolument continues par rapport à la mesure de Lebesgue, il n'est pas gênant de retirer un ensemble de mesure nulle, et partout où ce sera nécessaire dans la suite, on se permettra de négliger un nombre dénombrable d'hyperplans sans en faire la remarque (il suffit par exemple de ne considérer que les points dont les coordonnées sont rationnellement indépendantes).

Ce qui complique la recherche d'une mesure invariante pour  $f$ , c'est que ce n'est pas une application bijective, chaque point pouvant avoir plusieurs antécédents; notre idée est de rajouter des informations pour rendre la transformation bijective, c'est-à-dire, en notant  $\pi: \mathbb{R}^{2n} \rightarrow \mathbb{R}^n$  la projection naturelle qui à un point associe ses  $n$  premières coordonnées, de définir une transformation  $\tilde{F}$  sur un sous-cône du cône positif de  $\mathbb{R}^{2n}$  qui se projette par  $\pi$  sur  $F$ .

Pour expliquer comment on définit la transformation  $\tilde{F}$ , le plus simple est de donner une interprétation des fractions continues en termes de dynamique; rappelons tout d'abord la définition suivante:

**DÉFINITION.** — Soit  $h: X \rightarrow X$  une application, et  $E$  un sous-ensemble de  $X$ ; on appelle application induite de  $h$  sur  $E$ , notée  $h|_E$ , l'application  $E \rightarrow E$ ,  $x \mapsto h^{n_x}(x)$  où  $n_x$  est le plus petit entier strictement positif  $n$  tel que  $h^n(x) \in E$  (temps de premier retour en  $E$ ; on suppose que ce temps est partout défini).

Les algorithmes usuels de fractions continues consistent à induire des rotations sur des intervalles bien choisis, de façon à retrouver une rotation, ce qui permet d'itérer le procédé (voir section 3). Comme la transformation induite ne garde pas trace des temps de premier retour, on ne peut pas retrouver la transformation initiale; il est possible de conserver ces temps de retour en faisant la construction connue en théorie ergodique sous le nom de « tours », où l'on regarde, au lieu de l'application induite  $h|_E$ , l'application définie sur  $\{(x, n) | x \in E, 0 \leq n < n_x\}$  qui envoie  $(x, n)$  sur  $(x, n+1)$  si  $n+1 < n_x$ , et sur  $(h^{n_x}(x), 0)$  sinon. A partir d'une telle tour, il est possible de retrouver la transformation initiale; plus généralement, on peut itérer cette construction de tour, ce qui permet de remonter  $k$  étapes. Cependant, tant que l'on a des hauteurs entières, on ne peut remonter qu'un nombre fini d'étapes, d'où l'idée de remplacer formellement ces hauteurs par des nombres réels.

Nous prendrons sur  $\mathbb{R}^n$  des coordonnées  $(\alpha_1, \dots, \alpha_n)$ , et sur l'espace des rectangles  $\mathbb{R}^{2n}$  des coordonnées  $(\alpha_1, \dots, \alpha_n, \lambda_1, \dots, \lambda_n)$ .

En général, nous avons une application  $F$  définie sur une partie  $X$  de  $\mathbb{R}^n$  par une matrice  $A$ ; nous définissons l'application  $\tilde{F}$  sur  $\pi^{-1}(X) \subset \mathbb{R}^{2n}$  par la matrice

$$\begin{pmatrix} A & 0 \\ 0 & {}^t A^{-1} \end{pmatrix}$$

c'est-à-dire que, si  $F$  remplace  $(\alpha_i, \alpha_j)$  par  $(\alpha_i - \alpha_j, \alpha_j)$ , l'application  $\tilde{F}$  a le même effet sur les  $\alpha_i$  et remplace  $(\lambda_i, \lambda_j)$  par  $(\lambda_i, \lambda_j + \lambda_i)$ ; géométriquement, si  $F$  retranche à un intervalle de longueur  $\alpha_i$  un intervalle de longueur  $\alpha_j$ ,  $\tilde{F}$  retranche à un rectangle de base  $\alpha_i$  et de hauteur  $\lambda_i$  un rectangle de même hauteur et de base  $\alpha_j$ , pour le mettre au-dessus du rectangle de base  $\alpha_j$  et de hauteur  $\lambda_j$  (cf. Fig. 1). Il est clair que cette transformation

préserve la mesure de Lebesgue et la forme  $\sum_{i=1}^n \alpha_i \lambda_i$  (aire totale des tours).

Par ailleurs, on veut pouvoir normaliser le vecteur  $(\alpha_1, \dots, \alpha_n)$ ; on peut utiliser à cet effet le flot  $\Phi_t$  défini par:

$$(\alpha_1, \dots, \alpha_n, \lambda_1, \dots, \lambda_n) \mapsto (\alpha_1 e^t, \dots, \alpha_n e^t, \lambda_1 e^{-t}, \dots, \lambda_n e^{-t})$$

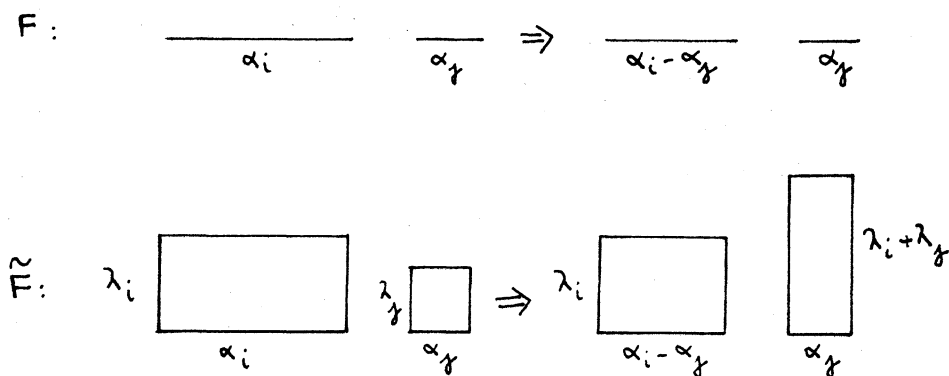


Fig. 1

Il est clair que  $\Phi_t$  préserve lui aussi la mesure de Lebesgue et l'aire totale des tours, et qu'il commute avec  $\tilde{F}$ .

On peut donc se restreindre à l'hypersurface donnée par la condition  $\sum \alpha_i \lambda_i = 1$ ; le flot et la fonction préservent sur cette hypersurface une mesure induite par la mesure de Lebesgue. On supposera dans la suite que l'on s'est donné un domaine  $D$  de cette hypersurface, invariant par  $\Phi_t$ , et sur lequel  $\tilde{F}$  est bijective.

Puisque  $\Phi_t$  et  $\tilde{F}$  commutent, on peut passer au quotient par  $\tilde{F}$ , et considérer le flot induit sur  $\Omega = D/\tilde{F}$  par  $\Phi_t$ . Prenons une surface de section  $\Sigma$  pour le flot (dans la suite, on prendra toujours l'hypersurface de  $D$  définie par  $\alpha_n = 1$ ); nous sommes intéressés par l'application  $\tilde{f}$  de premier retour du flot en  $\Sigma$ . Remarquons d'abord qu'il est inutile de préciser si la surface de section est dans  $\Omega$  ou dans  $D$ : en effet, si  $\Sigma$  est une hypersurface dans  $D$  qui coupe toute orbite du flot, on peut donner un domaine fondamental pour l'action de  $\tilde{F}$  sur  $D$  sous la forme :

$$\{ \Phi_t(x) \mid x \in \Sigma; 0 \leq t < t_x \}$$

où  $t_x$  est le temps de premier retour en  $\Sigma$ , et on peut assimiler  $\Omega$  à ce domaine (le temps de premier retour est presque partout défini dans le cas que nous regardons, puisque c'est  $t_x = -\log \alpha_n(\tilde{F}x)$ , où  $\alpha_n(x)$  est la  $n$ -ième coordonnée de  $x$ , et que  $\alpha_n(\tilde{F}x)$  est presque partout non nul).

Il est clair que l'application  $\tilde{f}$  se projette par  $\pi$  sur l'application  $f$ :

$$\begin{array}{ccc} \Sigma & \xrightarrow{\tilde{f}} & \Sigma \\ \pi \downarrow & & \downarrow \pi \\ \Delta & \xrightarrow{f} & \Delta \end{array}$$

puisque l'effet de  $\tilde{f}$  sur les  $n$  premières composantes est exactement d'appliquer  $F$  et de normaliser par le flot.

Par définition, le flot  $\Phi_t$  préserve une mesure explicitement donnée sur  $\Omega$ ; on peut donc en déduire par un calcul classique une mesure invariante pour l'application de premier retour  $\tilde{f}$ , et il suffira de la projeter sur  $\Delta$  pour obtenir la mesure invariante pour  $f$ . En fait,  $\Sigma$  est de dimension  $2n-2$ , on peut prendre sur cet ensemble les coordonnées  $(\alpha_1, \dots, \alpha_{n-1}, \lambda_1, \dots, \lambda_{n-1})$ , et dans tous les cas que nous calculerons on verra que  $\tilde{f}$  préserve la mesure de Lebesgue donnée par ces coordonnées; le calcul de la mesure invariante revient donc à calculer le volume de  $\pi^{-1}(\alpha_1, \dots, \alpha_{n-1})$ , c'est-à-dire à voir combien il y a de hauteurs possibles pour des bases données des rectangles.

Il convient peut-être d'insister sur le fait que la seule chose qui n'est pas algorithmique ici est le choix d'un domaine  $D$  sur lequel la fonction  $\tilde{F}$  est bijective; une fois qu'on a exhibé  $D$ , tout le reste n'est que du calcul, mais ce domaine semble en général difficile à trouver.

*Remarque.* — La transformation  $f$  possède beaucoup de mesures invariantes (en particulier, ses points périodiques sont denses); on montrera que la mesure définie ici est ergodique, c'est donc la seule mesure invariante équivalente à la mesure de Lebesgue. Quand on parlera ci-dessous de «la mesure invariante pour  $f$ », il s'agira d'un abus de langage pour désigner l'unique (à homothétie près) mesure invariante pour  $f$  absolument continue par rapport à la mesure de Lebesgue.

### 3. Un exemple simple : la mesure invariante pour l'algorithme des fractions continues

a) Un algorithme additif.

On considère l'application suivante, définie sur le cône  $0 < \alpha < \beta$  de  $\mathbb{R}^2$ :

$$\begin{aligned} F: (\alpha, \beta) &\mapsto (\alpha, \beta - \alpha) & \text{si } \alpha < \beta - \alpha \\ &\mapsto (\beta - \alpha, \alpha) & \text{si } \alpha > \beta - \alpha \end{aligned}$$

qui est un algorithme vectoriel additif correspondant à l'application

$$\begin{aligned} f: [0, 1] &\rightarrow [0, 1] \\ x &\mapsto \frac{1}{1-x} - 1 & \text{si } x < \frac{1}{2} \\ &\mapsto \frac{1}{x} - 1 & \text{si } x > \frac{1}{2} \end{aligned}$$

en posant  $x = \alpha/\beta$ .

C'est un algorithme très simple, qui correspond du point de vue dynamique à l'induction d'une rotation de l'intervalle, échangeant deux intervalles de longueurs  $\alpha$  et  $\beta$ , sur le complémentaire de l'image du petit intervalle, avec un changement d'orientation quand le petit intervalle devient le plus grand. L'itération de cet algorithme redonne l'algorithme classique des fractions continues. On représente cette induction, comme expliqué plus haut, par l'application de l'ensemble  $D = \{(\alpha, \beta, \lambda, \mu) \mid \alpha\lambda + \beta\mu = 1; 0 < \alpha < \beta, 0 < \lambda, 0 < \mu\}$

de  $\mathbb{R}^4$ :

$$\begin{aligned}\tilde{F}: (\alpha, \beta, \lambda, \mu) &\mapsto (\alpha, \beta - \alpha, \lambda + \mu, \mu) & \text{si } \alpha < \beta - \alpha \\ &\mapsto (\beta - \alpha, \alpha, \mu, \lambda + \mu) & \text{si } \alpha > \beta - \alpha\end{aligned}$$

Cette application se projette sur la précédente, elle est évidemment bijective (l'injectivité se déduit immédiatement du fait que la tour la plus haute est celle sur laquelle on vient d'induire, ce qui permet de trouver l'image réciproque) et elle préserve la mesure de Lebesgue. De même, le flot  $\Phi_t$  défini ci-dessus préserve cette mesure.

On peut prendre les nouvelles coordonnées  $a = \alpha\lambda$ ,  $b = \beta\lambda$ ,  $\tau = \log(\beta)$ ,  $e = \alpha\lambda + \beta\mu$ , on vérifie que l'action du flot dans ces coordonnées consiste à ajouter  $t$  à  $\tau$  et à laisser fixe les autres. D'autre part, la mesure de Lebesgue s'écrit  $1/b \, da \, db \, d\tau \, de$  dans ces coordonnées; puisque le flot et l'application préservent  $e$  (c'est la surface des tours) on peut se restreindre à l'hypersurface  $e = 1$ , sur laquelle on a une mesure invariante naturelle, et considérer, comme défini plus haut, l'application de premier retour à la surface de section  $\beta = 1$ , ou  $\tau = 0$ , qui possède la mesure invariante  $1/b \, da \, db$ ; on peut revenir aux coordonnées  $\alpha, \lambda$  en utilisant le fait que  $\beta = 1$  donc  $b = \lambda$ , et on trouve comme mesure invariante  $d\alpha \, d\lambda$  (on peut en fait donner directement l'application de premier retour et prouver qu'elle préserve la mesure de Lebesgue; le calcul précédent montre que l'on peut obtenir ce résultat de façon mécanique).

On veut maintenant projeter sur la coordonnée  $\alpha$  pour retrouver l'application de départ; il faut pour cela intégrer  $d\lambda$ , mais compte tenu de l'équation  $\alpha\lambda + \beta\mu = 1$ , on voit que  $\lambda$  doit être compris entre 0 et  $1/\alpha$ , ce qui donne finalement pour mesure invariante:

$$\frac{1}{\alpha} d\alpha$$

*Remarque.* — On aurait pu aussi présenter cet algorithme comme la fonction définie sur le cône positif de  $\mathbb{R}^2$  qui à  $(\alpha, \beta)$  associe  $(\alpha - \beta, \beta)$  si  $\alpha > \beta$ , et  $(\alpha, \beta - \alpha)$  sinon; on peut refaire exactement la même construction, et si l'on prend comme section la surface  $\Sigma$  d'équation  $\beta = 1$ , on trouve pour application  $f$  la fonction:

$$\begin{aligned}f: \mathbb{R} &\rightarrow \mathbb{R} \\ \alpha &\mapsto \alpha - 1 & \text{si } \alpha > 1 \\ &\mapsto \frac{\alpha}{1 - \alpha} & \text{si } \alpha < 1\end{aligned}$$

Cette fonction induit sur l'intervalle  $[0, 1]$  la « transformation inversée des fractions continues »,  $\alpha \mapsto \{1/(1 - \alpha)\}$  (cf. [AF1]), et le même calcul que ci-dessus montre qu'elle admet  $1/\alpha \, d\alpha$  comme mesure invariante, on retrouve ainsi le résultat de [AF1].

b) L'algorithme classique des fractions continues.

Considérons l'application :

$$F: \{(\alpha, \beta) \mid 0 < \alpha < \beta\} \rightarrow \{(\alpha, \beta) \mid 0 < \alpha < \beta\}$$

$$(\alpha, \beta) \mapsto \left( \beta - \left\lfloor \frac{\beta}{\alpha} \right\rfloor \alpha, \alpha \right)$$

qui correspond, en prenant pour coordonnées  $x = \alpha/\beta$ , à l'application  $f: x \mapsto \{1/x\}$ .

On peut, comme ci-dessus, représenter cette application par une induction (on retire autant de fois que possible l'intervalle  $\alpha$  à l'intervalle  $\beta$ , et on inverse les intervalles), et il faut remarquer que, dans cette construction, la hauteur de la première tour est toujours plus petite que celle de la seconde, ce qui explique la restriction du domaine de l'application ci-dessous, nécessaire pour la surjectivité :

$$\tilde{F}: \{(\alpha, \beta, \lambda, \mu) \mid 0 < \alpha < \beta, 0 < \lambda < \mu\} \rightarrow \{(\alpha, \beta, \lambda, \mu) \mid 0 < \alpha < \beta, 0 < \lambda < \mu\}$$

$$(\alpha, \beta, \lambda, \mu) \mapsto \left( \beta - \left\lfloor \frac{\beta}{\alpha} \right\rfloor \alpha, \alpha, \mu, \left\lfloor \frac{\beta}{\alpha} \right\rfloor \mu + \lambda \right)$$

L'application de premier retour à la surface  $\Sigma$  d'équation  $\beta = 1$  se calcule facilement : on part de  $(\alpha, 1, \lambda, \mu)$ , on applique  $\tilde{F}$ , ce qui donne  $(1 - [1/\alpha], \alpha, \mu, \lambda + [1/\alpha]\mu)$ , et on normalise par le flot pour obtenir  $(\{1/\alpha\}, 1, \alpha\mu, \alpha(\lambda + [1/\alpha]\mu))$ ; si l'on tient compte du fait que  $\alpha\lambda + \beta\mu = 1$ , donc  $\mu = 1 - \alpha\lambda$  sur la surface  $\Sigma$ , et de ce que  $\lambda < \mu$ , ce qui entraîne  $0 \leq \lambda < 1/(1 + \alpha)$  sur  $\Sigma$ , on voit que, en prenant les coordonnées  $(\alpha, \lambda)$  sur  $\Sigma$ , on a :

$$\tilde{f}: \Sigma \rightarrow \Sigma = \{(\alpha, \lambda) \mid 0 < \alpha < 1; 0 < \lambda < 1/(1 + \alpha)\}$$

$$(\alpha, \lambda) \mapsto \left( \left\{ \frac{1}{\alpha} \right\}, \alpha - \alpha^2 \lambda \right)$$

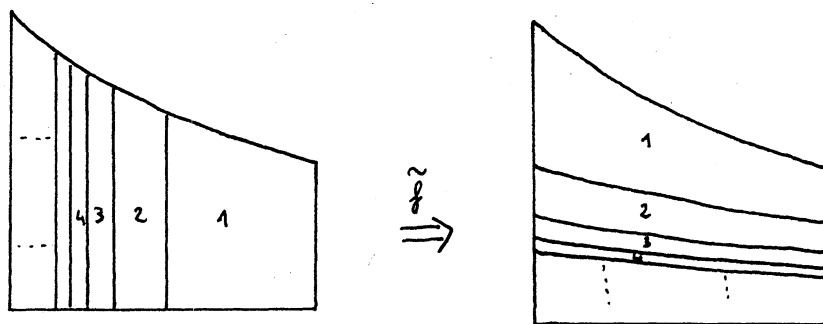


Fig. 2

Cette application est continue sur un nombre dénombrable de bandes verticales qui composent son domaine de définition, et envoie toutes ces bandes verticales sur des

bandes horizontales superposées mais disjointes (cf. Fig. 2). On voit directement en calculant le Jacobien qu'elle préserve la mesure de Lebesgue; en projetant sur la première coordonnée on trouve la transformation des fractions continues, et sa mesure invariante (mesure de Gauss) en donnant la mesure des  $\lambda$  admissibles pour  $\alpha$  fixé :

$$\frac{1}{1+\alpha} d\alpha$$

Cette mesure est de masse totale finie, contrairement à la précédente.

#### 4. Une interprétation géométrique : coordonnées pour le flot géodésique sur l'espace modulaire.

On sait (cf. [AF], [BKS], [Se]) que le flot géodésique sur la surface modulaire  $SL(2, \mathbb{Z}) \backslash \mathbb{H}$  est relié à la transformation des fractions continues; de fait, le flot  $\Phi_t$  que nous avons décrit dans la section précédente est presque le flot géodésique sur la surface modulaire: il suffit de faire un revêtement à 2 feuillets et de remplacer  $t$  par  $t/2$ .

Plus précisément, soit  $D = \{(\alpha, \beta, \lambda, \mu) \mid \alpha\lambda + \beta\mu = 1; (0 < \alpha < \beta \text{ et } 0 < \lambda < \mu) \text{ ou } (0 < \beta < \alpha \text{ et } 0 < \mu < \lambda)\}$  l'ensemble des couples de rectangles où le plus haut est aussi le plus large (mais on n'impose plus que le second soit le plus haut), et définissons la fonction  $\tilde{F}$  par :

$$\begin{aligned} (\alpha, \beta, \lambda, \mu) &\mapsto \left( \alpha, \beta - \left\lfloor \frac{\beta}{\alpha} \right\rfloor \alpha, \lambda + \left\lfloor \frac{\beta}{\alpha} \right\rfloor \mu, \mu \right) & \text{si } \alpha < \beta \\ (\alpha, \beta, \lambda, \mu) &\mapsto \left( \alpha - \left\lfloor \frac{\alpha}{\beta} \right\rfloor \beta, \beta, \lambda, \mu + \left\lfloor \frac{\alpha}{\beta} \right\rfloor \lambda \right) & \text{si } \beta < \alpha \end{aligned}$$

et le flot  $\Phi_t$  par :

$$\Phi_t(\alpha, \beta, \lambda, \mu) = (\alpha e^{t/2}, \beta e^{t/2}, \lambda e^{-t/2}, \mu e^{-t/2}).$$

On vérifie facilement que l'ensemble des points de  $D$  qui satisfont à  $\alpha < 1 \leq \beta$  ou à  $\beta < 1 \leq \alpha$  forme un domaine fondamental pour l'action de  $\tilde{F}$  sur  $D$ ; nous allons montrer que cet ensemble forme un système de coordonnées globales sur le fibré unitaire tangent à la surface modulaire, et que le flot  $\Phi_t$  est le flot géodésique exprimé dans ces coordonnées.

On peut prouver cela de diverses manières; nous allons utiliser le fait que la surface modulaire s'identifie à l'espace modulaire du tore (espace des structures plates sur le tore modulo les difféomorphismes, voir ci-dessous), et que le flot géodésique s'identifie au flot de Teichmüller; au prix de définitions un peu lourdes, on supprime ainsi tous les calculs, ce qui ne serait pas le cas dans d'autres approches.

Rappelons tout d'abord que l'espace de Teichmüller d'une surface est l'ensemble des structures conformes sur cette surface, modulo les difféomorphismes isotopes à l'identité (par le théorème d'uniformisation, c'est aussi l'ensemble des métriques à courbure constante 1, -1 ou 0 suivant le type topologique de la surface, modulo les difféomorphismes isotopes à l'identité, et à une constante multiplicative près dans le cas de la

courbure 0). Cet espace est muni d'une métrique naturelle, la distance de Teichmüller, donnée par le logarithme du minimum des rapports de quasi-conformité pour les applications quasi-conformes permettant de passer d'une classe de structures conformes à une autre. On peut ainsi définir des géodésiques et un flot géodésique naturel, le flot de Teichmüller, qui décrit une façon canonique de déformer des structures conformes. On peut montrer en étudiant ces déformations (« applications quasi-conformes extrémales ») que le fibré tangent s'identifie naturellement à l'ensemble des couples de feuilletages mesurés transverses à difféomorphisme isotope à l'identité près, et l'action du flot de Teichmüller sur le fibré tangent consiste, dans cette identification, à multiplier la mesure de l'un des feuilletages par  $e^{t/2}$  et l'autre par  $e^{-t/2}$ . Au lieu de prendre dans ces définitions des difféomorphismes isotopes à l'identité, on peut quotienter par tous les difféomorphismes : on obtient alors l'espace modulaire, quotient de l'espace de Teichmüller par le groupe modulaire de la surface (groupe des classes de difféomorphismes modulo les difféomorphismes isotopes à l'identité).

La démonstration consiste alors à rappeler que l'espace de Teichmüller du tore, muni de sa métrique naturelle, est isométrique au plan hyperbolique, que le groupe modulaire du tore est  $SL(2, \mathbb{Z})$ , la surface modulaire est donc aussi l'espace modulaire du tore, et le flot de Teichmüller est confondu avec le flot géodésique (pour une excellente introduction à ces résultats, voir [Mo]). L'identification du fibré tangent à l'espace modulaire à l'ensemble des couples de feuilletages mesurés transverses à difféomorphisme près, (et donc, sur le tore, à l'ensemble des couples de feuilletages mesurés linéaires transverses) va nous donner un moyen de paramétrer simplement le flot géodésique sur la surface modulaire.

On peut mettre des coordonnées sur ces couples de feuilletages de la façon suivante : appelons horizontales les feuilles du premier feuilletage, verticales les feuilles du second, orientons ces deux feuilletages (on vérifiera à la fin de la construction que l'orientation choisie n'a pas d'importance) et choisissons un point  $x$  sur le tore. On prend maintenant un segment horizontal de longueur 2 centré en  $x$ , on suit la feuille verticale issue de  $x$  jusqu'à rencontrer ce segment en un point  $y$ , on obtient ainsi un segment vertical de  $x$  à  $y$ . Le segment horizontal initial est séparé en deux parties par  $y$  ; on oublie la partie la plus courte, et on prolonge l'autre, qui contient le point  $x$ , jusqu'à recouper le segment vertical en un point  $z$ . Enfin on poursuit ce segment vertical jusqu'à recouper une deuxième fois le segment horizontal ainsi prolongé. Ces deux segments découpent le tore en deux rectangles, de base  $xy$  et  $xz$ , le plus haut étant aussi le plus large (cf. Fig. 3, où on a représenté les opérations décrites sur le revêtement universel  $\mathbb{R}^2$  du tore). On a bien associé à un couple de feuilletages un quadruplet dans le domaine fondamental ; réciproquement, à un tel quadruplet on associe un couple de rectangles qui donnent un pavage du plan avec deux feuilletages distingués, par les horizontales et les verticales, donc un couple de feuilletages sur le tore.

On sait que l'action du flot de Teichmüller sur les couples de feuilletages consiste à multiplier l'une des mesures par  $e^{t/2}$  et l'autre par  $e^{-t/2}$ , ce qui s'exprime bien de la façon donnée ci-dessus pour les coordonnées des rectangles associés.

*Remarque.* — Cette construction est en fait le cas le plus simple des « *zippered rectangles* » de Veech (cf. [Ve1]). On peut paramétrer d'une façon analogue le flot de

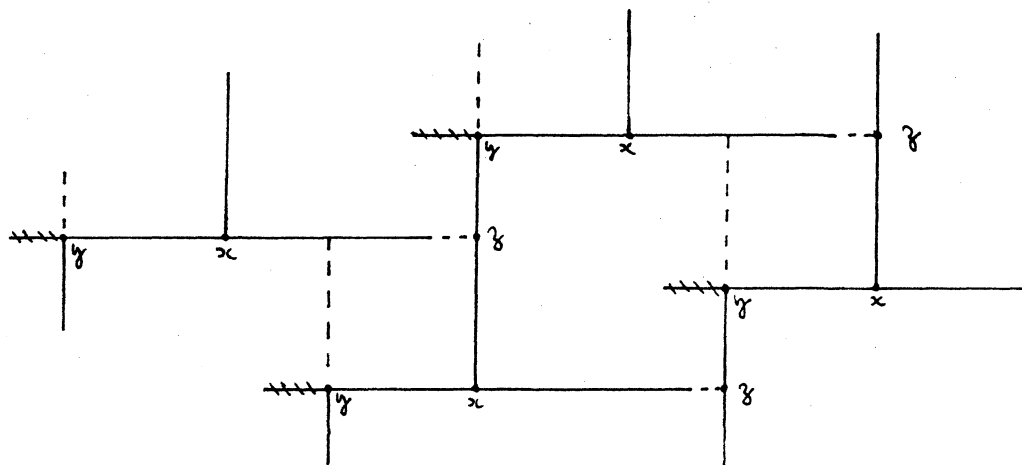


Fig. 3

Teichmüller sur une surface de genre quelconque, mais les calculs sont beaucoup plus complexes (cf. [Ve2]).

On préfère d'habitude représenter le fibré unitaire tangent à la surface modulaire comme  $\mathrm{PSL}(2, \mathbb{Z}) \backslash \mathrm{PSL}(2, \mathbb{R})$ ; on peut le faire ici, en associant au quadruplet  $(\alpha, \beta, \lambda, \mu)$  la matrice

$$\begin{pmatrix} \alpha & \mu \\ -\beta & \lambda \end{pmatrix}$$

qui est bien de déterminant 1. On vérifie facilement que  $\tilde{F}$  revient à une multiplication à gauche par une matrice du type  $\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$  ou  $\begin{pmatrix} 1 & 0 \\ n & 1 \end{pmatrix}$ , ce qui correspond au passage au quotient par  $\mathrm{PSL}(2, \mathbb{Z})$ , et que  $\Phi_t$  revient à une multiplication à droite par une matrice  $\begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$ , qui est bien la traduction matricielle du flot géodésique. La démonstration que nous avons faite ci-dessus nous évite d'avoir à prouver que les coordonnées que nous avons prises définissent un domaine fondamental pour l'action de  $\mathrm{PSL}(2, \mathbb{Z})$ . Ce domaine fondamental se représente mal géométriquement : ce n'est pas l'image réciproque d'un domaine fondamental pour l'action du groupe modulaire sur le plan hyperbolique, et la surface de section  $\Sigma$  ne se projette pas sur une courbe; mais c'est peut-être une condition nécessaire pour avoir un bon codage : contrairement à [AF], nous n'avons pas besoin de redresser l'application de premier retour pour avoir une partition markovienne (cf. Fig. 2).

**5. Une interprétation géométrique :  
la constante de Lévy comme moyenne des temps de premier retour en une section**

On note  $T(x) = \{1/x\}$  l'application associée au développement en fraction continue usuel ; rappelons que, pour tout  $x$  de  $[0, 1]$ , les quotients partiels de  $x$  sont les entiers  $a_n = [T^n(x)]$ , et les réduites  $p_n/q_n$  sont les meilleures approximations rationnelles de  $x$ , données par le début du développement en fraction continue de  $x$  :

$$\frac{p_n}{q_n} = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots + \frac{1}{a_n}}}}$$

Les dénominateurs  $q_n$  sont définis par la formule de récurrence  $q_{n+1} = a_{n+1}q_n + q_{n-1}$ ,  $q_0 = 1$ ,  $q_1 = a_1$ , où les  $a_i$  sont les quotients partiels du nombre considéré.

On sait (voir [Kh], [Le], [Bi]), que pour presque tout nombre réel les dénominateurs des réduites de ce nombre satisfont :

$$\lim_{n \rightarrow \infty} \frac{\log q_n}{n} \rightarrow \frac{\pi^2}{12 \log 2}$$

Nous allons donner une interprétation géométrique de cette constante.

Tout d'abord, les quotients partiels qui permettent de définir les  $q_n$  par récurrence interviennent aussi dans l'application  $\tilde{F}$ , avec les mêmes formules de récurrence, ce qui montre que, avec les notations du début de cette section,  $\tilde{F}^n(a, 1, 0, 1)$  est de la forme  $(x, y, q_n, q_{n+1})$  ou  $(x, y, q_{n+1}, q_n)$  suivant la parité de  $n$ .

D'autre part, si  $(\alpha, \beta, \lambda, \mu)$  appartient à la surface de section  $\Sigma$  définie par  $\sup(\alpha, \beta) = 1$ , on a  $1/2 < \sup(\lambda, \mu) < 1$ , puisque  $\alpha\lambda + \beta\mu = 1$ , et que  $\alpha > \beta$  si et seulement si  $\lambda > \mu$ .

Soit maintenant  $x = (\alpha, 1, 0, 1)$  un point de  $\Sigma$ , on s'intéresse à la suite  $\tilde{f}^n(x)$  des retours successifs de  $x$  en  $\Sigma$  ; on notera  $\tau(x)$  le temps de retour ; par définition, on a :

$$\tilde{f}^n(x) = \tilde{F} \Phi_{\tau(\tilde{f}^{n-1}(x))} \tilde{F} \dots \Phi_{\tau(\tilde{f}(x))} \tilde{F}(x)$$

Comme  $\Phi_t$  et  $\tilde{F}$  commutent, on peut encore écrire :

$$\tilde{f}^n(x) = \Phi_{\tau(x) + \tau(\tilde{f}(x)) + \dots + \tau(\tilde{f}^{n-1}(x))} \tilde{F}^n(x)$$

et la plus grande des deux hauteurs, d'après ce que nous avons vu, est

$$q_n e^{-(\tau(x) + \dots + \tau(\tilde{f}^{n-1}(x)))/2}$$

Mais cette hauteur est comprise entre  $1/2$  et  $1$  ; en prenant les logarithmes, on voit que  $(\log q_n)/n$  a même limite que la moitié de la somme de Birkhoff des temps de retour

$\frac{1}{n} \sum_{i=0}^{n-1} \tau(\tilde{f}^i(x))$ , si elle existe. Mais on sait que la mesure invariante pour  $\tilde{f}$  calculée ci-

dessus (c'est la mesure de Lebesgue dans les coordonnées  $\alpha, \lambda$ ) est ergodique; donc ces sommes de Birkhoff tendent presque partout vers la moyenne des temps de retour sur la surface  $\Sigma$  (en toute rigueur, on ne peut pas appliquer cela directement aux points  $(\alpha, 1, 0, 1)$ , qui sont de mesure nulle dans  $\Sigma$ , mais il est clair que ces sommes ne dépendent que de la première coordonnée, qui détermine complètement le temps de premier retour; Puisqu'on a convergence presque partout sur  $\Sigma$ , et que cette convergence ne dépend que de  $\alpha$ , par Fubini, on a convergence presque partout en  $\alpha$ ). C'est un fait général que l'intégrale du temps de retour sur la surface donne le volume de l'espace total, la moyenne est donc le quotient de ce volume par la masse de  $\Sigma$ .

L'aire de la surface modulaire est  $\pi/3$  (il y a un domaine fondamental qui est un triangle géodésique d'angles  $\pi/3, \pi/3$  et  $0$ , et on applique la formule de Gauss-Bonnet); le volume du fibré unitaire tangent est donc  $\pi^2/3$ , car la fibre en chaque point est de longueur  $\pi$  (L'action de  $SL(2, \mathbb{R})$  sur le plan hyperbolique n'est pas fidèle, car  $-Id$  agit trivialement, et le fibré unitaire tangent est en fait isomorphe à  $PSL(2, \mathbb{R})$ ; en particulier, la fibre en chaque point est isomorphe à  $PSO(2, \mathbb{R})$ , de longueur  $\pi$ , et non à  $SO(2, \mathbb{R})$ ; la surface  $\Sigma$  est un revêtement double de  $\{(\alpha, \lambda) \mid 0 < \alpha < 1; 0 < \lambda < 1/(1+\alpha)\}$ , donc de volume  $2 \log 2$  pour la mesure de Lebesgue, ce qui prouve bien que  $(\log q_n)/n$  converge presque partout vers  $\pi^2/(12 \log 2)$ .

*Remarque.* — A tout nombre quadratique de  $[0, 1]$  dont le développement en fraction continu est périodique, on peut faire correspondre sa hauteur, c'est-à-dire la longueur de la géodésique associée. On peut aussi montrer que, pour tout nombre de ce type, il existe une constante de Lévy, c'est-à-dire que  $\log q_n/n$  tend vers une constante, dépendant bien sûr du nombre. Faivre [Fa] a montré que la moyenne de ces constantes sur les nombres quadratiques périodiques purs de hauteur inférieures à  $C$  tend vers la constante de Lévy  $\pi^2/(12 \log 2)$  quand  $C$  tend vers l'infini. Ce résultat peut se déduire, dans notre approche, en remarquant que la constante de Lévy d'un nombre quadratique est la moyenne des temps de retour sur une orbite périodique, qu'il s'agit donc de faire la moyenne des temps de retour sur tous les nombres quadratiques de hauteur inférieure à  $C$ . Mais on sait (cf. [Po]) que ces nombres quadratiques sont répartis asymptotiquement suivant la mesure de Gauss; compte tenu du fait que la fonction temps de premier retour est continue par morceaux, on en déduit que la moyenne sur les nombres quadratiques doit tendre vers la moyenne sur l'intervalle suivant la mesure de Gauss, d'où le résultat.

## 6. Mesures invariantes pour les algorithmes de Brun et Selmer

Nous considérons le cas de trois nombres, et nous ferons les calculs en détail; le cas de  $n$  nombres se fait exactement par les mêmes techniques, mais la présentation est moins agréable, nous nous contenterons de donner les résultats.

a. *L'algorithme additif de Brun.*

Cet algorithme consiste, étant donnés trois nombres, à retirer le deuxième plus grand du plus grand, et à itérer le procédé; on a donc, en plaçant les nombres par ordre croissant, 3 cas possibles, suivant la valeur de  $\gamma - \beta$ , et la fonction correspondante est :

$$F: \{(\alpha, \beta, \gamma) \mid 0 < \alpha < \beta < \gamma\} \rightarrow \{(\alpha, \beta, \gamma) \mid 0 < \alpha < \beta < \gamma\}$$

$$\begin{aligned} (\alpha, \beta, \gamma) &\mapsto (\alpha, \beta, \gamma - \beta) && \text{si } \beta < \gamma - \beta \\ &\mapsto (\alpha, \gamma - \beta, \beta) && \text{si } \alpha < \gamma - \beta < \beta \\ &\mapsto (\gamma - \beta, \alpha, \beta) && \text{si } \gamma - \beta < \alpha \end{aligned}$$

Si l'on essaie de représenter cet algorithme avec des tours, on voit que la troisième tour doit être plus haute que la première, et que cette condition se conserve par induction, alors qu'il n'y a pas de condition sur la hauteur de la deuxième tour: c'est justement la hauteur de cette deuxième tour qui permet d'inverser la fonction. De façon plus explicite, l'algorithme inverse est le même que l'algorithme direct, mais sur les hauteurs: on prend la tour la plus haute, on découpe au sommet un rectangle de même hauteur que la tour intermédiaire, et on juxtapose ce rectangle à la tour intermédiaire, qui devient ainsi la plus large (cf. Fig. 4). On obtient donc la fonction suivante:

$$\tilde{F}: D \rightarrow D = \{(\alpha, \beta, \gamma, \lambda, \mu, \nu) \mid \alpha\lambda + \beta\mu + \gamma\nu = 1; 0 < \alpha < \beta < \gamma; 0 < \lambda < \nu; 0 < \mu\}$$

$$\begin{aligned} (\alpha, \beta, \gamma, \lambda, \mu, \nu) &\mapsto (\alpha, \beta, \gamma - \beta, \lambda, \mu + \nu, \nu) && \text{si } \beta < \gamma - \beta \\ &\mapsto (\alpha, \gamma - \beta, \beta, \lambda, \nu, \mu + \nu) && \text{si } \alpha < \gamma - \beta < \beta \\ &\mapsto (\gamma - \beta, \alpha, \beta, \nu, \lambda, \mu + \nu) && \text{si } \gamma - \beta < \alpha \end{aligned}$$

dont on vérifie facilement qu'elle est bijective sur le domaine  $D$ . On peut alors définir un flot  $\Phi_t$  et le faire passer au quotient  $\Omega = D/\tilde{F}$ , et considérer l'application de premier retour  $\tilde{f}$  sur la surface de section (ici de dimension 4) d'équation  $\gamma = 1$ .

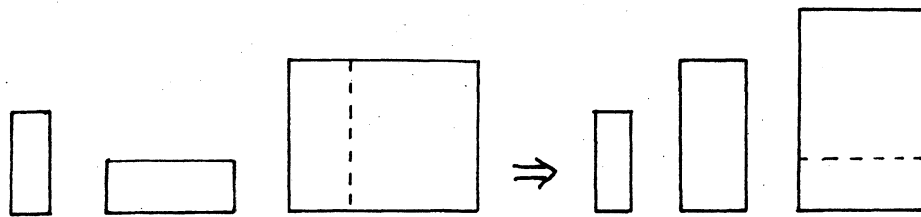


Fig. 4

Il n'est pas facile ici d'écrire  $\tilde{f}$  directement car il y a plusieurs cas. Il est plus simple de trouver mécaniquement la mesure invariante, en remplaçant les coordonnées de départ par  $(\alpha, \beta, \lambda, \mu, \tau = \log \gamma, e = \alpha\lambda + \beta\mu + \gamma\nu)$ . Dans ces coordonnées la mesure de Lebesgue s'écrit simplement  $d\alpha d\beta d\lambda d\mu d\tau de$ , le domaine  $D$  est contenu dans l'hypersurface  $e = 1$ , on a donc une mesure invariante sur  $D$  qui est simplement la mesure de Lebesgue dans

les coordonnées  $(\alpha, \beta, \lambda, \mu, \tau)$ . Puisque la surface de section  $\Sigma$  est d'équation  $\tau=0$  dans ces coordonnées, la mesure invariante pour  $\tilde{f}$  est la mesure de Lebesgue dans les coordonnées  $(\alpha, \beta, \lambda, \mu)$ , car l'effet du flot sur la coordonnée  $\tau$  est une translation de  $t$ .

On a donc trouvé la mesure invariante pour  $\tilde{f}$ ; pour trouver la densité de la mesure invariante pour  $f$ , il suffit de projeter cette mesure sur les deux premières coordonnées, c'est-à-dire de trouver la mesure de l'ensemble de  $(\lambda, \mu)$  admissibles pour  $(\alpha, \beta)$  donnés. Cet ensemble est caractérisé, compte tenu de ce que  $\gamma=1$  sur  $\Sigma$ , par les équations :

$$v > \lambda > 0, \quad \mu > 0, \quad \alpha\lambda + \beta\mu + v = 1$$

ou encore :

$$\lambda > 0, \quad \mu > 0, \quad (\alpha + 1)\lambda + \beta\mu < 1$$

c'est donc un triangle d'aire  $\beta^{-1}(1+\alpha)^{-1}$ , ce qui donne la densité de la mesure invariante pour  $f$ . La masse totale de cette mesure est finie, égale à :

$$\int_{0 < \alpha < \beta < 1} \frac{1}{\beta(1+\alpha)} d\alpha d\beta = \frac{\pi^2}{12}$$

On trouve bien pour la mesure de probabilité invariante pour  $f$  la densité annoncée dans l'introduction :

$$\frac{12}{\pi^2} \frac{1}{\beta(1+\alpha)}$$

#### b. L'algorithme multiplicatif de Brun

Il s'agit d'une version itérative du précédent : étant donnés 3 nombres  $\alpha < \beta < \gamma$ , on remplace  $\gamma$  par  $\gamma - [\gamma/\beta]\beta$  et on réordonne ; le relèvement à  $\mathbb{R}^6$  est :

$$\begin{aligned} \tilde{F}: D \rightarrow D = \{(\alpha, \beta, \gamma, \lambda, \mu, v) \mid \alpha\lambda + \beta\mu + \gamma v = 1; 0 < \alpha < \beta < \gamma; 0 < \lambda < v; 0 < \mu < v\} \\ (\alpha, \beta, \gamma, \lambda, \mu, v) \mapsto (\alpha, \gamma - [\gamma/\beta]\beta, \beta, \lambda, v, \mu[\gamma/\beta]v) \quad \text{si } \gamma - [\gamma/\beta]\beta > \alpha \\ (\alpha, \beta, \gamma, \lambda, \mu, v) \mapsto (\gamma - [\gamma/\beta]\beta, \alpha, \beta, v, \lambda, \mu[\gamma/\beta]v) \quad \text{sinon} \end{aligned}$$

On vérifie que  $\tilde{F}$  est bijective sur  $D$ , l'application réciproque de  $\tilde{F}$  consiste à appliquer le même algorithme, mais cette fois sur les hauteurs, et la condition supplémentaire par rapport à l'algorithme additif ( $\mu < v$ ) assure que l'on peut retrouver la partie entière  $[\gamma/\beta]$  à partir des hauteurs de l'image. En prenant comme section du flot  $\Sigma$  définie par  $\gamma=1$ , on vérifie comme ci-dessus qu'en prenant les coordonnées  $\alpha, \beta, \lambda, \mu$  sur  $\Sigma$  l'application de premier retour préserve la mesure de Lebesgue. Ces coordonnées sont soumises ici aux conditions  $\alpha < \beta < 1$ , et  $\alpha\lambda + \beta\mu + v = 1$ , avec  $0 < \lambda < v$  et  $0 < \mu < v$ .

On voit que, pour  $(\alpha, \beta)$  fixés, l'ensemble des  $(\lambda, \mu)$  admissibles est réunion de deux triangles (cf. Fig. 5), suivant que  $\lambda < \mu$  ou non,  $T_1$  d'équation  $0 < \lambda < \mu$ ,  $\alpha\lambda + (\beta+1)\mu < 1$  et d'aire  $(1+\beta)^{-1}(1+\alpha+\beta)^{-1}$  et  $T_2$  d'équation  $0 < \mu < \lambda$ ,  $(\alpha+1)\lambda + \beta\mu < 1$  et d'aire  $(1+\alpha)^{-1}(1+\alpha+\beta)^{-1}$ .

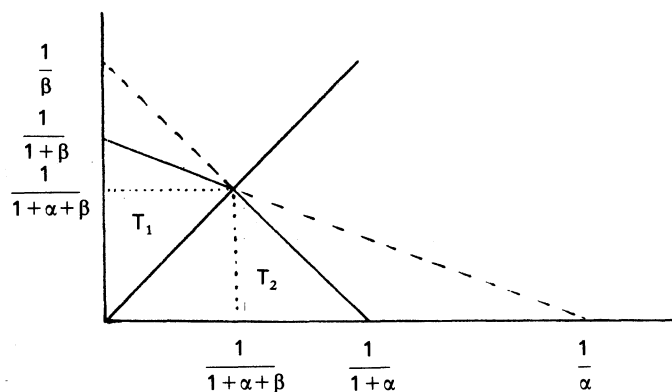


Fig. 5

Quand on projette sur les deux premières coordonnées, on obtient donc une mesure invariante de densité

$$\frac{1}{1+\beta} \frac{1}{1+\alpha+\beta} + \frac{1}{1+\alpha} \frac{1}{1+\alpha+\beta}$$

Cette densité est finie sur le domaine considéré, donc sa masse totale est finie.

c. *Généralisation en dimension n.*

On peut définir un algorithme multiplicatif de Brun en dimension  $n$ , qui consiste à prendre un  $n$ -uplet vérifiant  $0 < \alpha_1 < \dots < \alpha_{n-1} < \alpha_n$ , à remplacer  $\alpha_n$  par  $\alpha_n - [\alpha_n/\alpha_{n-1}] \alpha_{n-1}$  et à réordonner.

On peut généraliser la construction faite plus haut; la fonction  $\tilde{F}$  sera définie sur le domaine  $D$  d'équation :

$$\left\{ (\alpha_1, \dots, \alpha_n, \lambda_1, \dots, \lambda_n) \left| \sum_{i=1}^n \alpha_i \lambda_i = 1; 0 < \alpha_1 < \dots < \alpha_n; 0 < \lambda_i < \lambda_n, 1 \leq i < n \right. \right\}$$

On peut démontrer que  $\tilde{F}$  est bijective sur  $D$ . On prend comme ci-dessus la surface de section  $\Sigma$  d'équation  $\alpha_n = 1$ , on prouve de la même façon que l'application de premier retour en  $\Sigma$  du flot sur  $\Omega = D/\tilde{F}$  préserve la mesure de Lebesgue dans les coordonnées  $(\alpha_1, \dots, \alpha_{n-1}, \lambda_1, \dots, \lambda_{n-1})$ . Pour trouver la densité de la mesure invariante pour  $f$ , il s'agit encore de trouver la mesure de l'ensemble des hauteurs admissibles pour  $(\alpha_1, \dots, \alpha_{n-1})$  fixés. Cet ensemble est défini par les équations :

$$\sum_{i=1}^{n-1} \alpha_i \lambda_i + \lambda_n = 1$$

$$0 < \lambda_i < \lambda_n \quad i = 1, \dots, n-1.$$

Cet ensemble est partitionné en  $(n-1)!$  simplexes par les hyperplans  $\lambda_i = \lambda_j$ , on calcule le volume de chacun de ces simplexes (par exemple en prenant le déterminant de leurs sommets), ce qui donne une mesure invariante de densité :

$$\sum_{\sigma \in S_{n-1}} \prod_{i=1}^{n-1} \frac{1}{1 + \sum_{j=1}^i \alpha_{\sigma(j)}}$$

On peut remarquer que cette densité est bornée, donc la mesure est de masse totale finie.

On pourrait de même définir un algorithme additif de Brun en dimension  $n$ , et chercher la mesure invariante; les calculs sont les mêmes, la seule différence étant que l'on supprime l'équation  $\lambda_{n-1} < \lambda_n$  dans la définition de  $D$ , donc les simplexes dont il faut calculer le volume sont différents, et on trouve pour densité :

$$\frac{1}{\alpha_{n-1}} \sum_{\sigma \in S_{n-2}} \prod_{i=1}^{n-2} \frac{1}{1 + \sum_{j=1}^i \alpha_{\sigma(j)}}$$

Bien que cette densité ne soit pas bornée, la masse totale de la mesure est encore finie, sauf si  $n=2$ , cas de l'algorithme additif des fractions continues.

#### d. L'algorithme de Selmer.

Cet algorithme est une modification simple de celui de Brun : il consiste, étant donnés 3 nombres  $0 < \alpha < \beta < \gamma$ , à retirer le plus petit du plus grand. Il y a ici une difficulté supplémentaire, car on vérifie facilement que la région  $\Lambda$  définie par l'équation supplémentaire  $\gamma < \alpha + \beta$  est invariante, et que son complémentaire est une région transitoire, tout triplet finissant par arriver dans  $\Lambda$  (voir [Sc4] pour une étude plus détaillée); le support d'une mesure invariante finie est donc forcément contenu dans  $\Lambda$ .

En se restreignant à  $\Lambda$ , on peut appliquer la méthode précédente. On définit une application  $\tilde{F}$  sur  $\mathbb{R}^6$ , on vérifie facilement que cette application est bijective (le seul problème est de placer  $\gamma - \alpha$  par rapport à  $\alpha$ , et l'incertitude est levée par les hauteurs correspondantes), on se restreint au domaine d'équation  $\alpha\lambda + \beta\mu + \gamma\nu = 1$ , on prend pour hypersurface de section  $\gamma = 1$ , et on obtient une application de premier retour sur 4 variables qui conserve la mesure de Lebesgue et se projette sur l'application correspondant à l'algorithme de Selmer.

Les seules conditions sur les hauteurs sont  $\alpha\lambda + \beta\mu < 1$ , on intègre facilement et on trouve une mesure invariante de densité  $1/\alpha\beta$ .

Ce calcul se généralise sans peine à l'algorithme en dimension  $n$ , qui consiste à retirer le plus petit nombre du plus grand; comme ci-dessus, on montre facilement que la région  $\alpha_1 < \alpha_2 < \dots < \alpha_n < \alpha_1 + \alpha_2$  est invariante, et que l'algorithme projectif obtenu en posant  $\alpha_n = 1$  admet une mesure invariante de densité  $1/\alpha_1 \alpha_2 \dots \alpha_n$ ,

## 7. Ergodicité de la mesure invariante

Il est facile de montrer que la mesure invariante trouvée est ergodique lorsqu'elle est finie, en utilisant le fait que le flot  $\Phi_t$  est un flot du type d'Anosov, dont le feuilletage stable est donné par les bases  $(\alpha_1, \dots, \alpha_n) = \text{Cte}$ , et le feuilletage instable par les hauteurs  $(\lambda_1, \dots, \lambda_n) = \text{Cte}$ .

Les grandes lignes du raisonnement sont les suivantes : tout d'abord, la donnée des bases détermine l'application linéaire qui définit  $\tilde{F}$ , et les applications linéaires qui définissent ces puissances ; donc, bien que cette application ne soit pas continue, elle est continue et linéaire sur toute feuille du feuilletage stable. De plus, pour tous les algorithmes du type Brun, la matrice qui agit sur les hauteurs est à coefficients strictement positifs, et dès que  $\alpha_1$  est strictement positif, cette matrice est primitive à partir d'un certain rang.

Enfin, par application du théorème de Poincaré, puisque la mesure invariante est finie, et du fait que l'application linéaire associée à  $\tilde{F}$  est constante au voisinage d'un point, on va voir apparaître une infinité de fois la même matrice primitive ; donc, par application du théorème de Perron-Frobenius, pour des bases fixées, toutes les hauteurs vont tendre vers le même vecteur de hauteurs.

Il suffit alors de remarquer que, si l'on prend une fonction continue à support compact, les intégrales de Birkhoff le long du flot,  $1/T \int_0^T f(\Phi_t x) dt$ , qui tendent presque partout vers une limite, ont même limite le long d'une feuille du feuilletage stable ; de même, les intégrales de Birkhoff négatives,  $1/T \int_{-T}^0 f(\Phi_t x) dt$  ont même limite le long des feuilles du feuilletage instable. On en déduit que, localement, les intégrales de Birkhoff tendent presque partout vers la même limite.

Il suffit pour établir l'ergodicité de montrer que l'on peut joindre deux points quelconques par un chemin contenu dans des feuilles des deux feuilletages ; mais il est facile de vérifier dans tous les cas que l'on a étudiés que l'on peut toujours joindre un point de coordonnées  $(\alpha_1, \dots, \alpha_n, \lambda_1, \dots, \lambda_n)$  à un voisinage arbitraire du point  $(0, \dots, 0, \alpha_n, 0, \dots, 0, 1/\alpha_n)$ . Il suffit en laissant les bases fixes de faire diminuer les  $n-1$  premières hauteurs arbitrairement près de 0, puis en laissant les hauteurs fixes de diminuer les  $n-1$  premières bases arbitrairement près de 0 ; le point obtenu peut être amené près du point  $(0, \dots, 0, 1, 0, \dots, 0, 1)$  par le flot  $\Phi_t$ , ce qui montre la connexité cherchée.

*Remarque.* — Il ne suffit pas de dire que le flot consiste en une multiplication par  $e^t$  sur les bases et  $e^{-t}$  sur les hauteurs pour établir l'ergodicité ; en effet, nous considérons le flot sur  $\Omega = D/\tilde{F}$ , et comme  $\tilde{F}$  n'est pas a priori une isométrie de  $D$ , il n'y a pas de notion intrinsèque de distance sur  $\Omega$ , et l'on ne peut pas dire que  $\tilde{F}$  diminue les distances de  $e^t$  sur les feuilles stables ; ce n'est que dans le cas des fractions continues que nous avons une distance bien définie, à cause de la traduction géométrique.

### 8. Quelques questions ouvertes

La première question est bien sûr d'essayer d'appliquer cette méthode à d'autres algorithmes ; mais il n'est pas du tout évident de trouver un domaine  $D$  sur lequel  $\tilde{F}$  soit bijective, on a vu que même l'algorithme de Selmer pose des problèmes. Quant à l'algorithme classique de Jacobi-Perron, il se heurte à la difficulté supplémentaire qu'il faut, pour pouvoir revenir en arrière, retrouver non pas 1 mais 2 nombres entiers, et on ne voit pas comment extraire cette information des hauteurs.

On pourrait progresser dans cette matière si l'on arrivait, comme on le fait dans le cas  $n=2$ , à représenter géométriquement le flot ; il serait en particulier intéressant de trouver une représentation matricielle faisant apparaître une action de  $SL(n, \mathbb{Z})$ , puisque ce groupe apparaît forcément dans les généralisations classiques des fractions continues ; de plus, une telle représentation permettrait de définir des flots horosphériques commutant avec  $\tilde{F}$ , ce qui donnerait une preuve plus agréable de l'ergodicité. Mais la façon d'obtenir une matrice  $3 \times 3$  avec ces 6 nombres n'est pas claire ; peut-être faudrait-il rajouter des largeurs ?

Dans le cas de l'induction des échanges d'intervalles, qui est une autre forme de fraction continue généralisée, le raisonnement est guidé par l'interprétation dynamique du procédé. La méthode que nous proposons a aussi une interprétation dynamique, puisque l'on peut la considérer comme une construction par approximations successives (« tours ») d'une application sur les intervalles de base ; à la différence du cas considéré par Veech, l'application n'est pas donnée au départ, mais construite au fur et à mesure. Cette transformation possède-t-elle des propriétés particulières ? Son identification pourrait permettre de mieux comprendre les divers algorithmes.

### BIBLIOGRAPHIE

- [AF1] R. ADLER et L. FLATTO, *The Backward Continued Fraction Map and Geodesic Flow* (Ergod. Th. & Dynam. Sys., vol. 4, 1984, p. 487-492).
- [AF2] R. ADLER et L. FLATTO, *Geodesic Flows, Interval Maps and Symbolic Dynamics* (Bull. Amer. Math. Soc., vol. 25, 1991, p. 229-334).
- [Ar] G. ARSAC, *L'origine de la démonstration : essai d'épistémologie didactique*, *Recherches en didactique des mathématiques*, vol. 8, 1987, p. 267-312.
- [BKS] *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces*, edited by T. BEDFORD, M. KEANE and C. Series, Oxford Science Publications, Oxford, 1991.
- [Bi] P. BILLINGSLEY, *Ergodic Theory and Information*, John Wiley & Sons, New York, 1965.
- [Br] A. J. BRENTJES, *Multi-Dimensional Continued Fractions Algorithms*, Mathematical Centre Tracts 145, Amsterdam, 1981.
- [Fa] C. FAIVRE, *Distribution des constantes de Lévy des nombres quadratiques*, Thèse, Université de Provence, 1991.
- [INT] S. ITO, H. NAKADA et S. TANAKA, *On the Invariant Measure for the Transformations Associated with some Real Continued Fractions*, *Keio Engineering Reports*, vol. 30, 1977, p. 159-175.
- [Ja] H. JAGER, *The Distribution of Certain Sequences connected with the Continued Fraction* (Indag. Math., vol. 48, 1986, p. 61-69).
- [Kh] A. YA. KHINCHIN, *Continued Fractions*, The University of Chicago Press, Chicago, 1964.

- [Le] P. LEVY, *Sur le développement en fraction continue d'un nombre choisi au hasard*, *Compositio Mathematica*, vol. 3, 1936, p. 286-303.
- [Mo] B. MORIN, *Structure géométrique d'un ensemble de surfaces: l'ensemble des métriques à courbure nulle du tore*, dans *Encyclopédie philosophique*, Presses Universitaires de France, Paris.
- [Na] H. NAKADA, *Metrical Theory for a Class of Continued Fraction Transformations and Their Natural Extensions* (*Tokyo J. Math.*, vol. 4, 1981, p. 399-426).
- [Po] M. POLLICOTT, *Distribution of Closed Geodesics on the Modular Surface and Quadratic Irrationals* (*Bull. Soc. Math. Fr.*, vol. 114, 1986, p. 431-446).
- [Sc1] F. SCHWEIGER, *The Metrical Theory of Jacobi-Perron Algorithm* (*Lecture Notes Math.*, vol. 334, Springer Verlag, Berlin, 1973).
- [Sc2] F. SCHWEIGER, *Ergodische Eigenschaften der Algorithmen von Brun und Selmer* (*S. Ber. Österr. Akad. Wiss. Math.-naturw. Kl, Abt. II*, vol. 191, 1982, p. 325-329).
- [Sc3] F. SCHWEIGER, *Invariant Measures for Maps of Continued Fraction Type* [*J. Number Theory*, 1990, (à paraître)].
- [Sc4] F. SCHWEIGER, *Ergodic Properties of Fibered Systems*, Prépublication de l'université de Salzburg, 1989.
- [Se] C. SERIES, *Geometrical Markov Coding of Geodesics on Surfaces of Constant Negative Curvature*, (*Ergod. Th. & Dynam. Sys.*, vol. 6, 1986, p. 601-625).
- [Ve1] W. A. VEECH, *Gauss Measures for Transformations on the space of interval Exchange Maps* (*Ann. of Math.*, vol. 115, 1982, p. 201-242).
- [Ve2] W. A. VEECH, *The Teichmüller Geodesic flow* (*Ann. of Math.*, vol. 124, 1986, p. 441-530).

(Manuscrit reçu le 31 mars 1992.)

Pierre ARNOUX,  
Laboratoire de Mathématiques Discrètes, UPR 9016,  
163 avenue de Luminy,  
case 930, 13288 Marseille Cedex, France  
ARNOUX@LMD.UNIV-MRS.FR

Arnaldo NOGUEIRA,  
Instituto de Matemática,  
Universidade Federal de Rio de Janeiro,  
Caixa Postal 68530,  
21944 Rio de Janeiro, RJ Brasil.