

Trous dans \sqrt{n} modulo 1 et théorie ergodique

Louis Nebout

12 décembre 2011

Introduction à un domaine de recherche sous l'encadrement d'Yves Benoist.

Résumé

Le but de ce texte est de proposer une brève introduction à la théorie ergodique dans le cadre des actions de groupes de Lie. Pour cela, je commence par poser les définitions et les enjeux principaux de la théorie ergodique, avant d'énoncer un résultat plus spécifique aux groupes de Lie : le puissant théorème prouvé par Ratner en 1991, classifiant les mesures ergodiques pour les flots unipotents. Le reste de l'article est consacré à un exemple récent d'application concrète de ce théorème : le calcul de la répartition asymptotique des trous de la suite \sqrt{n} modulo 1, un résultat prouvé par Elkies et McMullen en 2004.

Table des matières

1	Théorie ergodique	2
1.1	Définition	2
1.2	Théorème de Birkhoff	2
1.3	Théorème de Ratner	3
2	Trous et réseaux affines	4
2.1	Énoncé du problème	4
2.2	Trous et triangles	5
2.3	Interprétation dynamique	6
3	Application du théorème de Ratner	7
3.1	Étude de la mesure limite	7
3.2	Calcul explicite	9

1 Théorie ergodique

1.1 Définition

Dans toute la suite, X désignera un ensemble, μ une mesure de probabilité sur cet ensemble et $f : X \rightarrow X$ une transformation μ -invariante.

Définition . On dit que f est ergodique (pour μ) si toute partie mesurable et f -invariante A (c'est-à-dire telle que $f^{-1}(A) = A$) est de mesure 0 ou 1.

Remarque 1.1. Quand la transformation f est sous-entendue, on dira que μ est ergodique. Cette définition s'étend naturellement au cas où un groupe G agit sur X de façon mesurable et μ -invariante.

Intuitivement, dire que f est ergodique veut dire que ses orbites balayent tout l'espace, une intuition illustrée par la proposition suivante :

Proposition 1.2. Soit X soit un espace métrique localement compact et séparable, μ une mesure de probabilité borélienne sur X , et f soit un homéomorphisme μ -invariant de X . Alors, si f est ergodique, pour μ -presque tout x , l'orbite $\{f^n(x), n \in \mathbb{Z}\}$ est dense dans le support de μ .

Démonstration. On peut supposer que X est le support de μ . Soit D une partie dénombrable dense de X , et \mathcal{D} l'ensemble des boules centrées en D de rayons rationnels. Alors pour toute boule B de \mathcal{D} , l'ensemble $A_B := \cup_{n \in \mathbb{Z}} f^n(B)$ des points dont l'orbite passe dans B est f -invariant, donc de mesure 1 par ergodicité. L'intersection dénombrable A des ensembles A_B est aussi de mesure 1, et l'orbite d'un point de A est dense dans X . \square

La propriété suivante, très importante, permet de relier mesures ergodiques et mesures invariantes :

Proposition 1.3. Supposons X compact et métrisable. Les mesures de probabilité f -ergodiques sur X sont exactement les points extrémaux de l'ensemble convexe des mesures de probabilité f -invariantes sur X . Toute mesure f -invariante possède une écriture comme combinaison convexe de mesures ergodiques de la forme suivante :

$$\mu = \int \nu \, dP(\nu),$$

pour P une certaine mesure de probabilité sur l'ensemble des mesures de probabilités ergodiques sur X .

Pour plus de détails sur cette décomposition, voir [HK], page 84.

1.2 Théorème de Birkhoff

Le résultat central de la théorie ergodique est le théorème de Birkhoff. Pour une preuve, voir par exemple [Be].

Théorème 1.4. *Pour tout ϕ dans $L^1(X, \mu)$, on note $\phi_n(x) = \sum_{0 \leq i \leq n} \phi(f^i(x))$. Alors la limite $c := \lim_{n \rightarrow \infty} \frac{1}{n} \phi_n(x)$ existe pour μ -presque tout x . De plus, si μ est ergodique, alors presque partout on a*

$$\phi_n(x) = \int_X \phi \, d\mu.$$

Ce théorème (qui s'étend aussi au cas d'une action continue et non discrète) veut dire que, dans le cas ergodique, les moyennes orbitales convergent presque sûrement vers la moyenne spatiale. Ainsi, la proportion de temps passée par une orbite dans une partie mesurable A converge presque sûrement vers la mesure de A . Ce théorème, à l'importance théorique centrale, a un gros défaut dans les cas pratiques : il ne donne que des convergences presque sûres, et ne permet donc rien de dire sur des orbites données. Il y a tout de même un cas où l'on sait facilement résoudre le problème : quand la mesure ergodique est unique (et donc quand la mesure invariante est unique).

Proposition 1.5. *Supposons que X admette une seule mesure f -invariante. Alors la convergence du théorème de Birkhoff a lieu pour tout x .*

Démonstration. Soit x dans X , considérons la suite $\mu_n := \frac{1}{n} \sum_{0 \leq i < n} \delta_{f^i(x)}$. C'est une suite de mesures de probabilités, qui admet une valeur d'adhérence (pour la topologie faible) par compacité. Une telle valeur d'adhérence étant f -invariante, c'est nécessairement μ . \square

On voit dans cet exemple que la connaissance précise de l'ensemble des mesures ergodiques sur X permet d'affiner les résultats donnés par le théorème de Birkhoff, ce qui pousse à essayer de classifier ces mesures ergodiques.

1.3 Théorème de Ratner

Le théorème de Ratner, prouvé dans [RA], permet justement de classifier les mesures ergodiques sur certains espaces homogènes :

Théorème 1.6. *Soit $\Gamma \subset G$ un sous-groupe discret d'un groupe de Lie connexe G , et soit N un sous-groupe unipotent de G (c'est-à-dire, tel que les adjoints des éléments de N aient pour seule valeur propre 1). Soit μ une mesure de probabilité N -invariante, N -ergodique sur G/Γ , et soit $J \subset G$ le groupe des éléments laissant μ invariante. Alors il existe x dans G/Γ tel que le support de μ soit l'orbite $J \cdot x$.*

Comme μ est J -invariante, une telle mesure est bien entièrement déterminée par son support $J \cdot x$ (moralement, c'est la mesure de Haar sur J). On a ainsi ramené le problème de classification des mesures ergodiques à un problème algébrique : trouver tous les sous-groupes J vérifiant $N \subset J \subset G$, problème qui se traite bien dans de nombreux cas concrets.

2 Trous et réseaux affines

Toute la suite de cet article sera consacrée à l'étude d'un tel cas concret : l'étude de la répartition asymptotique des trous de la suite \sqrt{n} modulo 1. En effet, cette répartition peut être interprétée en termes de théorie ergodique sur l'espace des réseaux du plan affine. Je donne des éléments de preuves relativement détaillés, car les méthodes utilisées pour faire apparaître un problème de théorie ergodique, puis les techniques utilisées pour résoudre ce problème, sont très représentatives, et peuvent être utilisées pour traiter bon nombre de problèmes similaires. Cette partie est consacrée à la mise en évidence de cette interprétation géométrique.

2.1 Énoncé du problème

Dans toute la suite, je note $\{x\}$ la partie fractionnaire du réel x . Il n'est pas difficile de montrer que la suite $(\{\sqrt{n}\})$ est équirépartie, c'est-à-dire que pour tout $a \leq b$ dans $[0, 1]$, la quantité

$$\frac{|\{0 \leq n \leq N-1 : \{\sqrt{n}\} \in [a, b]\}|}{N}$$

converge vers $b - a$ quand N tend vers l'infini. En ce sens, la suite $(\{\sqrt{n}\})$ se comporte comme une suite de nombres tirés au hasard.

Si l'on veut aller plus loin dans cette comparaison, une façon de faire est d'étudier les trous de la suite. L'ensemble des $(\{\sqrt{n}\})$, pour n compris entre 0 et $N-1$, divise \mathbb{R}/\mathbb{Z} en N intervalles, je note $T(N)$ l'ensemble de ces intervalles. La longueur moyenne de ces intervalles est de $\frac{1}{N}$, ce qui amène à étudier la limite de la quantité suivante :

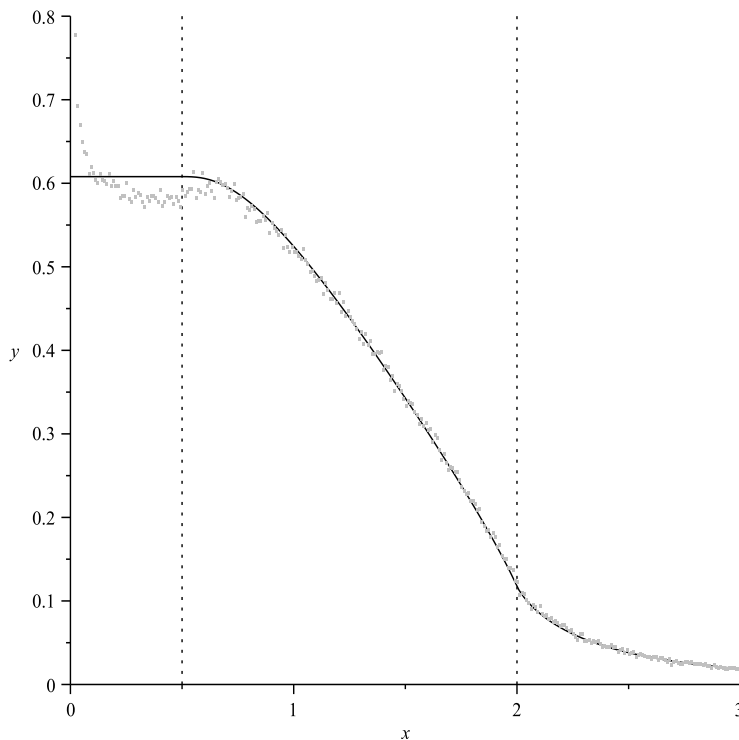
$$\frac{|\{T \in T(N) : |T| \in [\frac{a}{N}, \frac{b}{N}]\}|}{N}, \quad (2.1)$$

pour tout $a \leq b$ des réels positifs.

Le cas d'une suite de nombres aléatoires est facile à traiter : soit a dans $[0, 1]$, si l'on tire N nombres uniformément dans $[0, 1]$, la probabilité qu'un trou donné soit plus petit que a/N est de $1 - (1 - (a/N))^N$, et donc l'espérance du nombre de trous plus petits que $\frac{a}{N}$ est de $N(1 - (1 - (a/N))^N)$. La quantité (2.1) converge donc en espérance, quand N tend vers l'infini, vers $\int_a^b e^{-t} dt$, et la répartition asymptotique des trous suit une loi de Poisson.

Les tests numériques suggèrent que la suite $(\{n^\alpha\})$, pour α dans $]0, 1[$ différent de $\frac{1}{2}$, suit le même comportement : la répartition asymptotique des trous est décrite par une loi de Poisson. Par contre, la suite $(\{\sqrt{n}\})$ a un comportement radicalement différent. La répartition asymptotique des trous possède une densité $F(t)$, c'est-à-dire que la quantité précédente converge vers $\int_a^b F(t) dt$, où F est une fonction très différente d'une exponentielle. La figure suivante montre la courbe théorique F . Le nuage de points correspond quant à lui aux résultats expérimentaux obtenus en considérant les 10^6 premiers points de la suite $(\{n^\alpha\})$. La fonction F possède plusieurs propriétés surprenantes : elle est

constante, égale à $\frac{6}{\pi^2}$, sur $[0, \frac{1}{2}]$. Elle est lisse sauf aux points $\frac{1}{2}$ où elle n'est pas \mathcal{C}^3 , et 2 où elle n'est pas \mathcal{C}^2 . Enfin, quand t tend vers l'infini $F(t)$ est équivalent à $3/\pi^2 t^3$, ce qui veut dire qu'il y a comparativement beaucoup de gros trous par rapport à une répartition obtenue en tirant des nombres au hasard uniformément dans $[0, 1[$.



2.2 Trous et triangles

Soient $x \leq y$ dans $[0, 1]$. Il y a une donnée auquel nous avons facilement accès : à quelle condition sur x et y une valeur de $(\{\sqrt{n}\})$ tombe-elle dans $[x, y]$. Plus précisément, soit t dans $[0, 1[$, et δ un réel strictement positif. Pour simplifier, nous étudierons les N premières valeurs de la suite $(\{\sqrt{n}\})$, pour N un carré parfait : $N = s^2$, avec s un entier. On commence par étudier à quelle condition sur δ existe-t-il un entier n vérifiant $0 \leq n < N$ tel que $(\{\sqrt{n}\})$ appartienne à $[t, t + \frac{\delta}{N}]$. Le plus petit δ vérifiant cette condition nous donnera la longueur du trou (renormalisée) à droite de t . Puis, en faisant le même travail à gauche de t , nous obtiendrons la longueur du trou contenant t .

On a les équivalences

$$\begin{aligned} \{\sqrt{n}\} \in \left[t, t + \frac{\delta}{N} \right] &\iff \exists a \in \mathbb{N}, 0 \leq a < s, \sqrt{n} \in \left[a + t, a + t + \frac{\delta}{N} \right] \\ &\iff \exists a \in \mathbb{N}, 0 \leq a < s, n - a^2 \in \\ &\quad \left[2ta + t^2, 2ta + t^2 + 2(a + t)\frac{\delta}{N} + \frac{\delta^2}{N^2} \right]. \end{aligned}$$

Ici, a est le plus grand parfait carré inférieur à n , et $n - a^2$ est donc un entier, compris entre 0 et $2a$. Ainsi, nous avons obtenu une condition de nature géométrique sur $\delta : \{\sqrt{n}\}$ est dans $[t, t + \frac{\delta}{N}[$ si et seulement si la région suivante du plan \mathbb{R}^2 contient un point à coordonnées entières :

$$\{(x, y) \in \mathbb{R}^2, 0 \leq x < s, 0 \leq y \leq 2x : 2tx + t^2 \leq y < 2tx + t^2 + 2(x+t)\frac{\delta}{N} + \frac{\delta^2}{N^2}\}.$$

Des considérations élémentaires permettent de simplifier cette expression, ce qui nous ramène à l'étude des points entiers de l'ensemble suivant :

$$\{(x, y) \in \mathbb{R}^2, -t - \frac{\delta}{2N} < x < s - t : 2tx + t^2 \leq y < 2tx + t^2 + 2(x+t)\frac{\delta}{N} + \frac{\delta^2}{N^2}\}.$$

Cet ensemble est un triangle du plan. Malheureusement, on remarque l'apparition de termes d'erreurs, qui vont s'avérer pénibles dans la suite des calculs. Il y a un terme en $O(\frac{1}{N})$ dans les coordonnées de la pointe du triangle, et dans la hauteur $2\frac{s\delta}{N} + \frac{\delta^2}{N^2}$ du côté vertical du triangle, il y a un terme en $O(\frac{1}{N^2})$. Pour faciliter la suite de l'étude, nous allons tout de suite faire une approximation sur la taille des trous, approximation qui aura pour but de supprimer les termes d'erreurs pour aboutir à un triangle plus agréable, dépendant moins de N . Nous approximerons la taille du trou (renormalisé) contenant t par la quantité $L'_N(t)$ définie comme suit :

Définition . Supposons s strictement plus grand que 1, et soit t dans $[0, 1[$. On pose $N = s^2$. Alors $L'_N(t)$ est égal à l'aire $c_+ - c_-$ du plus grand triangle de la forme

$$\{(x, y) \in \mathbb{R}^2, -t \leq x < s - t : 2tx + t^2 + 2(x+t)\frac{c_-}{N} < y < 2tx + t^2 + 2(x+t)\frac{c_+}{N}\},$$

avec $c_- \leq 0 \leq c_+$ des réels, ne contenant pas de point à coordonnées entières. Si aucun tel triangle n'existe, alors $L'_N(t) = 0$.

Les estimations techniques justifiant le fait que cette approximation permet tout de même le calcul de la loi asymptotique de répartition des trous sont résumées dans le lemme suivant. Pour plus de détail, voir l'article originel [EM].

Lemme 2.1. *Supposons qu'il existe une fonction F positive de $[0, \infty[$ dans \mathbb{R} , continue et d'intégrale 1, telle que pour tout réel positif x , on ait la convergence*

$$\lim_{s \rightarrow \infty} \lambda(\{t \in [0, 1[, L'_{s^2}(t) < x\}) = \int_0^x tF(t) dt,$$

où λ désigne la mesure de Lebesgue, et que cette convergence soit uniforme pour x dans tout compact de $[0, \infty[$. Alors F est la densité de la répartition asymptotique des trous.

2.3 Interprétation dynamique

En appliquant des transformations affines du plan, on peut reformuler la définition de $L'_N(t)$: cette quantité est égale à l'aire du plus grand triangle de la forme

$$\Delta_{c_-, c_+} = \{(x, y) \in \mathbb{R}^2 : 0 < y < 1, 2c_-y < x < 2c_+y\}, \quad (2.2)$$

avec $c_- \leq 0 \leq c_+$, ne contenant pas de point du réseau affine $g(t, N)(\mathbb{Z}^2)$, où l'on a posé

$$g(t, N) := \begin{pmatrix} s & -2st & -st^2 \\ 0 & \frac{1}{s} & \frac{t}{s} \\ 0 & 0 & 1 \end{pmatrix}.$$

On remarque que

$$g(t, N) = a_s \sigma(t), \text{ où } a_s := \begin{pmatrix} s & 0 & 0 \\ 0 & \frac{1}{s} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ et } \sigma(t) := \begin{pmatrix} 1 & -2t & -t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}.$$

Notons E l'ensemble $ASL_2(\mathbb{R})/ASL_2(\mathbb{Z})$ des réseaux affines du plan de covolume 1, et L l'application partie linéaire, le projetant sur l'ensemble $B := SL_2(\mathbb{R})/SL_2(\mathbb{Z})$ des réseaux linéaires du plan de covolume 1. Notons f l'application qui à un réseau affine associe l'aire du plus petit triangle de la forme donnée par (2.2) ne contenant pas de point du réseau. Ainsi, d'après le lemme 2.1, la quantité que nous devons évaluer est la limite suivante :

$$\lim_{s \rightarrow \infty} \lambda(\{t \in [0, 1[, f(a_s \sigma(t)\mathbb{Z}^2) < x\}).$$

C'est une propriété de type dynamique : il nous faut étudier la façon dont la courbe $\{\sigma(t)\mathbb{Z}^2, t \in [0, 1]\}$ se comporte dans E quand on la pousse par a_s . Quand on regarde les parties linéaires, a_s agit comme le flot géodésique de B , qui possède de bonnes propriétés dynamiques, ce qui permet de penser que nos courbes vont elles aussi posséder de bonnes propriétés.

3 Application du théorème de Ratner

On munit E et B de leurs mesures de Haar à gauche, notées μ_E et μ_B . La courbe fermée de E , $\{a_s \sigma(t)\mathbb{Z}^2, t \in [0, 1]\}$, porte naturellement une mesure, image de la mesure uniforme sur $[0, 1[$. On note cette mesure μ_s . On passe d'une de ces mesures μ_s à une autre en poussant par $a_{s'}$, pour un s' bien choisi. Quand nous parlons de répartition de la courbe $\{a_s \sigma(t), t \in [0, 1]\}$, ce que nous voulons étudier, c'est la limite éventuelle de la suite de mesures de probabilité μ_s . Plus précisément, nous voudrions arriver au résultat suivant :

Théorème 3.1. *La suite de mesures de probabilité (μ_s) converge faiblement vers μ_E .*

3.1 Étude de la mesure limite

Je donne dans cette partie une ébauche des étapes importantes de la preuve, pour plus de détails se reporter à [EM].

Tout d'abord, par compacité, μ_s admet une valeur d'adhérence, que nous appellerons μ . Notre objectif est donc de montrer que cette valeur d'adhérence est nécessairement égale à μ_B . On a facilement accès à la partie linéaire de μ (c'est-à-dire à sa projection sur B par L). La connaissance de cette partie linéaire prouve de plus le fait important que la mesure limite μ est une mesure de probabilité.

Proposition 3.2. *Nous avons $L_*\mu = \mu_B$.*

Démonstration. On sait que le flot géodésique sur B est mélangeant, c'est à dire que pour tout sous-ensemble $X \subset B$ de mesure strictement positive, et toute fonction f dans $C_0(B)$, nous avons

$$\lim_{s \rightarrow \infty} \frac{1}{\mu_B(X)} \int_X f(a_s x) d\mu_B(x) = \int_E f d\mu_B$$

(c'est une version plus forte de l'ergodicité, et cela se prouve comme l'ergodicité à l'aide du théorème de décroissance des coefficients de Howe-Moore, voir par exemple [Be]). L'idée est de construire un petit cylindre de mesure strictement positive autour de notre courbe, de telle sorte que a_s ne déforme pas trop le cylindre, puis d'appliquer la propriété de mélange. \square

Définition . Soit t un réel, notons

$$n_t := \begin{pmatrix} 1 & t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Les matrices de la forme n_t forment un sous-groupe de $ASL_2(\mathbb{R})$, que nous noterons N .

Intuitivement, l'action de a_s étire de plus en plus l'horocycle σ dans la direction de N , et donc pousser la mesure $m(\sigma_s)$ par n_t , pour un t fixé, va peu modifier cette mesure pour s grand. Cette idée permet de démontrer facilement que μ est nécessairement N -invariante à gauche.

On ne peut pas appliquer directement le théorème Ratner à μ , qui n'a aucune raison d'être N -ergodique. Par contre, ce que l'on peut faire, c'est décomposer μ en somme de mesures ergodiques à l'aide de la proposition 1.3, et appliquer le théorème à ces mesures. On écrit donc

$$\mu = \int \nu dP(\nu), \tag{3.1}$$

et le théorème de Ratner nous dit, en notant $J(\nu)$ l'ensemble des éléments de $ASL_2(\mathbb{R})$ laissant ν invariante, que chaque mesure ν est portée par une $J(\nu)$ -orbite. En appliquant L à cette décomposition, on remarque de plus que pour P -presque tout ν , $L_*\nu = \mu_B$, et donc $L(J(\nu)) = SL_2(\mathbb{R})$.

On s'est donc ramené à un problème purement algébrique, trouver les sous-groupes de $ASL_2(\mathbb{R})$ contenant N et de partie linéaires $SL_2(\mathbb{R})$ tout entier. Le résultat des calculs algébriques est qu'un tel groupe est soit $ASL_2(\mathbb{R})$ tout entier (dans quel cas ν est la mesure de Haar μ_E espérée), soit un conjugué par une translation horizontale de $SL_2(\mathbb{R})$. Pour conclure il ne nous reste plus qu'à éliminer ce cas-là, c'est-à-dire à prouver que ces mesures dont le groupe des invariants est conjugué à $SL_2(\mathbb{R})$ ont un poids nul dans la décomposition de μ .

Cela se traite à la main, l'idée étant que ces mesures particulières ont une interprétation précise : l'application L de E dans B est une fibration à fibres toriques, et ces mesures sont précisément celles donnant un poids non nul à l'ensemble des translatés horizontaux des éléments de torsion de ces tores. Les coefficients de la matrice $\sigma(t)$ ne sont pas linéaires en t , ce qui permet de montrer que la mesure μ ne peut charger ces ensembles de points de torsion.

3.2 Calcul explicite

Attention, le travail n'est pas encore terminé : on a donc démontré que, asymptotiquement, le nombre de trous renormalisés de taille inférieure à x était égal à la mesure de l'ensemble des réseaux affines Λ vérifiant $f(\Lambda) < x$, à condition que cette expression (que l'on note $\tilde{f}(x)$) vérifie les bonnes conditions de régularité nécessaires pour que le lemme 2.1, justifiant l'approximation utilisée, fonctionne. Il faut donc forcément calculer cette quantité plus en détail pour clore la preuve. Notons p la fonction qui à un réel x positif associe la mesure de l'ensemble des réseaux affines Λ possédant un point dans un triangle d'aire x fixé à l'avance. Les calculs montrent que cette fonction p est de classe \mathcal{C}^2 , est reliée à notre expression précédente \tilde{f} par l'équation $\tilde{f}(x) = p(x) - xp'(x)$, et que les hypothèses du lemme 2.1 sont vérifiées en prenant $F = -p''$. Ainsi :

Théorème 3.3. *La densité de la répartition asymptotique des trous est $-p''$.*

Ces calculs permettent également de donner une expression explicite de cette densité, que je donne à titre indicatif :

Théorème 3.4. *Appelons F la densité de la loi asymptotique de répartition des trous. Alors,*

i) Si $x \leq 1/2$, alors $F(x) = 6/\pi^2$

ii) Si $1/2 \leq x \leq 2$, alors posons $r := 1/2x$, et

$$\psi(r) = \arctan\left(\frac{2r-1}{\sqrt{4r-1}}\right) - \arctan\left(\frac{1}{\sqrt{4r-1}}\right).$$

Alors,

$$F(x) = \frac{6}{\pi^2} \left(\frac{2}{3}(4r-1)^{3/2}\psi(r) + (1-6r)\log r + 2r - 1 \right).$$

iii) Si $x \geq 2$, alors posons $r := 1/2x$, et $\alpha := (1 - \sqrt{1-4r})/2$. Alors,

$$F(x) = \frac{6}{\pi^2} (4(1-4\alpha)(1-\alpha)^2 \log(1-\alpha) - 2(1-2\alpha)^3 \log(1-2\alpha) - 2\alpha^2).$$

Références

- [Be] Y. Benoist, *Réseaux des groupes de Lie*, <http://www.math.u-psud.fr/~benoist/prepubli/survey.html>, 2008.
- [EM] N. D. Elkies et C. T. McMullen, *Gaps in \sqrt{n} mod 1 and ergodic theory*, *Duke Math. J.* **123**(2004), 95-139.
- [HK] Édité par B. Hasselblatt et A. B. Katok, *Handbook of dynamical systems, Volume 1A*, North-Holland, Amsterdam, 2002.
- [RA] M. Ratner, *On Raghunathan's measure conjecture*, *Annals of Math.* **134**(1991), 545-607.