Sur la constante de Khinchin

Stéphane Flotron, Marc Hoyois, Ludovic Pirl

15 juin 2006

Table des matières

Introduction

En 1935, le mathématicien russe Aleksandr Yakovlevich Khinchin publiait dans le premier volume de *Compositio Mathematica* un résultat surprenant de théorie des nombres.

Une *fraction continue simple* est une expression finie ou infinie de la forme

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots}}$$

où a_0 est un entier quelconque et les a_i des entiers strictement positifs pour $i \geq 1$. Les entiers a_i sont les *éléments* de la fraction continue. Avec quelques conventions supplémentaires, tout nombre réel peut s'écrire de façon unique sous la forme d'une fraction continue simple. Réciproquement, toute fraction continue simple converge vers un nombre réel et le nombre ainsi représenté est noté $[a_0; a_1, a_2, \dots]$. Le résultat de Khinchin est le suivant.

THÉORÈME (Khinchin, 1935). *Il existe un nombre réel K_0 tel que, pour presque tout nombre réel $x = [a_0; a_1, a_2, \dots]$, la suite des moyennes géométriques partielles*

$$\left(\sqrt[n]{a_1 a_2 \dots a_n}\right)_{n=1}^{\infty}$$

des éléments de la fraction continue simple de x converge vers K_0 lorsque n tend vers l'infini.

Ici, dire que la propriété est vraie pour *presque tout* nombre réel signifie qu'elle est vraie sur le complémentaire d'un ensemble de mesure nulle (au sens de Lebesgue). Le nombre K_0 s'appelle la *constante de Khinchin* et vaut environ 2.685. On ne connaît que très peu de choses sur la nature de ce nombre, notamment on ignore s'il est irrationnel. Malgré cela, cette propriété n'a pas pu à ce jour être établie pour un nombre réel donné (bien que certains nombres, tels π ou la constante de Khinchin elle-même, semblent la vérifier). Parmi les nombres qui, au contraire, n'ont pas cette propriété sont les nombre rationnels, les racines de polynômes quadratiques à coefficients rationnels (par exemple le *nombre d'or* dont le développement en fraction continue simple est $[1; 1, 1, \dots]$), la base du logarithme naturel e et plus généralement tous les nombres dont nous connaissons le développement exact en fraction continue.

Le théorème de Khinchin nous révèle un aspect étonnant de la répartition des fractions continues sur la droite réelle. En effet, l'écriture d'un nombre en fraction continue simple induit une bijection entre ces fractions continues (qui peuvent être vues comme des suites de nombres naturels avec certaines propriétés) et les nombres réels. Seule une partie « infime » des fractions continues ont la propriété de Khinchin, mais ces fractions sont réparties sur la droite réelle de telle sorte qu'elles la recouvrent presque intégralement. Cette simple constatation montre de fait que l'ensemble des nombres réels ne satisfaisant pas le théorème, quoique de mesure nulle, a la puissance du continu.

Dans cet article, nous commencerons par étudier quelques propriétés élémentaires des fractions continues, puis nous donnerons deux démonstrations du théorème de Khinchin. La première est la preuve originale de Khinchin, pour laquelle nous développerons quelques aspects de la *théorie de la mesure des fractions continues.* La seconde s'inscrit dans le cadre plus général de la *théorie ergodique*, et en particulier le théorème de Khinchin apparaîtra comme une application aux fractions continues du très célèbre *théorème ergodique.*

1 Notions sur les fractions continues

1.1 Définitions

Une *fraction continue réelle finie* (de longueur n) est une expression de la forme

$$a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{\ddots + \cfrac{b_n}{a_n}}} \tag{1}$$

où n est un entier naturel et les a_i et les b_i sont des nombres réels. Une *fraction continue réelle infinie* est une expression formelle

$$a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2 + \cdots}} \tag{2}$$

où les a_i et les b_i sont des nombres réels. On dira aussi qu'une fraction continue infinie est de *longueur infinie* et que sa longueur est supérieure à celle d'une fraction continue finie.

On utilise le terme *fraction continue réelle* pour désigner une fraction continue réelle finie ou infinie. Pour $i \geq 1$, les nombres a_i et b_i s'appellent les *éléments* de la fraction continue. Le nombre a_0 est le *terme initial* de la fraction continue.

Les fractions continues (1) et (2) sont dites *régulières* si tous les b_i sont égaux à 1. Par la suite, nous considérerons uniquement les fractions continues réelles régulières, et par conséquent nous appellerons *éléments* les nombres a_1, a_2, ... exclusivement.

Remarquons que, pour le moment, nous ne nous préoccuppons pas des questions de convergence, et par conséquent nous ne pouvons, *a priori*, assigner une valeur à une fraction continue, qu'elle soit finie ou infinie.

Lorsqu'elles sont régulières, nous utilisons la notation $[a_0; a_1, a_2, \dots]$ pour les fractions continues (1) et (2). On se permettra en outre d'utiliser un symbole pour représenter une telle fraction, même si elle ne constitue, pour l'instant, qu'une expression formelle. Nous notons également $[a_0; a_1, \dots, a_n]$ pour préciser que la fraction continue considérée est finie ; $[a_1, \dots, a_n]$ et $[a_1, a_2, \dots]$ sont des abbréviations respectives de $[0; a_1, \dots, a_n]$ et $[0; a_1, a_2, \dots]$.

Soit k un entier naturel et $c = [a_0; a_1, a_2, \dots]$ une fraction continue (on suppose que si c est finie de longueur n, alors $k \leq n$). La fraction continue $s_k = [a_0; a_1, \dots, a_k]$ s'appelle le *segment initial d'ordre k* de c. La fraction continue $r_k = [a_k; a_{k+1}, a_{k+2}, \dots]$ s'appelle le *reste d'ordre k* de c.

1.2 Convergence d'une fraction continue

Tout segment initial $[a_0; a_1, \dots, a_k]$ d'une fraction continue, étant une fraction continue finie, peut s'écrire sous la forme d'une fraction usuelle dans laquelle le numérateur et le dénominateur sont des expressions polynomiales en a_0, a_1, ..., a_k à coefficients entiers. La notion de convergent répond à cette considération, en introduisant en outre des règles pour que le numérateur et le dénominateur de cette fraction soient univoquement déterminés.

Soit $c = [a_0; a_1, a_2, \dots]$ une fraction continue. Nous définissons les *coefficients de convergence* p_k et q_k par récurrence (pour k naturel dans le cas d'une fraction continue infinie et $k \leq n$ dans le cas d'une fraction continue finie de longueur n). Les coefficients de convergence d'ordre 0 de c sont par définition $p_0 = a_0$ et $q_0 = 1$. Supposons que nous ayons défini les coefficients de convergence d'ordre $k-1$. Si p' et q' sont les coefficients de convergence d'ordre $k-1$ de $[a_1; a_2, \dots]$, alors les coefficients de convergence d'ordre k de c sont par définition $p_k = a_0 p' + q'$ et $q_k = p'$. On appelle *convergent d'ordre k* de c le nombre p_k/q_k lorsque q_k est non nul, auquel cas on dit que le convergent d'ordre k *est défini* ; si q_k est nul, le convergent d'ordre k *n'est pas défini*.

Nous sommes maintenant capables de donner une définition commode de la convergence d'une fraction continue. Une fraction continue finie $c = [a_0; a_1, \dots, a_n]$ est *convergente* ou *converge* (vers le nombre réel α), si le convergent d'ordre n est défini et égal à α ; on écrit dans ce cas $\alpha = [a_0; a_1, \dots, a_n]$. Une fraction continue infine $c = [a_0; a_1, a_2, \dots]$ est *convergente* ou *converge* (vers le nombre réel α) si au plus un nombre fini de convergents ne sont pas définis et si $\lim_{k\to\infty} p_k/q_k = \alpha$ (le terme p_k/q_k étant défini pour k assez grand) ; on écrit $\alpha = [a_0; a_1, a_2, \dots]$. Dans les deux cas, le nombre α est la *valeur* de la fraction continue.

Supposons un instant que tous les convergents sont bien définis, pour donner une idée de leur signification. Puisque $[a_0; a_1, \dots, a_k] = a_0 + 1/[a_1; a_2, \dots, a_k]$ (cf. les propositions 2 et 3 ci-dessous), il s'ensuit que si $[a_1; a_2, \dots, a_k] = p'/q'$, alors $[a_0; a_1, \dots, a_k] = (a_0p' + q')/p'$. Lorsque cela a un sens, le convergent d'ordre k de c est donc la valeur du segment initial d'ordre k de c.

La distinction entre les fractions continues finies et infinies est analogue à celle qui existe entre une somme finie et une série infinie (si l'on fait abstraction des précautions à prendre, inévitables lorsque l'on a affaire à des quotients, pour qu'une fraction continue finie ait un sens). En effet, on considère en général une série infinie comme une expression formelle, mais lorsqu'elle converge vers un nombre on utilise la même expression pour représenter ce nombre. Toutefois, on peut toujours écrire une somme finie comme une série infinie avec au plus un nombre fini de termes non nuls, et ainsi il est possible de restreindre l'étude des sommes et des séries à l'étude des séries uniquement. Dans le cas des fractions continues, la situation est plus délicate, puisqu'il n'est pas possible de transformer trivialement une fraction continue finie en une fraction continue infinie de même valeur. Il ne serait donc pas pertinent de considérer uniquement les fractions continues infinies.

La suite des convergents d'une fraction continue infinie est l'analogue de la suite des sommes partielles d'une série, et la représentation du convergent d'ordre k comme quotient des coefficients de convergence d'ordre k correspond à la représentation de la k-ième somme partielle comme somme des k premiers termes de la série.

Remarquons que la définition de convergence que nous avons donnée n'implique pas qu'une fraction continue finie convergente ait un sens en tant qu'expression algébrique. Par exemple, la fraction continue

$$[2, 1, 0] = \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{0}}}$$

converge vers $1/2$ car $p_3 = 1$ et $q_3 = 2$, malgré l'absurdité de l'expression. Cependant, il est facile de voir que si une fraction continue finie représente un nombre en tant qu'expression algébrique, alors elle converge vers ce nombre.

Introduisons une dernière convention. Pour des raisons qui seront apparentes dans la proposition qui suit, nous posons $p_{-1} = 1$ et $q_{-1} = 0$ pour toute fraction continue.

Proposition 1. *Soient $c = [a_0; a_1, a_2, \dots]$ une fraction continue et p_k, q_k ses coefficients de convergence. Pour tout $k \geq 1$ (et $k \leq n$ si c est de longueur n),*

(i) $p_k = a_kp_{k-1} + p_{k-2}$,

(ii) $q_k = a_kq_{k-1} + q_{k-2}$,

(iii) $q_kp_{k-1} - p_kq_{k-1} = (-1)^k$,

(iv) $q_kp_{k-2} - p_kq_{k-2} = (-1)^{k-1}a_k$.

Preuve. Nous commençons par prouver (i) et (ii) simultanément. Si $k = 1$, la définition des coefficients de convergence donne $p_{-1} = 1$, $q_{-1} = 0$, $p_0 = a_0$, $q_0 = 1$, $p_1 = a_0a_1 + 1$ et $q_1 = a_1$, et le résultat est vérifié. Si $k = 2$, on a de même $p_0 = a_0$, $q_0 = 1$, $p_1 = a_0a_1 + 1$, $q_1 = a_1$, $p_2 = a_0a_1a_2 + a_0 + a_2$ et $q_2 = a_1a_2 + 1$ qui vérifient l'énoncé. Soit $n \geq 3$ un entier. Supposons les formules vraies pour l'entier $n - 1$, et notons c' la fraction continue $[a_1; a_2, a_3, \dots]$ avec coefficients de convergence p'_k et q'_k. L'hypothèse de récurrence appliquée aux coefficients de convergence de c' d'ordre $n - 1$ s'écrit $p'_{n-1} = a_np'_{n-2} + p'_{n-3}$ et $q'_{n-1} = a_nq'_{n-2} + q'_{n-3}$. Par conséquent,

$$\begin{aligned} p_n &= a_0p'_{n-1} + q'_{n-1} \\ &= a_0(a_np'_{n-2} + p'_{n-3}) + (a_nq'_{n-2} + q'_{n-3}) \\ &= a_n(a_0p'_{n-2} + q'_{n-2}) + (a_0p'_{n-3} + q'_{n-3}) \\ &= a_np_{n-1} + p_{n-2} \end{aligned}$$

et de même $q_n = p'_{n-1} = a_n p'_{n-2} + p'_{n-3} = a_n q_{n-1} + q_{n-2}$, ce qui prouve (i) et (ii).

En multipliant la formule (i) par q_{k-1}, la formule (ii) par p_{k-1} et en soustrayant la première à la seconde, on trouve $q_k p_{k-1} - p_k q_{k-1} = -(q_{k-1}p_{k-2} - p_{k-1}q_{k-2})$. Comme $q_0 p_{-1} - p_0 q_{-1} = 1$, a formule (iii) est démontrée par récurrence.

En multipliant la formule (i) par q_{k-2}, la formule (ii) par p_{k-2} et en soustrayant la première à la seconde, on trouve (en utilisant (iii)) $q_k p_{k-2} - p_k q_{k-2} = (q_{k-1}p_{k-2} - p_{k-1}q_{k-2})a_k = (-1)^{k-1}a_k$, ce qui prouve la formule (iv). $\square$

Nous démontrons maintenant deux propositions qui vont nous permettre, dans une certaine mesure, de manipuler les fractions continues comme des objets algébriques. Nous les utiliserons souvent implicitement par la suite.

Proposition 2. *Une fraction continue $c = [a_0; a_1, a_2, \dots]$ converge vers α si et seulement si la fraction continue $c' = [a_1, a_2, \dots]$ converge vers $\alpha - a_0$.*

Preuve. Notons p_k, q_k et p'_k, q'_k les coefficients de convergence respectifs de c et c'. De la définition des coefficients de convergence on déduit facilement que la suite des coefficients q_n ne dépend pas de a_0, c'est-à-dire que $q_n = q'_n$ pour tout $n \geq 0$. Si c ou c' est convergente, il existe un entier $k \geq 0$ tel que q_n est non nul dès que $n \geq k$. Si p et q sont les coefficients de convergence d'ordre $n-1$ de $[a_1; a_2, \dots]$, alors par définition, si $n \geq k$,

$$\frac{p_n}{q_n} = \frac{a_0 p + q}{p} = a_0 + \frac{q}{p} = a_0 + \frac{p'_n}{q'_n},$$

ce qui prouve l'énoncé. $\square$

La proposition 2 réduit ainsi le problème de la convergence de $[a_0; a_1, a_2, \dots]$ à celui de la convergence de $[a_1, a_2, \dots]$.

Proposition 3. *Une fraction continue $c = [a_1, a_2, \dots]$ converge vers un nombre non nul α si et seulement si la fraction continue $c' = [a_1; a_2, \dots]$ converge vers $1/\alpha$.*

Preuve. Notons p_k, q_k et p'_k, q'_k les coefficients de convergence respectifs de c et c'. Soit n un entier naturel. Par définition, $p_n = 0p'_{n-1} + q'_{n-1} = q'_{n-1}$ et $q_n = p'_{n-1}$. Ainsi, si p_n et q_n sont non nuls, alors

$$\frac{p_n}{q_n} = \frac{q'_{n-1}}{p'_{n-1}} \quad \text{et} \quad \frac{q_n}{p_n} = \frac{p'_{n-1}}{q'_{n-1}}.$$

Si c (respectivement c') converge vers un nombre non nul, alors p_n et q_n (respectivement p'_{n-1} et q'_{n-1}) sont non nuls lorsque n est assez grand (ou lorsque n et $n-1$ sont les longueurs de c et c' dans le cas de fractions continues finies), et l'énoncé est démontré. $\square$

1.3 Fractions continues à éléments strictement positifs

Les fractions continues dont les éléments sont strictement positifs présentent l'avantage d'avoir tous leurs convergents définis. En effet, vu la proposition 1, tous les coefficients de convergence d'une telle fraction continue sont strictement positifs, à l'exception de q_{-1} et éventuellement de p_0. En fait, tous les segments initiaux d'une telle fraction continue représentent un nombre en tant qu'expression algébrique, et ce nombre est égal au convergent correspondant.

Proposition 4. *Soit c une fraction continue à éléments strictement positifs. La suite (finie ou infinie) des convergents d'ordre pair est strictement croissante et celle des convergents d'ordre impair est strictement décroissante. De plus, tout convergent d'ordre pair est inférieur ou égal à tout convergent d'ordre impair et si c converge vers α, alors tout convergent d'ordre pair est inférieur ou égal à α et tout convergent d'ordre impair est supérieur ou égal à α.*

Preuve. La formule (iv) de la proposition 1 implique pour tout $k \geq 2$

$$\frac{p_{k-2}}{q_{k-2}} - \frac{p_k}{q_k} = \frac{(-1)^{k-1}a_k}{q_k q_{k-2}}.$$

Puisque a_k, q_k et q_{k-2} sont strictement positifs, la première partie de la proposition est démontrée. La formule (iii) de la proposition 1 donne de même pour tout $k \geq 1$

$$\frac{p_{k-1}}{q_{k-1}} - \frac{p_k}{q_k} = \frac{(-1)^k}{q_k q_{k-1}},$$

et ainsi tout convergent d'ordre impair est supérieur à son successeur d'ordre pair. La deuxième partie est alors un corollaire de la première. □

Lemme 5. *Si $c = [a_0; a_1, \ldots, a_n]$ est une fraction continue finie à éléments strictement positifs, et si $k \geq 1$ est un entier inférieur ou égal à n, alors*

$$c = \frac{p_{k-1} r_k + p_{k-2}}{q_{k-1} r_k + q_{k-2}}. \tag{3}$$

Preuve. La fraction continue c a la même valeur que $c' = [a_0; a_1, \ldots, a_{k-1}, r_k]$, car ces fractions représentent le même nombre en tant qu'expression algébrique. Les coefficients de convergence d'ordre k de c' sont $p'_k = p'_{k-1} r_k + p'_{k-2} = p_{k-1} r_k + p_{k-2}$ et $q'_k = q'_{k-1} r_k + q'_{k-2} = q_{k-1} r_k + q_{k-2}$ par la proposition 1. La valeur de c et c' est donc le convergent p'_k / q'_k, ce qui est exactement la formule (3). □

Nous généralisons maintenant le résultat du lemme 5 aux fractions continues infinies.

Proposition 6. *Soient $c = [a_0; a_1, a_2, \ldots]$ une fraction continue infinie dont les éléments sont strictement positifs, et $n \geq 0$ un entier.*

(i) Si le reste r_n d'ordre n de c est convergent, alors c est convergente.

(ii) Si la fraction continue est convergente, alors tous ses restes sont convergents.

Preuve. (i) Soient p_k et q_k les coefficients de convergence de c et p'_k et q'_k ceux de la fraction continue r_n. Le résultat étant évident si $n = 0$ (car $r_0 = c$), on peut supposer $n \geq 1$. Le lemme 5 appliqué à la fraction continue $c' = [a_0; a_1, \ldots, a_{n+k}]$ implique

$$c' = \frac{p_{n+k}}{q_{n+k}} = \frac{p_{n-1} \frac{p'_k}{q'_k} + p_{n-2}}{q_{n-1} \frac{p'_k}{q'_k} + q_{n-2}}. \tag{4}$$

Ainsi, le quotient p_{n+k} / q_{n+k} converge vers

$$\frac{p_{n-1} r_n + p_{n-2}}{q_{n-1} r_n + q_{n-2}}$$

lorsque k tend vers l'infini.

(ii) Supposons que la fraction c converge vers α et notons comme précédemment p'_k et q'_k les coefficients de convergence de r_n. On suppose dans un premier temps que n est non nul. Par la proposition 4, il est impossible que α soit égal à l'un des convergents de c. En particulier α est différent de p_{n-1}/q_{n-1}. Ainsi, de l'égalité (4), on obtient

$$\frac{p'_k}{q'_k} = \frac{p_{n-2} - \frac{p_{n+k}}{q_{n+k}} q_{n-2}}{\frac{p_{n+k}}{q_{n+k}} q_{n-1} - p_{n-1}}$$

et le dénominateur est certainement non nul si k est assez grand. Par conséquent, le reste r_n converge dans ce cas vers le nombre

$$\frac{p_{n-2} - \alpha q_{n-2}}{\alpha q_{n-1} - p_{n-1}}.$$

Si $n = 0$, alors $r_n = c$ et donc r_n converge par hypothèse. □

De la preuve de la proposition 6 on retiendra en particulier le résultat suivant :

SCHOLIE 7. *Soit $c = [a_0; a_1, a_2, \dots]$ une fraction continue à éléments strictement positifs. Si k est un entier supérieur ou égal à 1 et si le reste r_k d'ordre k de c est convergent, alors c converge vers*

$$\alpha = \frac{p_{k-1}r_k - p_{k-2}}{q_{k-1}r_k - q_{k-2}}.$$

Nous rappelons qu'un *produit infini* est une expression de la forme

$$\prod_{n=0}^{\infty} a_n, \tag{5}$$

où $(a_n)_{n=0}^{\infty}$ est une suite de nombres réels non nuls. On dit qu'un produit infini est *convergent* si la suite des produits partiels $(\prod_{k=0}^{n} a_k)_{n=0}^{\infty}$ converge vers une limite non nulle. Cette limite s'appelle le *produit* de (5).†

Le *terme général* a_n d'un produit infini convergent converge vers 1 lorsque $n \to \infty$. En particulier, si le produit infini (5) converge, il existe un nombre naturel N tel que $a_n > 0$ pour tout $n \geq N$. On dit dans ce cas que le produit infini est *absolument convergent* si la série

$$\sum_{n=N}^{\infty} \log(a_n)$$

est absolument convergente.

LEMME 8. *Soit $(a_n)_{n=0}^{\infty}$ une suite de nombres réels différents de -1. Le produit infini*

$$\prod_{n=0}^{\infty} (1 + a_n) \tag{6}$$

est absolument convergent si et seulement si la série

$$\sum_{n=0}^{\infty} a_n \tag{7}$$

est absolument convergente.

Preuve. Considérons dans un premier temps une série $\sum_{n=0}^{\infty} b_n$ dont le terme général est strictement plus grand que -1. Puisque $\log(1+x)/x \to 1$ lorsque $x \to 0$, on a pour tout $\varepsilon > 0$ les inégalités

$$(1-\varepsilon)|b_n| < |\log(1+b_n)| < (1+\varepsilon)|b_n|$$

dès que n est suffisamment grand. Il s'ensuit immédiatement que les séries

$$\sum_{n=0}^{\infty} |b_n| \quad \text{et} \quad \sum_{n=0}^{\infty} |\log(1+b_n)|$$

convergent simultanément.

Supposons que la série (7) est absolument convergente. Alors il existe un entier N tel que $1 + a_n > 0$ dès que $n \geq N$ et la série

$$\sum_{n=N}^{\infty} \log(1 + a_n) \tag{8}$$

converge absolument. Soit S la somme de la série (8). Dans ce cas, le produit (6) converge vers

$$e^S \prod_{n=0}^{N-1} (1 + a_n),$$

et la convergence est absolue par définition.

† Un produit infini est donc une série dans le groupe topologique multiplicatif $\mathbb{R}^*$.

Réciproquement, supposons que le produit (6) converge absolument. Alors il existe un entier N tel que $1 + a_n > 0$ dès que $n \geq N$, et la série

$$\sum_{n=N}^{\infty} \log(1 + a_n)$$

converge absolument par définition. Par conséquent, la série (7) converge absolument elle aussi. □

Théorème 9 (Critère de convergence). *Pour qu'une fraction continue infinie* $[a_1, a_2, \dots]$ *dont les éléments sont strictement positifs converge, il faut et il suffit que la série*

$$\sum_{k=1}^{\infty} a_k \tag{9}$$

diverge.

Preuve. Rappelons que tous les coefficients de convergence d'une fraction continue à éléments strictement positifs sont strictement positifs, à l'exception ici de p_0 et q_{-1} qui sont nuls.

Le point (iii) de la proposition 1 implique la formule

$$\frac{p_{k-1}}{q_{k-1}} - \frac{p_k}{q_k} = \frac{(-1)^k}{q_k q_{k-1}}$$

pour tout entier $k \geq 1$. La fraction continue c converge si et seulement si la suite des convergents p_k/q_k converge lorsque $k \to \infty$. En vertu du théorème 4, c'est le cas si et seulement si la suite des convergents d'ordre pair et celle des convergents d'ordre impair ont la même limite, ce qui est le cas uniquement si

$$q_k q_{k+1} \to \infty \tag{10}$$

lorsque k tend vers l'infini. La condition (10) est donc nécessaire et suffisante pour la convergence de la fraction continue $[a_1, a_2, \dots]$.

Supposons que la série (9) converge. Alors il existe un entier k_0 tel que $a_k < 1$ pour tout entier $k \geq k_0$. Puisque les coefficients de convergence de $[a_1, a_2, \dots]$ sont strictement positifs, on obtient par la formule (ii) de la proposition 1 que $q_k > q_{k-2}$ pour tout $k \geq 1$. Ainsi $q_k > q_{k-1}$ ou $q_{k-1} > q_{k-2}$. Dans le premier cas, $q_k < q_k a_k + q_{k-2}$ et on obtient, dès que $k \geq k_0$,

$$q_k < \frac{q_{k-2}}{1 - a_k}.$$

Dans le second cas, on obtient (lorsque $k \geq k_0$)

$$q_k < (1 + a_k) q_{k-1} < \frac{q_{k-1}}{1 - a_k}.$$

Dans tous les cas, si $k \geq k_0$, il existe donc un entier $l_1 < k$ tel que

$$q_k < \frac{q_{l_1}}{1 - a_k}.$$

Si $l_1 \geq k_0$, le même résultat s'applique à q_{l_1}, et on construit ainsi une suite d'indice l_1, l_2, ... strictement décroissante, avec chaque fois

$$q_{l_i} < \frac{q_{l_{i-1}}}{1 - a_{l_i}}.$$

La suite s'arrête, disons après n étapes, lorsque $l_n < k_0$, et alors

$$q_k < \frac{q_{l_n}}{(1 - a_k)(1 - a_{l_1}) \dots (1 - a_{l_{n-1}})}.$$

Puisque la série $\sum_{k=k_0}^{\infty} (-a_k)$ converge absolument, le lemme 8 assure que le produit infini

$$\prod_{k=k_0}^{\infty} (1 - a_k)$$

converge vers un nombre positif λ (car tous les produits partiels sont strictement positifs) qui est non nul par définition, et on a

$$(1-a_k)(1-a_{l_1})\dots(1-a_{l_{n-1}}) \geq \prod_{k=k_0}^{\infty}(1-a_k) = \lambda.$$

Si l'on note q le plus grand entier parmi q_0, q_1, ..., q_{k_0-1}, alors $q_k < q/\lambda$ dès que $k \geq k_0$, et ainsi $q_k q_{k+1} < q^2/\lambda^2$. La condition nécessaire (10) n'est pas vérifiée, donc la fraction continue $[a_1, a_2, \dots]$ diverge.

Réciproquement, supposons que la série (9) diverge. Par la formule (ii) de la proposition 1, on obtient que $q_k > q_{k-2}$ dès que $k \geq 1$. Ainsi les suites des coefficients de convergence $(q_{2k})_{k=0}^{\infty}$, d'une part, et $(q_{2k+1})_{k=0}^{\infty}$, d'autre part, sont strictement croissantes. Il s'ensuit que si c est le minimum de $\{q_0, q_1\}$, alors $q_k \geq c$ pour tout entier naturel k. La formule (ii) de la proposition 1 devient $q_k \geq q_{k-2} + c a_k$. En raisonnant par récurrence, nous obtenons les relations

$$q_{2k} \geq q_0 + c\sum_{n=1}^{k} a_{2n} \quad \text{et} \quad q_{2k+1} \geq q_1 + c\sum_{n=1}^{k} a_{2n+1}.$$

En sommant les deux relations, on trouve

$$q_{2k} + q_{2k+1} \geq q_0 + q_1 + \sum_{n=1}^{2k+1} a_n \quad \text{et} \quad q_{2k} + q_{2k-1} \geq q_0 + q_1 + \sum_{n=1}^{2k} a_n,$$

et puisque $q_0 = 1$, on a, pour tout entier $k \geq 1$ pair ou impair,

$$q_k + q_{k-1} > c\sum_{n=1}^{k} a_n.$$

En conclusion, l'un parmi q_k et q_{k-1} est supérieur à $\frac{c}{2}\sum_{n=1}^{k} a_n$, et puisque l'autre excède c, cela implique

$$q_k q_{k-1} > \frac{c^2}{2}\sum_{n=1}^{k} a_n.$$

Par la divergence supposée de la série (9), on obtient que la condition (10) est vérifiée, et par conséquent la fraction continue $[a_1, a_2, \dots]$ converge. $\square$

1.4 Fractions continues simples

Nous voulons imposer des restrictions sur la forme d'une fraction continue afin d'assurer, d'une part, la convergence et, d'autre part, l'unicité d'une fraction continue convergeant vers un nombre réel donné.

Nous constatons que la fraction continue $[\dots, a_k, 0, a_{k+1}, \dots]$ (avec $a_{k+1} \neq 0$), si elle est convergente, converge vers le même nombre que la fraction continue $[\dots, a_k + a_{k+1}, \dots]$. Il est donc naturel d'exiger qu'aucun élément d'une fraction continue (finie ou infinie) ne soit nul. Nous remarquons également que la fraction continue finie $[a_0; a_1, \dots, a_n, 1]$, si elle est convergente, converge vers le même nombre que la fraction continue $[a_0; a_1, \dots, a_n + 1]$. On pourra donc exiger que le dernier élément d'une fraction continue finie soit différent de l'unité. Précisons ces résultats.

Nous dirons qu'une fraction continue (finie ou infinie) $c = [a_0; a_1, a_2, \dots]$ est *simple*[†] si les propriétés suivantes sont vérifiées.

(i) Le terme initial est un entier et les éléments sont des entiers strictement positifs.

(ii) Si c est finie de longueur $n \geq 1$, alors $a_n > 1$.

[†]Les termes « régulier » et « simple » ne sont pas utilisés de façon consistante dans la littérature. Le choix de la terminologie que nous avons fait n'est donc pas universel.

La proposition suivante découle directement de la proposition 1.

Proposition 10. *Les suites (finies ou infinies) des coefficients de convergence p_n et q_n d'une fraction continue simple sont strictement croissantes dès que $n \geq 1$.*

Nous allons montrer que les fractions continues simples ont les propriétés que nous recherchions.

Proposition 11. *Toute fraction continue simple est convergente.*

Preuve. Toute fraction continue simple finie converge, car elle a un sens en tant qu'expression algébrique. Si $c = [a_0; a_1, a_2, \dots]$ est une fraction continue simple infinie, alors $a_n \geq 1$ pour tout entier $n \geq 1$, et par conséquent la série $\sum_{n=1}^{\infty} a_n$ diverge. Par le théorème 9 (et la proposition 2), la fraction continue simple c est convergente. □

Théorème 12. *Pour tout nombre réel α, il existe une unique fraction continue simple $[a_0; a_1, a_2, \dots]$ qui converge vers α. Cette fraction est finie si α est rationnel ; elle est infinie si α est irrationnel.*

Preuve. existence. Supposons dans un premier temps que α est irrationnel. Notons $a_0 = \lfloor \alpha \rfloor$ le plus grand entier inférieur ou égal à α et

$$r_1 = \frac{1}{\alpha - a_0},$$

puis définissons par récurrence

$$a_n = \lfloor r_n \rfloor \quad \text{et} \quad r_{n+1} = \frac{1}{r_n - a_n}. \tag{11}$$

Nous devons vérifier que $r_n - a_n$ n'est jamais nul pour que cela ait un sens. C'est une conséquence de l'irrationnalié de α. En effet, si l'on suppose que r_n est défini, alors

$$\alpha = [a_0; r_1] = [a_0; a_1, r_2] = \cdots = [a_0; a_1, \dots, a_{n-1}, r_n] ;$$

donc r_n n'est pas entier, car sinon α serait rationnel. Les suites $(a_n)_{n=0}^{\infty}$ et $(r_n)_{n=1}^{\infty}$ sont donc bien définies. Montrons que la fraction continue simple infinie $[a_0; a_1, a_2, \dots]$ converge vers α. Puique α est la valeur de la fraction continue finie $c_n = [a_0; a_1, \dots, a_n, r_{n+1}]$ à éléments strictement positifs dont le reste d'ordre n est r_n, le lemme 5 nous donne

$$\alpha = \frac{p_{n-1} r_n + p_{n-2}}{q_{n-1} r_n + q_{n-2}},$$

où p_k et q_k sont les coefficients de convergence d'ordre k de $[a_0; a_1, a_2, \dots]$ (qui sont aussi les coefficients de convergence d'ordre k de c_n si $k \leq n$). D'autre part, en vertu de la proposition 1, le convergent d'ordre n de c_n s'écrit

$$\frac{p_n}{q_n} = \frac{p_{n-1} a_n + p_{n-2}}{q_{n-1} a_n + q_{n-2}},$$

et on déduit à l'aide de la proposition 1 l'inégalité

$$\left| \alpha - \frac{p_n}{q_n} \right| = \left| \frac{(p_{n-1} q_{n-2} - q_{n-1} p_{n-2})(r_n - a_n)}{(q_{n-1} r_n + q_{n-2})(q_{n-1} a_n + q_{n-2})} \right| < \frac{1}{(q_{n-1} r_n + q_{n-2})(q_{n-1} a_n + q_{n-2})} < \frac{1}{q_n^2}.$$

Comme la suite $(q_n)_{n=0}^{\infty}$ est strictement croissante (cf. proposition 10), il s'ensuit que $p_n / q_n \to \alpha$ lorsque n tend vers l'infini.

Supposons maintenant que α est un nombre rationnel. Si α est un entier, alors α est la valeur de la fraction continue simple $[\alpha]$. Sinon, le processus (11) défini ci-dessus peut être appliqué tant que r_n n'est pas un entier. Si r_N est un entier pour un certain $N \geq 1$, alors la fraction continue finie $[a_0; a_1, \dots, a_{N-1}, r_N]$ converge vers α, et c'est une fraction continue simple car $r_N > 1$ par construction. Nous allons montrer qu'en fait il existe toujours un tel $N \geq 1$ pour lequel r_N est un entier. En effet, puisque α est rationnel, il est évident que si r_n est défini, alors r_n est rationnel. Soient a et b des entiers (strictement positifs) premiers entre eux tels que $r_n = a/b$. Si r_n n'est pas un entier, alors $r_n - a_n = c/b$ où $c = a - ba_n$ est un entier tel que $0 < c < b$, et donc $r_{n+1} = b/c$ avec b et c premiers entre eux. Par conséquent, la suite des dénominateurs canoniques des nombres r_n est strictement décroissante, et on en déduit qu'il existe $N \geq 1$ pour lequel r_N est un entier.

UNICITÉ. Soient $c = [a_0; a_1, a_2, \dots]$ et $c' = [a'_0; a'_1, a'_2, \dots]$ deux fractions continues simples convergeant vers le même nombre réel α. On suppose sans perte de généralité que la longueur de c est inférieure ou égale à la longueur de c'. Nous prétendons en général que si $[b_0; b_1, b_2, \dots]$ est une fraction continue simple convergeant vers β alors

$$b_0 = \lfloor \beta \rfloor. \tag{12}$$

Nous savons par la proposition 2 que $[b_0; b_1, b_2, \dots] = b_0 + [b_1, b_2, \dots]$. Il suffit donc de montrer que la valeur de $[b_1, b_2, \dots]$ est comprise dans $[0, 1[$. Tous les segments initiaux de $[b_1, b_2, \dots]$ prennent des valeurs dans $]0, 1[$.[†] Si $[b_1, b_2, \dots]$ est finie, alors sa valeur est également comprise dans $]0, 1[$. Si $[b_1, b_2, \dots]$ est infinie, on obtient par passage à la limite que $[b_1, b_2, \dots]$ converge vers un nombre dans $[0, 1]$. Il reste à remarquer que dans le cas d'une fraction continue infinie, puisque la suite des convergents possède une sous-suite strictement croissante et une sous-suite strictement décroissante (cf. proposition 4), sa valeur ne peut être ni 0 ni 1. L'assertion (12) est ainsi démontrée.

On a donc montré que $a_0 = a'_0 = \lfloor \alpha \rfloor$. Supposons par récurrence que $a_i = a'_i$ pour tout entier naturel $i \leq n$ (où n est strictement inférieur à la longueur de c si c est finie). Si r_k et r'_k dénotent les restes d'ordre k de c et c' respectivement, et si p_k et q_k sont les coefficients de convergence de c (et donc aussi ceux de c' si $k \leq n$), alors par la scholie 7

$$\alpha = \frac{p_n r_{n+1} + p_{n-1}}{q_n r_{n+1} + q_{n-1}} = \frac{p_n r'_{n+1} + p_{n-1}}{q_n r'_{n+1} + q_{n-1}},$$

d'où l'on déduit $r_{n+1} = r'_{n+1}$, et en utilisant l'assertion (12) on obtient que $a_{n+1} = a'_{n+1}$. Ainsi, le terme initial et tous les éléments de c coïncident avec le terme initial et les éléments correspondants de c'. Il reste à montrer que c' est de même longueur que c. Si c est infinie, alors c' est aussi infinie.

Supposons que c est finie de longueur n et par l'absurde que la longueur de c' est strictement supérieure à la longueur de c. On a, en vertu de la scholie 7, $[a_0; a_1, \dots, a_n] = [a_0; a_1, \dots, a_n + z_n]$ où z_n est la valeur de $[a'_{n+1}, \dots]$; mais ceci implique $z_n = 0$, et nous avons vu plus haut que ceci est impossible.

De ceci on conclut que c et c' sont soit finies de même longueur soit infinies et que leurs éléments coïncident. □

La proposition 11 et le théorème 12 établissent une bijection entre l'ensemble des nombres réels et l'ensemble des fractions continues simples, en associant à chaque fraction continue simple sa valeur. En général, on ne différencie pas une fraction continue simple du nombre vers lequel elle converge.

Nous avons besoin d'un dernier résultat élémentaire sur les fractions continues simples.

PROPOSITION 13. *Soit c une fraction continue simple infinie. Pour tout $k \geq 0$, on a l'estimation*

$$q_k > 2^{\frac{k-1}{2}}.$$

Preuve. Si $k = 0$, le résultat est vérifié car $q_k = 1$ et $2^{\frac{k-1}{2}} = 1/\sqrt{2}$. Sinon, supposons le résultat vrai pour tout entier naturel inférieur ou égal à $k - 1$. Comme $k > 0$, on a par hypothèse $a_k \geq 1$. Alors par la formule (ii) de la proposition 1, $q_k = q_{k-1} a_k + q_{k-2} > 2^{\frac{k-2}{2}} + 2^{\frac{k-3}{2}} = 2^{\frac{k-3}{2}}(\sqrt{2} + 1) > 2^{\frac{k-1}{2}}$. La proposition est ainsi démontrée par récurrence. □

[†] On utilise ici le fait que le dernier élément d'une fraction continue finie n'est pas égal à l'unité.

2 Théorie de la mesure des fractions continues

2.1 Résultats préliminaires de théorie de la mesure

Nous supposons que le lecteur est familier avec la théorie moderne de la mesure. Nous démontrons ici quelques résultats élémentaires que nous utiliserons par la suite.

Proposition 14. *Soient* $(X, \mathscr{A}, \mu)$ *un espace mesuré et* $(A_n)_{n=1}^{\infty}$ *une suite d'éléments de* $\mathscr{A}$*. Si la série*

$$\sum_{n=1}^{\infty} \mu(A_n) \tag{13}$$

converge dans $[0, \infty[$*, alors presque tout* $x \in X$ *n'appartient qu'à un nombre fini de* A_n*.*

Preuve. Notons A l'ensemble des nombres réels apparaissant dans un nombre infini de A_n. Nous voulons montrer que $\mu(A) = 0$. Observons que pour tout entier naturel m

$$A \subseteq \bigcup_{n=m}^{\infty} A_n$$

et en particulier

$$\mu(A) \leq \sum_{n=m}^{\infty} \mu(A_n).$$

Mais la convergence de la série (13) implique que pour tout $\varepsilon > 0$, il existe un entier M tel que $\sum_{n=M}^{\infty} \mu(A_n) < \varepsilon$. Ainsi, la mesure de A est arbitrairement petite, donc nulle. Cela signifie que presque tout x dans X n'appartient qu'à un nombre fini de A_n, ce qu'il fallait démontrer. □

Proposition 15. *Soient* $(X, \mathscr{A}, \mu)$ *un espace mesuré* σ*-fini et* Y *un élément de* $\mathscr{A}$*. S'il existe un nombre réel positif* C *tel que tout sous-ensemble mesurable de* Y *dont la mesure est finie a mesure inférieure à* C*, alors la mesure de* Y *est finie.*

Preuve. Soit $\{A_i \mid i \in \mathbb{N}\} \subseteq \mathscr{A}$ une partition de Y telle que $\mu(A_i)$ est fini pour tout entier naturel i. Une telle partition de Y existe par l'hypothèse de σ-finitude. Définissons une suite réelle $(u_n)_{n=0}^{\infty}$ par

$$u_n = \mu\left(\bigcup_{k=0}^{n} A_k\right).$$

Alors $u_n \leq C$ pour tout entier naturel n. Par conséquent, $\mu(Y) = \mu\big(\bigcup_{k=0}^{\infty} A_k\big) = \lim_{n\to\infty} u_n \leq C$, ce qui montre que $\mu(Y)$ est fini. □

Précisons ici les notations que nous allons utiliser. La σ-algèbre des sous-ensembles mesurables de $\mathbb{R}$ sera notée $\mathscr{M}$, celle des boréliens de $\mathbb{R}$ sera notée $\mathscr{B}$. Conformément à Khinchin [7], nous dénoterons par $\mathfrak{M}A$ la mesure de Lebesgue d'un élément A de $\mathscr{M}$.

Proposition 16. *La* σ*-algèbre borélienne de* $\mathbb{R}$ *est engendrée par* $\mathscr{E} = \{]-\infty, a] \mid a \in \mathbb{R}\}$*.*

Preuve. Notons $\mathscr{A}$ la σ-algèbre engendrée par $\mathscr{E}$. Puisque $\mathscr{E} \subset \mathscr{B}$, il est clair que $\mathscr{A} \subseteq \mathscr{B}$. Il suffit donc de montrer que tout ouvert de $\mathbb{R}$ (muni de la topologie usuelle) appartient à $\mathscr{A}$. Les ensembles de la forme $]a, b[$, pour deux nombres réels $a < b$, peuvent s'écrire

$$]a, b[= \left(\bigcup_{n=1}^{\infty}]-\infty, b - 1/n]\right) \cap (\mathbb{R} \setminus]-\infty, a]),$$

et par conséquent appartiennent à $\mathscr{A}$. Comme $\mathbb{R}$ est un espace métrique séparable, il existe un sous-ensemble dense et dénombrable D de $\mathbb{R}$. Alors $\{]a - 1/n, a + 1/n[\mid a \in D, n \in \mathbb{N} \setminus \{0\}\}$ est une base dénombrable de la topologie, c'est-à-dire que tout ouvert de $\mathbb{R}$ est la réunion dénombrable d'intervalles de la forme $]a, b[$, donc appartient à $\mathscr{A}$. □

Grâce la proposition 16, pour montrer qu'une fonction f est mesurable sur un espace mesurable $(X, \mathscr{A})$, il suffit, en vertu des propriétés de l'image réciproque par rapport aux réunions et aux complémentaires, de montrer que la préimage par f d'un ensemble de la forme $]-\infty, a]$ appartient à $\mathscr{A}$.

2.2 Les éléments d'une fraction continue simple comme fonctions

Comme seuls les éléments a_1, a_2, … d'une fraction continue simple $[a_0; a_1, a_2, \dots]$ nous intéressent, nous pouvons restreindre notre étude aux seules fractions continues simples représentant les nombres de l'intervalle $[0, 1[$. De plus, comme nous allons nous intéresser à la mesure d'ensembles définis par des fractions continues, on peut également oublier les nombres rationnels dont la mesure de Lebesgue est nulle. Ainsi, notre étude dans ce chapitre se portera sur les nombres réels dans

$$I = [0, 1[\setminus \mathbb{Q}.$$

En conséquence, par le théorème 12, toutes les fractions continues simples représentant des nombres de I seront infinies.† On dénotera par le terme *intervalle irrationnel* l'intersection d'un intervalle avec l'ensemble des nombres irrationnels. Il est clair que tout intervalle irrationnel est un borélien, en particulier mesurable au sens de Lebesgue.

L'espace mesuré dans lequel nous travaillerons sera le triple $(I, \mathscr{B}', \mathfrak{M}|_{\mathscr{B}'})$ où $\mathscr{B}' = \{B \cap I | B \in \mathscr{B}\}$ est la famille des boréliens de I. Vu que $\mathfrak{M}I = 1$, ce triple est un espace de probabilité. On pourra ainsi reformuler certains résultats exprimés en termes de mesure dans le langage plus intuitif (mais évidemment strictement équivalent) des probabilités. Lorsque nous considérerons des fonctions mesurables définies sur I, il s'agira de fonctions mesurables sur $(I, \mathscr{B}')$.

Dans ce chapitre, toutes les intégrales seront prises au sens de Lebesgue, et l'on omettra de le mentionner à chaque fois.

Nous définissons une suite de fonction à valeurs réelles $(a_i)_{i=1}^{\infty}$ sur l'intervalle irrationnel I par $a_i([k_1, k_2, \dots]) = k_i$. Nous voyons ainsi les éléments d'une fraction continue comme fonctions du nombre qu'elle représente. De manière analogue, nous définissons pour tout entier $k \geq 1$ des fonctions p_k, q_k, r_k et $z_k = r_k - a_k$ sur l'intervalle irrationnel I.

Si n_1, …, n_t, k_1, …, k_t sont des entiers strictement positifs, nous utilisons la notation

$$E\begin{pmatrix} n_1 & \dots & n_t \\ k_1 & \dots & k_t \end{pmatrix}$$

introduite par Khinchin [8] pour représenter l'ensemble des nombres α dans I pour lesquels $a_{n_1}(\alpha) = k_1$, …, $a_{n_t}(\alpha) = k_t$.

Soit k un entier naturel. Un nombre α dans I a la propiété $a_1(\alpha) = k$ si et seulement si $k \leq r_1(\alpha) < k + 1$, c'est-à-dire, puisque $1/\alpha = r_1(\alpha)$, si et seulement si $1/(k+1) < \alpha < 1/k$. Ainsi, l'ensemble

$$E\begin{pmatrix} 1 \\ k \end{pmatrix}$$

est l'intervalle irrationnel $[1/(k+1), 1/k] \setminus \mathbb{Q}$. Un tel intervalle s'appelle un *intervalle de premier rang.*

Considérons l'ensemble des nombres x de I ayant la propriété $a_1(x) = k$ pour un certain entier strictement positif k. Un nombre α dans cet ensemble est tel que $a_2(\alpha) = l$ si et seulement si $l < r_2(\alpha) < l + 1$, c'est-à-dire si et seulement si

$$\frac{1}{k + \frac{1}{l}} < \alpha < \frac{1}{k + \frac{1}{l+1}}.$$

Ainsi, l'ensemble

$$E\begin{pmatrix} 1 & 2 \\ k & l \end{pmatrix}$$

est l'intervalle irrationnel

$$\left[\frac{1}{k + \frac{1}{l}}, \frac{1}{k + \frac{1}{l+1}}\right] \setminus \mathbb{Q}.$$

Nous appellerons un tel intervalle un *intervalle de second rang.*

†Une autre approche possible consiste à considérer tout de même les nombres rationnels et à étendre par une convention arbitraire la suite des éléments de leurs fractions continues. Dans les deux cas, les nombres rationnels sont totalement insignifiants dans l'étude qui va suivre.

Nous allons généraliser ces résultats. Considérons l'ensemble J_n des nombres α de I ayant les propriétés $a_1(\alpha) = k_1$, $a_2(\alpha) = k_2$, …, $a_n(\alpha) = k_n$. En vertu de la scholie 7, un tel nombre α s'écrit

$$\alpha = \frac{p_n(\alpha)r_{n+1}(\alpha) + p_{n-1}(\alpha)}{q_n(\alpha)r_{n+1}(\alpha) + q_{n-1}(\alpha)}, \tag{14}$$

où le reste $r_{n+1}(\alpha)$ parcourt $]1, \infty[\setminus \mathbb{Q}$ lorsque α parcourt J_n, tandis que les coefficients de convergence d'ordre n et $n-1$ sont identiques pour tout α dans J_n, étant uniquement déterminés par les entiers k_1, …, k_n. Ainsi, J_n est l'ensemble des nombres α décrit par (14) lorsque $r_{n+1}(\alpha)$ varie dans $]1, \infty[\setminus \mathbb{Q}$. La fonction $r_{n+1}(\alpha) \mapsto \alpha$ définie par (14) étant monotone, on conclut que J_n est l'intervalle irrationnel dont les extrémités sont

$$\frac{p_n(\alpha) + p_{n-1}(\alpha)}{q_n(\alpha) + q_{n-1}(\alpha)} \quad \text{et} \quad \frac{p_n(\alpha)}{q_n(\alpha)}. \tag{15}$$

Un tel intervalle irrationnel est appelé un *intervalle de rang n*.

Il est immédiat que tous les intervalles de rang n forment une partition de I. Par convention, on dit que I est l'unique *intervalle de rang* 0, de telle sorte que les formules (15) étendent leur validité au cas $n = 0$. Chaque intervalle de rang n est lui-même partitionné par une quantité dénombrable d'intervalles de rang $n + 1$. Aussi l'ensemble des intervalles de rang n est-il dénombrable. Un ensemble de la forme

$$E\begin{pmatrix} n \\ k \end{pmatrix}$$

est donc la réunion d'intervalles de rang n pris dans chaque intervalle de rang $n-1$.

Si $u(a_1, \dots, a_n)$ est une expression dépendant des n premiers éléments d'une fraction continue infinie $[a_1, a_2, \dots]$, nous utiliserons le symbole de sommation

$$\sum^{(n)} u(a_1, \dots, a_n) \tag{16}$$

comme abbréviation de

$$\sum_{a_1,\dots,a_n \in \mathbb{N}\setminus\{0\}} u(a_1, \dots, a_n).$$

Lorsque l'on utilisera un tel symbole, il faudra bien sûr s'assurer que la famille

$$(u(a_1, \dots, a_n))_{a_1,\dots,a_n \in \mathbb{N}\setminus\{0\}}$$

est sommable (c'est-à-dire, puisque le nombre de termes est dénombrable, qu'une série correspondante est absolument convergente) pour que cela ait un sens. De façon équivalente, si A est un système de représentants des intervalles de rang n, la somme (16) peut aussi s'écrire

$$\sum_{\alpha \in A} u(a_1(\alpha), \dots, a_n(\alpha)).$$

On dit alors que la somme est prise *sur tous les intervalles de rang n*. Parmi les expressions dépendant de a_1, …, a_n, nous avons en particulier les coefficients de convergence p_k et q_k d'ordre k inférieur ou égal à n.

Proposition 17. *Quels que soient les entiers strictement positifs n_1, …, n_t, k_1, …, k_t, l'ensemble*

$$E\begin{pmatrix} n_1 & \dots & n_t \\ k_1 & \dots & k_t \end{pmatrix}$$

est un borélien. En particulier, il est mesurable au sens de Lebesgue.

Preuve. Il est clair que

$$E\begin{pmatrix} n_1 & \dots & n_t \\ k_1 & \dots & k_t \end{pmatrix} = \bigcap_{i=1}^{t} E\begin{pmatrix} n_i \\ k_i \end{pmatrix},$$

et chaque ensemble dans l'intersection est une réunion dénombrable d'intervalles irrationnels. L'ensemble considéré est ainsi une intersection finie de réunions dénombrables de boréliens, donc est un borélien. □

PROPOSITION 18. *Les fonctions a_i, p_i, q_i, r_i et z_i définies ci-dessus sont mesurables sur $(I, \mathscr{B}')$.*

Preuve. En vertu de la proposition 16, il suffit de montrer que pour tout x dans $\mathbb{R}$ l'image réciproque par ces fonctions de $]-\infty, x]$ est un borélien de I. Si $x < 0$, alors $a_i^{-1}(]-\infty, x])$ est l'ensemble vide qui est un borélien. Sinon

$$a_i^{-1}(]-\infty, x]) = \bigcup_{k=0}^{\lfloor x \rfloor} E\begin{pmatrix} i \\ k \end{pmatrix},$$

qui est un borélien en tant qu'union finie de boréliens. Les fonctions p_i et q_i sont par définition des sommes finies de produits finis des fonctions mesurables a_k avec $k \leq i$, et sont par conséquent mesurables.

Si $x \leq 1$, alors $r_i^{-1}(]-\infty, x])$ est l'ensemble vide qui est un borélien. Sinon,

$$r_i^{-1}(]-\infty, x]) = \bigcup_{a_1, \ldots, a_{i-1} \in \mathbb{N} \setminus \{0\}} \{[a_1, \ldots, a_{i-1}, r_i] \mid r_i \in]1, x] \setminus \mathbb{Q}\}.$$

Chaque terme dans cette réunion dénombrable est un intervalle irrationnel donné par les extrémités

$$\frac{p_n + p_{n-1}}{q_n + q_{n-1}} \quad \text{et} \quad \frac{p_n x + p_{n-1}}{q_n x + q_{n-1}},$$

donc un borélien. Finalement, $z_i = r_i - a_i$ est une fonction mesurable en tant que somme finie de fonctions mesurables. □

En termes de probabilité, les fonctions a_i, p_i, q_i, r_i et z_i sont donc des variables aléatoires sur l'espace de probabilité $(I, \mathscr{B}', \mathfrak{M}|_{\mathscr{B}'})$.

Une question naturelle qui se pose est celle de la mesure d'un ensemble de la forme

$$E\begin{pmatrix} n \\ r \end{pmatrix}.$$

Le cas $n = 1$ est facile car cet ensemble est un unique intervalle irrationnel :

$$\mathfrak{M}E\begin{pmatrix} 1 \\ r \end{pmatrix} = \frac{1}{r} - \frac{1}{r+1}.$$

Si $n = 2$, l'ensemble correspond à une réunion dénombrable d'intervalle irrationnels, dont la mesure est

$$\mathfrak{M}E\begin{pmatrix} 2 \\ r \end{pmatrix} = \sum_{k=1}^{\infty} \left(\frac{1}{k + \frac{1}{r+1}} - \frac{1}{k + \frac{1}{r}} \right).$$

La réponse devient vite extrémement compliquée. C'est pourquoi nous allons nous contenter d'une approximation simple exprimée dans le théorème suivant.

THÉORÈME 19. *Pour tous entiers strictement positifs n et r,*

$$\frac{1}{3r^2} < \mathfrak{M}E\begin{pmatrix} n \\ r \end{pmatrix} < \frac{2}{r^2}.$$

Preuve. Nous considérons un intervalle J_{n-1} de rang $n-1$ défini par

$$J_{n-1} = E\begin{pmatrix} 1 & \ldots & n-1 \\ a_1 & \ldots & a_{n-1} \end{pmatrix},$$

et un intervalle J_n^r de rang n définie par

$$J_n^r = E\begin{pmatrix} 1 & \ldots & n-1 & n \\ a_1 & \ldots & a_{n-1} & r \end{pmatrix},$$

avec en particulier $J_n^r \subset J_{n-1}$. Notons p_k et q_k les coefficients de convergence de la fraction continue $[a_1, \ldots, a_{n-1}]$. Nous avons déjà vu que les extrémités de l'intervalle irrationnel J_{n-1} sont les nombres

$$\frac{p_{n-1}+p_{n-2}}{q_{n-1}+q_{n-2}} \quad \text{et} \quad \frac{p_{n-1}}{q_{n-1}},$$

et que les éléments de J_{n-1} sont de la forme

$$\frac{p_{n-1}r_n+p_{n-2}}{q_{n-1}r_n+q_{n-2}}$$

pour un certain r_n dans $]1,\infty[\setminus\mathbb{Q}$. Puisque les nombres de J_{n-1} qui appartiennent également à J_n^r sont ceux pour lesquels $r \le r_n < r+1$, il s'ensuit que les extrémités de J_n^r sont les nombres

$$\frac{p_{n-1}r+p_{n-2}}{q_{n-1}r+q_{n-2}} \quad \text{et} \quad \frac{p_{n-1}(r+1)+p_{n-2}}{q_{n-1}(r+1)+q_{n-2}}.$$

Par conséquent, vu la proposition 1, la mesure de l'intervalle irrationnel J_n^r est donnée par

$$\begin{aligned}\mathfrak{M}J_n^r &= \left|\frac{p_{n-1}r+p_{n-2}}{q_{n-1}r+q_{n-2}} - \frac{p_{n-1}(r+1)+p_{n-2}}{q_{n-1}(r+1)+q_{n-2}}\right| \\ &= \frac{1}{(q_{n-1}r+q_{n-2})(q_{n-1}(r+1)+q_{n-2})} \\ &= \frac{1}{q_{n-1}^2 r^2\left(1+\dfrac{q_{n-2}}{rq_{n-1}}\right)\left(1+\dfrac{1}{r}+\dfrac{q_{n-2}}{rq_{n-1}}\right)}\end{aligned}$$

tandis que celle de J_{n-1} est

$$\mathfrak{M}J_{n-1} = \left|\frac{p_{n-1}+p_{n-2}}{q_{n-1}+q_{n-2}} - \frac{p_{n-1}}{q_{n-1}}\right| = \frac{1}{q_{n-1}(q_{n-1}+q_{n-2})} = \frac{1}{q_{n-1}^2\left(1+\dfrac{q_{n-2}}{q_{n-1}}\right)}.$$

En observant les inégalités évidentes

$$\left(1+\frac{q_{n-2}}{rq_{n-1}}\right)\left(1+\frac{1}{r}+\frac{q_{n-2}}{rq_{n-1}}\right) > 1, \quad 1+\frac{q_{n-2}}{q_{n-1}} \le 2, \quad \frac{1+\frac{q_{n-2}}{q_{n-1}}}{1+\frac{q_{n-2}}{rq_{n-1}}} \ge 1 \quad \text{et} \quad 1+\frac{1}{r}+\frac{q_{n-2}}{rq_{n-1}} < 3,$$

on obtient immédiatement une estimation de la densité de J_n^r dans J_{n-1} :

$$\frac{1}{3r^2} < \frac{\mathfrak{M}J_n^r}{\mathfrak{M}J_{n-1}} = \frac{1}{r^2}\frac{1+\dfrac{q_{n-2}}{q_{n-1}}}{\left(1+\dfrac{q_{n-2}}{rq_{n-1}}\right)\left(1+\dfrac{1}{r}+\dfrac{q_{n-2}}{rq_{n-1}}\right)} < \frac{2}{r^2},$$

ce qui s'écrit aussi

$$\frac{\mathfrak{M}J_{n-1}}{3r^2} < \mathfrak{M}J_n^r < \frac{2\mathfrak{M}J_{n-1}}{r^2}. \tag{17}$$

Si l'on prend la somme de (17) sur tous les intervalles de rang $n-1$, on trouve le résultat cherché. □

2.3 La distribution de Gauss–Kuz'min

Un résultat fondamental de la théorie de la mesure des fractions continues qui est entre autres à la base du théorème de Khinchin est que la suite des variables aléatoires $(a_i)_{i=1}^{\infty}$ définie plus haut converge en distribution, et la loi limite est donnée par une formule explicite.

Rappelons que nous avons défini, pour tout entier naturel k, des variables aléatoires r_k et $z_k = r_k - a_k$ sur l'espace de probabilité $(I, \mathscr{B}', \mathfrak{M}|_{\mathscr{B}'})$, et que $r_k(\alpha)$ est la valeur du reste d'ordre k de la fraction continue simple correspondant à α. Définissons sur $[0,1]$ une suite de fonction $(m_n)_{n=0}^{\infty}$ par

$$m_n(x) = \mathfrak{M}\{\alpha \in I \mid z_n(\alpha) \leq x\}\,;$$

en d'autres termes, m_n est la distribution sur $[0,1]$ de la variable aléatoire z_n (qui ne prend bien sûr que des valeurs dans $[0,1]$).

Dans une lettre au mathématicien Pierre-Simon Laplace, Carl Friedrich Gauss indiquait qu'il avait réussi à démontrer que

$$m_n(x) \to \frac{\log(1+x)}{\log 2}$$

lorsque $n \to \infty$. Sa preuve ne fut toutefois jamais publiée. Ce n'est qu'en 1928 que le mathématicien russe Rodion Osievich Kuz'min [10] démontrait le résultat, tout en donnant une estimation de la différence

$$m_n(x) - \frac{\log(1+x)}{\log 2}. \tag{18}$$

L'année suivante, le mathématicien français Paul Lévy [12] publiait indépendamment une preuve du même résultat avec une approximation encore meilleure de (18). Nous suivons ici la démonstration originale de Kuz'min, reprise par Khinchin [7].

Lemme 20. *Pour tout enier naturel n, z_n est une variable aléatoire continue.*

Preuve. Il suffit de montrer que l'ensemble $\{z_n = x\} = \{\alpha \in I \,|\, z_n(\alpha) = x\}$ est de mesure nulle pour tout x réel. La fonction z_n ne prend que des valeurs irrationnelles dans $]0,1[$, donc si $x \geq 1$, si $x \leq 0$ ou si x est rationnel, l'ensemble considéré est l'ensemble vide. Si $0 < x < 1$ et x est irrationnel, l'ensemble $\{z_n = x\}$ contient exactement les nombres représentés par des fractions continues de la forme $[a_1, a_2, \ldots, a_n + x]$, où les a_i sont des entiers strictement positifs. Ainsi l'ensemble $\{z_n = x\}$ est dénombrable, et par conséquent de mesure nulle. □

Proposition 21. *Pour tout entier naturel n et pour tout x dans $[0,1]$*

$$m_{n+1}(x) = \sum_{k=1}^{\infty} \left(m_n\left(\frac{1}{k}\right) - m_n\left(\frac{1}{k+x}\right)\right). \tag{19}$$

De plus chaque fonction m_n est différentiable sur $[0,1]$, et

$$m'_{n+1}(x) = \sum_{k=1}^{\infty} \frac{1}{(k+x)^2} m'_n\left(\frac{1}{k+x}\right). \tag{20}$$

Preuve. Fixons x dans $[0,1]$. Désignons par E l'ensemble des nombres α de I tels que $0 < z_{n+1}(\alpha) \leq x$, et pour tout entier $k \geq 1$ par E_k l'ensemble des nombre α de I tels que

$$\frac{1}{k+x} \leq z_n(\alpha) < \frac{1}{k}. \tag{21}$$

En vertu du lemme 20, le type d'inégalités dans (21) n'importe pas dans la mesure des ensembles E_k. Ainsi, les mesures de ces ensembles sont exactement

$$\mathfrak{M}E = m_{n+1}(x) \quad \text{et} \quad \mathfrak{M}E_k = m_n\left(\frac{1}{k}\right) - m_n\left(\frac{1}{k+x}\right).$$

Il est clair que, pour α dans I,

$$z_n(\alpha) = \frac{1}{a_{n+1}(\alpha) + z_{n+1}(\alpha)}.$$

Ainsi, α est dans E si et seulement s'il existe un entier k tel que α est dans E_k, c'est-à-dire

$$E = \bigcup_{k=1}^{\infty} E_k.$$

Les ensembles E_k étant deux à deux disjoints, on en déduit

$$\mathfrak{M}E = \sum_{k=1}^{\infty} \mathfrak{M}E_k,$$

ce qui est bien sûr l'équation fonctionnelle (19).

En dérivant formellement terme à terme la série (19), on obtient

$$\sum_{k=1}^{\infty} \frac{1}{(k+x)^2} m_n' \left(\frac{1}{k+x} \right). \tag{22}$$

Montrons que m_n est différentiable sur $[0,1]$ et que la série (22) converge uniformément sur $[0,1]$, ce qui établira (en utilisant le théorème sur la différentiabilité terme à terme des séries) la différentiabilité de m_{n+1} et l'égalité (20). Nous raisonnons par récurrence. Pour $n = 0$, on a clairement $m_0(x) = x$ et ainsi $m_0'(x) = 1$. Dans ce cas, le terme général de la série (22) est borné sur $[0,1]$ par $1/k^2$. Puisque la série numérique

$$\sum_{k=1}^{\infty} \frac{1}{k^2}$$

converge, la convergence uniforme de la série de fonctions en découle. Supposons que la fonction m_n est différentiable sur $[0,1]$ et que sa dérivée est bornée par μ. Alors le terme général de la série (22) est borné sur $[0,1]$ par μ/k^2, et on en déduit de même la convergence uniforme de la série de fonctions. De plus, m_{n+1}' est nécessairement bornée sur $[0,1]$ en tant que limite uniforme de fonctions bornées sur cet intervalle. $\square$

Nous allons maintenant utiliser l'équation fonctionnelle (20) pour déduire le résultat de Gauss et de Kuz'min. Nous commençons par établir un résultat technique, dont la preuve nécessitera plusieurs lemmes.

Théorème 22 (Kuz'min, 1928). *Soit $(f_n)_{n=0}^{\infty}$ une suite de fonctions réelles définies sur l'intervalle $[0,1]$. Supposons que pour tout $n \geq 0$ et tout x dans $[0,1]$*

$$f_{n+1}(x) = \sum_{k=1}^{\infty} \frac{1}{(k+x)^2} f_n \left(\frac{1}{k+x} \right). \tag{23}$$

Si f_0 est différentiable sur $[0,1]$ et s'il existe des nombres réels M et μ tels que, pour tout x dans $[0,1]$,

$$0 < f_0(x) < M \quad \textit{et} \quad |f_0'(x)| < \mu,$$

alors il existe des nombres réels $\lambda > 0$, $A > 0$ et une suite de fonctions $(\vartheta_n)_{n=0}^{\infty}$ tels que

$$f_N(x) = \frac{a}{1+x} + \vartheta_N(x) A e^{-\lambda\sqrt{N}} \quad \textit{et} \quad |\vartheta_N(x)| < 1$$

pour tout entier $N \geq 0$ et tout x dans $[0,1]$, où

$$a = \frac{1}{\log 2} \int_0^1 f_0(y)\, dy.$$

Les énoncés qui suivent reprennent les hypothèses et les notations du théorème 22.

Lemme 23. *Pour tout entier naturel n et pour tout x dans $[0,1]$*

$$f_n(x) = \sum^{(n)} f_0 \left(\frac{p_n + x p_{n-1}}{q_n + x q_{n-1}} \right) \frac{1}{(q_n + x q_{n-1})^2}. \tag{24}$$

Preuve. La preuve se fait par récurrence. Si $n = 0$, le résultat est trivial car $p_0 = 0$, $q_0 = 1$, $p_{-1} = 1$, $q_{-1} = 0$ et la somme (24) est prise sur l'unique intervalle de rang 0. Supposons le résultat vérifié pour un certain $n \geq 0$. Alors

$$\begin{aligned} f_{n+1}(x) &= \sum_{k=1}^{\infty} \frac{1}{(k+x)^2} f_n\left(\frac{1}{k+x}\right) \\ &= \sum_{k=1}^{\infty} \frac{1}{(k+x)^2} \sum^{(n)} f_0\left(\frac{p_n + \frac{1}{k+x}p_{n-1}}{q_n + \frac{1}{k+x}q_{n-1}}\right) \frac{1}{(q_n + \frac{1}{k+x}q_{n-1})^2} \\ &= \sum^{(n)} \sum_{k=1}^{\infty} f_0\left(\frac{p_n k + p_{n-1} + x p_n}{q_n k + q_{n-1} + x q_n}\right) \frac{1}{(q_n k + q_{n-1} + x q_n)^2} \\ &= \sum^{(n+1)} f_0\left(\frac{p_{n+1} + x p_n}{q_{n+1} + x q_n}\right) \frac{1}{(q_{n+1} + x q_n)^2}. \end{aligned}$$

Le passage de la série à l'intérieur de la somme est justifié par le fait que la série converge absolument (car elle converge par hypothèse et tous ses termes sont positifs). La dernière égalité est une conséquence des formules (i) et (ii) de la proposition 1. $\square$

LEMME 24. *Pour tout entier naturel n, la fonction f_n est différentiable sur $[0,1]$, et pour tout x dans $[0,1]$ on a*

$$|f'_n(x)| < \frac{\mu}{2^{n-3}} + 4M.$$

Preuve. Le résultat est immédiat si $n = 0$. Supposons donc que n est supérieur ou égal à 1. Si l'on dérive la somme (24) terme à terme, on obtient

$$\sum^{(n)} \left(f'_0(u)\frac{(-1)^{n-1}}{(q_n + xq_{n-1})^4} - 2f_0(u)\frac{q_{n-1}}{(q_n + xq_{n-1})^3}\right) \quad \text{où} \quad u = \frac{p_n + xp_{n-1}}{q_n + xq_{n-1}}. \tag{25}$$

Nous prétendons que la somme (25) est uniforme sur $[0,1]$. Vu que

$$\left|f'_0(u)\frac{(-1)^{n-1}}{(q_n + xq_{n-1})^4} - 2f_0(u)\frac{q_{n-1}}{(q_n + xq_{n-1})^3}\right| < \mu\frac{1}{q_n^4} + 2M\frac{1}{q_n^2},$$

il suffit pour le montrer de vérifier que la série numérique multiple

$$\sum_{a_1=1}^{\infty} \cdots \sum_{a_n=1}^{\infty} \left(\mu\frac{1}{q_n^4} + 2M\frac{1}{q_n^2}\right)$$

converge. Nous allons prouver plus généralement que la série multiple

$$\sum_{a_1=1}^{\infty} \cdots \sum_{a_n=1}^{\infty} \frac{1}{q_n^\tau}$$

converge pour tout nombre réel $\tau > 1$. Comme $n \geq 1$, la proposition 1 affirme que $q_n = a_n q_{n-1} + q_{n-2}$. Ainsi

$$\sum_{a_n=1}^{\infty} \frac{1}{q_n^\tau} \leq \sum_{a_n=1}^{\infty} \frac{1}{a_n^\tau q_{n-1}^\tau} = \frac{1}{q_{n-1}^\tau} \sum_{a_n=1}^{\infty} \frac{1}{a_n^\tau} = \frac{S}{q_{n-1}^\tau} \quad \text{où} \quad S = \sum_{k=1}^{\infty} \frac{1}{k^\tau}.$$

En prenant successivement les séries sur a_{n-1}, …, a_1 et en observant que $q_1 = a_1$, on obtient finalement

$$\sum_{a_1=1}^{\infty} \cdots \sum_{a_n=1}^{\infty} \frac{1}{q_n^\tau} \leq S^n,$$

ce qui établit notre assertion.

Par conséquent, f_n est différentiable sur $[0,1]$ et la somme (25) vaut $f_n'(x)$ pour tout x dans $[0,1]$. Vu la proposition 13,

$$\frac{1}{(q_n + xq_{n-1})^2} < \frac{2}{q_n(q_n + q_{n-1})} < \frac{2}{q_n^2} < \frac{1}{2^{n-2}}$$

pour tout x dans $[0,1]$, et par la formule (iii) de la proposition 1,

$$\sum^{(n)} \frac{1}{q_n(q_n + q_{n-1})} = \sum^{(n)} \left| \frac{p_n}{q_n} - \frac{p_n + p_{n-1}}{q_n + q_{n-1}} \right| = \mathfrak{M}I = 1.$$

Par conséquent, pour tout x dans $[0,1]$,

$$\begin{aligned} |f_n'(x)| &\leq \sum^{(n)} |f_0'(u)| \frac{1}{2^{n-2}} \frac{2}{q_n(q_n + q_{n-1})} + 2\sum^{(n)} |f_0(u)| \frac{q_{n-1}}{q_n + xq_{n-1}} \frac{2}{q_n(q_n + q_{n-1})} \\ &< \frac{\mu}{2^{n-3}} \sum^{(n)} \frac{1}{q_n(q_n + q_{n-1})} + 4M \sum^{(n)} \frac{1}{q_n(q_n + q_{n-1})} = \frac{\mu}{2^{n-3}} + 4M, \end{aligned}$$

ce qui conclut la démonstration. □

Lemme 25. *S'il existe deux nombres réels t et T et un certain entier n tels que pour tout x dans* $[0,1]$

$$\frac{t}{1+x} < f_n(x) < \frac{T}{1+x},$$

alors pour pour tout x dans $[0,1]$

$$\frac{t}{1+x} < f_{n+1}(x) < \frac{T}{1+x}. \tag{26}$$

Preuve. Soit τ un nombre réel quelconque. Alors

$$\begin{aligned} \sum_{k=1}^{\infty} \frac{1}{(k+x)^2} \frac{\tau}{1 + \frac{1}{k+x}} &= \tau \sum_{k=1}^{\infty} \frac{1}{(k+x)(k+x+1)} \\ &= \tau \sum_{k=1}^{\infty} \left(\frac{1}{k+x} - \frac{1}{k+x+1} \right) = \frac{\tau}{1+x}. \end{aligned} \tag{27}$$

En utilisant l'équation fonctionnelle (23), on obtient les inégalités

$$\sum_{k=1}^{\infty} \frac{1}{(k+x)^2} \frac{t}{1 + \frac{1}{k+x}} < f_{n+1}(x) < \sum_{k=1}^{\infty} \frac{1}{(k+x)^2} \frac{T}{1 + \frac{1}{k+x}}, \tag{28}$$

ce qui, en vertu de (27), est équivalent à (26). □

Lemme 26. *Pour tout entier naturel n, la fonction f_n est intégrable sur $[0,1]$ et*

$$\int_0^1 f_n(y)\, dy = \int_0^1 f_0(y)\, dy.$$

Preuve. Nous procédons par récurrence, le cas $n = 0$ étant tautologique. Si $n \geq 1$, vu que tous les termes de la série (23) sont positifs, on obtient à l'aide du théorème de convergence monotone†

$$\int_0^1 f_n(y)\, dy = \sum_{k=1}^{\infty} \int_0^1 f_{n-1}\left(\frac{1}{k+y}\right) \frac{dy}{(k+y)^2} = \sum_{k=1}^{\infty} \int_{\frac{1}{k+1}}^{\frac{1}{k}} f_{n-1}(u)\, du = \int_0^1 f_{n-1}(u)\, du,$$

et le lemme est démontré. □

† [3] p. 112, Theorem B.

Nous démontrons maintenant le théorème de Kuz'min en quatre étapes.

ÉTAPE 1. *Pour tout entier $n \geq 1$, il existe des nombres g_1 et G_1 tels que pour tout x dans $[0,1]$*

$$\frac{g_1}{1+x} < f_n(x) < \frac{G_1}{1+x}$$

De plus, on a l'estimation

$$G_1 - g_1 < (G-g)\delta + 2^{-n+2}(\mu + G) \quad \textit{où} \quad \delta = 1 - \frac{\log 2}{2} < 1, \tag{29}$$

et si n est assez grand, $G_1 < G$.

La fonction f_0 étant différentiable sur $[0,1]$, elle est continue sur ce même intervalle. Elle admet ainsi un minimum $m > 0$ sur cet intervalle. Notons $g = m/2$ et $G = 2M$. Pour tout x dans $[0,1]$, on a les inégalités

$$\frac{g}{1+x} < f_0(x) < \frac{G}{1+x}.$$

Nous définissons sur $[0,1]$ une nouvelle suite de fonctions $(\varphi_n)_{n=0}^{\infty}$ par

$$\varphi_n(x) = f_n(x) - \frac{g}{1+x}.$$

Alors clairement φ_0 est strictement positive, bornée sur $[0,1]$ et sa dérivée est bornée sur $[0,1]$. En remarquant que

$$\frac{g}{1+x} = g\sum_{k=1}^{\infty}\left(\frac{1}{k+x} - \frac{1}{k+x+1}\right) = g\sum_{k=1}^{\infty}\frac{1}{(k+x)(k+x+1)} = \sum_{k=1}^{\infty}\frac{1}{(k+x)^2}\frac{g}{1+\frac{1}{k+x}},$$

on constate que la suite de fonctions $(\varphi_n)_{n=0}^{\infty}$ vérifie également les hypothèses du théorème 22. Nous pouvons donc lui appliquer le lemme 23, qui nous donne, pour tout $n \geq 0$ et tout x dans $[0,1]$,

$$\varphi_n(x) = \sum^{(n)} \varphi_0(u)\frac{1}{(q_n + xq_{n-1})^2} \quad \text{où} \quad u = \frac{p_n + xp_{n-1}}{q_n + xq_{n-1}}. \tag{30}$$

En utilisant l'inégalité $q_n + xq_{n-1} \leq q_n + q_{n-1} < 2q_n$ valable pour tout x dans $[0,1]$, la relation (30) devient

$$\varphi_n(x) > \frac{1}{2}\sum^{(n)} \varphi_0(u)\frac{1}{q_n(q_n+q_{n-1})}. \tag{31}$$

Considérons un intervalle J_n de rang n apparaissant dans cette somme, disons $J_n = I_n \setminus \mathbb{Q}$, dont les extrémités sont

$$\frac{p_n + p_{n-1}}{q_n + q_{n-1}} \quad \text{et} \quad \frac{p_n}{q_n}.$$

Vu la formule (iii) de la proposition 1, la mesure d'un tel intervalle est

$$\left|\frac{p_n + p_{n-1}}{q_n + q_{n-1}} - \frac{p_n}{q_n}\right| = \frac{1}{q_n(q_n + q_{n-1})},$$

et par le théorème de la moyenne il existe u' dans I_n tel que

$$\int_{I_n} \varphi_0(y)\, dy = \varphi_0(u')\frac{1}{q_n(q_n+q_{n-1})}.$$

Lorsque l'on somme ce résultat sur tous les intervalles de rang n, on obtient

$$\int_0^1 \varphi_0(y)\, dy = \sum^{(n)} \varphi_0(u')\frac{1}{q_n(q_n+q_{n-1})}, \tag{32}$$

où il est entendu que u' est un nombre différent dans chaque terme, celui obtenu en appliquant le théorème de la moyenne sur l'intervalle correspondant. Les relations (31) et (32) nous donnent

$$\varphi_n(x) - \frac{1}{2}\int_0^1 \varphi_0(y)\, dy > \frac{1}{2}\sum^{(n)}(\varphi_0(u) - \varphi_0(u'))\frac{1}{q_n(q_n + q_{n-1})}. \tag{33}$$

Nous appliquons maintenant le théorème des accroissements finis à chaque terme $\varphi_0(u) - \varphi_0(u')$. Étant donné que pour tout x dans $[0, 1]$

$$|\varphi_0'(x)| = \left| f_0'(x) + \frac{g}{(1+x)^2}\right| < \mu + g,$$

on déduit du théorème des accroissements finis et de la proposition 13 que

$$|\varphi_0(u) - \varphi_0(u')| < (\mu + g)|u - u'| \leq (\mu + g)\frac{1}{q_n(q_n + q_{n-1})} < \frac{\mu + g}{q_n^2} < \frac{\mu + g}{2^{n-1}}.$$

En y injectant ce résultat, le relation (33) s'écrit

$$\begin{aligned}\varphi_n(x) &> \frac{1}{2}\int_0^1 \varphi_0(y)\, dy + \frac{1}{2}\sum^{(n)}\left(-\frac{\mu + g}{2^{n-1}}\right)\frac{1}{q_n(q_n + q_{n-1})} \\ &= \frac{1}{2}\int_0^1 \varphi_0(y)\, dy - \frac{\mu + g}{2^n}\sum^{(n)}\frac{1}{q_n(q_n + q_{n-1})} = \frac{1}{2}\int_0^1 \varphi_0(y)\, dy - \frac{\mu + g}{2^n},\end{aligned}$$

où la dernière égalité découle du fait que les intervalles de rang n (dont on somme ici les mesures) partitionnent l'intervalle irrationnel I dont la mesure est 1. Finalement, on obtient pour tout x dans $[0, 1]$

$$f_n(x) > \frac{g}{1+x} + l - \frac{\mu + g}{2^n} \geq \frac{g + l - 2^{-n+1}(\mu + g)}{1+x} = \frac{g_1}{1+x} \quad \text{où} \quad l = \frac{1}{2}\int_0^1 \varphi_0(y)\, dy.$$

En répétant tout ce qui a été fait précédemment pour la suite de fonction $(\psi_n)_{n=0}^{\infty}$ définie sur $[0, 1]$ par

$$\psi_n(x) = \frac{G}{1+x} - f_n(x),$$

nous trouvons similairement l'inégalité

$$f_n(x) < \frac{G - l' + 2^{-n+1}(\mu + G)}{1+x} = \frac{G_1}{1+x} \quad \text{où} \quad l' = \frac{1}{2}\int_0^1 \psi_0(y)\, dy.$$

Puisque les fonctions φ_0 et ψ_0 sont strictement positives sur l'intervalle $[0, 1]$, les intégrales l et l' sont également strictement positives. Si l'on choisit n suffisamment grand, disons si $n \geq \nu$, on a alors l'inégalité $G_1 < G$. Des relations

$$\begin{aligned}&G_1 - g_1 < G - g - (l + l') + 2^{-n+2}(\mu + G) \quad \text{et} \\ &\qquad l + l' = \frac{1}{2}\int_0^1 \left(f_0(y) - \frac{g}{1+y} + \frac{G}{1+y} - f_0(y)\right) dy = \frac{1}{2}\int_0^1 \frac{G - g}{1+y}dy = (G - g)\frac{\log 2}{2},\end{aligned}$$

on obtient, pour tout $n \geq 0$,

$$G_1 - g_1 < (G - g)\delta + 2^{-n+2}(\mu + G) \quad \text{où} \quad \delta = 1 - \frac{\log 2}{2} < 1,$$

ce qui complète la première étape.

Étape 2. *Fixons un entier $n \geq 1$. Définissons*

$$\mu_1 = \frac{\mu}{2^{n-3}} + 4M,$$

$$l_1 = l - g_1 \frac{\log 2}{2}, \qquad l_1' = G_1 \frac{\log 2}{2} - l',$$

$$g_2 = g_1 + l_1 - 2^{-n+1}(\mu_1 + g_1), \qquad G_2 = G_1 - l_1' + 2^{-n+1}(\mu_1 + G_1).$$

Alors pour tout x dans $[0, 1]$

$$\frac{g_2}{1+x} < f_{2n}(x) < \frac{G_2}{1+x}.$$

De plus, on a l'estimation

$$G_2 - g_2 < (G_1 - g_1)\delta + 2^{-n+2}(\mu_1 + G_1)$$

et si n est assez grand, $G_2 < G$.

Considérons la suite de fonctions $(h_m)_{m=0}^{\infty}$ définie par $h_m = f_{n+m}$. Pour cette nouvelle suite de fonctions, on a montré dans l'étape 1 que

$$\frac{g_1}{1+x} < h_0(x) < \frac{G_1}{1+x},$$

et par le lemme 24

$$|h_0'(x)| < \mu_1.$$

Ainsi la suite de fonctions $(h_m)_{m=0}^{\infty}$ vérifie les hypothèses du théorème. Observons que, en vertu du lemme 26,

$$\frac{1}{2}\int_0^1 \left(h_0(y) - \frac{g_1}{1+y}\right) dy = l_1 \quad \text{et} \quad \frac{1}{2}\int_0^1 \left(\frac{G_1}{1+y} - h_0(y)\right) dy = l_1'.$$

En répétant le raisonnement de l'étape 1, on obtient pour tout x dans $[0, 1]$ et pour tout $m \geq 1$

$$\frac{g_2}{1+x} < h_m(x) < \frac{G_2}{1+x} \quad \text{et} \quad G_2 - g_2 < (G_1 - g_1)\delta + 2^{-m+2}(\mu_1 + G_1)\,;$$

c'est vrai en particulier pour $m = n$, ce qui est le résultat cherché. Il reste à montrer que, pour n assez grand, G_2 est strictement inférieur à G. Développons l'expression de G_2 :

$$\begin{aligned} G_2 &= G_1 - l_1' + 2^{-n+1}(\mu_1 + G_1) \\ &= G - l' + 2^{-n+1}(\mu + G) - l_1' + 2^{-n+1}(\mu_1 + G - l' + 2^{-n+1}(\mu + G)). \end{aligned}$$

Comme l' et l_1' sont strictement positifs, on aura (en majorant $-l_1'$ par 0), pour autant que n soit assez grand (disons si $n \geq \nu'$), $G_2 < G$.

Étape 3. *Pour tout entier $i \geq 2$, définissons par récurrence*

$$\begin{aligned} \mu_i &= \frac{\mu}{2^{in-3}} + 4M, \\ l_i &= \frac{1}{2}\int_0^1 \left(f_{in}(y) - \frac{g_i}{1+y}\right) dy = l - g_i \frac{\log 2}{2}, \\ l_i' &= \frac{1}{2}\int_0^1 \left(\frac{G_i}{1+y} - f_{in}(y)\right) dy = G_i \frac{\log 2}{2} - l', \end{aligned}$$

$$g_{i+1} = g_i + l_i - 2^{-n+1}(\mu_i + g_i), \qquad G_{i+1} = G_i - l_i' + 2^{-n+1}(\mu_i + G_i).$$

Alors pour tout entier $r \geq 1$ et tout x dans $[0, 1]$

$$\frac{g_r}{1+x} < f_{rn}(x) < \frac{G_r}{1+x}. \tag{34}$$

De plus, on a l'estimation

$$G_r - g_r < (G_{r-1} - g_{r-1})\delta + 2^{-n+2}(\mu_{r-1} + G_{r-1}) \tag{35}$$

et si n est supérieur ou égal à $\max\{\nu, \nu'\}$, $G_r < G$.

Le résultat a déjà été démontré pour $r = 1$ et $r = 2$. Nous prouvons le cas général par récurrence. Supposons le résultat vérifié pour un certain entier $r-1$, avec $r \geq 3$. Puisque

$$\frac{g_{r-1}}{1+x} < f_{(r-1)n}(x) < \frac{G_{r-1}}{1+x} \quad \text{et} \quad |f'_{(r-1)n}(x)| < \mu_{r-1},$$

la suite de fonction $(f_{(r-1)n+m})_{m=0}^{\infty}$ vérifie les hypothèses du théorème, et les relations (34) et (35) en découlent par l'étape 1 comme dans l'étape 2.

Nous montrons maintenant que si $n \geq \max\{\nu, \nu'\}$, alors $G_r < G_{r-1}$, ce qui prouvera, vu que $G_2 < G$, la dernière assertion. Par hypothèse de récurrence, on a $G_{r-1} < G_{r-2}$. Ainsi,

$$\begin{aligned} G_r - G_{r-1} &= -l'_{r-1} + 2^{-n+1}(\mu_{r-1} + G_{r-1}) < -l'_{r-1} + 2^{-n+1}(\mu_{r-2} + G_{r-2}) \\ &= G_{r-1} - G_{r-2} + l'_{r-2} - l'_{r-1} = G_{r-1} - G_{r-2} + \frac{\log 2}{2}(G_{r-2} - G_{r-1}) \\ &\qquad < G_{r-1} - G_{r-2} + G_{r-2} - G_{r-1} = 0, \end{aligned}$$

ce qui conclut l'étape 3.[†]

Étape 4. *Conclusion du théorème.*

Résumons ce que nous avons démontré jusqu'à présent. Pour tout $n \geq \max\{\nu, \nu'\}$, nous avons construit des suites $(g_r(n))_{r=1}^{\infty}$ et $(G_r(n))_{r=1}^{\infty}$ (nous mettons maintenant en évidence la dépendence en n) avec les propriétés suivantes. Pour tout entier $r \geq 1$ et tout x dans $[0, 1]$

$$\frac{g_r(n)}{1+x} < f_{rn}(x) < \frac{G_r(n)}{1+x}, \tag{36}$$

$$G_{r+1}(n) - g_{r+1}(n) < (G_r(n) - g_r(n))\delta + 2^{-n+2}(\mu_r(n) + G_r(n)) \quad \text{et} \tag{37}$$

$$G_r(n) < G. \tag{38}$$

Pour un $n \geq \max\{\nu, \nu'\}$ donné, le résultat est en particulier vrai pour $r = n$, auquel cas les relation (36) et (37) deviennent

$$\frac{g_n(n)}{1+x} < f_{n^2}(x) < \frac{G_n(n)}{1+x} \quad \text{et} \tag{39}$$

$$G_n(n) - g_n(n) < (G_{n-1}(n) - g_{n-1}(n))\delta + 2^{-n+2}(\mu_{n-1}(n) + G_{n-1}(n)). \tag{40}$$

En appliquant récursivement l'inégalité (37) à partir de l'inégalité (40), et en utilisant finalement l'inégalité (29), on obtient

$$G_n(n) - g_n(n) < (G - g)\delta^n + 2^{-n+2} \sum_{i=0}^{n-2} \delta^i(\mu_{n-1-i}(n) + G_{n-1-i}(n)) + 2^{-n+2}\delta^{n-1}(\mu + G). \tag{41}$$

Mais

$$\mu_r(n) = \frac{\mu}{2^{rn-3}} + 4M,$$

et si n est choisi suffisamment grand, à savoir si

$$n \geq \nu'' = \frac{\log \mu - \log M}{\log 2} + 4,$$

on aura $\mu_r(n) < 5M$ pour tout $r \geq 1$. Rappelons que $G = 2M$ et $0 < \delta = 1 - (\log 2)/2 < 1$, et par conséquent $\delta^r \leq \delta$ si r est supérieur ou égal à 1. Ainsi

$$\begin{aligned} G_n(n) - g_n(n) &< (G - g)\delta^n + 2^{-n+2} \sum_{i=0}^{n-2} 7M\delta^i + 2^{-n+2}\delta(\mu + 2M) \\ &= (G - g)\delta^n + 2^{-n+2} 7M \left(\frac{1 - \delta^{n-2}}{1 - \delta}\right) + 2^{-n+2}\delta(\mu + 2M), \end{aligned}$$

[†]Le lecteur remarquera la nécessité de traiter séparément les cas $r = 1$ et $r = 2$. En effet, si l'on voulait inclure l'étape 2 dans l'étape 3, on aurait besoin de l'inégalité $\mu_1 < \mu$ qui n'est pas nécessairement vérifiée.

où l'on a utilisé la formule pour la somme d'une progression géométrique. Si l'on note $K = 7M/(1-\delta) + \delta(\mu + 2M)$, $L = G > (G - g)$, $Z = \min\{1/\delta, 2\} = 1/\delta > 1$, $\lambda = \log(Z) > 0$ et $B = e^{2\lambda}(L + K)$, on trouve successivement

$$G_n(n) - g_n(n) < L\left(\frac{1}{\delta}\right)^{-n} + 2^{-n+2}K < L\left(\frac{1}{\delta}\right)^{-n+2} + 2^{-n+2}K < (L+K)Z^{-n+2} = Be^{-\lambda n},$$

ce résultat étant vérifié dès que $n \geq \max\{\nu, \nu', \nu''\}$.

Nous pouvons maintenant conclure la démonstration. L'inégalité

$$G_n(n) - g_n(n) < Be^{-\lambda n} \tag{42}$$

implique dans un premier temps que les suites $(g_n(n))_{n=1}^{\infty}$ et $(G_n(n))_{n=1}^{\infty}$ sont convergentes et ont une même limite a :

$$\lim_{n\to\infty} g_n(n) = \lim_{n\to\infty} G_n(n) = a.$$

D'autre part, pour tout x dans $[0, 1]$,

$$\left| f_{n^2}(x) - \frac{a}{1+x} \right| < \frac{G_n(n)}{1+x} - \frac{g_n(n)}{1+x} \leq G_n(n) - g_n(n) < Be^{-\lambda n}, \tag{43}$$

ce qui implique, en prenant l'intégrale sur $[0, 1]$,

$$a\log 2 - Be^{-\lambda n} < \int_0^1 f_{n^2}(y)\,dy < a\log 2 + Be^{-\lambda n}.$$

Ainsi, par le lemme 26, le nombre a est bien celui de l'énoncé du théorème, à savoir

$$a = \frac{1}{\log 2}\int_0^1 f_0(y)\,dy.$$

Soit N un entier positif. Supposons que $n^2 \leq N < (n+1)^2$. L'inégalité (43) implique, pour tout x dans $[0, 1]$,

$$\frac{a - 2Be^{-\lambda n}}{1+x} < f_{n^2}(x) < \frac{a + 2Be^{-\lambda n}}{1+x},$$

et ainsi, en vertu du lemme 25,

$$\frac{a - 2Be^{-\lambda n}}{1+x} < f_N(x) < \frac{a + 2Be^{-\lambda n}}{1+x}.$$

En conclusion, on trouve pour tout entier $N \geq \max\{\nu, \nu', \nu''\}^2$

$$\left| f_N(x) - \frac{a}{1+x} \right| < 2Be^{-\lambda n} \leq Ae^{-\lambda(n+1)} < Ae^{-\lambda\sqrt{N}},$$

si A est supérieur à $2Be^{\lambda}$. Le résultat sera en outre vrai pour tout entier $N \geq 0$ si A est choisi suffisamment grand pour que, pour tout $r < \max\{\nu, \nu', \nu''\}^2$, on ait

$$\left| f_r(x) - \frac{a}{1+x} \right| < Ae^{-\lambda\sqrt{r}}.$$

Plus précisément, on peut par exemple choisir A égal à

$$\max\left\{2Be^{\lambda}, \left(MS^{\nu_0} + \frac{M}{\log 2}\right)e^{\lambda\sqrt{\nu_0}}\right\}, \quad \text{où } S = \sum_{k=1}^{\infty}\frac{1}{k^2} \text{ et } \nu_0 = \max\{\nu, \nu', \nu''\}^2. \tag{44}$$

Le théorème est démontré. □

Il est très important de noter que le nombre

$$\lambda = \log\left(\frac{2}{2 - \log 2}\right)$$

est une constante absolue et que le nombre A est entièrement déterminé par M et μ. Ceci découle de (44) et du fait aisément vérifié que B, ν, ν' et ν'' ne dépendent que de M et μ.

La suite de fonctions $(m_n')_{n=0}^{\infty}$ vérifie les hypothèses du théorème 22 car $m_0'(x) = 1$ et $m_0''(x) = 0$ pour tout x dans $[0,1]$. La conclusion du théorème est

$$m_n'(x) = \frac{1}{(1+x)\log 2} + \vartheta_n(x) A e^{-\lambda\sqrt{n}} \tag{45}$$

pour certaines constantes $A > 0$ et $\lambda > 0$ et avec $|\vartheta_n(x)| < 1$. Si l'on intègre ce résultat sur $[0,1]$, on obtient

$$\left| m_n(x) - \frac{\log(1+x)}{\log 2} \right| < A e^{-\lambda\sqrt{n}},$$

ce qui non seulement prouve l'assertion de Gauss mais nous donne une information sur la vitesse de convergence. Avec les notations de Landau, le résultat s'écrit

$$m_n(x) = \frac{\log(1+x)}{\log 2} + O(e^{-\lambda\sqrt{n}}), \quad n \to \infty.^{\dagger}$$

Nous utilisons maintenant ce résultat pour estimer la mesure de l'ensemble des nombres α dans I pour lesquels $a_n(\alpha) = k$. Il est clair que α appartient à cet ensemble si et seulement si

$$\frac{1}{k+1} < z_{n-1}(\alpha) < \frac{1}{k}, \tag{46}$$

et puisque la variable aléatoire z_{n-1} est continue, il s'ensuit que

$$\mathfrak{M}E\binom{n}{k} = m_{n-1}\left(\frac{1}{k}\right) - m_{n-1}\left(\frac{1}{k+1}\right) = \int_{\frac{1}{k+1}}^{\frac{1}{k}} m_{n-1}'(y)\,dy.$$

Ainsi, en intégrant (45) entre $1/(k+1)$ et $1/k$, on obtient

$$\left| \mathfrak{M}E\binom{n}{k} - \frac{\log\left(1 + \frac{1}{k(k+2)}\right)}{\log 2} \right| < \frac{A}{k(k+1)} e^{-\lambda\sqrt{n-1}}, \tag{47}$$

ce qui implique en particulier que

$$\mathfrak{M}E\binom{n}{k} \to \frac{\log\left(1 + \frac{1}{k(k+2)}\right)}{\log 2} \tag{48}$$

lorsque $n \to \infty$.

On peut reformuler ce résultat en adoptant le point de vue de la probabilité. Nous avons déjà vu que le triple $(I, \mathscr{B}', \mathfrak{M}|_{\mathscr{B}'})$, où $\mathscr{B}'$ est la σ-algèbre des boréliens de I, est un espace de probabilité sur lequel les fonctions a_i sont des variables aléatoires. La limite (48) signifie que la probabilité de l'événement $\{a_n = k\}$ converge vers

$$\frac{\log\left(1 + \frac{1}{k(k+2)}\right)}{\log 2}$$

lorsque n tend vers l'infini.

La fonction $g \in [0,1]^{\mathbb{R}}$ définie par

$$g(x) = \frac{\log\left(1 + \frac{1}{x(x+2)}\right)}{\log 2}$$

† Paul Lévy [12] utilise une méthode différente pour obtenir l'approximation

$$m_n(x) = \frac{\log(1+x)}{\log 2} + O(e^{-\lambda' n}), \quad n \to \infty,$$

pour une certaine constante $\lambda' > 0$ avec $\lambda < \lambda' < 1$.

si x est un entier strictement positif et $g(x) = 0$ sinon est une densité de probabilité discrète : c'est la loi limite de la suite de variables aléatoires $(a_i)_{i=1}^{\infty}$. On appelle la fonction g la *densité de Gauss–Kuz'min.*

Si l'on calcule numériquement cette densité, on trouve approximativement $g(1) \approx 0.415$, $g(2) \approx 0.170$, $g(3) \approx 0.093$, etc. Ce résultat semble ainsi apporter une explication au fait expérimental que la plupart des éléments d'une fraction continue représentant un nombre réel choisi « au hasard » sont égaux à 1, et que la proportion des éléments égaux à n diminue rapidement lorsque n croît. Nous trouverons par la suite une interprétation plus explicite de cette densité.

On peut généraliser ce résultat. Si des entiers strictement positifs k et $l_1, \ldots, l_k$ sont fixés, on considère la suite de fonctions $(M_n)_{n=0}^{\infty}$ définie sur $[0, 1]$ par

$$M_n(x) = \mathfrak{M}\{\alpha \in I \mid a_1(\alpha) = l_1, \ldots, a_k(\alpha) = l_k, z_{k+n}(\alpha) \leq x\}.$$

En d'autres termes, $M_n(x)$ est la mesure de l'ensemble des nombres α de I appartenant à un certain intervalle de rang k fixé et tels que $z_{k+n}(\alpha) \leq x$, ou de manière équivalente tels que

$$\frac{1}{a_n(\alpha) + x} \leq z_{k+n-1}(\alpha) < \frac{1}{a_n(\alpha)}.$$

On obtient ainsi une équation fonctionnelle identique à celle de la proposition 21, à savoir

$$M_n(x) = \sum_{r=1}^{\infty} \left(M_{n-1}\left(\frac{1}{r}\right) - M_{n-1}\left(\frac{1}{r+x}\right) \right) \tag{49}$$

pour tout entier $n \geq 1$ et tout x dans $[0, 1]$, et une preuve identique à celle de la proposition 21 montre que chaque M_n est différentiable sur $[0, 1]$ et que la suite $(M'_n)_{n=0}^{\infty}$ vérifie l'équation fonctionnelle (23) du théorème 22. La suite de fonctions $(\chi_n)_{n=0}^{\infty}$ définie sur $[0, 1]$ par

$$M_n(x) = \mathfrak{M}E \begin{pmatrix} 1 & \ldots & k \\ l_1 & \ldots & l_k \end{pmatrix} \chi_n(x)$$

différant de $(M_n)_{n=0}^{\infty}$ par un facteur constant, satisfait évidemment aussi l'équation fonctionnelle (49), et par conséquent la suite des dérivées (définies sur $[0, 1]$) satisfait l'équation (23). Nous allons montrer que la suite $(\chi'_n)_{n=0}^{\infty}$ satisfait en outre les autres hypothèses du théorème 22. Notons p_k et q_k les coefficients de convergence de la fraction continue $[l_1, \ldots, l_k]$. Nous savons que les nombres α appartenant à l'intervalle de rang k considéré sont ceux de la forme

$$\alpha = \frac{p_k r_{k+1} + p_{k-1}}{q_k r_{k+1} + q_{k-1}}$$

pour un certain r_{k+1} dans $]1, \infty[\setminus \mathbb{Q}$, ou de manière équivalente, puisque $z_k = 1/r_{k+1}$, les nombres de la forme

$$\alpha = \frac{p_k + z_k p_{k-1}}{q_k + z_k q_{k-1}} \tag{50}$$

pour un certain z_k appartenant à I. La fonction $z_k \mapsto \alpha$ définie par (50) étant monotone, il s'ensuit que l'ensemble $\{z_k \leq x\}$ est l'intervalle irrationnel défini par les extrémités

$$\frac{p_k}{q_k} \quad \text{et} \quad \frac{p_k + x p_{k-1}}{q_k + x q_{k-1}}.$$

Ainsi, d'après la formule (iii) de la proposition 1,

$$M_0(x) = \left| \frac{p_k}{q_k} - \frac{p_k + x p_{k-1}}{q_k + x q_{k-1}} \right| = \frac{x}{q_k(q_k + x q_{k-1})}.$$

D'autre part, la mesure de l'intervalle de rang k considéré est, comme nous l'avons déjà vu à plusieurs reprises,

$$\mathfrak{M}E \begin{pmatrix} 1 & \ldots & k \\ l_1 & \ldots & l_k \end{pmatrix} = \left| \frac{p_k}{q_k} - \frac{p_k + p_{k-1}}{q_k + q_{k-1}} \right| = \frac{1}{q_k(q_k + q_{k-1})},$$

et par conséquent, pour tout x dans $[0,1]$,

$$\chi_0(x) = \frac{(q_k + q_{k-1})x}{q_k + xq_{k-1}}, \quad \chi_0'(x) = \frac{q_k(q_k + q_{k-1})}{(q_k + xq_{k-1})^2} \quad \text{et} \quad \chi_0''(x) = -\frac{2q_kq_{k-1}(q_k + q_{k-1})}{(q_k + xq_{k-1})^3}.$$

De ces formules on conclut que $1/2 < \chi_0'(x) < 2$ et $|\chi_0''(x)| < 4$ pour tout x dans $[0,1]$, de sorte que, en posant $M = 2$ et $\mu = 4$, toutes les hypothèses du théorème 22 sont vérifiées pour la suite $(\chi_n')_{n=0}^{\infty}$. La conclusion du théorème est qu'il existe des nombres réels $A' > 0$, $\lambda > 0$ et une suite de fonction $(\vartheta_n)_{n=0}^{\infty}$ tels que, pour tout entier naturel n et tout x dans $[0,1]$,

$$\chi_n'(x) = \frac{M_n'(x)}{\mathfrak{M}E\begin{pmatrix}1 & \dots & k\\ l_1 & \dots & l_k\end{pmatrix}} = \frac{1}{(1+x)\log 2} + \vartheta_n(x)A'e^{-\lambda\sqrt{n}} \quad \text{et} \quad |\vartheta_n(x)| < 1, \tag{51}$$

où l'on a calculé

$$\int_0^1 \chi_0'(y)\,dy = \int_0^1 \frac{(q_k + q_{k-1})y}{q_k + yq_{k-1}}\,dy = 1.$$

Ici, les nombres A' et λ sont tous deux des constantes absolues, λ étant donné par le théorème 22 et A' ne dépendant que des constantes absolues 2 et 4.

Si $r \geq 1$ est un entier et si l'on intègre la relation (51) entre $1/(r+1)$ et $1/r$, on obtient

$$M_n\left(\frac{1}{r}\right) - M_n\left(\frac{1}{r+1}\right) = \left(\frac{\log\left(1 + \frac{1}{r(r+2)}\right)}{\log 2} + \eta_n\frac{A'}{r(r+1)}e^{-\lambda\sqrt{n}}\right)\mathfrak{M}E\begin{pmatrix}1 & \dots & k\\ l_1 & \dots & l_k\end{pmatrix} \tag{52}$$

avec $|\eta_n| < 1$. Il est clair d'après la relation (46) que

$$M_n\left(\frac{1}{r}\right) - M_n\left(\frac{1}{r+1}\right) = \mathfrak{M}E\begin{pmatrix}1 & \dots & k & k+n+1\\ l_1 & \dots & l_k & r\end{pmatrix},$$

et l'égalité (52) devient

$$\mathfrak{M}E\begin{pmatrix}1 & \dots & k & k+n+1\\ l_1 & \dots & l_k & r\end{pmatrix} = \left(\frac{\log\left(1 + \frac{1}{r(r+2)}\right)}{\log 2} + \eta_n\frac{A'}{r(r+1)}e^{-\lambda\sqrt{n}}\right)\mathfrak{M}E\begin{pmatrix}1 & \dots & k\\ l_1 & \dots & l_k\end{pmatrix}.$$

Si l'on somme cette égalité pour certains entiers parmi $l_1, l_2, \dots, l_k$ de 1 à l'infini, on fait disparaître ces éléments des deux côtés de l'égalité.

Nous avons donc montré le théorème suivant :

Théorème 27. *Il existe des nombres réels $A' > 0$ et $\lambda > 0$ tels que, quels que soient les entiers strictement positifs t, n_1, …, n_t, n_{t+1}, k_1, …, k_t, k avec $n_1 < \dots < n_t < n_{t+1}$,*

$$\left|\frac{\mathfrak{M}E\begin{pmatrix}n_1 & \dots & n_t & n_{t+1}\\ k_1 & \dots & k_t & k\end{pmatrix}}{\mathfrak{M}E\begin{pmatrix}n_1 & \dots & n_t\\ k_1 & \dots & k_t\end{pmatrix}} - \frac{\log\left(1 + \dfrac{1}{k(k+2)}\right)}{\log 2}\right| < \frac{A'}{k(k+1)}e^{-\lambda\sqrt{n_{t+1}-n_t-1}}.$$

L'interprétation probabiliste du théorème 27 est la suivante. La probabilité conditionnelle de l'événement $\{a_{n_{t+1}} = k\}$ sachant l'événement $\{a_{n_1} = k_1, \dots, a_{n_t} = k_t\}$ converge vers la densité de Gauss–Kuz'min en k lorsque n_{t+1} tend vers l'infini.

Remarque importante. Dans la suite de ce chapitre, λ représentera toujours la constante absolue des théorèmes 22 et 27, A sera la constante absolue obtenue du théorème 22 pour la suite de fonctions $(m_n')_{n=0}^{\infty}$, et A' la constante absolue du théorème 27.

2.4 Moyennes

Lemme 28 (Procédé de sommation de Cesàro). *Soient s un nombre réel et $(s_n)_{n=1}^{\infty}$ une suite de nombres réels. Si* $\lim_{n\to\infty} s_n = s$, *alors la suite* $(S_n)_{n=1}^{\infty}$ *définie par*

$$S_n = \frac{1}{n}\sum_{k=1}^{n} s_n$$

converge vers s. D'autre part, si $\lim_{n\to\infty} S_n = s$, *alors pour tout entier $m \geq 1$*

$$\frac{1}{n}\sum_{k=m}^{n+m-1} s_n \to s$$

lorsque $n \to \infty$.

Preuve. En remarquant que

$$\frac{(s_1 - s) + \cdots + (s_n - s)}{n} = \frac{s_1 + \cdots + s_n}{n} - s$$

on peut supposer $s = 0$. Soit $\varepsilon > 0$ un nombre réel et N un entier tel que $|s_n| < \varepsilon$ si $n \geq N$. Alors si $n \geq N$

$$|S_n| = \left|\frac{s_1 + \cdots + s_N}{n} + \frac{s_{N+1} + \cdots + s_n}{n}\right| < \left|\frac{s_1 + \cdots + s_N}{n}\right| + \varepsilon,$$

et donc $-\varepsilon \leq \lim_{n\to\infty} S_n \leq \varepsilon$. Mais ε était choisi arbitrairement, d'où $\lim_{n\to\infty} S_n = 0$.

Pour prouver la deuxième assertion, il suffit de remarquer que la distance entre les sommes

$$\frac{1}{n}\sum_{k=1}^{n} s_n \quad \text{et} \quad \frac{1}{n}\sum_{k=m}^{n+m-1} s_n$$

est aussi petite que l'on veut si n est assez grand. □

Théorème 29 (Khinchin, 1935). *Soit une fonction $f \in [0,\infty[^{\mathbb{N}\setminus\{0\}}$. Supposons qu'il existe des nombres réels $C > 0$ et $\delta > 0$ tels que*

$$f(r) < Cr^{1/2-\delta}$$

pour tout entier $r \geq 1$. Alors pour presque tout α dans I,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}(f\circ a_k)(\alpha) = \sum_{r=1}^{\infty} f(r)\frac{\log\left(1 + \frac{1}{r(r+2)}\right)}{\log 2}.$$

Remarquons que la série ci-dessus est convergente car $f(r) \leq \sqrt{r}$ par hypothèse et $\log(1+x) \leq x$ pour tout $x > -1$.

Preuve. Puisque a_k est mesurable sur $(I, \mathscr{B}')$, la fonction $f_k = f\circ a_k$ l'est aussi (car elle ne prend qu'un nombre dénombrable de valeurs sur des boréliens). Nous pouvons donc définir, pour tous entiers k, i, et n strictement positifs et tout β dans I,[†]

$$u_k = \int_I f_k(\alpha)\,d\alpha, \quad b_k = \int_I \big(f_k(\alpha) - u_k\big)^2\,d\alpha,$$

$$g_{ik} = \int_I \big(f_i(\alpha) - u_i\big)\big(f_k(\alpha) - u_k\big)\,d\alpha \quad \text{et} \quad s_n(\beta) = \sum_{j=1}^{n}\big(f_j(\beta) - u_j\big).$$

[†]Les nombres u_k, b_k et g_{ik} sont les espérances des variables aléatoires f_k, $f_k - u_k$ et $(f_i - u_i)(f_k - u_k)$ respectivement.

De plus, ces intégrales sont finies car pour tout entier $k \geq 1$ la fonction f_k appartient à $\mathscr{L}^2(\mathscr{B}', \mathfrak{M}|_{\mathscr{B}'}, \mathbb{R})$. En effet, vu que les fonctions f_k ne prennent qu'un nombre dénombrable de valeurs sur des boréliens et d'après le théorème 19,

$$\int_I f_k^2(\alpha)\, d\alpha = \sum_{s=1}^{\infty} f^2(s) \mathfrak{M}E \binom{k}{s} < 2C^2 \sum_{s=1}^{\infty} \frac{s^{1-2\delta}}{s^2} = C'$$

(la série étant convergente). En particulier, on obtient de l'inégalité de Cauchy-Schwarz

$$b_k = \int_I f_k^2(\alpha)\, d\alpha - 2u_k \int_I f_k(\alpha)\, d\alpha + u_k^2 = \int_I f_k^2(\alpha)\, d\alpha - u_k^2 < C' \quad \text{et} \quad u_k < \sqrt{C'}. \tag{53}$$

Notons également que nous avons les égalités

$$g_{ik} = \int_I f_i(\alpha) f_k(\alpha)\, d\alpha - u_i u_k = \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} f(r) f(s) \mathfrak{M}E \begin{pmatrix} i & k \\ r & s \end{pmatrix} - u_i u_k \tag{54}$$

$$\text{et} \quad u_i u_k = \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} f(r) f(s) \mathfrak{M}E \binom{i}{r} \mathfrak{M}E \binom{k}{s}. \tag{55}$$

En utilisant les théorèmes 27 et 19, on obtient, pour $k > i$, d'une part

$$\begin{aligned} \left| \mathfrak{M}E \begin{pmatrix} i & k \\ r & s \end{pmatrix} - \frac{\log\left(1 + \frac{1}{s(s+2)}\right)}{\log 2} \mathfrak{M}E \binom{i}{r} \right| &< \frac{A' e^{-\lambda\sqrt{k-i-1}}}{s(s+1)} \mathfrak{M}E \binom{i}{r} \\ &< 3A' e^{-\lambda\sqrt{k-i-1}} \mathfrak{M}E \binom{i}{r} \mathfrak{M}E \binom{k}{s} \end{aligned} \tag{56}$$

et d'autre part

$$\left| \mathfrak{M}E \binom{k}{s} - \frac{\log\left(1 + \frac{1}{s(s+2)}\right)}{\log 2} \right| < \frac{A e^{-\lambda\sqrt{k-1}}}{s(s+1)} < 3A e^{-\lambda\sqrt{k-1}} \mathfrak{M}E \binom{k}{s}. \tag{57}$$

En multipliant l'inégalité (57) par le nombre positif $\mathfrak{M}E\binom{i}{r}$, on obtient avec (56) en utilisant l'inégalité du triangle

$$\left| \mathfrak{M}E \begin{pmatrix} i & k \\ r & s \end{pmatrix} - \mathfrak{M}E \binom{i}{r} \mathfrak{M}E \binom{k}{s} \right| < 6B e^{-\lambda\sqrt{k-i-1}} \mathfrak{M}E \binom{i}{r} \mathfrak{M}E \binom{k}{s}, \tag{58}$$

où $B = \max\{A, A'\}$. En sommant pour tous r et s l'inégalité (58) multipliée par $f(r)f(s)$ et en utilisant la relation (54) on trouve

$$\begin{aligned} &\left| g_{ik} + u_i u_k - \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} f(r) f(s) \mathfrak{M}E \binom{i}{r} \mathfrak{M}E \binom{k}{s} \right| \\ &\qquad < 6B e^{-\lambda\sqrt{k-i-1}} \sum_{r=1}^{\infty} \sum_{s=1}^{\infty} f(r) f(s) \mathfrak{M}E \binom{i}{r} \mathfrak{M}E \binom{k}{s}, \end{aligned}$$

et finalement grâce à la relation (55),

$$|g_{ik}| < 6B e^{-\lambda\sqrt{k-i-1}} u_i u_k < 6C' B e^{-\lambda\sqrt{k-i-1}}.$$

Si on combine ce résultat avec la relation (53), on trouve pour tous entiers (strictement positifs) n

et m

$$\int_I \big(s_n(\alpha) - s_m(\alpha)\big)^2 \, d\alpha = \int_I \left(\sum_{k=m+1}^{n} (f_k(\alpha) - u_k) \right)^2 d\alpha$$
$$= \sum_{k=m+1}^{n} \int_I (f_k(\alpha) - u_k)^2 \, d\alpha + 2 \sum_{i=m+1}^{n} \sum_{k=i+1}^{n} \int_I (f_i(\alpha) - u_i)(f_k(\alpha) - u_k) \, d\alpha$$
$$= \sum_{k=m+1}^{n} b_k + 2 \sum_{i=m+1}^{n} \sum_{k=i+1}^{n} g_{ik} < C'(n-m) + 12BC \sum_{i=m+1}^{n} \sum_{k=i+1}^{\infty} e^{-\lambda\sqrt{k-i-1}}$$
$$< C''(n-m) \qquad (59)$$

(la série étant convergente).† Une démarche identique donne également

$$\int_I s_n^2(\alpha) \, d\alpha < C''n. \qquad (60)$$

Soit $\varepsilon > 0$ un nombre réel. Considérons les ensembles E_n ($n \geq 1$) contenant les nombres α de I pour lesquels $|s_n(\alpha)| \geq \varepsilon n$. Puisque s_n est mesurable sur $(I, \mathscr{B}')$ (car c'est une somme finie de fonctions mesurables), E_n appartient à $\mathscr{B}'$ en tant que préimage par s_n d'un borélien. Vu que $E_n \subseteq I$,

$$\int_I s_n^2(\alpha) \, d\alpha \geq \int_{E_n} s_n^2(\alpha) \, d\alpha \geq \varepsilon^2 n^2 \mathfrak{M} E_n,$$

et par l'inégalité (60)

$$\mathfrak{M} E_n \leq \frac{1}{\varepsilon^2 n^2} \int_I s_n^2(\alpha) \, d\alpha < \frac{C''}{\varepsilon^2 n}.$$

Par conséquent, la série $\sum_{n=1}^{\infty} \mathfrak{M} E_{n^2}$ converge dans $[0, \infty[$, et par la proposition 14, presque tout nombre de I n'appartient qu'à un nombre fini d'ensembles E_{n^2}. Ainsi, pour presque tout α dans I il existe un entier n_0 tel que, si $n \geq n_0$, $s_{n^2}(\alpha)/n^2 < \varepsilon$. Puisque ε est arbitraire on obtient que $s_{n^2}(\alpha)/n^2 \to 0$ presque partout lorsque $n \to \infty$. Soit N un entier tel que $n^2 \leq N < (n+1)^2$. Par la formule (59),

$$\int_I \big(s_N(\alpha) - s_{n^2}(\alpha)\big) \, d\alpha < C''(N - n^2) < C''(2n+1) \leq 3C''n.$$

Soit $\zeta > 0$ un nombre réel. Considérons maintenant, pour tous les entiers n et N avec $n \geq 1$ et $n^2 \leq N < (n+1)^2$, les ensemble E_n^N contenant les nombres α de I pour lesquels $|s_N(\alpha) - s_{n^2}(\alpha)| \geq \zeta n^2$, et définissons

$$F_n = \bigcup_{N=n^2}^{(n+1)^2-1} E_n^N.$$

Tous les ensembles E_n^N et F_n sont des boréliens, les premiers en tant que préimages de boréliens par une fonction mesurable et les seconds en tant qu'unions finies de boréliens. Nous avons comme précédemment

$$\int_I \big(s_N(\alpha) - s_{n^2}(\alpha)\big)^2 \, d\alpha \geq \int_{E_n^N} \big(s_N(\alpha) - s_{n^2}(\alpha)\big)^2 \, d\alpha \geq \zeta^2 n^4 \mathfrak{M} E_n^N,$$
$$\text{d'où} \quad \mathfrak{M} E_n^N < \frac{3C''}{\zeta^2 n^3} \quad \text{et} \quad \mathfrak{M} F_n < \frac{3C''}{\zeta^2 n^3}(2n+1) \leq \frac{9C''}{\zeta^2 n^2},$$

et par conséquent la série $\sum_{n=1}^{\infty} \mathfrak{M} F_n$ converge dans $[0, \infty[$. Nous obtenons ainsi que

$$\frac{s_N(\alpha)}{n^2} - \frac{s_{n^2}(\alpha)}{n^2} \to 0$$

† Il suffit de montrer que si $0 < q < 1$, alors $\sum_{n=0}^{\infty} q^{\sqrt{n}}$ converge. On peut par exemple utiliser le critère de condensation avec la suite n^2.

presque partout lorsque $n \to \infty$ car ζ était arbitraire (il est entendu que N est toujours supérieur ou égal à n^2 et strictement inférieur à $(n+1)^2$). Mais nous savons déjà que $s_{n^2}(\alpha)/n^2$ converge vers 0 presque partout, et par conséquent $s_N(\alpha)/n^2 \to 0$ presque partout lorsque $n \to \infty$ (pour $n^2 \leq N < (n+1)^2$), ce qui implique, vu que $N \geq n^2$, que $s_N(\alpha)/N \to 0$ presque partout lorsque $N \to \infty$. En revenant à la définition de s_N, cela signifie que

$$\frac{1}{N}\sum_{k=1}^{N} f_k(\alpha) - \frac{1}{N}\sum_{k=1}^{N} u_k \to 0 \tag{61}$$

presque partout lorsque $n \to \infty$. En vertu de l'inégalité (47), nous avons pour tout $k \geq 1$

$$\begin{aligned}\left| u_k - \sum_{r=1}^{\infty} f(r)\frac{\log\left(1+\frac{1}{r(r+2)}\right)}{\log 2}\right| &= \sum_{r=1}^{\infty} f(r)\left|\mathfrak{M}E\binom{k}{r} - \frac{\log\left(1+\frac{1}{r(r+2)}\right)}{\log 2}\right| \\ &< Ae^{-\lambda\sqrt{k-1}}\sum_{r=1}^{\infty}\frac{f(r)}{r(r+1)} = ASe^{-\lambda\sqrt{k-1}},\end{aligned}$$

(la série étant convergente d'après les restrictions imposées à f) de telle sorte que

$$u_k \to \sum_{r=1}^{\infty} f(r)\frac{\log\left(1+\frac{1}{r(r+2)}\right)}{\log 2}$$

lorsque $k \to \infty$ et par le procédé de sommation de Cesàro,

$$\frac{1}{N}\sum_{k=1}^{N} u_k \to \sum_{r=1}^{\infty} f(r)\frac{\log\left(1+\frac{1}{r(r+2)}\right)}{\log 2}$$

lorsque $N \to \infty$. De la formule (61), on obtient que

$$\frac{1}{N}\sum_{k=1}^{N} f_k(\alpha) \to \sum_{r=1}^{\infty} f(r)\frac{\log\left(1+\frac{1}{r(r+2)}\right)}{\log 2}$$

presque partout lorsque $N \to \infty$, ce qu'il fallait démontrer. □

En appliquant le théorème 29 à la fonction $r \mapsto \log(r)$ (qui vérifie clairement les hypothèses requises), on obtient que

$$\frac{1}{n}\sum_{k=1}^{n} \log(a_k(\alpha)) \to \sum_{r=1}^{\infty} \log(r)\frac{\log\left(1+\frac{1}{r(r+2)}\right)}{\log 2}$$

presque partout lorsque $n \to \infty$. En prenant l'exponentielle (qui est une fonction continue) de ces sommes, on obtient le théorème de Khinchin.

Théorème 30. *Pour presque tout α dans I,*

$$\lim_{n\to\infty} \sqrt[n]{a_1(\alpha)\dots a_n(\alpha)} = \prod_{r=1}^{\infty}\left(1+\frac{1}{r(r+2)}\right)^{\log r/\log 2}.$$

Le nombre

$$K_0 = \prod_{r=1}^{\infty}\left(1+\frac{1}{r(r+2)}\right)^{\log r/\log 2} \approx 2.685$$

est la *constante de Khinchin*. Le théorème de Khinchin est évidemment vrai pour l'ensemble de tous les nombres irrationnels, et même pour l'ensemble des nombres réels si l'on définit la suite de fonctions $(a_i)_{i=1}^{\infty}$ de manière arbitraire pour les nombres rationnels.

Un autre corollaire très intéressant du théorème 29 s'obtient en choisissant pour f la fonction caractéristique de l'ensemble $\{k\}$ pour un nombre entier $k \geq 1$ donné. Cette fonction satisfait trivialement les hypothèses du théorème. Dans ce cas, la somme

$$\frac{1}{n}\sum_{i=1}^{n}(f \circ a_i)(\alpha)$$

est la densité d'occurence de l'entier k dans les n premiers éléments de la fraction continue simple d'un nombre α dans I. Nous obtenons ainsi que cette somme converge presque partout vers le nombre

$$g(k) = \frac{\log\left(1 + \frac{1}{k(k+2)}\right)}{\log 2}$$

qui n'est autre que la densité de Gauss–Kuz'min (c'est d'ailleurs ce résultat qui lui donne le qualificatif de « densité »). Nous trouvons ainsi une deuxième signification à la densité de Gauss–Kuz'min, beaucoup plus significative. Si l'on choisit un nombre réel quelconque (en dehors d'un ensemble de mesure nulle) on trouvera dans son expression en fraction continue simple exactement $g(1) \approx 41.5\%$ d'éléments égaux à 1, $g(2) \approx 17.0\%$ d'éléments égaux à 2, $g(3) \approx 9.3\%$ d'éléments égaux à 3, etc.

3 Éléments de théorie ergodique

3.1 Le théorème ergodique ponctuel

Remarque préliminaire. Dans les théorèmes de cette section, tous les espaces mesurés seront supposés σ-finis. Il est nécessaire de faire cette supposition pour que ces théorèmes soient valides en toute généralité.

Soit $(X, \mathscr{A}, \mu)$ un espace mesuré. Une *transformation* de $(X, \mathscr{A}, \mu)$ est une application $(\mathscr{A}, \mathscr{A})$-mesurable de X dans lui-même. On dit qu'une transformation T de $(X, \mathscr{A}, \mu)$ *préserve la mesure* μ si $\mu(T^{-1}(A)) = \mu(A)$ pour tout élément A de $\mathscr{A}$, c'est-à-dire si la mesure image $T_*\mu$ est égale à μ. Si T est une transformation, une fonction f mesurable sur $(X, \mathscr{A}, \mu)$ à valeurs dans la droite réelle achevée est *T-invariante* si $(f \circ T)(x) = f(x)$ presque partout ; elle est *T-invariante partout* si $f \circ T = f$.

La *théorie ergodique* étudie les structures $(X, \mathscr{A}, \mu, G, \cdot, a)$ où $(X, \mathscr{A}, \mu)$ est un espace mesuré, $(G, \cdot)$ un monoïde et $a \in X^{G \times X}$ une action de $(G, \cdot)$ sur X telle que pour tout g dans G l'application $x \mapsto a(g, x)$ est une transformation qui préserve la mesure μ ; on dit aussi d'une telle action qu'elle préserve la mesure μ. Il arrive souvent que $(G, \cdot)$ soit un groupe lorsque l'on considère des transformations inversibles. Une telle structure est appelée un *système dynamique*, une terminologie qui trouve bien sûr son origine dans le monde de la physique où est née la théorie ergodique.

Un cas particulièrement intéressant est constitué des systèmes dynamiques de la forme $(X, \mathscr{A}, \mu, \langle T \rangle, a)$ où $\langle T \rangle$ est le monoïde engendré par une transformation T sur $(X, \mathscr{A}, \mu)$ préservant la mesure μ et où $a(T^k, x) = T^k(x)$, que nous appellerons des *systèmes dynamiques simples* et que nous noterons simplement $(X, \mathscr{A}, \mu, T)$.

Sous le nom de « théorème ergodique » on recense souvent divers théorèmes plus ou moins similaires qui ont une importance particulière dans la théorie ergodique. Le théorème qui nous intéresse ici est plus précisément le *théorème ergodique ponctuel* (aussi appelé le théorème ergodique *fort* ou *individuel*) découvert par George David Birkhoff.

Théorème 31 (Birkhoff, 1931). *Soient $(X, \mathscr{A}, \mu, T)$ un système dynamique simple et f une fonction réelle intégrable sur $(X, \mathscr{A}, \mu)$. Alors la suite*

$$\frac{1}{n} \sum_{k=0}^{n-1} (f \circ T^k)(x) \tag{62}$$

converge presque partout lorsque n tend vers l'infini. Si $f^*(x)$ en est la limite presque partout, alors f^* est intégrable et T-invariante. De plus, si $\mu(X)$ est fini,

$$\int_X f^* \, d\mu = \int_X f \, d\mu.$$

La fonction limite f^*, uniquement déterminée modulo μ, s'appelle la *moyenne temporelle* de f, et si $\mu(X)$ est fini et non nul, le nombre $\frac{1}{\mu(X)} \int_X f \, d\mu$ est la *moyenne spatiale* de f.

Pour illustrer le théorème ergodique ponctuel (et uniquement dans ce but), considérons le système dynamique simple $(X, \mathscr{P}(X), \mu, T)$ où X est un « ensemble de positions » que prennent des particules d'un gaz contenu dans un récipient (qu'on supposera fini), μ la mesure de comptage d'une partie du gaz et T la transformation du système par unité de temps, c'est-à-dire que $T^k(x)$ est la position après k unités de temps de la particule qui se trouvait originellement en x. Dans ce modèle physique, il est clair qu'à toute partie du gaz on peut associer le nombre de particules qu'elle contient, d'où le choix de $\mathscr{P}(X)$ pour la σ-algèbre des ensembles mesurables. En outre, toute partie non vide du gaz contient des particules ; en conséquence, les résultats « presque partout » seront ici valable partout. Il est évident dans cette situation que T préserve la mesure μ.[†] Définissons $f(x)$

[†] Nous avons ici simplifié le modèle physique à l'extrême en considérant un système à temps et espace discrets. Une modélisation plus réaliste de cette situation nécessite un sytème dynamique plus complexe, dans lequel le monoïde agissant sur X est constitué d'une famille continue $\{T_t \mid t \in [0, \infty[\}$, ou *flot*, de transformations telles que $T_t \circ T_s = T_{t+s}$. Il existe dans ce cas un théorème ergodique similaire au théorème ponctuel dans lequel la somme $\frac{1}{n} \sum_{k=0}^{n-1} (f \circ T^k)(x)$ est remplacée par l'intégrale $\frac{1}{\tau} \int_0^\tau (f \circ T_t)(x) \, dt$.

comme étant la vitesse de la particule se trouvant en x, de telle sorte que $(f \circ T^k)(x)$ est la vitesse de cette même particule après k unités de temps. La somme

$$\frac{1}{n}\sum_{k=0}^{n-1}(f \circ T^k)(x)$$

est la vitesse moyenne entre les instants 0 et $n-1$ de la particule se trouvant originellement en x tandis que l'intégrale

$$\int_X f\, d\mu$$

est la somme des vitesses de toutes les particules à l'instant initial. Le théorème ergodique ponctuel affirme d'une part que la vitesse moyenne d'une particule quelconque pendant n unités de temps depuis l'instant initial converge vers une certaine quantité lorsque n tend vers l'infini et d'autre part que cette quantité ne dépend pas de cet instant initial (on appellera cette quantité la *vitesse moyenne* de la particule). De plus, la somme des vitesses moyennes de toutes les particules est égale à la somme des vitesses de toutes les particules à un instant donné.

La preuve du théorème ergodique ponctuel que nous donnons est due à Frigyes Riesz [15]. Elle commence par établir deux lemmes.

Si x_1, …, x_n sont des nombres réels et si $m \leq n$ est un entier, on note $A_m(x_1, \ldots, x_n)$ l'ensemble des nombres x_k pour lesquels il existe un entier strictement positif $p \leq m$ tel que $x_k + \cdots + x_{k+p-1} \geq 0$. On utilisera ici la convention que la somme vide $\sum \varnothing$ est nulle, de sorte que pour deux parties A et B de $\mathbb{R}$ finies et disjointes la formule $\sum(A \cup B) = \sum A + \sum B$ reste valide si A ou B est vide.

Lemme 32. *Soient x_1, …, x_n des nombres réels et $m \leq n$ un entier strictement positif. Alors la somme $\sum A_m(x_1, \ldots, x_n)$ est positive ou nulle.*

Preuve. Procédons par récurrence. Si $n = 1$ et $A_1(x_1)$ est non vide, alors $\sum A_1(x_1) = x_1 \geq 0$. Soient x_1, …, x_n des nombres réels et $m \leq n$ un entier. En vertu de notre convention, on peut supposer $A_m(x_1, \ldots, x_n)$ non vide. Soit k le plus petit indice tel que x_k appartient à $A_m(x_1, \ldots, x_n)$, et soit $p \leq m$ le plus petit entier tel que $x_k + \cdots + x_{k+p-1} \geq 0$ comme dans la définition de $A_m(x_1, \ldots, x_n)$. Toutes les sommes $x_q + \cdots + x_{k+p-1}$, pour $k < q \leq k+p-1$, sont alors positives ou nulle, car sinon on aurait $x_k + \cdots + x_{q-1} > 0$ en contradiction avec la minimalité de p. Ainsi, tous les nombres x_k, …, x_{k+p-1} appartiennent à $A_m(x_1, \ldots, x_n)$. Mais $\sum A_m(x_1, \ldots, x_n) = x_k + \cdots + x_{k+p-1} + \sum A_m(x_{k+p}, \ldots, x_n) \geq \sum A_m(x_{k+p}, \ldots, x_n)$, et cette dernière somme est positive ou nulle par hypothèse de récurrence. $\square$

On notera dans la suite de la démonstration $f_i = f \circ T^i$. Remarquons qu'une conséquence immédiate du fait que T préserve la mesure μ est que l'intégrale de f sur X est égale à celle de f_i sur X. Le lemme suivant est le *théorème ergodique maximal*. Il reprend les hypothèse du théorème 31.

Lemme 33. *Soit E l'ensemble des x dans X pour lesquels il existe un entier $n \geq 1$ tel que $f_0(x) + \cdots + f_{n-1}(x) \geq 0$. Alors l'intégrale $\int_E f\, d\mu$ est positive ou nulle.*

Preuve. Pour $m \geq 0$, notons E_m l'emsemble des x dans X pour lesquels il existe un entier $p \leq m$ tel que $f_0 + \cdots + f_p \geq 0$. La suite $(E_m)_{m=0}^{\infty}$ est croissante et $\bigcup_{m=0}^{\infty} E_m = E$. Chaque ensemble E_m est évidemment mesurable, car c'est l'intersection d'un nombre fini de préimages de boréliens par des fonctions mesurables. Vu que

$$\int_E f\, d\mu = \lim_{m\to\infty} \int_{E_m} f\, d\mu,$$

il suffit de montrer que

$$\int_{E_m} f\, d\mu \geq 0$$

pour tout entier $m \geq 1$. Fixons des entiers strictement positifs m et n. Pour x dans X, on considère la somme $\sum A_m(f_0(x), \ldots, f_{n+m-1}(x)) = s(x)$, qui, nous le savons par le lemme 32, est positive ou nulle. Pour $k \geq 0$, notons D_k l'ensemble des x dans X pour lesquels $f_k(x)$ appartient à

$A_m(f_0(x),\dots,f_{n+m-1}(x))$, et χ_{D_k} sa fonction caractéristique. L'ensemble D_k est mesurable car c'est la préimage par f_k de $A_m(f_0(x),\dots,f_{n+m-1}(x))$ qui est fini. Il s'ensuit que s est intégrable, car

$$s = \sum_{k=0}^{n+m-1} f_k \chi_{D_k},$$

et son intégrale est positive ou nulle, c'est-à-dire

$$\int_X s\, d\mu = \sum_{k=0}^{n+m-1} \int_{D_k} f_k\, d\mu \geq 0. \tag{63}$$

Nous prétendons que si $1 \leq k \leq n-1$, alors $D_k = T^{-1}(D_{k-1})$, de sorte que $D_k = T^{-k}(D_0)$. En effet, $T(x)$ appartient à D_{k-1} si et seulement s'il existe un entier $p \leq m$ tel que $f_{k-1}(T(x)) + \cdots + f_{k-1+p-1}(T(x)) \geq 0$, c'est-à-dire tel que $f_k(x) + \cdots + f_{k+p-1}(x) \geq 0$, mais ceci est équivalent à dire que x appartient à D_k, ce qui prouve l'assertion. Par conséquent, la formule du changement de mesure[†] nous donne, pour $1 \leq k \leq n-1$,

$$\int_{D_k} f_k\, d\mu = \int_{T^{-k}(D_0)} (f \circ T^k)\, d\mu = \int_{D_0} f\, dT_*^k\mu = \int_{D_0} f\, d\mu.$$

Ceci implique que les n premiers termes de la somme (63) sont tous égaux à $\int_{D_0} f\, d\mu$. D'autre part, il est clair que $D_0 = E_m$, et ainsi

$$0 \leq \sum_{k=0}^{n+m-1} \int_{D_k} f_k\, d\mu = n \int_{E_m} f\, d\mu + \sum_{k=n}^{n+m-1} \int_{D_k} f_k\, d\mu \leq n \int_{E_m} f\, d\mu + m \int_X |f|\, d\mu.$$

En divisant la relation précédente par n, on obtient

$$\int_{E_m} f\, d\mu + \frac{m}{n} \int_X |f|\, d\mu \geq 0,$$

et, la deuxième intégrale étant finie, on conclut en faisant tendre n vers l'infini. □

Preuve du théorème 31. Soient a et b des nombres réels avec $a < b$, et notons

$$f_\infty = \liminf_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f_k \quad \text{et} \quad f^\infty = \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f_k.$$

Les fonctions f_∞ et f^∞ sont alors mesurables et, en vertu de la seconde assertion du lemme 28, T-invariantes partout. Soit Y l'ensemble des x dans X pour lesquels $f_\infty(x) < a < b < f^\infty(x)$. L'ensemble Y est mesurable en tant qu'intersection de deux préimages de boréliens par des fonctions mesurables. Vu que f_∞ et f^∞ sont T-invariantes partout, il est clair que $T^{-1}(Y) = Y$, ce qui implique $T(Y) \subseteq Y$. Nous voulons montrer que $\mu(Y) = 0$ et on peut donc supposer que Y est non vide.

Soit C un sous-ensemble mesurable de Y dont la mesure est finie, et soit χ_C sa fonction caractéristique. Comme C a mesure finie, χ_C est intégrable, et donc $f - b\chi_C$ l'est aussi. On peut ainsi appliquer le théorème ergodique maximal à la fonction $f - b\chi_C$. Si F est l'ensemble des x dans X pour lesquels il existe $n \geq 1$ tel que $((f - b\chi_C) \circ T^0)(x) + \cdots + ((f - b\chi_C) \circ T^{n-1})(x) \geq 0$, alors la conclusion du théorème ergodique maximal est que

$$\int_F (f - b\chi_C)\, d\mu \geq 0. \tag{64}$$

Si x appartient à Y, alors il existe un entier n (il en existe même une infinité) tel que

$$\frac{1}{n} \sum_{k=0}^{n-1} f_k(x) > b,$$

[†][3] p. 163, Theorem C.

et ceci implique que la somme

$$\sum_{k=0}^{n-1} \left(f_k(x) - b\chi_C(T^k(x)) \right)$$

est positive ou nulle, ce qui signifie que x appartient à F. Nous avons montré que Y est inclus dans F, et en particulier $C \subseteq F$. Par conséquent, $\int_F \chi_C \, d\mu = \mu(C)$, et de (64) on déduit que $\int_F |f| \, d\mu \geq \int_F f \, d\mu \geq b\mu(C)$. En répétant le raisonnement de ce paragraphe en remplaçant f par $-f$ et b par $-a$, on obtient de même $\int_F |f| \, d\mu \geq -a\mu(C)$. Tous les sous-ensembles mesurables de Y dont la mesure est finie ont donc non seulement une mesure finie mais bornée par $\frac{1}{b} \int_X |f| \, d\mu$ (si $b > 0$) ou par $\frac{1}{|a|} \int_X |f| \, d\mu$ (si $b \leq 0$, auquel cas $a < 0$), ce qui implique, par σ-finitude et par la proposition 15, que la mesure de Y est finie.

On considère maintenant la transformation $T|_Y$ de l'espace mesuré $(Y, \mathscr{A}', \mu|_{\mathscr{A}'})$, où $\mathscr{A}'$ est la famille des sous-ensembles mesurables de Y, bien définie vu que $T(Y) \subseteq Y$. Si l'on applique comme ci-dessus le théorème ergodique maximal à la fonction $f - b$ définie sur Y (dont on sait que la mesure est finie), l'ensemble correspondant à F est alors l'ensemble Y tout entier, de sorte que $\int_Y (f - b) \, d\mu \geq 0$. Un raisonnement indentique avec la fonction $a - f$ donne $\int_Y (a - f) \, d\mu \geq 0$. Nous obtenons ainsi que $\int_Y (a - b) \, d\mu \geq 0$, et puisque $a < b$ cela force $\mu(Y) = 0$.

Si nous prenons l'union dénombrable de tous les Y lorsque a et b prennent toutes les valeurs rationnelles avec $a < b$, on trouve que cette union est de mesure nulle, et par conséquent que $f_\infty(x) = f^\infty(x)$ presque partout. On obtient ainsi que $f_\infty(x)$ est la limite presque partout de la suite (62). De plus, puisque

$$\int_X \left| \frac{1}{n} \sum_{k=0}^{n-1} f_k \right| d\mu \leq \frac{1}{n} \sum_{k=0}^{n-1} \int_X |f_k| \, d\mu = \int_X |f| \, d\mu < \infty,$$

on déduit du lemme de Fatou† que

$$\int_X |f_\infty| \, d\mu \leq \int_X |f| \, d\mu < \infty,$$

c'est-à-dire que f_∞ est intégrable. Une fonction f^* est une limite de (62) presque partout si et seulement si f^* est équivalente à f_∞ modulo μ, auquel cas f^* est T-invariante et intégrable.

Finalement, supposons que $\mu(X)$ est fini. Pour des entiers k et n avec $n \geq 1$, notons $X(k,n)$ l'ensemble des x dans X tels que $k/n \leq f_\infty(x) < (k+1)/n$. Les ensembles $X(k,n)$ sont mesurables et tels que $T^{-1}(X(k,n)) = X(k,n)$ car f_∞ est T-invariante partout. De plus, si n est fixé, les ensembles $X(k,n)$ forment une partition de X. Pour presque tout x dans $X(k,n)$ et pour tout $\varepsilon > 0$, il existe un entier $p \geq 1$ tel que les sommes

$$\sum_{i=0}^{p-1} (f_i(x) - k/n + \varepsilon) \quad \text{et} \quad \sum_{i=0}^{p-1} ((f_\infty \circ T^i)(x) - k/n + \varepsilon)$$

sont positives ou nulles (c'est même vrai quel que soit p et quel que soit x pour la deuxième somme). Par le théorème ergodique maximal (appliqué à la transformation $T|_{X(k,n)}$ de $(X(k,n), \mathscr{A}'', \mu|_{\mathscr{A}''})$, où $\mathscr{A}''$ est la famille des sous-ensembles mesurables de $X(k,n)$, qui est bien définie car $T(X(k,n)) \subseteq X(k,n)$) nous obtenons

$$\int_{X(k,n)} f \, d\mu \geq (k/n - \varepsilon)\mu(X(k,n)) \quad \text{et} \quad \int_{X(k,n)} f_\infty \, d\mu \geq (k/n - \varepsilon)\mu(X(k,n)),$$

ce qui implique, ε étant arbitraire, $\int_{X(k,n)} f \, d\mu \geq k/n \, \mu(X(k,n))$ et $\int_{X(k,n)} f_\infty \, d\mu \geq k/n \, \mu(X(k,n))$. De même, vu que $f_\infty(x) < (k+1)/n$ pour tout x dans $X(k,n)$, on trouve $\int_{X(k,n)} f \, d\mu \leq (k+1)/n \, \mu(X(k,n))$ et $\int_{X(k,n)} f_\infty \, d\mu \leq (k+1)/n \, \mu(X(k,n))$. De ces quatre inégalités, on déduit la relation

$$-\frac{1}{n}\mu(X(k,n)) \leq \int_{X(k,n)} f \, d\mu - \int_{X(k,n)} f_\infty \, d\mu \leq \frac{1}{n}\mu(X(k,n)),$$

†[3] p. 113, Theorem F.

qui, sommée sur tous les entiers k, devient

$$\left|\int_X f\,d\mu - \int_X f_\infty\,d\mu\right| \leq \frac{1}{n}\mu(X). \tag{65}$$

Cette relation reste vraie si l'on remplace f_∞ par une fonction f^* équivalente modulo μ. On conclut en laissant n tendre vers l'infini dans (65). □

Soit $(X, \mathscr{A}, \mu)$ un espace mesuré muni d'une transformation T préservant la mesure. Supposons qu'il existe une partition $\{Y, Z\} \subset \mathscr{A}$ de X telle que $\mu(Y \triangle T^{-1}(Y)) = 0$ et $\mu(Z \triangle T^{-1}(Z)) = 0$. En d'autres termes, Y et Z sont invariants par T à un ensemble de mesure nulle près. On dit dans ce cas que Y et Z sont *T-invariants* et, si Y et Z sont non négligeables, que T est *décomposable*. L'étude du système dynamique simple $(X, \mathscr{A}, \mu, T)$ se réduit alors à l'étude de deux systèmes dynamiques simples indépendants sur Y et sur Z. Ces considérations motivent la définition suivante.

On dit qu'une transformation T d'un espace mesuré $(X, \mathscr{A}, \mu)$ est *ergodique* (ou *indécomposable*) si elle préserve la mesure et si tout élément T-invariant de $\mathscr{A}$ est de mesure nulle ou le complémentaire d'un ensemble de mesure nulle. Un système dynamique simple $(X, \mathscr{A}, \mu, T)$ où T est ergodique est appelé un *système ergodique*.

LEMME 34. *Soit $(X, \mathscr{A}, \mu, T)$ un système ergodique. Toute fonction réelle T-invariante sur $(X, \mathscr{A}, \mu)$ est constante presque partout.*

Preuve. Soit f une fonction réelle T-invariante (en particulier mesurable) sur $(X, \mathscr{A}, \mu)$. Pour tous entiers k et n avec $n \geq 1$, définissons $X(k, n)$ comme étant l'ensemble des éléments x de X tels que $k/n \leq f(x) < (k+1)/n$. Alors les $X(k, n)$ sont évidemment T-invariants (ils sont mesurables en tant que préimages par f de boréliens). Si n est fixé, les ensembles $X(k, n)$ forment une partition de X, et donc, puisque T est ergodique, tous sauf un ont mesure nulle, et le dernier, disons $X(k_n, n)$, est le complémentaire d'un ensemble de mesure nulle. En conséquence, le complémentaire de l'ensemble mesurable $Y = \bigcap_{n=1}^{\infty} X(k_n, n)$ a mesure nulle, et puisque f est constante sur Y (car $(k_n + 1)/n - k_n/n \to 0$ lorsque $n \to \infty$), le résultat est démontré. □

THÉORÈME 35. *Soient $(X, \mathscr{A}, \mu, T)$ un système ergodique et f une fonction réelle intégrable sur $(X, \mathscr{A}, \mu)$. Supposons que $\mu(X)$ est fini et non nul. Alors la moyenne temporelle de f est égale presque partout à la moyenne spatiale de f.*

Preuve. Puisque T est ergodique, la moyenne temporelle f^* de f est une constante presque partout car elle est T-invariante. Notons c cette constante. Si $\mu(X)$ est fini, alors $\int_X f\,d\mu = \int_X f^*\,d\mu = c\mu(X)$, et par conséquent si $\mu(X)$ est non nul, $c = \frac{1}{\mu(X)} \int_X f\,d\mu$. □

Revenons à l'exemple du système gazeux. Il est tout à fait raisonnable du point de vue physique de supposer que T est ergodique. En effet, on n'imagine pas qu'une partie du gaz ne se mélange pas avec son complémentaire pendant une unité de temps, sauf bien sûr si cette partie est vide ou est l'ensemble tout entier. Le théorème 35 a dans ce cas la signification très simple suivante. La vitesse moyenne d'une particule quelconque est égale à la moyenne des vitesses de toutes les particules à un instant quelconque.

Le théorème 35 a une telle importance en physique statistique qu'il est souvent référencé sous le nom du théorème ergodique.

3.2 Applications aux fractions continues simples

Les résultats de cette section sont essentiellement tirés de l'article de Czesław Ryll-Nardzewski [16].

Comment pouvons-nous appliquer les résultats de la théorie ergodique aux fractions continues ? Il faut tout d'abord définir un système dynamique simple $(X, \mathscr{A}, \mu, T)$ pour lequel on pourra appliquer le théorème ergodique ponctuel. Pour l'ensemble X nous prendrons évidemment l'intervalle irrationnel $I = [0, 1[\setminus \mathbb{Q}$. Pour la transformation T nous prendrons la fonction définie sur I par $T([a_1, a_2, \ldots]) = [a_2, a_3, \ldots]$, ou plus explicitement

$$T(\alpha) = \frac{1}{\alpha} - \left\lfloor \frac{1}{\alpha} \right\rfloor.$$

Cette transformation est parfois appelée l'*opérateur de Gauss–Kuz'min–Wirsing* ; elle a été introduite par Edward Marczewski [4]. On est tenté de prendre pour μ la mesure de Lebesgue comme dans le chapitre précédent, mais il s'avère que la transformation T ne préserve pas cette mesure. Nous définissons donc une nouvelle mesure μ sur $(I, \mathscr{B}')$ par[†]

$$\mu(E) = \frac{1}{\log 2} \int_E \frac{dx}{1+x}.$$

Le fait que μ est une mesure sur $(I, \mathscr{B}')$ découle des relations

$$\int_\varnothing f(x)\,dx = 0 \quad \text{et} \quad \int_{\bigcup_{n=0}^\infty E_n} f(x)\,dx = \sum_{n=0}^\infty \int_{E_n} f(x)\,dx$$

valables pour toute suite $(E_n)_{n=0}^\infty$ d'éléments de $\mathscr{B}'$ deux à deux disjoints et toute fonction f mesurable sur $(I, \mathscr{B}')$.

Proposition 36. *La mesure μ est équivalente à la mesure de Lebesgue sur $\mathscr{B}'$.*

Preuve. Observons que pour tout E dans $\mathscr{B}'$ nous avons les inégalités

$$\frac{1}{2\log 2}\mathfrak{M}E \leq \mu(E) \leq \frac{1}{\log 2}\mathfrak{M}E.$$

De la première inégalité on obtient que $\mathfrak{M}|_{\mathscr{B}'}$ est absolument continue par rapport à μ, et de la deuxième que μ est absolument continue par rapport à $\mathfrak{M}|_{\mathscr{B}'}$. Les deux mesures sont donc équivalentes. □

En particulier, une fonction est intégrable sur $(I, \mathscr{B}', \mu)$ si et seulement si elle est intégrable sur $(I, \mathscr{B}', \mathfrak{M}|_{\mathscr{B}'})$.

Proposition 37. *La transformation T préserve la mesure μ.*

Preuve. Montrons dans un premier temps que pour α dans $[0,1]$, la mesure de l'intervalle irrationnel $E = [0, \alpha] \setminus \mathbb{Q}$ est préservée par T. Or,

$$T^{-1}(E) = \bigcup_{n=1}^\infty \left[\frac{1}{n+\alpha}, \frac{1}{n}\right] \setminus \mathbb{Q},$$

et, la réunion étant disjointe, on obtient

$$\begin{aligned}\mu(T^{-1}(E)) &= \frac{1}{\log 2} \sum_{n=1}^\infty \int_{\left[\frac{1}{n+\alpha}, \frac{1}{n}\right]\setminus\mathbb{Q}} \frac{dx}{1+x} \\ &= \frac{1}{\log 2} \log\left(\prod_{n=1}^\infty \frac{(n+1)(n+\alpha)}{n(n+\alpha+1)}\right) = \frac{\log(1+\alpha)}{\log 2} = \frac{1}{\log 2}\int_E \frac{dx}{1+x} = \mu(E).\end{aligned}$$

Le résultat est également vérifié pour tous les intervalles irrationnels de la forme $]\alpha, \beta] \setminus \mathbb{Q}$, car $\mu(T^{-1}(]\alpha, \beta] \setminus \mathbb{Q})) = \mu(T^{-1}([0, \beta] \setminus \mathbb{Q})) - \mu(T^{-1}([0, \alpha] \setminus \mathbb{Q})) = \mu(]\alpha, \beta] \setminus \mathbb{Q})$. Puisque $T^{-1}(\{\alpha\})$ est dénombrable (si α appartient à I), on obtient que la mesure d'un intervalle quelconque est préservée.

Notons $\mathscr{I}$ le semi-anneau des intervalles irrationnels de I et $\mathscr{R}$ l'anneau engendré par $\mathscr{I}$. Nous avons montré que la mesure μ et son image $T_*\mu$ coïncident sur $\mathscr{I}$. Par conséquent,[‡] elles coïncident également sur l'anneau $\mathscr{R}$, et vu que $\mathscr{B}'$ est le σ-anneau engendré par $\mathscr{R}$, il s'ensuit[§] que $\mu = T_*\mu$. □

Si E est un borélien de I, notons $\mathfrak{M}_E$ la fonction de $\mathscr{B}'$ dans $[0, \mathfrak{M}E]$ qui associe à chaque borélien B le nombre $\mathfrak{M}(E \cap B)$. La proposition suivante est due à Konrad Knopp [9].

[†] Rappelons que $\mathscr{B}'$ est la famille des boréliens de I.
[‡] [3] p. 37, Exercise (5).
[§] [3], p. 54, Theorem A.

Proposition 38. *$(I, \mathscr{B}', \mu, T)$ est un système ergodique.*

Preuve. Puisque les mesures μ et $\mathfrak{M}|_{\mathscr{B}'}$ sont équivalentes, il suffit de montrer que la mesure de Lebesgue de tout élément T-invariant de $\mathscr{B}'$ est soit 0 soit 1. Soit $E \in \mathscr{B}'$ un ensemble T-invariant tel que $\mathfrak{M}E < 1$. Nous allons montrer que $\mathfrak{M}E = 0$. Notons χ_E la fonction caractéristique de l'ensemble E. Soit J_n l'intervalle de rang n contenant les nombres α de I pour lesquels $a_1(\alpha) = k_1$, $\ldots$, $a_n(\alpha) = k_n$, et notons p_k et q_k les convergents de la fraction continue $[k_1, \ldots, k_n]$, de sorte que l'intervalle de rang n considéré a pour extrémités

$$\frac{p_n}{q_n} \quad \text{et} \quad \frac{p_n + p_{n-1}}{q_n + q_{n-1}},$$

et a pour mesure $1/(q_n(q_n + q_{n-1}))$. Remarquons que, pour presque tout x dans I, la T-invariance de E implique

$$\chi_E\left(\frac{p_n + xp_{n-1}}{q_n + xq_{n-1}}\right) = \chi_E([k_1, \ldots, k_n + x]) = \chi_E(T^n([k_1, \ldots, k_n + x])) = \chi_E(x)$$

(pour la première égalité, voir la formule (50)). Par conséquent,

$$\begin{aligned} \frac{\mathfrak{M}_E J_n}{\mathfrak{M} J_n} &= q_n(q_n + q_{n-1}) \int_{J_n} \chi_E(y)\, dy \\ &= q_n(q_n + q_{n-1}) \int_I \chi_E\left(\frac{p_n + up_{n-1}}{q_n + uq_{n-1}}\right) \frac{du}{(q_n + uq_{n-1})^2} \\ &= q_n(q_n + q_{n-1}) \int_I \chi_E(u) \frac{du}{(q_n + uq_{n-1})^2}. \end{aligned}$$

Puisque la fonction $x \mapsto 1/(q_n + xq_{n-1})^2$ est décroissante, la dernière intégrale sera maximale si $E = [0, \mathfrak{M}E] \setminus \mathbb{Q}$, et donc

$$\begin{aligned} \frac{\mathfrak{M}_E J_n}{\mathfrak{M} J_n} &\leq q_n(q_n + q_{n-1}) \int_{[0,\mathfrak{M}E]\setminus\mathbb{Q}} \frac{du}{(q_n + uq_{n-1})^2} \\ &= 1 - \frac{q_{n-1}(1 - \mathfrak{M}E)}{q_n + \mathfrak{M}E q_{n-1}} \leq 1 - \frac{1 - \mathfrak{M}E}{1 + \mathfrak{M}E} < 1, \end{aligned} \tag{66}$$

car on a supposé $\mathfrak{M}E < 1$, et ceci est vrai pour tous les intervalles de rang n et pour tout entier $n \geq 0$. Puisque le diamètre de l'intervalle de rang n (qui coïncide ici avec sa mesure) contenant un point α de I est aussi petit que l'on veut si n est suffisamment grand (cf. proposition 13), l'inégalité (66) signifie que la dérivée inférieure (au sens général) de la fonction $\mathfrak{M}_E$ par rapport à la mesure $\mathfrak{M}$ est strictement inférieure à 1 en tout point α de I. Mais nous savons par le théorème de la densité de Lebesgue† que $\mathfrak{M}_E$ est presque partout dérivable et que $D\mathfrak{M}_E(\alpha) = \chi_E(\alpha)$ presque partout. Puisqu'ici $D\mathfrak{M}_E(\alpha) < \chi_E(\alpha)$ pour tout α dans E où la dérivée est définie, cela force $\mathfrak{M}E = 0$. □

Nous avons ainsi montré que $(I, \mathscr{B}', \mu, T)$ est un système ergodique, et puisque $\mu(I) = 1$ nous pouvons appliquer le théorème 35 à n'importe quelle fonction f intégrable sur $(I, \mathscr{B}', \mathfrak{M}|_{\mathscr{B}'})$. En prenant pour f la fonction caractéristique de l'ensemble $E\left(\frac{1}{r}\right)$ pour un entier $r \geq 1$, alors $f \circ T^i$ est la fonction caractéristique de l'ensemble $E\left(\frac{i+1}{r}\right)$, de sorte que la moyenne temporelle de f en α est la densité d'occurence de r parmi les éléments de la fraction continue simple de α. Cette moyenne temporelle est égale presque partout à la moyenne spatiale de f qui est le nombre

$$\int_I f\, d\mu = \mu\left(\left[\frac{1}{r+1}, \frac{1}{r}\right] \setminus \mathbb{Q}\right) = \frac{\log\left(1 + \frac{1}{r(r+2)}\right)}{\log 2},$$

soit la densité de Gauss–Kuz'min en r. En prenant pour f la fonction définie par

$$f(\alpha) = \log(a_1(\alpha)),$$

† [17] p. 117, Theorem (6.1).

le théorème 35 donne de même que la moyenne temporelle de f, à savoir

$$\frac{1}{n}\sum_{k=1}^{n}\log(a_k(\alpha)),$$

converge presque partout vers sa moyenne spatiale

$$\int_I f\, d\mu = \sum_{r=1}^{\infty}\log(r)\mu\left(\left[\frac{1}{r+1},\frac{1}{r}\right]\setminus\mathbb{Q}\right) = \sum_{r=1}^{\infty}\log(r)\frac{\log\left(1+\frac{1}{r(r+2)}\right)}{\log 2},$$

et on en déduit, en prenant comme précédemment l'exponentielle de ces sommes, le résultat de Khinchin.

Références

[1] L. V. Ahlfors, *Complex Analysis: an introduction to the theory of analytic functions of one complex variable*, second edition, McGraw-Hill, New York, 1966

[2] P. R. Halmos, *Lectures on Ergodic Theory*, Chelsea Publishing Company, New York, 1956

[3] P. R. Halmos, *Measure Theory*, Van Nostrand Company, Princeton (New Jersey), 1966

[4] S. Hartman, E. Marczewski, C. Ryll-Nardzewski, « Théorèmes ergodiques et leurs applications », *Colloquium Mathematicum II*, **2** (1950), 109–123

[5] M. Kac, *Statistical Independence in Probability, Analysis and Number Theory*, Mathematical Association of America, Washington DC, 1959

[6] A. Ya. Khinchin, „Metrische Kettenbruchprobleme“, *Compositio Mathematica*, **1** (1935), 361–382

[7] A. Ya. Khinchin, „Zur metrischen Kettenbruchprobleme“, *Compositio Mathematica*, **3** (1936), 276–285

[8] A. Ya. Khinchin, *Continued Fractions*, Dover Publications, New York, 1997

[9] K. Knopp, „Mengentheorische Behandlung einiger Probleme der diophantischen Approximationen und der transfiniten Wahrscheinlichkeiten“, *Mathematische Annalen*, **95** (1926), 409–426

[10] R. O. Kuz'min, «Ob odnoj zadache Gaussa», *Doklady Akademii Nauk SSSR*, **A** (1928), 375–380

[11] S. Lang, *Real Analysis*, Addison-Wesley Publishing Company, 1969

[12] P. Lévy, « Sur les lois de probabilité dont dépendent les quotients complets et incomplets d'une fraction continue », *Bulletin de la Société Mathématique de France*, **57** (1929), 178–194

[13] P. Lévy, « Sur le développement en fraction continue d'un nombre choisi au hasard », *Compositio Mathematica*, **3** (1936), 286–303

[14] O. Perron, *Die Lehre von den Kettenbrüchen*, zweite verbesserte auflage, Teubner, Leipzig, 1929

[15] F. Riesz, « Sur la théorie ergodique », *Commentarii Mathematici Helvetici*, **17** (1944–45), 221–239

[16] C. Ryll-Nardzewski, “On the ergodic theorems II (Ergodic theory of continued fractions)”, *Studia Mathematica*, **12** (1951), 74–79

[17] S. Saks, *Theory of the Integral*, second revised edition, Dover Publications, New York, 1964

[18] G. Valiron, *Théorie des fonctions*, Masson et C$^{\text{ie}}$, Paris, 1966