MATHEMATISCH CENTRUM
2e Boerhaavestraat 49
AMSTERDAM (O.)

Telefoon 51660
(Administratie, Bibliotheek, Afd. Statistiek,
Afd. Zuivere Wiskunde)

Telefoon 56643
(Afd. Toegepaste Wiskunde, Rekenafdeling)

Scriptum 1

DÉMONSTRATION ÉLÉMENTAIRE DU THÉORÈME SUR LA DISTRIBUTION DES NOMBRES PREMIERS.

J.G.van der Corput.

Scriptum 1

DÉMONSTRATION ÉLÉMENTAIRE DU THÉORÈME SUR LA DISTRIBUTION DES NOMBRES PREMIERS.

Déjà dans sa jeunesse Gauss conjecturait la formule

$$\lim_{x\to\infty} \frac{\pi(x)\,\log x}{x} = 1,$$

où $\pi(x)$ désigne le nombre des nombres premiers $\leqq x$ et où $\log x$ désigne le logarithme néperien de x. Jusqu'en 1896, donc à peu près pendant un siècle, cette proposition n'a pas été démontrée et à ce moment elle le fut par M.M. De la Vallée-Poussin et Hadamard. C'est seulement en 1948 qu'une démonstration élémentaire a été trouvée par M.M. A.Selberg et P.Erdös. Leur démonstration est la seule jusqu'à présent dans laquelle la théorie des fonctions de variables complexes ou réelles n'est pas employée.

Le 30 octobre 1948 M. P.Erdös a fait à Amsterdam une conférence sur ce sujet pour la "Wiskundig Genootschap". A propos de cette conférence et de celles, faites par M. Erdös au Centre Mathématique, M. van der Corput a pris les notes suivantes. Les pensées exprimées dans ces notes sont dûes à M.M. Selberg et Erdös, mais c'est M.van der Corput, qui est principalement responsable de la forme.

Pour comprendre le raisonnement suivant, le lecteur doit savoir ce qu'est une plus grande et une plus petite limite et ce qu'est un logarithme népe-rien et il doit connaître les éléments de la théorie des séries. Plus précisément, il doit savoir qu'il existe une constante positive τ telle que le produit

$$n^{\tau}\left\{ n \log\left(1 + \frac{1}{n}\right) - 1 \right\}$$

est borné pour tous les nombres naturels n. Cela est évident. En effet, $\tau = 1$ suffit, puisque $n \log(1 + \frac{1}{n})$ est égal à $\frac{1}{n}$, à un terme près de l'ordre de grandeur $\frac{1}{n^2}$.

Dans la théorie des nombres nous sommes habitués de considérer, non pas la fonction $\pi(x)$ qui n'est pas maniable, mais la fonction

$$\vartheta(x) = \sum_{p \leqq x} \log p,$$

qui est beaucoup plus facile à traiter; cette somme est étendue aux nombres premiers $p \leq x$. Il est suffisant de démontrer que $\frac{\vartheta(x)}{x}$ tend vers 1, si x croît indéfiniment. En effet, $\vartheta(x) \leq \pi(x) \log x$ et on a pour $1 \leqq y \leqq x$

$$\pi(x) - y \leqq \pi(x) - \pi(y) \leqq \frac{1}{\log y}\left\{\vartheta(x) - \vartheta(y)\right\} \leqq \frac{\vartheta(x)}{\log y}$$

donc

$$(1) \quad \frac{\vartheta(x)}{x} \leqq \frac{\pi(x) \log x}{x} \leqq \frac{y \log x}{x} + \frac{\log x}{\log y} \cdot \frac{\vartheta(x)}{x}$$

Choisissons $y = x^{\rho}$, où le nombre $\rho < 1$ dépendant de x tend vers 1 de manière assez lente pour que $\frac{\log x}{x^{1-\rho}}$ tende vers zéro, si x croît indéfiniment.

Alors les membres extrêmes de (1), donc aussi $\frac{\pi(x) \log x}{x}$ tend vers 1. Par conséquent, pour la démonstration du théorème concernant la distribution des nombres premiers il suffit de démontrer que $\frac{\vartheta(x)}{x}$ tend vers 1, si x croît indéfiniment.

La démonstration est composée de deux parties distinctes. La première partie est consacrée, presque exclusivement à la démonstration de la formule de limite démontrée par M. Selberg:

$$\frac{\vartheta(x)}{x} + \frac{2}{x \log x} \sum_{p \leq \sqrt{x}} \vartheta\left(\frac{x}{p}\right) \log p \to 2 \text{ pour } x \to \infty.$$

Dans la seconde partie nous déduirons de cette formule le théorème fondamental des nombres premiers.

DÉMONSTRATION DE LA FORMULE DE SELBERG.

Par $\mu(m)$ nous désignons la fonction de Möbius, qui est définie de la manière suivante:

$\mu(1) = 1$;

$\mu(m) = 0$, si m est divisible par un carré > 1

et dans tous les autres cas:

$\mu(m) = +1$ ou -1, selon que le nombre des facteurs premiers de m est pair ou impair.

Donc

$\mu(1) = 1$; $\mu(2) = -1$; $\mu(3) = -1$; $\mu(4) = 0$;
$\mu(5) = -1$; $\mu(6) = 1$; $\mu(7) = -1$; $\mu(8) = 0$;
$\mu(9) = 0$; $\mu(10) = 1$; $\mu(11) = -1$; $\mu(12) = 0$;
$\mu(13) = -1$; $\mu(14) = 1$; $\mu(15) = 1$.

La fonction $\mu(m)$ de Möbius est une fonction multiplicative, c'est-à-dire on a $\mu(ab) = \mu(a)\mu(b)$, si a et b sont des nombres naturels qui sont premiers entre eux. Cette assertion résulte immédiatement de la définition de la fonction de Möbius.

LEMME 1. Pour chaque nombre entier $h \geqq 0$ la fonction

$$\varphi_h(m) = \sum_{d/m} \mu(d) \log^h d,$$

(où la somme est étendue à tous les diviseurs de m, les diviseurs 1 et m inclus) est égale à zéro, si le nombre naturel m contient plus de h facteurs premiers différents.

REMARQUE: Nous emploierons cette proposition auxiliaire seulement pour $h = 0$, $h = 1$ et $h = 2$.

DÉMONSTRATION: Pour $h = 0$ la formule s'écrit

$$(2) \qquad \sum_{d/m} \mu(d) = 0$$

pour tout entier $m > 1$ (la somme est égale à 1 pour $m = 1$). Cette formule est évidente. En effet, soit $m = p_1^{\alpha_1} \dots p_r^{\alpha_r}$ la décomposition de m en facteurs premiers. Puisque $\mu(d) = 0$ pour chaque diviseur éventuel d de m qui est divisible par un carré > 1, seulement $2^r = 1 + \binom{r}{1} + \dots \binom{r}{r}$ diviseurs de m entrent ici en considération où $\binom{r}{h} = C_r^h$ désignent les coefficients du binôme. En premier lieu le diviseur $d = 1$ avec la contribution 1, puis les $\binom{r}{1}$ diviseurs $p_1, \dots, p_r$, chacun avec la contribution -1, ensuite les $\binom{r}{2}$ diviseurs $p_1p_2, p_1p_3, \dots, p_{r-1}p_r$, chacun avec la contribution $+1$, etc., de sorte que la contribution totale est égale à

$$1 - \binom{r}{1} + \binom{r}{2} - \dots \pm \binom{r}{r} = (1 - 1)^r = 0.$$

Ainsi l'assertion est démontrée pour $h = 0$, de sorte que nous pouvons supposer que h est $\geqq 1$ et que la démonstration est déjà donnée pour $h-1$ au lieu de h. Si nous posons $m = p^\alpha b$, où $\alpha \geqq 1$ et où l'entier b n'est pas divisible par le nombre premier p, nous obtenons

$$\varphi_h(m) = \sum_{d/bp^\alpha} \mu(d) \log^h d =$$

$$= \sum_{\substack{d_1/b \\ d_2/p^\alpha}} \mu(d_1 d_2)(\log d_1 + \log d_2)^h$$

$$= \sum_{n=0}^{h} \binom{h}{h} \sum_{d_1/b} \mu(d_1) \log^n d_1 \sum_{d_2/p^\alpha} \mu(d_2) \log^{h-n} d_2$$

$$= \sum_{n=0}^{h} \binom{h}{h} \varphi_n(b)\, \varphi_{h-n}(p^\alpha).$$

Puisque m contient plus de h facteurs premiers différents, b contient plus de $h-1$ facteurs premiers différents, de sorte que d'après notre hypothèse de récurrence $\varphi_n(b) = 0$ pour $n = 0, 1, \ldots, h-1$. Le terme restant $\varphi_h(b)\,\varphi_o(p^\alpha)$ s'annule, puisque le dernier facteur le fait, ainsi que nous l'avons déjà démontré. Donc le lemme est démontré.

<u>LEMME 2</u>: Soit $x > 0$ et posons

$$\lambda(d) = \mu(d) \log^2 \frac{x}{d} \quad \text{et} \quad f(m) = \sum_{d/m} \lambda(d)$$

(ces fonctions dépendent naturellement aussi de x). Alors on a

$$f(1) = \log^2 x;$$

$$f(p^\alpha) = -\log^2 p + 2(\log x)(\log p)$$

$$(p \text{ premier}; \ \alpha \geqq 1).$$

$f(p^\alpha q^\beta) = 2(\log p)(\log q)$ (p et q premiers et différents; $\alpha \geqq 1; \beta \geqq 1$).

$f(m) = 0$, si m contient plus de deux facteurs premier différents.

DÉMONSTRATION: La dernière assertion suit immédiatement du lemme précédent (appliqué avec $h = 0$, $h = 1$ et $h = 2$) et le reste est évident.

LEMME 3: Pour $x \geqq 2$ le quotient $\frac{\vartheta(x)}{x}$ est situé entre deux bornes positives fixes. (Théorème de Tchebycheff 1851-1852).

REMARQUE: Pour la démonstration de la relation de Selberg il suffit de savoir que la plus grande limite de $\frac{\vartheta(x)}{x}$ est finie. J'ajoute en même temps la démonstration que la plus petite limite est positive, parceque celle-ci présente une grande ressemblance avec la première et que ce résultat nous servira dans l'application de la formule de Selberg.

1. Démontrons d'abord que la plus grande limite de $\frac{\vartheta(x)}{x}$ est finie. Considérons l'entier $P = \frac{(2n)!}{n!\,n!}$, où $n! = 1\ 2 \ldots n$. Parceque les facteurs premiers $p > n$ et $\leqq 2n$ figurent dans le numérateur et non dans le dénominateur de P, le produit de ces nombres premiers est un division de P, donc au plus égal à $(1+1)^{2n} = 2^{2n}$. Le logarithme néperien de ce produit est égal à $\vartheta(2n) - \vartheta(n)$, d'où il suit

$$\vartheta(2n) - \vartheta(n) \leqq \log 2^{2n} = 2\,n \log 2.$$

Pour $x = 2^m$ (m entier $\geqq 1$) on obtient

$$\vartheta(x) = \{\vartheta(2^m) - \vartheta(2^{m-1})\} + \{\vartheta(2^{m-1}) - \vartheta(2^{m-2})\} + \ldots$$

$$< (2^m + 2^{m-1} + 2^{m-2} + \ldots + 2) \log 2 < 2^{m+1} \log 2 =$$

$$= 2\, x \log 2.$$

Pour $2^{m-1} < x \leqq 2^m$ on obtient donc

$$\vartheta(x) \leqq \vartheta(2^m) < 2^{m+1} \log 2 < 4x \log 2,$$

de sorte qu'on trouve pour chaque nombre $x > 1$

$$\frac{\vartheta(x)}{x} < 4 \log 2.$$

2. Démontrons ensuite que la plus petite limite de $\frac{\vartheta(x)}{x}$ est positive.

Le système $1, 2, \ldots, n$ contient $\left[\frac{n}{p}\right]$ multiples du nombre premier p (où $[u]$ désigne le plus grand entier contenu dans u), ensuite $\left[\frac{n}{p^2}\right]$ multiples de p^2, etc., de sorte que $n!$ contient précisément $\left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \ldots$ facteurs p. Par conséquent le nombre des facteurs p figurant dans P est précisément égal à

$$Q = \sum \left\{\left[\frac{2n}{p^\alpha}\right] - 2\left[\frac{n}{p^\alpha}\right]\right\},$$

où la somme est étendue aux nombres naturels α tels que $p^\alpha \leqq 2n$; en effet, les termes suivants s'annulent. Par conséquent le nombre des termes de Q est $\leqq \left[\frac{\log 2n}{\log p}\right]$.

$[2y] - 2[y]$ est une fonction de y de période 1 qui s'annule dans l'intervalle $0 \leqq y < 1/2$ et qui prend la valeur 1 dans l'intervalle $1/2 \leqq y < 1$. Par conséquent $[2y] - 2[y]$ est toujours $\leqq 1$, d'où il suit $Q \leqq \left[\frac{\log 2n}{\log p}\right]$, de sorte que P est un diviseur de l'entier U défini par

$$\log U = \sum_{p \leqq 2n} \left[\frac{\log 2n}{\log p}\right] \log p .$$

En vertu de

$$P = \frac{(n+1)(n+2)\ldots(2n)}{1.\ 2\ \ldots\ n} \geqq 2^n,$$

on obtient

$$n \log 2 \leqq \log P \leqq \log U = \sum_{p \leqq 2n} \left[\frac{\log 2n}{\log p}\right] \log p$$

On a

$$\left[\frac{\log 2n}{\log p}\right] \leqq \frac{\log 2n}{\log p} \text{ quelque soit } p,$$

$$\left[\frac{\log 2n}{\log p}\right] = 1 \text{ pour } p > \sqrt{2n},$$

donc

$$n \log 2 \leqq \sum_{p \leqq \sqrt{2n}} \log 2n + \sum_{p \leqq 2n} \log p$$

$$\leqq (\log 2n)\sqrt{2n} + \vartheta(2n)$$

Ainsi on trouve

$$\vartheta(2n) \geqq n \log 2 - (\log 2n)\sqrt{2n} \geqq 1/2(n+1) \log 2,$$

si n est suffisamment grand. Si x est suffisamment grand et si nous posons $2n \leqq x < 2n + 2$, nous obtenons

$$\vartheta(x) \geqq \vartheta(2n) \geqq 1/2(n+1) \log 2 > 1/4\ x \log 2,$$

d'où il suit que $\frac{\vartheta(x)}{x}$ possède pour $x > 2$ une borne inférieure positive.

REMARQUE: Le raisonnement précédent conduit encore à une autre formule dont nous n'aurons pas besoin pour la démonstration de la formule de Selberg, mais qui nous servira dans l'application de cette formule. Nous avons constaté que n! contient le facteur premier p exactement $\left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \ldots$ fois, donc

$$\log n! = \sum_{p \leqq n} \left(\left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \ldots \right) \log p.$$

En vertu de

$$\sum_{p \leqq n} \left(\left[\frac{n}{p^2}\right] + \left[\frac{n}{p^3}\right] + \ldots \right) \log p < \sum_{p \leqq n} \left(\frac{n}{p^2} + \frac{n}{p^3} \ldots \right) \log p$$

$$= n \sum_{p} \frac{\log p}{p(p-1)}$$

on trouve que log n! est approximativement égal à $\sum_{p \leqq n} \left[\frac{n}{p}\right] \log p$ et que leur différence est au plus de l'ordre de n. La dernière somme est approximativement égale à $\sum_{p \leqq n} \frac{n}{p} \log p$ et la différence

$\sum_{p \leqq n} \frac{n}{p} \log p - \sum_{p \leqq n} \left[\frac{n}{p}\right] \log p$ est au plus

$\sum_{p \leqq n} \log p = \vartheta(n)$, donc, d'après la première assertion du lemme précédent, au plus du même ordre que n. De cette façon on obtient que $\sum_{p \leqq n} \frac{\log p}{p}$ est égal à

$$\frac{1}{n} \log n! = \frac{1}{n} \sum_{h=2}^{n} \log h,$$

à un terme borné près. On a pour tout entier $h \geqq 2$

$$\log h = h \log h - (h-1) \log (h-1) - (h-1) \log\left(1 + \frac{1}{h-1}\right),$$

où le dernier terme est borné d'après la remarque, faite dans l'introduction. Par conséquent

$\sum_{h=2}^{n} \log h$ est approximativement égal à $n \log n$ et la différence est tout au plus du même ordre que n. Il s'ensuit que $\sum_{p \leqq n} \frac{\log p}{p}$ est approximativement égal à $\log n$ et la différence est bornée.

<u>LEMME 4</u>:

$$\frac{1}{x \log x} \sum_{p \leqq x} (\log p)(\log \frac{x}{p}) \to 0,$$

si x croît indéfiniment.

<u>DÉMONSTRATION</u>. Soit ε un nombre positif. On a $\log \frac{x}{p} < \log \frac{1}{\varepsilon}$ pour $p > \varepsilon x$, de sorte que l'expression en question est au plus égale à

$$\frac{1}{x \log x} \sum_{p \leqq \varepsilon x} (\log p)(\log x) +$$

$$+ \frac{1}{x \log x} \sum_{p \leqq x} (\log p)(\log \frac{1}{\varepsilon})$$

$$= \frac{\vartheta(\varepsilon x)}{x} + \frac{\vartheta(x)}{x} \frac{\log \frac{1}{\varepsilon}}{\log x} < c \left\{ \varepsilon + \frac{\log \frac{1}{\varepsilon}}{\log x} \right\},$$

où c est, d'après la première assertion du lemme 3, égal à une constante absolue. Quand ε tend vers zéro assez lentement pourque $\frac{\log \frac{1}{\varepsilon}}{\log x}$ tende vers zéro, nous obtenons le résultat demandé.

<u>LEMME 5</u>: $\frac{1}{x \log x} \sum_{p^{\alpha} \leqq x} \log p \to 0$ pour $x \to \infty$;

la somme étant étendue aux nombres premiers p et aux nombres naturels α tels que $p^{\alpha} \leqq x$.

DÉMONSTRATION: La somme figurant dans l'expression est égale à

$$\vartheta(x) + \vartheta(\sqrt{x}) + \vartheta(\sqrt[3]{x}) + \dots + \vartheta(\sqrt[k]{x}),$$

où k désigne l'entier le plus grand tel que $2^k \leqq x$, donc $k \leqq \frac{\log x}{\log 2}$. La somme est donc au plus $\vartheta(x) + k\vartheta(\sqrt{x})$ et chacun de ces termes, divisé par $x \log x$, tend vers zéro, si x croît indéfiniment.

LEMME 6:

$$\sum_{m \leqq x} f(m) = (\log x)\vartheta(x) + 2\sum_{p \leqq \sqrt{x}} \vartheta\left(\frac{x}{p}\right) \log p + o(x \log x);$$

le dernier terme désigne une fonction de x qui, après division par $x \log x$, tend vers zéro si x croît indéfiniment.

En général $o(g(x))$ désigne une fonction de x telle que $\frac{o(g(x))}{g(x)}$ tende vers zéro si x croît indéfiniment.

DÉMONSTRATION: Il suit du lemme 2 que

$$(3) \quad \sum_{m \leqq x} f(m) = \log^2 x + \sum_{p^\alpha \leqq x} \left\{-\log^2 p + 2\log x \log p\right\} +$$

$$+ 2 \sum_{\substack{p^\alpha q^\beta \leqq x \\ p < q}} (\log p)(\log q).$$

Considérons d'abord le premier terme dans le membre de droite. La contribution des termes avec $\alpha \geqq 2$ est au plus

$$2 \log^2 x \sum_{\substack{p^\alpha \leqq x \\ \alpha \geqq 2}} 1 = 2 \log^2 x \left\{\sqrt{x} + \sqrt[3]{x} + \dots + \sqrt[k]{x}\right\},$$

où k désigne le plus grand entier tel que $2^k \leqq x$,

de sorte que cette contribution est au plus
$2\sqrt{x}\, k \log^2 x \leqq \frac{2\sqrt{x}}{\log 2} \log^3 x = o(x \log x)$.

La contribution des termes avec $\alpha = 1$ est égale à

$$\sum_{p \leqq x} \left\{ -\log^2 p + 2\log x \log p \right\} = (\log x) \sum_{p \leqq x} \log p +$$
$$+ \sum_{p \leqq x} (\log p)(\log \frac{x}{p}) = (\log x)\vartheta(x) + o(x \log x)$$

d'après le lemme 4. Par conséquent le premier terme dans le membre de droite de (3) est égal à $(\log x)\vartheta(x) + o(x \log x)$.

Considérons finalement le second terme de ce membre. La contribution des termes avec $\beta \geqq 2$ et $\alpha \geqq 1$ est d'après le lemme précédent (appliqué avec $\frac{x}{q^\beta}$ au lieu de x) égale à

$$o \sum_{\substack{q^\beta \leqq x \\ \beta \geqq 2}} (\log q) \frac{x}{q^\beta} \log x = o(x \log x) \sum \frac{\log q}{q^\beta};$$

la dernière somme étant étendue à tous les nombres premiers q et à tous les entiers $\beta \geqq 2$ est convergente en vertu de

$$\sum_{\beta = 2}^{\infty} \frac{1}{q^\beta} = \frac{1}{q(q-1)} \leqq \frac{2}{q^2}$$

Par conséquent la contribution des termes avec $\beta \geqq 2$ (et également celle des termes avec $\alpha \geqq 2$) est égale à $o(x \log x)$, de sorte que le second terme dans le membre de droite de (3) est égal à

$$2 \sum_{\substack{pq \leqq x \\ p<q}} (\log p)(\log q) + o(x \log x)$$

Le premier terme est égal à

$$\sum_{pq \leqq x} (\log p)(\log q) - \sum_{p \leqq \sqrt{x}} \log^2 p$$

$$= \sum_{\substack{p \leqq \sqrt{x} \\ pq \leqq x}} (\log p)(\log q) + \sum_{\substack{q \leqq \sqrt{x} \\ pq \leqq x}} (\log p)(\log q) -$$

$$- \sum_{\substack{p \leqq \sqrt{x} \\ q \leqq \sqrt{x}}} (\log p)(\log q) - \sum_{p \leqq \sqrt{x}} \log^2 p$$

$$= \sum_{p \leqq \sqrt{x}} (\log p)\vartheta\left(\frac{x}{p}\right) + \sum_{q \leqq \sqrt{x}} (\log q)\vartheta\left(\frac{x}{q}\right) -$$

$$-\vartheta^2(\sqrt{x}) + o(\sqrt{x} \log^2 x)$$

$$= 2 \sum_{p \leqq \sqrt{x}} (\log p)\vartheta\left(\frac{x}{p}\right) + o(x \log x).$$

Ainsi nous trouvons que les deux termes dans le membre de droite de (3) sont egaux respectivement à $(\log x)\vartheta(x) + o(x \log x)$ et $2 \sum_{p \leqq \sqrt{x}} \vartheta\left(\frac{x}{p}\right)(\log p) +$

$$+ o(x \log x),$$

d'où résulte le lemme.

LEMME 7: On a pour chaque nombre naturel x

$$\left| \sum_{d=1}^{x} \frac{\mu(d)}{d} \right| \leqq 1 .$$

DÉMONSTRATION: En vertu de (2) nous obtenons

$$1 = \sum_{m=1}^{x} \sum_{d/m} \mu(d) = \sum_{d=1}^{x} \mu(d) \sum_{h} 1,$$ où h désigne

les multiples positifs $\leqq x$ de d.

Donc $\sum_{h} 1 = \left[\frac{x}{d}\right]$, d'où $1 = \sum_{d=1}^{x} \mu(d) \left[\frac{x}{d}\right]$

Par conséquent

$$\left| x \sum_{d=1}^{x} \frac{\mu(d)}{d} - 1 \right| = \left| \sum_{d=1}^{x} \mu(d) \left\{ \frac{x}{d} - \left[\frac{x}{d}\right] \right\} \right|$$

$$\leqq \sum_{d=1}^{x} \left| \frac{x}{d} - \left[\frac{x}{d}\right] \right| \leqq x - 1,$$

puisque chaque terme de la dernière somme est $\leqq 1$ et le dernier terme est égal à zéro. Ainsi nous trouvons

$$x \left| \sum_{d=1}^{x} \frac{\mu(d)}{d} \right| \leqq 1 + (x-1) = x,$$ d'où

découle le lemme.

LEMME 8: 1) Il existe une constante c_1 (pour notre but il n'est pas nécessaire de savoir que c'est la constante de Euler) telle que la fonction

$$\varepsilon(y) = \sum_{n \leqq y} \frac{1}{n} - \log y - c_1$$

tende vers zéro (si y croît indéfiniment), de manière assez rapide pourque $\varepsilon(y) \log y$ tende également vers zéro.

2) En outre nous trouverons que

$$\sum_{n \leqq y} \frac{\log n}{n} - \frac{1}{2} \log^2 y$$

tend vers une limite finie.

DÉMONSTRATION: 1) Soit z le plus petit entier > y. On a

$$\log z = \sum_{n=1}^{z-1} \left\{ \log(n+1) - \log n \right\} = \sum_{n=1}^{z-1} \log\left(1+\frac{1}{n}\right)$$

$$= \sum_{n=1}^{z-1} \frac{1}{n} + \sum_{n=1}^{z-1} \delta_n, \text{ où } \delta_n = \log\left(1+\frac{1}{n}\right) - \frac{1}{n}.$$

Comme nous l'avons fait remarquer dans l'introduction, δ_n a une valeur absolue $\leqq \frac{c}{n^{1+\tau}}$, où c et τ désignent des nombres positifs convenablement choisis, indépendants de n.

La contribution à la somme $\sum_{n=z}^{\infty} \delta_n$ des termes avec $2^h \leqq n < 2^{h+1}$ (où h désigne un entier $\geqq 0$) est en valeur absolue tout au plus $\frac{2^h c}{2^{(1+\tau)h}}$, puisque le nombre des termes considérés est égal à 2^h et chaque terme est en valeur absolue $\leqq \frac{c}{2^{(1+\tau)h}}$. Par conséquent

$$\left|\sum_{n=z}^{\infty} \delta_n\right| \leqq \sum_h \frac{c}{2^{\tau h}},$$

étendue aux entiers h tels que $2^{h+1} > z$. La somme de la progression géométrique est $\frac{2^\tau}{2^\tau - 1}$ fois le premier terme qui a l'ordre de grandeur de $\frac{1}{z^\tau}$, donc l'ordre de grandeur de $\frac{1}{y^\tau}$.

Par conséquent

$$(\log y) \sum_{n=z}^{\infty} \delta_n \to 0 \text{ pour } y \to \infty .$$

De cette manière nous trouvons que $\sum_{n=1}^{z-1} \delta_n$ est approximativement égale à

$$-c_1 = \sum_{n=1}^{\infty} \delta_n$$

et que la différence, multipliée par log y, tend vers zéro pour $y \to \infty$. Ainsi nous avons démontré 1.

2) Pour la démonstration de la deuxième assertion nous remplaçons de nouveau $\log(n+1)$ par $\log n + \frac{1}{n} + \delta_n$ et nous obtenons

$$\frac{1}{2}\log^2 z = \frac{1}{2}\sum_{n=1}^{z-1}\left\{\log^2(n+1) - \log^2 n\right\}$$

$$= \sum_{n=1}^{z-1}\frac{\log n}{n} + \sum_{n=1}^{z-1}\left\{\delta_n \log n + \frac{\delta_n}{n} + \frac{1}{2n^2} + \frac{1}{2}\delta_n^2\right\}$$

Puisque δ_n a tout au plus l'ordre de grandeur de $\frac{1}{n^{1+\tau}}$, où τ est positif, la dernière somme se transforme en une série convergente, si z est remplacé par ∞. Si nous désignons la somme de cette série convergente par c_2, nous obtenons que $\sum_{n\leq y}\frac{\log n}{n} + c_2 - \frac{1}{2}\log^2 z$, donc aussi

$$\sum_{n\leq y}\frac{\log n}{n} + c_2 - \frac{1}{2}\log^2 y$$

tend vers zéro pour $y \to \infty$.

<u>LEMME 9</u>: Si $\tau(n)$ désigne le nombre des diviseurs de n, on a

$$\sum_{n\leq y}\frac{\tau(n)}{n} = \frac{1}{2}\log^2 y + c_3 \log y + c_4 + o(1),$$

où c_3 et c_4 sont des constantes absolues convenablement choisies; $o(1)$ désigne ici une fonction de y qui tend vers zéro, si y croît indéfiniment.

<u>DÉMONSTRATION</u>: Puisque $\tau(n)$ est égal au nombre des paires de nombres naturels a et b tels que

$ab=n$, la somme $\sum_{n\leq y}\frac{\tau(n)}{n}$ est égale à la somme $\sum\frac{1}{ab}$, étendue aux nombres naturels a et b tels que $ab \leq y$.

Évaluons d'abord la contribution des paires a et b avec $a \leq \sqrt{y}$. Cette contribution est augmentée de la contribution (égale) des paires a et b avec $b \leq \sqrt{y}$, et le résultat obtenu doit être diminué de la contribution des paires a et b tels que $a \leq \sqrt{y}$ et $b \leq \sqrt{y}$.

Appliquons le lemme précédent. La contribution des paires a et b tels que $a \leq \sqrt{y}$ peut être écrite sous la forme

$$\sum_{a\leq\sqrt{y}}\frac{1}{a}\sum_{b\leq\frac{y}{a}}\frac{1}{b} = \sum_{a\leq\sqrt{y}}\frac{1}{a}\left\{\log\frac{y}{a} + c_1 + o\left(\frac{1}{\log y}\right)\right\}$$

$$= (\log y)\sum_{a\leq\sqrt{y}}\frac{1}{a} - \sum_{a\leq\sqrt{y}}\frac{\log a}{a} + c_1\sum_{a\leq\sqrt{y}}\frac{1}{a} + o\left(\frac{1}{\log y}\right)\sum_{a\leq\sqrt{y}}\frac{1}{a}$$

$$= (\log y)\left\{\log\sqrt{y} + c_1 + o\left(\frac{1}{\log y}\right)\right\} - \left\{\frac{1}{2}\log^2\sqrt{y} + c_5 + o(1)\right\} + c_1\left\{\log\sqrt{y} + c_1 + o(1)\right\} + o\left(\frac{1}{\log y}\right).(\log\sqrt{y}) = \frac{3}{8}\log^2 y + c_6\log y + c_7 + o(1).$$

Dans ce raisonnement c_5, c_6 et c_7 désignent des constantes absolues, convenablement choisies.

Finalement la contribution des termes avec $a \leq \sqrt{y}$ et $b \leq \sqrt{y}$ est d'après le lemme précédent égale à

$$\Big(\sum_{a\leq\sqrt{y}} \frac{1}{a}\Big)^2 = (\log\sqrt{y} + c_1 + \varepsilon(\sqrt{y}))^2$$

$$= \frac{1}{4}\log^2 y + c_1 \log y + c_1^2 + \varepsilon(\sqrt{y})\Big\{2\log\sqrt{y} + 2c_1 + + \varepsilon(\sqrt{y})\Big\}$$

$$= \frac{1}{4}\log^2 y + c_1 \log y + c_1^2 + o(1).$$

Ainsi nous avons démontré le lemme en vertu de

$$\frac{3}{8} + \frac{3}{8} - \frac{1}{4} = \frac{1}{2}$$

LEMME 10: On a

$$\sum_{d\leq x} \frac{\mu(d)}{d} \log\frac{x}{d} = o(\log x).$$

DÉMONSTRATION: D'après le lemme 8 le membre de gauche de l'assertion est égal à

$$\sum_{d\leq x} \frac{\mu(d)}{d}\Big\{\sum_{n\leq\frac{x}{d}} \frac{1}{n} + c_1 + \varepsilon\Big(\frac{x}{d}\Big)\Big\}$$

$$= \sum_{m\leq x} \frac{1}{m}\sum_{d|m}\mu(d) + c_1\sum_{d\leq x}\frac{\mu(d)}{d} + \sum_{d\leq x}\frac{\mu(d)}{d}\,\varepsilon\Big(\frac{x}{d}\Big);$$

(nous avons posé $dn=m$). Le premier terme est en vertu de (2) égal à 1 et le second terme est d'après le lemme 7 en valeur absolue $\leq |c_1|$, de sorte qu'il suffit de démontrer

$$\sum_{d\leq x} \frac{1}{d}\Big|\varepsilon\Big(\frac{x}{d}\Big)\Big| = o(\log x)$$

Posons $y = x^{\rho}$. où l'exposant $\rho<1$ dépend de x et tend vers 1 assez lentement pourque $x^{1-\rho}$ croîsse indéfiniment avec x. Pour les nombres $d \leq y$ on a

$\frac{x}{d} \geq x^{1-\rho}$, de sorte que $\varepsilon(\frac{x}{c})$ tend vers zéro et que la contribution de ces nombres d est égale à

$$o \sum_{d \leq x} \frac{1}{d} = o(\log x).$$

La contribution des nombres d tels que $y < d \leq x$ possède tout au plus l'ordre de grandeur de

$$\sum_{y < d \leq x} \frac{1}{d}$$

et cette somme est d'après la première partie du lemme 8 (appliquée deux fois, la second fois avec x au lieu de y) approximativement égale à $\log x - \log y = (1-\rho) \log x$ de telle façon que $\sum_{y<d\leq x} \frac{1}{d} - (1-\rho) \log x$ tende vers zéro. Ainsi on obtient le résultat voulu, puisque $1-\rho$ tend vers zéro.

<u>LEMME 11</u>: On a pour chaque nombre naturel k

$$\sum_{d|k} \mu(d)\ \tau(\frac{k}{d}) = 1.$$

<u>DÉMONSTRATION</u>: En vertu de $\tau(m) = \sum_{d_1 | m} 1$ le membre de gauche de l'assertion peut être écrit sous la forme $\sum_{d|k} \mu(d) \sum_{d_1 | \frac{k}{d}} 1 = \sum_{d_1 | k} \ \sum_{d | \frac{k}{d_1}} \mu(d)$

La contribution des diviseurs $d_1 < k$ de k s'annule d'après (2) et la contribution du diviseur k est égale à 1.

<u>LEMME 12</u>: On a $\sum_{d \leq x} \frac{\mu(d)}{d} \log^2 \frac{x}{d} = 2 \log x + o(\log x)$.

<u>DÉMONSTRATION</u>: D'après le lemme 9 (appliqué avec $y = \frac{x}{d}$) le membre de gauche de la formule à démontrer peut s'écrire sous la forme

$$2 \sum_{d \leqq x} \frac{\mu(d)}{d} \left\{ \sum_{m \leqq \frac{x}{d}} \frac{\tau(m)}{m} - c_3 \log \frac{x}{d} - c_4 \right\} + o(\log x)$$

(Voir la remarque à p. 32). D'après les lemmes 11 et 7 le prémier terme égale $U + o(\log x)$, où

$$U = 2 \sum_{d \leqq x} \frac{\mu(d)}{d} \sum_{m \leqq \frac{x}{d}} \frac{\tau(m)}{m},$$

donc (posons $k = md$)

$$U = 2 \sum_{k \leqq x} \frac{1}{k} \sum_{d/k} \mu(d) \ \tau(\tfrac{k}{d})$$

$$= 2 \sum_{k \leqq x} \frac{1}{k} \qquad \text{d'après le lemme 10}$$

$$= 2 \log x + o(\log x) \text{ d'après le lemme 8.}$$

Après ces considérations préliminaires nous passons à la démonstration de la formule de Selberg:

$$\frac{\vartheta(x)}{x} + \frac{2}{x \log x} \sum_{p \leqq \sqrt{x}} \vartheta(\tfrac{x}{p}) \log p \rightarrow 2 \text{ pour } x \rightarrow \infty$$

D'après le lemme 6 le membre de gauche possède à un terme près qui tend pour $x \rightarrow \infty$ vers zéro, la valeur $\frac{1}{x \log x} \sum_{m \leqq x} f(m)$.

Le problème est donc de démontrer que cette expression tend vers 2. On a d'après la définition de la fonction $f(m)$

$$\sum_{m\leq x} f(m) = \sum_{m\leq x} \sum_{d|m} \lambda(d) = \sum_{d\leq x} \lambda(d) \sum_{h} 1,$$

où h désigne les multiples $\leq$ x de d. Par conséquent $\sum_h 1$ est égal à $\frac{x}{d}$, à un terme près qui possède une valeur absolue ≤ 1. La somme $\sum_{m\leq x} f(m)$ est donc approximativement égale à

$$x \sum_{d\leq x} \frac{\lambda(d)}{d} = x \sum_{d\leq x} \frac{\mu(d)}{d} \log^2 \frac{x}{d}$$

$$= 2\, x \log x + o(x \log x)$$

d'après le lemme précédent, et l'erreur commise est au plus du même ordre de grandeur que

$$\sum_{d\leq x} |\lambda(d)| = \sum_{d\leq x} \log^2 \frac{x}{d},$$

de sorte qu'il suffit de démontrer que cette somme est égale à $o(x \log x)$.

Cela est très simple. La contribution des termes tels que $\frac{x}{2^{k+1}} < d \leq \frac{x}{2^k}$, où k désigne un entier ≥ 0, est au plus $(k+1)^2 \frac{x}{2^k}$, puisque le nombre des termes est au plus égal à $\frac{x}{2^h}$ et que chaque terme est $\leq (k+1)^2$. Par conséquent la somme considérée est inférieure à

$$\sum_{k=0}^{\infty} \frac{(k+1)^2 x}{2^k} = x \sum_{k=0}^{\infty} \frac{(k+1)^2}{2^k}$$

donc au plus de l'ordre de x.

De cette manière la relation fondamentale de Selberg est démontrée.

APPLICATION DE LA FORMULE DE SELBERG.

Dans la partie suivante de cette étude nous emploierons en premier lieu la formule de Selberg

$$\frac{\vartheta(x)}{x} + \frac{2}{x \log x} \sum_{p \leqq \sqrt{x}} \vartheta\left(\frac{x}{p}\right) \log p \to 2,$$

en outre la formule

$$\frac{1}{\log x} \sum_{p \leqq x} \frac{\log p}{p} \longrightarrow 1 \qquad \text{pour } x \to \infty$$

(démontrée dans la remarque adjointe au lemme 3) et finalement le fait, démontré dans le même lemme, que $\frac{\vartheta(x)}{x}$ possède pour $x > 2$ une limite inférieure positive.

Nous obtiendrons même le résultat suivant, qui est plus général que le théorème fondamental de la théorie des nombres premiers, savoir:

Considérons une suite de nombres $p_1, p_2, \ldots,$ tous $\geqq 1$ telle que la somme

$$\vartheta(x) = \sum_{p \leqq x} \log p,$$

étendue à tous les nombres $p \leqq x$ figurant dans la suite $p_1, p_2, \ldots,$ possède les propriétés

(4) $$\frac{\vartheta(x)}{x} + \frac{2}{x \log x} \sum_{p \leqq \sqrt{x}} \vartheta\left(\frac{x}{p}\right) \log p \to 2 \quad \text{pour } x \to \infty,$$

et

(5) $$\liminf_{x \to \infty} \frac{\vartheta(x)}{x} > 0$$

et que

(6) $$\frac{1}{\log x} \sum_{p \leqq x} \frac{\log p}{p} \to 1 \text{ pour } x \to \infty$$

Alors on a $\dfrac{\vartheta(x)}{x} \longrightarrow 1$ pour $x \longrightarrow \infty$.

Strictement la condition (6) est superflue, puisqu'on peut démontrer que (6) est conséquence de (4), mais pour obtenir ce résultat un raisonnement long et compliqué est nécessaire. C'est la raison pour laquelle nous préférons admettre également (6).

Dans la suite nous aurons seulement le droit d'utiliser les relations (4), (5) et (6) et le fait que les nombres p sont tous $\geqq 1$, donc $\log p \geqq 0$. Un trait remarquable de la démonstration suivante est qu'elle est directe et que la condition (5) est seulement employé dans la toute dernière phrase.

La formule (4) de Selberg nous apprend que $\dfrac{\vartheta(x)}{x}$ est borné, de sorte que les limites

$$A = \limsup_{x \to \infty} \frac{\vartheta(x)}{x} \quad \text{et} \quad a = \liminf_{x \to \infty} \frac{\vartheta(x)}{x}$$

existent. On a $0 \leqq a \leqq A$. La formule $\dfrac{\vartheta(x)}{x} \longrightarrow 1$ veut dire que $a = A = 1$.

Commençons avec le résultat partiel suivant:

<u>LEMME 13</u>: On a $A+a = 2$.

<u>DÉMONSTRATION</u>: Il est possible de faire croître x indéfiniment de telle façon que $\dfrac{\vartheta(x)}{x}$ tende vers A. Si δ désigne un nombre positif fixe, on a $\vartheta\left(\frac{x}{p}\right) > (a-\delta)\,\frac{x}{p}$ pour chaque nombre x suffisament grand et pour chaque nombre $p \leqq \sqrt{x}$, donc

$$\frac{2}{x \log x} \sum_{p \leqq \sqrt{x}} \vartheta\left(\frac{x}{p}\right) \log p \geqq \frac{2(a-\delta)}{\log x} \sum_{p \leqq \sqrt{x}} \frac{\log p}{p}$$

Le dernier terme tend vers $a-\delta$ en vertu de (6), de sorte que la formule de Selberg nous apprend

$A + a - \delta \leqq 2$. Ceci est valable pour chaque nombre δ positif fixe, donc $A + a \leqq 2$.

Si l'on change A et a dans ce raisonnement et on remplace δ par $-\delta$ et le signe $>$ par $<$, on trouve $A + a \geqq 2$, donc $A + a = 2$.

Dans le reste de ce raisonnement x croît indéfiniment de telle façon que $\frac{\vartheta(x)}{x}$ tend vers A, et le nombre δ désigne un nombre positif fixe.

LEMME 14: Pour chaque nombre fixe $\lambda > a$ on a

$$\frac{1}{\log x} \sum{}' \frac{\log p}{p} \longrightarrow 0,$$

si $\sum{}'$ est étendu aux nombres $p \leqq x$ tels que $\vartheta(\frac{x}{p}) \geqq \frac{\lambda x}{p}$.

REMARQUE: Le lemme 12 est un théorème de compensation: si $\vartheta(x)$ est grand, $\vartheta(\frac{x}{p})$ est petit pour "presque" chaque nombre $p \leqq x$.

DÉMONSTRATION: On a

$$\sum_{p \leqq x} \vartheta(\tfrac{x}{p}) \log p = \sum_{pq \leqq x} \log p \log q$$

$$= \sum_{p \leqq \sqrt{x}} \vartheta(\tfrac{x}{p}) \log p + \sum_{q \leqq \sqrt{x}} \vartheta(\tfrac{x}{q}) \log q - (\sum_{p \leqq \sqrt{x}} \log p)^2$$

$$= 2 \sum_{p \leqq \sqrt{x}} \vartheta(\tfrac{x}{p}) \log p - (\vartheta(\sqrt{x}))^2,$$ où p et q sont des éléments de la suite $p_1, p_2, \ldots$

Le dernier terme possède au plus l'ordre x en vertu de la formule de Selberg (appliquée avec $\sqrt{x}$ au lieu de x).

Par conséquent on peut écrire la formule de Selberg aussi sous la forme

$$\frac{\vartheta(x)}{x} + \frac{1}{x \log x} \sum_{p \leqq x} \vartheta\left(\frac{x}{p}\right) \log p \to 2.$$

Si $\frac{x}{p}$ est supérieur à une valeur u convenablement choisie dépendant de δ, on a $\vartheta\left(\frac{x}{p}\right) > (a-\delta)\,\frac{x}{p}$.
Il existe un nombre positif b dépendant de u, donc de δ, tel que

$$\vartheta\left(\frac{x}{p}\right) > (a-\delta)\,\frac{x}{p} - b$$

pour tous les p tels que $\frac{x}{p} \leqq u$. Ainsi nous trouvons la dernière inégalité pour tout $p \leqq x$.

Si l'on partage la somme $\sum\limits_{p \leqq x}$ en $\sum_1$ et $\sum_1'$, on obtient

$$\sum_{p \leqq x} \vartheta\left(\frac{x}{p}\right) \log p \geqq \lambda x \sum\nolimits_1 \frac{\log p}{p} + x(a-\delta) \sum\nolimits_1' \frac{\log p}{p} - b \sum_{p \leqq x} \log p$$

$$\geqq (a-\delta)\, x \sum_{p \leqq x} \frac{\log p}{p} + (\lambda - a)\, x \sum\nolimits_1 \frac{\log p}{p} - b\,\vartheta(x).$$

Si l'on substitue ce résultat dans la formule de Selberg, on obtient

$$A + a - \delta + (\lambda - a) \limsup \frac{1}{\log x} \sum\nolimits_1 \frac{\log p}{p} \leqq 2,$$

donc, en vertu de $A + a = 2$,

$$(\lambda - a) \limsup \frac{1}{\log x} \sum\nolimits_1 \frac{\log p}{p} \leqq \delta,$$

d'où suit le lemme, puisque $\lambda - a > 0$ et δ est un nombre arbitraire positif.

<u>LEMME 15</u>: On a pour tout nombre fixe $\mu < A$

$$\frac{1}{\log^2 x} \sum\nolimits_2 \frac{\log p}{p} \cdot \frac{\log q}{q} \to 0,$$

si $\sum_2$ est étendu aux paires de nombres p et q, figurant dans la suite $p_1, p_2, \ldots,$ tels que

$$p \leqq \sqrt{x}; \quad q \leqq \sqrt{\frac{x}{p}}; \; \vartheta\left(\frac{x}{pq}\right) \leqq \frac{\mu x}{pq}.$$

REMARQUE: Le théorème 15 est un théorème de double compensation, Si $\vartheta(x)$ est grand, $\vartheta(\frac{x}{p})$ est petit (d'après le lemme précédent) pour "presque" tout nombre $p \leqq x$, et $\vartheta(\frac{x}{pq})$ est grand pour"presque" chaque paire de nombres $p \leqq \sqrt{x}$ et $q \leqq \sqrt{\frac{x}{p}}$.

DÉMONSTRATION: Si l'on remplace dans la formule de Selberg x par $\frac{x}{p}$, on obtient

$$\vartheta\left(\frac{x}{p}\right) = \frac{2x}{p} + o\left(\frac{x}{p}\right) - \frac{2}{\log \frac{x}{p}} \sum_{q \leqq \sqrt{\frac{x}{p}}} \vartheta\left(\frac{x}{pq}\right) \log q,$$

La substitution de ce résultat dans la formule de Selberg fournit

$$\vartheta(x) = 2x + o(x) - \frac{2x}{\log x} \sum_{p \leqq \sqrt{x}} \{2+o(1)\} \frac{\log p}{p} + \frac{4V}{\log x}$$

où

$$V = \sum_{p \leqq \sqrt{x};\; q \leqq \sqrt{\frac{x}{p}}} \frac{1}{\log \frac{x}{p}} \vartheta\left(\frac{x}{pq}\right) (\log p)(\log q)$$

En vertu de (6) on trouve

$$\sum_{p \leqq \sqrt{x}} \frac{\log p}{p} = \frac{1}{2} \log x + o(\log x), \text{ donc } \vartheta(x) = \frac{4V}{\log x} + o(x$$

Dans chaque terme de la somme V on a

$$p \leqq \sqrt{x} \text{ et } q \leqq \sqrt{\frac{x}{p}}, \text{ donc } pq = p^{\frac{1}{2}}(pq^2)^{\frac{1}{2}} \leqq x^{\frac{3}{4}},$$

de sorte qu'on a $\vartheta(\frac{x}{pq}) < (A+\delta)\frac{x}{pq}$, si x est suffisamment grand. Si l'on partage la somme V en deux sommes $\sum_2$ et $\sum_2'$, on trouve donc

$$V \leq \mu x \sum_2 \frac{1}{\log\frac{x}{p}} \frac{\log p}{p} \cdot \frac{\log q}{q} + (A+\delta)x.$$

$$\sum_2' \frac{1}{\log\frac{x}{p}} \frac{\log p}{p} \cdot \frac{\log q}{q} =$$

$$= (A+\delta)xW - (A+\delta-\mu)x\sum_2 \frac{1}{\log\frac{x}{p}} \cdot \frac{\log p}{p} \cdot \frac{\log q}{q},$$

où

$$W = \sum_{p\leq\sqrt{x},\ q\leq\sqrt{\frac{x}{p}}} \frac{1}{\log\frac{x}{p}} \frac{\log p}{p} \cdot \frac{\log q}{q} =$$

$$= \sum_{p\leq\sqrt{x}} \frac{1}{\log\frac{x}{p}} \cdot \frac{\log p}{p} \sum_{q\leq\sqrt{\frac{x}{p}}} \frac{\log q}{q} =$$

$$= \sum_{p\leq\sqrt{x}} \frac{\log p}{p}\left\{\frac{1}{2} + o(1)\right\} \text{ en vertu de (6)}$$

$$= \frac{1}{4}\log x + o(\log x)$$

Ainsi nous obtenons

$$\vartheta(x) \leq (A+\delta)x - \frac{4}{\log x}(A+\delta-\mu)x\sum_2 \frac{1}{\log\frac{x}{p}} \frac{\log p}{p} \cdot \frac{\log q}{q} + o(x),$$

d'où

$$\frac{4}{\log^2 x}(A-\mu)\sum_2 \frac{\log p}{p} \cdot \frac{\log q}{q} \leq \frac{4}{\log x}(A+\delta-\mu).$$

$$\sum_2 \frac{1}{\log\frac{x}{p}} \frac{\log p}{p} \frac{\log q}{q} \leq A+\delta - \frac{\vartheta(x)}{x} + o(1).$$

Par conséquent

$$4(A-\mu) \text{ lim sup } \frac{1}{\log^2 x} \sum_2 \frac{\log p}{p} \cdot \frac{\log q}{q} \leqq \delta,$$

d'où résulte le lemme, puisque $A-\mu > 0$ et que δ est un nombre positif arbitraire.

<u>FIN DE LA DÉMONSTRATION</u>: Soit σ un nombre positif quelconque tel que $\sigma a < A$ et soit le nombre positif δ si petit que

$$(7) \qquad A - a\sigma \geqq \delta\sigma + 2\delta$$

Considérons la somme

$$S = \sum_3 \frac{\log p}{p} \cdot \frac{\log q}{q} \sum_4 \frac{\log r}{r}$$

$\sum_3$ est étendu aux paires des nombres p et q, figurant dans la suite $p_1, p_2, \ldots$ tels que

$$p \leqq \sqrt{x}; \quad q \leqq \sqrt{\frac{x}{p}}; \quad pq \geqq N; \quad \vartheta\left(\frac{x}{pq}\right) \geqq (A-\delta)\frac{x}{pq},$$

où N désigne un nombre naturel fixe quelconque; $\sum_4$ est étendu aux nombres r figurant dans la suite $p_1, p_2, \ldots$ tels que $\frac{pq}{\sigma} < r \leqq \sigma pq$.

Si $\sigma \leqq 1$, la somme $\sum_4$ est naturellement égale à zéro.

On a pour chaque terme figurant dans $\sum_4$

$$r \leqq \sigma pq = \sigma p^{\frac{1}{2}} (pq^2)^{\frac{1}{2}} \leqq \sigma x^{\frac{1}{4}} x^{\frac{1}{2}} = \sigma x^{\frac{3}{4}} \leqq x$$

si x est suffisamment grand.

Si x est suffisamment grand, tous les termes de $\sum_4$ satisfont à l'inégalité

$$(8) \qquad \vartheta\left(\frac{x}{r}\right) \geqq (a+\delta)\ \frac{x}{r}.$$

Cette inégalité est évidente pour les termes avec $r \leqq pq$, car alors on a

$$\vartheta(\frac{x}{r}) \geqq \vartheta(\frac{x}{pq}) \geqq (A-\delta)\frac{x}{pq} \geqq (A-\delta)\frac{x}{\sigma r} \geqq (a+\delta)\frac{x}{r}$$

en vertu de (7).

Considérons maintenant les termes avec $r > pq$.

Si l'on pose $\frac{x}{r} = u$ et $\frac{x}{pq} = v$, on obtient $u < v \leqq \sigma u$.

Si l'on remplace dans la formule de Selberg

$$(\log x)\,\vartheta(x) + 2\sum_{p \leqq \sqrt{x}} \vartheta(\frac{x}{p}) \log p = 2x \log x + o(x \log x)$$

x par v et par u, on obtient par soustraction

$$(\log v)\,\vartheta(v) - (\log u)\,\vartheta(u) \leqq 2v \log v - 2u \log u + o(u \log u),$$

donc

$$\vartheta(u) \geqq \frac{\log v}{\log u}\vartheta(v) - 2(v-u) + 2v\frac{\log v - \log u}{\log u} + o(u).$$

Dans le deuxième membre le premier terme est au moins $\vartheta(v) \geqq (A-\delta)v$ et le troisième terme est $o(u)$, donc

$$\vartheta(u) \geqq (A-\delta)v - 2(v-u) + o(u) = 2u - (2 - A + \delta)v + o(u)$$

$$\geqq 2u - (a+\delta)\sigma u + o(u) \geqq (a+2\delta)u + o(u)$$

en vertu de $A+a = 2$ et de (7).

Ainsi on trouve, si x est suffisamment grand,

$$\vartheta(\frac{x}{r}) = \vartheta(u) \geqq (a+\delta)u = (a+\delta)\frac{x}{r}$$

de sorte que (8) est valable pour chaque terme de $\sum_4$.

Par conséquent

$$S \leqq \sum_{\substack{r \leqq x \\ \vartheta(\frac{x}{r}) \geqq (a+\delta)\frac{x}{r}}} \frac{\log r}{r} \sum\nolimits_5 \frac{\log p}{p} \frac{\log q}{q},$$

où $\sum_5$ est étendu aux paires p et q telles que $p \leqq \sqrt{x}$; $q \leqq \sqrt{\frac{x}{p}}$; $\frac{r}{\sigma} \leqq pq < \sigma r$.

Ainsi on obtient

$$\sum\nolimits_5 \frac{\log p}{p} \frac{\log q}{q} \leqq \frac{\sigma}{r} \sum_{p \leqq \sqrt{x}} \log p \sum_{q < \frac{\sigma r}{p}} \log q$$

$$\doteq \frac{\sigma}{r} \sum_{p \leqq \sqrt{x}} (\log p)\, \vartheta\left(\frac{\sigma r}{p}\right)$$

$$< c_1 \sum_{p \leqq \sqrt{x}} \frac{\log p}{p} < c_2 \log x.$$

Dans ce raisonnement c_1 et c_2 désignent des constantes positives convenablement choisies.

Par conséquent

$$(9) \qquad S \leqq c_2 \log x \sum_{\substack{r \leqq x \\ \vartheta(\frac{x}{r}) \geqq (a+\delta)\frac{x}{r}}} \frac{\log r}{r} = o(\log^2 x)$$

d'après le lemme 14.

Introduisons maintenant la somme

$$T = \sum_{p \leqq \sqrt{x};\ q \leqq \sqrt{\frac{x}{p}};\ pq \geqq N} \frac{\log p}{p} \cdot \frac{\log q}{q}$$

qui est au moins égale à

$$\sum_{\sqrt{N} \leqq p \leqq \sqrt{x}} \frac{\log p}{p} \cdot \sum_{\sqrt{N} \leqq q \leqq \sqrt[4]{x}} \frac{\log q}{q}$$

Chacun de ces deux facteurs a, d'après (6), ordre de grandeur de log x, donc

$$T > c_3 \log^2 x,$$

où c_3 est un nombre positif indépendent de x.

Si l'on pose

$$T = \sum_3 \frac{\log p}{p} \cdot \frac{\log q}{q} + \sum_3' \frac{\log p}{p} \cdot \frac{\log q}{q},$$ où

chaque terme de la dernière somme satisfait à

$$p \leqq \sqrt{x}; \quad q = \sqrt{\frac{x}{p}}; \quad \vartheta\left(\frac{x}{pq}\right) < (A-\delta)\frac{x}{pq},$$

de sorte que le lemme 15 nous apprend que cette somme est égale à $o(\log^2 x)$. De cette manière nous trouvons $\sum_3 \frac{\log p}{p} \cdot \frac{\log q}{q} > \frac{1}{2} c_3 \log^2 x$,

si x est suffisamment grand.

Pour une valeur fixe de x on considère les paires p et q figurant dans la somme $\sum_3$ pour lesquelles $\sum_4 \frac{\log r}{r}$ prend la valeur minimum μ, où μ dépend alors de x seul.

Donc $S \geqq \mu \sum_3 \frac{\log p}{p} \frac{\log q}{q} > \frac{1}{2} \mu c_3 \log^2 x.$

En comparant ce résultat à (9), nous obtenons

$$(10) \qquad \mu = \sum_4 \frac{\log r}{r} \longrightarrow 0.$$

Par conséquent à tout nombre positif ε et à tout nombre naturel N correspond un nombre $t = pq \geqq N$ tel que

$$\sum_{\frac{t}{\sigma} < r \leqq \sigma t} \frac{\log r}{r} < \varepsilon,$$

donc

$$\frac{1}{\sigma t} \sum_{\frac{t}{\sigma} < r \leqq \sigma t} \log r < \varepsilon,$$

d'où $\vartheta(\sigma t) - \vartheta\left(\frac{t}{\sigma}\right) < \varepsilon \sigma t.$

Si N, donc aussi t est suffisamment grand, on a

$$\vartheta(\sigma t) > (a-\varepsilon)\sigma t \text{ et } \vartheta\left(\frac{t}{\sigma}\right) < (A+\varepsilon)\frac{t}{\sigma},$$

donc $(a-\varepsilon)\sigma - \frac{A+\varepsilon}{\sigma} < \varepsilon\sigma$

Cette inégalité vaut pour chaque nombre positif ε, d'où suit $a\sigma^2 - A \leqq 0$.

Ainsi chaque nombre positif σ tel que $a\sigma < A$ satisfait aussi à l'inégalité $a\sigma^2 \leqq A$. Ce résultat ne fournit rien de nouveau, si $a = C$, mais appliquons maintenant le fait que $\liminf \frac{\vartheta(x)}{x}$ est positif, c'est-à-dire que a est positif. Chaque nombre $\sigma < \frac{A}{a}$ possède la propriété $\sigma^2 < \frac{A}{a}$. Si σ tend vers $\frac{A}{a}$, on obtient $\left(\frac{A}{a}\right)^2 \leqq \frac{A}{a}$, donc $\frac{A}{a} \leqq 1$, d'où suit $A = a = 1$, en vertu de $a \leqq A$ et $A + a = 2$.

<u>Remarque</u> (Voir la première formule à p. 20). En effet, appliquons le lemme 8. La contribution au dernier terme des valeurs $\leqq \frac{x}{\log x}$ de d est en vertu de $\frac{x}{d} \geqq \log x$ égale à $o \sum_{d \leqq x} \frac{1}{d} = o(\log x)$ et la contribution des autres valeurs de d est au plus de l'ordre

$$\sum_{\frac{x}{\log x} < d \leqq x} \frac{1}{d} = \log x - \log \frac{x}{\log x} + o(\log x) =$$

$$= o(\log x).$$