

Le théorème des nombres premiers par Selberg

L'objectif de ce problème est de donner une preuve entièrement élémentaire du théorème des nombres premiers. La preuve est due à Alte Selberg, et le problème est très largement inspiré de l'article original de Selberg "*An elementary proof of the prime number theorem*" (1948).

On définit naturellement les fonctions de comptage des nombres premiers habituelles :

$$\pi(x) = \sum_{p \leq x} 1 \qquad \theta(x) = \sum_{p \leq x} \ln(p)$$

Avec la convention suivante, qu'on conservera tout au long du problème : les sommes et produits indexés par les lettres p, q portent uniquement sur les nombres premiers.

On définit également la fonction arithmétique de Moebius μ par :

$$\begin{aligned} \mu(n) &= 0 \quad \text{si } n \text{ est divisible par un carré parfait autre que } 1 \\ \mu(p_1 \dots p_r) &= (-1)^r \quad \text{si } p_1, \dots, p_r \text{ sont des nombres premiers distincts} \\ \mu(1) &= 1 \end{aligned}$$

On rappelle la relation d'inversion de Moebius, valable pour toute fonction $f : \mathbb{R}_+ \mapsto \mathbb{C}$

$$f(x) = \sum_{1 \leq n \leq x} \mu(n) \sum_{1 \leq m \leq \frac{x}{n}} f\left(\frac{x}{mn}\right)$$

1 Préliminaires

1. a) Soit $\epsilon > 0$. Démontrer l'encadrement suivant, pour tout $x > 1$:

$$(1 - \epsilon) \ln(x) (\pi(x) - \pi(x^{1-\epsilon})) \leq \theta(x) \leq \ln(x) \pi(x)$$

b) En déduire l'équivalent au voisinage de $+\infty$:

$$\theta(x) \sim \pi(x) \ln(x)$$

Ainsi, le théorème des nombres premiers se ramène à démontrer l'équivalent au voisinage de l'infini : $\theta(x) \sim x$.

2. a) Soit $n \in \mathbb{N}^*$. En remarquant que tous les nombres premiers de l'intervalle $[n, 2n]$ divisent le coefficient binomial $\binom{2n}{n}$, démontrer l'inégalité :

$$\theta(2n) - \theta(n) \leq n \ln(4)$$

b) En déduire l'estimation au voisinage de $+\infty$:

$$\theta(x) = O(x)$$

3. On note pour tout $n \in \mathbb{N}^*$, $d_n = \text{ppcm}(1, 2, \dots, n)$.

a) Soit $n \in \mathbb{N}^*$ et P un polynôme à coefficients entiers de degré inférieur ou égal à $n - 1$. Démontrer que :

$$d_n \int_0^1 P(t) dt \in \mathbb{Z}$$

b) Soit $n \in \mathbb{N}^*$. En prenant $P = X^n(1 - X)^n$ dans la question précédente, démontrer que :

$$d_{2n+1} \geq 4^n$$

c) En décomposant d_{2n+1} en facteurs premiers, montrer que $(2n+1)^{\pi(2n+1)} \geq d_{2n+1} \geq 4^n$.

d) Conclure quant à l'estimation asymptotique :

$$x = O(\theta(x))$$

Les trois questions précédentes permettent de définir les nombres strictement positifs suivants :

$$a = \liminf_{x \rightarrow +\infty} \frac{\theta(x)}{x} = \liminf_{x \rightarrow +\infty} \frac{\pi(x) \ln(x)}{x}$$

$$A = \limsup_{x \rightarrow +\infty} \frac{\theta(x)}{x} = \limsup_{x \rightarrow +\infty} \frac{\pi(x) \ln(x)}{x}$$

On note alors qu'on a trivialement $a \leq A$, et que démontrer le théorème des nombres premiers se ramène à démontrer que $a = A = 1$.

On va démontrer une formule asymptotique de plus sur les nombres premiers, qui pourra s'avérer utile au cours de la preuve du théorème. Cette formule est due à Mertens.

4. a) Soit $n \in \mathbb{N}^*$. En partant de la formule de Legendre pour la valuation p -adique de $n!$, démontrer que pour tout nombre premier $p \leq n$, on a :

$$\frac{n}{p} - 1 \leq v_p(n!) \leq \frac{n}{p-1}$$

b) En écrivant que $\ln(n!) = \sum_{p \leq n} \ln(p) v_p(n!)$, en déduire l'estimation asymptotique pour x tendant vers $+\infty$:

$$\sum_{p \leq x} \frac{\ln(p)}{p} = \ln(x) + O(1)$$

C'est la formule de Mertens.

2 La formule de Selberg

L'objectif de cette partie est de démontrer la formule asymptotique suivante, qui sera le point de départ de la preuve du théorème des nombres premiers :

$$\theta(x) \ln(x) + \sum_{p \leq x} \ln(p) \theta\left(\frac{x}{p}\right) = 2x \ln(x) + O(x) \quad (1)$$

2.1 Preuve de la formule de Selberg

La preuve de cette formule asymptotique repose sur l'estimation, de deux manières différentes, de la somme suivante :

$$S(x) = \sum_{1 \leq n \leq x} \sum_{d|n} \mu(d) \ln\left(\frac{x}{d}\right)^2$$

L'intérêt de cette somme $S(x)$ est, comme on va le voir, qu'elle charge en grande majorité les n qui sont des nombres premiers ou des produits de deux nombres premiers.

5. a) Soit $x > 1$ et $n \leq x$. Démontrer la formule suivante :

$$\sum_{d|n} \mu(d) \ln\left(\frac{x}{d}\right)^2 = \begin{cases} \ln(x)^2 & \text{si } n = 1 \\ -\ln(p)^2 + 2 \ln(x) \ln(p) & \text{si } n = p^h \text{ avec } p \text{ premier, } h \geq 1 \\ 2 \ln(p) \ln(q) & \text{si } n = p^h q^k \text{ avec } p, q \text{ premiers distincts et } k, h \geq 1 \\ 0 & \text{sinon} \end{cases}$$

b) Montrer que la contribution des termes qui ne sont pas de la forme pq ou p (p et q premiers) dans la somme $S(x)$ est en $O(x)$, et en déduire l'estimation :

$$S(x) = 2 \ln(x) \sum_{p \leq x} \ln(p) - \sum_{p \leq x} \ln(p)^2 + \sum_{pq \leq x} \ln(p) \ln(q) + O(x)$$

c) Conclure quant à la première estimation de $S(x)$:

$$S(x) = \theta(x) \ln(x) + \sum_{p \leq x} \theta\left(\frac{x}{p}\right) \ln(p) + O(x)$$

On va maintenant passer à la seconde manière d'estimer $S(x)$.

6. a) En appliquant l'inversion de Moebius à la fonction $x \mapsto 1$, démontrer l'estimation asymptotique suivante :

$$\sum_{n \leq x} \frac{\mu(n)}{n} = O(1)$$

b) De la même façon, appliquer l'inversion de Moebius à la fonction $x \mapsto x$ pour démontrer l'estimation :

$$\sum_{n \leq x} \frac{\mu(n)}{n} \ln\left(\frac{x}{n}\right) = O(1)$$

On pourra utiliser la a) ainsi que le développement asymptotique classique $\sum_{n \leq x} \frac{1}{n} = \ln(x) + \gamma + O\left(\frac{1}{x}\right)$ qu'on ne manquera pas de redémontrer.

c) Appliquer cette fois-ci l'inversion de Moebius à la fonction $x \mapsto x \ln(x)$ pour obtenir, par des méthodes similaires, l'estimation asymptotique :

$$\sum_{n \leq x} \frac{\mu(n)}{n} \ln\left(\frac{x}{n}\right)^2 = 2 \ln(x) + O(1)$$

On fera pour cela un bon usage du développement asymptotique $\sum_{n \leq x} \frac{x}{n} \ln\left(\frac{x}{n}\right) = \frac{x}{2} \ln(x)^2 + Ax \ln(x) + Bx + O(\ln(x))$ qu'on démontrera (A et B sont des constantes réelles), ainsi que des questions a) et b) qui précèdent.

Avec ce dernier calcul d'ordre de grandeur, on va pouvoir obtenir une deuxième estimation asymptotique de $S(x)$.

7. a) Soit $x > 1$. Démontrer l'égalité :

$$S(x) = \sum_{1 \leq d \leq x} \mu(d) \ln\left(\frac{x}{d}\right)^2 \left\lfloor \frac{x}{d} \right\rfloor$$

b) En déduire l'estimation :

$$S(x) = x \sum_{1 \leq d \leq x} \frac{\mu(d)}{d} \ln\left(\frac{x}{d}\right)^2 + O(x)$$

c) Avec la 6.c), en déduire la deuxième estimation asymptotique de S : $S(x) = 2x \ln(x) + O(x)$

8. a) Avec les deux estimations asymptotiques de S , démontrer la formule de Selberg :

$$\theta(x) \ln(x) + \sum_{p \leq x} \ln(p) \theta\left(\frac{x}{p}\right) = 2x \ln(x) + O(x) \quad (1)$$

b) Montrer qu'elle peut également s'écrire sous la forme :

$$\theta(x) + \sum_{pq \leq x} \frac{\ln(p) \ln(q)}{\ln(x)} = 2x + O\left(\frac{x}{\ln(x)}\right) \quad (2)$$

2.2 Une première application : $a + A = 2$

9. a) En considérant des grandes valeurs de x pour lesquelles $\theta(x) \leq (A - \epsilon)x$ (avec $\epsilon > 0$) dans la formule de Selberg, et en utilisant la question 4., démontrer que $A + a \leq 2$.

b) Démontrer de même que $a + A \geq 2$.

Ainsi, on a démontré la relation $a + A = 2$. Comme on sait que $a \leq A$, il suffit alors pour conclure que $a = A = 1$ de démontrer par exemple que $A \leq a$, ou bien que $A = 1$ ou encore que $a = 1$.

3 Premières estimations du reste $R(x) = \theta(x) - x$

On va dans cette partie s'intéresser plus explicitement au reste $R(x) = \theta(x) - x$. Pour conclure, il suffit évidemment de démontrer que $R(x) = o(x)$ au voisinage de $+\infty$.

On va commencer par donner deux estimations de $R(x)$ à partir de la formule de Selberg. La première est obtenue en question 10. et la seconde à la question 13.

10. a) Utiliser la formule de Selberg et la formule de Mertens (le résultat de 4.) pour démontrer l'estimation asymptotique :

$$R(x) \ln(x) + \sum_{p \leq x} R\left(\frac{x}{p}\right) \ln(p) = O(x)$$

11. a) En partant de la formule de Mertens (question 4.) démontrer l'estimation :

$$\sum_{pq \leq x} \frac{\ln(p) \ln(q)}{pq} = \frac{1}{2} \ln(x)^2 + O(\ln(x))$$

b) Par une transformation d'Abel (ou une intégration par parties), en déduire :

$$\sum_{pq \leq x} \frac{\ln(p) \ln(q)}{pq \ln(pq)} = \ln(x) + O(\ln(\ln(x)))$$

12. a) À partir de la formule de Selberg sous la forme (2), montrer :

$$\theta(x) + \sum_{pq \leq x} \frac{\ln(p) \ln(q)}{\ln(pq)} = 2x + O\left(\frac{x}{\ln(x)}\right)$$

b) Dédire de a) l'estimation :

$$\sum_{p \leq x} \theta\left(\frac{x}{p}\right) \ln(p) = 2x \ln(x) - \sum_{pq \leq x} \frac{\ln(p) \ln(q)}{\ln(pq)} \theta\left(\frac{x}{pq}\right) + O(x \ln(\ln(x)))$$

On pourra commencer par remplacer brutalement le θ dans le terme de gauche par son développement obtenu en a).

c) Avec b) et la formule de Selberg sous sa forme (1), en déduire :

$$\theta(x) \ln(x) = \sum_{pq \leq x} \frac{\ln(p) \ln(q)}{\ln(pq)} \theta\left(\frac{x}{pq}\right) + O(x \ln(\ln(x)))$$

13. Avec 11. et 12., obtenir la seconde estimation de $R(x)$:

$$R(x) \ln(x) = \sum_{pq \leq x} \frac{\ln(p) \ln(q)}{\ln(pq)} R\left(\frac{x}{pq}\right) + O(x \ln(\ln(x)))$$

En combinant ces deux estimations de R (questions 10. et 13.) on va pouvoir en obtenir une troisième, qui va s'avérer beaucoup plus maniable en ceci qu'elle ne fait plus intervenir les nombres premiers. Voici l'estimation en question, qui sera démontrée à la question 14.c) :

$$|R(x)| \leq \frac{1}{\ln(x)} \sum_{n \leq x} \left| R\left(\frac{x}{n}\right) \right| + O\left(\frac{x \ln(\ln(x))}{\ln(x)}\right) \quad (3)$$

14. a) Avec les deux estimations de $R(x)$ trouvées en 10. et 13., démontrer à l'aide d'une intégration par parties :

$$2R(x) \ln(x) \leq \sum_{n \leq x} \left(\theta(n) + \sum_{pq \leq n} \frac{\ln(p) \ln(q)}{pq} \right) \left(R\left(\frac{x}{n}\right) - R\left(\frac{x}{n+1}\right) \right) + O(x \ln(\ln(x)))$$

b) Utiliser 12.a) pour en déduire :

$$2R(x) \ln(x) \leq 2 \sum_{n \leq x} \left| R\left(\frac{x}{n}\right) \right| + O(x \ln(\ln(x)))$$

c) Conclure la preuve de la formule (3) :

$$|R(x)| \leq \frac{1}{\ln(x)} \sum_{n \leq x} \left| R\left(\frac{x}{n}\right) \right| + O\left(\frac{x \ln(\ln(x))}{\ln(x)}\right) \quad (3)$$

Cette formule sera celle qu'on utilisera dans la preuve du théorème des nombres premiers pour prouver que $R(x) = o(x)$. Son intérêt réside en ceci qu'elle ne fait pas intervenir de somme portant sur les nombres premiers.

4 Une propriété capitale de $R(x)$

L'objectif de cette partie est de démontrer (à partir de la formule obtenue à l'issue de la partie précédente) une propriété de $R(x)$ qui nous permettra de conclure la preuve du théorème des nombres premiers. Cette propriété est la suivante :

($\mathcal{P}$) : Il existe deux constantes $K_1, K_2 > 0$ telles que pour tout $\delta \in]0, 1[$ (δ a vocation à être petit), et tout $x > e^{K_2/\delta}$, l'intervalle $]x, e^{K_1/\delta}x[$ contienne un sous-intervalle de la forme $I =]y, e^{\delta/2}y[$, tel que pour tout $z \in I$ on ait $|R(z)| \leq 4\delta z$.

15. a) En intégrant par parties à partir de la formule de Mertens (question 4.), démontrer l'estimation suivante :

$$\sum_{n \leq x} \frac{R(n)}{n^2} = O(1)$$

b) En déduire l'existence d'une constante absolue $A > 0$ telle que pour tous $x, x' > 0$ avec $x \leq x'$, on ait :

$$\left| \sum_{x \leq n \leq x'} \frac{R(n)}{n^2} \right| < A$$

16. Soient $x, x' > 0$ tels que $x < x'$. On suppose que R ne change pas de signe dans l'intervalle $[x, x']$. Montrer l'existence de $y \in]x, x'[$ tel que :

$$\left| \frac{R(y)}{y} \right| < \frac{K_1}{\ln\left(\frac{x'}{x}\right)}$$

où K_1 est une constante absolue.

b) Montrer que, quitte à augmenter la valeur de K_1 , ceci reste vrai même si R change de signe sur l'intervalle $]x, x'[,$ On pourra remarquer que si R change de signe sur $]x, x'[,$ alors il existe un point y de ce même intervalle tel que $|R(y)| \leq \ln(y).$

c) Soit maintenant $\delta \in]0, 1[$ quelconque. Pour $x > 4$ quelconque, montrer avec ce qui précède l'existence de $y \in]x, e^{K_1/\delta}x[$ tel que :

$$R(y) \leq \delta y$$

17. a) À partir de la formule de Selberg dans sa forme (2), montrer l'existence d'une constante absolue $B > 0$ telle que pour tous $y, y' > 2$ avec $y \leq y'$, on ait :

$$0 \leq \theta(y') - \theta(y) \leq 2(y' - y) + B \frac{y'}{\ln(y')}$$

b) En déduire que pour tous $y, y' > 2$ avec $y \leq y'$, on a :

$$|R(y') - R(y)| \leq y' - y + B \frac{y'}{\ln(y')}$$

18. À l'aide des questions 16. et 17., montrer la propriété (P). On pourra, étant donné un intervalle de la forme $]x, e^{K_1/\delta}x[,$ prendre y comme en 16. et utiliser 17. pour montrer que R reste petit dans un voisinage convenable de $y.$

5 Preuve du théorème des nombres premiers

L'objectif de cette dernière partie est de montrer le théorème des nombres premiers :

$$\pi(x) \sim \frac{x}{\ln(x)}$$

En vertu de tout ce qui a été fait précédemment, ceci se ramène bien entendu à prouver que :

$$R(x) = o(x)$$

Et c'est ce que nous allons faire.

19. Montrer l'existence de :

$$\alpha = \limsup_{x \rightarrow +\infty} \frac{|R(x)|}{x}$$

On fixe maintenant $\epsilon > 0$ (qui a vocation à tendre vers 0) et on pose $\beta = \alpha + \epsilon.$ On fixe également x_0 tel que pour tout $x \geq x_0,$ on ait $|R(x)| \leq \beta x.$

Enfin, on prend $\delta = \frac{\beta}{8} > 0$ qui sera le δ auquel on va appliquer le résultat de la partie 4.

20. a) En utilisant la majoration de R obtenue à l'issue de la partie 3 (formule (3)), démontrer :

$$|R(x)| \leq \frac{x}{\ln(x)} \sum_{1 \leq n \leq \frac{x}{x_0}} \frac{1}{n} \left| \frac{n}{x} R\left(\frac{x}{n}\right) \right| + o(x)$$

b) On note $\rho = e^{K_1/\delta}$ (avec K_1 comme dans la partie 4, et $\delta = \frac{\beta}{8}$). En utilisant le résultat de la partie 4, ainsi que la majoration $R(y) \leq \beta y,$ obtenir la formule asymptotique :

$$|R(x)| \leq \beta \left(1 - \frac{\beta^2}{256K_2} \right) x + o(x)$$

c) En déduire :

$$\alpha \leq \alpha \left(1 - \frac{\alpha^2}{256K_2} \right)$$

d) Conclure.