

Travail de maturité - Mathématiques

PREUVES AUTOUR
DE L'EXISTENCE D'UNE INFINITÉ
DE NOMBRES PREMIERS
ET FONCTIONS D'APPROXIMATION

CHARLOTTE HEYD

MENTOR : ETIENNE GREZET

Lycée Blaise Cendrars
La Chaux-de-Fonds
24 janvier 2024

Table des matières

1	Introduction	1
2	Différentes preuves du fait qu'il existe une infinité de nombres premiers	3
2.1	Preuve historique d'Euclide	3
2.2	Variante de la preuve d'Euclide pour une progression arithmétique	4
2.2.1	Introduction au modulo	4
2.2.2	Preuve qu'il existe une infinité de nombres premiers congrus à 3 (mod 4)	6
2.3	Preuve par le produit eulérien	7
2.3.1	Preuve de la divergence de la série harmonique par Nicolas Oresme	7
2.3.2	Produit eulérien	9
2.4	Preuve de la divergence de la série des inverses des nombres premiers par Paul Erdős	12
3	Fonction de répartition et ses approximations	17
3.1	Fonction de compte : $\pi(x)$	17
3.2	Fonction d'approximation : $f(x) = \frac{x}{\ln(x)}$	19
3.2.1	Comparaison de $f(x) = \frac{x}{\ln(x)}$ avec $\pi(x)$	20
3.3	Fonction d'approximation : le logarithme intégral	22
3.3.1	Intégration numérique	23
3.3.2	Intégration par parties	24
3.3.3	Comparaison de $\text{li}(x)$ avec $\pi(x)$	27
3.4	Comparaison de $f(x) = \frac{x}{\ln(x)}$ et de $\text{li}(x)$ avec $\pi(x)$	29
4	Conclusion	34
5	Remerciements	36
6	Références	37
7	Annexes	38

1 Introduction

La première fois que j'ai entendu parler des nombres premiers, c'était lors de ma 9^e harmonie, en cours de mathématiques. L'exercice était le suivant : nous avions devant nous un tableau avec les nombres de 2 à 100, et nous devions tracer tous les multiples de 2, sauf 2 lui-même. Puis nous passions au prochain nombre non-tracé qui était 3, et nous tracions à nouveau tous les multiples de 3, sauf 3 lui-même. Nous devions continuer ainsi pour chaque nombre non-tracé jusqu'à 10. C'est ainsi que s'est dressée devant nous la liste des nombres premiers jusqu'à 100. Notre enseignant nous a alors expliqué que tous les nombres que nous n'avions pas tracés avaient exactement deux diviseurs, à savoir 1 et eux-mêmes, et que c'était ce que l'on appelait des nombres premiers. Nous ne le savions pas, mais ce que nous venions de faire n'était autre que le crible d'Eratosthène, un moyen efficace de tester la primalité de petits nombres. Je me souviens avoir dit à ma voisine de table que je trouvais cela intéressant, et j'ai gardé ceci dans un coin de ma tête.

Bien plus tard, j'ai regardé un documentaire qui parlait des nombres premiers et d'une hypothèse que le mathématicien Riemann a formulé à leur propos. Cette hypothèse n'a pas encore été prouvée et elle fait partie des problèmes du millénaire. J'ai été très impressionnée par le fait que des nombres qui présentaient une propriété aussi "simple" que de posséder exactement deux diviseurs puissent mener à des théorèmes aussi complexes et donner autant de fil à retordre aux mathématiciens.

Puis est venu le moment pour moi de devoir choisir un sujet pour mon travail de maturité. Il me paraissait évident que j'allais le faire en mathématiques, mais je ne savais pas exactement quel genre de sujets se traitaient dans un tel travail. J'ai donc pris connaissance de la liste des sujets proposée par mon enseignant et quand j'ai vu la catégorie "théorie des nombres", j'ai immédiatement pensé aux nombres premiers.

Une fois mon sujet accepté, il était temps de définir plus précisément ce à quoi j'allais m'intéresser dans mon travail. En repensant à l'exercice que j'avais dû faire à l'école secondaire, j'ai réalisé que ce qui m'intéresserait le plus serait d'élargir l'étude de l'apparition des nombres premiers à une bien plus grande échelle que celle à laquelle nous sommes habitués. En effet, les premiers nombres premiers sont bien connus.

Voici un tableau qui regroupe les nombres premiers jusqu'à 500 :

2	3	5	7	11	13	17	19	23	29	31	37	41
43	47	53	59	61	67	71	73	79	83	89	97	101
103	107	109	113	127	131	137	139	149	151	157	163	167
173	179	181	191	193	197	199	211	223	227	229	233	239
241	251	257	263	269	271	277	281	283	293	307	311	313
317	331	337	347	349	353	359	367	373	379	383	389	397
401	409	419	421	431	433	439	443	449	457	461	463	467
479	487	491	499									

TABLE 1 – Nombres premiers jusqu'à 500

Mais au fait, jusqu'où pourrions-nous étendre ce tableau ? Pour répondre à cette question, il est nécessaire de s'en poser d'abord une autre : existe-t-il un nombre fini ou un nombre infini de nombres premiers ?

Cette question sera donc le fil rouge de la première partie de mon travail. En effet, nous allons nous intéresser à plusieurs preuves du fait qu'il existe une infinité de nombres premiers. Nous débuterons par la preuve historique d'Euclide qui est une démonstration par l'absurde. Puis, nous ferons une variante de cette même preuve pour une progression arithmétique. La preuve d'après s'appuiera sur la divergence de la série harmonique ainsi que sur le produit eulérien de la fonction zêta de Riemann. Nous finirons avec une preuve de Paul Erdős concernant la divergence de la série des inverses des nombres premiers ; une preuve par l'absurde qui utilise un peu d'analyse combinatoire. Ces quatre preuves démontreront certes la même chose, mais elles requièrent toutes des outils mathématiques différents et variés. C'est assez passionnant de se rendre compte que l'on peut prouver un même théorème en empruntant plusieurs chemins bien distincts.

À ce stade-là, nous serons normalement convaincus du fait qu'il existe une infinité de nombres premiers. Mais une autre question se posera : à quelle fréquence est-ce que les nombres premiers apparaissent et comment sont-ils répartis ? En effet, on pourrait instinctivement penser que les nombres premiers se raréfient à mesure qu'on les cherche dans de grands nombres... Mais est-ce vraiment le cas, et si oui dans quelle mesure ? C'est là-dessus que portera la seconde partie de mon travail. Il existe une fonction qui "compte" les nombres premiers que l'on nomme $\pi(x)$. Cependant, elle devient difficilement manipulable lorsque l'on travaille avec de grands nombres. Nous allons donc également nous intéresser à deux fonctions d'approximation de $\pi(x)$: la fonction $f(x) = \frac{x}{\ln(x)}$ et le logarithme intégral. En regardant comment ces deux fonctions d'approximation se comportent et en les comparant avec $\pi(x)$, nous devrions être en mesure d'y voir plus clair et de nous faire ainsi une meilleure idée de la répartition des nombres premiers.

2 Différentes preuves du fait qu'il existe une infinité de nombres premiers

Dans cette première partie, nous allons tenter de prouver que dresser une liste exhaustive des nombres premiers est impossible. En d'autres termes, on aimerait prouver que l'on trouvera toujours des nombres premiers, même au sein de nombres gigantesques. Nous allons donc nous intéresser à quatre démonstrations employant toutes différents outils. Normalement, nous devrions parvenir à la même conclusion après les avoir toutes explorées.

2.1 Preuve historique d'Euclide

Cette démonstration s'appuie sur une preuve d'Euclide du fait qu'il existe une infinité de nombres premiers. C'est une preuve par l'absurde (démontrer qu'une affirmation est vraie en prouvant que son contraire est faux). Ici, il s'agit de prétendre qu'il existe un nombre fini de nombres premiers et de prouver que cette hypothèse débouche sur une contradiction.

Théorème 1

Il existe une infinité de nombres premiers.

Démonstration.

Supposons qu'il existe un nombre fini de nombres premiers que l'on note : $P_1, P_2, P_3, \dots, P_n$. Posons $M = \left(\prod_{i=1}^n P_i\right) + 1$, le produit des n nombres premiers additionné à 1.

Selon l'hypothèse de départ, M doit être décomposable en facteurs premiers listés plus haut. Cependant, on peut montrer que M n'est pas divisible par le nombre premier P_j , ceci $\forall j$:

$$\frac{M}{P_j} = \frac{\left(\prod_{i=1}^n P_i\right) + 1}{P_j} = \frac{\prod_{i=1}^n P_i}{P_j} + \frac{1}{P_j} = \left(\prod_{i=1, i \neq j}^n P_i\right) + \frac{1}{P_j}$$

En effet, comme $\frac{1}{P_j} \notin \mathbb{N}$, on a que $P_j \nmid M$

Ainsi, soit M est lui-même un nombre premier, soit il existe un facteur premier P non-listé qui lui, divise M .

Dans tous les cas, nous venons de démontrer que notre affirmation initiale est fausse. La liste des nombres premiers n'est pas finie et il existe donc une infinité de nombres premiers.

□

2.2 Variante de la preuve d'Euclide pour une progression arithmétique

Une progression arithmétique est une suite de nombres dont chacun des termes U_n est calculé en additionnant la raison r de la suite au terme U_{n-1} :

$$\begin{cases} U_0 = a \\ U_n = U_{n-1} + r \end{cases}$$

Exemples :

1, 2, 3, 4, 5, 6, ..., n est une suite arithmétique de raison 1 et de terme initial 1. On additionne 1 à chacun des termes pour trouver le suivant en partant de $U_0 = 1$

3, 7, 11, 15, 19, 23, ..., $4n - 1$ est une suite arithmétique de raison 4 et de terme initial 3.

Ce dernier exemple n'est pas choisi au hasard. Nous pouvons constater que bon nombre des termes de la suite arithmétique de raison 4 et de terme initial $U_0 = 3$ sont premiers. C'est donc cette progression arithmétique-là qui va nous intéresser afin d'établir une variante de la preuve historique d'Euclide.

2.2.1 Introduction au modulo

Pour mieux aborder la prochaine preuve qui se basera sur la progression arithmétique citée dans l'exemple précédent, il est nécessaire que nous introduisions la notion de modulo, plus précisément celle du modulo 4.

Si deux nombres entiers x et y ont un même reste positif lorsqu'on les divise respectivement par un même entier naturel n (division euclidienne), on dit alors que x et y sont **congrus modulo n** .

Ceci se note : $x \equiv y \pmod{n}$

Exemples :

$19 \equiv 10 \equiv 1 \pmod{9}$, car le reste de la division de ces nombres par 9 est 1.

$11 \equiv 7 \equiv 3 \pmod{4}$, car le reste de la division de ces nombres par 4 est 3.

Ainsi, nous pouvons utiliser la congruence modulo pour former ce que l'on appelle des classes d'équivalence. Dans notre cas, x et y sont liés par une relation d'équivalence s'ils ont le même reste de la division par n .

On note : $x \sim y$ si $x \equiv y \pmod{n}$

De cette manière, nous pouvons décider de travailler avec un seul représentant pour chaque classe, puisque les autres éléments qui y sont contenus sont équivalents à ce représentant.

Exemple :

Nous pourrions imaginer une classe d'équivalence qui regrouperait tous les nombres impairs, autrement dit tous les nombres congrus à 1 $\pmod{2}$.

On noterait : $\bar{1} = \{1, 3, 5, 7, 9, 11, \dots\}$

Le représentant de cette classe serait 1, et nous ne travaillerions donc qu'avec ce chiffre malgré que $\bar{1}$ contienne une infinité d'éléments.

On pourrait faire de même avec les nombres pairs, congrus eux à 0 $\pmod{2}$. On au-

rait $\bar{0} = \{0, 2, 4, 6, 8, 10, \dots\}$. Le représentant de cette classe serait 0. Sur l'ensemble des nombres naturels, nous ne travaillerions donc qu'avec 0 et 1. Dans cet exemple, nous avons en fait affaire au même principe que celui du système binaire.

Nous pouvons regrouper les classes d'équivalence en un ensemble $\overline{\mathbb{Z}_n} = \{\bar{0}, \bar{1}, \bar{2}, \dots, \overline{n-1}\}$. Comme seul le représentant de la classe suffit, nous pouvons travailler plus simplement avec $\mathbb{Z}_n = \{0, 1, 2, \dots, n-1\}$.

Mais avant de se plonger dans cette prochaine preuve, il est nécessaire de nous pencher sur les tables de Cayley de la congruence modulo 4 :

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

TABLE 2 – Table de Cayley pour \mathbb{Z}_4 avec addition

Cette première table montre donc la congruence modulo 4 du résultat de l'addition de la colonne de gauche à la ligne du haut.

Exemples :

$1 + 1 = 2 \equiv 2 \pmod{4}$ → l'intersection de la ligne 1 et de la colonne 1 est 2

$3 + 2 = 5 \equiv 1 \pmod{4}$ → l'intersection de la ligne 3 et de la colonne 2 est 1

·	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

TABLE 3 – Table de Cayley pour \mathbb{Z}_4 avec multiplication

Cette table respecte le même schéma que la précédente, mais elle montre cette fois-ci la congruence modulo 4 du résultat de la multiplication de la colonne de gauche avec la ligne du haut. Dans le théorème 2, nous allons nous intéresser aux nombres premiers congrus à 3 (mod 4).

Remarque 1

Nous remarquons, grâce à la table 3, que $3 \cdot 3 \equiv 1 \pmod{4}$, mais aussi que $3 \cdot 1 \equiv 3 \pmod{4}$. Ainsi, le produit d'un nombre pair de nombres congrus à 3 (mod 4) est congru à 1 (mod 4), et le produit d'un nombre impair de nombres congrus à 3 (mod 4) est lui aussi congru à 3 (mod 4). Donc si un produit est congru à 3 (mod 4), alors l'un de ses termes au moins est congru à 3 (mod 4).

Maintenant que nous avons clairement défini le modulo, revenons-en à la preuve historique d'Euclide que nous allons essayer d'appliquer à $\mathbb{Z}_4 = \{0, 1, 2, 3\}$.

2.2.2 Preuve qu'il existe une infinité de nombres premiers congrus à 3 (mod 4)

Cette preuve s'appuie donc sur le même principe que la précédente, sauf que nous resserons désormais notre champ de recherche en nous intéressant uniquement aux nombres premiers congrus à 3 (mod 4).

Théorème 2

Il existe une infinité de nombres premiers congrus à 3 (mod 4).

Démonstration.

Supposons qu'il existe un nombre fini de nombres premiers congrus à 3 (mod 4) que l'on note : $P_1, P_2, P_3, \dots, P_n$. Tout comme pour la preuve précédente, nous allons définir deux nombres M et M' tous deux congrus à 3 (mod 4). En effet, il existe deux cas : soit il y a un nombre pair de P_i (nous utiliserons alors M), soit il y a un nombre impair de P_i (nous utiliserons alors M').

$$M = \left(\prod_{i=1}^n P_i \right) + 2$$

$$M' = \left(\prod_{i=1}^n P_i \right) + 4$$

À la remarque 1, nous avons vu que le produit d'un nombre pair de P_i congrus à 3 (mod 4) est congru à 1 (mod 4) et que le produit d'un nombre impair de P_i congrus à 3 (mod 4) est congru à 3 (mod 4). Ainsi, en additionnant 2 au premier produit et 4 au second, M et M' sont tous deux congrus à 3 (mod 4).

Si la liste des P_i est finie, alors M et M' devraient être divisibles par au moins l'un des P_i (c.f. remarque 1). Mais tout comme pour la démonstration du théorème 1, nous pouvons montrer que ceci n'est pas possible :

$$\frac{M}{P_j} = \frac{\left(\prod_{i=1}^n P_i \right) + 2}{P_j} = \frac{\prod_{i=1}^n P_i}{P_j} + \frac{2}{P_j} = \left(\prod_{i=1, i \neq j}^n P_i \right) + \frac{2}{P_j}$$

$$\frac{M'}{P_j} = \frac{\left(\prod_{i=1}^n P_i \right) + 4}{P_j} = \frac{\prod_{i=1}^n P_i}{P_j} + \frac{4}{P_j} = \left(\prod_{i=1, i \neq j}^n P_i \right) + \frac{4}{P_j}$$

À nouveau, comme $\frac{2}{P_j} \notin \mathbb{N}$ et $\frac{4}{P_j} \notin \mathbb{N}$, on a que $P_j \nmid M$ et que $P_j \nmid M'$.

M et M' ne sont donc divisibles par aucun des P_i , mais ils doivent être décomposables en facteurs premiers, ce qui signifie qu'il existe des nombres premiers Q_j et Q'_j tels que :

$$M = \prod_{j=1}^m Q_j$$

$$M' = \prod_{j=1}^m Q'_j$$

Un produit de nombres non congrus à 3 (mod 4) n'est jamais congru à 3 (mod 4). Comme M et M' sont tous deux congrus à 3 (mod 4), l'un des Q_j et l'un des Q'_j qui les divisent sont eux forcément congrus à 3 (mod 4). Ces derniers ne sont donc pas listés, mais bien premiers et congrus à 3 (mod 4).

On se rend compte qu'il est impossible de faire une liste de nombres premiers congrus à 3 (mod 4) qui soit finie et cela prouve bien qu'il en existe une infinité. □

2.3 Preuve par le produit eulérien

Le produit eulérien de la fonction zêta de Riemann, que nous définirons plus tard, peut constituer une preuve du fait qu'il existe une infinité de nombres premiers. Mais la démonstration faite par Euler exige l'utilisation de la série harmonique et du fait qu'elle diverge. C'est pourquoi nous allons d'abord prouver ceci avant de nous lancer dans le produit eulérien.

2.3.1 Preuve de la divergence de la série harmonique par Nicolas Oresme

Nous devons cette preuve de la divergence de la série harmonique à Nicolas Oresme, un mathématicien français du Moyen-Âge.

Considérons la série harmonique :

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \dots$$

Afin de prouver qu'elle diverge, nous allons montrer qu'elle n'est pas majorée. Vu que la série harmonique est croissante, cela prouvera bien qu'elle est divergente.

Théorème 3

La série harmonique $\sum_{k=1}^{\infty} \frac{1}{k}$ diverge.

Démonstration.

Remplaçons chaque terme de la série harmonique par la plus petite puissance de $\frac{1}{2}$ inférieure ou égale à ce terme comme suit :

$$\begin{aligned} \sum_{k=1}^{\infty} \frac{1}{k} &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \dots \\ S &= 1 + \sum_{t=1}^{\infty} \frac{1}{2^t} = 1 + \frac{1}{2} + \overbrace{\frac{1}{4} + \frac{1}{4}}^{\frac{1}{2}} + \overbrace{\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}}^{\frac{1}{2}} + \dots \end{aligned}$$

Nous pouvons constater que les termes de la série S s'additionnent facilement pour former des groupes de $\frac{1}{2}$. Le deuxième terme forme le premier groupe de $\frac{1}{2}$ que l'on note $n = 1$, le groupe $n = 2$ est la somme des troisième et quatrième termes, $n = 3$ la somme des cinquième, sixième, septième et huitième termes, etc.

Regroupons maintenant les termes de la série harmonique de la même manière que pour notre série S :

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{\sum_{k=3}^4 \frac{1}{k}} + \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}}_{\sum_{k=5}^8 \frac{1}{k}} + \dots$$

Nous pouvons en déduire que chaque groupe s'exprime comme suit :

$$\sum_{k=2^{n-1}+1}^{2^n} \frac{1}{k}$$

La première étape de notre preuve consiste à prouver que le terme général de la série harmonique est plus grand que celui de la série S , donc que :

$$\sum_{k=2^{n-1}+1}^{2^n} \frac{1}{k} \geq \frac{1}{2}$$

Comme, si $2^n \geq k$ alors $\frac{1}{k} \geq \frac{1}{2^n}$, on peut voir que :

$$\begin{aligned} \sum_{k=2^{n-1}+1}^{2^n} \frac{1}{k} &\geq \sum_{k=2^{n-1}+1}^{2^n} \frac{1}{2^n} \\ &= \frac{1}{2^n} \cdot \sum_{k=2^{n-1}+1}^{2^n} 1 = \frac{1}{2^n} \cdot \overbrace{(2^n - 2^{n-1})}^{\text{nombre de termes}} = \frac{1}{2^n} \cdot 2^{n-1} \cdot (2 - 1) = \frac{1}{2^n} \cdot 2^{n-1} = \frac{1}{2} \end{aligned}$$

Même si c'était visuellement évident, nous venons de prouver que, $\forall n \geq 1$:

$$\sum_{k=2^{n-1}+1}^{2^n} \frac{1}{k} \geq \frac{1}{2}$$

Il nous reste maintenant à prouver que la série diverge, donc qu'elle n'est pas bornée.

On veut démontrer que $\forall M \in \mathbb{N}$:

$$\sum_{k=1}^{\infty} \frac{1}{k} \geq M$$

Pour faciliter les calculs, nous sortons le premier terme des sommes suivantes :

$$1 + \sum_{k=2}^{\infty} \frac{1}{k} > 1 + \sum_{k=2}^{2^{2M}} \frac{1}{k} = 1 + \sum_{n=1}^{2M} \left(\sum_{k=2^{n-1}+1}^{2^n} \frac{1}{k} \right)$$

Comme nous avons prouvé que $\sum_{k=2^{n-1}+1}^{2^n} \frac{1}{k} \geq \frac{1}{2}$, nous pouvons utiliser ce résultat pour poursuivre notre chaîne d'inégalité en notant que :

$$1 + \sum_{n=1}^{2M} \left(\sum_{k=2^{n-1}+1}^{2^n} \frac{1}{k} \right) \geq 1 + \sum_{n=1}^{2M} \frac{1}{2} = 1 + \frac{1}{2} \cdot 2M = M + 1 > M$$

Ainsi, comme $\sum_{k=1}^{\infty} \frac{1}{k} > M$, la série harmonique n'est pas majorée et elle est donc bien divergente, vu qu'elle est croissante.

□

2.3.2 Produit eulérien

Un produit eulérien est le produit d'un nombre infini de facteurs indexés par les nombres premiers. Il a été nommé ainsi grâce aux travaux d'Euler autour de la fonction zêta. Cette fonction est définie comme :

$$\zeta(s) = \sum_{q=1}^{\infty} \frac{1}{q^s} \quad (s \in \mathbb{R})$$

En remplaçant s par 1, la fonction zêta est équivalente à la série harmonique.

Donc, $\zeta(1) = \sum_{q=1}^{\infty} \frac{1}{q}$. C'est bien ce cas que nous allons traiter.

Soit P_i le i -ième nombre premier, Euler a montré que :

$$\zeta(1) = \prod_{i=1}^{\infty} \frac{1}{1 - \frac{1}{P_i}}$$

Remarque 2

Avant d'entamer la prochaine preuve, il est nécessaire de faire un petit point sur les séries géométriques

Une série géométrique est une somme de termes dont chacun est construit en multipliant le terme précédent par un même nombre r appelé raison de la suite.

Le terme général d'une telle série est donc :

$$u_n = a \cdot r^n, \quad a = u_0$$

Un exemple simple de série géométrique serait : $2 + 4 + 8 + 16 + 32 + \dots$

Ici, $a = 2$ et $r = 2$.

La somme partielle d'une série géométrique est donnée par :

$$S_n = \frac{a(1 - r^{n+1})}{1 - r}$$

Si nous voulons un nombre infini de termes et que $|r| < 1$, alors r^{n+1} tend vers 0, ce qui signifie que la série est convergente et que sa limite est la suivante :

$$\sum_{n=0}^{\infty} a \cdot r^n = \frac{a}{1 - r}$$

Nous allons maintenant démontrer l'affirmation d'Euler nous-même de deux manières différentes.

Théorème 4

Le produit eulérien de $\zeta(1)$ est égal à $\prod_{i=1}^{\infty} \frac{1}{1 - \frac{1}{P_i}}$

Démonstration.

Soit P_i un nombre premier. Comme $\forall i$ on a $P_i > 1$, alors $|\frac{1}{P_i}| < 1$, ce qui fait que l'on peut reprendre la définition de la limite d'une série géométrique pour $|r| < 1$ vue à la remarque 2 et écrire :

$$\frac{1}{1 - \frac{1}{P_i}} = \frac{a}{1 - r} = \sum_{n=0}^{\infty} a \cdot r^n$$

Ici, $a = 1$ et $r = \frac{1}{P_i}$, on peut donc poursuivre en écrivant que :

$$\frac{1}{1 - \frac{1}{P_i}} = \sum_{n=0}^{\infty} \left(\frac{1}{P_i}\right)^n = \sum_{n=0}^{\infty} \frac{1}{P_i^n}$$

Ainsi, comme $P_1 = 2$ et $P_2 = 3$, on a :

$$\frac{1}{1 - \frac{1}{2}} = \sum_{n=0}^{\infty} \frac{1}{2^n} = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots$$

$$\frac{1}{1 - \frac{1}{3}} = \sum_{n=0}^{\infty} \frac{1}{3^n} = 1 + \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \dots$$

Multiplions maintenant ces deux sommes :

$$\left(\frac{1}{1 - \frac{1}{2}}\right) \cdot \left(\frac{1}{1 - \frac{1}{3}}\right) = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{6} + \frac{1}{8} + \frac{1}{9} + \dots$$

En multipliant les deux premiers facteurs du produit eulérien, on obtient donc une somme de termes $\frac{1}{q}$, où q est décomposable en facteurs premiers P_1 et P_2 , à savoir 2 et 3.

Chaque $\frac{1}{q}$ peut donc s'écrire comme suit :

$$\frac{1}{q} = \frac{1}{P_1^{k_1}} \cdot \frac{1}{P_2^{k_2}}$$

k_m est ici le nombre de facteurs P_m qui décomposent q .

Exemples :

$$\frac{1}{6} = \frac{1}{2^1} \cdot \frac{1}{3^1}$$

$$\frac{1}{8} = \frac{1}{2^3} \cdot \frac{1}{3^0}$$

Maintenant que l'on a montré comment se comportait $\prod_{i=1}^2 \frac{1}{1 - \frac{1}{P_i}}$, on peut étendre ceci à un nombre infini de nombres premiers :

$$\prod_{i=1}^{\infty} \frac{1}{1 - \frac{1}{P_i}} = \left(\frac{1}{1 - \frac{1}{2}}\right) \cdot \left(\frac{1}{1 - \frac{1}{3}}\right) \cdot \dots \cdot \left(\frac{1}{1 - \frac{1}{P_m}}\right) \cdot \dots = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \dots$$

Chaque terme s'écrit donc :

$$\frac{1}{q} = \frac{1}{P_1^{k_1}} \cdot \frac{1}{P_2^{k_2}} \cdot \dots \cdot \frac{1}{P_m^{k_m}} \cdot \dots$$

Vu que la décomposition d'un nombre en facteurs premiers est unique et que l'on multiplie une infinité de $\frac{1}{1 - \frac{1}{P_i}}$ différents entre eux, la somme des $\frac{1}{q}$ qui résulte de ce produit emploie forcément q allant de 1 à l'infini.

Le produit eulérien est donc bien égal à la série harmonique qui est égale à $\zeta(1)$:

$$\zeta(1) = \sum_{q=1}^{\infty} \frac{1}{q} = \prod_{i=1}^{\infty} \frac{1}{1 - \frac{1}{P_i}}$$

□

Procédons maintenant à une deuxième manière de démontrer cette égalité. Nous allons cette fois-ci partir de $\zeta(1)$ pour arriver au produit eulérien, contrairement à la première démonstration.

Démonstration.

Considérons la fonction zêta de 1 et $P_1, P_2, \dots, P_i, \dots$ la liste croissante des nombres premiers :

$$\zeta(1) = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots$$

Divisons cette somme par $\frac{1}{P_1}$, à savoir $\frac{1}{2}$, pour créer une nouvelle somme :

$$\frac{1}{2} \cdot \zeta(1) = \frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \frac{1}{8} + \frac{1}{10} + \dots$$

Soustrayons ces deux sommes :

$$\zeta(1) - \frac{1}{2} \cdot \zeta(1) = 1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{9} + \dots$$

Mettons $\zeta(1)$ en évidence :

$$\left(1 - \frac{1}{2}\right) \cdot \zeta(1) = 1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{9} + \dots$$

On a donc enlevé du membre de droite, les termes $\frac{1}{q}$ où q est multiple de 2.

Répetons maintenant l'opération avec $\frac{1}{P_2}$, afin d'enlever du membre de droite les termes $\frac{1}{q}$ où q est multiple de 3 :

$$\frac{1}{3} \cdot \left(1 - \frac{1}{2}\right) \cdot \zeta(1) = \frac{1}{3} + \frac{1}{9} + \frac{1}{15} + \frac{1}{21} + \frac{1}{27} + \dots$$

$$\left(1 - \frac{1}{2}\right) \cdot \zeta(1) - \frac{1}{3} \cdot \left(1 - \frac{1}{2}\right) \cdot \zeta(1) = 1 + \frac{1}{5} + \frac{1}{7} + \frac{1}{11} + \frac{1}{13} + \dots$$

$$\left(1 - \frac{1}{3}\right) \cdot \left(1 - \frac{1}{2}\right) \cdot \zeta(1) = 1 + \frac{1}{5} + \frac{1}{7} + \frac{1}{11} + \frac{1}{13} + \dots$$

En procédant ainsi pour chaque P_i , on élimine du membre de droite tous les termes $\frac{1}{q}$ où q est multiple de P_i . En faisant cela une infinité de fois, il ne reste plus que 1 dans le membre de droite et on obtient :

$$\left(1 - \frac{1}{2}\right) \cdot \left(1 - \frac{1}{3}\right) \cdot \dots \cdot \left(1 - \frac{1}{P_m}\right) \cdot \dots \cdot \zeta(1) = 1$$

Donc :

$$\zeta(1) = \frac{1}{\left(1 - \frac{1}{2}\right) \cdot \left(1 - \frac{1}{3}\right) \cdot \dots \cdot \left(1 - \frac{1}{P_m}\right) \cdot \dots}$$

Ceci s'écrit :

$$\zeta(1) = \frac{1}{\prod_{i=1}^{\infty} \left(1 - \frac{1}{P_i}\right)}$$

Mais on peut aussi l'écrire :

$$\zeta(1) = \prod_{i=1}^{\infty} \frac{1}{1 - \frac{1}{P_i}}$$

On a donc bien l'égalité recherchée. □

Ce que nous venons de démontrer constitue bien une preuve du fait qu'il existe un nombre infini de nombres premiers. En effet, la série harmonique (ou $\zeta(1)$) diverge et est égale à un produit indexé par les nombres premiers. Il est impossible que ce même produit possède un nombre de facteurs fini puisqu'il est égal à une somme infinie (vu que la série harmonique diverge). Comme chaque facteur du produit eulérien est associé à un nombre premier différent et que le produit eulérien contient une infinité de facteurs, il y a bien une infinité de nombres premiers.

2.4 Preuve de la divergence de la série des inverses des nombres premiers par Paul Erdős

Cette preuve de la divergence de la série des inverses des nombres premiers a été formulée par Paul Erdős, un mathématicien hongrois du XX^e siècle.

Considérons la série des inverses des nombres premiers :

$$\sum_{i=1}^{\infty} \frac{1}{P_i} = \frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{11} + \frac{1}{13} + \frac{1}{17} + \frac{1}{19} + \dots$$

Nous allons prouver que cette série diverge, autrement dit que :

$$\sum_{i=1}^{\infty} \frac{1}{P_i} = +\infty$$

En effet, si cette série diverge, cela veut dire qu'elle comporte une infinité de termes, et qu'il y a donc une infinité de nombres premiers.

Définition 1

Définissons la fonction $N_m(x)$ comme le cardinal de l'ensemble des nombres naturels n inférieurs à x , tel que n ne soit divisible par aucun des P_i , ceci pour tout i supérieur à m .
 $N_m(x) = |\{n \in \mathbb{N} \mid n < x, P_i \nmid n \forall i > m\}|$

Faisons un exemple pour $N_m(x)$ avec $m = 4$ et $x = 25$.

Exemple :

Les P_i où i est supérieur à 4 sont donc $P_5 = 11, P_6 = 13, P_7 = 17, P_8 = 19, P_9 = 23$. On s'arrête à 23 étant donné que les nombres premiers suivants ne diviseront de toute manière aucun $n < 25$.

On a donc :

$$\begin{aligned} N_4(25) &= |\{n \in \mathbb{N} \mid n < 25, P_5, P_6, P_7, P_8, P_9 \nmid n\}| \\ &= |\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 15, 16, 18, 20, 21, 24\}| = 19 \end{aligned}$$

Lemme 1

$$N_m(x) \leq 2^m \cdot \sqrt{x}$$

Pour prouver ceci, nous allons recourir à un peu d'analyse combinatoire afin de déterminer combien d'éléments contient $N(x)$.

Démonstration.

Posons $n = k \cdot r^2$, où $k \in \mathbb{N}$ est sans facteur carré, c'est-à-dire que $k = P_1^{k_1} \cdot P_2^{k_2} \cdot \dots \cdot P_m^{k_m}$, avec $k_i \in \{0; 1\}$. Chaque k_i est donc soit 0 soit 1.

Comme le nombre d'arrangements possibles de 2 chiffres dans m cases est égal à 2^m , on en déduit que le nombre de k possibles est de 2^m .

De plus, on a que $x > n = k \cdot r^2 \geq r^2$, donc $r^2 \leq x$, et finalement $r \leq \sqrt{x}$.

En multipliant le nombre de k et de r différents, on obtient le maximum de n différents contenus dans $N_m(x)$.

Au final, on a donc bien : $N_m(x) \leq 2^m \cdot \sqrt{x}$.

□

Exemple :

Reprenons notre exemple précédent :

$$N_4(25) = |\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 15, 16, 18, 20, 21, 24\}| = 19$$

Comme attendu, notre lemme se vérifie ; on a bien $19 \leq 2^4 \cdot \sqrt{25} = 80$.

Lemme 2

$$x - N_m(x) \leq \sum_{i=m+1}^{\infty} \frac{x}{P_i}$$

Démonstration.

En écrivant $x - N_m(x)$, on obtient le cardinal de tous les entiers ne figurant pas dans $N_m(x)$, autrement dit tous les $n < x$ pour lesquels il existe un $i > m$ tel que P_i divise n :

$$x - N_m(x) = |\{n \in \mathbb{N} \mid n < x, \exists i > m, P_i \mid n\}|$$

D'autre part, écrivons la somme suivante :

$$\sum_{i=m+1}^{\infty} |\{n \in \mathbb{N} \mid n < x, P_i \mid n\}|$$

On somme ici pour chaque $i > m$, le nombre de $n < x$ divisibles par P_i . Nous remarquons que certains n sont potentiellement comptés plusieurs fois, puisque plusieurs P_i peuvent diviser n . Or, ceci n'est pas le cas pour $x - N_m(x)$.

On peut donc écrire :

$$x - N_m(x) \leq \sum_{i=m+1}^{\infty} |\{n \in \mathbb{N} \mid n < x, P_i \mid n\}|$$

Comme le nombre de $n < x$ divisibles par P_i est au maximum de $\frac{x}{P_i}$, on peut finalement écrire :

$$x - N_m(x) \leq \sum_{i=m+1}^{\infty} \frac{x}{P_i}$$

□

Exemple :

Illustrons encore une fois ce lemme par le même exemple.

$$x - N_4(25) = |\{11, 13, 17, 19, 22, 23\}| = 6$$

On a bien :

$$6 < \sum_{i=5}^{\infty} \frac{x}{P_i} = \overbrace{\frac{25}{11} + \frac{25}{13} + \frac{25}{17} + \frac{25}{19}}^{\sim 6,982} + \frac{25}{23} + \dots$$

Théorème 5

La série des inverses des nombres premiers diverge, autrement dit $\sum_{i=1}^{\infty} \frac{1}{P_i} = +\infty$

Démonstration.

Par l'absurde, supposons que cette série converge vers un nombre $\alpha \in \mathbb{R}$, donc que :

$$\sum_{i=1}^{\infty} \frac{1}{P_i} = \alpha$$

Si cette série est convergente, alors elle se "stabilise" en quelque sorte à mesure que ses termes s'additionnent. Pour tout $\varepsilon > 0$, il existe un M -ième terme à partir duquel la somme partielle des termes suivants est inférieure à ε . Donc, la valeur absolue de la différence entre la somme des M premiers termes de la série et sa limite α est strictement inférieure à ε .

Récapitulons ceci et définissons une valeur pour ε de la manière suivante :

Pour $\varepsilon = \frac{1}{4}$, $\exists M \in \mathbb{N}$ tel que :

$$\left| \sum_{i=1}^n \frac{1}{P_i} - \alpha \right| < \frac{1}{4} = \varepsilon, \forall n \geq M$$

Ceci revient au même que d'écrire que la valeur absolue de la somme partielle des termes pour i allant de $M + 1$ jusqu'à $+\infty$ est strictement plus petite que 2ε , puisque notre somme pourrait théoriquement varier de $\alpha - \varepsilon$ à $\alpha + \varepsilon$, ce qui fait bien 2ε en tout.

Donc, on a :

$$\left| \sum_{i=M+1}^{\infty} \frac{1}{P_i} \right| < \frac{1}{2} = 2\varepsilon$$

On peut se passer des valeurs absolues étant donné que notre somme est strictement positive :

$$\sum_{i=M+1}^{\infty} \frac{1}{P_i} < \frac{1}{2}$$

Considérons maintenant $N_M(x) = |\{n \in \mathbb{N} \mid n < x, P_i \nmid n \forall i > M\}|$.

Le lemme 2 nous montre que $x - N_M(x) \leq \sum_{i=M+1}^{\infty} \frac{x}{P_i}$. Mais maintenant que nous savons

que $\sum_{i=M+1}^{\infty} \frac{1}{P_i} < \frac{1}{2}$, nous pouvons poursuivre et écrire :

$$x - N_M(x) \leq \sum_{i=M+1}^{\infty} \frac{x}{P_i} = x \sum_{i=M+1}^{\infty} \frac{1}{P_i} < x \cdot \frac{1}{2} = \frac{x}{2}$$

Donc, on a :

$$x - N_M(x) < \frac{x}{2} \implies x - \frac{x}{2} < N_M(x) \implies \frac{x}{2} < N_M(x)$$

Comme le lemme 1 nous montre que $N_M(x) \leq 2^M \cdot \sqrt{x}$, on a pour finir que :

$$\frac{x}{2} < N_M(x) \leq 2^M \cdot \sqrt{x}$$

Donc :

$$\frac{x}{2} < 2^M \cdot \sqrt{x}$$

C'est là qu'apparaît notre contradiction. On peut vérifier avec $x = 2^{2M+2}$ que $\frac{x}{2}$ n'est pas strictement inférieur à $2^M \cdot \sqrt{x}$.

En effet, on a d'une part :

$$\frac{x}{2} = \frac{2^{2M+2}}{2} = 2^{2M+1}$$

et d'une autre :

$$2^M \cdot \sqrt{x} = 2^M \cdot \sqrt{2^{2M+2}} = 2^M \cdot 2^{M+1} = 2^{2M+1}$$

Maintenant que l'on est face à une contradiction, on peut donc affirmer que la série des inverses des nombres premiers ne converge pas vers α , et donc qu'elle diverge.

□

La preuve de la divergence de la série des inverses des nombres premiers constitue notre dernière preuve du fait qu'il existe un nombre infini de nombres premiers. En effet, il paraît évident que comme elle diverge, elle est composée d'une infinité de termes qui sont chacun l'inverse d'un nombre premier différent. On en arrive donc encore une fois à la conclusion qu'il existe une infinité de nombres premiers.

3 Fonction de répartition et ses approximations

La fréquence à laquelle apparaissent les nombres premiers est quelque peu mystérieuse. Nous n'avons pas de fonction qui nous renvoie directement le nombre de nombres premiers inférieurs ou égaux à x sans devoir littéralement les compter en vérifiant la primalité de chacun d'eux. Le seul moyen d'être précis en terme de répartition de ces nombres est donc bien de les compter, d'où la fonction de compte des nombres premiers : $\pi(x)$.

Cette fonction est définie comme : $\pi(x) = |\{p \in \mathbb{P} | p \leq x\}|$, où \mathbb{P} est l'ensemble des nombres premiers et x un nombre réel. En d'autres termes, l'image de la fonction est égale au nombre de nombres premiers inférieurs ou égaux à x .

Pour calculer un certain nombre de valeurs de $\pi(x)$, il est utile de créer un programme sur un ordinateur qui calcule de manière très machinale le nombre de nombres premiers inférieurs ou égaux à x . Mais cette méthode informatique a ses limites ; en effet, elle repose sur le fait de tester chaque nombre afin de déterminer s'il est premier ou non. Ainsi, plus x est grand, plus le temps de calcul de $\pi(x)$ augmente ; il est donc difficile de travailler avec des grandes valeurs.

C'est pourquoi il est intéressant de se pencher également sur des fonctions d'approximation de la répartition exacte des nombres premiers, dont nous pouvons évaluer l'erreur.

Ici, nous allons nous intéresser aux deux fonctions suivantes :

$$f(x) = \frac{x}{\ln(x)} \qquad \text{li}(x) = \int_0^x \frac{1}{\ln(t)} dt$$

Nous verrons qu'elles sont équivalentes à $\pi(x)$ quand x tend vers l'infini.

Notons que si nous avons une fonction d'approximation parfaite (ou plutôt exacte) de $\pi(x)$, celle-ci en deviendrait aussi un test de primalité : si $f(x) = y$ et que $f(x+1) = y+1$, alors $x+1$ est premier, et ce $\forall x$.

3.1 Fonction de compte : $\pi(x)$

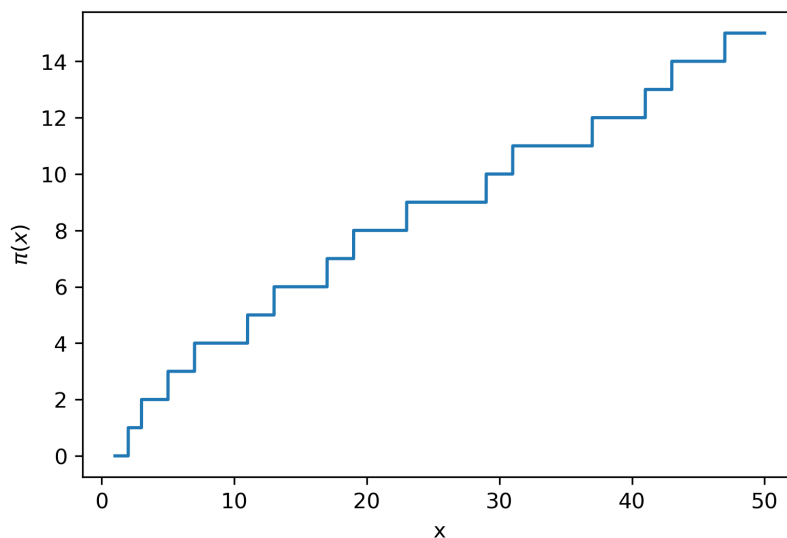


FIGURE 1 – Graphe de $\pi(x)$ pour x allant de 1 à 50

Ce premier graphe de $\pi(x)$ pour x allant de 1 à 50 se base sur un programme *Python*. Ce dernier repose principalement sur deux boucles "while" et une liste de nombres premiers qui s'allonge petit à petit. L'une des deux boucles fait office de test de primalité; elle va tester si un nombre k se divise par les nombres se trouvant dans la liste des nombres premiers inférieurs à k , mais seulement pour ceux allant de 2 à \sqrt{k} (c.f. remarque 3). Si après avoir testé tout cela, il s'avère que k ne se divise par aucun des facteurs premiers inférieurs ou égaux à \sqrt{k} qui se trouvent dans la liste, cela veut dire que k est lui-même premier : le programme ajoute donc k dans la liste. Évidemment, dès que k est divisible par l'un des éléments de la liste, on sort de la boucle sans tester les autres éléments de la liste pour limiter le temps de calcul au maximum. Cette boucle "while" est contenue dans une autre boucle "while", qui s'occupe elle de faire varier k de 1 à n (n étant la plus grande valeur de x souhaitée), afin de parcourir l'intégralité des coordonnées x du graphe. Les coordonnées de chaque point du graphe sont donc (k ; longueur de la liste des nombres premiers pour k en question).

Remarque 3

Lorsque l'on trouve un diviseur a de n , on trouve forcément un diviseur $b = \frac{n}{a}$.

On peut alors écrire que :

$$n = a \cdot b$$

Comme $n = \sqrt{n} \cdot \sqrt{n}$, alors sans perdre de généralité, on a : $a \leq \sqrt{n}$ et $b \geq \sqrt{n}$.

Étant donné qu'à part 1 et lui-même, un nombre non-premier a forcément deux diviseurs a et b , et que a est plus petit ou égal à \sqrt{n} , un test allant de 2 à \sqrt{n} pour chercher un éventuel diviseur de n suffit pour déterminer si n est premier.

Remarque 4

Notons que la fonction $\pi(x)$ est discontinue mais que nous avons représenté le graphe ci-dessus en escalier pour une meilleure visualisation.

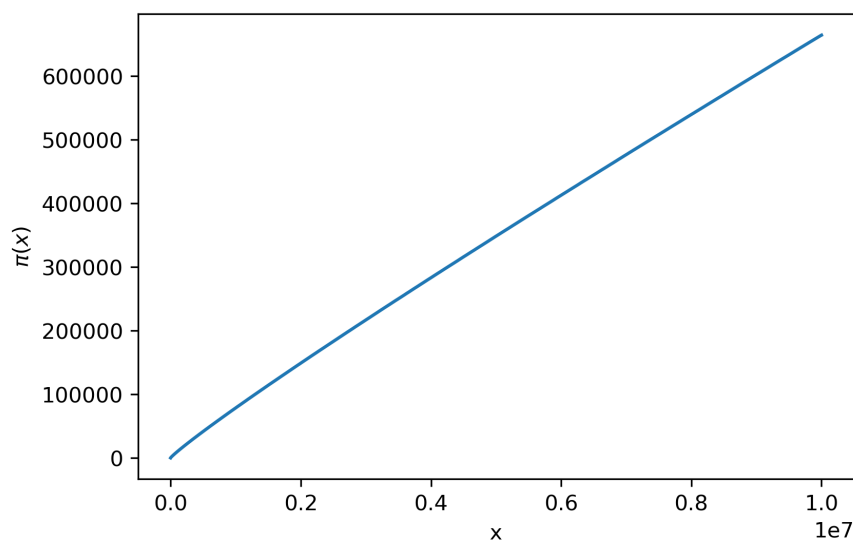


FIGURE 2 – Graphe de $\pi(x)$ jusqu'à $x = 10^7$

Si nous entrons un x plus grand dans notre programme, le graphe qui en ressort porterait à croire qu'il s'agisse d'une fonction linéaire. Nous ne voyons plus l'escalier du graphe de la fonction, mais ce dernier est bien entendu toujours présent.

Cependant, comme mentionné plus haut, le temps de calcul de ce graphe a considérablement augmenté par rapport à celui de $\pi(50)$: pour $x = 10^7$, il a fallu une dizaine de minutes, comparé à moins d'une seconde pour $x = 50$.

3.2 Fonction d'approximation : $f(x) = \frac{x}{\ln(x)}$

Vers la fin du XVIII^e siècle, Gauss et Legendre ont postulé que la fonction $f(x) = \frac{x}{\ln(x)}$ était très proche de la fonction de compte $\pi(x)$. Cela a été prouvé plus tard, en 1896, par Hadamard et La Vallée Poussin grâce à la fonction zêta de Riemann. Nous ne traiterons pas de cette preuve qui est quelque peu complexe, mais nous contenterons d'énoncer le théorème auquel elle a donné suite ; le théorème des nombres premiers.

Selon ce théorème, $\pi(x)$ est équivalente à $\frac{x}{\ln(x)}$ quand x tend vers l'infini :

Théorème 6

$$\pi(x) \underset{x \rightarrow \infty}{\sim} \frac{x}{\ln(x)}$$

C'est-à-dire :

$$\lim_{x \rightarrow +\infty} \frac{\pi(x)}{\frac{x}{\ln(x)}} = 1 \implies \lim_{x \rightarrow +\infty} \pi(x) \cdot \frac{\ln(x)}{x} = 1$$

À présent, regardons le graphe de $f(x) = \frac{x}{\ln(x)}$ pour x allant de 2 à 10^7 :

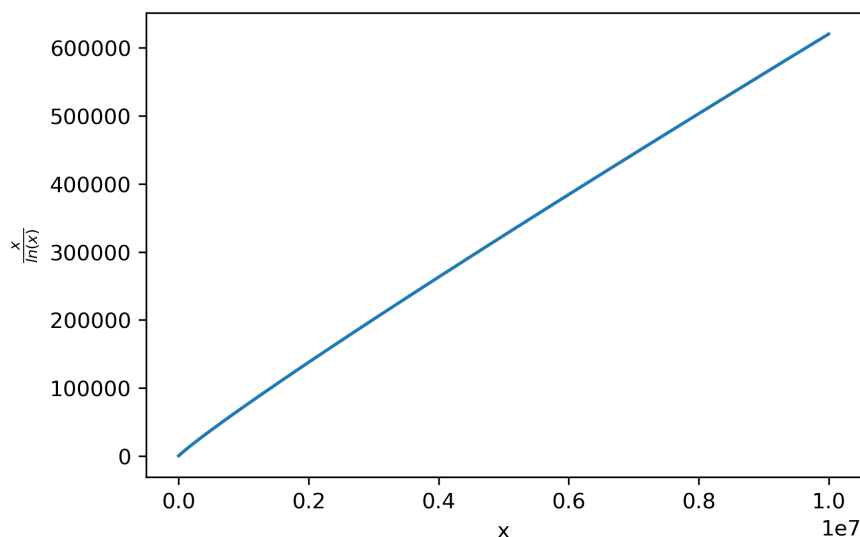


FIGURE 3 – Graphe de $f(x) = \frac{x}{\ln(x)}$ jusqu'à $x = 10^7$

En effet, ce graphe paraît plutôt semblable à celui de $\pi(x)$. Mais comparons-les plus en détail.

3.2.1 Comparaison de $f(x) = \frac{x}{\ln(x)}$ avec $\pi(x)$

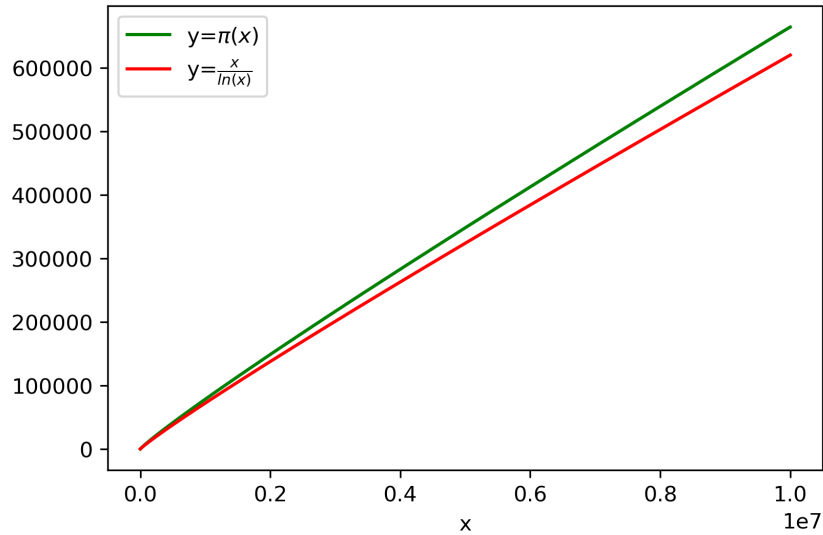


FIGURE 4 – Graphe de comparaison de $f(x) = \frac{x}{\ln(x)}$ et de $\pi(x)$ jusqu'à $x = 10^7$

À première vue, les deux fonctions semblent évoluer de manière plus ou moins similaire. Néanmoins, on constate que même si elles progressent d'une façon comparable, un écart non négligeable se creuse entre les deux courbes et atteint une valeur d'environ 50'000 pour $x = 10^7$.

Afin de mieux se représenter ce que signifie cet écart, il est judicieux d'observer les graphes de l'erreur absolue ϵ (écart brut entre les deux fonctions) et de l'erreur relative σ (quotient de l'écart entre les deux fonctions par la valeur de la fonction de compte)

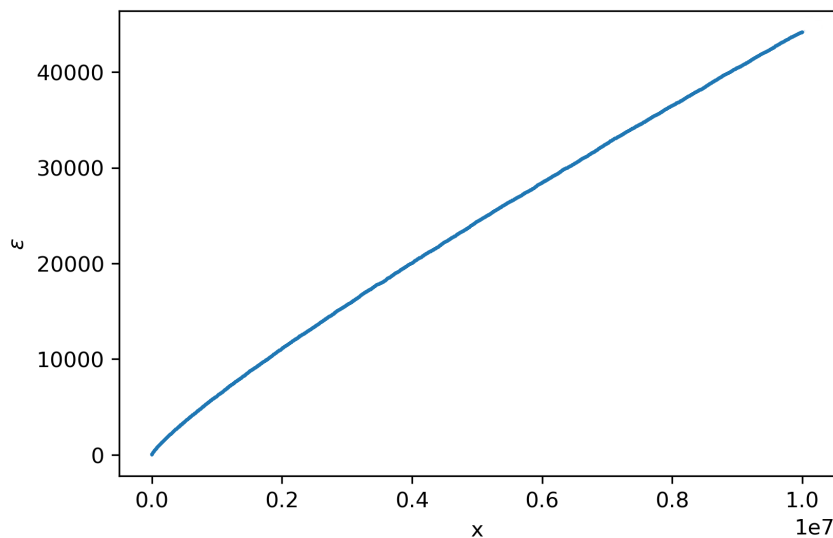


FIGURE 5 – Graphe de l'erreur absolue de $f(x) = \frac{x}{\ln(x)}$ par rapport à $\pi(x)$

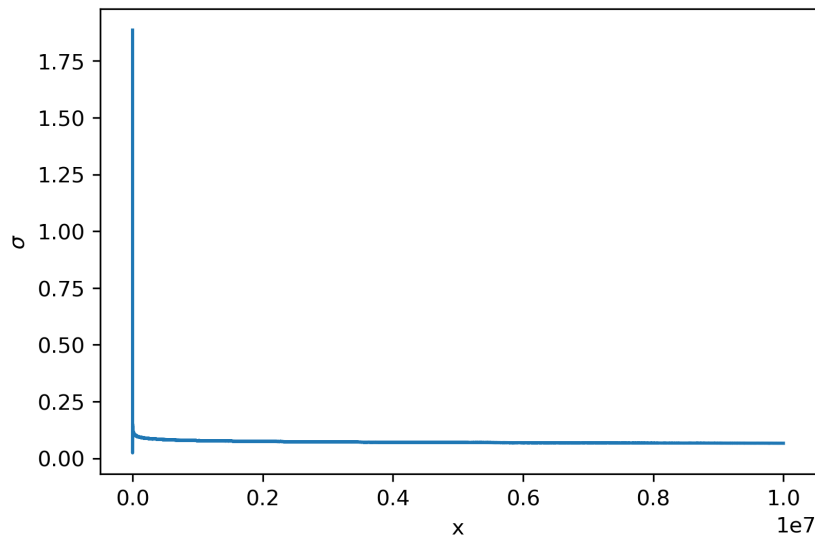


FIGURE 6 – Graphe de l’erreur relative de $f(x) = \frac{x}{\ln(x)}$ par rapport à $\pi(x)$

Si le premier graphe nous montre que l’écart entre $\pi(x)$ et $\frac{x}{\ln(x)}$ grandit à mesure que la valeur de x augmente, le deuxième, lui, montre bien que l’erreur relative est nettement plus petite mais surtout qu’elle diminue.

Constatons que ce second graphe contient une partie verticale pour les petites valeurs de x ; il y a donc une grande erreur relative pour ces valeurs. En effet, si nous prenons l’exemple de $x = 2$: $\pi(2) = 1$, tandis que $\frac{2}{\ln(2)} \simeq 2,89 \implies$ l’erreur est d’environ 1,89.

Mais intéressons-nous plutôt à de bien plus grandes valeurs de x . Pour cela, penchons-nous sur un autre graphe qui ne commence qu’à $x = 500$ (les nombres premiers inférieurs à 500 sont bien connus), afin de nous représenter l’erreur relative à une meilleure échelle :

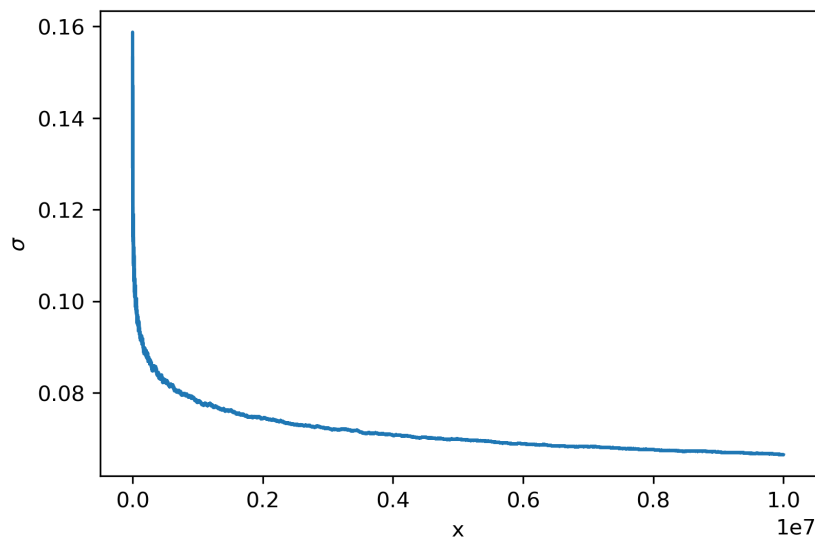


FIGURE 7 – Graphe de l’erreur relative de $f(x) = \frac{x}{\ln(x)}$ par rapport à $\pi(x)$, sans les 500 premières valeurs

Nous voyons désormais bien mieux que l'erreur relative σ entre $\pi(x)$ et $\frac{x}{\ln(x)}$ diminue à mesure que x augmente, mais surtout qu'elle semble tendre vers 0. Ceci est effectivement confirmé par le théorème des nombres premiers cité précédemment.

En effet, le théorème des nombres premiers nous dit que :

$$\lim_{x \rightarrow +\infty} \frac{\pi(x)}{\frac{x}{\ln(x)}} = 1$$

Et en ce qui concerne l'erreur relative :

$$\lim_{x \rightarrow +\infty} \sigma(x) = \lim_{x \rightarrow +\infty} \frac{\frac{x}{\ln(x)} - \pi(x)}{\pi(x)} = \lim_{x \rightarrow +\infty} \frac{\frac{x}{\ln(x)}}{\pi(x)} - 1$$

Comme, par symétrie, $\lim_{x \rightarrow +\infty} \frac{\frac{x}{\ln(x)}}{\pi(x)}$ est égale à $\lim_{x \rightarrow +\infty} \frac{\pi(x)}{\frac{x}{\ln(x)}}$ qui est égale à 1, on a finalement :

$$\lim_{x \rightarrow +\infty} \sigma(x) = \overbrace{\lim_{x \rightarrow +\infty} \frac{\frac{x}{\ln(x)}}{\pi(x)}}^1 - 1 = 0$$

Vu que leur erreur relative tend vers 0, $\pi(x)$ et $f(x) = \frac{x}{\ln(x)}$ sont bien équivalentes lorsque x tend vers l'infini.

3.3 Fonction d'approximation : le logarithme intégral

La fonction logarithme intégral $\text{li}(x)$ est définie comme suit, pour tout $x \in \mathbb{R}_+^* - \{1\}$:

$$\text{li}(x) = \int_0^x \frac{1}{\ln(t)} dt$$

Toujours selon le théorème des nombres premiers, cette fonction est équivalente à $\pi(x)$ quand x tend vers l'infini.

Théorème 7

$$\pi(x) \underset{x \rightarrow \infty}{\sim} \text{li}(x)$$

Donc :

$$\lim_{x \rightarrow +\infty} \frac{\pi(x)}{\text{li}(x)} = 1$$

Pour nous éviter des complications dans l'intégration de $\frac{1}{\ln(t)}$ au vu de l'asymptote que cette fonction présente en $t = 1$, nous considérerons le logarithme intégral comme :

$$\text{li}(x) = \underbrace{\text{li}(2)}_{\sim 1,0451638} + \underbrace{\text{Li}(x)}_{\int_2^x \frac{1}{\ln(t)} dt}$$

En effet, la valeur de $\text{li}(2)$ est bien connue, et procéder ainsi nous facilitera la tâche. Tout comme nous l'avons fait pour $f(x) = \frac{x}{\ln(x)}$, il serait intéressant de pouvoir tracer le graphe de $\text{li}(x)$ afin de pouvoir comparer cette deuxième fonction d'approximation à $\pi(x)$.

Comme $\frac{1}{\ln(t)}$ n'a pas de primitive élémentaire, nous allons devoir recourir à deux méthodes d'intégration alternatives pour pouvoir représenter correctement $\text{li}(x)$, avec le plus de précision possible.

Tout d'abord, nous procéderons à une intégration numérique, puis à une multiple intégration par parties. Nous comparerons ensuite nos deux méthodes d'approximation entre elles. Enfin, nous jugerons de leur précision en les comparant à $\pi(x)$.

3.3.1 Intégration numérique

Une méthode assez efficace d'intégration numérique est celle de Simpson.

Elle consiste à découper l'intervalle de la fonction $f(x)$ à intégrer en n parties égales : $x_1, x_2, x_3, \dots, x_n$. En prenant l'image des trois premiers x_i par f et en y faisant passer une courbe, on obtient une fonction quadratique aisément intégrable de x_1 à x_3 . On répète l'opération avec une nouvelle courbe de x_3 à x_5 , puis avec les trois x_i d'après, et ceci jusqu'à arriver à x_n . En formulant tout cela correctement et en le développant suffisamment, on obtient :

$$\int_a^b f(x) dx \simeq \frac{h}{3} \left(f(a) + 4 \sum_{i \text{ impair}} f(x_i) + 2 \sum_{i \text{ pair}} f(x_i) + f(b) \right)$$

Précisons que n doit forcément être pair, et que $h = \frac{b-a}{n}$ (écart entre chaque x_i), et que évidemment, plus n est grand, moins l'erreur est grande.

Remarque 5

La formule de Simpson ci-dessus est générale et emploie donc x comme variable. Même si nous allons continuer de l'utiliser, il ne faut pas perdre de vue que c'est $\frac{1}{\ln(t)}$ que nous souhaitons intégrer, et que x n'est que la borne supérieure de notre intégrale. Ainsi, lorsque nous parlons de découpage en x_i , il faut garder en tête que dans notre cas précis, il s'agit plutôt d'un découpage en t_i .

La formule en main, nous pouvons facilement créer un programme *Python* qui calcule une approximation de la valeur de l'intégrale en fonction de f , des bornes a et b , et du découpage n .

Le principe est simple : il s'agit une boucle "while" qui passe en revue chaque i jusqu'à n , à l'intérieur de laquelle sont sommées les $f(x_i)$ entre elles, en fonction de si i est pair ou impair (condition "if"). Une fois que le programme a calculé les deux sommes en question (celle pour les i pairs et celle pour les i impairs), il ne lui reste plus qu'à entrer toutes les données dans la formule de Simpson ci-dessus, ainsi que d'y ajouter la valeur connue de $\text{li}(2)$.

On peut ensuite créer un deuxième programme qui, en appelant le premier, calcule la valeur de chaque point entre 2 et une valeur souhaitée pour en faire un graphe.

Voici un graphe de $\text{li}(x)$ jusqu'à $x = 10^6$:

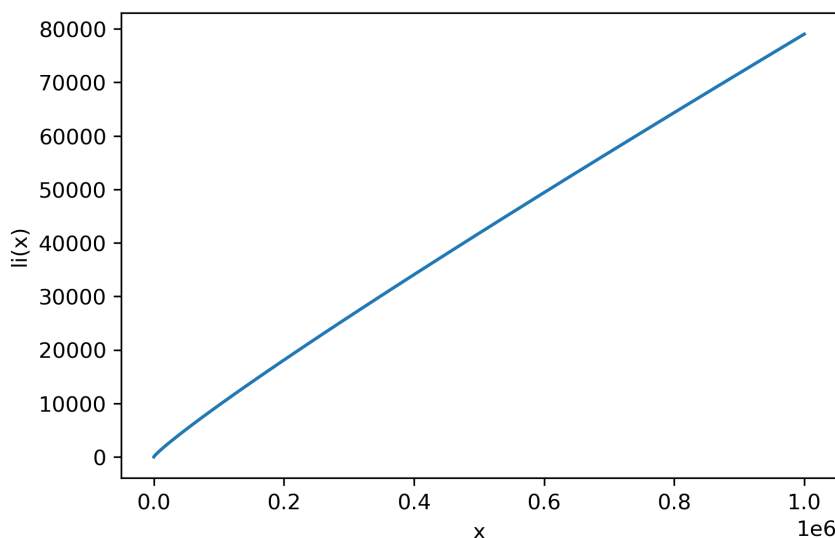


FIGURE 8 – Graphe de $\text{li}(x)$ avec la méthode de Simpson pour $n = 1000$

Ce graphe a été réalisé avec un découpage $n = 1000$ pour chaque intervalle $[2 ; x]$. La précision de $\text{li}(x)$ pour les grandes valeurs de x est donc largement diminuée. En effet, le résultat d'un découpage d'un intervalle de 1 en 1000 est bien plus précis que celui d'un intervalle de 10^6 en 1000...

Si l'on veut en revanche calculer une seule valeur de $\text{li}(x)$ et pas celle de chaque point du graphe, on peut se permettre d'augmenter considérablement n sans que le temps de calcul soit trop long.

Sur le graphe, avec $n = 1000$, on a $\text{li}(10^6) \simeq 79030, 21$. Si l'on prend maintenant $n = 10^6$, on a :

$$\text{li}(10^6) \simeq 78626, 51$$

Sachant que le site de WolframAlpha nous donne $\text{li}(10^6) \simeq 78627, 55$, on s'aperçoit bien que l'augmentation de n va de pair avec celle de la précision.

3.3.2 Intégration par parties

Une autre manière d'intégrer $\frac{1}{\ln(t)}$ est de concevoir cette opération de la manière suivante :

$$\int \frac{1}{\ln(t)} dt = \int \frac{1}{\ln(t)} \cdot 1 dt$$

Ainsi, la fonction à intégrer devient un produit que l'on peut intégrer par parties selon la formule suivante :

$$\int f(t) \cdot g'(t) dt = f(t) \cdot g(t) - \int f'(t) \cdot g(t) dt$$

Dans le cas de $\int \frac{1}{\ln(t)} \cdot 1 dt$, on considère donc :

$$f(t) = \frac{1}{\ln(t)} \Rightarrow f'(t) = \frac{-1}{t \ln^2(t)} \text{ et } g'(t) = 1 \Rightarrow g(t) = t$$

En appliquant une première fois la formule d'intégration par parties, on obtient :

$$\int \frac{1}{\ln(t)} dt = \frac{t}{\ln(t)} + \int \frac{1}{\ln^2(t)} dt$$

Nous sommes donc face à une nouvelle intégrale. Sur le même principe qu'avant, considérons :

$$f(t) = \frac{1}{\ln^2(t)} \Rightarrow f'(t) = \frac{-2}{t \ln^3(t)} \text{ et } g'(t) = 1 \Rightarrow g(t) = t$$

On applique la formule une seconde fois et on obtient :

$$\int \frac{1}{\ln(t)} dt = \frac{t}{\ln(t)} + \frac{t}{\ln^2(t)} + \int \frac{2}{\ln^3(t)} dt$$

Pour la troisième intégrale, on considère dans la même logique à nouveau :

$$f(t) = \frac{2}{\ln^3(t)} \Rightarrow f'(t) = \frac{-6}{t \ln^4(t)} \text{ et } g'(t) = 1 \Rightarrow g(t) = t$$

Et cela nous donne :

$$\int \frac{1}{\ln(t)} dt = \frac{t}{\ln(t)} + \frac{t}{\ln^2(t)} + \frac{2t}{\ln^3(t)} + \int \frac{6}{\ln^4(t)} dt$$

On commence à comprendre que l'on pourrait continuer ainsi pendant un bon moment... Mais le faire quelques fois seulement nous suffit à deviner la formule générale de cette intégration par parties :

$$\int \frac{1}{\ln(t)} dt = \frac{t}{\ln(t)} + \frac{t}{\ln^2(t)} + \frac{2t}{\ln^3(t)} + \frac{6t}{\ln^4(t)} + \frac{24t}{\ln^5(t)} + \dots + \frac{(k-1)! \cdot t}{\ln^k(t)} + \dots$$

N'oublions pas que nous souhaitons intégrer notre fonction de 2 à x . Donc, en ajoutant nos bornes, on obtient finalement que :

$$\text{Li}(x) \simeq \left[\sum_{k=1}^n \frac{(k-1)! \cdot t}{\ln^k(t)} \right]_2^x$$

Donc :

$$\text{Li}(x) \simeq \frac{x}{\ln(x)} + \frac{x}{\ln^2(x)} + \dots + \frac{(n-1)! \cdot x}{\ln^n(x)} - \left(\frac{2}{\ln(2)} + \frac{2}{\ln^2(2)} + \dots + \frac{(n-1)! \cdot 2}{\ln^n(2)} \right)$$

À ce stade-là il est temps de se poser la question de la précision de ce développement. Quel n minimum devrions-nous prendre afin d'être suffisamment précis? Pour cela, il est utile de faire un graphe qui compare $\text{li}(x)$ selon différents n , ainsi qu'avec $\pi(x)$ afin d'avoir un point de repère.

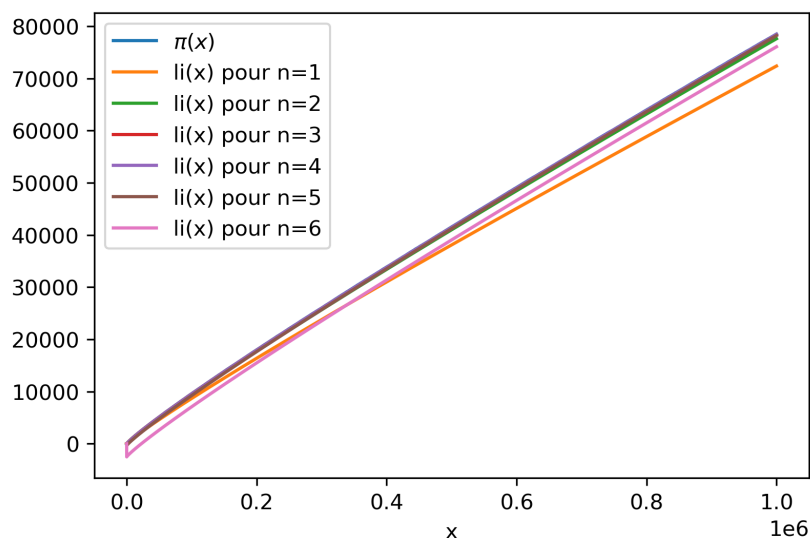


FIGURE 9 – Graphe de $li(x)$ pour différents n et de $\pi(x)$, jusqu'à $x = 10^6$

Nous remarquons que toutes les courbes semblent être à cette échelle plus ou moins superposées, ou en tout cas semble se chevaucher, sauf celle de $n = 1$, ce qui semble assez normal étant donné qu'elle est le résultat d'une seule intégration par parties.

Remarque 6

L'intégration par parties de $\frac{1}{\ln(t)}$ de 2 à x avec $n=1$ n'est autre que $\frac{x}{\ln(x)} - \frac{\overbrace{2}^{\simeq 2,89}}{\ln(2)}$. Elle est donc à peu près égale à $f(x) = \frac{x}{\ln(x)}$. Le logarithme intégral est en fait une sorte d'amélioration de $f(x) = \frac{x}{\ln(x)}$

Il nous donne donc une approximation de $\pi(x)$ avec une meilleure précision que le fait la fonction $f(x) = \frac{x}{\ln(x)}$.

Toutefois, relevons le fait que l'approximation avec $n = 1$ a l'air de bien concorder avec les premières valeurs de $\pi(x)$, même mieux que l'approximation avec $n = 6$ qui est elle plus en accord avec de plus grandes valeurs de $\pi(x)$. En effet, la courbe rose de $n = 6$ semble se rapprocher très lentement de celle de $\pi(x)$.

Mais comme cela ne nous avance pas énormément dans notre quête du n idéal, faisons un grand zoom sur ce graphe pour tenter de distinguer chacune des courbes correctement :

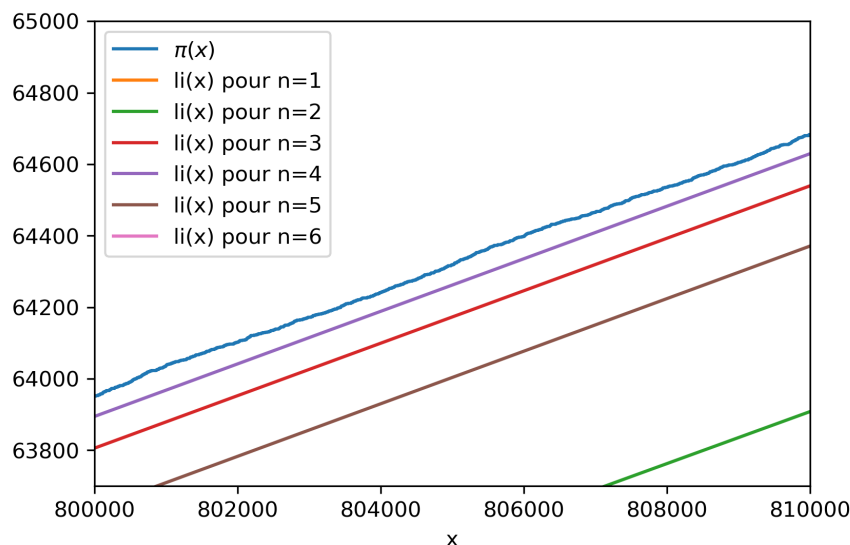


FIGURE 10 – Graphe de $\text{li}(x)$ pour différents n et de $\pi(x)$, entre $x = 8 \cdot 10^5$ et $x = 8.1 \cdot 10^5$

Ce deuxième graphe nous avance bien mieux quant au n à adopter pour notre approximation de $\text{Li}(x)$. C'est donc $n = 4$ qui suit à cette échelle le mieux et de plus près la fonction $\pi(x)$. Si nous prenions une échelle encore bien plus grande que celle-ci, ce ne serait peut-être pas ce même n qui nous donnerait la meilleure approximation, mais comme c'est à peu près cet ordre de grandeur dont nous nous servons pour comparer le logarithme intégral à $\pi(x)$, $n = 4$ nous conviendra parfaitement.

N'oublions pas d'ajouter la valeur connue de $\text{li}(2)$ à $\text{Li}(x)$ et retenons comme approximation finale :

$$\text{li}(x) \simeq 1.0456138 + \frac{x}{\ln(x)} + \frac{x}{\ln^2(x)} + \frac{2x}{\ln^3(x)} + \frac{6x}{\ln^4(x)} - \left(\frac{2}{\ln(2)} + \frac{2}{\ln^2(2)} + \frac{4}{\ln^3(2)} + \frac{12}{\ln^4(2)} \right)$$

Ainsi, nous pouvons déterminer la valeur que prend notre approximation de $\text{li}(x)$ en $x = 10^6$, ce qui nous donne :

$$\text{li}(10^6) \simeq 78474,78$$

3.3.3 Comparaison de $\text{li}(x)$ avec $\pi(x)$

Afin de montrer que le logarithme intégral est équivalent à $\pi(x)$ quand x tend vers l'infini, regardons directement le graphe de l'erreur relative σ entre les deux fonctions. En effet, nous avons vu avec $f(x) = \frac{x}{\ln(x)}$ que l'erreur absolue n'était pas vraiment représentative de l'équivalence à $\pi(x)$.

D'autre part, nous pouvons à nouveau enlever les 500 premières valeurs qui nous intéressent moins :

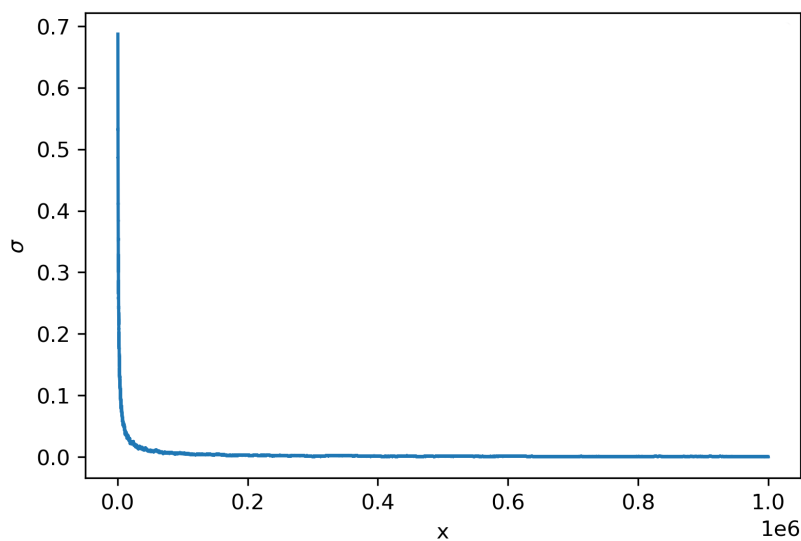


FIGURE 11 – Graphe de l’erreur relative de notre approximation de $\text{li}(x)$ par rapport à $\pi(x)$, sans les 500 premières valeurs

Comme on pouvait s’y attendre, l’erreur relative semble tendre vers 0. Mais ceci paraît encore plus évident si l’on enlève non pas les 500 mais les 10000 premières valeurs de $\text{li}(x)$ et de $\pi(x)$:

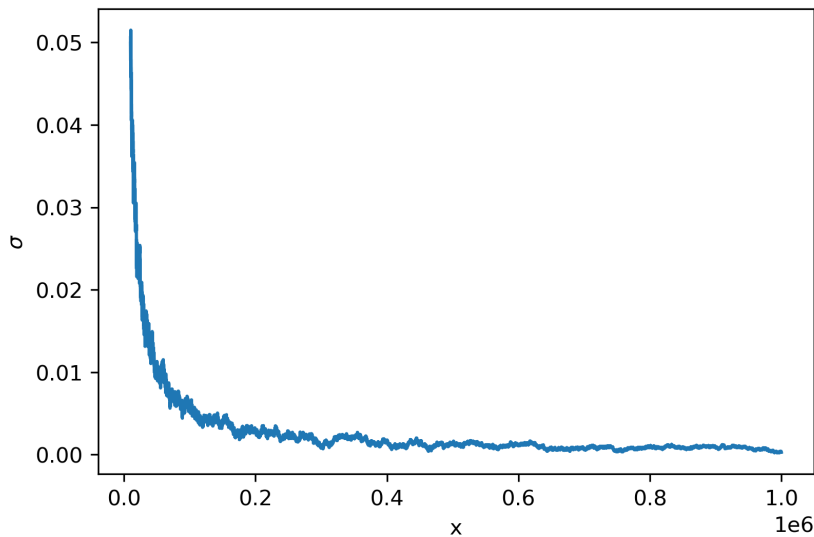


FIGURE 12 – Graphe de l’erreur relative de notre approximation de $\text{li}(x)$ par rapport à $\pi(x)$, sans les 10000 premières valeurs

Pour finir, assurons-nous que l’erreur relative σ tende bien vers 0 en calculant sa limite quand x tend vers l’infini :

$$\lim_{x \rightarrow +\infty} \sigma(x) = \lim_{x \rightarrow +\infty} \frac{\text{li}(x) - \pi(x)}{\pi(x)} = \lim_{x \rightarrow +\infty} \frac{\text{li}(x)}{\pi(x)} - 1$$

Comme $\lim_{x \rightarrow +\infty} \frac{\pi(x)}{\text{li}(x)} = 1$ selon le théorème des nombres premiers, on a de nouveau que, par symétrie :

$$\lim_{x \rightarrow +\infty} \frac{\text{li}(x)}{\pi(x)} = 1$$

Donc :

$$\lim_{x \rightarrow +\infty} \sigma(x) = \lim_{x \rightarrow +\infty} \frac{\text{li}(x)}{\pi(x)} - \frac{\pi(x)}{\pi(x)} = 1 - 1 = 0$$

Ceci nous permet bien d'affirmer que $\pi(x)$ et $\text{li}(x)$ sont des fonctions équivalentes quand x tend vers l'infini.

3.4 Comparaison de $f(x) = \frac{x}{\ln(x)}$ et de $\text{li}(x)$ avec $\pi(x)$

Maintenant que nous avons parcouru plusieurs fonctions d'approximation de $\pi(x)$, il est temps de conclure ce chapitre en les comparant les unes aux autres, afin de déterminer dans quelle mesure elles sont précises.

Tout d'abord, regardons ce tableau récapitulatif des images de chaque fonction pour quelques grandes valeurs de x entre 10^5 et 10^9 (le fait de ne pas tracer le graphe de chaque fonction mais de ne sélectionner que quelques valeurs qui nous intéressent nous permet d'étudier les images de plus grandes valeurs de x que précédemment) :

x	10^5	10^6	10^7	10^8	10^9
$\pi(x)$	9592	78'498	664'579	5'761'455	50'847'534
$f(x) = \frac{x}{\ln(x)}$	8685,89	72'382,41	620'420,69	5'428'681,02	48'254'942,43
$\text{li}(x)$ avec Simpson	9629,81	78'627,55	664'918,41	5'762'211,11	50'849'268,64
$\text{li}(x)$ avec $n = 2$	9434,33	77'615,62	658'906,87	5'723'380,80	50'583'475,90
$\text{li}(x)$ avec $n = 3$	9553,38	78'362,07	663'671,13	5'755'366,06	50'808'190,96
$\text{li}(x)$ avec $n = 4$	9535,55	78'474,78	664'508,13	5'760'525,16	50'840'671,56
$\text{li}(x)$ avec $n = 5$	9247,42	78'222,47	664'428,75	5'761'356,74	50'846'650,99
$\text{li}(x)$ avec $n = 6$	7088,57	76'075,72	662'333,19	5'759'499,88	50'846'002,05

TABLE 4 – Comparaison des images de $\pi(x)$ et de plusieurs fonctions d'approximation pour différentes valeurs de x

Précisons que nous avons ici utilisé $n = 10^7$ pour le calcul de $\text{li}(x)$ avec la méthode de Simpson (découpage en n intervalles).

Pour nous faire une meilleure idée de la précision de chacune des fonctions, utilisons ce tableau comme base afin de représenter graphiquement les erreurs absolues en utilisant simplement les cinq valeurs de x du tableau (ceci nous suffira à nous faire une idée globale de l'évolution de l'erreur).

Remarque 7

Afin d'assurer une visualisation optimale des graphes suivants, ceux-ci comportent une

échelle logarithmique. Ainsi, tous les points se trouvent à la même distance les uns des autres.

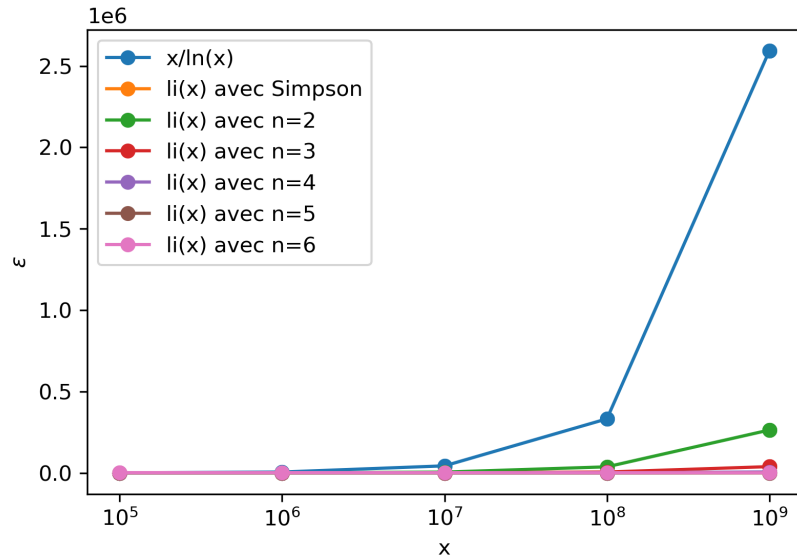


FIGURE 13 – Graphe de l’erreur absolue de plusieurs fonctions d’approximation de $\pi(x)$

Ce qui saute aux yeux sur ce graphe, c’est que l’erreur absolue de $f(x) = \frac{x}{\ln(x)}$ augmente extrêmement plus rapidement que celle de toutes les autres fonctions d’approximation. Mise à part $li(x)$ avec $n = 2$ qui semble également un peu moins bonne, les erreurs absolues des autres fonctions sont entassées vers le bas du graphe, et sont pratiquement indiscernables les unes des autres. Faisons donc un zoom sur ce graphe pour mieux les différencier.

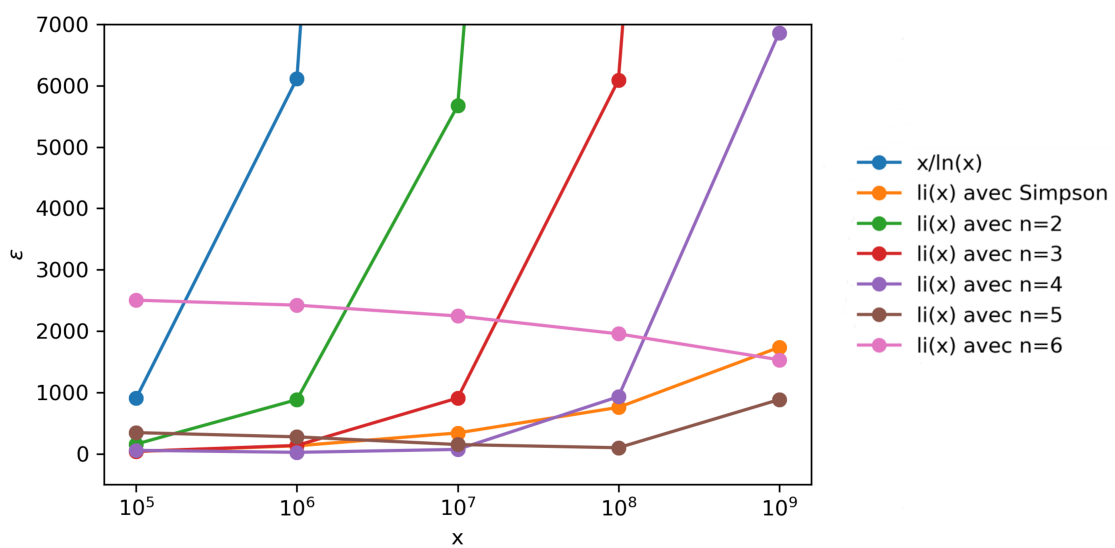


FIGURE 14 – Graphe agrandi de l’erreur absolue de plusieurs fonctions d’approximation de $\pi(x)$

Deux éléments semblent importants à relever sur ce graphe.

Premièrement, à cette échelle-là et en ce qui concerne le logarithme intégral par parties, ce n'est plus $\text{li}(x)$ avec $n=4$ qui semble le mieux correspondre à $\pi(x)$, mais bien avec $n = 5$.

Deuxièmement, l'erreur absolue de $\text{li}(x)$ avec $n = 6$, bien qu'elle soit gigantesque dans les petites valeurs de x , est la seule qui ne fait que diminuer à mesure que la valeur de x augmente, à tel point qu'elle est même plus petite que celle de Simpson pour $x = 10^9$.

Regardons maintenant les erreurs relatives de nos fonctions d'approximation.

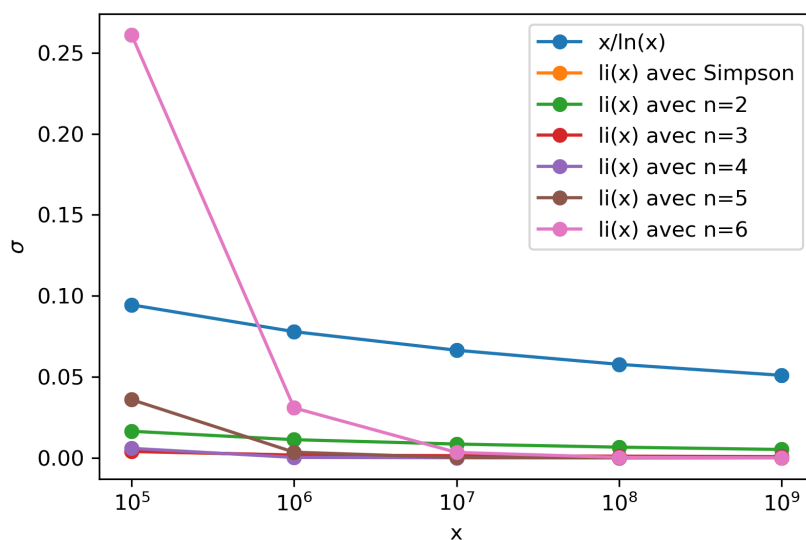


FIGURE 15 – Graphe de l'erreur relative de plusieurs fonctions d'approximation de $\pi(x)$

Ce graphique des erreurs relatives confirme nos précédentes observations : $f(x) = \frac{x}{\ln(x)}$ semble être la fonction la moins précise sur toute la ligne, tandis que $\text{li}(x)$ avec $n = 6$ n'est pas précise du tout pour de petites valeurs, mais gagne en précision assez rapidement, bien plus vite qu'elle n'en gagnait en terme d'erreur absolue.

Les erreurs des autres fonctions sont toujours difficiles à discerner ; faisons à nouveau un zoom sur ce graphe.

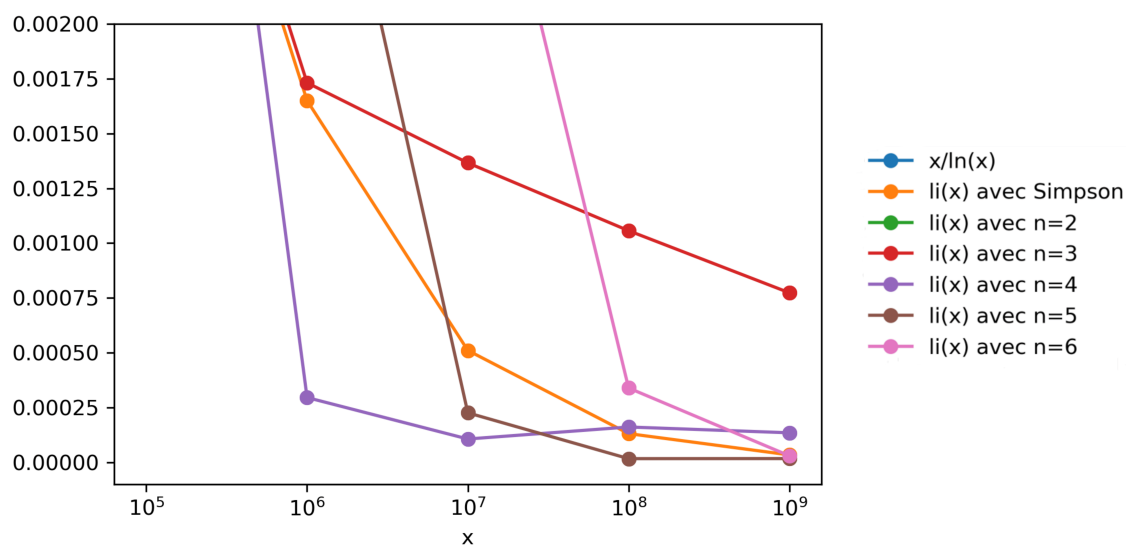


FIGURE 16 – Graphe agrandi de l’erreur relative de plusieurs fonctions d’approximation de $\pi(x)$

Ce graphe présente une petite bizarrerie qui mérite d’être relevée : l’erreur relative de $\text{li}(x)$ avec $n = 4$ augmente légèrement entre $x = 10^7$ et $x = 10^8$. Ceci est certes interpellant, mais rappelons-nous que nous travaillons à très petite échelle, ce qui fait que cette variation est en réalité assez faible. De plus, si nous calculons l’erreur pour $x = 10^{10}$ qui n’est pas représentée sur le graphe, nous trouvons une valeur plus petite encore toutes les précédentes, ce qui nous confirme que la courbe de l’erreur relative de $\text{li}(x)$ avec $n = 4$ continue bien de descendre malgré cette petite anomalie.

On voit également que la fonction d’approximation de $\text{li}(x)$ avec Simpson est celle dont l’erreur varie le moins, avec une précision relativement bonne. On observe aussi que celles avec $n = 4$ et $n = 5$ sont très rapidement à peu près constantes dans leur erreur relative. D’autre part, l’erreur relative de $\text{li}(x)$ avec $n = 6$ en $x = 10^9$ se confond avec celle de Simpson et celle de $n = 5$.

Nous pourrions donc faire l’hypothèse que l’erreur relative de $\text{li}(x)$ avec $n = 6$ continuerait de diminuer ainsi jusqu’à ce que, peut-être, $n = 7$ devienne plus précise, puis peut-être $n = 8$, puis $n = 9$ bien plus loin encore,...

Revenons à notre échelle du milliard ; il serait judicieux de déterminer laquelle des trois fonctions dont les erreurs relatives ont l’air semblable en $x = 10^9$, est la plus précise pour cette valeur.

Pour cela, regardons un dernier tableau où figurent les erreurs relatives pour chaque fonction :

$f(x) = \frac{x}{\ln(x)}$	$5,099 \cdot 10^{-2}$
li(x) avec Simpson	$3,411 \cdot 10^{-5}$
li(x) avec $n = 2$	$5,193 \cdot 10^{-3}$
li(x) avec $n = 3$	$7,737 \cdot 10^{-4}$
li(x) avec $n = 4$	$1,350 \cdot 10^{-4}$
li(x) avec $n = 5$	$1,737 \cdot 10^{-5}$
li(x) avec $n = 6$	$3,013 \cdot 10^{-5}$

TABLE 5 – Valeurs de l’erreur relative pour différentes fonctions d’approximation de $\pi(x)$ pour $x = 10^9$

C’est donc li(x) avec $n = 5$ qui semble être la meilleure fonction d’approximation pour cet ordre de grandeur du milliard, avec une erreur relative de $1,737 \cdot 10^{-5}$ en $x = 10^9$. Cependant, Simpson avec un grand découpage en n intervalles reste une valeur sûre à n’importe quelle échelle, là où $n = 6$ semble être légèrement plus précise, mais uniquement pour de très grandes valeurs, et là où $n = 4$ était idéal pour notre ordre de grandeur du million.

4 Conclusion

Dans ce travail, nous avons exploré quatre différentes manières de répondre à notre première question introductive : "Existe-t-il un nombre fini ou un nombre infini de nombres premiers ?"

La première preuve, celle d'Euclide, est la plus triviale mais surtout la plus connue.

Nous avons prétendu qu'il existait un nombre fini de nombres premiers et les avons listés. Puis, nous avons créé un nombre M et avons démontré qu'il possédait forcément un diviseur premier hors-liste, ce qui nous a prouvé que la liste des nombres premiers n'était donc pas finie.

Nous avons ensuite poursuivi notre chemin dans les raisonnements par l'absurde et avons adapté cette preuve afin de démontrer qu'en plus du fait qu'il existe une infinité de nombres premiers, il en existe également une infinité possédant une même propriété ; la congruence à 3 (mod 4). Ceci nous a permis de faire une petite parenthèse sur le modulo ainsi que sur les classes d'équivalence.

Ensuite, nous avons totalement changé de répertoire et nous sommes saisis de la série harmonique comme outil de preuve.

Nous avons tout d'abord démontré sa divergence, dont nous allions avoir besoin plus tard. Cette preuve de Nicolas Oresme utilise le critère de comparaison avec une autre série plus petite que l'on sait diverger.

Ce joli détour nous a permis de faire le lien avec le produit eulérien que nous avons démontré de deux manières différentes. La première nous a donné l'occasion de poursuivre notre chemin auprès des séries en utilisant la formule de convergence d'une série géométrique. La seconde nous a permis une approche plus instinctive, plus visuelle. Au final, les deux se sont très bien complétées. Dans tous les cas, nous avons pu démontrer que le produit eulérien était bien composé d'une infinité de termes dont chacun employait un nombre premier différent.

Enfin, nous nous sommes attaqués à la dernière preuve de cette première partie : celle de la divergence de la série des inverses des nombres premiers par Paul Erdős. Cette preuve par l'absurde nous a fait faire un petit crochet par l'analyse combinatoire. Il s'agissait certainement de la plus coriace car elle requérait un certain niveau d'abstraction.

Dans cette première partie, nous avons pu progresser avec toute une panoplie d'outils mathématiques extrêmement variés, tout cela pour prouver au final une seule affirmation. Il est très intéressant de se rendre compte qu'autant de chemins peuvent être pris pour répondre à une seule et même question. Le fait de ne pas se focaliser sur une réponse obtenue mais de pousser, d'étendre la réflexion à la recherche d'autres moyens d'y parvenir, est assez spécifique aux mathématiques.

La deuxième partie de mon travail se distingue drastiquement de la première ; on quitte l'approche très théorique adoptée précédemment et on se lance dans quelque chose qui relève plutôt de l'approximation, une méthode bien plus expérimentale qui requiert quelques concessions quant à la rigueur théorique adoptée précédemment.

Tout d'abord, nous avons pu prendre connaissance de la fonction de compte des nombres premiers : $\pi(x)$. Programmer cette fonction sur *Python* a été assez enrichissant de par le fait qu'il a fallu essayer de trouver le code le plus adéquat pour limiter le temps de

calcul de l'ordinateur au maximum. Ceci nous a contraint à repenser le test de primalité dans le but de l'optimiser.

Nous nous sommes ensuite penchés sur deux fonctions d'approximation de $\pi(x)$, dans le but de pouvoir nous faire une idée du nombre de nombres premiers inférieurs ou égaux à x , sans devoir tester la primalité de tous ces nombres (processus qui devient long pour de grandes valeurs de x ...).

Nous nous sommes donc appuyés sur le théorème des nombres premiers (en nous contentant de l'énoncer), qui démontre l'équivalence de $f(x) = \frac{x}{\ln(x)}$ et de $\text{li}(x)$ avec $\pi(x)$ quand x tend vers l'infini. Sans prouver ce théorème, nous avons tenté de montrer au mieux ce qu'il contient.

En ce qui concerne $f(x) = \frac{x}{\ln(x)}$, il n'a pas été trop compliqué de la comparer à $\pi(x)$; le graphe de son erreur relative a suffi à nous convaincre de l'équivalence de ces deux fonctions.

Quant au logarithme intégral, le chemin pour y parvenir a été plus long, mais aussi nettement plus divertissant. Le fait que la fonction à intégrer ne possède pas de primitive élémentaire nous a conduits à deux autres méthodes d'intégration bien différentes l'une de l'autre : l'intégration numérique et l'intégration par parties.

En ce qui concerne l'intégration numérique, nous nous sommes penchés sur la méthode de Simpson et avons pu constater qu'elle était très utile et précise pour calculer une valeur isolée et modérément grande du logarithme intégral, mais qu'elle était tout de suite moins efficace pour afficher des graphes incluant de grandes valeurs.

Nous avons donc transitionné sur de multiples intégrations par parties qui pourrait théoriquement s'effectuer à l'infini. Afin de déterminer combien de fois il était nécessaire de répéter l'opération pour avoir la meilleure approximation à l'échelle que nous étudions, nous avons représenté différentes approximations de cette méthode graphiquement dans le but de pouvoir visualiser laquelle nous serait la plus utile. Nous avons pu en conclure que la fonction d'approximation la plus adaptée variait selon l'ordre de grandeur des valeurs de x étudié.

Finalement, nous avons montré que l'erreur relative de chacune des approximations du logarithme intégral par rapport à $\pi(x)$ tendait bien vers zéro.

Cette deuxième partie, bien que son point de départ soit une fonction relativement simple, nous a permis d'à nouveau élargir le champ des possibles en passant par de l'intégration numérique, pas mal de codage, de l'intégration "à la main", des calculs de limites, des calculs d'erreurs, de l'analyse de graphes, etc.

Plus généralement, ce travail nous a amenés à étudier des concepts sous différents angles, nous faisant découvrir par la même occasion des chemins mathématiques d'une grande variété. Bien entendu, le sujet pourrait encore être longuement approfondi, étant donné le nombre de notions qui émergent à chaque fois que l'on creuse ici ou là. On pourrait se demander par exemple : Quelles autres manières de prouver le fait qu'il existe une infinité de nombres premiers pourrions-nous trouver ? Pourrait-on prouver l'existence de l'infinité de certains couples de nombres premiers tels que les nombres premiers jumeaux, cousins, et sexy (différence de 2, 4, et 6 entre chacun) ? Quelle fonction d'approximation de $\pi(x)$ serait la plus précise si nous élargissions notre étude à des valeurs de x encore plus grandes ?

5 Remerciements

Tout d'abord, je souhaite remercier chaleureusement mon mentor, Etienne Grezet, pour le suivi de mon travail qu'il a mené avec rigueur, implication et enthousiasme.

Ensuite, je voudrais exprimer ma gratitude envers mon professeur de piano, Simon Peguiron, pour la relecture méticuleuse de l'ensemble de mon travail et pour ses commentaires très pertinents sur celui-ci. Je tiens également à remercier l'un de mes enseignants de l'école secondaire, Nathan Jeanmonod, qui m'a donné le goût des mathématiques et qui a rigoureusement relu une partie de ce travail. Merci aussi à mes parents d'avoir corrigé mon orthographe et ma syntaxe.

Je me dois bien-sûr de remercier Tryphon Bettex qui a réglé mes problèmes informatiques et patiemment essuyé mes larmes quand L^AT_EX me résistait. Enfin, un grand merci à mon acolyte Samsara Seghaïria Calame pour toutes nos précieuses discussions mathématiques, parfois entretenues dans un état de légère euphorie festive.

6 Références

- [1] Wikipédia. Congruence sur les entiers.
https://fr.wikipedia.org/wiki/Congruence_sur_les_entiers
(consulté le 17.09.2023).
- [2] Khan Academy. Proof : harmonic series diverges (vidéo).
<https://www.khanacademy.org/math/ap-calculus-bc/bc-series-new/bc-10-6/v/harmonic-series-divergent>
(consulté le 19.09.2023).
- [3] Wikipédia. Produit eulérien.
https://fr.wikipedia.org/wiki/Produit_eul%C3%A9rien
(consulté le 28.10.2023).
- [4] Bibmath. Divergence de la série des inverses des nombres premiers.
https://www.bibmath.net/dico/index.php?action=affiche&quoi=./s/serie_prem.html
(consulté le 13.11.2023).
- [5] Wikipédia. Fonction de compte des nombres premiers.
https://fr.wikipedia.org/wiki/Fonction_de_compte_des_nombres_premiers#:~:text=En%20math%C3%A9matiques%2C%20la%20fonction%20de,valeurs%20enti%C3%A8res%20de%20la%20variable.
(consulté le 30.08.2023).
- [6] Wikipédia. Logarithmic integral function.
https://en.wikipedia.org/wiki/Logarithmic_integral_function
(consulté le 26.11.2023).

7 Annexes

Voici un programme codé en *Python* sur Spyder 5 pour le calcul de la fonction $\pi(x)$:

```
1 def pi(x):
2     L=[2]
3     k=3
4     i=0
5     while k<=x :
6         while i<len(L) and L[i]<=sqrt(k) and k<=x:
7             if k%L[i]==0:
8                 k+=1
9                 i=0
10            else:
11                i+=1
12        if k<=x:
13            L.append(k)
14        k+=1
15        i=0
16    return len(L)
```

Et voilà le programme de l'intégration numérique de $\frac{1}{\ln(t)}$ avec la méthode de Simpson, également codé en *Python* sur Spyder 5 :

```
1 def simpson(f, a, b, n):
2     h = (b-a)/n
3     x = a+h
4     i = 1
5     simpair = 0
6     spair = 0
7     if n % 2 == 1:
8         return False
9     while i < n:
10        if i % 2 == 1:
11            simpair = f(x)+simpair
12            x = x+h
13            i = i+1
14        else:
15            spair = f(x)+spair
16            x = x+h
17            i = i+1
18    return h/3*(f(a)+4*simpair+2*spair+f(b))
```