École Normale Supérieure

Écarts entre les nombres premiers consécutifs et fonctions multiplicatives sur des petits intervalles

Introduction au domaine de recherche

Élie Goudout
sous la direction de
Régis De La Bretèche

Table des matières

1 La théorie des nombres

La théorie des nombres a pour but d'étudier les propriétés des nombres entiers. Nous savons depuis l'antiquité que tout entier naturel $n > 1$ se décompose de manière unique, à l'ordre des facteurs près, sous forme d'un produit de nombres premiers. À l'aide de ce théorème, Euclide parvient à montrer qu'il existe une infinité de nombres premiers ; c'est-à-dire, notant $\pi(x)$ le nombre d'entiers premiers inférieurs ou égaux à x,

$$\pi(x) \underset{x\to\infty}{\longrightarrow} \infty. \tag{1}$$

L'intérêt des mathématiciens à étudier les nombres premiers vient naturellement de la structure multiplicative des entiers. Durant mon mémoire, je me suis intéressé à des travaux récents sur la conjecture des nombres premiers jumeaux, selon laquelle $p + 2$ serait premier pour une infinité de nombres premiers p. Dans beaucoup de cas, les structures additive et multiplicative des entiers interagissent peu, et cette conjecture est un très bon exemple de la complexité à justifier rigoureusement ce type d'heuristique. Dans un premier temps je présente ci-dessous une approche de cette conjecture développée au début du XX$^{\text{e}}$ siècle, et qui n'a toujours pas livré tous ses secrets : la théorie du crible. Je parlerai ensuite très brièvement de la théorie multiplicative des nombres sur les petits intervalles qui a connu une très belle avancée en janvier 2015 grâce à Matomäki et Radziwiłł, et qui est désormais mon domaine de recherche.

2 Le crible

2.1 Première tentative

Depuis Euclide, les mathématiciens s'attachent à essayer de préciser quantitativement la relation qualitative (1). Le premier pas consiste naturellement à essayer de trouver une méthode pour générer tous les premiers. Ératosthène y parvient à l'aide du très célèbre *crible d'Ératosthène*, profitant du fait qu'un entier inférieur à x n'admettant aucun facteur premier inférieur à $\sqrt{x}$ est premier.

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96

Premiers :
2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89

Figure 1 – En jaune, les entiers ayant « survécu » au crible d'Ératosthène : ils sont premiers !

On peut alors essayer de calculer formellement le nombre de nombres premiers inférieurs à x grâce à cette méthode. Pour $z > 0$, notons $P_z = \prod_{\substack{p \leq z \\ p \text{ premier}}} p$. On a alors, pour $x \geq 1$,

$$\pi(x) - \pi(\sqrt{x}) + 1 = \#\left\{1 \leq n \leq x \mid \operatorname{pgcd}(n, P_{\sqrt{x}}) = 1\right\}.$$

Pour calculer le cardinal de cet ensemble, il suffit d'estimer la taille du complémentaire en utilisant la *formule du crible* pour une union d'ensembles. Par exemple, pour savoir combien d'entiers inférieurs à x sont divisibles par 2 ou 3, il faut compter les entiers pairs (il y en a $\left\lfloor \frac{x}{2} \right\rfloor$) et les entiers divisibles par 3 ($\left\lfloor \frac{x}{3} \right\rfloor$), mais en enlevant les entiers divisibles par 2 et par 3, comptés 2 fois ($\left\lfloor \frac{x}{6} \right\rfloor$). On aboutit alors à la formule suivante :

$$\pi(x) - \pi(\sqrt{x}) + 1 = \sum_{d \mid P_{\sqrt{x}}} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor \tag{2}$$

où μ est la fonction de Möbius, à support sur les entiers sans facteur carré et valant -1 puissance le nombre de facteurs premiers.

2.2 L'échec d'Ératosthène, la réussite (partielle) de Brun

Dans (2), on est tentés de remplacer $\left\lfloor \frac{x}{d} \right\rfloor$ par $\frac{x}{d} - \left\{ \frac{x}{d} \right\}$. Lorsque l'on fait cela, on arrive à l'estimation

$$\pi(x) - \pi(\sqrt{x}) + 1 = x \prod_{p \leq \sqrt{x}} \left(1 - \frac{1}{p}\right) + \mathcal{O}\left(2^{\pi(\sqrt{x})}\right),$$

le terme d'erreur provenant de la majoration $\left\{\frac{x}{d}\right\} < 1$. D'après une formule de Mertens, $\prod_{p\leq\sqrt{x}}\left(1-\frac{1}{p}\right) = \frac{2e^{-\gamma}+o(1)}{\log x}$. Deux remarques en découlent : d'une part, le terme d'erreur est beaucoup trop grand car il comporte beaucoup trop de termes. Le problème provient des grands nombres premiers, qui introduisent trop de termes d'erreur par rapport au entiers qu'ils éliminent. On peut alors obtenir une majoration plus fine en se contentant de cribler avec les nombres premiers inférieurs à un certain paramètre z. D'autre part, le « terme principal » nous montre, compte tenu du théorème des nombres premiers, que les termes d'erreur dus aux $\left\{\frac{x}{d}\right\}$ contribuent de manière non négligeable. Il semblerait donc que même en adaptant convenablement notre méthode, on ne pourra jamais prouver le théorème des nombres premiers de cette manière. Cependant, il ne paraît pas impossible d'avoir une estimation correcte à une constante près.

Au début du XX$^{\text{e}}$ siècle, Viggo Brun pense à ne cribler que partiellement, en s'arrêtant à une étape paire, afin d'avoir une majoration. Par exemple :

$$\pi(x) \leq x - \left\lfloor\frac{x}{2}\right\rfloor - \left\lfloor\frac{x}{3}\right\rfloor - ... + \left\lfloor\frac{x}{2\cdot 3}\right\rfloor + \left\lfloor\frac{x}{2\cdot 5}\right\rfloor ...$$

et c'est tout ! On obtient bien une majoration grâce à la formule suivante :

$$\delta \leq 1 * \mu_{2\ell} \tag{3}$$

où $*$ désigne la convolution classique [1] entre fonctions multiplicatives, δ la fonction indicatrice de $\{1\}$ et $\mu_{2\ell}$ est la fonction μ tronquée aux entiers avec moins de 2ℓ facteurs premiers.

Avec cette méthode, en optimisant le paramètre z introduit précédemment et l'étape 2ℓ à laquelle on s'arrête de cribler, on obtient assez facilement :

$$\pi(x) \ll \frac{x\log\log x}{\log x}.$$

Ce résultat peut sembler faible, compte tenu du fait qu'au XX$^{\text{e}}$ siècle le théorème des nombres premiers était déjà connu. Cependant la méthode de Brun peut s'attaquer à une vaste catégorie de problèmes. En effet, on peut s'intéresser aux ensembles du type

$$\{n \in \mathcal{A} \mid \mathrm{pgcd}\,(n, P_z) = 1\},$$

où $\mathcal{A}$ peut être un ensemble plus général, non plus forcément $[1, x]$, du moment qu'on sait estimer $\mathcal{A}_d := \{n \in \mathcal{A} \mid d|n\}$. Par exemple, en prenant $\mathcal{A} = \{n(n+2) \mid n \leq x\}$, Brun [2] montre en 1919 le théorème suivant.

Théorème 1 (Brun).

$$\#\{p \leq x \mid p \textit{ et } p+2 \textit{ sont premiers}\} \ll \frac{x(\log\log x)^2}{(\log x)^2},$$

d'où l'on déduit que la somme des inverses des nombres premiers jumeaux converge.

De ce théorème on ne parvient malheureusement pas à déduire qu'il existe une infinité de nombres premiers jumeaux comme on le conjecture. Brun arrivera cependant à montrer qu'il existe une infinité d'entiers n tels que $n(n+2)$ ait moins de 9 facteurs premiers.

1. Pour f et g deux fonctions définies sur $\mathbb{N}^*$ et $n \geq 1$, on pose $(f * g)(n) = \sum_{d|n} f(d)g(n/d)$.

2.3 Le crible de Selberg

En 1947 Atle Selberg [12] s'inspire de Brun, mais remplace (3) par :

$$\delta(n) \leq \left(\sum_{d|n} \lambda_d \right)^2$$

où les λ_d sont des réels et $\lambda_1 = 1$. De manière générale, après optimisation sur les λ_d cette majoration donne de meilleurs résultats que ceux de Brun, et permet notamment de majorer asymptotiquement le cardinal d'ensembles tels que celui des nombres premiers jumeaux inférieurs à x avec les bornes conjecturées, à un facteur 4 près (contre un $(\log \log x)^2$ pour le crible de Brun dans ce cas là). Dans les années 1970, Chen [3] montre grâce au crible de Selberg qu'il existe une infinité de nombres premiers p tels que $p+2$ soit premier ou produit de 2 facteurs premiers (ce résultat est jugé optimal pour ces méthodes).

3 Écarts entre les nombres premiers consécutifs

3.1 Une nouvelle approche de la conjecture des nombres premiers jumeaux

Dans mon mémoire, je me suis intéressé aux travaux très récents initiés par les progrès de Maynard en novembre 2013 sur la conjecture des nombres premiers jumeaux. Ils prennent appui sur les travaux de Goldston, Pintz et Yıldırım.

3.1.1 L'approche de Goldston-Pintz-Yıldırım

La conjecture des nombres premiers jumeaux peut s'énoncer ainsi :

$$\liminf_{n \to +\infty} p_{n+1} - p_n = 2.$$

Ainsi, un progrès de nature différente de celle des travaux de Chen peut être imaginé : on sait (théorème des nombres premiers) qu'en moyenne $p_{n+1} - p_n \sim \log p_n$, mais on peut chercher à montrer que cette différence est plus petite infiniment souvent, et pourquoi pas égale à 2 pour une infinité d'entiers n ? En 2005, Goldston, Pintz et Yıldırım [5] (qu'on notera désormais GPY) travaillent dans cette direction et montrent pour la première fois qu'infiniment souvent, $p_{n+1} - p_n = o(\log p_n)$. On décrit désormais brièvement leur méthode, qui utilise les estimations de Selberg.

De manière générale, on conjecture que pour tous $k \geq 1, h_1 < h_2 < ... < h_k$ entiers, il existe une infinité d'entiers naturels n tels que tous les $n + h_i$ soient simultanément premiers, à moins bien sûr qu'une raison « évidente » rende ceci impossible. On entend par là qu'il faut exclure les cas comme n et $n+1$ dont l'un des deux est toujours pair. Une version faible de cette conjecture très forte consiste à essayer de montrer que pour k assez grand, il y a une infinité d'entiers n tels qu'au moins plusieurs des $n + h_i$ soient premiers. Notant $\mathcal{H} = \{h_1, ..., h_k\}$, on veut montrer qu'infiniment souvent, $n + \mathcal{H}$ contient plusieurs nombres premiers. GPY tentent de faire ceci en construisant une mesure w sur $[N, 2N]$ pour laquelle

$$\mathbb{E}_w \left(|(n + \mathcal{H}) \cap \mathbb{P}| \right)$$

est grande. Pour ce faire, ils choisissent la mesure grâce au crible de Selberg,

$$w(n) \propto \left(\sum_{\substack{d \mid \prod_{i=1}^{k}(n+h_i) \\ d \mid P_z}} \lambda_d \right)^2 . \tag{4}$$

Malheureusement, leur méthode une fois optimisée n'aboutit qu'à

$$\mathbb{E}_w \left(|(n+\mathcal{H}) \cap \mathbb{P}| \right) \geq 1 - \varepsilon. \tag{5}$$

En uniformisant leurs estimations ils arrivent quand même à prouver le résultat précité, et à l'améliorer 2 années plus tard [6]. Mais l'histoire ne s'arrête pas là.

3.1.2 Renforts de Zhang et Maynard

En mai 2013, Yitang Zhang [13] reprend la preuve de GPY qui s'appuie sur le crible de Selberg mais aussi sur un théorème de Bombieri-Vinogradov, assez rigide. Il parvient cependant à l'améliorer légèrement dans le cadre de ce problème et passe alors de $1-\varepsilon$ à $1+\delta$, pour avoir

Théorème 2 (Zhang)**.**

$$\liminf_{n \to +\infty} p_{n+1} - p_n \leq 70.000.000.$$

Le progrès est spectaculaire ! Terence Tao initie alors le projet Polymath8a [10] pour optimiser toutes les constantes, et à l'aide de plusieurs mathématiciens, ils abaissent le 70.000.000 à 4.680 !

Contre toute attente, en novembre de cette même année, un progrès encore plus fulgurant est apporté par James Maynard [9]. Au lieu d'utiliser le crible de Selberg classique pour w, il pense à cribler individuellement les $n+h_i$ en prenant w de la forme suivante :

$$w(n) \propto \left(\sum_{\substack{d_i \mid (n+h_i), \forall 1 \leq i \leq k \\ \prod_{i=1}^{k} d_i \mid P_z}} \lambda_{d_1,\dots,d_k} \right)^2 . \tag{6}$$

Cette version k-dimensionnelle du crible de Selberg lui permet d'obtenir les résultats suivants :

Théorème 3 (Maynard)**.**

$$\liminf_{n \to +\infty} p_{n+1} - p_n \leq 600 \tag{7}$$

$$\liminf_{n \to +\infty} p_{n+m} - p_n \ll m^3 e^{4m}. \qquad (m \geq 1) \tag{8}$$

À nouveau, ces résultats ont été légèrement améliorés par le projet polymath8b [11]. La méthode de Maynard a l'avantage d'être assez souple, et permet de montrer beaucoup de choses et souvent avec une uniformité assez confortable. Cela permet par exemple de trouver plusieurs nombres premiers consécutifs dans une même progression arithmétique donnée et très proches les uns des autres [8].

Un autre résultat intéressant est que la suite des écarts normalisés $\frac{p_{n+1}-p_n}{\log p_n}$ a pour valeur d'adhérence au moins $X\left(\frac{1}{8}+o(1)\right)$ réels $0 \leq x \leq X$ [1].

3.2 Grands écarts entre les nombres premiers

Il est aussi intéressant de chercher à comprendre les grandes valeurs de $p_{n+1} - p_n$, si elles existent. Grâce au crible développé par Maynard, en se basant sur une méthode d'Erdős-Rankin et avec un peu de théorie des graphes, Ford, Green, Konyagin, Maynard et Tao [4] ont montré en décembre 2014 que

Théorème 4 (FGKMT)**.**

$$\max_{p_{n+1}\leq X} p_{n+1} - p_n \gg \frac{(\log X)(\log\log X)(\log\log\log\log X)}{\log\log\log X}.$$

Il gagnent ainsi un facteur $\log\log\log X$ sur le précédent résultat datant de 1938 (dû à Rankin). Sur des bases probabilistes, Cramér a conjecturé que les plus grands écarts étaient de l'ordre de $(\log X)^{2+o(1)}$, mais pour le moment, cela semble bien loin de nos techniques.

4 Théorie multiplicative des nombres

Étant donnée f une fonction multiplicative, en général on sait assez bien estimer $\frac{1}{x}\sum_{n\sim x} f(n)$ ($n \sim x$ signifie $x < n \leq 2x$) grâce aux formules de Perron. Il n'est d'ailleurs pas nécessaire que f soit multiplicative si sa série de Dirichlet

$$F(s) := \sum_{n\geq 1} \frac{f(n)}{n^s}$$

est une fonction que l'on sait bien traiter. Par exemple, le théorème des nombres premiers découle de l'estimation de $\sum_{n\sim x} \Lambda(n)$ où $\sum_{n\geq 1} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)}$. Il est cependant beaucoup plus difficile d'estimer des sommes « courtes » de fonctions multiplicatives, par exemple

$$\sum_{X<n\leq X+X^\delta} f(n).$$

Si l'on savait le faire pour la fonction Λ de Von Mangoldt pour $\delta < \frac{1}{2}$, on pourrait montrer qu'il y a toujours un nombre premier en n^2 et $(n+1)^2$ par exemple. Ce genre de question se pose aussi pour les entiers friables par exemple. Jusqu'à l'année dernière, peu de résultats étaient connus sur les sommes courtes, et ceux-ci ne traitaient souvent que des sommes de longueur X^c pour $c > 0$ pas très petit. En janvier 2015, Kaisa Matomäki et Maksym Radziwiłł [7] ont montré un théorème très puissant.

Théorème 5. *Soit $f : \mathbb{N} \to [-1, 1]$ une fonction multiplicative. Il existe des constantes absolues $C, C' > 1$ telles que pour tout $2 \leq h \leq X$ et $\delta > 0$,*

$$\left| \frac{1}{h} \sum_{x<n\leq x+h} f(n) - \frac{1}{X} \sum_{X<n\leq 2X} f(n) \right| \leq \delta + C' \frac{\log\log h}{\log h}$$

pour tout $x \sim X$ sauf au plus

$$CX \left(\frac{(\log h)^{1/3}}{\delta^2 h^{\delta/25}} + \frac{1}{\delta^2 (\log X)^{1/50}} \right).$$

À un $o(1)$ près, certes pas très petit — essentiellement une puissance de $\log h$ —, on est désormais capables d'avoir certaines estimations sur des intervalles aussi petits qu'on veut, du moment que leur taille tend vers l'infini avec X, pour presque tout $x \sim X$. Leur méthode laisse entendre qu'il sera possible d'estimer en moyenne $\pi_k(x+h) - \pi_k(x)$ pour k assez grand, où $\pi_k(x) = \{n \leq x \mid n \text{ a } k \text{ facteurs premiers distincts}\}$, et d'obtenir un résultat similaire au théorème d'Erdős-Kac, qui estime la proportion des entiers n tels que $|\omega(n) - \log\log n| \leq \tau\sqrt{\log\log n}$, mais sur des petits intervalles.

Références

[1] W. D. Banks, T. Freiberg, and J. Maynard. On the limit points of the sequence of normalized prime gaps. *eprint arXiv :1404.5094*, Octobre 2014.

[2] V. Brun. La série 1/5+1/7+1/11+1/13+1/17+1/19+1/29+1/31+1/41+1/43+1/59+1/61+..., où les dénominateurs sont nombres premiers jumeaux est convergente ou finie. *Bull. Sci. Math.*, 43 :100–104, 124–128, 1919.

[3] J. R. Chen. On the representation of a large even integer as the sum of a prime and the product of at most two primes. *Sci. Sinica*, 16(157-176), 1973.

[4] K. Ford, B. Green, S. Konyagin, J. Maynard, and T. Tao. Long gaps between primes. preprint `arxiv.org/abs/1412.5029`.

[5] D. A. Goldston, J. Pintz, and C. Y. Yıldırım. Primes in tuples I. *Ann. of Math*, 170(2) :819–862, Sept 2009.

[6] D. A. Goldston, J. Pintz, and C. Y. Yıldırım. Primes in tuples II. *Springer*, 204(1) :1–47, Mars 2010.

[7] K. Matömaki and M. Raziwiłł. Multiplicative functions in short intervals. preprint `arxiv.org/abs/1501.04585`, Janvier 2015.

[8] J. Maynard. Dense clusters of primes in subsets. preprint `arxiv.org/abs/1405.2593`, Mai 2014.

[9] J. Maynard. Small gaps between primes. *Ann. of Math*, 181(1) :383–413, Jan 2015.

[10] D. H. J. Polymath. New equidistribution estimates of Zhang type, and bounded gaps between primes. unpublished at `arxiv.org/abs/1402.0811v2`.

[11] D. H. J. Polymath. Variants of the Selberg sieve, and bounded intervals containing many primes. *Res. Math. Sci.*, 1(12), Oct 2014.

[12] A. Selberg. On an elementary method in the theory of primes. *Norske Vid. Selsk. Forh. Trondheim*, 19(64-67), 1947.

[13] Y. Zhang. Bounded gaps between primes. *Ann. of Math*, 179(3) :1121–1174, Mai 2014.