Variations modernes sur la suite des nombres premiers

De la densité de la suite $\sin p$ lorsque p parcourt l'ensemble des nombres premiers

Olivier Ramaré

9 février 2008

Olivier Ramaré est chercheur au CNRS et actuellement en poste à l'université de Lille 1. Sa spécialité est sans surprise : il s'agit des nombres premiers !

Introduction

Si les nombres premiers ont toujours fasciné les apprentis mathématiciens, il est frappant que le théorème des nombres premiers constitue en général le point ultime des connaissances en ce domaine. Ce résultat donne un équivalent du nombre de nombres premiers inférieurs à une borne donnée (voir (3.6)), ce qui est bien sûr l'information minimale que l'on puisse espérer, et date de 1896. Depuis, la théorie s'est énormément étoffée, sans que ces améliorations ne trouvent la voie sinon de la célébrité, du moins celle du savoir classique d'un mathématicien. Nous tentons ici d'amorcer cette migration. Notre fil directeur et notre prétexte est de démontrer que la suite des nombres réels formées des $\sin p$, où p parcourt l'ensemble des nombres premiers, est *dense* dans l'intervalle $[-1,1]$, c'est à dire que pour tout point de $[-1,1]$ et tout intervalle ouvert I aussi petit qui soit qui contienne ce point, nous pouvons trouver un nombre premier p tel que $\sin p$ appartienne à I. Le résultat en lui-même n'a probablement pas d'autre intérêt que celui de répondre à un défi, mais en cours de route, nous montrerons comment manipuler ces nombres premiers, comment démontrer des théorèmes les concernant, ce qui de fait nous éclairera sur la nature de ces nombres.

Le chemin que nous suivons est résolûment moderne et résulte des divers travaux du vingtième siècle sur ces questions, depuis la technique de crible introduite par Viggo Brun en 1916, son utilisation par Ivan Vinogradov en 1937 pour explorer la structure des nombres premiers jusqu'à la période 1960/2000 marquée par l'affinement de ces méthodes et leur meilleure compréhension.

Concernant les prérequis nécessaires, notre démarche a été d'écrire un texte qui ressemble à un problème de classe préparatoire, pour ce qui est de la structure, que nous avons ensuite scindé en petits chapitres tout en étoffant le texte. Si les notions préalables sont réduites au minimum, cette réduction reste nécessairement limitée tant il est illusoire de vouloir introduire en deux lignes les logarithmes et espérer ensuite que le lecteur les manipule ... Toutefois, rappeler notions et notations est utile, ne serait ce que pour rafraîchir la mémoire de la lectrice, tout comme plus haut, nous avons précisé ce que "dense" signifiait. Cette alchimie est assez délicate et je dois remercier Xavier Caruso, Nicole Garnier et Lazare Vidiani pour leurs conseils précieux. Puis Stéphane Louboutin et Jean-François Burnol, ainsi que d'autres lecteurs, pour les erreurs ou la trop grande rapidité d'exposition par endroit qu'ils

ont découverts dans la première version de cette monographie. Le parti est aussi pris d'expliciter autant que faire se peut les constantes numériques et de ne pas recourir à la notation $\mathcal{O}$ qui simplifierait pourtant les preuves. Tout simplement parce que cette notation demande de la pratique et que les calculs en eux-mêmes permettent de comprendre ce qui est en jeu. Ce n'est que lorsque le lecteur dispose d'une maîtrise complète de ces calculs qu'il peut s'en dispenser et s'appuyer sur la notation $\mathcal{O}$. Nous recourrons à la notation $\mathcal{O}^*$, moins connue mais beaucoup plus facile d'usage. Écrire $f_1 = f_2 + \mathcal{O}^*(g)$, c'est tout simplement écrire $|f_1 - f_2| \leq g$; c'est à dire que $f_1(x)$ ne diffère en valeur absolue de $f_2(x)$ qu'au plus de $g(x)$. Il s'agit donc d'un $\mathcal{O}$ avec une constante égale à 1. Au passage, tous les problèmes d'uniformité par rapport aux paramètres disparaissent. Cette notation permet en outre d'utiliser rigoureusement des approximations, comme dans $\pi = 3,14 + \mathcal{O}^*(0,01)$.

Si l'idée directrice de ce livre est de ne rien cacher des difficultés, nous avons parfois laissé des parties faciles mais longues à la charge de la lectrice. Nous garantissons qu'il n'y a là aucun dragon caché et que les preuves dites *faciles* ne sont qu'au pire pénibles ...

Enfin, comme il semble que les femmes aient du mal à établir leur place de droit dans le monde des mathématiques, nous avons décidé de nous adresser alternativement à la lectrice et au lecteur, mais que nul.le ne se sente exclu.e par un vocable ou l'autre !

Quelques notations et des rappels

La première des notations qu'il nous faut détailler est le symbole $\sum$ (lire Sigma, ou Somme, au choix) qui nous permet de noter une sommation sur des *entiers*. Par exemple

$$\sum_{4<n\leq 7} a_n$$

désigne la somme sur tous les indices entiers n entre 4 exclus et 7 inclus, i.e. $a_5 + a_6 + a_7$. Parfois nous aurons besoin de spécifier plus avant le domaine de sommation :

$$\sum_{\substack{4<n\leq 7 \\ 2|n}} a_n$$

désigne cette fois-ci la somme sur les entiers n de l'intervalle $]4, 7]$, mais qui vérifient aussi la seconde condition, c'est à dire qu'ils doivent être pairs : il ne reste que $n = 6$ et la somme en question vaut donc a_6. Nous exploiterons énormément cette notation avec des domaines d'appartenance pour n plus ou moins compliqués. Les conditions portées sur les variables peuvent être telles qu'aucune ne les vérifie ! La somme porte alors sur un ensemble vide, et une telle somme vaut 0. Il faut bien voir que l'ordre d'énumération n'a

aucune espèce d'importance, ce pourquoi la notation plus usuelle

$$\sum_{n=5}^{7} a_n$$

n'est pas adaptée. De plus cette dernière notation devient vite illisible dés que le domaine est un tant soit peu complexe.

Lorsque la variable de sommation est notée p, la somme est restreinte aux *nombres premiers* p.

Nous notons $C(P)$ le nombre de nombres premiers dans l'intervalle $]P/4, P]$. Il s'agit d'une notation non standard mais qui sera bien pratique ici. Si l'on traduit cette définition en termes de sommation, cela donne :

$$C(P) = \sum_{P/4<p\leq P} 1.$$

Il faut remarquer que P n'est pas nécessairement un entier et, pour fixer les idées, $C(12) = 3$.

Nous utiliserons beaucoup une forme particulière des exponentielles complexes, nommément $\exp(2i\pi x)$ quand x est un réel et que nous notons $e(x)$. Une façon de définir cette fonction est de dire $e(x) = \cos(2\pi x) + i\sin(2\pi x)$. Elle a surtout l'avantage d'avoir une formule d'addition bien plus simple que celle du cosinus ou du sinus : $e(x + x') = e(x)e(x')$. Le complexe $e(x)$ est de module 1 et correspond par conséquent dans le plan complexe à un point sur le cercle unité, de centre 0 et de rayon 1.

L'inégalité $\sin u \geq 2u/\pi$ a lieu si $0 \leq u \leq \pi/2$ ce qui nous permet d'obtenir $|\sin(\pi v)| \geq 2\|v\|$ où $\|v\|$ désigne la distance au plus proche entier et ensuite

$$\left|\sum_{n\in I} e(nv)\right| \leq \frac{1}{2\|v\|} \tag{1.1}$$

pour tout intervalle I si v n'est pas congru à 0 modulo 1. En effet, cela est évident si I ne contient aucun point entier et sinon il suffit de sommer les termes d'une progression géométrique entre $n = n_1$ et n_2 où n_1 et n_2 sont respectivement le plus petit et le plus grand entier que contient I. Le module du résultat est alors $|\sin(\pi(n_2 - n_1 + 1)v)/\sin(\pi v)|$ où la minoration de sinus donnée ci-dessus permet de conclure. Si I est grand et v pas trop proche de 0, l'inégalité (1.1) nous dit que les $e(nv)$ sont bien répartis sur le cercle unité de sorte que lorsqu'on les ajoute, leur contribution a tendance à se compenser. L'exemple où $v = 1/2$ est emblématique de ce qui se passe, où nous sommons simplement $(-1)^n$.

Les accolades et les crochets sont réservés à un usage particulier : $[x]$ est la partie entière du réel x et $\{x\}$ sa partie fractionnaire, de sorte que $x = [x] + \{x\}$. Comme cet ouvrage est assez petit, nous nous astreignons à ne pas utiliser les accolades et les crochets comme des parenthèses ... Ce

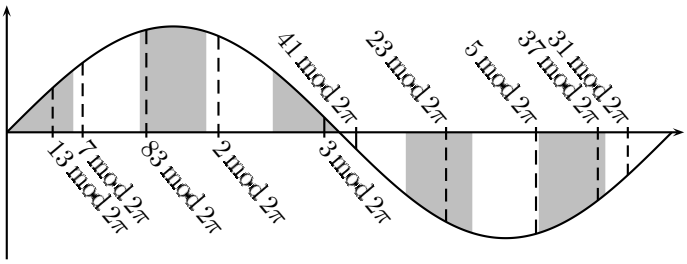
qui, du coup, nous réduit à distinguer les niveaux de parenthèses uniquement selon leur taille ! Une exception notoire concerne les intervalles : $]a, b]$ désigne l'intervalle de bornes a et b, ouvert en a et fermé en b, soit encore l'ensemble des réels x tels que $a < x \leq b$. Ce que la littérature anglo-saxonne note $(a, b]$. La même convention nous fait noter $[a, b]$ l'intervalle fermé en a et b, etc. Par contre nous noterons le coefficient binomial C_n^k sous la forme d'origine anglo-saxonne $\binom{n}{k}$, car c'est celle qui prévaut à l'heure actuelle en combinatoire. Il s'agit bien sûr de $n!/(k!(n-k)!)$. Enfin, l'auteur que je suis est d'un autre temps où le logarithme usuel, celui que l'on qualifie aussi de népérien d'après John Neper qui écrivit en 1614 un mémoire sur ce qui allait devenir cette fonction (mais la notion de fonction était alors mal comprise), où le logarithme usuel donc se notait Log. La notation log était réservée au logarithme en base 10... Aujourd'hui ln est préféré à Log, simplement parce qu'à une époque, il a été difficile d'inscrire Log sur les calculatrices. Cette époque est révolue et rien ne s'oppose au retour de Log, notation qui est effectivement la plus usitée malgré la courte période ln-iste. Nous utiliserons en conséquence Log ici.

Enfin, une comparaison à une intégrale (un exemple de cette technique se trouve page 8) donne l'estimation suivante valable pour tout réel $N \geq 3$:

$$\sum_{1 \leq n \leq N} 1/n \leq \operatorname{Log} N + 1 \leq 2 \operatorname{Log} N. \tag{1.2}$$

Le plan de la bataille

Essayons tout d'abord de nous faire une idée du chemin à parcourir. Commençons par découper l'intervalle $[0, 2\pi]$ en 10 parties égales et regardons si nous pouvons trouver des nombres premiers p tels que $p \bmod 2\pi$ tombe dans chacun de ces intervalles. Ici $p \bmod 2\pi$ désigne l'unique réel x tel que, d'une part x appartienne à $[0, 2\pi[$ et, d'autre part, tel que p et x ne diffèrent que d'un multiple entier de 2π. Dit autrement, x est égal à $2\pi \left\{ \frac{p}{2\pi} \right\}$. Nous trouvons effectivement de tels nombres premiers et obtenons par exemple :



Revenons à notre problème initial, soit la densité des sinus p dans $[-1, 1]$. Partons de u dans $[-1, 1]$. Nous pouvons l'écrire sous la forme $u = \sin x$, pour un certain x appartenant à $[0, 2\pi]$. Il existe un de nos $p \bmod 2\pi$, disons x' qui est à moins de $2\pi/10$ de x. Par conséquent, et grâce à l'inégalité des accroissements finis, $\sin x'$ est aussi à moins de $2\pi/10$ de $u = \sin x$, ce qui établi que tout intervalle de $[-1, 1]$ de longueur $\geq 2\pi/10$ contient un $\sin p$. Il ne s'agit certes pas d'une immense victoire, car $2\pi/10 = 0.628 + \mathcal{O}^*(10^{-3})$ n'est pas très petit, mais le lecteur constate que nous pourrions remplacer 10 par 1 000 ou plus... Ceci ramène le problème[1] à montrer que les $p \bmod 2\pi$ sont denses dans $[0, 2\pi]$, ou encore que les $\left\{ \frac{p}{2\pi} \right\}$ le sont dans $[0, 1]$.

L'architecture

Pour tout intervalle $[a, b] \subset [0, 1]$ non réduit à un point, nous montrons qu'il existe un nombre premier p tel que la partie fractionnaire de $p/(2\pi)$ appartienne à $[a, b]$. Cela nous permettra d'en déduire que l'ensemble des valeurs prises par $\sin p$ est dense dans $[-1, 1]$. *Un paramètre important tout au*

[1] mais ne lui est pas équivalent, puisque la fonction sinus prend toutes les valeurs possibles sur, par exemple, $[\pi/2, 3\pi/2]$. Nous démontrons un peu plus que nécessaire mais cela va en fait simplifier l'exposition.

long de cette preuve sera $\varepsilon = b - a > 0$. Nous commençons par montrer qu'il y a suffisamment de nombres premiers ce que traduira l'inégalité $C(P) \geq 0.08P/\operatorname{Log} P$ valable pour $P \geq 2$. Nous approcherons ensuite $\alpha = 1/(2\pi)$ par une suite de rationnels r_n/s_n avec un s_n tendant vers l'infini et r_n premier à s_n, ce qui est possible puisque α est irrationnel. Nous allons trouver des nombres premiers répondant à notre question dans l'intervalle $]P_n/4, P_n]$ pourvu que n soit assez grand, où $P_n = s_n^{3/2}$. Pour cela, nous établissons tout d'abord l'existence d'un entier M dépendant de a et b mais pas de n tel qu'il suffisse d'établir que

$$\forall m \in \{1, \ldots, M\}, \qquad \sum_{P_n/4 < p \leq P_n} e(imr_n p/s_n)/C(P_n) \underset{n\to\infty}{\longrightarrow} 0 \tag{H}$$

où, rappelons-le une dernière fois, p désigne un nombre premier. *Pour simplifier la typographie, nous utiliserons l'abbréviation* $L_n = \operatorname{Log} P_n$ *tout au long des chapitres ultérieurs.* Nous approchons alors la suite des nombres premiers de notre intervalle par la suite des entiers ℓ qui sont premiers à tous les entiers $\leq z$. Bien sûr, si $z = \sqrt{P_n}$, la suite approximante serait égale à la suite de départ, aussi nous prenons z beaucoup plus petit. Cette approximation est effectuée grâce à une identité combinatoire et le terme reste a une structure particulière : il s'agit d'une somme sur deux variables de termes oscillants. Nous exploitons alors cette structure en appliquant l'inégalité de Cauchy-Schwarz ce qui établira que ce terme est effectivement négligeable. Il nous reste alors à montrer que la suite des entiers ℓ vérifie une condition analogue à (H), mais avec une somme portant sur ℓ et non sur p, ce que nous faisons en exprimant la fonction caractéristique de cet ensemble en termes de la fonction de Möbius.

Estimations classiques sur les nombres premiers

Nous allons évidemment avoir besoin de quantifier le nombre de nombres premiers dans un intervalle. De telles estimations existent depuis longtemps, et le travail consistant à les rendre explicite numériquement a débuté dans les années quarante. Nous disposons maintenant de bonnes estimations pour les quantités simples, mais elles sortent du cadre de ce livre. Nous nous proposons ici d'employer une approche classique afin de rendre la preuve de la densité des $\sin p$ complète, ce qui nous préparera au passage pour les derniers chapitres. Nous en profitons aussi pour introduire deux techniques simples et efficaces.

La fonction de von Mangoldt

Commençons par introduire la fonction dite de van Mangoldt, d'après le mathématicien allemand Hans von Mangoldt. Celui-ci a publié en 1894 un mémoire important sur les nombres premiers où il introduit notamment cette fonction, mais sous la forme $L(n)$; la notation d'usage à l'heure actuelle est $\Lambda(n)$ et c'est celle que nous emploierons. Voici la définition de cette fonction :

$$\Lambda(n) = \begin{cases} \operatorname{Log} p & \text{si } n = p^a \text{ avec } a \geq 1, \\ 0 & \text{sinon.} \end{cases} \tag{3.1}$$

Par exemple $\Lambda(2) = \Lambda(4) = \operatorname{Log} 2$ et $\Lambda(15) = 0$. Notons explicitement que $\Lambda(1) = 0$. L'apparition de cette fonction n'est en rien mystérieuse si l'on raisonne en termes de séries de Dirichlet mais nous évitons ce point de vue ici. Du coup, la justification de son introduction vient de deux aspects. Tout d'abord, elle permet d'isoler les puissances des nombres premiers des autres entiers tout en leur attribuant le poids assez peu fluctuant $\operatorname{Log} p$, et nous verrons au lemme 3.5 que la contribution des p^2, p^3, ... est négligeable devant celle des nombres premiers p.

Cela étant, son intérêt véritable résulte de l'identité :

$$\forall n \text{ (entier) } \geq 1, \quad \sum_{d|n} \Lambda(d) = \operatorname{Log} n, \tag{3.2}$$

où la somme porte sur tous les diviseurs $d \geq 1$ de n.

PREUVE. Pour $n = 1$, cette identité est évidente. Pour n plus grand, nous le décomposons en facteurs premiers $n = p_1^{a_1} p_2^{a_2} \cdots p_K^{a_K}$ où les p_i sont des nombres premiers distincts et les a_i des entiers ≥ 1. Il vient alors

$$\operatorname{Log} n = a_1 \operatorname{Log} p_1 + a_2 \operatorname{Log} p_2 + \cdots + a_K \operatorname{Log} p_K$$

et les diviseurs d de n pour lesquels $\Lambda(d) \neq 0$ sont les $p_1^{b_1}$ avec $1 \leq b_1 \leq a_1$ (il y en a a_1 de cette forme), puis les $p_2^{b_2}$ avec $1 \leq b_2 \leq a_2$ (il y en a a_2 de cette forme), etc. Et bien sûr, $a_1 \operatorname{Log} p_1 = \sum_{b_1} \operatorname{Log} p_1$, ce qui permet de clore la preuve. ◇◇◇

De la fonction Log à la fonction Λ

Commençons par un lemme classique d'analyse :

Lemme 3.1 *Pour $x \geq 2$, nous avons*

$$\sum_{1 \leq n \leq x} \operatorname{Log} n = x \operatorname{Log} x - x + \mathcal{O}^*(\operatorname{Log}(2x)).$$

PREUVE. Soit N la partie entière de x. Nous procédons par comparaison à une intégrale, c'est à dire que nous utilisons les inégalités

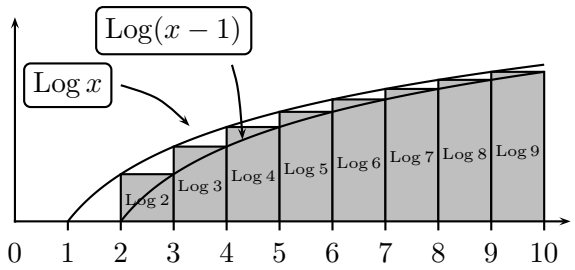
$$\int_{n-1}^{n} \operatorname{Log} t \, dt \leq \operatorname{Log} n \leq \int_{n}^{n+1} \operatorname{Log} t \, dt,$$

lesquelles sont une simple conséquence du caractère croissant du logarithme. En les sommant, nous obtenons

$$\int_{1}^{N} \operatorname{Log} t \, dt \leq \sum_{2 \leq n \leq N} \operatorname{Log} n \leq \int_{2}^{N+1} \operatorname{Log} t \, dt \tag{3.3}$$

et un peu de travail permet de conclure, moyennant de se souvenir que $x \mapsto x \operatorname{Log} x - x$ est une primitive de $x \mapsto \operatorname{Log} x$.

Voici une interprétation graphique de l'encadrement :



La somme cumulée des aires des petits rectangles est la quantité qui nous intéresse. Nous constatons graphiquement qu'elle est majorée par l'intégrale de la fonction, soit le membre de droite de (3.3), et minorée par l'intégrale de cette fonction translatée d'une unité vers la gauche, soit le membre de gauche de (3.3). ◇◇◇

Nous écrivons alors

$$\sum_{n\leq x}\operatorname{Log} n=\sum_{n\leq x}\sum_{d|n}\Lambda(d)=\sum_{d\leq x}\Lambda(d)[x/d]$$

et l'idée de tout ce qui suit est basée sur l'égalité :

$$\sum_{d\leq x}\Lambda(d)[x/d]=x\operatorname{Log}x-x+\mathcal{O}^*(\operatorname{Log}(2x)). \tag{3.4}$$

Une majoration à la Chebyshev

Théorème 3.2 *Pour tout $x\geq 1$, la majoration suivante est valide :*

$$\sum_{x/2<d\leq x}\Lambda(d)\leq 7x/10.$$

Preuve. Une utilisation directe de (3.4) donne

$$\sum_{d\leq x}\Lambda(d)\big([x/d]-2[x/(2d)]\big)=x\operatorname{Log}2+\mathcal{O}^*(2\operatorname{Log}(2x)).$$

Nous remarquons maintenant que $[x]-2[x/2]\geq 0$ pour tout x réel : en effet, cette quantité vaut 0 si x est dans $[0,1[$, puis 1 si x est dans $[1,2[$ et enfin est périodique de période 2. Par conséquent, l'équation ci-dessus implique

$$\sum_{x/2<d\leq x}\Lambda(d)\big([x/d]-2[x/(2d)]\big)\leq x\operatorname{Log}2+2\operatorname{Log}(2x).$$

Pour les d entre $x/2$ et x, nous avons $[x/d]-2[x/(2d)]=1$ ce qui aboutit à

$$\sum_{x/2<d\leq x}\Lambda(d)\leq x\operatorname{Log}2+2\operatorname{Log}(2x).$$

Cette inégalité permet de prouver le théorème si $x\geq 1150$ et une vérification numérique (nous parlerons plus avant de cet aspect dans le chapitre suivant) permet d'étendre ce résultat à tout x réel ≥ 1. ◇◇◇

En découpant l'intervalle $[1,x]$ entre $]x/2,x]$ union $]x/4,x/2]$ union etc, nous obtenons le corollaire classique :

Corollaire 3.3 *Nous avons $\sum_{d\leq x}\Lambda(d)\leq 7x/5$ pour tout $x\geq 1$.*

Pafnouty Chebyshev est le premier à avoir établi en 1848 une telle estimation, par une méthode d'ailleurs proche de celle que nous avons développée. Notons ici que John Rosser a montré en 1941 que le maximum de la fonction $\sum_{d\leq x}\Lambda(d)/x$ était atteint en $x=113$ et était un peu inférieur à 1.04.

Ceci nous donne aussi une majoration du nombre de nombres premiers inférieurs à une borne donnée :

Corollaire 3.4 *Pour tout $x \geq 1$, le nombre de nombres premiers inférieurs à x est au plus $3x/(2\operatorname{Log} x)$.*

Signalons les notations traditionnelles $\pi(x) = \sum_{p\leq x} 1$ et $\psi(x) = \sum_{d\leq x} \Lambda(d)$ dont nous n'userons pas.

PREUVE. Remarquons tout d'abord que ce nombre de nombres premiers s'écrit aussi $\sum_{p\leq x} 1$. Or, nous tirons du corollaire précédent la majoration : $\sum_{p\leq x} \operatorname{Log} p \leq 7x/5$. Nous nous débarrassons alors du poids $\operatorname{Log} p$ par une technique que l'on appelle *la sommation par parties* du fait que, dans le formalisme de l'intégrale de Stieltjes, il s'agit effectivement de l'extension de la technique du même nom standard au niveau du calcul intégral. Une version souple et élémentaire s'obtient en écrivant :

$$\frac{1}{\operatorname{Log} p} = \frac{1}{\operatorname{Log} x} + \int_p^x \frac{dt}{t\operatorname{Log}^2 t}.$$

Il vient alors

$$\sum_{p\leq x} \frac{\operatorname{Log} p}{\operatorname{Log} p} = \sum_{p\leq x} \operatorname{Log} p \left(\frac{1}{\operatorname{Log} x} + \int_p^x \frac{dt}{t\operatorname{Log}^2 t} \right)$$

ce qui nous donne

$$\begin{aligned} \sum_{p\leq x} \frac{\operatorname{Log} p}{\operatorname{Log} p} &= \frac{\sum_{p\leq x} \operatorname{Log} p}{\operatorname{Log} x} + \int_2^x \Big(\sum_{p\leq t} \operatorname{Log} p \Big) \, \frac{dt}{t\operatorname{Log}^2 t} \\ &\leq \sum_{p\leq x} \frac{\operatorname{Log} p}{\operatorname{Log} p} \leq \frac{7x}{5\operatorname{Log} x} + \int_2^x \frac{7dt}{5\operatorname{Log}^2 t}. \end{aligned}$$

Pour la dernière intégrale, nous commençons par remarquer qu'une intégration par parties (classique !) implique, pour $k \geq 0$, que :

$$J_k = \int_2^x \frac{dt}{\operatorname{Log}^k t} \leq \frac{x}{\operatorname{Log}^k x} + \int_2^x \frac{k\,dt}{\operatorname{Log}^{k+1} t}$$

d'où nous déduisons que $(\operatorname{Log} x - 3)(\operatorname{Log}^2 x)J_3 \leq x$ et $J_2 \leq x/\operatorname{Log}^2 x + J_3$. Ce qui à terme nous donne le résultat annoncé si $x \geq \exp(5)$. Un calcul finit. ◇◇◇

La fonction Λ donne aussi un poids non nul à des entiers qui ne sont pas des nombres premiers mais des puissances de ceux-ci. Leur contribution est la plupart du temps négligeable grâce au lemme suivant :

Lemme 3.5 *Pour $x \geq 1$, nous avons*

$$\sum_{\substack{d\leq x \\ d \text{ non premier}}} \Lambda(d) \leq 3\sqrt{x}.$$

PREUVE. En effet, les d comptés s'écrivent p^a avec $a \geq 2$ et $p \leq \sqrt{x}$. Un nombre premier va apparaître en p, puis p^2, et caetera jusqu'à p^a où $a \leq (\operatorname{Log} x)/\operatorname{Log} p$. Le corollaire précédent conclut. ◇◇◇

Un théorème à la Mertens

Nous en arrivons à un théorème dans l'esprit d'un résultat de Franz Mertens issu d'un mémoire de 1874. Il contient en essence une *minoration* du nombre de nombres premiers comme nous le montrons ci-après.

Théorème 3.6 *Pour $x \geq 2$, l'égalité suivante a lieu*

$$\sum_{d\leq x} \Lambda(d)/d = \operatorname{Log} x - \tfrac{1}{4} + \mathcal{O}^*(\tfrac{4}{5})$$

PREUVE. Nous partons toujours de (3.4) et écrivons cette fois-ci $[x] = x - \{x\}$ où nous majorons la partie fractionnaire par 1 et la minorons par 0. Il vient

$$\frac{1}{x}\sum_{d\leq x} \Lambda(d) + \operatorname{Log}(2x)/x \geq \sum_{d\leq x} \Lambda(d)/d - \operatorname{Log} x + 1 \geq -\operatorname{Log}(2x)/x.$$

Pour $x \geq 125$ et moyennant d'invoquer encore le Corollaire 3.3, cela donne la borne supérieure $\operatorname{Log} x + \frac{4}{9}$ et la borne inférieure $\operatorname{Log} x - \frac{21}{20}$, ce qui est meilleur que le résultat annoncé. Pour x plus petit, une vérification numérique conclut. ◇◇◇

Un résultat de type postulat de Bertrand

Joseph Bertrand conjecturait en 1845 qu'il y a toujours un nombre premier dans l'intervalle $[n, 2n-3]$ si n est un entier ≥ 4, conjecture qui devait être démontrée par Chebyshev en 1850. Les résultats que nous avons établis sont un peu plus faibles que ceux dont disposaient Chebyshev mais nous permettent de démontrer un résultat du même genre, à savoir :

Théorème 3.7 *Pour $x \geq 2$, nous avons*

$$C(x) = \sum_{x/4<p\leq x} 1 \geq 2x/(25 \operatorname{Log} x).$$

Il existe donc un nombre premier dans l'intervalle $]x/4, x]$ pour tout $x \geq 2$. Pour arriver au postulat de Bertrand, nous pourrions rechercher une inégalité sur les parties entières autre que $[x] - 2[x/2] \geq 0$ utilisée dans la preuve du théorème 3.2 ; c'est le chemin que suivit Chebyshev autour de 1850. Ou inclure dans ce théorème 3.2 la contribution des entiers entre $x/4$ et $x/8$ et modifier conséquemment le théorème 3.6.

PREUVE. En appliquant le théorème 3.6 en x et $x/4$, nous obtenons

$$\sum_{x/4<d\leq x} \Lambda(d)/d \geq \operatorname{Log} 4 - 1 \tag{3.5}$$

et par conséquent $\sum_{x/4<d\leq x}\Lambda(d)\geq x(\operatorname{Log}4-1)/4\geq x/11$ pour $x\geq 16$. Nous étendons cette inégalité à $x\geq 2$ par le calcul et il s'agit ensuite de passer de $\Lambda(d)$ à une somme sur les nombres premiers, ce que nous effectuons à l'aide du lemme 3.5, obtenant

$$\sum_{x/4<p\leq x}\operatorname{Log}p\geq\frac{x}{11}-3\sqrt{x}\geq 2x/25$$

si $x\geq 76\,000$, d'où le théorème dans ce cas. Un calcul numérique (détaillé au chapitre suivant) permet d'étendre ce résultat. ◇◇◇

Concernant de bonnes approximations de la somme des nombre premiers $\leq x$, signalons ici que dans la continuité des travaux de Rosser, Pierre Dusart a établi en 1999 que, pour $x\geq 598$

$$1+\frac{0.992}{\operatorname{Log}x}<\frac{\operatorname{Log}x}{x}\sum_{p\leq x}1<1+\frac{1.2762}{\operatorname{Log}x}. \tag{3.6}$$

Nous aurons encore besoin à la fin de ce livre d'un dernier résultat que voici.

Lemme 3.8 *Pour $x\geq 2$, nous avons*

$$\sum_{p\leq x}1/p=\operatorname{Log}\operatorname{Log}x-1+\mathcal{O}^*(7/6).$$

PREUVE. Le théorème de Mertens implique que

$$-\tfrac{21}{20}-\sum_{k\geq 2,p\geq 2}\frac{\operatorname{Log}p}{p^k}\leq\sum_{p\leq x}\frac{\operatorname{Log}p}{p}-\operatorname{Log}x\leq\tfrac{11}{20}.$$

En sommant d'abord sur k et en utilisant un peu de calcul numérique, nous montrons que la somme sur k et p qui apparaît vaut au plus 0.8. Pour obtenir le lemme, nous utilisons une sommation par parties comme page 10. Il en résulte la majoration

$$\sum_{p\leq x}1/p\leq\operatorname{Log}\operatorname{Log}x+1+\frac{11}{20\operatorname{Log}2}-\operatorname{Log}\operatorname{Log}2$$

et la minoration $\operatorname{Log}\operatorname{Log}x+1-\frac{21}{20\operatorname{Log}2}-\operatorname{Log}\operatorname{Log}2$. ◇◇◇

Approximation par des rationnels

Pour montrer que les parties fractionnaires des $p/(2\pi)$ sont denses dans $[0,1]$, il nous faut un argument qui dise que $1/(2\pi)$ est irrationnel. En effet et à titre d'exemple, les seules valeurs que prenne $\{p/5\}$ sont 0, 1/5, 2/5, 3/5 et 4/5 ; cette suite ne saurait en aucun cas être dense dans $[0,1]$! Il nous faut ensuite traduire cette irrationnalité de façon plus quantitative, ce que nous réaliserons à l'aide des approximations rationnelles de ce nombre. Pour l'exposition, nous nous intéressons d'abord aux propriétés d'approximation pour conclure par une démonstration de l'irrationalité de π^2, dont découle évidemment celle de $1/(2\pi)$.

Approximations rationnelles sur le cercle

Le théorème suivant constitue la base de l'approximation diophantienne. L'une de ses particularités tient à sa preuve qui est particulièrement limpide lorsqu'elle prend un appui géométrique :

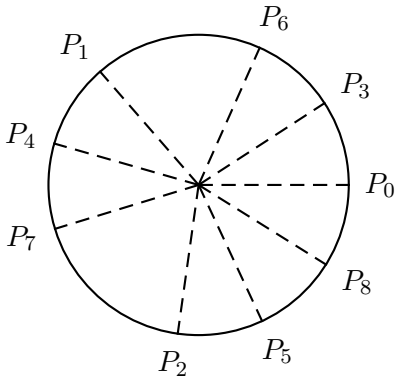
Théorème 4.1 *Soit q un entier ≥ 1 et β un nombre réel. Il existe deux entiers u et v tels que*

$$|v\beta - u| \leq \frac{1}{q+1} \qquad 1 \leq v \leq q.$$

Preuve.

La preuve est on ne peut plus simple et repose sur une transcription géométrique du problème. Considérons un cercle de circonférence 1. Nous y fixons un point P_0 qui est notre 0 et avançons de β sur le cercle, quitte à faire plus d'un tour. Nous atteignons un point P_1. Puis nous recommençons et construisons

Avec $\beta = 0.3660563$:



ainsi P_2, P_3, ... ,P_q, lequel correspond donc à $q\beta$. Nous disposons de $q+1$ point sur une circonférence de longueur 1, il y en a par conséquent deux qui sont espacés de moins de $1/(q+1)$ car sinon, la longueur totale du cercle serait $> (q+1)/(q+1) = 1$. Disons qu'il s'agisse de P_ℓ et $P_{\ell'}$ avec $\ell < \ell'$. Dire que P_ℓ et $P_{\ell'}$ sont à moins de $1/(q+1)$ signifie que

$$\big|\{\ell'\beta\} - \{\ell\beta\}\big| \leq \frac{1}{q+1}$$

soit encore qu'il existe un entier u tel que

$$\big|(\ell'-\ell)\beta - u\big| \leq \frac{1}{q+1}.$$

Nous posons $v = \ell' - \ell$, ce qui prouve notre théorème. ◇◇◇

D'où nous déduisons :

Corollaire 4.2 *Soit β un nombre irrationnel. Il existe deux suites $(r_n)_{n\geq 1}$ et $(s_n)_{n\geq 1}$, d'entiers strictement positifs pour s_n et d'entiers pour r_n et telles que*

$$\left|\beta - \frac{r_n}{s_n}\right| < \frac{1}{s_n^2} \quad , \quad \text{pgcd}(r_n, s_n) = 1 \quad , \quad s_n \to \infty.$$

PREUVE. Pour tout entier n, nous prenons $q = n-1$ dans le théorème précédent et obtenons donc des entiers u_n et v_n tels que

$$\left|\beta - \frac{u_n}{v_n}\right| \leq \frac{1}{nv_n} \quad , \quad 1 \leq v_n < n.$$

Nous réduisons u_n/v_n en fraction irréductible, i.e. écrivons $u_n/v_n = r_n/s_n$ avec $\text{pgcd}(r_n, s_n) = 1$ et $r_n \geq 1$. Bien sûr s_n est aussi inférieur à v_n. Il vient

$$\left|\beta - \frac{r_n}{s_n}\right| \leq \frac{1}{ns_n} \leq \frac{1}{s_n^2}$$

et il nous faut à présent montrer que la suite (s_n) tend vers l'infini. Supposons que ce ne soit pas le cas. Cela signifie qu'il existe une borne entière B et une infinité de n tels que $s_n \leq B$. Comme il y a au plus B entiers inférieurs à B, il existe une valeur $b \leq B$ pour laquelle $s_n = b$ a lieu pour un infinité de nos n. Disons que cela a lieu pour $n \in \mathcal{N}$ pour un certain ensemble $\mathcal{N}$. Nous avons alors

$$\forall n \in \mathcal{N}, \quad |b\beta - r_n| \leq \frac{1}{n}.$$

En faisant tendre n vers l'infini mais en restant dans $\mathcal{N}$, nous constatons que r_n tend vers $b\beta$, qui est par conséquent entier et donc $b\beta = r_n$ dès que $n \geq 2$ et dans $\mathcal{N}$. Ceci signifie que β est rationnel, contrairement à nos hypothèses. ◇◇◇

Il est aussi possible d'atteindre ce corollaire en passant par la théorie des *fractions continues*. C'est une théorie plus précise que notre approche puisqu'elle donne des expressions pour les s_n et r_n, mais qui est aussi plus longue à mettre en place. De façon rapide, on part d'un nombre, disons $\beta = \beta_0 = 1/(2\pi)$ et on l'écrit sous la forme

$$\beta_0 = [\beta_0] + \frac{1}{\beta_1}$$

où $[\beta_0]$ désigne, rappelons-le, la partie entière de β_0. Puis on recommence avec β_1. Par exemple, ici, $[\beta_0] = 0$ et $\beta_1 = 2\pi$, dont cette fois-ci la partie entière est 6, ce qui nous donne

$$\beta_0 = \frac{1}{6 + 1/\beta_2}.$$

En remplaçant $1/\beta_2$ par 0, nous obtenons notre première approximation : $1/6$; la théorie nous garantit alors la borne

$$|\beta - 1/6| \leq 1/6^2$$

que nous vérifions numériquement. Ensuite la partie entière de $1/\beta_2$ est 3 ce qui nous donne

$$\beta_0 = \frac{1}{6 + \frac{1}{3+1/\beta_3}}.$$

En remplaçant $1/\beta_3$ par 0, cela nous donne la fraction $3/19$ et nous avons

$$|\beta - 3/19| \leq 1/19^2.$$

Nous continuons ainsi et la théorie dit que les approximations ainsi générées sont excellentes, en ce qu'elles vérifient les hypothèses de notre corollaire.

L'irrationnalité de $\alpha = 1/(2\pi)$

Pour pouvoir appliquer notre corollaire, il faut établir que $\alpha = 1/(2\pi)$ est irrationnel, ce qui équivaut évidemment à montrer que π est irrationnel. Nous établissons ici ce fait en suivant la simplification dûe à Ivan Niven en 1947 de la preuve originelle de Johann Lambert qui elle date de 1761. En l'occurence, nous montrons plus, à savoir que π^2 n'est pas rationnel, dont il découle évidemment que π ne saurait être rationnel.

Nous commençons par une inégalité obtenue par une méthode simple et efficace et que l'on appelle *la méthode de Rankin additive* en liaison avec *la méthode de Rankin multiplicative* qui apparaîtra plus loin :

Lemme 4.3 *Pour $n \geq 1$, nous avons $n! \geq (n/e)^n$.*

PREUVE. Donnons-nous un paramètre x réel positif. Nous écrivons

$$\frac{x^n}{n!} \leq \sum_{m\geq 0} \frac{x^m}{m!} = e^x$$

ce qui résulte en $n! \geq x^n e^{-x}$ pour tout $x \geq 0$. Qu'il nous suffit de choisir au mieux de nos intérêts ! Nous prenons $x = n$ ce qui clôt la preuve. ◇◇◇

Théorème 4.4 *π^2 est irrationnel.*

PREUVE. Nous déployons la preuve en deux actes. Nous considérons tout d'abord la fonction

$$f(x) = \frac{x^n(1-x)^n}{n!} = \frac{1}{n!}\sum_{k=0}^{n} \binom{n}{k} (-1)^k x^{n+k}$$

pour un paramètre entier n que nous choisirons en fin de démonstration. Cette fonction vérifie $0 < f(x) < 1/n!$ si x est dans $]0,1[$. La première expression de f nous garantit que $f^{(s)}(0)$, la dérivée d'ordre s en zero, est nulle si s est un entier entre 0 et $n-1$ alors que la seconde expression nous donne (pour $m = 0, \ldots, n$)

$$f^{(n+m)}(x) = \frac{1}{n!}\sum_{k=m}^{n} \binom{n}{k} \frac{(n+k)!}{(k-m)!} (-1)^k x^{k-m}$$

d'où nous concluons que $f^{(n+m)}(0)$ est un entier. Finalement, les dérivées successives de f en 0 sont des entiers et comme $f(1-x) = f(x)$, il en va de même des dérivées successives en 1. Voilà qui clôt le premier acte.

Le second acte s'ouvre avec une hypothèse : supposons que π^2 soit rationnel, c'est à dire s'écrive a/b avec des entiers strictement positifs a et b. La suite consiste en deux remarques qui concernent l'intégrale

$$I = \int_0^1 a^n \pi f(x) \sin(\pi x) dx.$$

Des encadrements de f donnés plus haut, nous tirons $0 < I < \pi a^n/n!$, qui est < 1 si n est assez grand en vertu du lemme précédent. Mais, et c'est là la clé de la preuve, nous montrons ci-après que cette intégrale est un entier ! Ce qui établira notre théorème. En effet définissons

$$F(x) = b^n \Big(\pi^{2n} f(x) - \pi^{2n-2} f''(x) + \pi^{2n-4} f^{(4)}(x) - \cdots + (-1)^n f^{(2n)}(x) \Big).$$

Nous constatons, grâce au premier acte, que $F(0)$ et $F(1)$ sont des entiers. Cette fonction est liée à notre problème par la relation

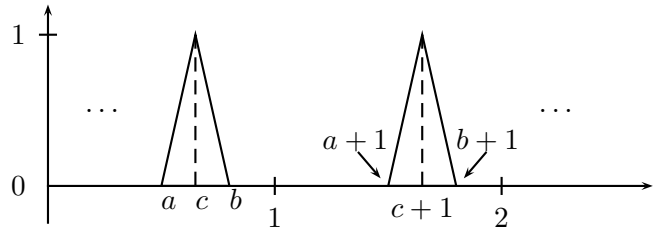
$$\frac{d}{dx}\left(F'(x)\sin\pi x - \pi F(x)\cos\pi x\right) \\ = \left(F''(x) + \pi^2 F(x)\right)\sin\pi x = \pi^2 a^n f(x)\sin\pi x$$

tant et si bien que $I = F(0) + F(1)$. Et la preuve est terminée. ◇◇◇

On peut se demander ici quelle est la définition de π utilisée, et la lecture de la preuve précédente est limpide : π est le plus petit zéro strictement positif de la fonction sinus.

Un peu d'analyse de Fourier

Nous considérons la fonction f triangulaire périodique de période 1 et dont le graphe est (avec $c = \frac{a+b}{2}$) :



où, de façon analytique :

$$\begin{cases} f(t) = 0 & \text{si } t \le a \text{ ou } t \ge b \text{ modulo } 1, \\ f(t) = 2\frac{t-a}{b-a} & \text{si } a \le t \le c \text{ modulo } 1, \\ f(t) = 2\frac{b-t}{b-a} & \text{si } c \le t \le b \text{ modulo } 1. \end{cases} \tag{5.1}$$

Il s'agit en fait du noyau de Leopold Fejér périodique. Rappelons que $\varepsilon = b-a$. Nous développons f en série de Fourier et tronquons les termes $|m| > M$, où M est un paramètre entier ≥ 1 à choisir :

$$f(t) = \frac{\varepsilon}{2} + \sum_{0<|m|\le M} \frac{A(m)}{m^2} e(mt) + \mathcal{O}^*\left(\frac{4}{M\varepsilon}\right), \tag{5.2}$$

$A(m)$ étant un coefficient borné par $2/\varepsilon$ et donné par (5.4). Rappelons aussi que $e(x) = \exp(2i\pi x)$.

Preuve. Comme f est une fonction continue C^1 par morceaux, la convergence de la série de Fourier usuelle vers f ne pose pas de problème, mais par souci de complétude, une preuve directe est proposée dans la section suivante. Les coefficients $A(m)$ sont donnés pour $m \neq 0$ par

$$A(m)/m^2 = \int_0^1 f(u)e(-mu)du \tag{5.3}$$

et un petit calcul en utilisant (5.1) aboutit à

$$A(m) = 2\sin^2(m\pi\varepsilon/2)e(-mc)/\varepsilon. \tag{5.4}$$

Comme $\sin^2 \leq 1$, nous constatons que la contribution des termes pour lesquels $|m| > M$ est majorée par

$$2 \sum_{M+1 \leq m} \frac{2}{\varepsilon m^2} \leq \frac{4}{\varepsilon} \int_M^\infty \frac{dt}{t^2} = \frac{4}{M\varepsilon}.$$

◇◇◇

Digression : une approche directe

Nous avons utilisé dans la section précédente le fait que la série de Fourier convergeait vers la fonction si celle-ci était continue avec des coefficients de Fourier tendant assez vite vers 0. Nous proposons ici une preuve directe de

$$f(t) = \frac{b-a}{2} + \sum_{0<|m|} A(m)e(mt)/m^2. \tag{5.5}$$

Appelons $h(t)$ le membre de droite. Considérons

$$h_M(t) = \varepsilon/2 + \sum_{0<|m|\leq M} \frac{A(m)}{m^2}\left(1 - \frac{|m|}{M}\right)e(mt) \tag{5.6}$$

que nous allons montrer converger simultanément vers $h(t)$ et vers $f(t)$, quand M tend vers l'infini. Ceci nous donnera bien évidemment $h(t) = f(t)$.

Commençons par la partie la plus facile et qui consiste à établir que cette suite tend vers $h(t)$. En effet, le calcul qui mène à (5.2) nous garantit tout d'abord que, pour tout M' entier ≥ 1, nous avons

$$h(t) = \frac{\varepsilon}{2} + \sum_{0<|m|\leq M'} \frac{A(m)}{m^2} e(mt) + \mathcal{O}^*\left(\frac{4}{M'\varepsilon}\right).$$

Pour tout M' entier ≥ 1 mais de surcroît $\leq M$, cette même chaîne de calcul nous permet aussi d'écrire

$$h_M(t) = \frac{\varepsilon}{2} + \sum_{0<|m|\leq M'} \frac{A(m)}{m^2}\left(1 - \frac{|m|}{M}\right)e(mt) + \mathcal{O}^*\left(\frac{4}{M'\varepsilon}\right)$$

tout simplement parce que le facteur additionnel $(1 - \frac{|m|}{M})$ reste inférieur à 1 en valeur absolue (cela nous suffit, mais il est vrai que ce facteur est même positif !). Il vient alors

$$\left|h(t) - h_M(t)\right| \leq \sum_{0<|m|\leq M'} \frac{|A(m)|}{m^2}\frac{|m|}{M} + \frac{8}{M'\varepsilon}.$$

Nous majorons alors $|m|/M$ par M'/M et remarquons que la démonstration qui mène à (5.2) nous assure encore que la somme sur tous les m non nuls de $|A(m)|/m^2$ est majorée à $4/\varepsilon$. Ceci nous permet finalement d'écrire en choisissant $M' = [M] + 1$

$$\big|h(t) - h_M(t)\big| \leq \frac{4M'}{M\varepsilon} + \frac{8}{M'\varepsilon} \leq \frac{12}{\sqrt{M}\varepsilon} + \frac{4}{M\varepsilon}$$

qu tend bien vers 0 quand M tend vers l'infini.

Abordons à présent le problème de la convergence de $h_M(t)$ vers $f(t)$. Le premier point consiste à remarquer que

$$\sum_{-M\leq m\leq M} \big(M - |m|\big)e(mu) = \left|\sum_{0\leq \ell\leq M} e(\ell u)\right|^2 = \left|\frac{\sin \pi M u}{\sin \pi u}\right|^2.$$

Une fois cela noté, nous partons de l'expression, pour m non nul :

$$A(m)/m^2 = \int_0^1 f(u)e(-mu)du$$

(et similairement pour $m = 0$) ; celle-ci nous amène à

$$h_M(t) = \int_0^1 f(u) \sum_{-M\leq m\leq M} \left(1 - \frac{|m|}{M}\right)e(m(t-u))du$$

soit encore, en invoquant l'identité ci-dessus et avec $v = t - u$,

$$h_M(t) = \int_{t-1}^t f(t-v)\left|\frac{\sin \pi M v}{\sin \pi v}\right|^2 dv/M.$$

En déroulant dans l'autre sens les égalités qui ont mené à cette expression de $h_M(t)$, nous constatons que

$$\int_{t-1}^t \left|\frac{\sin \pi M v}{\sin \pi v}\right|^2 dv/M = 1.$$

Par ailleurs f vérifie $|f(t-v) - f(t)| \leq 2\|v\|/\varepsilon$, d'où

$$\int_{t-1}^t |f(t-v) - f(t)|\left|\frac{\sin \pi M v}{\sin \pi v}\right|^2 \frac{dv}{M} \leq \int_{t-1}^t \|v\|\left|\frac{\sin \pi M v}{\sin \pi v}\right|^2 \frac{2dv}{M\varepsilon}$$

quantité qui est inférieure à 2 fois la même intégrale mais entre 0 et 1 ou à 4 fois cette intégrale mais entre 0 et 1/2. En définitive, nous majorons $|h_M(t) - f(t)|$ par

$$\int_0^{1/2} v\left|\frac{\sin \pi M v}{\sin \pi v}\right|^2 \frac{8dv}{M\varepsilon} \leq \int_0^{1/2} \frac{|\sin \pi M v|}{v} \frac{2dv}{M\varepsilon}$$

grâce à l'inégalité $|\sin \pi v| \geq 2|v|$ si $v \leq 1/2$. Cette dernière expression est inférieure à $2(3 + \operatorname{Log} M)/(M\varepsilon)$ comme le lecteur le montrera en séparant selon que $v \leq 1/M$ ou non.

Addendum : une approche alternative

Jean-François Burnol, professeur à l'université de Lille 1, a mis au point une jolie démonstration de (5.2), plus élémentaire dans l'esprit, plus pédestre dans sa mise en place, et que la lectrice trouvera peut être plus à son goût. Cette approche repose sur une identité trigonométrique qu'il faut établir au préalable.

Une identité trigonométrique

Nous établissons ici l'identité suivante, valable pour tout réel x de $[0, 2\pi]$:

$$\sum_{k\geq 1} \frac{\cos(kx)}{k^2} = \frac{1}{4}(x-\pi)^2 - \frac{\pi^2}{12}. \tag{5.7}$$

Nous commençons par un lemme.

Lemme 5.1 *Pour tout $n \geq 1$, nous avons*

$$\Big|\pi - u - 2\sum_{1\leq k\leq n} \frac{\sin(ku)}{k}\Big| \leq \begin{cases} \pi & si\ 0 \leq u \leq \frac{2\pi}{2n+1}, \\ \frac{4\pi}{(2n+1)u} & si\ 0 < u \leq \pi. \end{cases}$$

Preuve. Pour ce faire, nous introduisons la fonction, que l'on nomme aussi le *noyau de Dirichlet*, définie comme suit :

$$D_n(x) = 1 + 2\sum_{1\leq k\leq n} \cos(kx) = \frac{\sin((n+\frac{1}{2})x}{\sin(x/2)}.$$

Soit alors u dans $]0, \pi]$. Nous commençons par remarquer que

$$\pi - u - 2\sum_{1\leq k\leq n} \frac{\sin(ku)}{k} = \int_u^\pi D_n(x)dx.$$

La fonction $h(x) = 1/\sin(x/2)$ est C^1, positive et décroissante sur $]0, \pi]$. Nous intégrons par parties le membre de droite de l'équation précédente pour obtenir

$$\begin{aligned}\int_u^\pi D_n(x)dx &= \frac{2\cos((n+\frac{1}{2})u)}{2n+1}g(u) + \int_u^\pi \frac{2\cos((n+\frac{1}{2})x)}{2n+1}g'(x)dx \\ &\leq \frac{2}{2n+1}g(u) + \int_u^\pi \frac{2}{2n+1}|g'(x)|dx.\end{aligned}$$

Comme $|g'(x)| = -g'(x)$, nous pouvons intégrer le dernier terme et en déduire la majoration $\frac{4}{2n+1}g(u)$. La lectrice vérifiera que cette quantité est bien

inférieure à $4\pi/((2n+1)u)$ comme annoncé. Ceci prouve la seconde inégalité du lemme. En ce qui concerne la première, nous écrivons cette fois-ci

$$u + 2\sum_{1\le k\le n} \frac{\sin(ku)}{k} = \int_0^u D_n(x)dx.$$

Nous supposons ici $0 \le (n+\frac{1}{2})u \le \pi$, ce qui implique que l'intégrand $D_n(x)$ reste positif. Il est par ailleurs toujours $\le 2n+1$ et le lecteur complètera aisément la preuve en combinant ces deux remarques. ◇◇◇

Nous sommes à présent en mesure de montrer (5.7). Tout d'abord, nous remarquons que

$$\Delta_n(x) = \frac{-(\pi-x)^2+x^2}{2} + 2\sum_{1\le k\le n} \frac{\cos(kx)-1}{k^2}$$

s'écrit aussi

$$\Delta_n(x) = \int_0^x \Big(\pi - u - 2\sum_{1\le k\le n} \frac{\sin(ku)}{k}\Big)du.$$

Ceci va nous permettre de montrer qu'il tend vers 0 quand n tend vers l'infini. En effet, nous majorons $|\Delta_n(x)|$ en introduisant les valeurs absolues à l'intérieur de l'intégrale et en invoquant les majorations du lemme précédent. Cela nous garantit que $|\Delta_n(x)|$ est inférieur à

$$\int_0^{\frac{2\pi}{2n+1}} \pi du + \int_{\frac{2\pi}{2n+1}}^{\pi} \frac{4\pi}{(2n+1)u}du \le \frac{2\pi^2 + 4\pi\,\mathrm{Log}(n+\frac{1}{2})}{2n+1}$$

qui tend bien vers 0 comme voulu. Il nous faut pour conclure déterminer la valeur de $B = \sum_{k\ge1} 1/k^2$. Il est bien sûr très classique que cette valeur vaut $\pi^2/6$, mais nous disposons de tout le matériel pour l'établir. Pour cela, nous d'appliquons notre identité en $x = \pi$ et remarquons que

$$\sum_{k\ge1} \frac{(-1)^k}{k^2} + B = \sum_{\ell\ge1} \frac{2}{(2\ell)^2} = B/2.$$

Nous laissons la lectrice finir les calculs nécessaires !

Ajoutons pour finir que nous avons à présent démontré la validité de (5.7) pour x dans $[0,\pi]$; ce domaine se prolonge directement à $[0,2\pi]$ du fait que les deux fonctions, celle du membre de gauche et celle du membre de droite, ont toutes deux des graphes symétriques par rapport à $x = \pi$.

Conclusion

Nous nous contentons d'établir le développement de Fourier de la fonction 2π-périodique $F(x) = f(c + \frac{1}{2\pi}x)$ et laissons au lecteur le soin d'en déduire celui de f.

Soit $g(x)$ la fonction paire et 2π-périodique définie par (5.7). La dernière clé de cette preuve consiste à considérer la fonction, elle aussi paire, $k(x) = g(x+\pi\varepsilon) + g(x-\pi\varepsilon) - 2g(x)$. En séparant selon que x appartient à l'intervalle $[0,\pi\varepsilon]$ ou à $[\pi\varepsilon,\pi]$ et en remarquant que la quantité $\cos(k(x+\pi\varepsilon))+\cos(k(x-\pi\varepsilon))-2\cos(kx)$ est égale à $2\cos(kx)(\cos(k\pi\varepsilon)-1)$, soit encore à $-4\cos(kx)\sin^2(\frac{1}{2}k\pi\varepsilon)$, nous atteignons

$$-\sum_{k\geq 1}\frac{4\sin^2(\frac{1}{2}k\pi\varepsilon)}{k^2}\cos(kx) = \begin{cases}\frac{1}{2}\pi^2\varepsilon^2 - \pi(\pi\varepsilon - x) & (0\leq x\leq \pi\varepsilon),\\ \frac{1}{2}\pi^2\varepsilon^2 & (\pi\varepsilon\leq x\leq\pi).\end{cases}$$

Il est immédiat d'en déduire

$$\frac{\varepsilon}{2}+\sum_{k\geq 1}\frac{4\sin^2(\frac{1}{2}k\pi\varepsilon)}{\pi^2k^2\varepsilon}\cos(kx) = \begin{cases}1-\frac{x}{\pi\varepsilon} & (0\leq x\leq\pi\varepsilon),\\ 0 & (\pi\varepsilon\leq x\leq\pi),\end{cases}$$

ce qui était précisément le développement recherché pour F, puisqu'il nous suffit de l'étendre à x négatif en invoquant la parité de notre fonction. Il suffit ensuite de tronquer cette série et d'exprimer F en terme de f pour conclure la preuve.

Voici qui termine notre détour au royaume des séries de Fourier et nous retournons dès le chapitre suivant à notre problème principal !

Preuve principale : Première étape

Preuve conditionnelle du résultat principal

En vertu du corollaire 4.2, nous disposons de deux suites $(r_n)_{n\geq 1}$ et $(s_n)_{n\geq 1}$ d'entiers strictement positifs tels que

$$\left|\alpha - \frac{r_n}{s_n}\right| \leq \frac{1}{s_n^2} \quad , \quad \text{pgcd}(r_n, s_n) = 1 \quad , \quad s_n \to \infty. \tag{6.1}$$

Nous définissons

$$P_n = s_n^{3/2}. \tag{6.2}$$

La preuve repose sur la somme

$$F = \sum_{P_n/4<p\leq P_n} f(\alpha p) \tag{6.3}$$

pour la fonction f définie en (5.1), somme qui porte sur $C(P_n)$ nombres premiers. Montrer qu'elle est strictement positive pour un certain n montrera qu'il existe un nombre premier p tel que $f(\alpha p) > 0$. En particulier, la partie fractionnaire de cet αp appartiendra à notre intervalle initial $[a, b]$.

Le développement (5.2) de f nous conduit à

$$F = \frac{\varepsilon C(P_n)}{2} + \mathcal{O}^*\left(\frac{4C(P_n)}{M\varepsilon}\right) + \sum_{0<|m|\leq M} \frac{A(m)}{m^2} \sum_{P_n/4<p\leq P_n} e(m\alpha p).$$

(Rappelons que $\varepsilon = b - a$). Nous choisissons M de telle sorte que le $\mathcal{O}^*$ ci-dessus soit inférieur à $\varepsilon C(P_n)/5$, c'est à dire que nous prenons pour M la partie entière de $20/\varepsilon^2$ à laquelle nous ajoutons 1. Comme $|A(m)/m^2| \leq 2/\varepsilon$, il vient

$$F \geq \frac{3\varepsilon C(P_n)}{10} - \sum_{1\leq m\leq M} \frac{4}{m^2\varepsilon}\left|\sum_{P_n/4<p\leq P_n} e(m\alpha p)\right|.$$

Nous allons établir dans les chapitres ultérieurs le fait suivant :

Fait 1. Soit $\varepsilon' > 0$ un nombre réel. Il existe un entier $n_0(\varepsilon')$ tel que, pour tout $n \geq n_0(\varepsilon')$ et tout entier $m \in [1, M]$, nous avons

$$\left| \sum_{P_n/4 < p \leq P_n} e(m\alpha p) \right| \leq \varepsilon' C(P_n).$$

Admettons-le ici et poursuivons notre démonstration. Nous prenons $\varepsilon' = \varepsilon^2/35$ dans cet énoncé et

$$n = n_0(\varepsilon').$$

Pour ce n (et tous ceux qui lui sont supérieurs d'ailleurs), nous avons

$$F/C(P_n) \geq \frac{3\varepsilon}{10} - \sum_{1 \leq m \leq M} \frac{4\varepsilon'}{m^2 \varepsilon} \geq \varepsilon/10 > 0$$

tout simplement en étendant la somme sur m à tous les entiers. Ce qui termine la preuve du théorème principal.

Préparation à la preuve du fait 1

Jusqu'à présent nous n'avons vu aucune explication à notre choix de P_n mais la raison va apparaître ici.

Nous remplaçons $m\alpha$ par $\alpha_{m,n} = mr_n/s_n$ que nous écrivons aussi $\alpha_{m,n} = r_{m,n}/s_{m,n}$ avec $\operatorname{pgcd}(r_{m,n}, s_{m,n}) = 1$. Le nouveau dénominateur est assez grand car m est borné indépendamment de n :

$$P_n^{2/3} \geq s_{m,n} \geq P_n^{2/3}/M. \tag{6.4}$$

Comme $\left|e(m\alpha p) - e(\alpha_{m,n} p)\right| \leq 2\pi\, p\, |m\alpha - \alpha_{m,n}|$, nous obtenons

$$\left| \sum_{P_n/4 < p \leq P_n} \big(e(m\alpha p) - e(\alpha_{m,n} p)\big) \right| \leq 2\pi m C(P_n) P_n / s_n^2$$

que nous majorons encore par $7mP_n^{-1/3}C(P_n)$, lequel est largement assez petit pour notre propos. Par exemple pour le rendre inférieur à $C(P_n)\varepsilon'/2$, il suffit d'imposer $P_n \geq (14M/\varepsilon')^3$. Soit $n_1(\varepsilon')$ le plus petit indice n à partir duquel cette inégalité est vérifiée pour tous les indices plus grands. Le fait 1 est alors une conséquence de

Fait 2. Soit $\varepsilon' > 0$ un nombre réel. Il existe un entier $n_2(\varepsilon')$ tel que, pour tout $n \geq n_2(\varepsilon')$ et tout entier $m \in [1, M]$, nous avons

$$\left| \sum_{P_n/4 < p \leq P_n} e(\alpha_{m,n} p) \right| \leq \varepsilon' C(P_n)/2.$$

En effet il nous suffit de poser $n_0(\varepsilon') = \max(n_1(\varepsilon'), n_2(\varepsilon'))$ pour en déduire le fait 1.

La fonction de Möbius

August Möbius introduisit en 1831 une fonction qui devait garder son nom et qui se définit par

$$\mu(d) = \begin{cases} 0 & \text{si } \exists p \ / \quad p^2 | d, \\ (-1)^k & \text{si } d = p_1 \cdots p_k, \text{ les } p_i \text{ premiers et distincts.} \end{cases}$$

Voici le début des valeurs de cette fonction :

d	1	2	3	4	5	6	7	8	9	10
$\mu(d)$	1	-1	-1	0	-1	1	-1	0	0	1

La propriété fondamentale que nous utiliserons est

Théorème 7.1

$$\sum_{d|\ell} \mu(d) = \begin{cases} 1 & \textit{si } \ell = 1, \\ 0 & \textit{si } \ell > 1, \end{cases}$$

où la somme en d porte sur tous les diviseurs ≥ 1 de ℓ.

PREUVE. Nous procédons par récurrence sur le nombre k de facteurs premiers de ℓ. Si $k = 0$, i.e. $\ell = 1$ la propriété est vérifée. Maintenant supposons qu'elle soit vérifiée pour k et montrons-la pour $k+1$. Soit donc

$$\ell = p_1^{a_1} \cdots p_k^{a_k} p_{k+1}^{a_{k+1}} = \ell' \cdot p_{k+1}^{a_{k+1}}$$

où les p_i sont des nombres premiers distincts et les a_i sont des entiers ≥ 1. Tout diviseur d de ℓ s'écrit de façon unique $d' p_{k+1}^a$ avec $d' | \ell'$ et $0 \leq a \leq a_{k+1}$. D'ailleurs d' est le pgcd de d et de ℓ', alors que p_{k+1}^a est celui de d et de $p_{k+1}^{a_{k+1}}$. À partir de ce moment, nous avons

$$\begin{aligned} \sum_{d|\ell} \mu(d) &= \sum_{d'|\ell'} \sum_{0 \leq a \leq a_{k+1}} \mu\left(d' p_{k+1}^a\right) \\ &= \sum_{d'|\ell'} \mu(d') \sum_{0 \leq a \leq a_{k+1}} \mu\left(p_{k+1}^a\right). \end{aligned}$$

Par conséquent, si $\ell' \neq 1$, la première somme est nulle, prouvant notre assertion. Si $\ell' = 1$, alors $a_{k+1} \geq 1$ et

$$\sum_{0 \leq a \leq a_{k+1}} \mu\left(p_{k+1}^a\right) = 1 - 1 + 0 + \cdots + 0 = 0$$

comme annoncé. ◇◇◇

Sommes sur nombres premiers

Il est essentiel de savoir comment manipuler des sommes du style $\sum_p g(p)$ pour des fonctions $g : \mathbb{N} \to \mathbb{C}$ bornées en module. D'une certaine façon, nous pouvons même dire que toute l'étude des nombres premiers se résume à celle de telles sommes. Nous établissons ici une identité qui permet d'une part de donner une majoration de la somme en question si g est positive ou nulle, et cela constitue la base du *crible* de Brun de 1916 dont nous donnerons un aperçu au chapitre et d'autre part, de traiter le cas de sommes oscillantes, comme lorsque $g(p) = \exp(ip)$ par exemple. L'idée revient ici à Vinogradov en 1937, la présentation très simplifiée ci-dessous étant elle due à l'auteur.

Théorème 8.1 *Nous nous donnons deux paramètres réels z et P tels que $4 \leq z^2 \leq P$. Soit $r(n)$ le nombre de facteurs premiers de l'entier n qui sont dans l'intervalle $]z, \sqrt{P}]$ et $Q = \prod_{p \leq z} p$. Nous définissons ensuite $\rho(n) = 1/(1 + r(n))$ si n est premier à Q et 0 sinon. Alors*

$$\sum_{P/4<p\leq P} g(p) = \sum_{\substack{P/4<\ell\leq P \\ \mathrm{pgcd}(\ell,Q)=1}} g(\ell) - \sum_{z<p\leq\sqrt{P}} \sum_{\frac{P}{4p}<d\leq\frac{P}{p}} \rho(d)g(dp) + R$$

avec $|R| \leq 3P/(2z)$ si $|g(n)| \leq 1$ pour tout n.

Les terminologies ont évoluées et il existe plusieurs méthodes combinatoires qui permettent d'atteindre de telles identités. Le point commun est d'avoir une première somme que l'on sait étudier (ce que lecteur découvrira aux prochains chapitres !) et une seconde somme qui porte sur deux variables. Vinogradov parlait initialement de sommes de type I et de sommes de type II ; la terminologie plus moderne tend à parler de partie linéaire, ce qui anticipe sur le traitement que nous lui donnerons, ou de partie criblée et appelle partie bilinéaire l'autre terme. Notons finalement que la variable d reste première à Q puisque $\rho(d)$ s'annule sinon.

Preuve. Nous détectons les nombres premiers parmi les entiers ℓ qui sont premiers à Q en enlevant à cette suite ceux qui admettent un facteur premier p dans $]z, \sqrt{P}]$, i.e. qui s'écrivent $\ell = dp$. Mais il faut aussi diviser par le nombre de telles écritures, soit $r(dp)$. Cela nous donne

$$\sum_{P/4<p\leq P} g(p) = \sum_{\substack{P/4<\ell\leq P \\ \mathrm{pgcd}(\ell,Q)=1}} g(\ell) - \sum_{z<p\leq\sqrt{P}} \sum_{\substack{P/(4p)<d\leq P/p \\ \mathrm{pgcd}(d,Q)=1}} \frac{g(dp)}{r(dp)}.$$

Comme $r(dp) = r(d) + 1$ dès que d n'est pas divisible par p, nous pouvons remplacer $r(dp)$ par $r(d) + 1$ pourvu que nous corrigions la formule pour les dp de la forme tp^2. Cela nous donne précisément la formule annoncée avec

$$R = \sum_{z<p\leq\sqrt{P}} \sum_{\frac{P}{4p^2}<t\leq\frac{P}{p^2}} \frac{\rho(tp^2)g(tp^2)}{r(tp^2)}. \tag{8.1}$$

Il nous suffit de majorer ce terme qui doit être regardé comme un terme d'erreur. Pour cela nous étendons la sommation sur p à tous les entiers, la simplifions en majorant $|g(tp^2)/r(tp^2)|$ par 1 et $|\rho(tp^2)|$ par $1/2$, puis concluons en comparant la somme résultante à une intégrale :

$$|R| \leq \frac{3P}{4} \sum_{z<p\leq\sqrt{P}} \frac{1}{p^2} \leq \frac{3P}{2z}. \tag{8.2}$$

◇◇◇

Preuve principale : Preuve du fait 2

Reprenons la preuve principale où nous l'avons laissée page 26. Nous appliquons l'identité du théorème 8.1 avec $P = P_n$ et $g(p) = e(\alpha_{m,n}p)$ où m est un entier $\leq M$. En ce qui concerne z, nous prenons

$$\operatorname{Log} z = L_n^{1/4}, \qquad (L_n = \operatorname{Log} P_n) \tag{9.1}$$

qui vérifie $z \geq 10$ si $P_n \geq B_0 = 10^{14}$. Le paramètre z croît plus vite que n'importe quelle puissance de $L_n = \operatorname{Log} P_n$, ce qui est numériquement visible. Il tend vers l'infini, ce qui cette fois-ci est numériquement difficilement décelable ! ! Nous écrivons

$$\sum_{P_n/4<p\leq P_n} e(\alpha_{m,n}p) = S_Q(\alpha_{m,n}) - B(\alpha_{m,n}) + R \tag{9.2}$$

avec

$$S_Q(\alpha_{m,n}) = \sum_{\substack{P_n/4<\ell\leq P_n \\ \operatorname{pgcd}(\ell,Q)=1}} e(\alpha_{m,n}\ell), \tag{9.3}$$

et

$$B(\alpha_{m,n}) = \sum_{z<p\leq\sqrt{P_n}} \sum_{\frac{P_n}{4p}<d\leq\frac{P_n}{p}} \rho(d)e(\alpha_{m,n}dp). \tag{9.4}$$

Quant à R, il est majoré par (8.2), qui résulte en

$$|R|/C(P_n) \leq \exp\bigl(-L_n^{1/4}/2\bigr) \tag{9.5}$$

si $P_n \geq B_1 = \exp(2\,000\,000)$ où B_1 est effectivement plus grand que B_0.

Preuve. En effet, $|R|/C(P_n)$ est majoré par

$$\exp\bigl(-\tfrac{1}{2}L_n^{1/4} + \operatorname{Log} L_n + \operatorname{Log}\frac{3}{2\times 0.08}\bigr)\exp\bigl(-\tfrac{1}{2}L_n^{1/4}\bigr)$$

et nous vérifions que l'argument de la première exponentielle est négatif si L_n est supérieur à $2\,000\,000$. Le lecteur et la lectrice, curieux et scrupuleuse, ou l'inverse, ou les deux, calculeront la valeur minimale valable au lieu de cette majoration un peu grossière. ◇◇◇

La preuve se scinde alors en deux. D'un part, nous établirons le fait suivant :

Fait 3. Pour $m \in [1, M]$, et $P_n \geq B_1$, nous avons

$$|S_Q(\alpha_{m,n})| \leq \exp\bigl(-L_n^{1/3}/3\bigr)C(P_n).$$

Ce qui règlera le traitement de S_Q. Puis nous continuerons avec la véritable clé de cette preuve :

Fait 4. Pour $m \in [1, M]$, et $P_n \geq B_2 = \exp(10^{17})$, nous avons

$$|B(\alpha_{m,n})| \leq \exp\bigl(-L_n^{1/4}/3\bigr)C(P_n).$$

Établir le fait 2 est alors simple : nous partons de la majoration

$$\Bigl|\sum_{P_n/2<p\leq P_n} e(\alpha_{m,n}p)\Bigr|/C(P_n) \leq \exp\bigl(-L_n^{1/3}/3\bigr) + \exp\bigl(-L_n^{1/4}/3\bigr) + \exp\bigl(-L_n^{1/4}/2\bigr)$$

si $P_n \geq B_2$ et $m \in [1, M]$. Chaque sommant du membre de droite est inférieur à $\varepsilon'/6$ si, respectivement

$$\begin{cases} P_n \geq \exp\bigl((3\,\mathrm{Log}(6/\varepsilon'))^3\bigr), \\ P_n \geq \exp\bigl((3\,\mathrm{Log}(6/\varepsilon'))^4\bigr), \\ P_n \geq \exp\bigl((2\,\mathrm{Log}(6/\varepsilon'))^4\bigr), \end{cases}$$

ce qui se réduit à $P_n \geq \exp\bigl(81\,\mathrm{Log}^4(6/\varepsilon')\bigr)$. Nous prenons donc $n_2(\varepsilon')$ comme étant l'indice n à partir duquel cette inégalité est satisfaite, de même que $P_n \geq B_2 = \exp(10^{17})$.

Étude de la partie criblée : Preuve du fait 3

Comme nous comparons la suite des nombres premiers à celle des entiers premiers à Q, il faut bien évidemment tout d'abord étudier cette dernière, c'est à dire S_Q définie en (9.3) par

$$S_Q(\alpha_{m,n}) = \sum_{\substack{P_n/4<\ell\leq P_n \\ \operatorname{pgcd}(\ell,Q)=1}} e(\alpha_{m,n}\ell).$$

Nous commençons par remarquer que

$$\sum_{\substack{k|\ell \\ k|Q}} \mu(k) = \begin{cases} 1 & \text{si } \operatorname{pgcd}(\ell,Q)=1 \\ 0 & \text{sinon.} \end{cases} \tag{10.1}$$

car la condition de sommation se résume à $k|\operatorname{pgcd}(Q,\ell)$. En introduisant une telle expression dans la définition de $S_Q(\alpha_{m,n})$, nous obtenons

$$S_Q(\alpha_{m,n}) = \sum_{k|Q} \mu(k) \sum_{\substack{P_n/4<\ell\leq P_n \\ k|\ell}} e(\alpha_{m,n}\ell). \tag{10.2}$$

Il se trouve que Q est beaucoup plus grand que P_n ; nous nous en sortons en tronquant la somme selon le nombre $\omega(k)$ de facteurs premiers de k, i.e. nous écrivons

$$|S_Q(\alpha_{m,n})| \leq \sum_{\substack{k|Q \\ \omega(k)\leq r}} \left| \sum_{\substack{P_n/4<\ell\leq P_n \\ k|\ell}} e(\alpha_{m,n}\ell) \right| + R' \tag{10.3}$$

avec

$$R' = \sum_{\substack{k|Q \\ \omega(k)>r}} \left| \sum_{\substack{P_n/4<\ell\leq P_n \\ k|\ell}} e(\alpha_{m,n}\ell) \right| \tag{10.4}$$

où

$$r = (\operatorname{Log} z)^2 = L_n^{1/2}. \tag{10.5}$$

Remarquons que $z^r < P_n^{1/4} < s_{m,n}$ si $P_n \geq B_1$.

PREUVE. En effet

$$z^r = \exp\bigl(rL_n^{1/4}\bigr) = \exp\bigl(L_n^{3/4}\bigr)$$

et l'argument de l'exponentielle est $\leq L_n/4$ dès que $L_n \geq 4^4$ ce qui est bien plus petit que $\operatorname{Log} B_1 = 2\,000\,000$. ◇◇◇

Dans le premier terme, $k = q_1 \cdots q_r$ où chaque q_i est ou bien 1, ou bien un nombre premier inférieur à z. Ce k est alors majoré par z^r, c'est à dire trop petit pour que $k\alpha_{m,n}$ puisse être congru à 0 modulo 1. La somme sur ℓ est alors une progression géométrique et à l'aide de (1.1), nous constatons qu'elle vaut au plus $s_{m,n}/2$. Ensuite le nombre de k possibles est majoré par z^r puisque qu'il y a au plus z nombres q_i pour chaque i. Cela nous donne

$$|S_Q(\alpha_{m,n})| \leq s_{m,n} z^r + R' \leq P_n^{11/12} + R'. \tag{10.6}$$

La méthode de Rankin multiplicative

La méthode que nous exposons sur cet exemple est due à Robert Rankin en 1947 dans un article où il étudie les grandes différences entre nombres premiers consécutifs. Elle est étonamment simple et flexible, à peine une remarque, alors qu'elle a permis de rendre beaucoup de résultats accessibles.

Afin d'étudier R' nous pouvons restreindre la sommation à $k \leq P_n$, car sinon aucun entier ℓ ne répond à la condition. Il vient

$$R' \leq \sum_{\substack{k|Q,\\ \omega(k)>r, k\leq P_n}} P_n/k. \tag{10.7}$$

Nous oublions maintenant la condition $k \leq P_n$. Pour traiter la condition $\omega(k) > r$, nous introduisons un paramètre réel $x \geq 1$ et écrivons

$$\sum_{\substack{k|Q\\ \omega(k)>r}} \frac{1}{k} \leq \sum_{\substack{k|Q\\ \omega(k)>r}} \frac{x^{\omega(k)-r}}{k} \leq \sum_{k|Q} \frac{x^{\omega(k)-r}}{k}$$

en oubliant cette fois-ci la condition $\omega(k) > r$. Nous recourrons ici à une technique qui nous sera utile plusieurs fois et qui est la *mise sous forme de produit eulérien*, du nom de Leonhard Euler, l'un des mathématiciens les plus prolixes du XVIII$^{\text{e}}$ siècle. Il s'agit de remarquer que

$$\sum_{k|Q} \frac{x^{\omega(k)}}{k} = \prod_{p\leq z} (1 + x/p) \tag{10.8}$$

tout simplement parce qu'en développant le membre de droite, nous obtenons une et une seule fois tous les termes du membre de gauche. Nous poursuivons

alors la preuve principale en employant l'inégalité $\operatorname{Log}(1+x/p) \le x/p$:

$$R'/P_n \le x^{-r} \exp\Big(\sum_{p\le z} x/p\Big) \le x^{-r} \exp(2x \operatorname{Log} z).$$

Cette dernière majoration vient tout simplement de ce que

$$\sum_{p\le z} 1/p \le \sum_{n\le z} 1/n \le 2 \operatorname{Log} z$$

comme nous l'avons remarqué dans l'introduction en (1.2). Nous pourrions bien sûr améliorer nos estimations en invoquant le lemme 3.8 mais nous essayons de nous appuyer sur le moins de matériel possible. La valeur optimale de x est alors $x = r/(2 \operatorname{Log} z) \ge 1$. En partant de (10.6), cela nous permet d'établir la majoration

$$|S_Q(\alpha_{m,n})|/C(P_n) \le \exp\big(-L_n^{1/4}/3\big) \tag{10.9}$$

si $P_n \ge B_1 = \exp(2\,000\,000)$.

Preuve. En effet, nous avons atteint la majoration

$$\Big(P_n^{11/12} + P_n \exp\Big(-L_n^{1/2} \operatorname{Log} \frac{L_n^{1/4}}{2e}\Big)\Big)/\Big(0.08 P_n/L_n\Big)$$

pour $|S_Q(\alpha_{m,n})|/C(P_n)$. Nous montrons successivement que

$$\begin{cases} \frac{1}{3} L_n^{1/4} - \frac{1}{12} L_n + \operatorname{Log} L_n \le \operatorname{Log}\big(\frac{1}{2} \times 0.08\big), \\ \frac{1}{3} L_n^{1/4} - L_n^{1/2} \operatorname{Log} \dfrac{L_n^{1/4}}{2e} + \operatorname{Log} L_n \le \operatorname{Log}\big(\frac{1}{2} \times 0.08\big) \end{cases}$$

le premier pour $L_n \ge 108$ et le second pour $L_n \ge 2\,550$, ce qui établit bien l'inégalité demandée moyennant quelques manipulations que nous laissons à la lectrice attentive. ◇◇◇

Étude de la partie bilinéaire : Preuve du fait 4

Voici la dernière brique construisant la preuve de la densité des $\sin p$: l'étude de la partie bilinéaire $B(\alpha_{m,n})$ définie en (9.4) par

$$B(\alpha_{m,n}) = \sum_{z<p\leq\sqrt{P_n}} \sum_{\frac{P_n}{4p}<d\leq\frac{P_n}{p}} \rho(d)e(\alpha_{m,n}dp).$$

Notons que, de ρ définie au théorème 8.1, nous n'utiliserons que le fait qu'elle est bornée en module par 1. La preuve comprend deux étapes.

Localisation de p

Nous restreignons tout d'abord le domaine de variations de p. Posons

$$B(X, X') = \sum_{X<p\leq X'} \sum_{\frac{P_n}{4p}<d\leq\frac{P_n}{p}} \rho(d)e(\alpha_{m,n}dp).$$

Nous décomposons B en la somme de tels termes avec $X' \leq 4X$ en considérant les intervalles $]z, 4z]$, puis $]4z, 4^2z]$ etc, jusqu'à ce que la borne supérieure dépasse $\sqrt{P_n}$, i.e. $4^k z \geq \sqrt{P_n}$. Soit au plus $1+\mathrm{Log}(\sqrt{P_n}/z)/\log 4 \leq L_n$ telles sommes. Tout ceci nous donne

$$|B(\alpha_{m,n})| \leq L_n \ \max |B(X, X')| \tag{11.1}$$

où le maximum est pris sur tous les couples de réels (X, X') qui vérifient

$$z \leq X < X' \leq \min(4X, \sqrt{P_n}). \tag{11.2}$$

Utiliser l'inégalité de Cauchy-Schwarz

Nous appliquons l'inégalité de Cauchy-Schwarz :

$$|B(X, X')|^2 \leq 4X \sum_{X<p\leq X'} \Big| \sum_{\frac{P_n}{4p}<d\leq\frac{P_n}{p}} \rho(d)e(\alpha_{m,n}dp)\Big|^2$$

car $\sum_{X<p\leq X'} 1 \leq X'-X+1 \leq 4X$ en étendant la somme à tous les entiers. Le point fondamental consiste à étendre la somme sur p dans la seconde sommation en une somme sur tous les entiers que l'on note t, ce qui a pour effet de remplacer une variable qui évolue dans une suite plutôt inconnue (celle des nombres premiers) par une variable simple à manipuler. Ensuite nous développons le module au carré en utilisant $|z|^2 = z\overline{z}$. Il résulte de tout ceci que $|B(X,X')|^2$ est majoré par

$$4X \sum_{\frac{P_n}{4X'}<d,\, d'\leq\frac{P_n}{X}} \sum \rho(d)\rho(d') \sum_{t\in I(d,d')} e(\alpha_{m,n}t(d-d'))$$

où $I(d,d')$ est un intervalle qui contient moins de P_n/d points entiers. Nous simplifions notre majorant et séparons en classes modulo $s_{m,n}$. Il en ressort que $|B(X,X')|^2$ est inférieur à

$$4X \sum_{d\leq P_n/X} \sum_{0\leq a\leq s_{m,n}-1} \sum_{\substack{d'\leq P_n/X \\ d'\equiv d+a[s_{m,n}]}} \Big| \sum_{t\in I(d,d')} e(\alpha_{m,n}ta)\Big|.$$

Si $a=0$, nous majorons la somme interne par P_n/d et sinon, grâce à (1.1),

$$\Big| \sum_{t\in I(d,d')} e(\alpha_{m,n}ta)\Big| \leq \frac{1}{2\|ar_{m,n}/s_{m,n}\|}.$$

Ces deux majorations sont indépendantes de d' pour la contribution duquel nous utilisons

$$\sum_{\substack{d'\leq P_n/X \\ d'\equiv d+a[s_{m,n}]}} 1 \leq 1+P_n/(Xs_{m,n}) \leq 2P_n/(Xs_{m,n})$$

Comme $a\mapsto b=ar_{m,n}$ est une bijection sur $(\mathbb{Z}/s_{m,n}\mathbb{Z})\setminus\{0\}$ du fait que $r_{m,n}$ est premier à $s_{m,n}$, cela nous donne comme contribution

$$4X \sum_{d\leq P_n/X} \frac{2P_n}{Xs_{m,n}}\Big(\frac{P_n}{d} + \sum_{1\leq b\leq s_{m,n}/2} \frac{s_{m,n}}{b}\Big). \tag{11.3}$$

L'inégalité (1.2) donnée dans l'introduction garantit simultanément $\sum_{d\leq P_n/X} 1/d \leq L_n$ et $\sum_{1\leq b\leq s_{m,n}/2} 1/b \leq L_n$, et en conséquence (11.3) est inférieur à

$$\frac{8P_n^2}{s_{m,n}}L_n + \frac{8P_n^2}{X}L_n.$$

En remarquant que $s_{m,n}\geq z$, nous simplifions cela :

$$|B(X,X')|^2 \leq 16P_n^2L_n/z \tag{11.4}$$

que nous remettons dans (11.1) :

$$|B(\alpha_{m,n})| \leq 4P_nL_n^{3/2}\exp\big(-L_n^{1/4}/2\big). \tag{11.5}$$

Conclusion

Tout cela résulte en

$$|B(\alpha_{m,n})| \leq C(P_n) \exp\bigl(-L_n^{1/4}/3\bigr). \tag{11.6}$$

pour $P_n \geq B_2 = \exp(10^{17})$.

PREUVE. En effet, il s'agit de montrer que

$$4L_n^{5/2} \exp\bigl(-\tfrac{1}{6}(L_n)^{1/4}\bigr)/0.08$$

est inférieur à 1 si $P_n \geq B_2$. Nous écrivons cette quantité

$$\exp\bigl(-\tfrac{1}{6}L_n^{1/4} + \tfrac{5}{2}\operatorname{Log} L_n + \operatorname{Log}(4/0.08)\bigr)$$

Il est à présent facile de montrer que l'argument de l'exponentielle est effectivement négatif si $L_n \geq 10^{17}$. ◇◇◇

Bien sûr, cette valeur de B_2 est astronomique, mais c'est malheureusement celle que donne la démonstration et l'une des faiblesse de la théorie actuelle.

Ici s'achève la preuve du résultat principal : la suite des $(\sin p)$ est dense dans $[-1, 1]$ et nous savons le démontrer ! Nous poursuivons l'exposition avec deux chapitres qui éclairent la méthode. Tout d'abord, si nous avons parlé de *crible*, nous n'en avons pas encore utilisé, ni montré comment le théorème 8.1 y est relié. Le prochain chapitre répare cette lacune. Le dernier chapitre prend un peu de recul par rapport au théorème 8.1 et présente plusieurs décompositions partie linéaire/partie bilinéaire pour des sommes sur des nombres premiers. Chacune de ces décompositions a ses propres avantages et inconvénients, mais nous laisserons les lecteurs à cet endroit, libres à eux de continuer l'exploration !

Variation no 2 : Du crible de Brun pur

Ici, et dans la chapitre qui suit, nous analysons plus avant la nature du théorème 8.1. Nous avons parlé de partie *criblée* sans expliquer ce que c'était, ni en quoi cette théorie était reliée à notre propos et nous réparons cette lacune dans ce chapitre.

Une approche consiste à appliquer l'identité du théorème 8.1 à des sommes non-oscillantes. Le problème dans cette utilisation vient de la partie bilinéaire, car nous ne pouvons espérer montrer que cette partie est négligeable si la fonction g est positive ou nulle. Mais, et pourvu de nous contenter d'une borne supérieure, nous pouvons abandonner carrément ce terme du fait qu'il est négatif ! Il vient dans ce cas

$$\sum_{P/4<p\leq P} g(p) \leq \sum_{\substack{P/4<\ell\leq P\\ \operatorname{pgcd}(\ell,Q)=1}} g(\ell) + 3P/(2z). \tag{12.1}$$

Une preuve directe : remarquez simplement que la suite des entiers ℓ de $]P/4, P]$ qui sont premiers à Q contient celle des nombres premiers de cet intervalle ! L'inégalité est donc valable sans terme d'erreur et nous aurions égalité avec $Q = \sqrt{P}$. L'idée de considérer cette égalité remonte à Adrien-Marie Legendre vers 1830 et constitue ce que l'on appelle « le crible de Legendre ». Nous disposons dans notre présentation d'une part du paramètre z (ou Q), d'autre part de la méthode de Rankin. Ce qui va nous permettre de tirer beaucoup d'informations de cette simple inégalité.

Commençons par un lemme facile dont l'idée est due à Brun :

Lemme 12.1 *Si r est un entier pair ≥ 1, alors*

$$\sum_{\substack{d|\ell\\ \omega(d)\leq r}} \mu(d) \geq \begin{cases} 1 & \text{si } \ell = 1,\\ 0 & \text{en général.}\end{cases}$$

Preuve. Le cas $\ell = 1$ est facile à régler, de même que le cas où ℓ n'admet qu'un seul facteur premier. Ensuite nous écrivons $\ell = p_1^{a_1} \dots p_K^{a_K}$ avec les a_i des entiers ≥ 1, les p_i des nombres premiers distincts, et $K = \omega(\ell)$. Nous

vérifions que

$$\sum_{\substack{d|\ell\\ \omega(d)\leq r}} \mu(d) = \sum_{0\leq k\leq r} (-1)^k \binom{K}{k}$$

car il y a exactement $\binom{K}{k}$ diviseurs de ℓ sans facteurs carrés qui ont exactement k facteurs premiers. Cette somme se calcule par récurrence sur r et vaut $(-1)^r\binom{K-1}{r}$ (soit 0 si $r=K$). La conclusion est immédiate. ◇◇◇

À l'aide de ce lemme, nous écrivons alors directement, pour tout entier r pair :

$$\sum_{P/4<p\leq P} g(p) \leq \sum_{P/4<\ell\leq P} \sum_{\substack{d|Q\\ d|\ell\\ \omega(d)\leq r}} \mu(d)g(\ell). \tag{12.2}$$

Si le lecteur compare au traitement de S_Q donné au chapitre précédent, il constatera que cette expression nous dispense complètement du traitement du terme d'erreur causé par la troncation à $\omega(\ell)\leq r$. Que la lectrice remarque aussi qu'en prenant r impair, nous obtiendrions une borne inférieure. La méthode de Rankin va encore être utile ici car il nous faut à présent calculer la somme résultante. Nous explicitons les étapes sur l'exemple utilisé par Brun en 1916 et qui a eu l'effet d'un coup de tonnerre dans le monde des arithméticiens.

Peu de premiers jumeaux par intervalle

Considérons le nombre $J(P)$ de nombres premiers p de $]P/4,P]$ qui sont tels que $p+2$ est aussi premier, comme 17 ou 41. Depuis le milieu du dix neuvième siècle, on conjecture qu'il y a une infinité de tels nombres, dits *jumeaux*, conjecture qui reste très largement hors de notre portée. Le fait est que nous savons très mal travailler avec cette condition supplémentaire sur p, et même montrer que de tels nombres ne sont pas en proportion positive dans la suite des nombres premiers est longtemps resté indémontré ! Le théorème de Brun ci-dessous répare cette lacune de façon éclatante.

Nous prenons pour $g(n)$ la fonction qui vaut 1 si $n+2$ n'admet aucun facteur premier $\leq z$ et 0 sinon, de sorte que

$$J(P) \leq \sum_{P/4<p\leq P} g(p).$$

Pour appréhender le membre de droite, le mieux est de reprendre le raisonnement menant à (12.2) et d'écrire

$$\sum_{P/4<p\leq P} g(p) \leq \sum_{\substack{P/4<\ell\leq P\\ \operatorname{pgcd}(\ell(\ell+2),Q)=1}} 1 \leq \sum_{P/4<\ell\leq P} \sum_{\substack{d|Q\\ d|\ell(\ell+2)\\ \omega(d)\leq r}} \mu(d).$$

Nous échangeons alors les sommations en ℓ et en d

$$\sum_{P/4<p\le P} g(p) \le \sum_{\substack{d|Q\\ \omega(d)\le r}} \mu(d) \sum_{\substack{P/4<\ell\le P\\ d|\ell(\ell+2)}} 1.$$

Il nous faut alors évaluer la dernière somme que nous décomposons d'abord en classes modulo d (qui est sans facteurs carrés en tant que diviseur de Q) :

$$\sum_{\substack{P/4<\ell\le P\\ d|\ell(\ell+2)}} 1 = \sum_{\substack{c \bmod d\\ c(c+2)\equiv 0[d]}} \sum_{\substack{P/4<\ell\le P\\ \ell\equiv c[d]}} 1.$$

La somme interne vaut $3P/(4d) + \mathcal{O}^*(2)$ et la somme externe est calculée dans le lemme suivant.

Lemme 12.2 *Soit d un entier sans facteurs carrés. Le nombre de classes c modulo d qui vérifient $c(c+2) \equiv 0[d]$ vaut $\kappa_d 2^{\omega(d)}$ où $\kappa_d = 1/2$ si d est pair et $\kappa(d) = 1$ sinon.*

PREUVE. En effet le lemme chinois nous garantit que r_1 solutions de la congruence $c(c+2) \equiv 0$ modulo m_1 et r_2 solutions de cette même congruence, mais modulo m_2, donnent lieu à $r_1 r_2$ solutions modulo $m_1 m_2$ si m_1 et m_2 sont premiers entre eux. Il suffit alors de compter le nombre de solutions modulo un nombre premier p car d est supposé sans facteurs carrés. La conclusion est facile. ◇◇◇

Tout ceci nous donne

$$\sum_{\substack{P/4<\ell\le P\\ d|\ell(\ell+2)}} 1 = 2^{\omega(d)}\kappa_d\Big(\frac{3P}{4d} + \mathcal{O}^*(2)\Big). \tag{12.3}$$

D'où il résulte

$$J(P) \le \frac{3P}{4} \sum_{\substack{d|Q\\ \omega(d)\le r}} \frac{\mu(d)2^{\omega(d)}\kappa_d}{d} + 2 \sum_{\substack{d|Q\\ \omega(d)\le r}} 2^{\omega(d)}.$$

Le dernier terme est simplement $\le 2\big(1 + 2\sum_{p\le z} 1\big)^r$, qui se majore encore par $2(2z)^r$. L'évaluation de la première somme sur d passe par la méthode de Rankin multiplicative. Nous écrivons, pour un paramètre $x \ge 1$ à choisir :

$$\begin{aligned}\sum_{\substack{d|Q\\ \omega(d)\le r}} \frac{\mu(d)2^{\omega(d)}\kappa_d}{d} &= \Big(\sum_{d|Q} - \sum_{\substack{d|Q\\ \omega(d)>r}}\Big)\frac{\mu(d)2^{\omega(d)}\kappa_d}{d}\\ &= \tfrac{1}{2}\prod_{3\le p\le z}\big(1-2/p\big) + \mathcal{O}^*\Big(x^{-r}\sum_{d|Q}(2x)^{\omega(d)}/d\Big),\end{aligned}$$

en écrivant le premier terme en produit eulérien. Nous modelons le traitement du second terme sur celui présenté page 36, mais il nous faut ici recourir à l'estimation plus précise de la somme des $1/p$ donnée au lemme 3.8. Nous prenons $x = r/(\operatorname{Log}\operatorname{Log} z + 13/6)$, puis pour r l'entier pair immédiatement supérieur à $10\operatorname{Log}\operatorname{Log} z$. Lorsque $\operatorname{Log} z \geq 7$, le $\mathcal{O}^*$ ci-dessus est majoré par

$$\exp\biggl(-10(\operatorname{Log}\operatorname{Log} z)\operatorname{Log}\frac{10(\operatorname{Log}\operatorname{Log} z)/e}{\operatorname{Log}\operatorname{Log} z + 13/6}\biggr)$$

qui est inférieur à $(\operatorname{Log} z)^{-8}$. En utilisant cette fois-ci la minoration du lemme 3.8 à laquelle il ôtera la contribution du terme $p = 2$, le lecteur montrera en passant aux logarithmes que

$$\tfrac{1}{2}\prod_{3\leq p\leq z}(1-2/p) \leq \frac{\exp(4/3)/2}{\operatorname{Log}^2 z}$$

et par conséquent

$$J(P) \leq \frac{3P}{2(\operatorname{Log} z)^2} + 2(2z)^{11\operatorname{Log}\operatorname{Log} z}.$$

Nous prenons

$$\operatorname{Log} z = \frac{\operatorname{Log} P}{24\operatorname{Log}\operatorname{Log} P} \tag{12.4}$$

et $P \geq \exp(1\,200)$ pour garantir $\operatorname{Log} z \geq 7$. Finalement

Théorème 12.3 *Pour $P \geq \exp(1\,200)$, nous avons*

$$J(P) = \sum_{\substack{P/4<p\leq P\\ p+2 \text{ premier}}} 1 \leq 870\frac{P(\operatorname{Log}\operatorname{Log} P)^2}{\operatorname{Log}^2 P}.$$

Preuve. Pour conclure la preuve, il nous faut majorer

$$\tfrac{3}{2} + \exp\Bigl(11(\operatorname{Log}\operatorname{Log} z)\operatorname{Log}(2z) + \operatorname{Log}(2/P) + 2\operatorname{Log}\operatorname{Log} z\Bigr)$$

et il nous suffit de remarquer que l'argument de l'exponentielle est (vraiment très !) négatif si $\operatorname{Log} P \geq 1\,200$. ◇◇◇

La borne $P \geq \exp(1\,200)$ est évidemment gigantesque, la constante 870 aussi, mais ce théorème montre que le nombre de nombres premiers jumeaux dans $]P/4, P]$ est asymptotiquement nettement inférieur à celui des nombres premiers. Fait que Brun a exprimé en 1916 en énonçant :

Théorème 12.4 (Viggo Brun) *La somme des inverses des nombres premiers jumeaux est finie ou convergente.*

Alors que la somme des inverses des nombres premiers est, elle, divergente.

PREUVE. Soit $P_0 = 4^{900} \geq \exp(1\,200)$. Le théorème précédent nous garantit que

$$\sum_{\substack{4^k < p \leq 4^{k+1} \\ p+2 \text{ premier}}} 1/p \leq 870 \frac{\operatorname{Log}^2(k \operatorname{Log} 4)}{k^2 \operatorname{Log}^2 4}$$

et par conséquent

$$\sum_{\substack{p \geq 2 \\ p+2 \text{ premier}}} 1/p \leq \sum_{p \leq P_0} 1/p + \sum_{k \geq 900} 870 \frac{\operatorname{Log}^2(k \operatorname{Log} 4)}{k^2 \operatorname{Log}^2 4}.$$

Comme la dernière série est convergente, le théorème suit. ◇◇◇

Sommes sur nombres premiers et identités

Nous prenons ici un peu de recul par rapport à la preuve principale pour analyser plus avant la technique qui émerge. Le théorème 8.1 joue un rôle tout particulier dans notre démonstration, en ce que c'est le seul véritable endroit où nous avons requis que p soit un nombre premier. La preuve utilise aussi une borne inférieure pour $C(P)$ mais une borne faible suffit. La construction d'identités constitue dès lors un point crucial de la théorie.

La caractéristique de ces identités est d'écrire une somme sur des nombres premiers $\sum_{P/4<p\leq P} g(p)$ en une combinaison linéaire tout d'abord de termes que l'on sait calculer, ici c'était la partie criblée, et de un ou plusieurs autres termes de la forme

$$\sum_{\ell m\leq P}\sum a_\ell b_m g(\ell m)$$

où les a_ℓ et les b_m sont mal connus mais bornés[1]. Dans notre identité, a_ℓ vaut 1 si ℓ est un nombre premier dans l'intervalle $]\sqrt{z},\sqrt{P}]$ et 0 ailleurs, alors que les b_m sont encore plus mystérieux. Mais la structure de forme bilinéaire nous permet de recourir à l'inégalité de Cauchy-Schwarz et de n'utiliser de ces suites que des bornes supérieures. Dont cette fois-ci nous disposons. Une dernière remarque : si ℓ ou m est autorisé à être grand, disons entre $P/10$ et P, alors la sommation sur l'autre variable va être ridiculement courte, au mieux entre 1 et 10. Il n'est plus alors correct de penser notre somme comme étant bilinéaire, entrainant par là qu'aucun bénéfice ne saurait résulter d'une utilisation de l'inégalité de Cauchy-Schwarz ! Il est donc capital de borner ces variables supérieurement, et nous avions par exemple ℓ entre $\sqrt{z}$ et $\sqrt{P}$ et m entre $\sqrt{P}/4$ et P/z.

Nous supposons dans la suite, quitte à remplacer g par son produit avec la fonction caractéristique de l'intervalle, que $g(n)=0$ hors de $]P/4,P]$. Ceci nous dispense des conditions de taille sur la variable n.

[1] Il faut souvent affaiblir cette condition et accepter des fonctions de diviseurs. De telles fonctions sont bornées *en moyenne*, ce qui nous suffit.

Le problème du terme principal

Il s'agit donc de remplacer la fonction caractéristique de la suite des nombres premiers, restreinte à un intervalle, par une autre, issue de la partie linéaire ou criblée, à laquelle nous ajoutons une partie bilinéaire. L'idéal serait bien sûr que cette partie bilinéaire ait une contribution négligeable, car cela nous permettrait alors de *calculer* précisément la somme sur les nombres premiers.

Il suffit pour mesurer cela de comparer les sommes en prenant $g = 1$, i.e. dans notre cas comparer $C(P)$ avec

$$\sum_{\substack{P/4<n\leq P\\ \mathrm{pgcd}(n,Q)=1}} 1. \qquad (Q=\prod_{p\leq z} p)$$

La lectrice dispose du matériel nécessaire pour montrer que cette somme se comporte comme $3\prod_{p\leq z}(1-1/p)P/4$ si $\mathrm{Log}\, z$ est assez petit, disons pour ne pas avoir de problèmes $\mathrm{Log}\, z \leq \sqrt{\mathrm{Log}\, P}$. D'où l'on déduit qu'elle est comprise entre deux constantes strictement positives fois $P/\mathrm{Log}\, z$. Nous concluons de cette petite étude que notre partie bilinéaire *contient encore une partie du terme principal* !

Une identité de Linnik

Yu Linnik introduisit en 1960 une autre identité qui cette fois-ci repose sur la fonction de von Mangoldt définie en (3.1) et que voici. Pour K entier pair, $P \geq 4$ et $g \geq 0$, nous avons :

$$\sum_n \frac{\Lambda(n)}{\mathrm{Log}\, n} g(n) \geq \sum_{1\leq k\leq K} \frac{(-1)^{k+1}}{k} \sum_n \tau_k^*(n) g(n)$$

où $\tau_k^*(n)$ est le nombre de k-uplets $(d_1, d_2, \ldots, d_k)$ d'entiers *strictement* supérieurs à 1 tels que $d_1 d_2 \ldots d_k = n$. Pour K impair, l'inégalité est simplement inversée. Ce sont ces identités qui ont fait dire « si l'on arrive à comprendre les fonctions de diviseurs, on comprendra les nombres premiers ». Notons que les d_i ne sont ni minorés, ni majorés, ce qui peut entrainer des complications ; nous verrons plus loin comment éviter cet écueil.

Preuve. Nous commençons par l'inégalité :

$$\Lambda(n) \geq T(n) = \Lambda(n) - \sum_{\ell a=n} \Lambda(\ell)\tau_K^*(a).$$

Le lemme 13.1 ci-dessous permet de transformer $T(n)$:

$$\begin{aligned} T(n) &= \sum_{bc=n} \mu(b)\,\mathrm{Log}\, c - \sum_{bca=n} \mu(b)\tau_K^*(a)\,\mathrm{Log}\, c \\ &= \sum_{cd=n} \Big(\mu(d) - \sum_{ba=d} \mu(b)\tau_K^*(a)\Big)\,\mathrm{Log}\, c. \end{aligned}$$

Pour la parenthèse interne, nous écrivons $a = qa'$. Il vient

$$\sum_{ba=d} \mu(b)\tau_K^*(a) = \sum_{\substack{bqa'=d,\\ q>1}} \mu(b)\tau_{K-1}^*(a').$$

Nous *collons* alors b et q en une variable b' et établissons

$$\sum_{\substack{bq=b',\\ q>1}} \mu(b) = \begin{cases} 0 & \text{si } b' = 1, \\ -\mu(b') & \text{sinon} \end{cases}$$

grâce au théorème 7.1. Du coup

$$\sum_{ba=d} \mu(b)\tau_K^*(a) = \tau_{K-1}^*(d) - \sum_{b'a'=d} \mu(b')\tau_{K-1}^*(a').$$

La somme du membre de droite est du même style que celle que nous venons d'étudier mais avec $K-1$ au lieu de K. Nous sautons les étapes de la récurrence qui nous mène, pour K pair, à

$$\mu(d) - \sum_{ba=d} \mu(b)\tau_K^*(a) = \sum_{k=0}^{K-1} (-1)^k \tau_k^*(d)$$

où $\tau_0^*(d)$ vaut 1 si $d = 1$ et 0 sinon.

Ensuite nous établissons que

$$\sum_{cd=n} \tau_k^*(d) \operatorname{Log} c = \tau_{k+1}^*(n) \operatorname{Log}(n)/(k+1)$$

en écrivant $\operatorname{Log} n = \operatorname{Log}(d_1 \cdots d_{k+1})$ dans le membre de droite, ce qui permet de clore la preuve. ◇◇◇

Lemme 13.1 *Pour $n \geq 1$, nous avons*

$$\sum_{b\ell=n} \mu(b) \operatorname{Log} \ell = \Lambda(n).$$

PREUVE. Nous exprimons simplement Log en termes de Λ à l'aide de (3.2) :

$$\sum_{b\ell=n} \mu(b) \operatorname{Log} \ell = \sum_{b\ell=n} \mu(b) \sum_{md=\ell} \Lambda(m) = \sum_{bmd=n} \sum \mu(b)\Lambda(m)$$

où la sommation a lieu sur b, m et d. Nous *collons* b et m en posant $t = bm$. Maintenant $\sum_{bm=t} \mu(b)$ vaut 1 ou 0 selon que t vaut 1 ou non, comme établi au théorème 7.1, ce qui termine la preuve. ◇◇◇

Une autre identité

Les autres identités sont pour la plupart basées sur la relation (3.2) qui relie la fonction Λ à la fonction Log. Par exemple, en 1996, Hedi Daboussi obtenait l'identité

$$\sum_{\text{pgcd}(n,Q)=1} \Lambda(n)g(n) = \sum_{\text{pgcd}(n,Q)=1} g(n)\operatorname{Log} n - \sum_{\substack{z<\ell,\, m\le P/z \\ \text{pgcd}(\ell m,Q)=1}}\sum g(\ell m)\Lambda(m),$$

où $Q=\prod_{p\le z} p$. Bien sûr et moyennant quelques hypothèses faibles sur g, le lemme 3.5 dit que la somme considérée diffère peu de $\sum_p g(p)\operatorname{Log} p$. Qui plus est, le lecteur, au fait de la technique de sommation par parties, recouvrira facilement $\sum_p g(p)$.

La somme $\sum_m g(\ell m)\Lambda(m)$ apparaît au membre de droite, somme qui, à ℓ fixé, est du même type que la somme initiale : nous pouvons alors réutiliser l'identité ! Et jouer avec le second paramètre z qui n'a pas vocation à rester fixe ... Ce livre touche à sa fin et nous laissons la lectrice explorer elle-même ces pistes.

En prenant $g=1$, la somme de départ est de taille P alors que la somme "linéaire" (celle où apparaît le $\operatorname{Log} n$) est elle de l'ordre de $P(\operatorname{Log} P)/\operatorname{Log} z$, ce qui fait que la partie bilinéaire contient encore une partie du terme principal. Notons que le paramètre ℓ est borné inférieurement par z. Nous pourrions ajouter cette même condition dans l'identité de Linnik, garantissant par là la minoration $d_i > z$.

Récupérer le terme principal dans la partie linéaire ...

Notre incapacité à traiter la condition $\text{pgcd}(n,Q)=1$ complètement dès que z est une puissance de X nous force à prendre z plus petit dans les identités que nous avons vues jusqu'à présent. Dans la partie bilinéaire, une des variables peut alors être relativement petite. En 1976, Robert Vaughan introduisit une identité qui évite ce problème. La somme $\sum_n \Lambda(n)g(n)$ égale

$$\sum_n (u_n - v_n)g(n) + \sum_{\substack{yz<\ell\le P/z \\ yz<m\le P/(yz)}}\sum w_\ell g(\ell m)$$

avec $P > 4z$ et

$$u_n = \sum_{\substack{b|n \\ b\le y}} \mu(b)\operatorname{Log}(n/b)\ ,\ v_n = \sum_{\substack{bc|n \\ b\le y, c\le z}}\sum \mu(b)\Lambda(c)$$

$$w_\ell = \sum_{\substack{bc=\ell \\ b>y, c>z}} \mu(b)\Lambda(c).$$

Il est assez courant de prendre $y = z = P^{1/5}$. Les deux premières sommes, avec u_n et v_n, sont souvent faciles à calculer du fait que la partie "difficile" (le $\mu(b)$ dans u_n ou le $\mu(b)\Lambda(c)$ dans v_n) ne porte que sur des variables convenablement bornées.

PREUVE. La somme sur ℓ ci-dessus s'écrit aussi

$$S_1 = \sum_{b>y,\, c>z,\, m}\sum\sum \mu(b)\Lambda(c)g(bcm),$$

écriture que nous modifions maintenant en :

$$\Big(\sum_{b>y,\, c,\, m}\sum\sum - \sum_{b>y,\, c\le z,\, m}\sum\sum\Big)\mu(b)\Lambda(c)g(bcm) = S_2 - S_3.$$

Dans la première somme, nous *collons* b et m en posant $\ell = bm$ et invoquons (3.2), ce qui nous donne

$$S_2 = \sum_{b>y,\, \ell}\sum \mu(b)\operatorname{Log}\ell\, g(b\ell).$$

Nous inversons cette fois-ci l'inégalité en b en écrivant

$$S_2 = \Big(\sum_{b,\, \ell}\sum - \sum_{b\le y,\, \ell}\sum\Big)\mu(b)\operatorname{Log}\ell\, g(b\ell)$$

où la première somme est $\sum_n \Lambda(n)g(n)$ grâce au lemme 13.1, et la seconde $\sum_n u_n g(n)$. Nous traitons S_3 de façon similaire en inversant l'inégalité sur c et en *collant* c et m en cours de route. Signalons au lecteur qu'il existe une preuve utilisant des séries de Dirichlet qui permet d'y voir plus clair, mais elle est hors de la thématique de ce livre. ◇◇◇

Cette identité est singulière : le terme principal est effectivement porté par la partie linéaire ! Mais pour le montrer il faut utiliser deux évaluations que nous n'avons pas établie :

$$\begin{cases} \displaystyle\sum_{b\le y} \frac{\mu(b)}{b} = \mathcal{O}((\operatorname{Log} y)^{-2}), \\ \displaystyle -\sum_{b\le y} \frac{\mu(b)\operatorname{Log} b}{b} = 1 + \mathcal{O}((\operatorname{Log} y)^{-2}). \end{cases}$$

... et pour la partie bilinéaire, éviter l'inégalité de Cauchy-Schwarz.

Maintenant que nous avons localisé le terme principal dans la partie linéaire, il faut réussir à montrer que la partie bilinéaire ne contribue effectivement qu'au terme d'erreur pour une classe assez large de fonctions g. Il reste un écueil : dans notre traitement de cette partie, nous ne nous appuyons que sur une majoration de w_ℓ, ce qui est essentiellement équivalent à remplacer le $\mu(b)$ qui y apparaît par 1 et, partant, w_ℓ par $\operatorname{Log}\ell$. Nous perdons son caractère oscillant et la forme bilinéaire attenante est alors de la taille du terme principal, ruinant tous nos efforts !

Transformer d'abord ce terme s'avère par conséquent capital. C'est dans cette voie que se sont engagés les chercheurs des années 1970/2000 où il nous faut signaler les noms de Henryk Iwaniec, Matti Jutila et d'Étienne Fouvry, notamment pour ces deux auteurs dans un travail en commun avec Henryk Iwaniec, ainsi que le nom de Glyn Harman. Il a fallu attendre 1998 et une collaboration de John Friedlander et d'Iwaniec pour la formulation d'une hypothèse correcte pour le terme d'erreur, assortie d'une forme bilinéaire y donnant accès. Ils montrent aussi que cette hypothèse est réaliste en donnant une formule asympotique pour le nombre de nombres premiers de la forme a^2+b^4 inférieurs à une borne donnée dans le théorème exceptionnel suivant :

Théorème 13.2 (Friedlander & Iwaniec) *Nous avons*

$$\sum_{a^2+b^4\le P}\sum \Lambda(a^2+b^4) = \kappa^* P^{3/4}(1+o(1))$$

où $\kappa^* = \sqrt{2}\Gamma(1/4)^2/(3\pi^{3/2})$ *et où* a *et* b *parcourent les entiers positifs.*

Le caractère exceptionnel vient de ce que la suite d'où l'on extrait ces nombres premiers, c'est à dire la suite des valeurs prises par les a^2+b^4, contient très peu d'éléments, nommément environ $P^{3/4}$ jusqu'à P. Avant ce théorème, les seules suites pour lesquelles on ait eu accès à une formule asymptotique contenaient au moins $P/(\operatorname{Log} P)^A$ membres jusqu'à P, pour une certain $A \ge 0$, à l'exception toutefois des suites de Pyateski-Shapiro. Ce dernier a en effet montré dés 1953 que le nombre de nombres premiers de la forme « partie entière de n^c » et inférieurs à une borne P, ceci pour un c entre 1 et 12/11, est asymptotique à $P^{1/c}/\operatorname{Log} P$. La nouveauté de la suite de Friedlander & Iwaniec vient de ce qu'elle est plus difficile ... Comment donc mesurer cela autrement qu'au juger ? Tout simplement en disant que la méthode de Pyatetski-Shapiro ne permettait de traiter qu'une collection très restreinte de suites, alors que celle de Friedlander & Iwaniec a un domaine d'action beaucoup plus large. Elle s'applique notamment à des suites dont la définition est plus « arithmétique ».

Ces deux auteurs s'appuyent sur l'identité de Vaughan qu'ils modifient et dans laquelle ils surimposent une condition similaire à la condition $\text{pgcd}(n, Q) = 1$ du chapitre précédent. Nous quittons la lectrice ici et la laissons continuer seule sur ce chemin !

Problèmes

Bien que le but soit de comprendre en général la structure des nombres premiers, la tradition en arithmétique consiste à se concentrer sur des problèmes précis. Nous en déduisons l'information générale ultérieurement, tout comme ici nous avons étudié la densité des $\sin p$ avant de dégager les notions de parties linéaires et de parties bilinéaires.

Voici alors en vrac une liste de questions dans le même esprit que la preuve que nous venons de quitter. Dans ces énoncés, p désigne un nombre premier :

1. Que dire de la densité dans $[-1,1]$ de la suite des $\sin(2p+1)$?
2. Que dire de la densité dans $\mathbb{R}$ de la suite des $\tan(3p)$?
3. Que dire de la densité dans $[-1,1]$ de la suite des $\sin(p^2)$?
4. Étant donné un point x de $[0,1]$, existe t-il un nombre premier p tel que $\big|x - \|p/\pi\|\big| \leq 1/\operatorname{Log} p$? En existe-t-il une infinité ? Peut on remplacer le $1/\operatorname{Log} p$ par $1/p^{0.01}$?
5. Montrer que la suite $\mu(n)\sin n$ est elle aussi dense dans $[-1,1]$.

Quelques réponses : (1) Cette suite est dense dans $[-1,1]$, (2) cette suite est aussi dense dans $\mathbb{R}$, (3) cette suite est encore dense $[-1,1]$ mais la preuve devient difficile, (4) la réponse est oui au deux questions, mais je ne sais pas quelle identité donne le meilleur résultat, ni quel est le meilleur résultat possible. Signalons ici que Georges Rhin a démontré en 1974 l'équirépartition de toute suite de la forme $(f(p) \bmod 1)_p$, où p parcourt les nombres premiers, si f appartient à un ensemble très large de fonctions entières. (5) On traite la fonction de Möbius de la même façon que celle de van Mangoldt c'est à dire que l'on cherche une bonne écriture à l'aide d'une forme bilinéaire ...

Ressources bibliographiques

Nous donnons ici quelques références, essentiellement en accès libre sur le web.

Tout d'abord, `http://math-doc.ujf-grenoble.fr/` le portail documentaire de la cellule MathDoc donne accès à beaucoup de ressources, des thèses, des livres numérisées, des prépublications et toute une foultitude de documents. Nous y trouvons par exemple le livre de Legendre « Théorie des Nombres » où le lecteur découvrira la théorie de fractions continues comme exposée au début du 19ième siècle et les débuts du crible dans la partie consacrée au « nombre d'entiers inférieurs à N et premiers à N ». La prépublication de Friedlander & Iwaniec citée dans le dernier chapitre s'y trouve aussi. Elle est stockée sur le serveur ArXiv sous le titre « Asympotic sieve for primes ». L'article « The polynomial $X^2 + Y^4$ captures its primes » des mêmes auteurs contient la preuve du théorème 13.2.

`http://www.unilim.fr/laco/theses/1998/T1998_01.pdf` est l'adresse d'une version électronique de la thèse de Pierre Dusart, soutenue à Limoges en 1998. Attention, ce document fait 173 pages ... La lectrice y trouvera les plus récentes approximations des fonctions de comptage sur les nombres premiers.

Le site `http://www.dpmms.cam.ac.uk/Number-Theory-Web/` entretenu par Keith Matthews contient beaucoup de matériel, mais essentiellement en anglais. Si cela ne déroute pas le lecteur, nous conseillons alors la sous-rubrique `http://www.dpmms.cam.ac.uk/Number-Theory-Web/N4.html`

`http://www.mast.queensu.ca/~murty/erat.dvi` est un article de Ram Murty et Natarajan Saradha en anglais, où le crible d'Erathosène est repris de façon particulièrement élémentaire et élégante à l'aide d'une autre version de la méthode de Rankin.

Sur internet, le groupe de discussions `fr.sci.maths` est accueillant et un bon endroit où poser des questions, échanger en français avec d'autres personnes intéressées par les mathématiques et ce livre trouve en fait son origine dans une question posée sur ce groupe, ainsi que sur un site roumain, en juin 2004. D'ailleurs, j'avais ensuite pensé à publier une solution dans l'excellente *Revue de Mathématiques Spéciales*, mais il m'est apparu opportun de développer plus avant le matériel et donc de lui consacrer une monographie entière.

Le *Journal de Théorie de Nombres de Bordeaux* propose des volumes en consultations gratuites. Les articles publiés dans ce journal sont générale-

ment des articles de recherche, dont certains sont en français. Son adresse électronique :

`http://almira.math.u-bordeaux.fr/jtnb/`

Nous recommandons par exemple l'article de Jean-Pierre Massias et Guy Robin de 1996 sur des évaluations liées à la taille du $k$ième nombre premier, et l'article de Jean-Pierre Kahane de 1997 sur les nombres premiers généralisés.

Concernant les livres accessibles en français, il faut noter celui de Enrico Bombieri *Le grand crible dans la théorie analytique des nombres* publié par la Société Mathématique de France sous le numéro 18 et réédité en 1984. Il s'agit certes d'un livre d'un niveau nettement plus élevé que celui-ci, mais il est merveilleusement écrit.

Index

Table des matières

S'il est classique que la suite des $(\sin n)_n$ est dense dans $[-1,1]$ lorsque n parcourt l'ensemble des entiers relatifs, il est moins connu mais tout aussi vrai que la suite des $(\sin p)_p$ lorsque cette fois p est réduit à parcourir seulement l'ensemble des nombres premiers positifs a cette même propriété. Mais pour le démontrer, le chemin à parcourir est plus difficile. Nous accompagnons ici le lecteur le long de ce parcours qui permettra de comprendre plus avant la structure des nombres premiers. Il contient une preuve complète de la propriété annoncée à partir de connaissances du niveau de la première année d'université. Les diverses techniques auxiliaires sont décrites en détail afin que le lecteur puisse faire sienne cette démonstration et continuer seul l'exploration de ce domaine.