

## 44 nombres premiers consécutivement !

par Pierre GERMAIN-LACOUR [\*]

### Résumé.

La formule d'Euler  $x^2+x+41$  fournit 40 nombres premiers pour  $x$  entier allant de 0 à 39. Cette formule célèbre a été suivie de beaucoup d'études et même en 2006 d'un concours international. L'auteur expose ici les recherches qui l'ont conduit à trouver une autre formule donnant 44 nombres premiers, ce qui dépasse les formules analogues connues.

### I Les origines

Léonard Euler, 1707-1783, a effectué de nombreuses avancées dans les divers domaines des mathématiques. En 1772 il a publié la formule  $x^2+x+41$  qui fournit consécutivement 40 nombres premiers, tous positifs et tous différents, pour  $x$  entier allant de 0 à 39.

Cette formule étonnante et simple a fait ensuite l'objet de nombreuses études complémentaires. Il est possible, mais difficile, d'obtenir une formule du même type donnant plus de nombres premiers dans les mêmes conditions.

### II Des coefficients rationnels

Il y a plusieurs façons d'obtenir une autre formule qui fournit plus de nombres premiers. La plus simple est d'accepter des coefficients fractionnaires pourvu que le polynôme donne toujours des valeurs entières pour des valeurs entières de la variable. Par exemple :  $(1/2)x^2+(1/2)x = x(x+1)/2$  est toujours entier quelque soit  $x$ , puisque soit  $x$  soit  $x+1$  est pair. Le polynôme suivant :  $(x^6-78x^5+2455x^4-39762x^3+349948x^2-1610076x+3168684)/36$  fournit consécutivement 43 nombres premiers, tous positifs et tous différents, pour  $x$  entier allant de 0 à 42. Il y en a très probablement d'autres ayant encore de meilleurs scores. Mais pour rester au plus près de la formule d'Euler je n'accepte pas les coefficients fractionnaires. Si on choisit 50 ou 60 nombres premiers quelconques et que l'on calcule par la méthode de Lagrange le polynôme qui interpole le premier

[\*] <http://pgl10.chez.com/mathematiques.html>

nombre pour  $x=0$  puis le deuxième nombre pour  $x=1$ , etc. jusqu'au dernier, on obtient bien un polynôme à coefficients fractionnaires donnant consécutivement de nombreux nombres premiers. Mais, avec ce subterfuge, le degré du polynôme serait très élevé.

### III Des nombres premiers négatifs

La deuxième façon la plus courante pour obtenir un score avantageux est d'accepter en tant que nombres premiers obtenus les nombres négatifs qui sont premiers en valeur absolue. Wroblewski et Meyrignac ont donné dans le concours international référencé en [6] l'exemple suivant :  $x^5-99x^4+3588x^3-56822x^2+348272x-286397$  qui fournit consécutivement 47 nombres premiers en valeur absolue, tous différents, pour  $x$  allant de 0 à 46. Et cette formule :  $(x^6-141x^5+8311x^4-262983x^3+4729720x^2-46034028x+190148424)/24$  a un score de 49. La formule d'Euler ne fournit que des nombres positifs. Les nombres premiers obtenus consécutivement dans les formules admises seront tous positifs.

### IV Des nombres premiers redondants

Une autre façon d'obtenir un bon résultat est d'accepter un nombre premier obtenu plusieurs fois. Dans le concours référencé en [6] Marc Beyleveld a donné l'exemple suivant :  $x^4-97x^3+3294x^2-45458x+213589$  qui fournit consécutivement 50 nombres premiers en valeur absolue, parmi lesquels 9 sont négatifs et la valeur -271 est obtenue deux fois. On peut aussi noter que le polynôme d'Euler  $x^2+x+41$  symétrique par rapport à  $x = -1/2$  fournit les mêmes 40 nombres premiers pour  $x$  allant de -1 à -40 que ceux pour  $x$  allant de 0 à 39 ce qui pourrait faire dire qu'il fournit 80 nombres premiers pour  $x$  entier allant de -40 à +39. Si un nombre premier est obtenu plusieurs fois, faudrait-il le compter plusieurs fois ? De manière à rester fidèle à la formule d'Euler, les formules ayant des nombres premiers obtenus plusieurs fois ne sont pas conservées.

### V Tous les coups sont permis

Le concours référencé en [6] a rassemblé 118 participants venant de nombreux pays. Il comprenait trois catégories. Dans la première catégorie toutes les règles refusées ici étaient interdites. Et les gagnants en degrés 3 à 6 étaient :

$$f_3(x) = 42x^3+270x^2-26436x+250703$$

$$f_4(x) = 45x^4-3416x^3+96738x^2-1212769x+5692031$$

$$f_5(x) = x^5-61x^4+1339x^3-12523x^2+42398x+11699$$

$$f_6(x) = x^6-119x^5+5850x^4-152072x^3+2205416x^2-16929506x+53822339$$

Le meilleur résultat,  $f_4(x)$ , dû à Kazmenko et Trofimov, a un score de 42 et les autres, dus à Wroblewski et Meyrignac, ont un score de 40 ou 41. A l'inverse, on peut aussi accepter les règles qui sont refusées ici. C'est ce qui était admis dans les autres catégories du même concours. On a déjà cité l'exemple de Wroblewski et Meyrignac, avec ces règles favorables, ayant un score de 47, de même pour l'exemple de Marc Beyleveld ayant le score de 50. Il y en a d'autres. Quand on fait une exploration longue et difficile en acceptant ces règles on peut avoir un résultat encore meilleur. Dress et Landreau ont publié en 2012 le polynôme :  $(1/72)x^6 - (5/24)x^5 - (1493/72)x^4 + (1027/8)x^3 + (100471/18)x^2 - (11971/6)x - 57347$  qui fournit consécutivement 58 nombres premiers en valeur absolue, tous différents, pour  $x$  allant de -42 à +15 ce qui est reconnu comme un record.

## VI Une nouvelle formule

Ce qui précède montre bien la difficulté qu'il y a à découvrir des formules meilleures que celle d'Euler, tout particulièrement avec les règles qui sont acceptées ici. Depuis la publication antérieure, voir [3], j'avais trouvé ou retrouvé les formules suivantes ayant un score de 43 :

$$p_1(x) = 108x^4 - 9777x^3 + 331416x^2 - 4984701x + 28080037$$

$$p_2(x) = 53x^4 - 3738x^3 + 97338x^2 - 1097691x + 4521977$$

$$p_3(x) = 4x^4 - 112x^3 + 3822x^2 - 87104x + 624451$$

Et j'ai trouvé en mars 2018 la formule suivante :

$$q_1(x) = 15x^4 - 1203x^3 + 38454x^2 - 569418x + 3243881$$

qui fournit consécutivement 44 nombres premiers, tous positifs et tous différents, pour  $x$  entier allant de 0 à 43.

## VII Point de vue algorithmique

Le logiciel qui effectue cette exploration est programmé en C++ et fonctionne avec les entiers codés sur 64 bits. L'emploi d'une bibliothèque comme GMP, MPIR, NTL ou autres pouvant utiliser des entiers aussi grands que l'on veut est à déconseiller parce que les exécutions nécessaires sont très longues : il faut éviter tout ce qui pourrait les ralentir. Dans une première phase, relativement rapide, on effectue un crible d'Ératosthène entre 1 et 32000000000. Dans la seconde phase, pour chaque nombre premier  $p_n$  on forme les diverses suites de 5 nombres premiers  $p_0$  à  $p_4$  tels que :  $11 \leq p_0 < p_1 < p_2 < p_3 < p_4 = p_n$  où il est inutile de commencer avant 11. Pour chaque suite on calcule le polynôme du 4-ième degré qui vaut  $p_0$  pour  $x=0$ ,  $p_1$  pour  $x=1$ ,  $p_2$  pour  $x=2$ ,  $p_3$  pour  $x=3$  et  $p_4$  pour  $x=4$ . S'il n'est pas à coefficients entiers on passe à la suite suivante. S'il est à coefficients entiers on compte les nombres premiers obtenus consécutivement avant  $p_0$  et après  $p_4$  et si leur nombre est assez grand on fait le changement d'origine pour avoir  $x=0$  au premier nombre premier obtenu. Cette méthode a plusieurs

avantages. Les polynômes examinés ont tous au moins 5 valeurs premières consécutivement, les autres polynômes ne sont pas examinés. L'avantage essentiel de la méthode est son aptitude à stopper la session à la fin du traitement du nombre premier  $p_n$  en cours puis à continuer les calculs en commençant par le nombre premier  $p_n$  suivant dans une autre session. De plus, cela permet aussi d'organiser facilement plusieurs explorations en parallèle avec des intervalles différents pour le nombre  $p_n$ . A noter que cette méthode fait une impasse : on pourrait choisir pour chacun des  $p_0, p_1, p_2$  et  $p_3$  toutes les valeurs comprises entre 11 et  $p_4=p_n$  mais cela ralentirait très fortement l'exploration et ce n'est fait qu'exceptionnellement. La formule trouvée  $q_1(x)$  au score de 44 est obtenue pendant le traitement de  $p_n=50147$  et la première formule antérieure  $p_1(x)$  au score de 43 est obtenue avec  $p_n=19387$ . La taille du domaine associé à chaque nombre  $p_n$  à traiter a une très forte croissance, il est donc de plus en plus long de traiter les  $p_n$  les plus grands. Le nombre premier 50147 est le 5150-ième nombre premier et pour son traitement il faut environ 12 heures avec un ordinateur ayant un Intel Core i7. Par contre, la formule  $f_4(x)$  déjà citée de Kazmenko et Trofimov, ayant un score de 42, est obtenue avec  $p_n=7109$  et son traitement dure seulement une minute. Les formules ayant un score élevé sont rares, mais il y a aussi des cas exceptionnels où l'exploration du domaine associé au nombre  $p_n$  permet d'obtenir deux formules différentes ayant un score élevé, c'est ce qui arrive pour 50147 et 53593.

Le parcours des valeurs de  $p_0, p_1, p_2$  et  $p_3$  pour chaque valeur de  $p_4=p_n$  correspond à plusieurs boucles imbriquées de programmation. On peut avantageusement éviter les valeurs inappropriées de  $p_0$  où  $50p_0+6p_4$  n'est pas un multiple de 8 parce que le coefficient du premier degré de  $p(x)$  que l'on veut entier vaut :  $(-50p_0+96p_1-72p_2+32p_3-6p_4)/24$ . Et de même, on peut éviter les valeurs de  $p_1$  où  $p_0-4p_1+p_4$  est un multiple de 3 parce que le coefficient du quatrième degré de  $p(x)$  vaut :  $(p_0-4p_1+6p_2-4p_3+p_4)/24$ .

Pour examiner plus rapidement le domaine associé à chaque valeur de  $p_n$  on peut aussi, si on veut, augmenter l'impasse déjà effectuée en commençant l'exploration concernant  $p_n$  à  $p_0=p_n-\text{delta}$ , ou le premier nombre premier suivant, au lieu de 11, avec une valeur arbitraire pour delta. Cette variante permet d'aller plus loin. En explorant jusqu'à  $p_n=69677$  avec  $\text{delta}=8000$  et sans conserver la croissance de  $p_0$  à  $p_4$  on a une autre formule ayant un score de 44 :

$$q_2(x)=177x^4-22213x^3+1041433x^2-21610727x+167461447$$

## VIII Autres recherches amusantes

La recherche expliquée ci-dessus comporte plusieurs variantes. Il est possible de supposer qu'avec l'une de ces variantes ou bien avec une autre méthode on trouvera une formule vérifiant les mêmes règles ayant un score encore meilleur.

En attendant, on peut indiquer quelques recherches ou quelques résultats déjà obtenus qui méritent d'être signalés.

La formule  $(127x^4 - 8638x^3 + 216761x^2 - 2377778x + 9638452)/4$  fournit 42 nombres premiers tous différents pour  $x$  allant de 0 à 41, ils sont tous positifs sauf un seul, le 22-ième : -179.

La formule  $110x^5 - 1650x^4 + 9350x^3 - 24750x^2 + 30140x + 4704593$  fournit consécutivement 23 nombres premiers seulement, tous positifs. Mais le nombre premier 4717793 est obtenu 5 fois de suite pour  $x$  égal 1 à 5. Les autres sont tous différents. L'explication de ce résultat exceptionnel est la suivante : la formule vaut  $110*(x-1)*(x-2)*(x-3)*(x-4)*(x-5) + 4717793$ .

L'exploration des nombres premiers offre la possibilité d'y faire toutes sortes de recherches. On observe facilement que 19, 199 et 1999 sont des nombres premiers. Le nombre 5613099946 est peut-être le plus petit entier tel que en lui ajoutant une fois, deux fois, ... , dix fois le chiffre 9 à la fin on obtienne dix nombres premiers. Cette recherche nécessite l'utilisation d'entiers ayant une taille qui dépasse celle des entiers codés sur 64 bits.

Il y a un moyen très simple pour obtenir une formule qui fournit consécutivement beaucoup de nombres premiers. Soient :  $a(t)=(1-t)(2-t)/2$   $b(t)=t(2-t)$  et  $c(t)=t(t-1)/2$ . La fonction  $f(x) = a(x \bmod 3)p(x/3) + b(x \bmod 3)q((x-1)/3) + c(x \bmod 3)r((x-2)/3)$  fournit :  $f(0)=p(0)$   $f(1)=q(0)$   $f(2)=r(0)$   $f(3)=p(1)$   $f(4)=q(1)$  ... etc. On peut très facilement choisir pour  $p(x)$ ,  $q(x)$  et  $r(x)$  trois polynômes qui donnent chacun 35 nombres premiers différents pour  $x$  allant de 0 à 34. Il en résulte que la formule combinée  $f(x)$  fournit consécutivement 105 nombres premiers pour  $x$  allant de 0 à 104. On peut même combiner de cette manière plus de trois polynômes simples. Mais la formule combinée  $f(x)$  n'est pas un polynôme en raison de l'opération modulo.

L'exploration des nombres premiers suscite à la fois des recherches individuelles et des recherches collectives. Le théorème de Green et Tao nous dit que la suite des nombres premiers contient des suites arithmétiques arbitrairement longues. Il en démontre l'existence mais il n'indique rien pour aider à les trouver. Les plus longues suites arithmétiques de nombres premiers connues n'ont que 26 éléments. La plus ancienne à 26 éléments fut trouvée en 2010 par un français, Benoît Perichon, grâce au projet PrimeGrid qui associe en parallèle un très grand nombre d'ordinateurs personnels dans le monde entier.

## Références

[1] Jean-Paul Delahaye, Merveilleux nombres premiers, 2012, Belin, pages 146 à 158.

[2] François Dress et Bernard Landreau, <http://arxiv.org/pdf/1402.7312.pdf>, 2012, 28 pages.

[3] Pierre Germain-Lacour, Une nouvelle formule de nombres premiers, Quadrature n° 94, octobre 2014.

[4] Paulo Ribenboim, The Little Book of Bigger Primes, 2004, Springer, pages 138 à 151.

[5] David Wells, Prime Numbers, 2005, Wiley, pages 77-79.

[6] Al Zimmermann, Al Zimmermann's Programming Contests, <http://recmath.com/contest/PGP/index.php>

[7] Daniel Perrin, Pourquoi  $n^2+n+41$  a-t-il beaucoup de nombres premiers ? <https://www.math.u-psud.fr/~perrin/journeedu2311/redaction2311e.pdf>