Etudier des nombres d'écarts (Denise Vella-Chemla, 31.7.2016)

Ici, on fournit simplement des nombres d'écarts entre nombres premiers et certaines constatations les concernant.

Dans le tableau ci-dessous, sont fournis pour n une puissance de 10 :

- en seconde colonne, $\pi(n)$, le nombre de nombres premiers inférieurs ou égaux à n ;
- dans les colonnes $e = k$ le nombre de couples de nombres premiers dont le plus petit des deux est inférieur ou égal à n et dont la différence est k (d'écart k).

n	$\pi(n)$	$e=2$	$e=4$	$e=6$	$e=8$	$e=10$	$e=12$
10^2	25	8	9	16	9	11	15
10^3	168	35	41	74	38	51	70
10^4	1 229	205	203	411	208	270	404
10^5	9 592	1 224	1 216	2 447	1 260	1 624	2 421
10^6	78 498	8 169	8 144	16 386	8 242	10 934	16 378
10^7	664 579	58 980	58 622	117 207	58 595	78 211	117 486
10^8	5 761 455	440 312	440 258	879 908	439 908	586 811	880 196

n	$e=14$	$e=16$	$e=18$	$e=20$	$e=22$	$e=24$	$e=26$	$e=28$	$e=30$
10^2	10	9	14	10	7	15	9	8	18
10^3	48	39	74	48	41	79	42	41	99
10^4	245	200	417	269	226	404	240	248	536
10^5	1 488	1 233	2 477	1 645	1 351	2 475	1 348	1 468	3 329
10^6	9 878	8 210	16 451	10 972	9 171	16 343	8 928	9 784	21 990
10^7	70 463	58 606	117 463	78 218	65 320	117 342	64 203	70 440	156 517
10^8	528 095	441 055	880 444	586 267	489 085	880 927	480 567	528 313	1 173 934

On constate que les écarts multiples de 6, et parmi eux l'écart 30, sont bien plus fréquents que les autres (à cause de la nécessité pour les nombres premiers d'appartenir à telle ou telle classe de congruence selon les modules 2 et 3).

Pour les couples de nombres premiers dont le plus petit est inférieur à 100, on s'est intéressé aux nombres d'occurrence de tel ou tel écart de valeur paire comprise entre 2 et 96.

Mais il y a redondance dans les comptages : si deux écarts de longueur 4 se suivent, on comptera également un écart de longueur 8. On peut représenter les divisibilités entre les écarts, divisibilités qui entraînent des comptages redondant, par des chaînes (ces chaînes sont autant de branches d'un arbre de racine 2, 2 divisant tout nombre pair). On préfère une telle structure d'arbre plutôt que de graphe pour la relation de divisibilité, en "accrochant" un nombre à un père unique qui est son plus grand diviseur strict.

Voici les chaînes de divisibilité :

$2 \mid 4 \mid 8 \mid 16 \mid 32 \mid 64$	$2 \mid 54$
$2 \mid 6 \mid 12 \mid 24 \mid 48 \mid 96$	$2 \mid 58$
$2 \mid 6 \mid 18 \mid 36 \mid 72$	$2 \mid 62$
$2 \mid 6 \mid 42 \mid 84$	$2 \mid 66$
$2 \mid 10 \mid 20 \mid 40 \mid 80$	$2 \mid 70$
$2 \mid 10 \mid 30 \mid 60$	$2 \mid 74$
$2 \mid 10 \mid 50$	$2 \mid 78$
$2 \mid 14 \mid 28 \mid 56$	$2 \mid 82$
$2 \mid 22 \mid 44 \mid 88$	$2 \mid 86$
$2 \mid 26 \mid 52$	$2 \mid 94$
$2 \mid 34 \mid 68$	
$2 \mid 38 \mid 76$	
$2 \mid 46 \mid 92$	

En s'intéressant à la chaîne de divisibilité $2 \mid 4 \mid 8 \mid 16 \mid 32 \mid 64$, on a l'idée de comparer la somme des écarts de longueurs 2, 8 et 32 et la somme des écarts de longueur 4, 16 et 64. On constate que les colonnes

$e = 2$, $e = 4$, $e = 8$, $e = 16$, $e = 32$ et $e = 64$ sont sensiblement les mêmes. Ces proximités ont pour conséquence que les sommes de colonnes sont sensiblement égales aussi (les nombres à comparer sont en bleu).

n	$e=2$	$e=8$	$e=32$	$\sum$	n	$e=4$	$e=16$	$e=64$	$\sum$
10^2	9	9	6	24	10^2	9	9	7	25
10^3	36	38	37	111	10^2	41	39	36	116
10^4	206	208	196	610	10^2	203	200	201	604
10^5	1 225	1 260	1 204	3 689	10^2	1 216	1 233	1 255	3 704
10^6	8 170	8 242	8 196	24 608	10^2	8 144	8 210	8 261	24 615
10^7	58 981	58 595	58 565	176 141	10^2	58 622	58 606	58 664	175 892
10^8	440 313	439 908	439 524	1 319 745	10^2	440 258	441 055	440 328	1 321 641

De même, en étudiant la chaîne 2 | 6 | 12 | 24 | 48 | 96, la somme des écarts de longueurs 6 et 24 est proche de la somme des écarts de longueur 12 et 48.

n	$e=6$	$e=24$	$\sum$	n	$e=12$	$e=48$	$\sum$
10^2	16	15	31	10^2	15	13	28
10^3	74	79	153	10^2	70	72	142
10^4	411	404	815	10^2	404	409	813
10^5	2 447	2 475	4 922	10^2	2 421	2 483	4 904
10^6	16 386	16 343	32 729	10^2	16 378	16 501	32 879
10^7	117 207	117 342	234 549	10^2	117 486	117 313	234 799
10^8	879 908	880 927	1 760 835	10^2	880 196	880 138	1 760 334

On constate (un peu paradoxalement) que les colonnes $e = 3k$ sont sensiblement les doubles des colonnes $e = k$ (par exemple, pour $e = 4$ ou bien $e = 8$ ou encore $e = 10$, à mettre en relation avec $e = 12$, $e = 24$ ou encore $e = 30$) mais que cela n'est pas le cas pour les colonnes $e = 6$ et $e = 3 \times 6 = 18$ qui sont sensiblement les mêmes.

Enfin, deux dernières constatations très troublantes :

- observons les colonnes $e = 2$ et $e = 10$ d'une part, $e = 4$ et $e = 20$ d'autre part, et enfin $e = 6$ et $e = 30$. On multiplie l'écart dont on comptabilise les occurrences par 5 dans les 3 cas. Les nombres dans les colonnes quant à eux semblent être multipliés par $4/3$, ce rapport semblant s'améliorer au fur et à mesure ;
- de ce fait, on observe alors les colonnes $e = 2$ et $e = 14$ d'une part, et $e = 4$ et $e = 28$ d'autre part, dans le cas où les écarts sont multipliés par 7 et là, surprise, les contenus des colonnes semblent multipliés par $6/5$.

Les données concernant les puissances de 2 plutôt que les puissances de 10 semblent confirmer cela : ci-dessous les nombres d'occurences des écarts jusqu'à $2^{20} = 1\,048\,576$.

$\Delta =$ 2 : 8535	$\Delta =$ 26 : 9305	$\Delta =$ 50 : 11463	$\Delta =$ 74 : 8783	$\Delta =$ 98 : 10221
$\Delta =$ 4 : 8500	$\Delta =$ 28 : 10192	$\Delta =$ 52 : 9281	$\Delta =$ 76 : 9043	$\Delta =$ 100 : 11376
$\Delta =$ 6 : 17064	$\Delta =$ 30 : 22887	$\Delta =$ 54 : 17097	$\Delta =$ 78 : 18593	
$\Delta =$ 8 : 8588	$\Delta =$ 32 : 8543	$\Delta =$ 56 : 10232	$\Delta =$ 80 : 11362	
$\Delta =$ 10 : 11382	$\Delta =$ 34 : 9111	$\Delta =$ 58 : 8884	$\Delta =$ 82 : 8777	
$\Delta =$ 12 : 17056	$\Delta =$ 36 : 17115	$\Delta =$ 60 : 22752	$\Delta =$ 84 : 20498	
$\Delta =$ 14 : 10266	$\Delta =$ 38 : 9074	$\Delta =$ 62 : 8913	$\Delta =$ 86 : 8739	
$\Delta =$ 16 : 8545	$\Delta =$ 40 : 11416	$\Delta =$ 64 : 8594	$\Delta =$ 88 : 9494	
$\Delta =$ 18 : 17119	$\Delta =$ 42 : 20642	$\Delta =$ 66 : 19034	$\Delta =$ 90 : 22788	
$\Delta =$ 20 : 11443	$\Delta =$ 44 : 9575	$\Delta =$ 68 : 9194	$\Delta =$ 92 : 9074	
$\Delta =$ 22 : 9523	$\Delta =$ 46 : 8945	$\Delta =$ 70 : 13701	$\Delta =$ 94 : 8806	
$\Delta =$ 24 : 17004	$\Delta =$ 48 : 17170	$\Delta =$ 72 : 17118	$\Delta =$ 96 : 17161	

Il semblerait que les écarts apparaissant le moins souvent entre deux nombres premiers soient les puissances de 2, puis les écarts de la forme $2^k.p$ avec p premier, puis les écarts de la forme $2^k.5^{k'}$, puis les écarts de la forme $6k$, les écarts ayant beaucoup de diviseurs semblant bien plus nombreux que les autres.

On a alors l'idée de calculer par programme des nombres d'écarts entre nombres premiers tous les deux inférieurs à n ainsi que les sommes alternées de ces écarts (alors que les nombres fournis dans les tableaux ci-dessus comptent des écarts pour lesquels seul le plus petit des deux nombres premiers est inférieur à n).

On les fournit pour $n = 100$. On constate que les sommes alternées des nombres d'occurrence sont proches (on fait la somme des nombres d'occurrences pour les écarts doubles de pairs d'une part (apparaissant dans les colonnes impaires du tableau) et la somme des nombres d'occurrences pour les écarts doubles d'impairs d'autre part (dans les colonnes paires du tableau)).

écart	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38
nb occur.	8	8	15	9	10	13	9	8	12	8	6	12	8	6	13	5	5	11	6

écart	40	42	44	46	48	50	52	54	56	58	60	62	64	66	68	70	72	74	76
nb occur.	6	9	5	4	7	6	4	7	6	3	7	3	3	6	4	3	3	2	3

écart	78	80	82	84	86	88	90	92	94	96
nb occur.	3	2	1	2	2	0	1	1	1	0

Ci-dessous sont fournies les sommes alternées des nombres d'occurrences pour n une puissance de 10.

n	*écarts* $4k$	*écarts* $4k+2$
10^2	143	133
10^3	6 901	6 960
10^4	376 407	376 971
10^5	22 992 181	22 996 664

Les différences sont trop grandes pour être approfondies.

Enfin, comme on a constaté dans les tableaux précédents que certaines colonnes vérifient "presque" certaines relations de proportionnalité (par exemple, la colonne $e = 6$ est "presque" la somme des colonnes $e = 2$ et $e = 4$[1], on essaie d'obtenir des puissances de matrices 2×2 dont les valeurs approcheraient les valeurs des colonnes mais on n'y parvient pas.

On fournit ci-après les puissances successives de la matrice $D = \begin{pmatrix} 2 & 1 \\ 1 & 1/2 \end{pmatrix}$.

$D^2 = \begin{pmatrix} 5 & 2.5 \\ 2.5 & 1.25 \end{pmatrix}$	$D^9 = \begin{pmatrix} 3051.76 & 1525.88 \\ 1525.88 & 762.939 \end{pmatrix}$
$D^3 = \begin{pmatrix} 12.5 & 6.25 \\ 6.25 & 3.125 \end{pmatrix}$	$D^{10} = \begin{pmatrix} 7629.39 & 3814.7 \\ 3814.7 & 1907.35 \end{pmatrix}$
$D^4 = \begin{pmatrix} 31.25 & 15.625 \\ 15.625 & 7.8125 \end{pmatrix}$	$D^{11} = \begin{pmatrix} 19073.5 & 9536.74 \\ 9536.74 & 4768.37 \end{pmatrix}$
$D^5 = \begin{pmatrix} 78.125 & 39.0625 \\ 39.0625 & 19.5312 \end{pmatrix}$	$D^{12} = \begin{pmatrix} 47683.7 & 23841.9 \\ 23841.9 & 11920.9 \end{pmatrix}$
$D^6 = \begin{pmatrix} 195.312 & 97.6562 \\ 97.6562 & 48.8281 \end{pmatrix}$	$D^{13} = \begin{pmatrix} 119209 & 59604.6 \\ 59604.6 & 29802.3 \end{pmatrix}$
$D^7 = \begin{pmatrix} 488.281 & 244.141 \\ 244.141 & 122.07 \end{pmatrix}$	$D^{14} = \begin{pmatrix} 298023 & 149012 \\ 149012 & 74505.8 \end{pmatrix}$
$D^8 = \begin{pmatrix} 1220.7 & 610.352 \\ 610.352 & 305.176 \end{pmatrix}$	$D^{15} = \begin{pmatrix} 745058 & 372529 \\ 372529 & 186265 \end{pmatrix}$

Le 1220 de D^8 est "presque" $\pi(10^4)$.

[1]Ceci se disant de la façon suivante avec le langage mnémotechnique associé aux nombres premiers "On compte environ autant de premiers jumeaux et cousins mis ensemble que de premiers sexy !".

Voici les puissances successives de la matrices $T = \begin{pmatrix} 3 & 1 \\ 1 & 1/3 \end{pmatrix}$.

$T^2 = \begin{pmatrix} 10 & 3.33333 \\ 3.33333 & 1.11111 \end{pmatrix}$	$T^7 = \begin{pmatrix} 4115.23 & 1371.74 \\ 1371.74 & 457.247 \end{pmatrix}$
$T^3 = \begin{pmatrix} 33.3333 & 11.1111 \\ 11.1111 & 3.7037 \end{pmatrix}$	$T^8 = \begin{pmatrix} 13717.4 & 4572.47 \\ 4572.47 & 1524.16 \end{pmatrix}$
$T^4 = \begin{pmatrix} 111.111 & 37.037 \\ 37.037 & 12.3457 \end{pmatrix}$	$T^9 = \begin{pmatrix} 45724.7 & 15241.6 \\ 15241.6 & 5080.53 \end{pmatrix}$
$T^5 = \begin{pmatrix} 370.37 & 123.457 \\ 123.457 & 41.1523 \end{pmatrix}$	$T^{10} = \begin{pmatrix} 152416 & 50805.3 \\ 50805.3 & 16935.1 \end{pmatrix}$
$T^6 = \begin{pmatrix} 1234.57 & 411.523 \\ 411.523 & 137.174 \end{pmatrix}$	$T^{11} = \begin{pmatrix} 508053 & 169351 \\ 169351 & 56450.3 \end{pmatrix}$

Voici enfin les puissances successives de la matrices $S = \begin{pmatrix} 6 & 1 \\ 1 & 1/6 \end{pmatrix}$.

$S^2 = \begin{pmatrix} 37 & 6.166676.16667 & 1.02778 \end{pmatrix}$	$S^8 = \begin{pmatrix} 2034719.6 & 339119.9 \\ 339119.9 & 56520.0 \end{pmatrix}$
$S^3 = \begin{pmatrix} 228.2 & 38.0 \\ 38.0 & 6.3 \end{pmatrix}$	$S^9 = \begin{pmatrix} 12547437.5 & 2091239.6 \\ 2091239.6 & 348539.9 \end{pmatrix}$
$S^4 = \begin{pmatrix} 1407.0 & 234.5 \\ 234.5 & 39.1 \end{pmatrix}$	$S^{10} = \begin{pmatrix} 77375864.4 & 12895977.4 \\ 12895977.4 & 2149329.6 \end{pmatrix}$
$S^5 = \begin{pmatrix} 8676.7 & 1446.1 \\ 1446.1 & 241.0 \end{pmatrix}$	$S^{11} = \begin{pmatrix} 477151163.6 & 79525193.9 \\ 79525193.9 & 13254199.0 \end{pmatrix}$
$S^6 = \begin{pmatrix} 53506.1 & 8917.7 \\ 8917.7 & 1486.3 \end{pmatrix}$	$S^{12} = \begin{pmatrix} 2942432175.3 & 490405362.5 \\ 490405362.5 & 81734227.1 \end{pmatrix}$
$S^7 = \begin{pmatrix} 329954.5 & 54992.4 \\ 54992.4 & 9165.4 \end{pmatrix}$	$S^{13} = \begin{pmatrix} 18144998414.3 & 3024166402.4 \\ 3024166402.4 & 504027733.7 \end{pmatrix}$