INTRODUCING FERMAT SEQUENCES

Ridlo W. Wibowo[1,2]

[1]**Astronomy Department, Institut Teknologi Bandung, Bandung, Indonesia**
[2]**European Southern Observatory, Santiago, Chile**
ridlo.w.wibowo@gmail.com

Abstract. This paper focuses on simple observation of Fermat's factorization method (FFM) applied on the first 5×10^6 odd number. The number of iteration used in the FFM is related to the factor that the method found for each odd number. A new collection of sequences or sets of numbers (we called it Fermat-d sequences) can be generated based on the factor that we found using this method. For example, the first member (Fermat-1) in this collection of sequences is odd prime number sequence, which has factor–of course–of 1 and has highest number of iteration in the FFM, which is ∼half of the number itself. The next member of these sequences (Fermat-3, Fermat-5, and so on) are always start from d^2, continue with arithmetic progression of $2d$, but, with additional 'jumps' related to the prime number distribution. We present the first exploration of these new sequences (not registered in oeis.org yet) and the beauty of the location of these sequences in the Ulams and Sacks spiral.

Key words and Phrases: Fermat's factorization method, new integer sequences, prime number.

1. FERMAT'S FACTORIZATION

An odd integer N can always be represented by the difference between two square of integers (Eq 1), a and b for instance. This property is used in Fermat's factorization method to find the factors: c and d which are also odd integers,

$$N = a^2 - b^2 = (a + b)(a - b) = c \cdot d. \tag{1}$$

Other relation can also be derived: $a = (c + d)/2$ and $b = (c - d)/2$.

This method starts from ceiling the square root value of the number, $a = \lceil\sqrt{N}\rceil$. Iteration of $a \mathrel{+}= 1$ is then carried out until we find its pair, $b = a^2 - N$, to be a perfect square [2]. We will get Fig 1, if we plot the number of iteration needed to find the factor of odd integers using this method. I found this graph in 2013 when I play with this method in a homework from a Computational Science class, yet at that time I can not explain this graph immediately and not too interested on the subject. Lately, I understand how this plot produced and also I found that the sequences resulted from this method is interesting.

FFM is efficient if c/d is close to 1. One example is if N an odd square, then this method only needs 0 step. On the other hand, it requires a considerable number of trials when c/d is much larger than 1. The most extreme case would be if N is a large prime number. FFM will spend $(N + 1)/2 - \lceil\sqrt{N}\rceil$ steps to finish the procedure. This is not efficient to perform a primality test ($\mathcal{O}(N)$ steps). In

2000 Mathematics Subject Classification:
Received: dd-mm-yyyy, accepted: dd-mm-yyyy.

general, the required total number of steps (k) can be approximated by

$$k = a_{\text{found}} - \lceil\sqrt{N}\rceil \approx \frac{(c+d)}{2} - \sqrt{N} = \frac{(\sqrt{c}-\sqrt{d})^2}{2} = \frac{(\sqrt{N}-d)^2}{2d}. \tag{2}$$

Evidently, when $N \ggg d$, the number of steps in this method becomes

$$k \approx N/2d. \tag{3}$$

Eq. 3 explain the "line" structures in Fig 1.

Now, we can try to solve quadratic equation in Eq 2,

$$\begin{aligned} k &\approx \frac{(\sqrt{N}-d)^2}{2d} \\ 0 &\approx d^2 - 2(\sqrt{N}+k)d + N \\ c \text{ and } d &\approx \left(\sqrt{N}+k\right) \pm \left(\sqrt{2\sqrt{N}k+k^2}\right) \end{aligned} \tag{4}$$

Two terms in the parenthesis are a and b. Eq 4 gives us bounds and relation between the odd number N, factors (c and d), and the number of iteration in Fermat's method. If $k = 0$, N is a perfect square. Next, if c differs less than $\approx (4N)^{1/4}$ from $\lceil\sqrt{N}\rceil$, the method requires only 1 step (independent from how large is N).

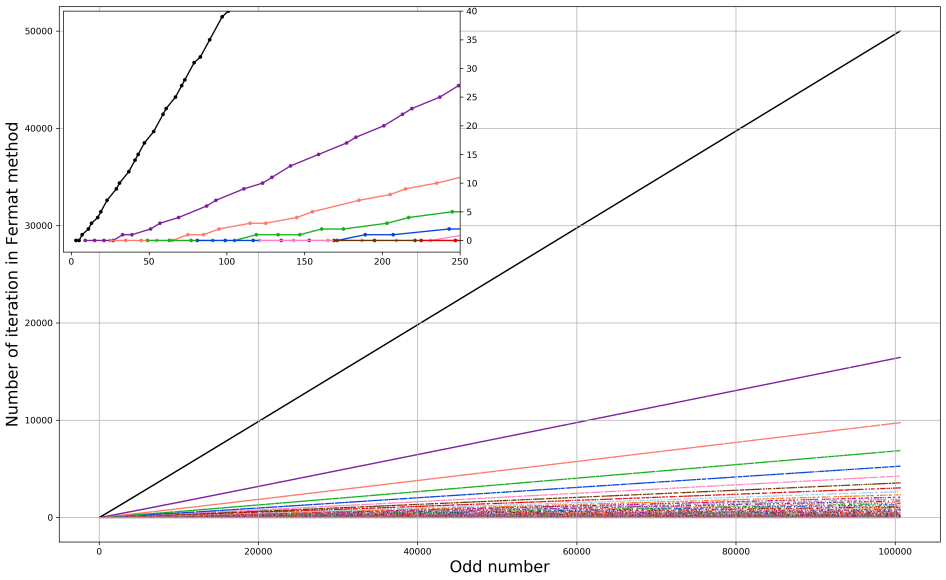

FIGURE 1. Number of iteration (k) in FFM for each odd number N. Each point are colored based on the factor of d found using FFM. For $d = 1$ (black), all the member are odd prime number; the next color are for $d = 3, 5, 7, 9, 11, \ldots$. As N become larger, the trend $k \approx N/2d$ appear in this plot.

If we make a sequence from number of steps k needed to factorize odd integers using Fermat's method, we will get,

$$\text{a(n)} = (0), 0, 0, 1, 0, 2, 3, 0, 4, 5, 0, 7, 0, 0, 9, 10, 1, 0, 12, 1, 14, 15, 0, 17, 0, \ldots$$

This sequence is similar to A078753; the only difference is that it is start from $N = 3$ and the early step before any increment of a is included as the first iteration.

2. FERMAT SEQUENCES/SETS

If we divide the odd integers based on the factor d that the Fermat's method found (i.e. coloring in Fig 1), we can make several "sequences" or "sets" of number. I will call it Fermat sequences, and more specifically Fermat-d sequences ($F_{[d]}$, d is odd number). Several member of these sequences are listed below:

$$
\begin{aligned}
F_{[1]} &= (1), 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, \ldots \\
F_{[3]} &= 9, 15, 21, 27, 33, 39, 51, 57, 69, 87, 93, 111, 123, 129, 141, 159, 177, 183, \ldots \\
F_{[5]} &= 25, 35, 45, 55, 65, 75, 85, 95, 115, 125, 145, 155, 185, 205, 215, 235, 265, \ldots \\
F_{[7]} &= 49, 63, 77, 91, 105, 119, 133, 147, 161, 175, 203, 217, 245, 259, 287, 301, \ldots \\
F_{[9]} &= 81, 99, 117, 135, 153, 171, 189, 207, 243, 261, 279, 333, 369, 387, 423, \ldots \\
F_{[11]} &= 121, 143, 165, 187, 209, 231, 253, 275, 297, 319, 341, 363, 385, 407, 451, \ldots \\
F_{[13]} &= 169, 195, 221, 247, 273, 299, 325, 351, 377, 403, 429, 455, 481, 507, 533, \ldots
\end{aligned}
$$

N	1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35	37	39
1	1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35	37	39
3	3	9	15	21	27	33	39	45	51	57	63	69	75	81	87	93	99	105	111	117
5	5	15	25	35	45	55	65	75	85	95	105	115	125	135	145	155	165	175	185	195
7	7	21	35	49	63	77	91	105	119	133	147	161	175	189	203	217	231	245	259	273
9	9	27	45	63	81	99	117	135	153	171	189	207	225	243	261	279	297	315	333	351
11	11	33	55	77	99	121	143	165	187	209	231	253	275	297	319	341	363	385	407	429
13	13	39	65	91	117	143	169	195	221	247	273	299	325	351	377	403	429	455	481	507
15	15	45	75	105	135	165	195	225	255	285	315	345	375	405	435	465	495	525	555	585
17	17	51	85	119	153	187	221	255	289	323	357	391	425	459	493	527	561	595	629	663
19	19	57	95	133	171	209	247	285	323	361	399	437	475	513	551	589	627	665	703	741
21	21	63	105	147	189	231	273	315	357	399	441	483	525	567	609	651	693	735	777	819
23	23	69	115	161	207	253	299	345	391	437	483	529	575	621	667	713	759	805	851	897
25	25	75	125	175	225	275	325	375	425	475	525	575	625	675	725	775	825	875	925	975
27	27	81	135	189	243	297	351	405	459	513	567	621	675	729	783	837	891	945	999	1053
29	29	87	145	203	261	319	377	435	493	551	609	667	725	783	841	899	957	1015	1073	1131
31	31	93	155	217	279	341	403	465	527	589	651	713	775	837	899	961	1023	1085	1147	1209
33	33	99	165	231	297	363	429	495	561	627	693	759	825	891	957	1023	1089	1155	1221	1287
35	35	105	175	245	315	385	455	525	595	665	735	805	875	945	1015	1085	1155	1225	1295	1365
37	37	111	185	259	333	407	481	555	629	703	777	851	925	999	1073	1147	1221	1295	1369	1443
39	39	117	195	273	351	429	507	585	663	741	819	897	975	1053	1131	1209	1287	1365	1443	1521
41	41	123	205	287	369	451	533	615	697	779	861	943	1025	1107	1189	1271	1353	1435	1517	1599
43	43	129	215	301	387	473	559	645	731	817	903	989	1075	1161	1247	1333	1419	1505	1591	1677
45	45	135	225	315	405	495	585	675	765	855	945	1035	1125	1215	1305	1395	1485	1575	1665	1755
47	47	141	235	329	423	517	611	705	799	893	987	1081	1175	1269	1363	1457	1551	1645	1739	1833
49	49	147	245	343	441	539	637	735	833	931	1029	1127	1225	1323	1421	1519	1617	1715	1813	1911
51	51	153	255	357	459	561	663	765	867	969	1071	1173	1275	1377	1479	1581	1683	1785	1887	1989
53	53	159	265	371	477	583	689	795	901	1007	1113	1219	1325	1431	1537	1643	1749	1855	1961	2067
55	55	165	275	385	495	605	715	825	935	1045	1155	1265	1375	1485	1595	1705	1815	1925	2035	2145

FIGURE 2. Multiplication table of odd numbers. Fermat-d are listed as columns with blue color. Fermat-1 are all odd prime numbers, Fermat-3 are multiple of 3 with $N \geq 3$, Fermat-5 are multiple of 5 with $N \geq 5$, and so on. The 'missing terms' are the one with white color. The position of the missing terms has 'rows structure'. The missing terms are appear in previous row.

These Fermat-d sequences always start from d^2. Each of the Fermat-d sequence progress with an arithmetic progression of $2d$, except for $F_{[1]}$ which are odd primes. Pseudo-Fermat sequences are sequences with formula:

$$\text{Pseudo–}F_{[d]}(i) = d^2 + 2d \cdot i = d(d+2i) \qquad d = 3, 5, 7, 9, ... \qquad i = 0, 1, 2, 3, ... \quad (5)$$

Fermat-d sequences ($d > 1$) follow this formula, however some of the terms are missing and appear in other sequences, e.g. 45 and 75 are missing from $F_{[3]}$ but appear in $F_{[5]}$, 63 appear in $F_{[7]}$, 81 appear in $F_{[9]}$, and so on. Generally, the missing terms are non-semiprime numbers; especially for Fermat-d when d is prime.

We can also say that these d-sequence are product of odd number d and odd number equal or greater than d, but missing some of the terms. In that sense, we can look for the behavior of these sequences in multiplication table of odd number like shown in Fig 2. Odd integers are union of Fermat-d sequences. From the multiplication table we can see the the missing terms on each Fermat-d sequences seems in the similar (and related) to the Fermat-1 or prime number distribution.

3. DISCUSSION

3.1. **On Other Factorization method.** As seen from the plot in Fig. 1, Fermat's factorization method are effective if the factors c and d are near $\sqrt{N}$. For some big odd integer N, if the number of iteration $k \gtrsim N/6$, then N must be prime number (we can stop the iteration) [1].

High number of iteration are needed not only for prime numbers, but also for composite number which has 2 factors; a very big and a very small number (early Fermat-d sequences). Therefore, the easiest way to improve this method is by combining it with trial division of some small terms of prime numbers. In the Fig. 1, by simply doing trial division of 3, we can exclude a numbers which are member of Fermat-3 sequence from the Fermat's core iteration; in which if we use original Fermat's method it will need a large number of iteration. The question is how many terms we need to do trial division to make this method more efficient to factorize an odd number N? To design an efficient algorithm for large number factorization, we always need to consider how fast the elementary operation can be carried out in a computer, e.g. addition, subtraction, multiplication, division and the extraction of a square root.

In addition to do trial division, we can also make a better method by making larger and "smarter jump" (not by increment of 1). Other useful trick is by multiplying N with some rational number u, in hope that uN is a product of two nearby integers, and then $\gcd(uN, N)$ may be taken to obtain the factorization of N [3]. The fundamental ideas of FFM are the basis of the best-known algorithms for factoring large semiprimes: quadratic sieve and general number field sieve [1]. The relation of Fermat-d sequences and these methods will be not discussed here.

3.2. **Prime counting function.** Prime-counting function is the function counting the number of prime numbers less than or equal to some real number x. It is denoted by $\pi(x)$. From prime number theorem (PNT), it is said that the distribution of prime number is asymptotic to the logarithmic integral,

$$\lim_{x\to\infty} \pi(x) = \mathrm{Li}(x) = \int_2^x \frac{1}{\ln t} dt.$$

We can make a plot of counting function not only for prime number, but also for each Fermat-d sequences. By counting the number of member of sequence below a value (x), we will get a plot like in Fig 3. It seems that as N getting bigger, the

counting function of Fermat-d sequences producing a similar distribution as of the primes number (Fermat-1, see Fig 3).

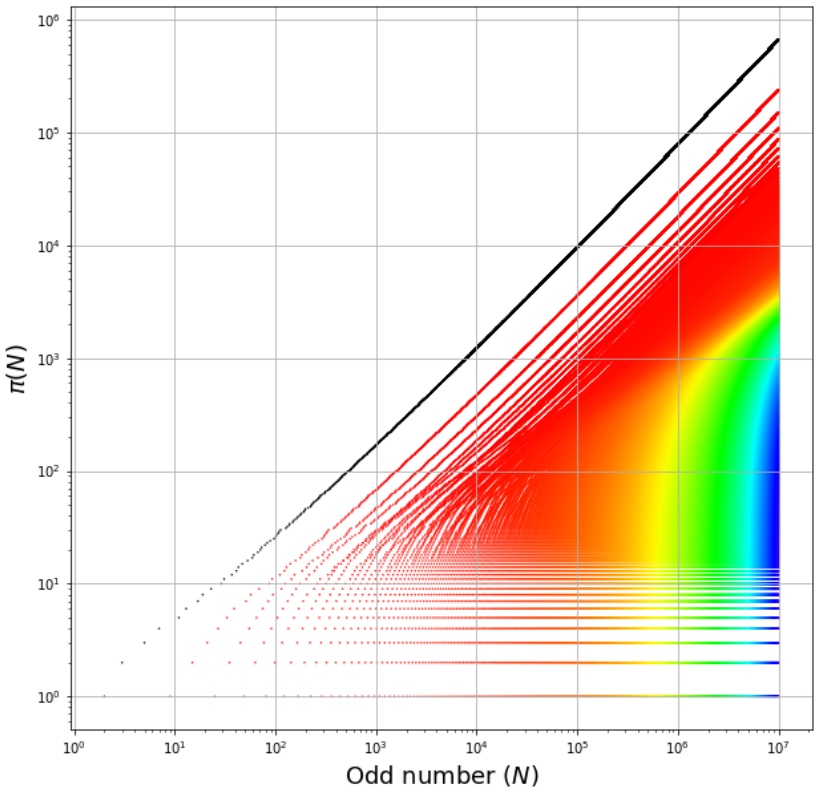

FIGURE 3. Prime counting together with 'Fermat' counting in log–log plot. Black for prime number (Fermat-1), asymptotic to logarithmic integral function; rainbow color for other Fermat-d sequences. It seems that the distribution of each Fermat-d sequences is similar to the prime number distribution (Fermat-1).

4. ULAM AND SACKS SPIRAL

Started by Stanislaw Ulam[5], we can make a graphical depiction of the set of prime numbers by writing the positive integers in a square spiral and specially marking the prime numbers. However, in this work, I also add odd composite numbers and marking it with different color for each Fermat-d sequences (Fig. 4). By writing non-negative integers on a ribbon and roll it up with zero at the center and arrange it so that all the perfect squares line up in a row on the right side, we can also generate other graphical depiction called Sacks spiral [4]. The position of Fermat-d sequences on Sacks spiral also presented in Fig. 5. Prime numbers are

like atoms in the world of numbers. However, in this spiral, numbers are like stars in our Galaxy and prime numbers are like the missing matter [1].

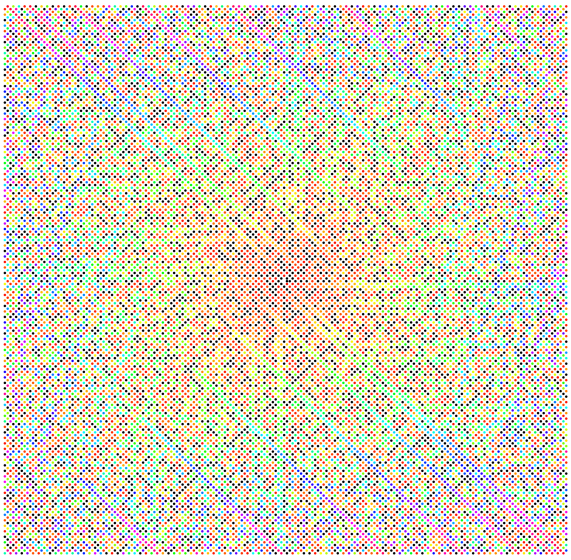

Figure 4. Fermat sequences plotted on Ulam spiral. Black circle are for prime numbers (original Ulam Spiral; Fermat-1 sequence, together with 2, but excluding 1). Rainbow color circles are for other Fermat-d sequences started from Fermat-3 (in red). The white spaces are for even numbers.

Acknowledgement. This work is mainly done in free time (sometimes not) during author's PhD life, which is supported by ESO studentship and Voucher ITB scholarship.

REFERENCES

[1] Crandall, R. and Pomerance, C., *Prime Numbers: A Computational Perspective*, 2nd ed., Springer-Verlag, New York, 2005.
[2] Hardy, G. H. and Wright, E. M., *An Introduction to the Theory of Numbers*, 4th ed., Oxford Univ. Press, London, 1960.
[3] Lehman, R. Sherman, "Factoring Large Integers", *Mathematics of Computation* **28** (1974): 637 - 646.
[4] Sacks, R., "Number Spiral", 2003, web site published online at: `http://www.numberspiral.com/` 2003 - 2007 by Robert Sacks
[5] Stein, M. L., Ulam, S. M., Wells, M. B., "A Visual Display of Some Properties of the Distribution of Primes", *American Mathematical Monthly*, **71**, 5 (1964): 516 - 520

[1] Video of Ulam and Sacks spiral based on Fermat-d sequences.

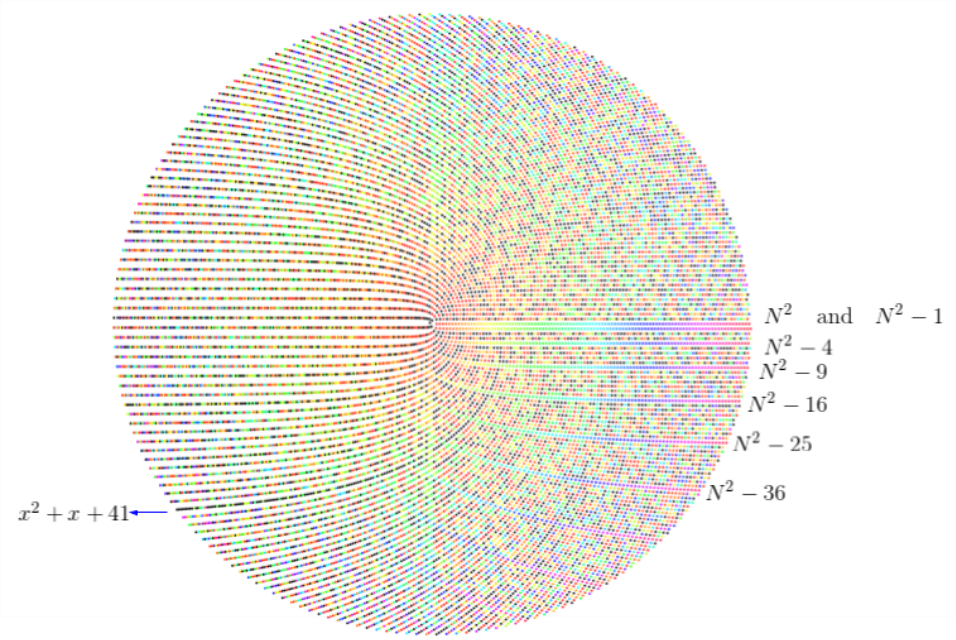

FIGURE 5. Fermat sequences plotted on Sacks spiral. Same configuration with Fig. 4. A curve from prime-generating formula discovered by Euler in 1772 is marked. 'Rainbow' curves found in this spiral are in the form of $N^2 - n^2$ (see S-curve in numberspiral.com).