

A Sieve for the Twin Primes

H.L. Mitchell

H. L. Mitchell

hlmitchell88@yahoo.com

CUNY-The City College

160 Convent avenue New York, NY 10031

Abstract. This paper presents a sieve for the twin primes using a double application of the Sieve of Eratosthenes to the k index for multiples of 6 centered around the primes ranging from 5 to \sqrt{N} . Considering the set $\{(6k-1, 6k+1)\}$, we look at the exact pattern in which multiples of primes p occur in the set with respect to the values of k . By deleting the values of k such that $6k-1$ is composite, we get the set of all pairs such that $6k-1$ is prime, then from these we delete those in which $6k+1$ is composite. That leaves us with the set of all twin primes less than N .

We derive a mathematical formula accordingly using the Euler product formula, the Prime Number Theorem, Mertens' Theorem and the Hardy-Littlewood Conjecture.

The main results are:

- 1) A sieve for the twin primes similar to the sieve of Eratosthenes for primes
- 2) A formula for the approximate number of twin primes less than N in terms of the number of primes
- 3) The asymptotic formula for the number of twin primes less than N

2010 Mathematics Subject Classification: 11A07; 11A41; 11B25; 11N05; 11N25; 11N36

Keywords and phrases: Primes, Twin Primes, Congruence, Sieve methods

1. Introduction

The main result of this paper is a sieve for the twin primes akin to the sieve of Eratosthenes for primes. It is applied twice to the set of all natural numbers k where $\{k\}$ is the *index* for all multiples of 6 but the range for the primes is from $p = 5$ to $p < N^{1/2}$.

We consider the set of all pairs $(6k - 1, 6k + 1)$ which are less than N and delete the values of k such that $6k - 1$ is composite. This leaves us with the pairs for which $6k - 1$ is prime. From these we delete the values of k such that $6k + 1$ is composite and that leaves us with the twin primes less than N .

Using the Euler product formula, The Prime number theorem and the Hardy-Littlewood Conjecture, we derive a formula for the approximation of $\pi_2(N)$ (the number of primes p less than N such that $p + 2$ is also prime) in terms of $\pi(N)$ (the number of primes less than N) and also the asymptotic formula for $\pi_2(N)$.

2. Deriving the formula and some set theory

All the twin primes except $\{3, 5\}$ are of the form $\{6k-1, 6k+1\}$

Let $T = \{(6k-1, 6k+1) \mid k = 1, 2, 3, \dots\}$

Let $u_k = 6k - 1$ and let $v_k = 6k + 1$

And define $t_k = (u_k, v_k)$

Listed below are the first few members of the set T . (the composite numbers are underlined)

5,7 11,13 17,19 23,25 29,31 35,37 41,43 47,49 53,55 59,61 65,67 71,73 77,79 83,85

89, 91 95, 97 101,103 107,109 113,115 119,121 125,127 131,133 137,139 143,145 149,151

$k \equiv \pm 1 \pmod{5} \Rightarrow t_k$ contains a multiple of 5 and is therefore not a pair of twin primes.

Let $S_p = \{t_k \mid t_k \text{ contains a multiple of prime } p\}$

$T_p = T \setminus S_p = \{t_k \mid t_k \text{ does not contain a multiple of prime } p\}$

$S_5 = \{t_4, t_6, t_9, t_{11}, t_{14}, t_{16}, t_{19}, \dots\} = \{(23, 25), (35, 37), (53, 55), (65, 67), \dots\}$

$T_5 = \{t_1, t_2, t_3, t_5, t_7, t_8, t_{10}, t_{12}, \dots\}$

The values of k in $S_5 = \{4, 6, 9, 11, 14, 16, 19, 21, \dots\} = \{k \in \mathbb{Z}^+ \mid k \equiv \pm 1 \pmod{5}\}$

The values of k in $T_5 = \{1, 2, 3, 5, 7, 8, 10, 12, 13, 15, 17, 18, 20, 22, 23, 25, 27, 28, 30, \dots\}$

$k \equiv \pm 1 \pmod{7} \Rightarrow t_k$ contains a multiple of 7

The values of k in $S_7 = \{6, 8, 13, 15, 20, 22, 27, 29, 36, 38, \dots\} = \{k \in \mathbb{Z}^+ \mid k \equiv \pm 1 \pmod{7}\}$

The values of k in $T_7 = \{1, 2, 3, 4, 5, 7, 9, 10, 11, 12, 14, 16, 17, 18, 19, \dots\}$

Since $6(2) - 1 = 11$, $k \equiv \pm 2 \pmod{11} \Rightarrow t_k$ contains a multiple of 11

The values of k in $S_{11} = \{9, 13, 20, 24, 31, 35, 42, 46, \dots\} = \{k \in \mathbb{Z}^+ \mid k \equiv \pm 2 \pmod{11}\}$

The values of k in $T_{11} = \{1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 14, 15, 16, 17, 18, 19, 21, 22, 23, 25, \dots\}$

If p is prime and $p = 6a \pm 1$, t_k contains a multiple of $p \Leftrightarrow k \equiv \pm a \pmod{p}$

Let \mathcal{P} = the set of all primes

Define T_w as $\{t_k = (u_k, v_k), k \in \mathbb{Z}^+ \mid u_k \in \mathcal{P} \text{ and } v_k \in \mathcal{P}\}$

$T_w = T \setminus \bigcup_{p \geq 5} S_p$; by De Morgan's Law $T_w = \bigcap_{p \geq 5} T_p$

Lemma 1:

Define k_p as the value of k for prime p such that $p = 6k + 1$ or $p = 6k - 1$.

and T_w as the set of all twin prime pairs.

Given a large integer N and $6k + 1 < N$,

$t_k \notin T_w \Leftrightarrow k \equiv \pm k_p \pmod{p}$ for some prime p , $5 \leq p < \sqrt{N}$.

As in the sieve of Eratosthenes, we delete $\{t_k \mid k = np \pm k_p \text{ for primes } p = 6k_p \pm 1\}$

$n = \{1, 2, 3, \dots\} \forall p \ni 5 \leq p < N$

Let $\pi_2(N)$ = the number of primes p less than N such that $p + 2$ is also prime.

$$(1) \quad \pi_2(N) \approx \frac{N}{6} \prod_5^V (1 - 2/p)$$

V = maximum prime p less than \sqrt{N}

Example : $N = 529$, $V = 19$

$$\pi_2(N) \approx \frac{529}{6} \prod_5^{19} (1 - 2/p) \approx 21$$

Actually, $\pi_2(N) = 25$

Let $\pi_2(N)$ = the number of primes p less than N such that $p + 2$ is also prime.

Where V is maximum prime $p < \sqrt{N}$

Table 1 ($\pi_2(N)$ compared to the formula)

N	$\pi_2(N)$	$\frac{N}{6} \prod_{p \leq \sqrt{N}} (1 - \frac{2}{p})$ Approx.
529	25	21
1000	35	31
2500	72	64
5000	126	111
7500	169	150
10000	205	191
15000	272	261
20000	342	328
25000	408	394
30000	467	456
35000	539	520
40000	591	570
50000	705	700
75000	958	968

This estimate becomes excessive for large values of N because for some primes p and interval $\{1, \frac{N}{6}\}$, the number of multiples of p in each of the sets $\{k: p | 6k - 1\}$ and $\{k: p | 6k + 1\}$

$= \lfloor \frac{N}{6p} + 1 \rfloor$ where $\lfloor \rfloor$ is the greatest function, some composites will not be sifted out and more primes will remain in the set.

3. Assessing $\pi_2(N)$ as N goes to infinity

We will consider Mertens 3rd Theorem, The Prime Number Theorem and The Hardy Littlewood Conjecture to arrive at the final results in this paper.

Mertens' 3rd theorem

$\lim_{n \rightarrow \infty} \ln n \prod_{p \leq n} (1 - \frac{1}{p}) = e^{-\gamma}$, where γ is the Euler - Mascheroni constant

$$\gamma = 0.5772156649\dots,$$

$$e^\gamma = 1.7810724\dots \text{ and } e^{-\gamma} = 0.561459\dots \quad (\text{Lagarias [2]})$$

The Prime Number Theorem

The Prime Number Theorem gives an asymptotic formula for the number of primes less than N as $\pi(n) \sim n / \ln n$ and for a better approximation $\pi(n) \sim \int_2^n 1 / \ln t \, dt$ which is the logarithmic integral $\text{li}(n)$.

The Hardy- Littlewood Conjecture

$$\pi_2(N) \sim 2C_2 \frac{N}{(\ln N)^2} \text{ where } C_2 \text{ is the twin prime constant } \prod_{p>2} \frac{p(p-2)}{(p-1)^2},$$

$$C_2 \approx 0.6601618458168\dots$$

with a better approximation $\pi_2(N) \sim 2C_2 \text{li}_2(N)$, where $\text{li}_2(N) = \int_2^N \frac{1}{(\ln t)^2} \, dt$.

We should take note the following identity:

$$1 - \frac{2}{p} = \frac{p(p-2)}{(p-1)^2} \left(1 - \frac{1}{p}\right)^2$$

Where the first factor on the right is used in the twin prime constant.

Except for 3, all lesser twin primes are of the form $6k - 1$.

Consider the set $\{u_k \mid u_k = 6k - 1, k \in \mathbb{Z}^+\}$.

Lemma 2:

Given $u_k < N$, $u_k \notin T_w \Leftrightarrow k \equiv \pm k_p \pmod{p}$ for some prime p , $5 \leq p < \sqrt{N}$.

i.e. (u_k, v_k) is not a pair of twin primes $\Leftrightarrow k \equiv \pm k_p \pmod{p}$ for some prime p , $5 \leq p < \sqrt{N}$.

Out of every p elements in the set $\{u_k\}$, (p prime and $p \geq 5$),

exactly one is a multiple of p and one precedes a $6k+1$ multiple of p .

If we list the elements of $\{u_k \mid u_k = 6k - 1, k \in \mathbb{Z}^+\}$ and delete every u_k in which $k \equiv \pm 1 \pmod{5}$ or $\pm 1 \pmod{7}$ or $\pm 2 \pmod{11}$ or $\pm 2 \pmod{13}$ or $k \equiv \pm 3 \pmod{17}$ or $\pm 3 \pmod{19}$... $\pm k_p \pmod{p}$ up to $p < \sqrt{N}$

The remaining terms are all twin primes.

We use this method to find twin primes in the table below by deleting all $k \equiv \pm 1 \pmod{5}$

or $\pm 1 \pmod{7}$, $\pm 2 \pmod{11}$ or $\pm 2 \pmod{13}$, since $13 = \max p < \sqrt{179}$

Table 2 (Values of k (not deleted) such that $6k-1$ is a lesser twin prime)

k	1	2	3	4	5	6	7	8	9	10
u_k	5	11	17	23	29	35	41	47	53	59
k	11	12	13	14	15	16	17	18	19	20
u_k	65	71	77	83	89	95	101	107	113	119
k	21	22	23	24	25	26	27	28	29	30
u_k	125	131	137	143	149	155	161	167	173	179

The u_k 's that correspond to the undeleted values of k are the lesser of twin primes

i.e.: 5, 11, 17, 29, 41, 59, 71, 101, 107, 137, 149, 179

We can demonstrate this sieve method by the following procedure:

first we cross out all values of k such that $k \equiv k_p \pmod{p}$, if $p \equiv -1 \pmod{6}$, i.e. $p = 6k_p - 1$ and

all values of k such that $k \equiv -k_p \pmod{p}$, if $p \equiv 1 \pmod{6}$, i.e. $p = 6k_p + 1$ up to $p < \sqrt{N}$

so that we are left with the set $\{k \in \mathbb{Z}^+ \mid k < \frac{N}{6} \text{ and } (6k-1) \text{ is prime}\}$.

Table 3 Values of k such that $6k-1$ is a prime (not deleted)

k	1	2	3	4	5	6	7	8	9	10
u_k	5	11	17	23	29	35	41	47	53	59
k	11	12	13	14	15	16	17	18	19	20
u_k	65	71	77	83	89	95	101	107	113	119
k	21	22	23	24	25	26	27	28	29	30
u_k	125	131	137	143	149	155	161	167	173	179

We then cross out the elements of the set $\{(k \in \mathbb{Z}^+ \mid k < \frac{N}{6} \text{ and } (6k+1) \text{ is composite})\}$ i.e.

$\{k \mid k \equiv -k_p \pmod{p} \text{ if } p \equiv -1 \pmod{6}\} \cup \{k \mid k \equiv k_p \pmod{p} \text{ if } p \equiv 1 \pmod{6}\}$. This leaves us with

the set of all twin primes less than N . (see Table 2)

This can be expressed mathematically as

$$(2) \quad \pi_2(N) \approx \frac{N}{6} \prod_5^V \frac{p-1}{p} \prod_5^V \frac{p-2}{p-1} \approx \frac{\pi(N)}{2} \prod_5^V \frac{p(p-2)}{(p-1)^2} \frac{p-1}{p}, \quad V = \max p < \sqrt{N}$$

By the Prime Number Theorem, $\pi(N) \sim \frac{N}{\ln N}$ and by Mertens' Theorem,

$$\prod_2^V \frac{p-1}{p} \sim \frac{2e^{-\gamma}}{\ln N} = \frac{1.122\dots}{\ln N} \quad \text{which overestimates the true ratio } \frac{\pi(N)}{N}. \quad [4] \text{ (Polya)}$$

By using $\frac{1}{\ln N}$, which is a lower bound for $\frac{\pi(N)}{N}$ [5] (Rosser and Schoenfeld)

and a little bit of algebra, we obtain

$$\pi_2(N) \sim \frac{N}{2 \ln N} \times \frac{4}{3} C_2 \times 3 \times \frac{1}{\ln N}$$

$$(3) \quad \pi_2(N) \sim 2C_2 \frac{N}{(\ln N)^2}$$

Hardy and Littlewood [2] gave a better approximation

$$(4) \quad \pi_2(N) \sim 2C_2 \int_2^N \frac{1}{(\ln t)^2} dt, \text{ also based on PNT}$$

Formula (4) is naturally equivalent to (3) but

$$(5) \quad \int_2^n \frac{1}{(\ln t)^2} dt = \frac{n}{(\ln n)^2} \left(1 + \frac{2!}{\ln n} + \frac{3!}{(\ln n)^2} + \dots\right),$$

and the second factor on the right hand side is (*for the values of n that we have to consider*) far from negligible. [2] (Hardy and Wright)

This suggests $2C_2 \frac{N}{(\ln N)^2} < \pi_2(N)$, for large enough values of N

M. Wolf noted, that the following formula

$$(6) \quad \pi_2(N) \sim 2C_2 \frac{[\pi(N)]^2}{N} \quad [6] \text{ (Wolf)}$$

agrees with numerical data much better than formula (3) but not as well as formula (4).

From equation (2) and the fact that $\frac{N}{6} \prod_5^N \frac{p-1}{p}$ is an approximation of $\frac{\pi(N)}{2}$,

we obtain (after multiplying the right side by $\frac{N}{6}$ and $\frac{6}{N}$) :

$$(7) \quad \pi_2(N) \approx \frac{\pi(N)}{2} \cdot \frac{4}{3} C_N \cdot \frac{6}{N} \cdot \frac{\pi(N)}{2} \text{ where } C_N = \prod_{3 < p < \sqrt{N}} \frac{p(p-2)}{(p-1)^2}$$

$$(8) \quad \pi_2(N) \approx 2C_N \frac{[\pi(N)]^2}{N} .$$

$$\lim_{N \rightarrow \infty} C_N = C_2 \wedge \pi(N) \sim \frac{N}{\ln N} \Rightarrow$$

$$(9) \quad \pi_2(N) \sim 2C_2 \frac{N}{(\ln N)^2}$$

From Rosser and Schoenfeld [5],

$$(10) \quad \frac{N}{\ln N} < \pi(N) \quad \forall N \geq 17 \quad \Rightarrow$$

$$(11) \quad 2C_2 \frac{N}{(\ln N)^2} < 2C_2 \frac{[\pi(N)]^2}{N} \quad \text{for large enough values of } N.$$

Table 4

(Values of $\pi_2(N)$ compared to logarithmic integral and ratio formulas) [1] (Caldwell)

N	$\pi_2(N)$	$2C_2 \operatorname{li}_2(N)$	$2C_2 \frac{N}{(\ln N)^2}$
10^6	8169	8248	6917
10^7	58980	58754	50822
10^8	440312	440368	389107
10^9	3424506	3425308	3074425
10^{10}	27412679	27411417	24902848
10^{11}	224376048	224368865	205808661
10^{12}	1870585220	1870559867	1729364449
10^{13}	15834664872	15834598305	14735413063
10^{14}	135780321665	135780264894	127055347335
10^{15}	1177209242304	1177208491861	1106793247903

$$\text{Let } W(N) = 2C_2 \frac{[\pi(N)]^2}{N}$$

Table 5

(limit $\frac{W(N)}{\pi_2(N)}$ approaching 1 as N increases)

N	$\pi(N)$	$W(N)$	$\frac{W(N)}{\pi_2(N)}$
10^6	78498	8136	0.9959603...
10^7	664579	58314	0.99251114...
10^8	5761455	438273	0.995242615...
10^9	50847534	3413659	0.9968325...
10^{10}	455052511	27340309	0.99735998...
10^{11}	4118054813	223905433	0.99790256...
10^{12}	37607912018	1867406346	0.998300599...
10^{13}	346065536839	15812374441	0.99859230168...
10^{14}	3204941750802	135619040528	0.99881219...
10^{15}	29844570422669	1176010096499	0.998981365...

References:

1. Caldwell, C.: An Amazing Prime Heuristic
utm.edu/staff/caldwell/preprints/Heuristics.pdf
2. Hardy, G. H., Littlewood, J. E.: Some Problems of 'Partitio Numerorum' III: On the Expression of a Number as a Sum of Primes. *Acta Mathematica* **44**, (1923) pp.42–44
3. Lagarias, J.: Euler's Constant: Euler's Work and Modern Developments. *Bulletin of the American Mathematical Society* **50** (4), (2013) pp.527- 628
4. Polya, G.: Heuristic Reasoning in the Theory of Numbers. *The American Mathematical Monthly* **66** (5), (1959) p. 383
5. Rosser, J. Barkley, Schoenfeld, Lowell: "Approximate formulas for some functions of prime numbers". *Illinois J. Math.* **6**, (1962) p. 69
6. Wolf, M.: Some Heuristics on the Gaps between Consecutive Primes.
ArXiv1102.0481 [math. NT] (2011) p. 8