A New Sieve for the Twin Primes

and how the number of twin primes is related to the number of primes

by

H.L. Mitchell

Department of mathematics CUNY-The City College

160 Convent avenue New York, NY 10031 USA

e-mail: hlmitchell88@yahoo.com

Abstract. We introduce a sieve for the number of twin primes less than *n* by sieving through the set $\{k \in \mathbb{Z}^+ \mid 6k < n\}$. We derive formula accordingly using the Euler product and the Brun Sieve.

We then use the Prime Number Theorem and Mertens' Theorem.

The main results are:

1) A sieve for the twin primes similar to the sieve of Eratosthenes for primes involving only the values of *k*, the indices of the multiples of 6, ranging over $k = p$,$5 \leq p < \sqrt{n}$.It shows the uniform distribution of the pairs (6*k*-1,6*k*+1) that are not twin primes and the decreasing frequency of multiples of *p* as *p* increases.

2) A formula for the approximate number of twin primes less than *N* in terms of the number of primes less than *n*

3) The asymptotic formula for the number of twin primes less than *n* verifying the Hardy Littlewood Conjecture.

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

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

1. Introduction

The twin primes have been studied by a number of mathematicians over the past 3 centuries and thus far it is not known whether there exist infinitely many of them.

Hardy and Littlewood proposed their famous conjecture in 1923 giving a formula for the number of twin primes less than a given integer n.

We introduce a sieve for the twin primes less than n similar to the sieve of Eratosthenes for primes. It is applied twice to the set of all natural numbers k such that $k < 6n$ and the range for the primes is $p = 5$ to $p < \sqrt{n}$.

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 Brun Sieve, The Prime number theorem and Mertens' 3rd Theorem, we derive a formula for the approximation of $\pi_2(n)$ in terms of $\pi(n)$ (the number of primes less than n) and the asymptotic formula for $\pi_2(n)$ to verify the Hardy Littlewood Conjecture.

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) | k = 1, 2, 3…}

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

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

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

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

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

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

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

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…}

$k \equiv \pm 1 \bmod 7 \Rightarrow t_k$ contains a multiple of 7

The values of k in S_7 = {6, 8, 13.15, 20, 22, 27, 29, 36, 38…} = $\{k \in \mathbb{Z}^+ \mid k \equiv \pm 1 \bmod 7\}$

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

Since 6(2) – 1= 11, $k \equiv \pm 2 \bmod 11 \Rightarrow t_k$ contains a multiple of 11

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

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…}

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

Let $\mathbb{P}$ = the set of all primes

Let T$_w$ = $\{t_k = (u_k, v_k) \mid u_k \in \mathbb{P}$ and $v_k \in \mathbb{P}, k \in \mathbb{Z}^+\}$

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

Lemma 1:

Define k_p as the value of k for primes $p = 6k + 1$ 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 $\{k \mid k = np \pm k_p$ for primes $p = 6k_p \pm 1\}$

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

Consider the set $K = \{k \in \mathbb{Z}^+ \mid k < 6N\}$.

In every interval $I \in K$ such that $I = \{np, (n+1)p\}$, $n \in \mathbb{Z}^+$ and p is prime, $5 \leq p < \sqrt{N}$

$\exists$ exactly 2 values of k ,(i.e. $k = np + k_p$ and $k = (n+1)p - k_p$) ,such that t_k contains a multiple of p

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

By the Brun Sieve we have:

(1) $$\pi_2(N) = \frac{N}{6} \prod_5^V (1 - 2/p) + R_p$$ where R_p is the error term

and V= maximum prime $p < \sqrt{N}$

Example 1:

$N = 529$, V=19

$$\boldsymbol{\pi_2(N)} \approx \frac{529}{6} \prod_5^{19} (1 - 2/p) = 20.6521\ldots$$

Actually, $\pi_2(N) = 25$ so $R_p \approx 4.3$

Let π_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 (π_2 (n) compared to the formula)

$\boldsymbol{n}$	$\boldsymbol{\pi_2(n)}$	$\frac{\boldsymbol{n}}{\boldsymbol{6}}\prod_{\boldsymbol{5}}^{\boldsymbol{V}}(\boldsymbol{1-2/p})$ 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 exceeds the actual number of twin prime pairs for large values of n because for some primes $p \notin T_w$ and elements in $\{k : 1 \leq k < \frac{n}{6}\}$, the number of elements in each of the sets $\{k : p \mid 6k - 1\}$ and $\{k: p \mid 6k + 1\} = \lfloor \frac{n}{6p} + 1 \rfloor$ where $\lfloor\ \rfloor$ is the greatest integer function, therefore some composites will not be sifted out by the product formula given above. The formula can be refined by rewriting it as a two-part sieve formula that represents the application of Eratosthenes' Sieve first to 6k-1 type numbers then to the 6k +1 type .

See equation (2).

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

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

Lemma 2:

Given $u_k < N$, $u_k \notin T_w \Leftrightarrow k \equiv \pm k_p (\text{mod } p)$ *for some prime p*, $5 \le p < \sqrt{N}$.

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

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

exactly one is a multiple of *p* and one precedes a 6*k*+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 ≡ ±1(mod 5) or ±1(mod 7) or ±2(mod 11) or ±2(mod 13) or k≡ ±3(mod 17) or ±3(mod 19) … $\pm k_p$(mod 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 $k \equiv \pm 1$ mod 5

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

Table 2 (Values of *k* (not deleted) such that 6*k*-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	1~~3~~	~~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 $k \equiv k_p \bmod p$ if $p \equiv -1 \bmod 6$ (i.e. $p = 6k_p - 1$)and all values of k such that $k \equiv -k_p \bmod p$ if $p \equiv 1 \bmod 6$ (i.e. $p = 6k_p + 1$) up to $p < \sqrt{n}$ so that we are left with the set $\{k \in \mathbb{Z}^+ \mid k < \frac{n}{6}$ and $(6k-1)$ 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 $\{(k \in \mathbb{Z}^+ \mid k < \frac{N}{6}$ and $(6k+1)$ is composite$\}$ i.e.

$\{k \mid k \equiv -k_p \bmod p$ if $p \equiv -1 \bmod 6\} \cup \{k \mid k \equiv k_p \bmod p$ if $p \equiv 1 \bmod 6\}$. This leaves us with the set of all twin primes less than N . (Table 2)

This can be expressed as approximation formula that follows :

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

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

(3) $\prod_2^V \frac{p-1}{p} \sim \frac{2e^{-\gamma}}{\ln N} = \frac{1.122\ldots}{\ln N}$ which overestimates the true ratio $\frac{\pi(N)}{N}$

and $\gamma =$ 5772156649 … is the Euler-Mascheroni constant. [4] (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:

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

$C_2 = \prod_{p\geq 3} \frac{p(p-2)}{(p-1)^2} = 0.6601618\ldots$ is the twin prime constant and

(5) $\pi_2(N) \sim 2C_2 \frac{N}{(\ln N)^2}$ which is the Hardy - Littlewood Conjecture.

Hardy and Littlewood [2] also conjectured a better approximation :

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

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

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

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 integers N

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

(because for some primes $p \notin T_w$, $5 \le p < \sqrt{N}$, $|\{k : p \nmid (6k-1) \setminus k_p\}|$

$= \lfloor \frac{N}{6}\left(\frac{p-1}{p} - \frac{6}{N}\right) \rfloor$ and likewise for $\{k : p \nmid (6k+1) \setminus k_p\}$) .

After multiplying the right side of equation (2) by $\frac{N}{6}$ and $\frac{6}{N}$ we obtain:

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

which includes $\lim_{N\to\infty}\left(\frac{N(p-2)-6(p-1)}{N(p-1)}\right)\left(\frac{Np}{N(p-1)-6p}\right)$,for some primes p, $5 \le p < \sqrt{N}$.

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

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

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

As shown by Rosser and Schoenfeld [5],

$\frac{N}{\ln N} < \pi(N) \quad \forall N \ge 17$ which gives us :

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

Table 4

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

n	$\pi_2\ (n)$	$2C_2\ \text{li}_2(n)$	$2C_2\ \frac{n}{(\ln n)^2}$
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

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…

Let $t_k = (6k - 1, 6k + 1)$ and T_w = the set of all twin prime pairs.

The occurrence of twin primes may be summarized as follows:

$\forall k > 3$ and primes p,

$t_k \in T_w \Leftrightarrow k \equiv 0, 2 \text{ or } 3 \pmod 5 \wedge k \not\equiv \pm k_p \pmod p \ \forall p > 5$

Where k_p is the value of k for the primes $p = 6k + 1$ or $p = 6k - 1$

References:

1.Caldwell, C.: An Amazing Prime Heuristic
utm.edu/staff/caldwell/preprints/Heuristics.pdf

2. Halberstam, H., Richert, H. E.: Sieve Methods. Academic Press (1974)

3. 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

4. Lagarias, J.: Euler's Constant: Euler's Work and Modern Developments. Bulletin of the American Mathematical Society **50** (4), (2013) pp.527- 628

5. Polya, G.: Heuristic Reasoning in the Theory of Numbers. The American Mathematical Monthly **66** (5)**,** (1959) p. 383

6. Rosser, J. Barkley, Schoenfeld, Lowell: Approximate formulas for some functions of prime numbers. Illinois J. Math. **6**, (1962) p. 69

7. Wolf, M.: Some Conjectures on the Gaps between Consecutive Primes

www.researchgate.net/publication/2252793 (1998) p.5