A Proof of the Twin Prime Conjecture

Abstract

The traditional definition of the twin prime conjecture is that there is an infinite number of twin primes. The traditional definition of a twin prime is a pair of odd primes separated by one even number, e.g., 29 and 31.

We give a distinctly new approach to the twin prime conjecture. We don't use the methods of analytic number theory. Instead, we use Eratosthenes' sieve[1]. We are not interested in the primes that are uncovered. Instead, we find in the sparse sequences of natural numbers that remain after each implementation of the sieve, useful structures that we call Eratosthenes' Patterns. These structures reveal ~~a number of~~ twin primes, and with each implementation of the sieve, these structures reveal increasing numbers of twin primes.

The essence of our proof is to show that the number of twin primes between p_n and p_n^2 approaches infinity as n approaches infinity.

Moreover, we have found a second constellation of primes, two primes that differ by four. In each of the Eratosthenes Patterns these primes are equal in number to the twin primes. We suggest that these are as significant as the twin primes. They are also infinite in number.

In order, we describe Eratosthenes' Sieve, Eratosthenes' Patterns, and the twins. Then we give the proof.

Before we begin, we cover two nomenclature topics: a definition and a mathematical notation.

Definition of prime number We restrict our definition to the natural numbers, $\mathbb{N}$.
Any number that has exactly two divisors is prime.
Any number that has three or more divisors is composite.
1 is the only number that has exactly one divisor and is neither prime nor composite. □

We have a possible confusion in our notation. We use (a,b) for the greatest common divisor of a and b. This notation is commonly used in Number Theory. We also use (a,b) for an open set of numbers, bounded by a and b. This notation is commonly used in Set Theory. The reader should be aware of the context.

The Sieve of Eratosthenes

Eratosthenes' sieve is a method for finding prime numbers. Eratosthenes lived circa 200 BCE.

In the Sieve of Eratosthenes, primes and multiples of primes are removed from the natural numbers. One starts with the natural numbers and creates a second set by removing all multiples of 2. Then one creates the third set by removing all multiples of 3 from the second set. This process is continued as long as desired. We name these sets A_1, A_2, A_3, ..., A_n with A_1 being the natural numbers.

We show portions of these sets here.

A_1	1 2 3 4 5 6 7 8 ...	the natural numbers
A_2	1 3 5 7 9 11 13 15 ...	multiples of 2 removed (odd numbers)
A_3	1 5 7 11 13 17 19 23 25 29 31 35 37 ...	multiples of 3 removed from A_2
A_4	1 7 11 13 17 19 23 29 31 37 41 43 47 49 53 59 ...	multiples of 5 removed from A_3

$\vdots$

All of the primes, however many there are, are in A_1
All of the primes greater than 2 are in A_2, in order.
All of the primes greater than 3 are in A_3, in order.
All of the primes greater than 5 are in A_4, in order.
$\vdots$

These statements seem obvious and trivial, but they will prove necessary below.

In each set, the first member is 1 and the second member is the nth prime.
In each set, the first composite is p_n^2. *** first composites are critical to this proof ***
In each set, apart from the 1, all of the numbers less than p_n^2 are prime.
When multiples of p_n are removed from A_n in order to create A_{n+1}, an infinity of composites is removed but only one prime is removed.
A_n contains all of the members of A_{n+i} where i is any positive integer.

By choosing the 2nd member of each set, A_1 through A_n, we construct a list of the first n primes.

All of the information given above is thousands of years old, except for the names of the sets, which we have chosen.

Instead of analyzing the primes that have been found, we have analyzed the numbers that are left behind after primes and their multiples have been removed from the natural numbers and we have found interesting and useful structures that we have named Eratosthenes' Patterns.

Eratosthenes' Patterns[2]

We show A_3 and A_4 rearranged ~~so as~~ to demonstrate the patterns that we have found.

A_3 original multiples of 2 and 3 have been removed from A_1

1 5 7 11 13 17 19 23 25 29 31 35 37 ...

A_3 rearranged into patterns

1	7	13	19	25	31	37	43	49	55	61 ...
5	11	17	23	29	35	41	47	53	59	65 ...

A_4 original multiples of 5 have been removed from A_3

1 7 11 13 17 19 23 29 31 37 41 43 47 49 53 59 61 67 ...

A_4 rearranged into patterns

1	31	61	91	121	151	181	211	241	271	301	331	361	391	421 ...
7	37	67	97	127	157	187	217	247	277	307	337	367	397	427 ...
11	41	71	101	131	161	191	221	251	281	311	341	371	401	431 ...
13	43	73	103	133	163	193	223	253	283	313	343	373	403	433 ...
17	47	77	107	137	167	197	227	257	287	317	347	377	407	437 ...
19	49	79	109	139	169	199	229	259	289	319	349	379	409	439 ...
23	53	83	113	143	173	203	233	263	293	323	353	383	413	443 ...
29	59	89	119	149	179	209	239	269	299	329	359	389	419	449 ...

As stated above, the first member in the first column is 1 and the second number is the nth prime. We call the columns patterns (to be explained below). The first pattern is the 'fundamental pattern'. Note the spaces after every p_n patterns. These delineate what we call extended patterns (to be explained below). The first extended pattern is the 'fundamental extended pattern'.

We use three parameters to describe the patterns, λ, ν, and τ. λ is used for the length of a pattern. Note that $\lambda_n \times p_n = \lambda_{n+1}$. ν is used for the number of members in a pattern, and

[2]We have not found these structures in the literature

τ is used for the number of twins in a pattern. (We explain twins below.) p is used for prime numbers. Note that the boundaries of the patterns are not members of the set. For example, in A_4, the boundaries are at 0, 30, 60, 90, ... Note that $\lambda_4 = 30$. The length of an extended pattern in A_4 is 210. $\lambda_4 \times p_4 = 210$.

The creation of A_{n+1} from A_n involves removing all multiples of p_n from A_n. The extended patterns in A_n become the patterns of A_{n+1}. When an extended pattern of A_n becomes a pattern of A_{n+1}, the patterns of A_n become subpatterns of the patterns of A_{n+1}. For example, the 2nd extended pattern of A_4, which is bounded by 210 and 420, becomes the second pattern of A_5. The patterns of the 2nd ~~extened~~ pattern of A_4 become subpatterns in the second pattern of A_5. These subpatterns are bounded by 210, 240, 270, 300, ... None of these boundaries is a member of A_5. We speak of 'stepping' from A_n to A_{n+1}.

Next, we show the sets A_1 through A_5 with the parameters added and the multiples of p_n in boldface. The distribution of the multiples of p_n is critical.

A_1 $p_1 = 2$, $\lambda_1 = 1$, $\nu_1 = 1$

1	**2**	3	**4**	5	**6**	7	**8**	9 ...

A_2 $p_2 = 3$, $\lambda_2 = 2$, $\nu_2 = 1$

1	**3**	5	7	**9**	11	13	**15**	17	19 ...

A_3 $p_3 = 5$, $\lambda_3 = 6$, $\nu_3 = 2$, $\tau_3 = 1$

1	7	13	19	**25**	31	37	43	49	**55**	61 ...
5	11	17	23	29	**35**	41	47	53	59	**65** ...

A_4 $p_4 = 7$, $\lambda_4 = 30$, $\nu_4 = 8$, $\tau_4 = 3$

1	31	61	**91**	121	151	181	211	241	1	**301**	331	361	391	421 ...
7	37	67	97	127	157	187	**217**	247	277	307	337	367	397	**427** ...
11	41	71	101	131	**161**	191	221	251	281	311	341	**371**	401	431 ...
13	43	73	103	**133**	163	193	223	253	283	313	**343**	373	403	433 ...
17	47	**77**	107	137	167	197	227	257	**287**	317	347	377	407	437 ...
19	**49**	79	109	139	169	199	229	**259**	289	319	349	379	409	439 ...
23	53	83	113	143	173	**203**	233	263	293	323	353	383	**413**	443 ...
29	59	89	**119**	149	179	209	239	269	299	**329**	359	389	419	449 ...

A_5 $\quad p_5 = 11,\ \lambda_5 = 210,\ \nu_5 = 48,\ \tau_5 = 15$

1	211	421	631	841	1051	1261	1471	1681	1891	**2101**	2311	2521 ...
11	221	431	641	851	1061	1271	1481	1691	1901	2111	**2321**	2531 ...
13	223	433	643	853	1063	1273	1483	1693	**1903**	2113	2323	2533 ...
17	227	437	647	857	**1067**	1277	1487	1697	1907	2117	2327	2537 ...
19	229	439	**649**	859	1069	1279	1489	1699	1909	2119	2329	2539 ...
23	233	443	653	863	1073	1283	1493	1703	1913	**2123**	2333	2543 ...
29	239	449	659	**869**	1079	1289	1499	1709	1919	2129	2339	2549 ...
31	241	**451**	661	871	1081	1291	1501	1711	1921	2131	2341	2551 ...
37	247	457	667	877	1087	1297	**1507**	1717	1927	2137	2347	2557 ...
41	251	461	**671**	881	1091	1301	1511	1721	1931	2141	2351	2561 ...
43	**253**	463	673	883	1093	1303	1513	1723	1933	2143	2353	**2563** ...
47	257	467	677	887	1097	1307	1517	**1727**	1937	2147	2357	2567 ...
53	263	**473**	683	893	1103	1313	1523	1733	1943	2153	2363	2573 ...
59	269	479	689	899	1109	1319	**1529**	1739	1949	2159	2369	2579 ...
61	271	481	691	901	**1111**	1321	1531	1741	1951	2161	2371	2581 ...
67	277	487	697	907	1117	1327	1537	1747	1957	**2167**	2377	2587 ...
71	281	491	701	911	1121	**1331**	1541	1751	1961	2171	2381	2591 ...
73	283	493	703	**913**	1123	1333	1543	1753	1963	2173	2383	2593 ...
79	289	499	709	919	1129	1339	1549	1759	**1969**	2179	2389	2599 ...
83	293	503	713	923	**1133**	1343	1553	1763	1973	2183	2393	2603 ...
89	299	509	719	929	1139	1349	1559	1769	1979	**2189**	2399	2609 ...
97	307	**517**	727	937	1147	1357	1567	1777	1987	2197	2407	2617 ...
101	311	521	731	941	1151	1361	1571	1781	**1991**	2201	2411	2621 ...
103	313	523	733	943	1153	1363	**1573**	1783	1993	2203	2413	2623 ...
107	317	527	**737**	947	1157	1367	1577	1787	1997	2207	2417	2627 ...
109	**319**	529	739	949	1159	1369	1579	1789	1999	2209	2419	**2629** ...
113	323	533	743	953	1163	1373	1583	**1793**	2003	2213	2423	2633 ...
121	331	541	751	961	1171	1381	1591	1801	2011	2221	**2431**	2641 ...
127	337	547	757	967	**1177**	1387	1597	1807	2017	2227	2437	2647 ...
131	**341**	551	761	971	1181	1391	1601	1811	2021	2231	2441	**2651** ...
137	347	557	767	977	1187	**1397**	1607	1817	2027	2237	2447	2657 ...
139	349	559	769	**979**	1189	1399	1609	1819	2029	2239	2449	2659 ...
143	353	563	773	983	1193	1403	1613	1823	2033	2243	**2453**	2663 ...
149	359	569	779	989	**1199**	1409	1619	1829	2039	2249	2459	2669 ...
151	361	571	**781**	991	1201	1411	1621	1831	2041	2251	2461	2671 ...
157	367	577	787	997	1207	1417	1627	**1837**	2047	2257	2467	2677 ...
163	373	**583**	793	1003	1213	1423	1633	1843	2053	2263	2473	2683 ...
167	377	587	797	1007	1217	1427	1637	1847	**2057**	2267	2477	2687 ...
169	379	589	799	1009	1219	1429	**1639**	1849	2059	2269	2479	2689 ...
173	383	593	**803**	1013	1223	1433	1643	1853	2063	2273	2483	2693 ...
179	389	599	809	1019	1229	1439	1649	**1859**	2069	2279	2489	2699 ...
181	391	601	811	1021	1231	**1441**	1651	1861	2071	2281	2491	2701 ...
187	397	607	817	1027	1237	1447	1657	1867	2077	2287	**2497**	2707 ...
191	401	611	821	1031	1241	1451	**1661**	1871	2081	2291	2501	2711 ...
193	403	613	823	1033	**1243**	1453	1663	1873	2083	2293	2503	2713 ...
197	**407**	617	827	1037	1347	1457	1667	1877	2087	2297	2507	**2717** ...
199	409	619	829	1039	1249	1459	1669	1879	2089	**2299**	2509	2719 ...
209	419	629	839	1049	1259	1469	1679	1889	2099	2309	**2519**	2729 ...

Note the repetition in the distribution of the numbers in boldface (multiples of p_n) across the extended patterns. This is ~~due to the fact~~ that a member of an extended pattern plus the length of an extended pattern ($\lambda_n \times p_n$) gives a corresponding member in the next extended pattern. If one of these is a multiple of p_n then the other must also be a multiple. Also, we find that in any row of an extended pattern, there is exactly one multiple of p_n. We give a theorem to explain this.

The Equal Spacing Theorem

Before stating the theorem, we give an example of its use.

Consider the sequence 13, 18, 23, 28, 33, 38, 43, 48, 53, 58 which has a spacing of 5. In any three consecutive members, exactly one is a multiple of 3. In any eight consecutive members exactly one is a multiple of 8. However, in ten consecutive members, there is no multiple of ten since ten and five are not relatively prime.

Theorem 1 In any n consecutive members of a sequence of equally spaced positive integers, with spacing b, exactly one will be divisible by n, provided that b and n are relatively prime.

Proof Let b be the spacing between members of the sequence. Let a be the beginning of the sequence. We wish to show that $a + bx = ny$. This is a linear Diophantine equation[3] in two unknowns, x and y and is solvable since the statement of the theorem requires that b and n be relatively prime. x gives the number of b spacings and y is the desired multiple of n. There is an infinte number of solutions. Choose the solution pair with the smallest positive value of x, say, x_0, y_0. Then the solutions are $x = x_0 + nt$ and $y = y_0 + bt$ with t ranging over all integers. We show that there is exactly one multiple of n in a sequence of n integers. When $t = 0$, $x = x_0$, and $y = y_0$ and ny_0 is a number in the sequence. When $t = \pm 1, y = y_0 \pm b$, and $n(y_0 \pm b)$ are numbers that differ from ny_o by more than n. □

We apply Theorem 1 to a row in an extended pattern in A_n. The spacing between the members in a row is λ_n. The number of members in the row is p_n. $(\lambda_n, p_n) = 1$. Thus, there is exactly one multiple of p_n in any row of an extended pattern.

From the above description of the construction of the sets, it is clear that any member of A_n can have no factors among the sequence of consecutive primes, $(2, ..., p_{n-1})$. Since λ_n is the product of all the consecutive primes, $(2, ..., p_{n-1})$, all members of A_n are relatively prime to λ_n. This implies that all of the factors of a member of A_n must be, apart from 1, greater than or equal to p_n, and, any natural number that contains no factors among the set of consecutive primes, $(2, ..., p_{n-1})$ is in A_n.

Symmetries in the patterns and extended patterns

We start with a useful feature of common divisors.

Theorem 2 $(m, a) = (m - a, a)$, where $m > a$, and $m, a \in \mathbb{N}$
Proof Let $g_1 = (m, a)$ and $g_2 = (m - a, a)$. g_1 also divides $m - a$ and $g_1 | g_2$. g_2 also divides $m - a + a = m$ and $g_2 | g_1$. Thus $g_1 = g_2$. □

Let a be a member of A_n that is less than $\lambda_n/2$ (a is relatively prime to λ_n). Theorem 2 shows that $\lambda_n - a$ is also relatively prime to λ_n and is therefore a member of A_n. a and $\lambda_n - a$ are equidistant from $\lambda/2$. This shows that the ~~distribuion~~ of the members of the upper half of the

[3] See for example: NIVEN and ZUCKERMAN, section 5.2

fundamental pattern is a mirror image of the distribution of the members of the lower half. If a is a multiple of p_n, then since λ_n is not a multiple of p_n, $\lambda_n - a$ will not be a multiple of p_n.

In an extended pattern, the multiples of p_n are s~~ymmtrically~~ diagonally, arranged about the center of an extended pattern.

Let a be a member of the ~~fundmental~~ extended pattern with $a < \lambda_n p_n/2$. All of the factors of a are greater than or equal to p_n. $\lambda_n p_n$ is the upper boundary of the fundamental extended pattern and is the product of the consecutive primes, $2, ..., p_n$. Thus, $\lambda_n p_n - a$ has no factors less than p_n and is a member of A_n. Note that a and $\lambda_n p_n - a$ are equidistant from $\lambda_n p_n/2$. Also note that if a is multiple of p_n, then $5\lambda_n p_n - a$ is also a multiple of p_n.

These are the reasons for our choice of p_n patterns as the size of the extended patterns.

We have found an expression that can be used to generate any member of $A_n, n \geq 2$. This allows one to build portions of a set without having previous sets available.

$$a = (\lambda_n/2) \pm 2^j p_{k_1}^{b_1} p_{k_2}^{b_2} ... \;, \qquad j > 0,\; k_i \geq n,\; b_i \geq 0,\; n \geq 2,\; \lambda_n/2 = \prod_{i=2}^{n-1} p_i \tag{1}$$

In the minus case of the above plus or minus operator, the parameters must be restricted so that the values of a are positive.

The first term on the right side is the product of all the odd primes, 3, ..., $\lambda/2$. The second term must contain a power of 2, making a an odd number. In addition, the second term may contain powers of odd primes and they must be a unique collection and they must not contain any of the primes in the first term. Otherwise a would not be a member of the set.

Primes and Composites

Recall from the above that any member of A_n less than p_n^2 is prime. Also note that all composites between p_n^2 and p_{n+1}^2 are multiples of p_n. This is used in the proof below.

When one studies the patterns, in any set, and observes the infinitude of numbers that have been removed from the natural numbers, one should keep in mind the fact that only composite numbers have been removed, with one exception: the gap between 1 and p_n. The first two members in a fundamental pattern ~~is~~ the only ~~place~~ where primes have been removed. In fact, in any gap between two consecutive members of any set, no primes have been eliminated, apart from the first gap of the fundamental pattern. The following is an example.

1073 and 1087 are two consecutive members of A_{10}, ($p_{10} = 29$). The 13 natural numbers that fall between them are all composite. We give the factors of the odd members. Note that the non-members all have at least one factor less than 29.

$29, 37|1073$ a member of A_{10}
$5, 43|1075$
$3, 359|1077$
$13, 83|1079$
$23, 47|1081$
$3, 19|1083$
$5, 7, 31|1085$
1087 is prime a member of A_{10}

When stepping from one set to the next, ~~and~~ an extended pattern in A_n becomes a pattern in A_{n+1}, the ratio of primes to composites in the pattern in A_{n+1} is greater than the ratio in the extended pattern in A_n.

Growth of the gaps between primes

For the proof given below we need to show that the average gap between the primes (not twins) increases slowly without limit.

$\pi(x)$ is commonly used to designate the number of primes that are less than or equal to x.

According to the prime number theorem:[4]

$$\lim_{x\to\infty} \pi(x)\frac{log(x)}{x} = 1.$$

If we let $\pi^*(x) = \frac{x}{logx}$, then

$$\lim_{x\to\infty} \pi^*(x)\frac{log(x)}{x} = 1.$$

And since

$$\lim_{x\to\infty} \pi(x)\frac{log(x)}{x} = 1,$$

we have

$$\lim_{x\to\infty} \pi(x) = \lim_{x\to\infty} \pi^*(x).$$

Now, $\dfrac{d\pi^*(x)}{dx} = \dfrac{logx-1}{log^2x}$. This approaches $\dfrac{1}{logx}$ as x approaches infinity.

We now have,

$$\lim_{x\to\infty} \frac{d\pi(x)}{dx} = \lim_{x\to\infty} \frac{d\pi^*(x)}{dx} = \frac{1}{logx}$$

[4] references 5,6,7,8,9

As n approaches infinity, the average gap between two consecutive primes approaches $\log p_n$, which approaches infinity.[5]

We use g_n for the gap between primes: $g_n = p_{n+1} - p_n$.

Number Theory

We step aside to show the reader that these sets have some interesting features which one finds in the study of Number Theory.

The members of the fundamental pattern in A_n are a reduced residue system[6] of λ_n.

The members of the fundamental pattern in A_n are a multiplicative group modulo λ_n[7]

The number of members in the fundamental pattern in A_n, ν_n, is Euler's totient[8] for λ_n.

As we search through the extended patterns of A_n we see that there is a multiple of p_n in the same position of ~~a~~ every extended pattern. These particular multiples are separated by $p_n \times \lambda_n$. Thus, they are a residue class[9] in A_n.

These features are not necessary for the proof which follows, but we find this interesting.

The Twins

We define two types of twins: two-twins and four-twins[10]. Historically, the term, twins, has referred to what we call two-twins. Two-twins are two odd numbers which differ by two, such as 17 and 19. The even number that falls between these two odd numbers is a multiple of six (an even multiple of three). Four-twins are two odd numbers that differ by four, such as 19 and 23. The odd number that falls between these is an odd multiple of three (~~explantion~~ given below). It will become apparent that all twins are created in A_3.

We repeat a portion of A_3

A_3 $\quad p_3 = 5,\ \lambda_3 = 6,\ \nu_3 = 2,\ \tau_3 = 1$

1	7	13	19	25	31	37	43	49	55	61 ...
5	11	17	23	29	35	41	47	53	59	65 ...

When interested in two-twins in A_3, we notice that all members are of the form $6n \pm 1$. When interested in four-twins, we notice that all members are the form $6m + 3 \pm 2$. It is also clear that any member of A_3 can be expressed in either format. In general, both $6n \pm 1, n \geq 1$, and

[5] D. Koukoulopoulos, Chapters 28 & 29
[6] See, for example, Apostol, section 5.2
[7] See, for example, Niven, section 2.11
[8] See, for example, Niven, section 2.1
[9] See, for example, Apostol section 5.2
[10] We have not found this terminology in the literature

$6m+3\pm 2, m \geq 0$ give identical sets of numbers, i.e., the set A_3

Since all primes greater than or equal to 5 are in the set, A_3, we can justify the following theorem.

Theorem 3 All primes ≥ 5 are in one of the forms, $6n \pm 1$, and, are also in one of the forms, $6m+3\pm 2$. □

In the case of a two-twin, the two defining members of the twin cannot be multiples of three (they are members of A_3), so the number between them must be a multiple of three and it must be even, hence it is a multiple of six.

In the case of a four-twin, the two members that define the twin cannot be multiples of three (they are members of A_3), so the number centered between them must be a multiple of three and it must be odd. The other two numbers between the defining members are even.

In general, there can be no members of the set between the two defining members of a two-twin or of a four-twin. This implies that there ~~are~~ no primes between the defining members.

To designate a 2-twin, we use the multiple of 6 (even multiple of 3) on which the twin is centered. For example, the 2-twin, 17,19, is designated as 18. To designate a 4-twin, we use the odd multiple of 3 on which the twin is centered. For example, the four-twin, 19,23 is designated as 21. A twin with two prime members is called a prime twin. A twin with one prime member is called a combination twin, and a twin with two composite members is called a composite twin.

We have adopted the convention that 3,5 is not a two-twin, since 3 is not a member of A_3.

A number can only be a member of a unique two-twin. A number can only be a member of a unique four-twin. However, a number can simultaneously be a member of both a two-twin and a four-twin, but cannot be a member of more than one two-twin or more than one four-twin. For example, 19 is a member of the two-twin, 18, and is also a member of the four twin, 21. But 19 is not a member of any other two-twin nor any other four-twin. This is clear when one looks at the alternating two-twins and four-twins in A_3. Also, it is clear, from the structure of A_3, that the spacing between any two twins, either 2-twins or 4-twins, is a multiple of 6.

In the steps to higher numbered sets, two-twins and four-twins can only be destroyed and cannot be created. Thus, the density of either type of twin in A_{n+1} must be less than that of A_n.

All twins, two-twins and four-twins, are created in A_3.

Next, we look at another feature that adds to the claim that the four-twins are as significant as the two-twins. We will show that the number of two-twins equals the number of four-twins.

First we give an example of the process used in the proof.

We recall a portion of A_4: $p_4 = 7$, $\lambda_4 = 30$, $\nu_4 = 8$, $\tau_4 = 3$

1	31	61	**91**	121	151	181	211	241	271	**301**	331	361	391	421 ...
7	37	67	97	127	157	187	**217**	247	277	307	337	367	397	**427** ...
11	41	71	101	131	**161**	191	221	251	281	311	341	**371**	401	431 ...
13	43	73	103	**133**	163	193	223	253	283	313	**343**	373	403	433 ...
17	47	**77**	107	137	167	197	227	257	**287**	317	347	377	407	437 ...
19	**49**	79	109	139	169	199	229	**259**	289	319	349	379	409	439 ...
23	53	83	113	143	173	**203**	233	263	293	323	353	383	**413**	443 ...
29	59	89	**119**	149	179	209	239	269	299	**329**	359	389	419	449 ...

There are 3 two-twins and 3 four-twins in any pattern of A_4 and 21 of each in any extended pattern. The 21 two-twins occupy 6 rows and 6 are eliminated when the multiples of 7 are removed in the step to A_5. The 21 four-twins also occupy 6 rows and 6 are eliminated in the step to A_4. In some cases a multiple of 7 simultaneously eliminates both a two-~~twin~~ and a four-twin, but there are still 6 of each eliminated in the step to A_5. 15 of each will remain in the step to A_5 where they will be in the fundamental pattern.

Theorem 4 In any pattern, the number of two-twins is equal to the number of four-twins.

Proof. Basis step. Use the example above where A_4 has 21 two-twin and 21 four-twins in each extended pattern.

Induction step. There are τ_n two-twins and τ_n four-twins in any pattern of A_n and $p_n\tau_n$ of each in any extended pattern. $2\tau_n$ of each will be eliminated in each extended pattern in the step to A_{n+1}. This leaves $\tau_n(p_n - 2)$ of each in the patterns of A_{n+1}. $\square$

We show A_5 with a count of the twins added.

A_5 $\quad$ $p_5 = 11,\ \lambda_5 = 210,\ \nu_5 = 48,\ \tau_5 = 15$

2-twin													4-twin
15	1	211	421	631	841	1051	1261	1471	1681	1891	**2101**	2311 ...	
1	**11**	221	431	641	851	1061	1271	1481	1691	1901	2111	**2321** ...	
1	13	223	433	643	853	1063	1273	1483	1693	**19.5**	2113	2323 ...	1
2	17	227	437	647	857	**1067**	1277	1487	1697	1907	2117	2327 ...	1
2	19	229	439	**649**	859	1069	1279	1489	1699	1909	2119	2329 ...	2
	23	233	443	653	863	1073	1283	1493	1703	1913	**2123**	2333 ...	2
3	29	239	449	659	**869**	1079	1289	1499	1709	1919	2129	2339 ...	
3	31	241	**451**	661	871	1081	1291	1501	1711	1921	2131	2341 ...	
	37	247	457	667	877	1087	1297	**1507**	1717	1927	2137	2347 ...	3
4	41	251	461	**671**	881	1091	1301	1511	1721	1931	2141	2351 ...	3
4	43	**253**	463	673	883	1093	1303	1513	1723	1933	2143	2353 ...	4
	47	257	467	677	887	1097	1307	1517	**1727**	1937	2147	2357 ...	4
	53	263	**473**	683	893	1103	1313	1523	1733	1943	2153	2363 ...	
5	59	269	479	689	899	1109	1319	**1529**	1739	1949	2159	2369 ...	
5	61	271	481	691	901	**1111**	1321	1531	1741	1951	2161	2371 ...	
	67	277	487	697	907	1117	1327	1537	1747	1957	**2167**	2377 ...	5
6	71	281	491	701	911	1121	**1331**	1541	1751	1961	2171	2381 ...	5
6	73	283	493	703	**913**	1123	1333	1543	1753	1963	2173	2383 ...	
	79	289	499	709	919	1129	1339	1549	1759	**1969**	2179	2389 ...	6
	83	293	503	713	923	**1133**	1343	1553	1763	1973	2183	2393 ...	6
	89	299	509	719	929	1139	1349	1559	1769	1979	**2189**	2399 ...	
	97	307	**517**	727	937	1147	1357	1567	1777	1987	2197	2407 ...	7
7	101	311	521	731	941	1151	1361	1571	1781	**1991**	2201	2411 ...	7
7	103	313	523	733	943	1153	1363	**1573**	1783	1993	2203	2413 ...	8
8	107	317	527	**737**	947	1157	1367	1577	1787	1997	2207	2417 ...	8
8	109	**319**	529	739	949	1159	1369	1579	1789	1999	2209	2419 ...	9
	113	323	533	743	953	1163	1373	1583	**1793**	2003	2213	2423 ...	9
	121	331	541	751	961	1171	1381	1591	1801	2011	2221	**2431** ...	
	127	337	547	757	967	**1177**	1387	1597	1807	2017	2227	2437 ...	10
	131	**341**	551	761	971	1181	1391	1601	1811	2021	2231	2441 ...	10
9	137	347	557	767	977	1187	**1397**	1607	1817	2027	2237	2447 t ...	
9	139	349	559	769	**979**	1189	1399	1609	1819	2029	2239	2449 ...	11
	143	353	563	773	983	1193	1403	1613	1823	2033	2243	**2453** ...	11
10	149	359	569	779	989	**1199**	1409	1619	1829	2039	2249	2459 ...	
10	151	361	571	**781**	991	1201	1411	1621	1831	2041	2251	2461 ...	
	157	367	577	787	997	1207	1417	1627	**1837**	2047	2257	2467 ...	
	163	373	**583**	793	1003	1213	1423	1633	1843	2053	2263	2473 ...	12
11	167	377	587	797	1007	1217	1427	1637	1847	**2057**	2267	2477 ...	12
11	169	379	589	799	1009	1219	1429	**1639**	1849	2059	2269	2479 ...	13
	173	383	593	**803**	1013	1223	1433	1643	1853	2063	2273	2483 ...	13
12	179	389	599	809	1019	1229	1439	1649	**1859**	2069	2279	2489 ...	
12	181	391	601	811	1021	1231	**1441**	1651	1861	2071	2281	2491 ...	
	187	397	607	817	1027	1237	1447	1657	1867	2077	2287	**2497** ...	14
13	191	401	611	821	1031	1241	1451	**1661**	1871	2081	2291	2501 ...	14
13	193	403	613	823	1033	**1243**	1453	1663	1873	2083	2293	2503 ...	15
14	197	**407**	617	827	1037	1347	1457	1667	1877	2087	2297	2507 ...	15
14	199	409	619	829	1039	1249	1459	1669	1879	2089	**2299**	2509 ...	
15	**209**	419	629	839	1049	1259	1469	1679	1889	2099	2309	**2519** ...	

Here are the parameters for the patterns for the first 10 sets.

n	p_n	p_n^2	λ_n	ν_n	τ_n
1	2	4	1	1	-
2	3	9	2	1	-
3	5	25	6	2	1
4	7	49	30	8	3
5	11	121	210	48	15
6	13	169	2310	480	135
7	17	289	30030	5760	1485
8	19	361	510510	92160	22275
9	23	529	9699690	1658880	378675
10	29	841	223092870	36495360	7952175

Formulas

Here we give formulas for calculating λ, ν, and τ.

The length of a pattern in A_n is: $$\lambda_n = \prod_{i=1}^{n-1} p_i, \quad n \geq 2 \tag{2}$$

$$\lambda_{n+1} = \lambda_n p_n$$

The number of members in a pattern in A_n is: $$\nu_n = \prod_{i=1}^{n-1} (p_i - 1), \quad n \geq 2 \tag{3}$$

$$\nu_{n+1} = \nu_n (p_n - 1)$$

The number of twins in a pattern in A_n is: $$\tau_n = \prod_{i=2}^{n-1} (p_i - 2), \quad n \geq 3 \tag{4}$$

$$\tau_{n+1} = \tau_n (p_n - 2)$$

~~Clearly, there~~ is an infinite number of twins in any set greater than A_2. Our goal is to show that among these there is an ~~infinte~~ number of prime twins.

In what follows, we refer to the multiples of p_n. In the fundamental extended pattern, the ν_n multiples of p_n result from multiplying each member of the ~~fundamntal~~ pattern by p_n. Also, recall that p_n^2 is the first composite in A_n, p_{n+1}^2 is the first composite in A_{n+1}, etc. We will be interested in the distances between the first composites as we step through the sets in the proof below.

We speak of a twin as being a multiple of p_n. By this we mean that one of the members of the twin is a multiple of p_n.

On the Distribution of Twins in a Set

We will cover several features of the distribution of the twins. But first we want to comment on extended patterns. We are concerned about possible confusion with our ~~terminolgy~~. In a set, A_n, an extended pattern consists of p_n patterns and ν_n rows. The patterns are vertical and the rows are horizontal. Refer to the listings, above, of the first five sets in the section on Eratosthenes Patterns. We will speak of a 'row of twins' which will ~~actually~~ be two rows of members of the set. When we want to speak of an actual single row we will use 'row of members' or just 'row'.

The first feature is the average distance between the twins in a set.

The average distance between twins in a pattern of A_n is λ_n/τ_n. The number of twins in an extended pattern of A_n is $\tau_n p_n$ and the average distance between twins in ~~an~~ extended pattern of A_n is $\lambda_n p_n/(\tau_n p_n)$ also equal to λ_n/τ_n.

λ_n/τ_n increases at a ~~decreaseing~~ rate. $\frac{\lambda_{n+1}}{\tau_{n+1}} = \frac{p_n}{(p_{n-2})}(\frac{\lambda_n}{\tau_n})$.
As n approaches infinity $\frac{\lambda_{n+1}}{\tau_{n+1}}$ approaches $\frac{\lambda_n}{\tau_n}$.
The inverse, τ_n/λ_n, is the average density in patterns and in extended patterns.

The twins occupy $2\tau_n$ rows of an extended pattern and there is exactly one multiple of p_n in each row. 2 twins are eliminated from each row of twins and $2\tau_n$ twins are eliminated from each extended pattern in the step to A_{n+1}. Also, $\tau_n(p_n - 2)$ twins in each extended pattern of A_n are not eliminated, giving an average distance of ~~of~~ $\lambda_n p_n/(\tau_n(p_n - 2))$ in each pattern of A_{n+1}. Now, since the distance between twins in the patterns of A_{n+1} is λ_{n+1}/τ_{n+1} we have: $\lambda_{n+1}/\tau_{n+1} = \frac{p_n}{(p_n - 2)}\ (\lambda_n/\tau_n)$. As n approaches infinity, λ_{n+1}/τ_{n+1} approaches λ_n/τ_n, i.e., for large n, when stepping from one set to the next, the change in distance between twins is ~~neglilgible~~.

In step from A_n to A_{n+1}, the fraction of twins removed from an extended pattern is $2\tau_n/(p_n\tau_n) = 2/p_n$. This approaches zero as n approaches infinity.

We are also interested in the distance between first composites. In A_n the first composite is p_n^2, in A_{n+1}^2 the first composite is p_{n+1}^2, etc. As we step through the sets, we step from one first composite to the next and note the distances between them, e.g., $p_{n+1}^2 - p_n^2$.

Recall from above that g_n is the distance between primes: $g_n = p_{n+1} - p_n$

$$p_{n+1}^2 = (p_n + g_n)^2 = p_n^2 + 2p_n g_n + g_n^2 \ , \quad p_{n+1}^2 - p_n^2 = 2p_n g_n + g_n^2 \tag{5}$$

We see that the distance between first composites is $2p_n g_n + g_n^2$ and, on average, this increases at an increasing rate. As n approaches ~~infinity~~, the distance between first composites approaches infinity. The distance between first composites is critical in the proof below.

The next feature in our discussion of the distribution of twins is a parameter called 'Vulnerable Twin'. These are the twins that are eliminated in the step from one set to the next. There are $2\tau_n$ vulnerable twins in an extended pattern. We now have a name for the twins that are eliminated as we discussed above. We are interested in the distance between vulnerable twins. It is $p_n\lambda_n/(2\tau_n)$. As shown above, λ_n/τ_n increases at a ~~decreaseing~~ rate. Therefore, the p_n

term dominates and the distance between vulnerable twins approaches infinity as n approaches infinity.

Since the average distance between twins is λ_n/τ_n the average number of twins that fall between two consecutive vulnerable twins is $\frac{\lambda_n p_n}{2\tau_n} \Big/ \frac{\lambda_n}{\tau_n} = \frac{p_n}{2}$. The twins that fall between two consecutive vulnerable twins are all multiples of primes greater than p_n. The number of twins that fall between two consecutive vulnerable twins approaches infinity as n approaches infinity.

The next feature is a parameter we call constellations of twins.[11]

A constellation of twins is a sequence of members of a set that consists only of two-twins and four-twins. Any two consecutive members must be either a two~~-twin~~ or a four-twin. We will look at two types, the pattern boundary two-twins and the pattern center hextuples. As we step through the sets, the distribution of these constellations becomes more uniform.

We give examples of ~~the~~ these two constellations.

We start with the pattern boundary two-twins. This constellation consists of a single two-twin. They are created in A_3, but we find it easier to study them in A_4 where the boundaries of the patterns are the two-twins, 30, 60, 90, 120, 150, ... For example, the twin, 60, is the pattern boundary two~~-twin~~ between the 2nd and 3rd pattern. 59 is in the 8th row of the extended pattern and 61 is in the 1st row. Neither is a multiple of 7 and 60 will not be eliminated in the step to A_5.

In the 1st row, 91 is the one multiple of 7 and in the 8th row 119 is the one multiple of 7. In the step to A_5, the pattern boundary two-twins that are eliminated are 90 and 120. None of the other 5 pattern boundary two-twins in the fundamental extended pattern is eliminated.

In the 8th row, 7th pattern, (the upper boundary of the fundamental extended pattern) there's a special situation. Neither of the members of 210 (a multiple of 7) is a multiple of 7, preventing it from being eliminated. The upper boundary of every extended pattern in A_4 is a multiple of 7 and this boundary is a two twin that will not be eliminated in the step to A_5. This guarantees that there will be pattern boundary two-twins in each pattern of A_5.

In general, all but 2 of the pattern boundary two-twins in each extended pattern of A_n will remain intact in the step to A_{n+1}. In A_{n+1} there are p_n subpatterns and all but two of them will contain pattern boundary two-twins. As n approaches infinity, the two subpatterns that are missing pattern boundary two-twins become negligible. The distribution of pattern boundary two-twins approaches uniformity, i.e. constant spacing throughout the sets.

Next we look at the pattern center ~~hextuples~~. A hextuple is a constellation of numbers consisting of 3 four-twins and 2 two-twins with the following form. Let a be the center of a hextuple. a is not a member of the set. The members are $a \pm 2, \pm 4, \pm 8$.

For example, in A_4, 15 is at the center of the fundamental pattern and $15 = \lambda_4/2$. Also 15 is a

[11]references 11 and 12. Wolfram gives an ~~exstensive~~ list of constellations and a long list of references. We do not always agree with other researchers' definitions. We suggest that the use of ~~Eratosthene~~ Patterns makes ~~is~~ easier to see the construction of ~~consellations~~

four-twin. The members of this four-twin are 13 and 17. Next are two 2-twins, 12 and 18 the members of which are generated by 15 ± 2 and 15 ± 4. Next are 2 four-twins, 9 and 21, the members of which are generated by 15 ± 4 and 15 ± 8.

07 Here is the first ~~hextuple~~ in A_4
11
13 Refer to the listing of A_4 above. Note the appearance
17 of ~~hextuples~~ at the center of all of the patterns of the set.
19
23 the two-twins are 12 and 18, the four-twins are 9, 15, and 21

There are 7 patterns in each ~~exended~~ pattern of A_4, each with a pattern center hextuple. In 6 of the patterns one member of the hextuple will have a multiple of 7 and these 6 will be eliminated in the step to A_5. The other hextuple will remain intact and symmetry demands that it will be the center pattern. This leads to a hextuple in the center subpattern of each of the 11 patterns in an extended pattern of A_5. There will also be remnants of hextuples in the other 6 subpatterns. These remnants will have at least one two-~~twin~~ and at least two four-twins. Thus, there are twins in each subpattern.

In the step ~~to~~ A_6, 6 of the ~~hextuples~~ in the 6 center rows of an extended pattern will be eliminated and 5 will remain intact. These 5 will be passed to the fundamental pattern of A_6 and symmetry requires that one is in the center subpattern. The other 4 will be symmetrically arranged about ~~he~~ center ~~subpatterm~~. The other subpatterns will each have multiple twins.

This process is continued as we step through the sets. Consider the hextuple at the center of the ~~fundamental~~ extended pattern of A_n. $\lambda p_n/2$ will be at the center of the hextuple and it is a multiple of p_n. The next nearest multiples of p_n are $\lambda p_n/2 \pm 2p_n$, both of which are outside of the hextuple. Thus, none of the members of this hextuple is divisible by p_n and it will be passed to the center of the fundmental pattern of A_{n+1}.

For any $n \geq 4$ all but six of the pattern center hextuples in A_n will remain intact in the step to A_{n+1}. In A_{n+1} there are p_n subpatterns and all but six of them will contain pattern center hextuples. For those six incomplete ~~hextuples~~ each will have at least one two-twin and at least two four-twins. As n approaches infinity, the distribution of the pattern center hextuples in the subpatterns approaches uniformity, i.e., constant spacing throughout the sets.

We are concerned that our use of the phrase "eliminate a constellation" is not ~~self~~ explanatory. When one of the members of a ~~hextuple~~ is removed there are still five members remaining with at least one two-~~twin~~ and two four-twins. Five members ~~does~~ not constitute a hextuple and the original hextuple no longer exists, however a fraction of it remains.

We briefly mention three other constellations of twins each of which is a subset of the ~~hextuples~~. They are the triples, quadruples, and quintuples. Here are examples from the first hextuple in A_4 given above.

Triples 7 11 13, 11 13 17, 13 17 19, 17 19 23

Quadruples 7 11 13 17, 11 13 17 19, 13 17 19 23

Quintuples 7 11 13 17 19, 11 13 17 19 23

These are formed when, in the elimination process, one of the members of a hextuple is removed. There are several ways that these constellations can be formed. For example, removing the fourth member of a quintuple in A_n creates a triple in A_{n+1}. We work within the confines of a fundamental extended pattern, ~~obvserving~~ the removal of ~~mutltiples~~ of p_n in rows of twins which are members of a constellation. The new smaller constellations ~~which~~ are created will reside in the fundamental pattern of A_{n+1}. In that fundamental pattern these constellations will be at the head of rows of constellations. Let r be the size of the constellation. In those rows of ~~constelltions~~ all but r of them will remain intact in the step to the next set. As in the examples above this will result in all but r of the ~~subpastterns~~ of A_{n+2} will have intact instances of the ~~constellatioin~~ at hand. As n approaches ~~infinituy~~ these constellations will approach a uniform distribution.

We have shown that the distribution of pattern ~~bounary~~ two-twins, pattern center ~~hextuples~~, and the three other constellations approach uniformity as we step through the sets. The distribution of the twins is the same in every pattern of a set. We make use of this nearly uniform distribution in the proof that follows.

Note: As we step through sets, the ~~elimnation~~ of constellations creates an increasing number of single twins that are not members of any constellation. They appear to be randomly distributed. Thus we have a random distribution of single twins superimposed with near uniform distributions of constellations of twins.

The Proof of the Twin Prime Conjecture

We will show that the number of twins in the range, p_n to p_n^2 in A_n, approaches infinity as n approaches infinity. As shown above, all twins in this range are prime.

In A_n there are τ_n twins in the fundamental pattern and, as shown above, there are near uniform distributions of some these twins. Also, we showed that there are multiple twins in each subpattern.

Of primary importance in this proof is the extreme difference between the distance between first composites and the distance between twins.

Above, we found that the distance between first composites is $2p_ng_n + g_n^2$ and this approaches infinity as n approaches infinity.

The average distance between twins is λ_n/τ_n. This distance increases at a ~~decreaseing~~ rate and for large n the increase is negligible.

The average number of twins that fall between two consecutive first composites is proportional to the distance between them and this number approaches infinity as n approaches infinity. The fraction of these twins that is removed in the step to the next set is $2/p_n$. This approaches zero as n approaches infinity. This implies that, for large n, the change in density of the twins is negligible.

From the above discussion of the vulnerable twins, we learned that the distance between vulnerable twins ~~approaches~~ infinity as n approaches infinity. We also learned that the number of twins between consecutive vulnerable twins approaches infinity and ~~that~~ they are all multiples of primes greater than p_n.

In the step from A_n to A_{n+1}, p_n is the only member removed from the range, 1 to p_n^2, and only multiples of p_n are removed from the range, p_n^2 to p_{n+1}^2. These two ranges merge to form the range, 1 to p_{n+1}^2 in A_{n+1}.

We see that in each set, A_r, the range, 1 to p_r^2, continues to grow in size with each step.

Given the above, as we step through the sets, we find the following. The distance between consecutive first composites increases. The number of twins between consecutive first composites increases. The distance between vulnerable twins increases. The number of twins between vulnerable twins increases and all of these twins are multiples of primes greater than p_n. The fraction of twins that ~~is~~ eliminated from a set in the step to the next set approaches zero.

~~Clearly the~~ number of twins between 1 and p_n^2 approaches infinity as n approaches infinity and these are all prime.

References

1. T. Apostol *Introduction to Analytic Number Theory*, Springer-Verlag, 1976

2. J. De Koninck and F. Luca *Analytic Number Theory*
American Mathematical Society, 2012

3. G. Harman *Prime-Detecting Sieves*, Princeton University Press, 2007

4. D. Koukoulopoulos, *The Distribution of Prime Numbers*,
American Mathematical Society, 2019

5. I. Niven and H.S. Zuckerman
An Introduction to the Theory of Numbers, 4th Ed., Wiley, 1980

Comment on $\pi(n)$. References 6 and 7 were written almost simultaneously by different researchers who were not in contact with each other. Both used analytical methods.
References 8 and 9 were similar discoveries and used elementary methods.
Reference 10 is a more recent use of analytic methods.

6. J. Hadamard *Sur la distribution des zeros de la fonction $\zeta(s)$ et ses consequences arithmetiques*
Bull. Soc. Math. France, 24, pp. 199-220, 1896

7. Ch.-J. de la vallee poussin
La Fonction $\zeta(s)$ de Riemann et les Nombres Premiers en General
Annales de la Societe Scienifique de Bruxelles, 20, pp. 185-256, 1896

8. P. Erdos *On a New Method in Elementary Number Theory Which Leads to an Elementary Proof of the Prime Number Theory*
Proceedings of the National Academy of Sciences, Vol 35, No. 7, pp. 374-384, July 1949

9. A. Selberg *An Elementary Proof of the Prime-Number Theorem*
The Annals of Mathematics, Vol. 50, N0. 2, pp. 303-313, April, 1949

10. D. Newman *Simple Analytic Proof of the Prime Number Theorem*
American Mathematical Monthly 87, 693 - 696, (Nov., 1980)

11. https://mathworld.wolfram.com/PrimeConstellation.html

12. https://archive.lib.msu.edu/crcmath/math/math/p/p573.htm