PRIME PYTHAGOREAN TRIANGLES

HARVEY DUBNER AND TONY FORBES

Abstract. A prime Pythagorean triangle has three integer sides of which the hypotenuse and one leg are primes. In this article we investigate their properties and distribution. We are also interested in finding chains of such triangles, where the hypotenuse of one triangle is the leg of the next in the sequence. We exhibit a chain of seven prime Pythagorean triangles and we include a discussion of a general primality proof to deal with the larger elements (up to 2310 digits) of the associated set of eight primes.

1. Introduction

While investigating the distribution of special forms of primes, the first author accidently came across a conjecture about Pythagorean triangles (right triangles with integral sides). The conjecture, based on the famous Conjecture (H) of Siepiński and Schinzel, states that there is an infinite number of Pythagorean triangles which have a leg and hypotenuse both prime [9, page 408].

Pythagorean triangles have been the subject of much recreational material [1] as well as the basis of some of the most important and fundamental topics in number theory. However, we could not find any significant references to such 2-prime Pythagorean triangles, and hoping that we had found a new topic to study we enthusiastically:

1. started developing appropriate theory and computer programs;
2. started searching for large 2-prime triangles;
3. started searching for sequences of 2-prime triangles where the hypotenuse of the previous triangle becomes the leg of the next one.

The largest 2-prime Pythagorean triangle that was found had a leg of 5357 digits and an hypotenuse of 10713 digits. It soon became apparent that finding sequences of triangles was exceptionally interesting and challenging. Eventually a sequence of seven triangles was found. More significant than the seven triangles is the improvement by the second author of the general method, APRCL, for primality proving so that the seventh hypotenuse of 2310 digits could be proved prime.

2. Theory

A 2-prime Pythagorean triangle, $A^2 + B^2 = C^2$, must be primitive so that

$$A = u^2 - v^2, \qquad B = 2uv, \qquad C = u^2 + v^2,$$

Date: January 20, 2000.

1991 *Mathematics Subject Classification.* Primary 11A41.

Key words and phrases. prime numbers, primality proving .

with $\gcd(u,v)=1$, and u,v of different parity. Since $A=(u+v)(u-v)$, for A to be prime it is necessary that $(u-v)=1$ so that

$$A=2v+1, \qquad B=2v^2+2v, \qquad C=2v^2+2v+1.$$

Thus,

$$C=\frac{A^2+1}{2}. \tag{2.1}$$

Note that the even leg is only one less than the hypotenuse. The triangles get quite thin as A increases.

To find 2-prime Pythagorean triangles it is necessary to find pairs of primes A,C that satisfy the above equation. Table 1 lists the smallest 2-prime Pythagorian triangles.

Table 1. Pythagorean triangles with 2 prime sides

rank	prime leg	even leg	hypotenuse
1	3	4	5
2	5	12	13
3	11	60	61
4	19	180	181
5	29	420	421
6	59	1740	1741
7	61	1860	1861
8	71	2520	2521
9	79	3120	3121
10	101	5100	5101
100	4289	9197760	9197761
1000	91621	4197203820	4197203821

Small triangles are easy to find by a simple search, but finding large triangles with thousands of digits is complicated by the difficulty of proving true primality of the hypotenuse, C. However, if $(C-1)$ has many factors then it is easy to prove primality using [2], assuming that the factored part of $(C-1)$ exceeds $\sqrt[3]{C}$. Since

$$C-1=\frac{A^2+1}{2}-1=(A^2-1)/2=(A-1)(A+1)/2, \tag{2.2}$$

by picking an appropriate form for A, then $(A-1)$ can be completely factored so that $(C-1)$ will be about 50% factored.

Using the form, $A=k\cdot 10^n+1$, a computer search of a few days gave the following large triangle:

$$A=491140\cdot 10^{1300}+1, \quad 1306 \text{ digits}, \qquad C=2612 \text{ digits}.$$

A few days after this result was posted to the NMBRTHRY list we received a message from Iago Camboa announcing a much larger triangle:

$$A=1491\cdot 2^{17783}+1, \quad 5357 \text{ digits}, \qquad C=10713 \text{ digits}.$$

He cleverly used a previously computed list of primes as a source for A thus eliminating the large amount of time required to find the first prime.

3. 2-prime Pythagorean triangle sequences

It is possible to find a series of primes, $P_0, P_1, P_2, ...P_k, ...P_n$ such that

$$P_{k+1} = \frac{P_k^2 + 1}{2}. \tag{3.1}$$

This represents a sequence of n 2-prime triangles where P_k is the hypotenuse of the k'th triangle and the leg of the $(k+1)$'th triangle. Each P has about twice the number of digits as the previous P. Table 2 is a list of the smallest sets of two sequential prime Pythagorean triangles.

Table 2. Two sequential prime Pythagorean triangles

	triangle 1			triangle 2		
1	3	4	5	5	12	13
2	11	60	61	61	1860	1861
3	19	180	181	181	16380	16381
4	59	1740	1741	1741	1515540	1515541
5	271	36720	36721	36721	674215920	674215921
6	349	60900	60901	60901	1854465900	1854465901
7	521	135720	135721	135721	9210094920	9210094921
8	929	431520	431521	431521	93105186720	93105186721

Table 3 is a list of the starting primes for the smallest prime Pythagorean sequences for two, three, four and five triangles. These were found by straight forward unsophisticated searching and took about 10 computer-days (Pentium/200), mostly for finding five triangles.

Finding the starting prime for the smallest prime sequence of six triangles took about 120 computer days.

$$P_0 \text{ for 6 triangles} = 2500282512131.$$

Next, we attempted to derive the number of n triangle sequences that could be expected. If the $(n+1)$ numbers that make up the n triangles were selected randomly but were of the proper size then the probability that P is the start of n triangles is

$$Q(P,n) = \prod_0^n \frac{1}{\log P_i} = \prod_0^n \frac{1}{(2^i)(\log P)} = \frac{1}{(2^{(n(n+1)/2)})(\log P)^{n+1}}. \tag{3.2}$$

However, there are correlations between the primes that affect the prime probabilities. It is easy to show from equation (2.1) that P_0 can only end in 1 or 9, which elininates half the possible P_0's, and assures that all subsequent potential primes cannot be divisible by 2, 3 or 5. Thus, the probability of each subsequent number being prime is increased by the factor $(2/1)(3/2)(5/4) = 3.75$. The probability that P is the start of n prime triangles now becomes,

$$Q(P,n) = \frac{0.5(3.75)^n}{(2^{(n(n+1)/2)})(\log P)^{n+1}}. \tag{3.3}$$

Table 3. Starting prime for smallest prime Pythagorean sequences

	2 triangles	3 triangles	4 triangles	5 triangles
1	3	271	169219	356498179
2	11	349	1370269	432448789
3	19	3001	5965699	5380300469
4	59	10099	15227879	10667785241
5	271	11719	17750981	11238777509
6	349	12281	19342559	12129977791
7	521	25889	21828601	23439934621
8	929	39901	24861761	28055887949
9	1031	46399	27379621	33990398249
10	1051	63659	34602049	34250028521
11	1171	169219	39844619	34418992099
12	2381	250361	48719711	34773959159
13	2671	264169	50049281	34821663421
14	2711	287629	51649019	36624331189
15	2719	289049	52187371	40410959231
16	3001	312581	52816609	43538725229
17	3499	353081	58026659	47426774869
18	3691	440681	73659239	48700811941
19	4349	473009	79782821	49177751131
20	4691	502501	86569771	59564407571

The expected number of prime triangles up to a given P_0 is

$$E(P_0,n)=\sum_{P=3}^{P_0} Q(P,N)=\frac{0.5(3.75)^n}{(2^{(n(n+1)/2)})}\sum_{P=3}^{P_0}\frac{1}{(\log P)^{n+1}}. \tag{3.4}$$

The last summation can be approximated by an integral, which after integrating by parts becomes,

$$R(P,n)=\frac{1}{n!}Li(P)-\frac{1}{n!}\frac{P}{\log P}-\cdots-\frac{1}{n(n-1)}\frac{P}{(\log P)^{(n-1)}}-\frac{1}{n}\frac{P}{(\log P)^n},$$

where $Li(P)$ is the logarithmic integral. Equation (3.4) now becomes

$$E(P_0,n)=\frac{0.5(3.75)^n}{2^{(n(n+1)/2)}}R(P_0,n)(1.3)^n\ . \tag{3.5}$$

Note the inclusion of a correction factor, $(1.3)^n$. As is discussed in the following section on sieving, there are other correlations between the primes which affect the expectation. These are difficult to derive theoretically so we determined it empirically. Table 4 compares the estimated and actual number of triangles found. The corrected estimate appears adequate to assist in estimating the search time for seven Prime Pythagorean triangles.

Next, we use equation (3.5) to estimate the smallest P_0 that will give seven triangles. The following table shows we can expect that P_0 for seven triangles will be about 6700 times larger than P_0 for six triangles. Using performance data from the search for six triangles, this means that the search for the smallest sequence of

TABLE 4. Estimated and actual number of Prime Pythagorean triangles

triangles n	P_0	actual	estimate	corrected estimate
1	130000	1302	1090	1420
2	1980000	1005	741	1252
3	10^8	953	469	1030
4	$18 \cdot 10^8$	205	53	151
5	$63 \cdot 10^9$	21	4	15
6	$28 \cdot 10^{12}$	1	0.14	0.7

seven Prime Pythagorean triangles could be expected to take about 200 computer-years!

n	P_0 for expectation=1	actual P_0
2	28	3
3	1,350	271
4	1,000,000	169,219
5	$1.5 \cdot 10^9$	$3.5 \cdot 10^8$
6	$4.0 \cdot 10^{12}$	$2.5 \cdot 10^{12}$
7	$2.7 \cdot 10^{16}$	

It was clear that the search for the smallest sequence of seven triangles as presently constituted was impractical. For every P_0 the search method included testing by division to see if each of the $(n+1)$ potential primes was free of small factors. The second author then proposed an efficient sieving method that limited the search to sequences that had a high probability of success. This made a search for seven triangles reasonable.

4. The Sieve

To find a set of seven Pythagorean triangles with the desired properties we need to search for a chain of eight primes, $P_0, P_1, \ldots, P_7$, linked by the condition $P_{i+1} = (P_i^2 + 1)/2$, $i = 0, 1, \ldots, 6$.

The purpose of the sieve is to eliminate from further consideration numbers P_0 for which either P_0 itself or one of the numbers P_i, $i = 1, 2, \ldots, 7$, is divisible by a small prime. Let q be an odd prime and suppose P is to be considered as a possible value of P_0. Clearly, we can reject P if $P \equiv 0 \pmod{q}$. Furthermore, we can reject P if P_1 is divisible by q, that is, if

$$P \equiv \sqrt{-1} \pmod{q},$$

on the assumption that $\left(\frac{-1}{q}\right) = 1$. Continuing in this way, we can reject P if P_2 is divisible by q, the equivalent condition on P being

$$P \equiv \sqrt{2\sqrt{-1} - 1} \pmod{q};$$

or if P_3 is divisible by q,

$$P \equiv \sqrt{2\sqrt{2\sqrt{-1} - 1} - 1} \pmod{q};$$

and so on, provided that the various square roots $\pmod q$ exist. In each case, where there is a square root $\pmod q$ there are two possible values and hence two extra residues $\pmod q$ that can be eliminated.

If $q \equiv 3 \pmod 4$, we do not proceed beyond the first stage, and for these primes we reject P only if $P \equiv 0 \pmod q$. On the other hand, if $q \equiv 1 \pmod 4$, we always have at least three forbidden residues, 0, $\sqrt{-1}$ and $q - \sqrt{-1} \pmod q$.

In general, we can compute the set $E(q)$ of forbidden residues $\pmod q$ for any odd prime q as follows. Start with $E_0(q) = \{0\}$. Given $E_i(q)$, let

$$E_{i+1} = \left\{ \pm\sqrt{2e-1} \pmod q : e \in E_i \text{ and } \left(\frac{2e-1}{q}\right) = 1 \right\}.$$

Then $E(q)$ is the union of $E_0(q)$, $E_1(q)$, $\ldots$, $E_7(q)$. In Table 5 we list $E(q)$ for the first few primes $q \equiv 1 \pmod 4$.

Now let

$$P = NQ + H,$$

where Q is the product of small primes and H is allowed to run through all the permitted residues $\pmod Q$. We sieve the numbers N. That is, we start with an interval $N_0 \le N < N_1$ and for each sieving prime q, $\gcd(q, Q) = 1$, we remove all those $N \in [N_0, N_1)$ for which $NQ + H$ is divisible by q.

We split Q into pairwise-coprime divisors m_0, m_1, $\ldots$, m_r. For each divisor m_j of Q, $j = 0, 1, \ldots, r$, we make a list of the permitted residues $\pmod{m_j}$; h is a permitted residue $\pmod{m_j}$ if h is not zero $\pmod{m_j}$ and if the function $h \to (h^2+1)/2 \pmod{m_j}$ does not produce zero $\pmod{m_j}$ during the first seven iterations. The permitted residues $H \pmod Q$ are constructed from permitted residues $h \pmod{m_j}$ using the Chinese Remainder Theorem. It works well if Q is the product of primes which have small percentages of permitted residues. From this perspective the best primes, in descending order of merit, turn out to be: 29 (34%), 5 (40%), 2 (50%), 17 (59%), 13 (62%), 3 (67%), 53 (68%), 101 (71%), 89 (74%) and 233 (77%).

For the actual search we chose the parameter $Q = 21342962305470$, with divisors $6630 = 2 \cdot 3 \cdot 5 \cdot 13 \cdot 17$, 29, 89, 101, 53, and 233. The number of values of $H \pmod Q$ turns out to be $320 \cdot 10 \cdot 66 \cdot 72 \cdot 36 \cdot 180 = 98537472000$, the indicated factors of this product being the numbers of permitted residues modulo the corresponding factors of Q. The last two divisors, 53 and 233, of Q were in fact used for parcelling out the work to various computers that were engaged in the search. Each computer was given a specific residue $\pmod{53 \cdot 233}$ and left to work through the 15206400 residues $\pmod{6630 \cdot 29 \cdot 89 \cdot 101}$.

The construction of the sieve and the method of computing $H \pmod Q$ were based on computer programs designed for finding prime k-tuplets and written by the second of the authors; see Forbes [6] for the details. We set up a table of sieving primes q together with pre-computed values of $-1/Q \pmod q$ as well as, for $q \equiv 1 \pmod 4$, $e/Q \pmod q$ for each pair e, $q - e$ in $E(q)$. We can then easily calculate the index of the first N to be removed from the sieve array for $P \equiv e \pmod q$, namely $e/Q - H/Q - N_0 \pmod q$.

The program also allows us to limit the size of primes $q \equiv 3 \pmod 4$ used by the sieve. One reason for doing so would be to give priority to primes $q \equiv 1 \pmod 4$; they have more residues for sieving and therefore one would expect them to be in some sense more efficient. In fact it was found by experiment that if P has about

Table 5. $E(q)$

q	$E(q)$
5	$\{0, 2, 3\}$
13	$\{0, 3, 5, 8, 10\}$
17	$\{0, 3, 4, 5, 12, 13, 14\}$
29	$\{0, 2, 3, 5, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 24, 26, 27\}$
37	$\{0, 6, 8, 14, 23, 29, 31\}$
41	$\{0, 9, 32\}$
53	$\{0, 4, 14, 16, 17, 18, 19, 22, 23, 30, 31, 34, 35, 36, 37, 39, 49\}$
61	$\{0, 11, 50\}$
73	$\{0, 23, 27, 46, 50\}$
89	$\{0, 9, 26, 27, 30, 34, 37, 38, 39, 40, 41, 44, 45, 48, 49, 50, 51, 52, 55, 59, 62, 63, 80\}$
97	$\{0, 7, 22, 25, 72, 75, 90\}$
101	$\{0, 7, 10, 12, 15, 16, 22, 23, 25, 26, 34, 35, 37, 38, 50, 51, 63, 64, 66, 67, 75, 76, 78, 79, 85, 86, 89, 91, 94\}$
109	$\{0, 33, 76\}$
113	$\{0, 2, 15, 46, 54, 59, 67, 98, 111\}$
137	$\{0, 22, 37, 100, 115\}$
149	$\{0, 44, 105\}$
157	$\{0, 10, 28, 31, 126, 129, 147\}$
173	$\{0, 32, 80, 93, 141\}$
181	$\{0, 2, 9, 19, 30, 33, 41, 47, 54, 56, 64, 78, 80, 88, 93, 101, 103, 117, 125, 127, 134, 140, 148, 151, 162, 172, 179\}$
193	$\{0, 57, 81, 112, 136\}$
197	$\{0, 14, 37, 94, 103, 160, 183\}$
229	$\{0, 18, 19, 30, 48, 54, 59, 69, 91, 107, 110, 119, 122, 138, 160, 170, 175, 181, 199, 210, 211\}$
233	$\{0, 3, 5, 7, 12, 13, 16, 21, 25, 27, 30, 42, 43, 44, 48, 52, 53, 55, 61, 67, 71, 80, 85, 89, 101, 104, 115, 118, 129, 132, 144, 148, 153, 162, 166, 172, 178, 180, 181, 185, 189, 190, 191, 203, 206, 208, 212, 217, 220, 221, 226, 228, 230\}$
241	$\{0, 64, 177\}$
257	$\{0, 16, 51, 206, 241\}$
269	$\{0, 82, 187\}$
277	$\{0, 8, 52, 60, 106, 171, 217, 225, 269\}$
281	$\{0, 53, 228\}$
293	$\{0, 4, 121, 138, 155, 172, 289\}$
313	$\{0, 7, 21, 25, 92, 221, 288, 292, 306\}$
317	$\{0, 17, 23, 24, 31, 44, 50, 52, 56, 74, 97, 114, 115, 126, 130, 134, 141, 142, 145, 153, 164, 172, 175, 176, 183, 187, 191, 202, 203, 220, 243, 261, 265, 267, 273, 286, 293, 294, 300\}$
337	$\{0, 21, 31, 34, 50, 71, 73, 90, 110, 114, 116, 144, 148, 153, 157, 162, 175, 180, 184, 189, 193, 221, 223, 227, 247, 264, 266, 287, 303, 306, 316\}$

19 digits, a sieve limit $L_0 = 20000$ for $q \equiv 3 \pmod 4$ and 480000 for $q \equiv 1 \pmod 4$ was approximately optimal.

Further performance improvements are possible by limiting the influence of primes $q \equiv 1 \pmod 4$. For each P that survives the sieve we do a probable-primality test, $2^P \equiv 2 \pmod P$, on P as well as, if P turns out to be a probable-prime, the numbers that follow P in the chain, stopping as soon as a composite is found. The effort required to perform the probable-primality test increases by a factor of about eight as we move from P_i to P_{i+1}. Therefore it might be better if priority were given to sieving primes q and residues $e \pmod q$ which would eliminate composite numbers from the larger elements of the chain. (The numbers we are dealing with are usually too small to benefit from sophisticated techniques for squaring, such as the fast Fourier transform; the factor, 8, represents the proportional increase in the cost of performing the probable-primality test on x when $\log x$ doubles.)

For controlling the effect of primes $q \equiv 1 \pmod 4$ we provided a set of parameters $L_1, L_2, \ldots$. If $q \equiv 1 \pmod 4$ is a sieving prime and $e \in E_i(q)$ then we do not use residue $e \pmod q$ for sieving unless $q < L_i$. As a result of a certain amount of experimentation we found that the optimum sieving rate occurs with the limits set approximately as follows: $L_1 = 120000$, $L_2 = 240000$, $L_3 = 360000$, together with a limit $L_0 = 20000$ for primes $\equiv 3 \pmod 4$ and an overall sieve limit of 480000. From these values we can compute an expected survival rate of

$$\eta = \prod_{q \text{ prime}} \frac{q - \nu_q}{q} = \frac{1}{3770},$$

approximately, where ν_q is the number of residues $\pmod q$ used by the sieve.

During the development of the program it became apparent that the problem of optimizing the set of sieving primes q and the sets sieving residues $\pmod q$ is one of some difficulty.

5. Eight Primes

In September 1999 the search was successful and this chain of eight probable primes was found:

$$\begin{aligned}
P_0 &= 2185103796349763249 && \text{(19 digits)},\\
P_1 &= (P_0^2+1)/2 && \text{(37 digits)},\\
P_2 &= (P_1^2+1)/2 && \text{(73 digits)},\\
P_3 &= (P_2^2+1)/2 && \text{(145 digits)},\\
P_4 &= (P_3^2+1)/2 && \text{(289 digits)},\\
P_5 &= (P_4^2+1)/2 && \text{(579 digits)},\\
P_6 &= (P_5^2+1)/2 && \text{(1155 digits)},\\
P_7 &= (P_6^2+1)/2 && \text{(2310 digits)}.
\end{aligned}$$

The program, written in a combination of C / Assembler Language by the second author, was designed to run on standard IBM PCs. We employed about 15 such machines, with clock speeds ranging from 200 MHz to 400 MHz. The faster computers were sieving and testing numbers at rates of about ten billion per hour.

During the search we also found 174 additional chains of seven (probable) primes.

6. Primality Proofs

The first five numbers, P_0, P_1, P_2, P_3 and P_4 (as well as P_{278} and other small primes mentioned in the discussion that follows) are easily verified, for example, by the UBASIC program APRT-CLE. For the others, we attempt to factorize

$$P_i - 1 = (P_0 - 1) \prod_{j=0}^{i-1} \frac{1}{2} (P_j + 1) ,$$

$i = 5, 6, 7$. We have

$$\begin{aligned}
P_0 - 1 &= 2^4 \cdot 233 \cdot 586132992583091, \\
(P_0 + 1)/2 &= 3^2 \cdot 5^3 \cdot 13 \cdot 761 \cdot 19087 \cdot 5143087, \\
(P_1 + 1)/2 &= 7^2 \cdot 1063 \cdot 189043 \cdot 7552723 \cdot 113558719 \cdot 141341652553, \\
(P_2 + 1)/2 &= 7058053 \cdot 5848063479673576700713235221 \\
&\quad \cdot 34520041584369005634844907730019249777, \\
(P_3 + 1)/2 &= 2179 \cdot 1847645923 \cdot C_{132}, \\
(P_4 + 1)/2 &= 307 \cdot 769 \cdot 262513 \cdot P_{278}, \\
(P_5 + 1)/2 &= 108139 \cdot 11360649709 \cdot 5586562264501 \cdot C_{550}, \\
(P_6 + 1)/2 &= 4177 \cdot 1372052449 \cdot 5098721569 \cdot 84098816095916212867 \cdot C_{1113},
\end{aligned}$$

where C_{132}, C_{550} and C_{1113} are composite numbers of 132, 550 and 1113 digits, respectively, and P_{278} is a 278-digit prime:

$$\begin{aligned}
P_{278} = \ & 66505518540598996114987486506055236521044267373138 \\
& 69473288000457727001877127498646545001634613677898 \\
& 53932112480508999228232340454335875401889420451888 \\
& 17780482079524485531037464472393979852934170207932 \\
& 02663155485302406204947222346461607409301255277393 \\
& 4788467292248055697961196019.
\end{aligned}$$

The 28-digit factor of P_2+1 and the 20-digit factor of P_6+1 were found by Manfred Toplic and Paul Zimmermann.

Since we have a partial factorization (77%) of P_5-1 we can establish the primality of P_5 by the methods of Brillhart, Lehmer and Selfridge [2]. Similarly, a 41% factorization of $P_6 - 1$ enables us to construct a primality proof for P_6.

It remains to deal with P_7.

We do not have enough prime factors of P_7-1 for a simple primality proof based on [2]; so we use a combination of methods. Suppose d is a non-trivial prime factor of P_7. The proof that no such d exists proceeds in several stages.

Pocklington's Theorem. Gathering together the factors listed above, let

$$\begin{aligned} F_1 \quad = \quad & 11364028773118678645863393880225035110068188490680 \\ & 74284625807644534721210969640169863192044176288720 \\ & 57382836214336492569310719940321645143241641366672 \\ & 31704620613678520580684280352992373327229897947340 \\ & 09917692032743575475918022578947700337216860293874 \\ & 96561498464943981086970289943873321681460108830000 \\ & 00131801406514260770804840415255291401064877989705 \\ & 76202962420323563098312300324091122817224414751412 \\ & 15123765209184430598589590008879997663256918503367 \\ & 07250451432160496252649191808276871593840887080642 \\ & 91103468534974000 \text{ (517 digits).} \end{aligned}$$

Then F_1 divides $P_7 - 1$ and F_1 can be compeletly factorized into primes. After confirming that the conditions of Pocklington's Theorem [8] hold, we have

$$d \equiv 1 \pmod{F_1}. \tag{6.1}$$

Morrison's Theorem. Let $F_2 = 43^2 \cdot 73 = 134977$. Then F_2 divides $P_7 + 1$. After confirming that the conditions of Morrison's Theorem [2, Theorem 16] hold (with F_2 replaced by $2F_2$), we have

$$d \equiv \pm 1 \pmod{F_2}. \tag{6.2}$$

The APRCL test. We confirm that the appropriate conditions for the APRCL test (see, for example, Cohen and Lenstra [5]) are satisfied with the following collection of prime powers p^k: $\{2^5, 3^3, 5^2, 7, 11, 13\}$, and primes q: {11, 17, 19, 23, 29, 31, 37, 41, 53, 61, 67, 71, 79, 89, 97, 101, 109, 113, 127, 131, 151, 157, 181, 199, 211, 241, 271, 281, 313, 331, 337, 353, 379, 397, 401, 421, 433, 463, 521, 541, 547, 601, 617, 631, 661, 673, 701, 757, 859, 881, 911, 937, 991, 1009, 1051, 1093, 1171, 1201, 1249, 1301, 1321, 1801, 1873, 1951, 2003, 2017, 2081, 2161, 2311, 2341, 2377, 2521, 2731, 2801, 2861, 2971, 3121, 3169, 3301, 3361, 3433, 3511, 3697, 3851, 4159, 4201, 4621, 4951, 5281, 5851, 6007, 6301, 6553, 7151, 7393, 7561, 7723, 8009, 8191, 8317, 8581, 8737, 9241, 9829, 9901, 11551, 11701, 12601, 13729, 14561, 14851, 15121, 15401, 15601, 16381, 16633, 17551, 18481, 19801, 20021, 20593, 21601, 21841, 23761, 24571, 25741, 26209, 28081, 30241, 34651, 36037, 38611, 39313, 42901, 47521, 48049, 50051, 51481, 54601, 55441, 65521, 66529, 70201, 72073, 79201, 81901, 92401, 93601, 96097, 103951, 108109, 109201, 110881, 118801, 120121, 123553, 131041, 140401, 150151, 151201, 180181, 193051, 196561, 200201, 216217, 218401, 257401, 270271, 300301, 332641, 393121, 415801, 432433, 450451}. The actual implementation follows steps (h) to (k) of the algorithm described in Cohen and Lenstra [4].

Let T be the product of the p^ks and let S be the product of the qs, multiplied by 11, the only q that divides T. Then $T = 21621600$ and

$$\begin{aligned} S \quad = \quad & 898080082225113217822223481720216007388686822330 24\\ & 44539522938364666709284347969136134439971184078953\\ & 43448926110251734618866677160265385666702172524877\\ & 14880656314884111327138509681893187202637968286542\\ & 83293015406063187900321990063760228444306508719210\\ & 87374318335464280164671112699177078381998108933220\\ & 08595569363494379320612699391931488477880423743996\\ & 04982905147776427399328285077729661659607325950022\\ & 44652659557331648190118182727751315532905547543602\\ & 46016541006719320534163395247117363557783380631924\\ & 67962869074581133724035035597332553162773278185621\\ & 96526723803946778487744027151656295879855418736649\\ & 746686392165044145832625043609441849 \text{ (636 digits).} \end{aligned}$$

The result of the APRCL test is that

$$d \equiv P_7^i \pmod{S} \text{ for some } i = 1, 2, \ldots, T-1. \tag{6.3}$$

The Chinese Remainder Theorem. Let $G = F_1F_2S$. Since F_1, F_2 and S are pairwise coprime, we can combine (6.1), (6.2) and (6.3) by the Chinese Remainder Theorem to obtain

$$\begin{aligned} d \quad \equiv \quad & \left(\frac{1}{F_2S} \bmod F_1\right) F_2S + \left(\frac{e}{F_1S} \bmod F_2\right) F_1S \\ & + \left(\frac{P_7^i}{F_1F_2} \bmod S\right) F_1F_2 \pmod{G} \end{aligned} \tag{6.4}$$

for some $e = \pm 1$ and $i = 1, 2, \ldots, T-1$.

Trial division. For $e = -1$ and 1, and for each i from 1 to $T-1$, we compute the value of d, $0 < d < G$, that satisfies (6.4) and confirm that it does not divide P_7. Hence if d is a prime factor of P_7, $d \geq G > \sqrt{P_7}$. Therefore P_7 is prime.

7. A General Primality Prover

Many IBM PC users are familiar with UBASIC [3], the large-integer arithmetic software, and the program APRT-CLE that comes as part of the package. APRT-CLE is an implementation of the primality proof described in Cohen and Lenstra [4] and has been a popular choice for verifying the primality of numbers which have no special structure.

However, APRT-CLE was unsuitable for the primality proof of P_7; it could not deal with numbers greater than about 10^{844} and there was no obvious mechanism for using information provided by the 517-digit partial factorization of $P_7 - 1$. Rather than write specific extensions to APRT-CLE for establishing the primality of P_7, we decided that a simple but powerful general-purpose primality prover for the IBM PC would be a very useful tool for other investigations. So we created a new computer program, VFYPR, which combines the APRCL test together with the theorems of Pocklington and Morrison.

Let N be given and suppose d is a prime divisor of N. If F_1 divides $N-1$, if F_1 is completely factorized into primes and if the conditions of Pocklington's Theorem hold, we have

$$d \equiv 1 \pmod{F_1}. \tag{7.1}$$

If F_2 divides $N+1$, if F_2 is completely factorized into primes and if the conditions of Morrison's Theorem hold,

$$d \equiv \pm 1 \pmod{F_2}. \tag{7.2}$$

Let $\{{p_1}^{k_1}, {p_2}^{k_2}, ...\}$ be a finite set of prime powers and let $T = {p_1}^{k_1}{p_2}^{k_2}...$ be their product. Let $\{q_1, q_2, ...\}$ be a set of primes that do not divide F_1F_2 and such that $\operatorname{lcm}(q_1-1, q_2-1, ...)$ divides T. Let $S_0 = q_1q_2...$ and let $S = S_0 \cdot \gcd(S_0, T)$. Suppose the conditions of the APRCL test hold. Then

$$d \equiv N^i \pmod{S} \text{ for some } i < T. \tag{7.3}$$

Let $G = F_1F_2S/2$. By (7.1), (7.2) and (7.3) and the Chinese Remainder Theorem, d must belong to one of at most $2T$ residue classes $\pmod{G}$. After it has verified the conditions leading to (7.1), (7.2) and (7.3) (for suitable F_1, F_2, S and T), VFYPR computes the representatives in $[1, G-1]$ of the possible residue classes $\pmod{G}$ and tries them one by one as possible non-trivial divisors of N. If there are none, we can conclude that $d > G$. If $G > \sqrt{N}$, N is prime. Otherwise, let F be the greater of F_1 and F_2. We now require $G > N^{1/3}$ and $mGF^2 > N$, where m is not too large. VFYPR then completes the primality proof by means of one of the following theorems (which are similar to Theorems 7 and 19, respectively, of [2].)

Theorem 7.1. *Suppose N is composite, suppose $N-1 = FR$, where F is even, R is odd, $R > 1$ and $\gcd(F, R) = 1$. Suppose also that for each prime factor f of F there is an integer a, $1 < a < N-1$, such that $a^{N-1} \equiv 1 \pmod{N}$ and $\gcd(a^{(N-1)/f} - 1, N) = 1$. If every prime factor of N is greater than G and $N^{1/3} < G < N^{1/2}$ then N is the product of two primes, $bF+1$ and $cF+1$. Furtheremore, if*

$$N < G(2AF^2 - G + rF + 2) \tag{7.4}$$

for some $A \geq 1$, where r and s are defined by $R = 2Fs + r$, $0 < r < 2F$, then $2(s-A) < bc \leq 2s$.

Proof. The first part follows from Pocklington's Theorem. Thus $N = (bF+1)(cF+1) = bcF^2 + (b+c)F + 1$ with $b, c > 0$ and bc even (since R is odd and F is even). Write $\lambda = s - \frac{bc}{2}$. Then $2F\lambda + r = b + c > 0$ and, since $r < 2F$, we have $\lambda \geq 0$. On the other hand, by Lemma 1 of [2], $(bF+1) + (cF+1) \leq G + \frac{N}{G}$ and therefore, by (7.4), $\lambda < A$.

Theorem 7.2. *Suppose N is composite, suppose $N+1 = FR$, where F is even, R is odd, $R > 1$ and $\gcd(F, R) = 1$. Suppose there exists an integer D for which $(D/N) = -1$ and also, for each prime factor f of F, a Lucas sequence $\{U_i\}$ with discriminant D such that N divides U_{N+1} and $\gcd(U_{(N+1)/f}, N) = 1$. If every prime factor of N is greater than G and $N^{1/3} < G < N^{1/2}$, then N is the product of two primes, $bF+e$ and $cF-e$, where $e = 1$ or -1. Furtheremore, if*

$$N < G(2AF^2 + G - |rF - 2|) \tag{7.5}$$

for some $A \geq 1$*, where* r *and* s *are defined by* $R = 2Fs + r$, $|r| < F$*, then* $2(s - A) < bc < 2(s + A)$.

Proof. The first part follows from Morrison's Theorem. Thus $N = (bF + e)(cF - e) = bcF^2 + e(c - b)F - 1$ with $e = \pm 1$, $b, c > 0$ and bc even. Write $\lambda = s - \frac{bc}{2}$. Then $2\lambda F^2 + rF - 2 = FR - bcF^2 - 2 = e(c - b)F - 2$ and, since it is the difference between the two prime divisors of N, $|e(c - b)F - 2|$ is bounded above by $\frac{N}{G} - G$. Hence $|2\lambda F^2 + rF - 2| \leq \frac{N}{G} - G < 2AF^2 - |rF - 2|$ by (7.5). Therefore $-A < \lambda < A$.

8. Acknowledgements

We would like to thank Jeremy Humphries, Manfred Toplic and Paul Zimmermann for contributing their own computer resources to the search for seven prime Pythagorean triangles. We are specifically grateful to Paul Zimmermann, who also made his Elliptic Curve program available to us for the partial factorization of $P_7 - 1$.

References

1. A. H. Beiler, *Recreations In the Theory of Numbers*, 2nd ed., Dover Publications, New York, ch. XIV, 1966.
2. John Brillhart, D. H. Lehmer and J. L. Selfridge, *New primality criteria and factorizations of* $2^m \pm 1$, Math. Comp., **29** (1975), 620-647.
3. C. K. Caldwell, *UBASIC*, J. Recreational Math., **25** (1993), 47-54.
4. H. Cohen and A. K. Lenstra, *Implementation of a new primality test*, Math. Comp., **48** (1987), 103-121.
5. H. Cohen and H. W. Lenstra, *Primality testing and Jacobi sums*, Math. Comp., **42** (1984), 297-330.
6. Tony Forbes, *Prime clusters and Cunningham chains*, Math. Comp., **68** (1999), 1739-1747.
7. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 5th ed., Oxford University Press, 1979.
8. H. C. Pocklington, *The determination of the prime or composite nature of large numbers by Fermat's theorem*, Proc. Cambridge Philos. Soc., **18** (1914-16), 29-30.
9. P. Ribenboim, *The New Book of Prime Number Records*, 3rd ed., Springer-Verlag, New York, 1995.

449 Beverly Road, Ridgewood, New Jersey 07450
E-mail address: `hdubner1@compuserve.com`

22 St. Albans Road, Kingston upon Thames, Surrey, KT2 5HQ England
E-mail address: `tonyforbes@ltkz.demon.co.uk`