Computing Euclid's Primes

Samuel S. Wagstaff, Jr.

Department of Computer Sciences
Purdue University
West Lafayette, Indiana 47907
U.S.A.

In Proposition 20 of Book IX of his *Elements*, Euclid gave a proof like the following that there are infinitely many primes. Suppose that $p_1, \ldots, p_n$ are all the primes we know about. Let $P_n = \prod_{i=1}^{n} p_i$. Then $1 + P_n$ is not divisible by any of the primes $p_1, \ldots, p_n$, so the prime factors of $1 + P_n$ are new to us. Hence, the number of primes is unbounded. If we "discover" just the smallest prime factor p_{n+1} of $1 + P_n$ and if we begin with $p_1 = 2$, then we are lead in a natural way to the sequence $p_2 = 3$, $p_3 = 7$, $p_4 = 43$, $p_5 = 13$, etc. Shanks [**8**] has conjectured that this sequence contains all primes. He gave a heuristic argument which makes this conjecture plausible.

We have computed p_n as far as $p_{43} = 4357$. We have factored $1+P_n$ completely for all n up to 27 and for several larger n. Our results support Shanks' conjecture. Guy and Nowakowski [**2**] studied $\{p_n\}$ and several related sequences. We extend the computation of some of their sequences and answer a question of Mullin.

Euclid's proof does not specify which prime factor(s) of 1 plus the product of those found so far should be "discovered". If only the largest one is discovered, then we would obtain the sequence $q_1 = 2$, $Q_n = \prod_{i=1}^{n} q_i$, q_{n+1} = the largest prime factor of $1 + Q_n$, with $q_2 = 3$, $q_3 = 7$, $q_4 = 43$, $q_5 = 139$, etc. Many difficult factorizations must be done to compute the sequences $\{p_n\}$ and $\{q_n\}$. The sequences $\{p_n\}$ and $\{q_n\}$ appear in Sloane's *Handbook* [**9**] as sequences number 329 and 330, respectively.

If one feels that *all* prime factors of 1 plus the product of those found so far are "discovered", then one is lead to the sequence $a_1 = 2$, $A_n = \prod_{i=1}^{n} a_i$, $a_{n+1} = 1 + A_n$. The terms of this sequence can be computed without any factoring since $a_{n+1} = a_n(a_n - 1) + 1$. We do not consider this sequence further because Guy and Nowakowski [**2**] have already investigated it thoroughly.

Provided that one begins with the prime 3, Euclid's proof will work if one *subtracts* 1 from the product of the primes found so far. This modification leads to these two sequences: $r_1 = 3$, $R_n = \prod_{i=1}^{n} r_i$, $r_{n+1} =$ the smallest prime factor of $R_n - 1$, so that $r_2 = 2$, $r_3 = 5$, $r_4 = 29$, $r_5 = 11$, etc., and $s_1 = 3$, $S_n = \prod_{i=1}^{n} s_i$, $s_{n+1} =$ the largest prime factor of $S_n - 1$, so that $s_2 = 2$, $s_3 = 5$, $s_4 = 29$, $s_5 = 79$, etc. Computing these sequences requires much factoring.

The values of these four sequences which are known to me are presented in Tables 1 to 6. Guy and Nowakowski [**2**] gave them up to p_{14}, q_9, r_{19} and s_{10}. Naur [**6**] computed the first eleven q_i.

The sequences $\{p_n\}$ and $\{r_n\}$ clearly are not monotonic. Guy and Nowakowski [**2**] found that $s_6 > s_7$ so that $\{s_n\}$ is not monotonic. Mullin [**5**] asked whether $\{q_n\}$ is monotonic. We see from Table 3 that $q_9 > q_{10}$ so that $\{q_n\}$ is not monotonic either.

Cox and van der Poorten [**1**] showed that some primes (including 5, 11, 13, 17, 19, 23, 29, 31, 37, 41 and 47) do not appear in $\{q_n\}$. Selfridge (see [**2**]) showed that some primes (including 7, 11, 13, 17, 19 and 23) are absent from $\{s_n\}$.

On the other hand, there is good reason to believe that $\{p_n\}$ and $\{r_n\}$ contain all primes. Shanks [**8**] gave a heuristic argument that $\{p_n\}$ contains all primes. Here is the analog of his argument for $\{r_n\}$: Let q be the smallest prime that has not occured up to r_N. Let a and b be the least non-negative residues modulo q of R_{N-1} and r_N, respectively. Then q does not divide ab since q has not occured yet. But $q = r_{N+1}$ if and only if

$$ab \equiv 1 \pmod{q}. \tag{1}$$

The product ab modulo q can *a priori* be any residue between 1 and $q-1$. If (1) fails, then we can replace N by $N+1$, $N+2$, etc. After $k(q-1)$ values of N, each residue between 1 and $q-1$ will be represented by ab modulo q an average of k times. As $k \to \infty$ it is highly unlikely that (1) will never happen. When it does happen, q appears and (1) can never happen again since q divides a ever after.

Of course, we have not proved the approximate equidistribution of ab among the non-zero residue classes modulo q. The only hint I know that this hypothesis might fail is a tiny one. Sometimes $R_n - 1$ is prime, so that $r_{n+1} = R_n - 1$. (This happens for $n = 1, 2, 3, 8$ and 10, for example.) In this situation we have

THEOREM. *If $n > 1$ and $r_{n+1} = R_n - 1$, then $r_{n+2} \equiv 1$ or $9 \pmod{10}$.*

Proof: We have $R_{n+1} = R_n r_{n+1} = R_n^2 - R_n$, so that $4(R_{n+1} - 1) = (2R_n - 1)^2 - 5$. Thus 5 is a quadratic residue of any factor of $R_{n+1} - 1$ and, in particular, of its smallest prime factor r_{n+2}. When $n = 1$, $r_{n+2} = 5$. But when $n > 1$, 5 divides R_{n+1} and so not $R_{n+1} - 1$. Thus $r_{n+2} \equiv 1$ or $4 \pmod 5$. The conclusion follows because r_{n+2} is odd.

I expect that prime values of $R_n - 1$ are so rare that this theorem will not affect the heuristic argument above. As you can see from Tables 4 and 5, when $R_n - 1$ is composite r_{n+2} may have 3 or 7 for its unit's digit. The Theorem is analogous to one which Shanks [**8**] proved for $\{p_n\}$.

Shanks [**8**] noted that 31, 41, 47, 59, 67 and 73 are the first few primes which have not yet known appeared in $\{p_n\}$. We have computed $\{r_n\}$ a bit further than $\{p_n\}$. The first primes which have not yet appeared in $\{r_n\}$ are 53, 59, 61, 67, 71 and 73.

Most of the factoring was done by a program written by Peter Montgomery. Methods of factoring used included trial division (to 10000), Pollard's $p - 1$ method [**7**] and Lenstra's elliptic curve method [**3**].

In the tables, when a number is asserted to be the greatest or least prime factor of another number, some proof is required. In each case when p is claimed to the greatest prime factor of P, I have factored P completely. These complete factorizations are given in the early parts of Tables 3 and 6. The bulky factorizations of large numbers at the ends of these tables are given in Table 7. In some lines of Table 7 a long factorization is broken at a center dot.

When a small prime p (less than 10^8, say) is supposed to be the least prime factor of P, this fact may be checked easily by trial division. In most cases when we say that a larger prime p is the least prime factor of P, we give the complete factorization of P in Table 1, 4, 7 or 8. One difficult proof of this type was that the ten-digit prime factor

$p = 3143065813$ of $1 + P_{31}$ is indeed p_{32}. We showed this by a novel application of the elliptic curve method (ECM). Suppose that $1 + P_{31}$ had a prime factor $q < p$. Our goal was to run ECM on $(1 + P_{31})/p$ once and either discover q certainly or show that there was no such divisor q. Suppose we run ECM with limits L_1 for Step 1 and L_2 for Step 2 and assume that $10 < L_1 < L_2$. ECM begins by choosing a random elliptic curve whose order over $GF(q)$ is e. This run of ECM will discover q provided that the greatest prime factor of e is $< L_2$ and all other prime factors of e are $< L_1$. (Montgomery's program [**4**] uses high powers of small primes to allow for any possible repeated prime factors of e.) Although e is unknown to us, we do know that $e < q + 2\sqrt{q} - 1$. Hence, $e < p + 2\sqrt{p} - 1 < 3143200000$.

Now it is possible when starting ECM to insure that the unknown order e is divisible by 12 (see [**4**]). Let $m = e/12$. Then $m < 262000000$. This run of ECM will discover q provided that the largest prime factor of m is $< L_2$ and all other prime factors of m are $< L_1$. These conditions are satisfied provided we choose $L_2 > 262000000$ and $L_1 > \sqrt{262000000}$ or $L_1 > 16187$. The run was made with $L_1 = 20000$ and $L_2 = 270000000$. Since no factor was found, it was shown that p is the smallest prime factor of $1 + P_{31}$, so that $p_{32} = p$.

In a similar fashion, it was shown that the smallest prime factors of $R_{25} - 1$, $R_{28} - 1$ and $R_{49} - 1$ are r_{26}, r_{29} and r_{50}, respectively. However, we could not show without undue effort that the twelve-digit divisor of $R_{53} - 1$ was actually r_{54}. That is why we stopped computing $\{r_n\}$ with r_{53}.

Table 1.

$p_1 = 2$, $P_n = \prod_{i=1}^{n} p_i$, p_{n+1} = least prime factor of $1 + P_n$.

n	p_n	$1 + P_n$
1	2	3
2	3	7
3	7	43
4	43	$1807 = 13 \cdot 139$
5	13	$23479 = 53 \cdot 443$
6	53	$1244335 = 5 \cdot 248867$
7	5	6221671 (prime)
8	6221671	38709183810571 (prime)
9	38709183810571	1498400911280533294827535471
		$= 139 \cdot 25621 \cdot 420743244646304724409$
10	139	20827772666799412798102743033 1
		$= 2801 \cdot 2897 \cdot 489241 \cdot 119812279 \cdot 437881957$
11	2801	58338591239705155247485783235433 1
		$= 11 \cdot 1009 \cdot 241139351 \cdot 2179736509396276989 19$
12	11	641724503636756707722343615589763 1
		$= 17 \cdot 1949 \cdot 193681376161759185018665262907$
13	17	1090931656182486403127984146502597 11
		$= 5471 \cdot 1994026057727081709245081605744 1$
14	5471	596848709097438311151320126551570873411
		$= 52662739 \cdot 113334156261306179147142370728 49$
15	52662739	31431687789685319348762761330032346946392869991
		$= 23003 \cdot 9481141 \cdot 144119457035843546516309623213989617$
16	23003	723023114226131400979589798874734076807875188379971
		$= 30693651606209 \cdot 23556112628836625540740261445212918019$

Table 2.

$p_1 = 2,\ P_n = \prod_{i=1}^{n} p_i,\ p_{n+1} =$ least prime factor of $1 + P_n$.

n	p_n
17	30693651606209
18	37
19	1741
20	1313797957
21	887
22	71
23	7127
24	109
25	23
26	97
27	159227
28	643679794963466223081509857
29	103
30	1079990819
31	9539
32	3143065813
33	29
34	3847
35	89
36	19
37	577
38	223
39	139703
40	457
41	9649
42	61
43	4357

Table 3.

$q_1 = 2$, $Q_n = \prod_{i=1}^{n} q_i$, q_{n+1} = greatest prime factor of $1 + Q_n$.

n	q_n	$1 + Q_n$
1	2	3
2	3	7
3	7	43
4	43	$1807 = 13 \cdot 139$
5	139	$251035 = 5 \cdot 50207$
6	50207	$12603664039 = 23 \cdot 1607 \cdot 340999$
7	340999	$4297836833293963 = 23 \cdot 79 \cdot 2365347734339$
8	2365347734339	$10165878616190575459068761119$ $= 17 \cdot 127770091783 \cdot 4680225641471129$
9	4680225641471129	
10	1368845206580129	
11	889340324577880670089824574922371	
12	20766142440959799312827873190033784610984957267051218394040721	
13	348654613352373829454902145370501700873487314509264314920485482\16\ 1426646699863760337897225492334460782554524464800179 9	

Table 4.

$r_1 = 3$, $R_n = \prod_{i=1}^{n} r_i$, r_{n+1} = least prime factor of $R_n - 1$.

n	r_n	$R_n - 1$
1	3	2
2	2	5
3	5	29
4	29	$869 = 11 \cdot 79$
5	11	$9569 = 7 \cdot 1367$
6	7	$66989 = 13 \cdot 5153$
7	13	$870869 = 37 \cdot 23537$
8	37	32222189 (prime)
9	32222189	$1038269496173909 = 131 \cdot 1610899 \cdot 4920061$
10	131	136013303998782209 (prime)
11	136013303998782209	18499618864665144581031859013701889
		$= 31 \cdot 41 \cdot 181 \cdot 499 \cdot 8870749 \cdot 18166774231909276189$
12	31	573488184804619482011987629424758589
		$= 197 \cdot 3221 \cdot 903789983570098326830409620597$
13	197	112977172406510037956361562996677442229
		$= 19 \cdot 2154611 \cdot 9547427 \cdot 49532972059 \cdot 5835626580317$
14	19	2146566275723690721170869696936871402369
		$= 157 \cdot 769 \cdot 2543 \cdot 271338827 \cdot 25766771512898971353713$
15	157	337010905288619443223826542419088810172089
		$= 17 \cdot 452704788101 \cdot 43790504143967027283161477717$
16	17	5729185389906530534805051221124509772925529
		$= 8609 \cdot 32183 \cdot 8907623 \cdot 23214098064220105304253412 09$

Table 5.

$r_1 = 3$, $R_n = \prod_{i=1}^{n} r_i$, r_{n+1} = least prime factor of $R_n - 1$.

n	r_n
17	8609
18	1831129
19	35977
20	508326079288931
21	487
22	10253
23	1390043
24	18122659735201507243
25	25319167
26	9512386441
27	85577
28	1031
29	3650460767
30	107
31	41
32	811
33	15787
34	89
35	68168743
36	4583
37	239
38	1283
39	443
40	902404933
41	64775657
42	2753
43	23
44	149287
45	149749
46	7895159
47	79
48	43
49	1409
50	184274081
51	47
52	569
53	63843643

Table 6.

$s_1 = 3$, $S_n = \prod_{i=1}^{n} s_i$, s_{n+1} = greatest prime factor of $S_n - 1$.

n	s_n	$S_n - 1$
1	3	2
2	2	5
3	5	29
4	29	$869 = 11 \cdot 79$
5	79	68729 (prime)
6	68729	$4723744169 = 61 \cdot 139 \cdot 149 \cdot 3739$
7	3739	$17662079451629 = 2839019 \cdot 6221191$
8	6221191	$109879169725765491329 = 83 \cdot 8423 \cdot 157170297801581$
9	157170297801581	$41 \cdot 5955703423 \cdot 70724343608203457341903$
10	70724343608203457341903	
11	46316297682014731387158877659877	
12	78592684042614093322289223662773	
13	181891012640244955605725966274974474087	

Table 7. Auxiliary Factorizations.

Notation: Pxx is a prime of xx digits, Cxx is a composite of xx digits

Number	Factorization
$1 + P_{17}$	$37 \cdot 8109973 \cdot 1049918455514883211 \cdot P38$
$1 + P_{18}$	$1741 \cdot 2687771 \cdot P57$
$1 + P_{19}$	$1313797957 \cdot 1587086232579380268953381 \cdot P36$
$1 + P_{20}$	$887 \cdot 6599 \cdot 1630146233 \cdot 2993625319460509818171977729 \cdot P36$
$1 + P_{21}$	$71 \cdot 3299661004790609 \cdot 11782243278281460747007953787 \cdot P35$
$1 + P_{22}$	$7127 \cdot 352201 \cdot 155354729501063 \cdot 11654246919591371 \cdot P44$
$1 + P_{23}$	$109 \cdot 85669 \cdot 232047887 \cdot 2824330157926317541 \cdot P54$
$1 + P_{24}$	$23 \cdot P88$
$1 + P_{25}$	$97 \cdot 191 \cdot 474716141 \cdot 65748525431 \cdot P67$
$1 + P_{26}$	$159227 \cdot 1067159 \cdot 43497281 \cdot 2527540905245931542309 \cdot P53$
$1 + P_{27}$	$643679794963466223081509857 \cdot 2496022367830647867616317307 \cdot P44$
$1 + P_{28}$	$103 \cdot 31336667 \cdot 36591209 \cdot C108$
$1 + P_{29}$	$1079990819 \cdot 2434978091641012135177 \cdot P96$
$1 + P_{30}$	$9539 \cdot 245433668891 \cdot 979752962034735781 \cdot 847371699114699802 7\cdot$ $\cdot 26294987506338782316507217723423 \cdot P52$
$1 + P_{31}$	$3143065813 \cdot C130$
$1 + P_{32}$	$29 \cdot 10429 \cdot 165047 \cdot C139$
$1 + P_{33}$	$3847 \cdot 2607917067290207 \cdot P132$
$1 + P_{34}$	$89 \cdot 191 \cdot 677371128232689991 \cdot 336373220775307632 47 \cdot C113$
$1 + P_{35}$	$19 \cdot 787 \cdot 7757 \cdot 28006756507 \cdot 1022974063703 \cdot C126$
$1 + P_{36}$	$577 \cdot P155$
$1 + P_{37}$	$223 \cdot 5393 \cdot 74673192479 \cdot P143$
$1 + P_{38}$	$139703 \cdot 43085355700150267667 \cdot P138$
$1 + P_{39}$	$457 \cdot 37179386588269 \cdot 159834478959851 \cdot P137$
$1 + P_{40}$	$9649 \cdot 319466050329395719 \cdot P149$
$1 + P_{41}$	$61 \cdot 6827978951 \cdot 66042713762390953740707 \cdot C140$
$1 + P_{42}$	$4357 \cdot 7027 \cdot C169$
$1 + P_{43}$	$C180$
$1 + Q_9$	$89 \cdot 839491 \cdot 556266121 \cdot 836312735653 \cdot 1368845206580129$
$1 + Q_{10}$	$1307 \cdot 5603023948537038280588 7 \cdot 8893403245778806700898245749223 71$
$1 + Q_{11}$	$11 \cdot 2535627899784285829626317277 29 \cdot P62$
$1 + Q_{12}$	$739 \cdot 2311 \cdot 2019993928879340834647669995 29 \cdot P118$
$1 + Q_{13}$	$11 \cdot 13 \cdot 107536547 \cdot C261$
$S_{10} - 1$	$7 \cdot 349 \cdot 449 \cdot 112939 \cdot 9937441 \cdot 21420649 \cdot P32$
$S_{11} - 1$	$7 \cdot 257 \cdot 521 \cdot 682511 \cdot 10829594203 \cdot 50852665316801 \cdot$ $\cdot 20431584153688937909 39 \cdot P32$
$S_{12} - 1$	$7 \cdot 11 \cdot 17 \cdot 86599 \cdot 294757 \cdot 933418660159 \cdot 9669562218961751 \cdot$ $\cdot 2289336175732053683 \cdot 35403807765085882291423 \cdot P39$
$S_{13} - 1$	$11 \cdot 204249779 \cdot C150$

Table 8. More Auxiliary Factorizations.

Notation: Pxx is a prime of xx digits, Cxx is a composite of xx digits

Number	Factorization
$R_{17}-1$	$1831129 \cdot 96593227 \cdot 395499093031447 \cdot 705073635630813269$
$R_{18}-1$	$35977 \cdot 30902882521913 \cdot 12326099580658421 \cdot 6590447658135399749$
$R_{19}-1$	$508326079288931 \cdot 8888176173420238273 \cdot 7191747396675796605978 43$
$R_{20}-1$	$487 \cdot 4783 \cdot 317419 \cdot P61$
$R_{21}-1$	$10253 \cdot 112687 \cdot 24025694597 \cdot P56$
$R_{22}-1$	$1390043 \cdot 836498713878858549845338160532 7 \cdot P42$
$R_{23}-1$	$18122659735201507243 \cdot P66$
$R_{24}-1$	$25319167 \cdot 5211496051 \cdot 58429754491680845821 \cdot P68$
$R_{25}-1$	$9512386441 \cdot C102$
$R_{26}-1$	$85577 \cdot C117$
$R_{27}-1$	$1031 \cdot 1787 \cdot 274100051 \cdot 2353368011777399 \cdot C97$
$R_{28}-1$	$3650460767 \cdot C121$
$R_{29}-1$	$107 \cdot 1636358697177293 \cdot C122$
$R_{30}-1$	$41 \cdot C140$
$R_{31}-1$	$811 \cdot 86085747863 \cdot C130$
$R_{32}-1$	$15787 \cdot 1763431 \cdot P136$
$R_{33}-1$	$89 \cdot 12211 \cdot 1577027 \cdot P138$
$R_{34}-1$	$68168743 \cdot 2880625453 \cdot 2119710631572329177 \cdot P117$
$R_{35}-1$	$4583 \cdot 630175649 \cdot 13723021380961 \cdot C135$
$R_{36}-1$	$239 \cdot C162$
$R_{37}-1$	$1283 \cdot 23059 \cdot C159$
$R_{38}-1$	$443 \cdot C167$
$R_{39}-1$	$902404933 \cdot 8037715351 \cdot 29371574741 \cdot P143$
$R_{40}-1$	$64775657 \cdot 385983277 \cdot C165$
$R_{41}-1$	$2753 \cdot C185$
$R_{42}-1$	$23 \cdot 40904021 \cdot C183$
$R_{43}-1$	$149287 \cdot 172969 \cdot 1588051 \cdot C177$
$R_{44}-1$	$149749 \cdot 33807989 \cdot C186$
$R_{45}-1$	$7895159 \cdot C197$
$R_{46}-1$	$79 \cdot 137 \cdot 367 \cdot C204$
$R_{47}-1$	$43 \cdot 61 \cdot 991 \cdot 14821 \cdot 60077 \cdot C197$
$R_{48}-1$	$1409 \cdot 218131 \cdot 293847231283 \cdot C194$
$R_{49}-1$	$184274081 \cdot C209$
$R_{50}-1$	$47 \cdot 547 \cdot 1571 \cdot 4621 \cdot C215$
$R_{51}-1$	$569 \cdot C225$
$R_{52}-1$	$63843643 \cdot 1037601959 \cdot C213$
$R_{53}-1$	$111973205287 \cdot C227$

References

1. C. D. Cox and A. J. van der Poorten, "On a sequence of prime numbers," *J. Austral. Math. Soc.* **8** (1968), 571–574. MR 37 # 3998.
2. Richard Guy and Richard Nowakowski, "Discovering primes with Euclid," *Delta* **5** (1975), 49–63. MR 52 # 5548.
3. H. W. Lenstra, Jr., "Factoring integers with elliptic curves," *Ann. of Math. (2)* **126** (1987), 649–673. MR 89g:11125.
4. Peter L. Montgomery, *An FFT Extension of the Elliptic Curve Method of Factorization*, Ph. D. thesis at the University of California, Los Angeles, 1992.
5. Albert A. Mullin, "Recursive function theory (a modern look at a Euclidean idea)," *Bull. Amer. Math. Soc.* **69** (1963), 737.
6. Thorkil Naur, *Integer Factorization*, DAIMI Report PB-144, University of Aarhus, 1982.
7. J. M. Pollard, "Theorems on factorization and primality testing," *Proc. Camb. Phil. Soc.* **76** (1974), 521–528. MR 50 # 6992.
8. Daniel Shanks, "Euclid's primes," *Bull. Inst. Combinatorics and its Applications* **1** (1991), 33–36.
9. N. J. A. Sloane, *A Handbook of Integer Sequences*, Academic Press, New York-London, 1973. MR 50 # 9760.