The primes that Euclid forgot

Enrique Treviño

joint work with Paul Pollack

AMS/MAA Joint Math Meetings
January 16, 2014

LAKE FOREST COLLEGE

There are infinitely many primes

Start with $q_1 = 2$. Supposing that q_j has been defined for $1 \leq j \leq k$, continue the sequence by choosing a prime q_{k+1} for which

$$q_{k+1} \mid 1 + \prod_{j=1}^{k} q_j.$$

Then 'at the end of the day', the list $q_1, q_2, q_3, \ldots$ is an infinite sequence of distinct prime numbers.

Euclid-Mullin sequences

Since the sequence in the previous slide is not unique, Mullin suggested two possible unique sequences.

- The first is to take $q_1 = 2$, then define recursively q_k to be the **smallest** prime dividing $1 + q_1 q_2 \ldots q_{k-1}$.
- i,e. 2, 3, 7, 43, 13, 53, 5, 6221671, 38709183810571, 139, 2801, 11, 17, 5471, 52662739, ...
- It is conjectured that the first Mullin sequence touches all the primes eventually.
- Not much is known of this sequence.

Second Euclid-Mullin Sequence

- The second Mullin sequence is to take $q_1 = 2$, then define recursively q_k to be the **largest** prime dividing $1 + q_1 q_2 \dots q_{k-1}$.
- i.e. 2, 3, 7, 43, 139, 50207, 340999, 2365347734339, 4680225641471129, ...
- Cox and van der Poorten (1968) proved 5, 11, 13, 17, 19, 23, 29, 31, 37, 41, 47, and 53 are missing from the first Euclid-Mullin sequence.
- Booker in 2012 showed that infinitely many primes are missing from the sequence.
- Booker's proof uses deep theorems from analytic number theory such as the Burgess inequality.

5 is not in the second Euclid-Mullin sequence

- Suppose 5 is in the Euclid-Mullin sequence.
- Therefore there exists n such that $5|q_n = 1 + q_1q_2 + \ldots q_{n-1}$ and with 5 being the largest prime divisor of q_n.
- Since $q_1 = 2$ and $q_2 = 3$, then $(q_n, 6) = 1$.
- Therefore $q_n = 5^\alpha$ for some $\alpha \geq 1$.
- Now $5^\alpha \equiv 1 \pmod 4$ while $1 + q_1q_2 \ldots q_{n-1} \equiv 3 \pmod 4$.
- Contradiction!

5 is not in the second Euclid-Mullin sequence

- Suppose 5 is in the Euclid-Mullin sequence.
- Therefore there exists n such that $5|q_n = 1 + q_1 q_2 + \ldots q_{n-1}$ and with 5 being the largest prime divisor of q_n.
- Since $q_1 = 2$ and $q_2 = 3$, then $(q_n, 6) = 1$.
- Therefore $q_n = 5^\alpha$ for some $\alpha \geq 1$.
- Now $5^\alpha \equiv 1 \pmod 4$ while $1 + q_1 q_2 \ldots q_{n-1} \equiv 3 \pmod 4$.
- Contradiction!

5 is not in the second Euclid-Mullin sequence

- Suppose 5 is in the Euclid-Mullin sequence.
- Therefore there exists n such that $5|q_n = 1 + q_1 q_2 + \dots q_{n-1}$ and with 5 being the largest prime divisor of q_n.
- Since $q_1 = 2$ and $q_2 = 3$, then $(q_n, 6) = 1$.
- Therefore $q_n = 5^\alpha$ for some $\alpha \geq 1$.
- Now $5^\alpha \equiv 1 \pmod 4$ while $1 + q_1 q_2 \dots q_{n-1} \equiv 3 \pmod 4$.
- Contradiction!

Squares

Consider the sequence

$$2, 5, 8, 11, \ldots$$

Can it contain any squares?

- Every positive integer n falls in one of three categories: $n \equiv 0$, 1 or 2 $\pmod 3$.
- If $n \equiv 0 \pmod 3$, then $n^2 \equiv 0^2 = 0 \pmod 3$.
- If $n \equiv 1 \pmod 3$, then $n^2 \equiv 1^2 = 1 \pmod 3$.
- If $n \equiv 2 \pmod 3$, then $n^2 \equiv 2^2 = 4 \equiv 1 \pmod 3$.

Squares

Consider the sequence

$$2, 5, 8, 11, \ldots$$

Can it contain any squares?

- Every positive integer n falls in one of three categories: $n \equiv 0, 1 \text{ or } 2 \pmod 3$.
- If $n \equiv 0 \pmod 3$, then $n^2 \equiv 0^2 = 0 \pmod 3$.
- If $n \equiv 1 \pmod 3$, then $n^2 \equiv 1^2 = 1 \pmod 3$.
- If $n \equiv 2 \pmod 3$, then $n^2 \equiv 2^2 = 4 \equiv 1 \pmod 3$.

Squares and non-squares

Let n be a positive integer. For $q \in \{0, 1, 2, \ldots, n-1\}$, we call q a square mod n if there exists an integer x such that $x^2 \equiv q \pmod{n}$. Otherwise we call q a non-square.

- For $n = 3$, the squares are $\{0, 1\}$ and the non-square is 2.
- For $n = 5$, the squares are $\{0, 1, 4\}$ and the non-squares are $\{2, 3\}$.
- For $n = 7$, the squares are $\{0, 1, 2, 4\}$ and the non-squares are $\{3, 5, 6\}$.
- For $n = p$, an odd prime, there are $\frac{p+1}{2}$ squares and $\frac{p-1}{2}$ non-squares.

Least non-square

How big can the least non-square be?
Let $g(p)$ be the least non-square modulo p.

p	Least non-square
3	2
5	2
7	3
11	2
13	2
17	3
19	2
23	5
29	2
31	3

p	Least non-square
7	3
23	5
71	7
311	11
479	13
1559	17
5711	19
10559	23
18191	29
31391	31
422231	37
701399	41
366791	43
3818929	47

An elementary bound for $g(p)$

Let $g(p)$ be the least non-square mod p.

Theorem

$g(p) \leq \sqrt{p} + 1$.

Proof.

Suppose $g(p) = q$ with $q > \sqrt{p} + 1$. Let k be the ceiling of p/q. Then $p < kq < p + q$, so $kq \equiv a \bmod p$ for some $0 < a < q$, and therefore kq is a square modulo p. Since $q > \sqrt{p} + 1$, then $p/q < \sqrt{p}$, so k is at most the ceiling of $\sqrt{p} < \sqrt{p} + 1 < q$. Therefore k is a square modulo p. But if k and kq are squares modulo p, then q is a square modulo p. Contradiction! □

Consecutive squares or non-squares

Let $H(p)$ be the largest string of consecutive nonzero squares or non-squares modulo p.
For example, with $p = 7$ we have that the nonzero squares are $\{1, 2, 4\}$ and the non-squares are $\{3, 5, 6\}$. Therefore $H(7) = 2$.

p	$H(p)$
11	3
13	4
17	3
19	4
23	4
29	4
31	4
37	4
41	5

An elementary bound for $H(p)$

Sketch of a proof that $H(p) < 2\sqrt{p}$.

- The largest string of non-squares is $< 2\sqrt{p}$.
- Suppose $\{a+1, a+2, \ldots, a+H\}$ are all squares mod p.
- For n a non-square, $na+n, \ldots, na+Hn$ are non-squares.
- If $Hn > p$, then $H(p) \leq n-1$. Therefore $H(p) \leq \max\{p/n, n-1, 2\sqrt{p}\}$.
- If $n \in (\sqrt{p}/2, 2\sqrt{p}]$ we have $H(p) < 2\sqrt{p}$.
- Let k be the largest integer such that $k^2 g(p) \leq \sqrt{p}/2$.
- $(k+1)^2 g(p) > 2\sqrt{p} \geq 4k^2 g(p)$ implies $(2k+1) > 3k^2$ which is false for each $k \geq 1$. Therefore there is a non-square in the interval $(\sqrt{p}/2, 2\sqrt{p}]$, yielding $H(p) < 2\sqrt{p}$.

An elementary bound for $H(p)$

Sketch of a proof that $H(p) < 2\sqrt{p}$.

- The largest string of non-squares is $< 2\sqrt{p}$.
- Suppose $\{a+1, a+2, \ldots, a+H\}$ are all squares mod p.
- For n a non-square, $na+n, \ldots, na+Hn$ are non-squares.
- If $Hn > p$, then $H(p) \leq n-1$. Therefore $H(p) \leq \max\{p/n, n-1, 2\sqrt{p}\}$.
- If $n \in (\sqrt{p}/2, 2\sqrt{p}]$ we have $H(p) < 2\sqrt{p}$.
- Let k be the largest integer such that $k^2 g(p) \leq \sqrt{p}/2$.
- $(k+1)^2 g(p) > 2\sqrt{p} \geq 4k^2 g(p)$ implies $(2k+1) > 3k^2$ which is false for each $k \geq 1$. Therefore there is a non-square in the interval $(\sqrt{p}/2, 2\sqrt{p}]$, yielding $H(p) < 2\sqrt{p}$.

An elementary bound for $H(p)$

Sketch of a proof that $H(p) < 2\sqrt{p}$.

- The largest string of non-squares is $< 2\sqrt{p}$.
- Suppose $\{a+1, a+2, \ldots, a+H\}$ are all squares mod p.
- For n a non-square, $na+n, \ldots, na+Hn$ are non-squares.
- If $Hn > p$, then $H(p) \leq n-1$. Therefore $H(p) \leq \max\{p/n, n-1, 2\sqrt{p}\}$.
- If $n \in (\sqrt{p}/2, 2\sqrt{p}]$ we have $H(p) < 2\sqrt{p}$.
- Let k be the largest integer such that $k^2 g(p) \leq \sqrt{p}/2$.
- $(k+1)^2 g(p) > 2\sqrt{p} \geq 4k^2 g(p)$ implies $(2k+1) > 3k^2$ which is false for each $k \geq 1$. Therefore there is a non-square in the interval $(\sqrt{p}/2, 2\sqrt{p}]$, yielding $H(p) < 2\sqrt{p}$.

An elementary bound for $H(p)$

Sketch of a proof that $H(p) < 2\sqrt{p}$.

- The largest string of non-squares is $< 2\sqrt{p}$.
- Suppose $\{a+1, a+2, \ldots, a+H\}$ are all squares mod p.
- For n a non-square, $na+n, \ldots, na+Hn$ are non-squares.
- If $Hn > p$, then $H(p) \leq n-1$. Therefore $H(p) \leq \max\{p/n, n-1, 2\sqrt{p}\}$.
- If $n \in (\sqrt{p}/2, 2\sqrt{p}]$ we have $H(p) < 2\sqrt{p}$.
- Let k be the largest integer such that $k^2 g(p) \leq \sqrt{p}/2$.
- $(k+1)^2 g(p) > 2\sqrt{p} \geq 4k^2 g(p)$ implies $(2k+1) > 3k^2$ which is false for each $k \geq 1$. Therefore there is a non-square in the interval $(\sqrt{p}/2, 2\sqrt{p}]$, yielding $H(p) < 2\sqrt{p}$.

An elementary bound for $H(p)$

Sketch of a proof that $H(p) < 2\sqrt{p}$.

- The largest string of non-squares is $< 2\sqrt{p}$.
- Suppose $\{a+1, a+2, \dots, a+H\}$ are all squares mod p.
- For n a non-square, $na+n, \dots, na+Hn$ are non-squares.
- If $Hn > p$, then $H(p) \leq n-1$. Therefore $H(p) \leq \max\{p/n, n-1, 2\sqrt{p}\}$.
- If $n \in (\sqrt{p}/2, 2\sqrt{p}]$ we have $H(p) < 2\sqrt{p}$.
- Let k be the largest integer such that $k^2 g(p) \leq \sqrt{p}/2$.
- $(k+1)^2 g(p) > 2\sqrt{p} \geq 4k^2 g(p)$ implies $(2k+1) > 3k^2$ which is false for each $k \geq 1$. Therefore there is a non-square in the interval $(\sqrt{p}/2, 2\sqrt{p}]$, yielding $H(p) < 2\sqrt{p}$.

The primes that Euclid forgot

Theorem

Let $Q_1, Q_2, \ldots Q_r$ be the smallest r primes omitted from the second Euclid-Mullin sequence, where $r \geq 0$. Then there is another omitted prime smaller than

$$12^2 \left(\prod_{i=1}^{r} Q_i \right)^2 .$$

Using the deep results of Burgess, Booker showed that the exponent can be replaced with any real number larger than $\frac{1}{4\sqrt{e}-1} = 0.178734\ldots$, provided that 12^2 is also replaced by a possibly larger constant.

Proof Sketch

Let $X = 12^2(\prod_{i=1}^{r} Q_i)^2$. Assume there is no prime missing from $[2, X]$ besides $Q_1, \ldots, Q_r$. Let p be the prime in $[2, X]$ that is last to appear in the sequence $\{q_i\}$.

Let n be such that $q_n = p$. Then $1 + q_1 \ldots q_{n-1} = Q_1^{e_1} \ldots Q_r^{e_r} p^e$.

Let d be the smallest number satisfying the following conditions:

(i) $d \equiv 1 \pmod 4$,

(ii) $d \equiv -1 \pmod{Q_1 \ldots Q_r}$

(iii) d and -1 are either both squares mod p or both non-squares mod p.

- Using the Chinese Remainder Theorem and the bound on $H(p)$ yields that $d \leq X$.
- Given the conditions on d and using that $d \leq X$ shows that d is both a square and a non-square mod $1 + q_1 q_2 \ldots q_{n-1}$. Contradiction!

Thank you!