Sifting the Primes

Gihan Marasingha
University of Oxford

18 March 2005

Irreducible forms:

$$q_1(x,y) := a_1x^2 + 2b_1xy + c_1y^2,$$
$$q_2(x,y) := a_2x^2 + 2b_2xy + c_2y^2,$$

$a_i, b_i, c_i \in \mathbb{Z}$.

Variety V defined by:

$$V: \begin{array}{l} q_1(x,y) = u^2 + v^2 \\ q_2(x,y) = s^2 + t^2 \end{array}$$

The Sieve of Eratosthenes

2	3	4	5	6	7
8	9	10	11	12	13
14	15	16	17	18	19
20	21	22	23	24	25
26	27	28	29	30	31
32	33	34	35	36	37

Prime numbers $n \leq N$. If n is composite, then it has a prime factor p with

$$p \leq \sqrt{n}.$$

Thus, having struck out multiples of primes $p \leq \sqrt{N}$, we've extinguished all the composite numbers less that N.

The Sieve of Eratosthenes–Legendre

$\pi(N) :=$ the number of primes $p \leq N$.

$$S(N, r) := \#\{n \leq N : 2, 3, 5, \ldots, p_r \nmid n\},$$

where $p_r < N$ is the r-th prime.

Then we have the relationship:

$$\pi(N) \leq p_r + S(N, r).$$

Reason: suppose p counted by $\pi(N)$, so that $p \leq N$. Either $p \leq p_r$ or $p > p_r$. In the first case, $p \in \{1, \ldots, p_r\}$, so p is counted by p_r, otherwise p is counted by $S(N, r)$.

Define

$$N_k := \#\{n \leq N : k|n\},$$

then

$$\begin{aligned} S(N, r) = N &- N_2 - N_3 - \ldots - N_{p_r} \\ &+ N_6 + N_{10} + \ldots + N_{p_i p_j} \\ &- \sum N_{p_i p_j p_k} + \ldots \pm N_{p_1 \ldots p_r} \end{aligned}$$

Now N_k counts $k, 2k, 3k, \ldots qk$ where $qk \leq N < (q+1)k$, so

$$N_k \leq \frac{N}{k} < N_k + 1.$$

$$\begin{aligned} S(N,r) = N &- \frac{N}{2} - \frac{N}{3} - \ldots - \frac{N}{p_r} \\ &+ \frac{N}{6} + \frac{N}{10} + \ldots + \frac{N}{p_i p_j} \\ &- \sum \frac{N}{p_i p_j p_k} + \ldots \pm \frac{N}{p_1 \ldots p_r} + \text{error}, \end{aligned}$$

where $|\text{error}| \leq 2^r \leq 2^{p_r}$. We'll write error$= O(2^{p_r})$, where $f(n) = O(g(n))$ means that there exists a constant C such that $|f(n)| \leq Cg(n)$. In our case $C = 1$. So

$$\begin{aligned} S &= N\left(1 - \sum_{i \leq r} \frac{1}{p_i} + \sum_{i \neq j \leq r} \frac{1}{p_i p_j} - \ldots \pm \frac{1}{p_1 \ldots p_r}\right) \\ &\quad + O(2^{p_r}). \\ &= N \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right) + O(2^{p_r}) \end{aligned}$$

Notation: $f(n) \sim g(n)$ means that $f(n)/g(n) \to 1$ as $n \to \infty$.

Fact: there is a constant C such that

$$\prod_{p<z}\left(1-\frac{1}{p}\right)\sim C\frac{1}{\log z}.$$

Choose r such that $p_r < \log N \le p_{r+1}$, then

$$\begin{aligned}S(N,r) &= N\prod_{p<\log N}\left(1-\frac{1}{p}\right)+O(2^{\log N})\\ &\sim C\frac{N}{\log\log N}+N^{\log 2}\\ &= O(\frac{N}{\log\log N}+N^{0.7})\\ &= O(\frac{N}{\log\log N}).\end{aligned}$$

Thus:

$$\begin{aligned}\pi(N)\le S(N,r)+p_r &= O(\frac{N}{\log\log N}+\log N)\\ &= O(\frac{N}{\log\log N}).\end{aligned}$$

Theorem (Hadamard, de la Vallée Poussin, 1896).

$$\pi(N)\sim\frac{N}{\log N}.$$

Theorem (Hadamard, de la Vallée Poussin, 1896).

$$\pi(N) \sim \frac{N}{\log N}.$$

Conjecture (Goldbach, 1750). *Let N be an even number greater that 2, then*

$$N = p + q,$$

for some primes p and q.

How many representations? That is what is

$$\#\{p \leq N : N - p \text{ is prime}\}?$$

Heuristically, it's

$$\begin{aligned}
&\sum_{p \leq N} \text{prob. that } N - p \text{ is prime} \\
&\approx \sum_{p \leq N} \frac{1}{\log(N-p)} \\
&\approx \sum_{p \leq N} \frac{1}{\log N} \\
&\approx \frac{1}{\log N}\pi(N) \sim \frac{1}{\log N}\frac{N}{\log N} = \frac{N}{(\log N)^2}.
\end{aligned}$$

Definition: We say $n \in \mathbb{N}$ is a k-almost prime and write that n is P_k if n has at most k prime factors.

Theorem (Chen Jing-Run, 1974). *For all sufficiently large even N, one has that*

$$N = p + P_2,$$

for p a prime. More precisely, there exists a (computable) constant C such that for all sufficiently large even N,

$$|\{p : p \leq N, N - p = P_2\}| > C\frac{N}{(\log N)^2}.$$

Idea: Get a good lower bound for representations $N = p + P_3$ then take away the representations $N = p + p_1p_2p_3$ by deducing an upper bound. What's left are the representations $N = p + P_2$.

Essentially, Chen is interested in calculating a lower bound for the number of 2-almost primes in the set

$$\mathcal{A} := \{N - p : p \neq N\},$$

much as in our heuristic development of the Goldbach conjecture. Chen's primary innovation in the solution of this problem was the "reversal of rôles", with which he relates $|\mathcal{A}|$ to $|\mathcal{B}|$, where

$$\begin{aligned}\mathcal{B} := \{N - p_1p_2p_3 : p_1p_2p_3 < N,\\ N^{1/10} \leq p_1 < N^{1/3} \leq p_2 < p_3\},\end{aligned}$$

so that $|\mathcal{B}|$ refers to the number of representations of $N - p$ as the sum of a prime and a product of exactly three primes.

It is an upper bound for the number of primes in $\mathcal{B}$ which is 'taken away' from the number of representations $N = p + P_3$ to provide our lower bound for the number of almost prime in $\mathcal{A}$.

Conjecture (Twin Primes). *There exist infinitely many primes p such that $p+2$ is prime.*

Theorem (Chen Jing-Run, 1974). *Let h be an even natural number. Then there exist infinitely many primes p such that*

$$p + h = P_2.$$

Theorem (Brun, 1912). *The sum*

$$\sum_{\substack{p \\ p+2 \text{ prime}}} \frac{1}{p} + \frac{1}{p+2}$$

is convergent, its value being referred to as Brun's constant.

1995: Thomas Nicely computed prime twins up to 10^{14}, and found a bug in the Intel Pentium!

Conjecture (Euler, 1752). *There exist infinitely many primes p of the form $p = x^2 + 1$.*

Theorem (Dirichlet, 1837). *If a, b are coprime integers, then there exist infinitely many primes p of the form*

$$p = ax + b.$$

Hypothesis H (Schinzel, Sierpinski, 1958). *Let $F_1(x), \ldots, F_n(x)$ be distinct irreducible polynomials with integer coefficients, then under a certain condition on the product, there exist infinitely many x such that each $F_i(x)$ is prime.*

Eratosthenes–Legendre:
primes in $\mathcal{A} := \{n : n \leq N\}$. Introduced $N_k := |\mathcal{A}_k|$ with $\mathcal{A}_k := \{n \leq N : k|n\}$. Approximated N_k by N/k and found $R_k = N_k - N/k$ is bounded by $|R_k| \leq 1$.

Theorem (Halberstam & Richert, 1972). *Let $\mathcal{A}$ be a set of integers with $|\mathcal{A}| \approx X$. Then, under certain conditions, one can find constants $r \in \mathbb{N}$, κ, $\delta > 0$ and $C \geq 1$ such that*

$$|\{P_r : P_r \in \mathcal{A}\}| \geq \delta \frac{X}{(\log X)^\kappa} \left(1 - \frac{C}{\sqrt{\log X}}\right).$$

The crucial condition in the determination of the least number of almost primes r is a good bound for the error term R_k. We look for a condition something like:

$$\sum_{d<X^\alpha} |R_k| \leq C \frac{X}{(\log X)^\kappa}.$$

If we can find an estimate with a large value of α, then we may correspondingly use a small value of r.

Theorem (H&R, 1972). *Let Q_1, Q_2 be irreducible quadratic polynomials over the integers such that Q_1Q_2 has no fixed prime divisor, then there exist infinitely many integers n such that*

$$Q_1(n)Q_2(n) = P_9.$$

Theorem (Iwaniec, 1972). *Let $F(x, y)$ be a quadratic polynomial. Then, under a certain simple condition on the coefficients, the number of primes $p \leq N$ represented by $F(x, y)$ is of order $N/(\log N)^{3/2}$.*

Theorem (M, 2005). *Let q_1, q_2 be irreducible binary quadratic forms over the integers, then subject to certain conditions on the forms, then there exist infinitely many pairs of integers (n, m) such that*

$$q_1(n,m)q_2(n,m) = P_6.$$

Old problem: investigate the variety

$$V: \begin{array}{l} q_1(x,y) = u^2 + v^2 \\ q_2(x,y) = s^2 + t^2 \end{array}$$

Count $N(X)$, the points $(x, y, u, v, s, t) \in \mathbb{Z}^6$ with $|x|, |y| < X$, and derive asymptotic formula as $X \to \infty$.

In calculating the asymptotic formula, one needs to evaluate quantities of the type

$$\sum_{k<X} |R_k|,$$

as in Halberstam and Richert's theorem.

More Questions:

Extend results of Marasingha? Pairs of irreducible cubic forms? Triples of quadratic forms?

There are sieve methods for calculating $\pi(N)$ in time $O(N^{2/3+\epsilon})$, which don't require the computation of all the primes, but it seems that to compute $\pi_2(N) :=$ the number of twin primes $\leq N$, we need to compute all the twin primes, taking time $O(N)$. Can we improve on this?