Lecture 6

January 15, 2018

1 Sieves

1.1 The sieve of Eratosthenes

The Inclusion-Exclusion principle, or the Möbius inversion formula, can be used – at least theoretically – to calculate $\pi(x)$. For a sufficiently large x, let us write

$$P = \prod_{p \le \sqrt{x}} p.$$

Then an integer n with $\sqrt{x} < n < x$ is prime if and only if $(n, P) = 1$. Thus, we can write

$$\pi(x) - \pi(\sqrt{x}) + 1 = \sum_{n \le x} E((n, P))) = \sum_{n \le x} \sum_{\substack{d|n \\ d|P}} \mu(d) = \sum_{d|P} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor, \quad (1)$$

where, as we know,

$$E(n) := \sum_{d|n} \mu(d)$$

is 1 if $n = 1$ and 0 otherwise (see Theorem 1 (i) of Lecture 4). If at this stage we insert the simple estimate

$$\left\lfloor \frac{x}{d} \right\rfloor = \frac{x}{d} + O(1)$$

in (1), we obtain

$$\pi(x) - \pi(\sqrt{x}) + 1 = x \prod_{p \le \sqrt{x}} \left(1 - \frac{1}{p}\right) + O(2^{\pi(\sqrt{x})}). \quad (2)$$

By the estimate of Problem 5 of Lecture 2, the first term of the right hand side of (2) is

$$\sim 2e^{-\gamma}\frac{x}{\log x} \qquad \text{as } x \to \infty,$$

while by Chebyshev's estimates, the error term in (2) can be seen to be larger than any power of x, thus showing that the error term in (2) can in fact be larger than the main term, thereby spoiling our goal to obtain something worthwhile by this approach.

The above calls for two comments. On the one hand, the exact formula (1) – called the *sieve formula of Eratosthenes* or at times the *Legendre formula* – involves too many terms for any reasonable practical estimate. On the other hand, the estimate of the main term itself shows, taking into account the Prime Number Theorem and the fact that $e^{-\gamma} \neq 1$, that the "error terms" created by replacing $\lfloor x/d \rfloor$ by x/d have made a global contribution of the same order of magnitude as the "main term". This suggests that this method, even suitably adapted, will never allow for a proof of the Prime Number Theorem. However, it can provide Chebyshev type estimates in a wide context.

In order to obtain a nontrivial result starting from formula (1), one may introduce a parameter y, $2 \leq y \leq x$, and bound $\pi(x) - \pi(y) + 1$ by the number of integers $n \leq x$ having no prime factor $p \leq y$. With the same calculations we get

$$\begin{aligned} \pi(x) &\leq x\prod_{p\leq y}\left(1-\frac{1}{p}\right) + O(2^y) \\ &= \frac{x(e^{-\gamma}+o(1))}{\log y} + O(2^y) \ll \frac{x}{\log\log x}, \end{aligned} \tag{3}$$

where we chose $y = \log x$.

With the aim of improving the efficiency of the above method, Viggo Brun invented the combinatorial sieve between 1917 and 1924.

2 The Brun pure sieve

The Eratosthenes sieve rests on the identity

$$E(n) = \sum_{d|n} \mu(d)$$

Brun's idea was to introduce two auxiliary functions μ_1 and μ_2 satisfying

$$\sum_{d|n} \mu_1(d) \leq 0 \leq \sum_{d|n} \mu_2(d) \tag{4}$$

for $n > 1$ (and $\mu_1(1) = \mu_2(1) = 1$) and vanishing often enough so that the number of nonzero terms in the resulting formula analogous with (1) is not overwhelming. Brun's initial choice lead to what is now called *Brun's pure sieve* and is the following.

Theorem 1. *Denote by χ_t the characteristic function of the set of integers n such that $\omega(n) \leq t$. Then for each integer $h \geq 0$, the functions defined by*

$$\mu_i(n) = \mu(n)\chi_{2h+2-i}(n) \qquad (i = 1, 2)$$

satisfy inequalities (4).

Proof. Since $\sum_{d|n} \mu_i(d)$ depends only on the kernel of n, we may assume that $\mu(n) \neq 0$. If $\omega(n) = k$, then, for each r with $0 \leq r \leq k$, it is clear that n has exactly $\binom{k}{r}$ divisors d with $\omega(d) = r$. For any given $t \geq 0$, we can thus write

$$\chi_t * 1(n) = \sum_{\substack{d \,|\, n \\ \omega(d) \leq t}} \mu(d) = \sum_{0 \leq r \leq t} (-1)^r \binom{k}{r} = (-1)^t \binom{k-1}{t},$$

where the last equality is easily obtained by induction over t. □

The above result immediately yields the following corollary.

Corollary 1. *Let $\mathcal{A}$ be a finite set of integers and let $\mathcal{P}$ be a set of prime numbers. Write*

$$\begin{aligned} \mathcal{A}_d &= \#\{a \in \mathcal{A} : a \equiv 0 \pmod{d}\}, \\ P(y) &= \prod_{\substack{p \leq y \\ p \in \mathcal{P}}} p, \\ \mathcal{S}(\mathcal{A}, \mathcal{P}, y) &= \#\{a \in \mathcal{A} : (a, P(y)) = 1\}. \end{aligned}$$

Then, for each integer $h \geq 0$,

$$\sum_{\substack{d \,|\, P(y) \\ \omega(d) \leq 2h+1}} \mu(d)\mathcal{A}_d \leq \mathcal{S}(\mathcal{A}, \mathcal{P}, y) \leq \sum_{\substack{d \,|\, P(y) \\ \omega(d) \leq 2h}} \mu(d)\mathcal{A}_d.$$

Let us see how the above result helps us to considerably improve the upper bound of $\pi(x)$ obtained by the Eratosthenes sieve (see (3)).

In Corollary 1, we chose $\mathcal{A} = \{n : n \le x\}$, $\wp = \{\text{all primes}\}$ and $P = P(y) = \prod_{p \le y} p$. Then $\mathcal{S}(\mathcal{A}, \mathcal{P}, y)$ is the number of positive integers $n \le x$ having no prime factor $p \le y$, so that

$$\begin{aligned}
\mathcal{S}(\mathcal{A}, \mathcal{P}, y) &\le \sum_{\substack{d \mid P(y) \\ \omega(d) \le 2h}} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor \\
&= x \sum_{\substack{d \mid P(y) \\ \omega(d) \le 2h}} \frac{\mu(d)}{d} + O\left(\sum_{\substack{d \mid P(y) \\ \omega(d) \le 2h}} 1 \right) \\
&= x \prod_{p \le y} \left(1 - \frac{1}{p}\right) + O\left(\sum_{\substack{d \mid P(y) \\ \omega(d) \le 2h}} 1 + x \sum_{\substack{d \mid P(y) \\ \omega(d) > 2h}} \frac{1}{d} \right), \qquad (5)
\end{aligned}$$

and similarly

$$\begin{aligned}
\mathcal{S}(\mathcal{A}, \mathcal{P}, y) &\ge \sum_{\substack{d \mid P(y) \\ \omega(d) \le 2h+1}} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor \\
&= x \prod_{p \le y} \left(1 - \frac{1}{p}\right) + O\left(\sum_{\substack{d \mid P(y) \\ \omega(d) \le 2h+1}} 1 + x \sum_{\substack{d \mid P(y) \\ \omega(d) > 2h+1}} \frac{1}{d} \right). \qquad (6)
\end{aligned}$$

The first of the two error terms appearing either at (5) or at (6) does not exceed y^{2h+1} since this is an upper bound for all integers d such that $d \mid P(y)$ and $\omega(d) \le 2h + 1$. The d-sums arising in the second error terms are bounded, namely, say for the second error term in (5), by

$$\sum_{\substack{d \mid P(y) \\ \omega(d) > 2h}} \frac{1}{d} \le \sum_{k > 2h} \frac{1}{k!} \left(\sum_{p \le y} \frac{1}{p} \right)^k \le \sum_{k > 2h} \frac{1}{k!} (\log\log y + c_0)^k.$$

Using the weak form of Stirling's formula $k! \ge (k/e)^k$, together with $y < x$, we get

$$\sum_{\substack{d \mid P(y) \\ \omega(d) > 2h}} \frac{1}{d} \le \sum_{k > 2h} \frac{1}{k!} (\log\log x + c_0)^k \le \sum_{k > 2h} \left(\frac{e \log\log x + e c_0}{k} \right)^k.$$

Choosing an integer $h \geq e \log\log x + ec_0$, we obtain

$$\sum_{\substack{d|P(y)\\ \omega(d)>2h}} \frac{1}{d} \leq \left(\frac{1}{2}\right)^{2h} \left(1 + \frac{1}{2} + \frac{1}{4} + \cdots\right) \ll \frac{1}{(\log x)^{2e\log 2}} \ll \frac{1}{(\log x)^2},$$

because $2e\log 2 = e\log 4 > e > 2$. For this choice of h, we impose that $y^{2h+1} \leq x/(\log x)^2$, which for $h > 1$ is implied by

$$y \leq \frac{x^{1/(2h+1)}}{\log x} \leq \exp\left(\frac{\log x}{2e\log\log x + c_1} - \log\log x\right),$$

where we can take $c_1 = 2ec_0 + 1$. Since $1/2e > 1/10$, it follows that we may choose

$$y = \exp\left(\frac{\log x}{10\log\log x}\right), \tag{7}$$

in which case the inequality $y^{2h+1} \ll x/(\log x)^2$ holds for all x. With this choice of y, we have that

$$\prod_{p\leq y}\left(1 - \frac{1}{p}\right) \asymp \frac{1}{\log y} \asymp \frac{\log\log x}{\log x},$$

while the error terms in (5) and (6) are $O(x/(\log x)^2)$. Thus, we have proved that

$$\mathcal{S}(\mathcal{A}, \mathcal{P}, y) = x\prod_{p\leq y}\left(1 - \frac{1}{p}\right)\left(1 + O\left(\frac{1}{\log y}\right)\right). \tag{8}$$

Since

$$\mathcal{S}(\mathcal{A}, \mathcal{P}, y) \geq \pi(x) - \pi(y) \geq \pi(x) - y \geq \pi(x) + O(x^{1/2}),$$

we immediately deduce that

$$\pi(x) \ll x^{1/2} + x\prod_{p\leq y}\left(1 - \frac{1}{p}\right) \ll \frac{x\log\log x}{\log x},$$

which, although much weaker than Chebyshev's estimate, is remarkable because of the simplicity and generality of the argument.

To summarize, we have just proved the following result:

Theorem 2. *Letting*

$$\Phi(x, y) = \#\{n \leq x : p(n) > y\},$$

then, for $y \leq \exp\left(\dfrac{\log x}{10 \log\log x}\right)$,

$$\Phi(x,y) = x \prod_{p \leq y} \left(1 - \frac{1}{p}\right) \left\{1 + O\left(\frac{1}{\log y}\right)\right\}.$$

3 Twin Primes

Now we expose another remarkable application of Brun's pure sieve, namely the fact that the sum of the reciprocal of the twin primes is convergent.

Proposition 1. *Let* $\mathcal{J} = \{p : \ p$ *and* $p+2$ *are both primes*$\}$ *and set* $\mathcal{J}(x) = \#\{p \leq x : p \in \mathcal{J}\}$. *Then*

$$\mathcal{J}(x) \ll \frac{x(\log\log x)^2}{(\log x)^2}.$$

Proof. In Corollary 1, set $\mathcal{A} = \{n(n+2) : n \leq x\}$. Again, let $\mathcal{P}$ stand for the set of all primes and let y be a parameter to be chosen later. To understand $\#\mathcal{A}_d$, we look at

$$\rho(d) = \#\{0 \leq n \leq d-1 \ : \ n(n+2) \equiv 0 \pmod{d}\}.$$

Let us first show that $\rho(d)$ is multiplicative. Indeed, if u and v are coprime and $c \pmod{uv}$ is some congruence class modulo uv such that $n(n+2) \equiv 0 \pmod{uv}$, then certainly $c \pmod{u}$ ($c \pmod{v}$, respectively) is a congruence class modulo u (modulo v, respectively) such that $n(n+2) \equiv 0 \pmod{u}$ ($n(n+2) \equiv 0 \pmod{v}$, respectively). Conversely, if $a \pmod{u}$ and $b \pmod{v}$ are congruence classes for n modulo u and v which are solutions to $n(n+2) \equiv 0 \pmod{u}$ and $n(n+2) \equiv 0 \pmod{v}$, respectively, then by the Chinese Remainder Theorem, there exists a class $c \pmod{uv}$ (which is unique) such that $c \equiv a \pmod{u}$ and $c \equiv b \pmod{v}$. Hence, $n(n+2) \equiv 0 \pmod{u}$ and $n(n+2) \equiv 0 \pmod{v}$, and since u and v are coprime, we get that $n(n+2) \equiv 0 \pmod{uv}$. This shows that $\rho(uv) = \rho(u)\rho(v)$. Note that $\rho(2) = 1,\ \rho(4) = 2,\ \rho(2^k) = 4$ for $k \geq 3$ and $\rho(p^k) = 2$ if $p > 2$ is odd. In particular, if d is squarefree, then $\rho(d) = 2^{\omega(d)}$ if d is odd and $\rho(d) = 2^{\omega(d)-1}$ if d is even. Since there are precisely $\rho(d)$ solutions n to the congruence $n(n+2) \equiv 0 \pmod{d}$ in any interval of length d, and since the interval $[1,x]$ is made up of $\lfloor x/d \rfloor$ intervals of length d and (maybe) one shorter interval, we get that

$$\begin{aligned}\mathcal{A}_d &= \#\{n \le x \,:\, d \mid n(n+2)\} = \rho(d)\left(\left\lfloor \frac{x}{d} \right\rfloor + O(1)\right)\\ &= \frac{x\rho(d)}{d} + O(\rho(d)) = \frac{x\rho(d)}{d} + O(2^{\omega(d)}).\end{aligned}$$

Upon noting that if $p, p+2$ are twin primes, then either $p \le y$ or $p \in \mathcal{S}(\mathcal{A}, \mathcal{P}, y)$, we have, by Corollary 1, that

$$\begin{aligned}\mathcal{J}(x) &\le \pi(y) + \sum_{\substack{d \mid P(y)\\ \omega(d) \le 2h}} \mu(d)\mathcal{A}_d\\
&= \sum_{\substack{d \mid P(y)\\ \omega(d) \le 2h}} \mu(d)\left(\frac{x\rho(d)}{d} + O(2^{\omega(d)})\right) + O(y)\\
&= x \sum_{\substack{d \mid P(y)\\ \omega(d) \le 2h}} \frac{\mu(d)\rho(d)}{d} + O\left(y + \sum_{\substack{d \mid P(y)\\ \omega(d) \le 2h}} 2^{\omega(d)}\right)\\
&= x \sum_{d \mid P(y)} \frac{\mu(d)\rho(d)}{d} + O\left(y + 2^{2h} \sum_{\substack{d \mid P(y)\\ \omega(d) \le 2h}} 1 + x \sum_{\substack{d \mid P(y)\\ \omega(d) > 2h}} \frac{2^{\omega(d)}}{d}\right)\\
&= x \prod_{p \le y}\left(1 - \frac{\rho(p)}{p}\right) + O\left(y + 2^{2h} \sum_{\substack{d \mid P(y)\\ \omega(d) \le 2h}} 1 + x \sum_{\substack{d \mid P(y)\\ \omega(d) > 2h}} \frac{2^{\omega(d)}}{d}\right)\\
&= \frac{x}{2} \prod_{3 \le p \le y}\left(1 - \frac{2}{p}\right) + O\left(y + (2y)^{2h} + x \sum_{\substack{d \mid P(y)\\ \omega(d) > 2h}} \frac{2^{\omega(d)}}{d}\right). \qquad (9)\end{aligned}$$

Using the combinatorial fact that

$$\sum_{\substack{d \mid P(y)\\ \omega(d) > 2h}} \frac{2^{\omega(d)}}{d} \le \sum_{k > 2h} \sum_{\substack{d \mid P(y)\\ \omega(d) = k}} \frac{2^k}{d} \le \sum_{k > 2h} \frac{1}{k!}\left(\sum_{p \le y} \frac{2}{p}\right)^k,$$

together with Mertens' formula and the weak Stirling estimate, we get

$$\sum_{\substack{d\,|\,P(y)\\ \omega(d)>2h}} \frac{2^{\omega(d)}}{d} < \sum_{k>2h} \frac{1}{k!}(2\log\log x + 2c_0)^k < \sum_{k>2h} \left(\frac{2e\log\log x + c_1}{k}\right)^k,$$

where $c_1 = 2ec_0$. Hence, we see that if we choose h to be twice as large as in the proof of Theorem 2, that is, the minimal positive integer h larger than $2e\log\log x + c_1$, we then get

$$\sum_{\substack{d\,|\,P(y)\\ \omega(d)>2h}} \frac{2^{\omega(d)}}{d} < \frac{1}{2^{2h}}\left(\sum_{l\geq 0}\frac{1}{2^l}\right) = \frac{2}{2^{2h}} \ll \frac{1}{(\log x)^{4e\log 2}} < \frac{1}{(\log x)^2}. \tag{10}$$

Choosing $y = \exp(\log x/(20\log\log x))$, we obtain that

$$(2y)^{2h} < 2^{2h}\exp\left(\frac{2h\log x}{20\log\log x}\right) = \exp\left(\frac{(1+o(1))4e\log x}{20}\right) \ll \frac{x}{(\log x)^2}. \tag{11}$$

Inserting estimates (10) and (11) into (9), we get

$$\mathcal{J}(x) \ll x\prod_{p\leq y}\left(1-\frac{2}{p}\right) + \frac{x}{(\log x)^2}.$$

Finally, using Problem 7 of Lecture 2 with $\kappa = -2$, we have that

$$\begin{aligned}\prod_{p\leq y}\left(1-\frac{2}{p}\right) &= \frac{c_2}{(\log y)^2}(1+o(1)) = c_2\left(\frac{20\log\log x}{\log x}\right)^2(1+o(1))\\ &= (1+o(1))\frac{400c_2(\log\log x)^2}{(\log x)^2},\end{aligned}$$

so that

$$\mathcal{J}(x) \ll \frac{x(\log\log x)^2}{(\log x)^2},$$

which is what we wanted to prove. $\square$

Corollary 2. *The series*

$$\sum_{p,\ p+2\ \text{primes}} \frac{1}{p} < \infty.$$

Proof. Since

$$\mathcal{J}(n) - \mathcal{J}(n-1) = \begin{cases} 1 & \text{if } n \text{ and } n+2 \text{ are both primes,} \\ 0 & \text{otherwise,} \end{cases}$$

then, in light of Proposition 1,

$$\begin{aligned} \sum_{p,p+2 \text{ primes}} \frac{1}{p} &= \sum_{n=2}^{\infty} \frac{\mathcal{J}(n) - \mathcal{J}(n-1)}{n} = \sum_{n=1}^{\infty} \mathcal{J}(n) \left(\frac{1}{n} - \frac{1}{n+1} \right) \\ &= \sum_{n=1}^{\infty} \frac{\mathcal{J}(n)}{n(n+1)} \ll \sum_{n \geq e}^{\infty} \frac{n(\log\log n)^2}{(\log^2 n)\, n(n+1)} \\ &< \sum_{n \geq e}^{\infty} \frac{(\log\log n)^2}{n(\log n)^2} \ll \int_e^{\infty} \frac{(\log\log t)^2}{t \log^2 t}\, dt < \infty, \end{aligned}$$

as requested. □