Small Sieves, Part I: The Brun Sieve

1. Introduction. In this set of notes we introduce some ideas of Sieve Theory by looking at one of the first, and easiest, applications of the method. This is Brun's proof that there are not too many twin primes. Brun's method is appealing in that it is based on a strong combinatorial intuition, but subsequent methods, particularly the Selberg Sieve, lead to better results.

Our attack on the problem of finding an upper bound for $\pi_2(N)$ (the number of twin primes less than N) rests on the rather obvious observation that, if p is a twin prime, the number $f(p) = p(p+2)$ does not have any small prime factors. Let $\mathcal{A}$ denote the sequence $f(1), \ldots, f(N)$, and write $\mathcal{A}_d$ for the number of elements of $\mathcal{A}$ which are divisible by d. Let

$$\mathcal{P} \;=\; \{p_1 < p_2 < \ldots\}$$

be the set of prime numbers. Then the inclusion-exclusion principle gives us an exact formula for $S(\mathcal{A}, \mathcal{P}, z)$, the number of elements of $\mathcal{A}$ with no prime divisor less than z:

$$S(\mathcal{A}, \mathcal{P}, z) \;=\; \sum_{s=0}^{r}(-1)^s \sum_{1 \le i_1 < \cdots < i_s \le r} \mathcal{A}_{p_{i_1} p_{i_2} \cdots p_{i_s}}, \tag{1}$$

where p_r is the greatest prime not exceeding z. The main obstacle between us and an accurate estimation of this sum is our poor knowledge of $\mathcal{A}_d$ when d is large. When $d = p$ is prime we fairly clearly have

$$\mathcal{A}_p \;=\; \frac{\epsilon(p)N}{p} + R_p$$

where $\epsilon(2) = 1$, $\epsilon(p) = 2$ for $p > 2$ and $|R_p| \le 2$. This extends fairly easily to give

$$\mathcal{A}_{p_{i_1} p_{i_2} \cdots p_{i_s}} \;=\; \frac{\epsilon(p_{i_1}) \ldots \epsilon(p_{i_s})}{p_{i_1} p_{i_2} \ldots p_{i_s}} N + R_{p_{i_1} p_{i_2} \cdots p_{i_s}}, \tag{2}$$

where $|R_{p_{i_1} p_{i_2} \cdots p_{i_s}}| \le 2^s$. Any attempt to substitute this into (1) is fairly clearly doomed to failure, because the contributions from the error terms R will dominate if z is at all large (unless there is some cancellation, but it is very unclear how one might prove that).

2. Brun's Method. Brun's brilliantly simple way of circumventing the problem of large error terms rests on the well known observation (sometimes referred to in the probability literature as Bonferroni's Inequality) that if we truncate the formula (1) at an *even* index $t = 2k$, we get an upper bound for $S(\mathcal{A}, \mathcal{P}, z)$. In other words

$$S(\mathcal{A}, \mathcal{P}, z) \;\le\; \sum_{s=0}^{t}(-1)^s \sum_{1 \le i_1 < \cdots < i_s \le r} \mathcal{A}_{p_{i_1} p_{i_2} \cdots p_{i_s}} \tag{3}$$

for any $0 \le t = 2k \le r$. Substituting (2) into this gives

$$S(\mathcal{A}, \mathcal{P}, z) \;\le\; N \sum_{s=0}^{t}(-1)^s \sum_{1 \le i_1 < \cdots < i_s \le r} \frac{\epsilon(p_{i_1}) \ldots \epsilon(p_{i_s})}{p_{i_1} p_{i_2} \ldots p_{i_s}} \;+\; \sum_{s=0}^{t} \binom{r}{s} 2^s. \tag{4}$$

The remainder of the argument consists, essentially, of playing about with this expression and choosing an optimal value of k. To see what our first manipulation should be, let us conduct a short heuristic experiment to guess at the value of $S(\mathcal{A}, \mathcal{P}, z)$. Assuming that p divides $n \in \mathcal{A}$ with probability $\epsilon(p)/p$, and that these events are independent over different p (these assumptions amount to assuming that the error R in (2) is zero), we have

$$S(\mathcal{A}, \mathcal{P}, z) \;=\; N \prod_{p<z} \left(1 - \frac{\epsilon(p)}{p}\right).$$

It would be nice, then, to have something like $N \prod_{p<z} \left(1 - \frac{\epsilon(p)}{p}\right)$ as the "main term" of (4). This can be achieved quite easily by adding in and taking off some terms, then using the triangle inequality. Indeed (2) gives

$$\begin{aligned} S(\mathcal{A}, \mathcal{P}, z) \;&\leq\; N \sum_{s=0}^{r} (-1)^s \sum_{1 \leq i_1 < \cdots < i_s \leq r} \frac{\epsilon(p_{i_1}) \ldots \epsilon(p_{i_s})}{p_{i_1} p_{i_2} \ldots p_{i_s}} \\ &\qquad + N \sum_{s=t+1}^{r} \sum_{1 \leq i_1 < \cdots < i_s \leq r} \frac{2^s}{p_{i_1} p_{i_2} \ldots p_{i_s}} \;+\; \sum_{s=0}^{t} \binom{r}{s} 2^s \\ &=\; N \prod_{p<z} \left(1 - \frac{\epsilon(p)}{p}\right) \;+\; N \sum_{s=t+1}^{r} \sum_{1 \leq i_1 < \cdots < i_s \leq r} \frac{2^s}{p_{i_1} p_{i_2} \ldots p_{i_s}} \;+\; \sum_{s=0}^{t} \binom{r}{s} 2^s. \\ &=\; N \prod_{p<z} \left(1 - \frac{\epsilon(p)}{p}\right) \;+\; E_1 \;+\; E_2. \end{aligned} \tag{5}$$

We must now estimate the error terms E_1 and E_2.

Lemma 1 *Suppose that $t \geq 8e \log\log z$ and that z is sufficiently large. Then*

$$E_1 \;\leq\; 2N \left(\frac{4e \log\log z}{t}\right)^t.$$

Proof We begin by observing that

$$\sum_{1 \leq i_1 < \cdots < i_s \leq r} \frac{2^s}{p_{i_1} p_{i_2} \ldots p_{i_s}} \;\leq\; \frac{1}{s!} \left(\frac{2}{p_1} + \cdots + \frac{2}{p_r}\right)^s. \tag{6}$$

A standard result from elementary prime number theory gives that $\sum_{p \leq z} p^{-1} \leq 2 \log\log z$ provided that z is greater than some absolute constant. Using this in (6) gives

$$E_1 \;\leq\; N \sum_{s>t} \frac{(4e \log\log z)^s}{s!}.$$

It is easy to see that, for $t \geq 8e \log\log z$, the ratio of each term in this sum to the next is at least 2. Combining this obesrvation with the inequality $s! \geq (s/e)^s$ gives the required estimate. $\square$

Lemma 2 *Let z be sufficiently large and suppose that $t < r$. Then*

$$E_2 \ \leq \ 2\left(\frac{4ez}{t\log z}\right)^t.$$

Proof The inequality $\binom{r}{s} \leq r^s/s!$ gives

$$E_2 \ \leq \ \sum_{s=0}^{t} \frac{(2r)^s}{s!}.$$

If $t < r$ then one may check that the ratio of each term in this sum to the next is at most $1/2$. This, together with the observation that $r \leq \frac{2z}{\log z}$ for z sufficiently large, gives the stated result. $\square$

In our final lemma we give an upper bound for the so-called "main term" in (5).

Lemma 3

$$\prod_{p<z}\left(1 - \frac{\epsilon(p)}{p}\right) \ \ll \ (\log z)^{-2}.$$

Proof We have

$$\begin{aligned}
\log \prod_{p<z}\left(1 - \frac{\epsilon(p)}{p}\right) \ &= \ -\log 2 + \sum_{3\leq p<z} \log\left(1 - \frac{2}{p}\right) \\
&\leq \ -\log 2 - 2\sum_{3\leq p<z} p^{-1} \\
&\leq \ -2\log\log z + C.
\end{aligned}$$

This completes the proof. $\square$

Now if we could take z to be a power of N whilst keeping the error terms E_1 and E_2 sufficiently small we could get a bound

$$S(\mathcal{A}, \mathcal{P}, z) \ \ll \ N(\log N)^{-2}.$$

Since any twin prime lies in $S(\mathcal{A}, \mathcal{P}, z)$ this immediately gives the bound

$$\pi_2(N) \ \ll \ N(\log N)^{-2},$$

which agrees with the obvious heuristic bound up to a multiplicative constant. Unfortunately this turns out not to be possible. Working out the best possible values for z and t is a rather tedious business. A reasonably good choice is $z = N^{(\log\log N)^{-2}}$, $t = (\log\log N)^2$. We leave it to the reader to check that the conditions of the three lemmas are satisfied for large N, and that one winds up with a bound

$$\pi_2(N) \ \ll \ N(\log N)^{-2+\epsilon}$$

for any $\epsilon > 0$. This is already enough to give Brun's famous result that the sum of the reciprocals of the twin primes converges.

The above discussion was devoted to finding an upper bound for $S(\mathcal{A}, \mathcal{P}, z)$, and it is natural to ask whether the same method could give a lower bound for the same quantity. The answer is that it can, and it is only necessary to make minimal changes to the argument. Such a lower bound would translate into lower bounds for the number of $n \leq N$ for which $n(n+2)$ has at most, say, $(\log \log N)^2$ prime factors. Such a statement is interesting, but we do not yet know anything about the number of $n \leq N$ for which $n(n+2)$ has, say, at most 2^{1000} prime factors. The Brun Sieve in the form we have developed it is incapable of answering such a question, and we must turn to the Selberg Sieve. As we shall see Selberg's upper bound for $S(\mathcal{A}, \mathcal{P}, z)$ is not particularly hard to understand. Unfortunately, however, the more interesting lower bound lies rather deeper and does not bear such close resemblance to the upper bound as in Brun's method.

References In preparing this exposition I used Nathanson's *Additive Number Theory: The Classical Bases* as well as (unpublished) lecture notes from a course by Tim Gowers.