Prime k-tuples

This unit begins the third part of the course, in which we apply the results gathered in the first two parts in order to say something about the extent to which primes cluster together in short intervals.

Reference cited below: P.X. Gallagher, On the distribution of primes in short interval, *Mathematika* **23** (1976), 4–9; corrigendum, *ibid.* **28** (1981), 86. (I couldn't find this online.)

1 The Hardy-Littlewood k-tuples conjecture

Let $\mathcal{H}$ denote a k-tuple of distinct integers. What does one expect about the distribution of the integers n such that $n + h$ is prime for each $h \in \mathcal{H}$?

Here is a rather simple-minded guess. The prime number theorem suggests that if one chooses a random integer of size x, it will be prime with probability $1/(\log x)$. If one then chooses k distinct integers of size x, and there is no obvious reason why they cannot all be prime, then one might expect them to be simultaneously prime with probability $\log^{-k} x$, and the number of such tuples with terms bounded by x should be asymptotic to $x \log^{-k} x$, with the constant 1.

However, this turns out not to be the correct constant, as is easily verified against experimental evidence in the case of twin primes. The reason is perhaps obvious: the facts that the different $n+h$ are coprime to a fixed prime p are not independent, and one needs to account for this. Here is the recipe for doing so proposed by Hardy-Littlewood (and mentioned by Ben Green in his guest lecture).

Fix a prime p. The probability that k randomly chosen integers are all not divisible by p is $(1 - 1/p)^k$. On the other hand, the probability that the $n + h$ are all coprime to p is $1 - v_{\mathcal{H}}(p)/p$, where $v_{\mathcal{H}}(p)$ is the number of residue classes modulo p represented by elements of $\mathcal{H}$. We thus set

$$\mathfrak{S}(\mathcal{H}) = \prod_p \left(1 - \frac{v_{\mathcal{H}}(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}$$

(called the "singular series", because it occurs in the Hardy-Littlewood circle method as a series summed over singularities of some integral) and conjecture as follows.

Conjecture 1 (Hardy-Littlewood)**.** *Suppose that $v_{\mathcal{H}}(p) < p$ for all p. Then the number of integers $n \leq x$ such that $n + h$ is prime for each $h \in \mathcal{H}$ is asymptotic to $\mathfrak{S}(\mathcal{H}) x \log^{-k} x$.*

Of course if $v_{\mathcal{H}}(p) = 0$ for some p, then there is a trivial obstruction created by divisibility mod p, so you only get finitely many prime k-tuples of that shape. On the other hand, if $v_{\mathcal{H}}(p) < p$ for all p, then the product converges absolutely and so $\mathfrak{S}(\mathcal{H}) > 0$.

Convention: it will be convenient later to take the same definition for $\mathfrak{S}(\mathcal{H})$ even if $\mathcal{H}$ does not have distinct entries.

2 k-tuples and prime gaps

If one is only interested in looking for primes which are close together, without specifying exactly what the gaps are, one could go back to the probabilistic model (attributed to Cramér, more famous for his rule for solving linear systems). It suggests that the distribution of $\pi(n+h) - \pi(n)$, for $n \leq N$ and $h \sim \lambda \log N$ with λ fixed, should approach a Poisson distribution with parameter λ as $N \to \infty$. The fact that this follows from a suitably uniform version of the k-tuples conjecture is due to Gallagher; the main part of the argument is the following result, which we will need later.

Theorem 2 (Gallagher)**.** *We have*

$$\sum_{\mathcal{H} \in \{1,\dots,x\}^k} \mathfrak{S}(\mathcal{H}) \sim x^k.$$

In other words, the fudge factor $\mathfrak{S}(\mathcal{H})$ between the probabilistic model and the Hardy-Littlewood prediction averages out to 1, so the prediction based on the probabilistic model is consistent with Hardy-Littlewood. (Note: the contribution from tuples not having distinct entries is $O(x^{k-1})$, so it doesn't matter whether we include them or not.)

Here is a sketch of Gallagher's proof, with the missing details left as exercises. (Throughout, keep k fixed.) Put

$$a(p,m) = \left(1 - \frac{m}{p}\right)\left(1 - \frac{1}{p}\right)^{-k} - 1$$
$$a_{\mathcal{H}}(p) = a(p, v_{\mathcal{H}}(p))$$

so that

$$\mathfrak{S}(\mathcal{H}) = \prod_p (1 + a_{\mathcal{H}}(p)).$$

Extend a by multiplicativity to squarefree arguments d, so that

$$\mathfrak{S}(\mathcal{H}) = \sum_d a_{\mathcal{H}}(d)$$

with the sum on the right being absolutely convergent.

We can truncate the sum over d by showing that for each fixed $\epsilon > 0$,

$$\sum_{\mathcal{H} \in \{1,\dots,x\}^k} \mathfrak{S}(\mathcal{H}) = \sum_{d \leq y} \sum_{\mathcal{H}} a_{\mathcal{H}}(d) + O(x^k (xy)^\epsilon / y) \qquad (1)$$

with the constant depending only on k, ϵ and not on x, y (exercise).

For any given d, we can rewrite the inner sum of (1) as a sum

$$\sum_v \left(\prod_{p|d} a(p, v(p)) \right) f_d(x, v),$$

where v runs over vectors indexed by the prime factors of d, with $v(p) \in \{1, \ldots, p\}$ for each $p|d$, and $f_d(x, v)$ counts k-tuples $\mathcal{H} \in \{1, \ldots, x\}^r$ which occupy exactly $v(p)$ residue classes modulo p for each $p|d$.

Write $\left\{ {a \atop b} \right\}$ for the number of partitions of an a-element set into b unordered parts (Stirling number of the second kind). If we set

$$A(d) = \sum_v \prod_{p|d} a(p, v(p)) \binom{p}{v(p)} v(p)! \left\{ {k \atop v(p)} \right\}$$

$$B(d) = \sum_v \prod_{p|d} |a(p, v(p))| \binom{p}{v(p)} v(p)! \left\{ {k \atop v(p)} \right\}$$

$$C(d) = \sum_v \prod_{p|d} |a(p, v(p))|,$$

then

$$\sum_{\mathcal{H}} a_{\mathcal{H}}(d) = (x/d)^k A(d) + O((x/d)^{k-1} B(d)) + O(x^{k-1} C(d)). \quad (2)$$

From this, plus the identities

$$\sum_{v=1}^{p} \binom{p}{v} v! \left\{ {k \atop v} \right\} = p^k \quad (3)$$

$$\sum_{v=1}^{p} v \binom{p}{v} v! \left\{ {k \atop v} \right\} = p^{k+1} - (p-1)^k p, \quad (4)$$

it is not difficult to deduce Theorem 2.

Exercises

1. Prove that for k a positive integer,

$$\int_2^x \log^{-k} t \, dt \sim x \log^{-k} x.$$

2. Prove (1). (If you get stuck, see the hint for problem 5.)

3. Prove (2). (Hint: it might help to think in terms of counting lattice points.)

4. Prove the identities (3), (4).

5. Complete the proof of Theorem 2 from (1) and (2). (Hint: first use the Stirling number identities to calculate $A(d)$. Then estimate $B(d)$ and $C(d)$, using the bound

$$|a(p, m)| \leq \begin{cases} c(k)(p-1)^{-2} & m = k \\ c(k)(p-1)^{-1} & m < k. \end{cases}$$

That is, the constant $c(k)$ depends on k but not on p or m. Finally, take $y = x^{1/2}$ in (1).)