Sieving an interval

It is interesting that our sieving apparatus that we have developed is robust enough to handle some problems on the distribution of prime numbers that would otherwise be unattackable. For example, we saw that if

\[ A = \{ \text{integers } n : n \in (0,x) \}, \quad P = \text{the set of primes}, \quad y = x^{1/(10 \log \log x)}, \]

then

\[ S(A, P, y) \ll \frac{x}{\log y} \ll \frac{x \log \log x}{\log x}. \]

As a corollary we have

\[ \pi(x) \ll \frac{x \log \log x}{\log x}. \]

Now say we make a very small change. We alter the sequence \( A \) to

\[ A' = \{ \text{integers } n : n \in (z, z+x) \}, \]

and we leave everything else unchanged. What allowed us to estimate \( S(A, P, y) \) was the estimate

\[ A_d = \sum_{a \in A \atop d|a} 1 = \left\lfloor \frac{x}{d} \right\rfloor = \frac{x}{d} + O(1). \]

What about \( A'_d \)? This is

\[ A'_d = \sum_{a \in A' \atop d|a} 1 = \left\lfloor \frac{x+z}{d} \right\rfloor - \left\lfloor \frac{z}{d} \right\rfloor = \frac{x+z}{d} - \frac{z}{d} + O(1) = \frac{x}{d} + O(1). \]

That is, we have essentially the same estimate for \( A'_d \) as we had for \( A_d \). So, exactly the same calculation as before gives us

\[ S(A', P, y) \ll \frac{x}{\log y} \ll \frac{x \log \log x}{\log x}. \]

And we have the corollary that

\[ \pi(z+x) - \pi(z) \ll \frac{x \log \log x}{\log x}. \]

All of our sieve results work in this more general way, allowing us to sift an interval.
2 The prime convexity conjecture

In a paper from 1923, Hardy and Littlewood laid out a great many conjectures on the distribution of prime numbers, giving them an analytic cast. For example, instead of just conjecturing that there are infinitely many twin primes, they conjectured that

$$\pi_2(x) \sim 2C_2 \int_2^x \frac{dt}{(\log t)^2},$$

where $C_2$ is the twin prime constant discussed earlier. They had similar strong conjectures for prime $k$-tuples, and they also had a strong version of Goldbach’s conjecture (that even numbers starting at 4 are the sum of two primes). They also presented a new conjecture, now known as the prime convexity conjecture:

If $x, z \geq 2$, then $\pi(z + x) \leq \pi(z) + \pi(x)$.

That is, the champion (or co-champion) interval of length $x$ for having prime numbers is the initial interval $(0, x]$. The interval $(z, z + x]$ has $\pi(z + x) - \pi(z)$ primes, and the prime convexity conjecture asserts that this count cannot exceed $\pi(x)$, the number of primes in $(0, x]$.

It came as somewhat of a shock in 1973 when Hensley and Richards showed that the prime convexity conjecture is incompatible with the prime $k$-tuples conjecture. Here is their idea.

Suppose that $\{a_1, a_2, \ldots, a_k\}$, $0 < a_1 < a_2 < \cdots < a_k$ is admissible. Recall that this means that for every prime $p$, the $a_i$’s do not contain a complete residue system modulo $p$. The prime $k$-tuples conjecture then asserts that there are infinitely many integers $n$ with $n + a_1, n + a_2, \ldots, n + a_k$ all prime. Each time this happens we have

$$\pi(n + a_k) - \pi(n) \geq k.$$ 

What Hensley and Richards found was a large admissible set with $a_k < p_k$, where $p_k$ denotes the $k$-th prime. The prime convexity conjecture with $z = n$ and $x = a_k$ would assert that

$$k \leq \pi(z + x) - \pi(z) \leq \pi(x) = \pi(a_k) \leq \pi(p_k - 1) = k - 1,$$

a contradiction.

So, which conjecture is to be believed? Most people seem to put their faith (if you want to call it that) in the prime $k$-tuples conjecture. The admissible set that Hensley and Richards came up with is quite concrete, though fairly large. To disprove the prime convexity conjecture, all one would need to do is to find some one integer $n$ where all of $n + a_i$ are prime. However, this is not such an easy task, since the admissible set is fairly large. People have also looked at trying to reduce the size of the admissible set in the proof, and there has been some success there. I believe it is the case still that none of the conjectures in the original Hardy and Littlewood paper have been settled. However, we do know for sure that at least one conjecture in the paper is false!
3 Periodic functions

We have seen that a Dirichlet character $\chi \mod k$ is a periodic function with period $k$ that is also completely multiplicative. Let us look now more generally at the set $\mathcal{P}_k$ of all arithmetic functions $f : \mathbb{Z} \to \mathbb{C}$ which are periodic with period $k$. It is clear that $\mathcal{P}_k$ is closed under addition of functions and under scalar multiplication (with the field of scalars being $\mathbb{C}$). It is a complex vector space. Since the values $f(1), f(2), \ldots, f(k)$ determine $f$ and may be arbitrarily prescribed, it is clear that $\mathcal{P}_k$ has dimension $k$. In particular, it is naturally isomorphic to $\mathbb{C}^k$, the vector space of ordered $k$-tuples of complex numbers.

We look now for a convenient basis for the vector space $\mathcal{P}_k$. We of course have the natural one corresponding to the standard basis of $\mathbb{C}^k$. But we would like a basis that more overtly reflects the periodicity modulo $k$. Mathematics has long used the trig functions sin and cos to model periodic functions, and we can try the same thing. First, some convenient notation. It is cumbersome to write out $e^{2\pi i \theta}$ all the time, so we abbreviate this as $e(\theta)$. So, in this notation there is a little bit of ambiguity between the number “$e$”, and now the function “$e$”. To lessen this ambiguity, let us also define

$$e_k(\theta) = e(\theta/k) = e^{2\pi i \theta/k}.$$  

Note that $e_k(\theta_1) = e_k(\theta_2)$ if and only if $\theta_1 - \theta_2$ is an integer divisible by $k$.

For $j = 0, 1, \ldots, k-1$, let

$$f_j(n) = e_k(jn) = e^{2\pi i jn/k}.$$  

Then $f_j$ is a periodic function with period $k$, so it is in $\mathcal{P}_k$. We show that $f_0, f_1, \ldots, f_{k-1}$ form a basis for $\mathcal{P}_k$ and we do this by showing that every member $g \in \mathcal{P}_k$ can be written as a linear combination of $f_0, f_1, \ldots, f_{k-1}$. That is, we wish to show that there are complex numbers $c_0, c_1, \ldots, c_{k-1}$ with

$$g = c_0 f_0 + c_1 f_1 + \cdots + c_{k-1} f_{k-1}.$$  

In fact, there is a neat formula for these coefficients $c_j$:

$$c_j = \frac{1}{k} \sum_{m=0}^{k-1} g(m) f_m(-j) = \frac{1}{k} \sum_{m=0}^{k-1} g(m) e_k(-jm).$$  

Let’s check it out:

$$\sum_{j=0}^{k-1} c_j f_j(n) = \frac{1}{k} \sum_{j=0}^{k-1} e_k(jn) \sum_{m=0}^{k-1} g(m) e_k(-jm) = \frac{1}{k} \sum_{m=0}^{k-1} g(m) \sum_{j=0}^{k-1} e_k(j(n-m)),$$

since $e_k(\theta_1) e_k(-\theta_2) = e_k(\theta_1 - \theta_2)$. The inner sum here can be evaluated, in fact it is a sum of a geometric progression. If $n \equiv m \mod k$, then $e_k(j(n-m)) = 1$ and the inner sum is $k$, while otherwise, the inner sum is $0$. Hence the total expression pulls out that value of $g(m)$ where $m \equiv n \mod k$, and gives it weight $\frac{1}{k} = 1$. So the above sum is $g(n)$ and we indeed have

$$g = c_0 f_0 + c_1 f_1 + \cdots + c_{k-1} f_{k-1}.$$
4 The Fourier coefficients of a Dirichlet character

If we have a Dirichlet character mod \( k \), we recognize it as a periodic function mod \( k \) that has some extra properties. Forgetting for a moment those extra properties, we can use the approach of the last section to write \( \chi \) as a linear combination of the functions \( e_k(jn) \) for \( j = 0, 1, \ldots, k - 1 \). As we saw, the coefficients \( c_j \) are given by the expression

\[
\frac{1}{k} \sum_{m=0}^{k-1} \chi(m)e_k(-jm).
\]

There is a special notation for the sum here: Let

\[
G(j, \chi) = \sum_{m=0}^{k-1} \chi(m)e_k(jm).
\]

This is a “Gauss sum,” so the letter \( G \) is used. From what we did in the previous unit, we have

\[
\chi(n) = \frac{1}{k} \sum_{j=0}^{k-1} G(-j, \chi)e_k(jn), \quad n \in \mathbb{Z}. \tag{1}
\]

We now try and figure out some properties of Gauss sums. By trying to compute a few examples we notice that sometimes \( G(j, \chi) = 0 \) and sometimes not. For some characters \( \chi \) this seems to depend on whether \( j \) is coprime to \( k \), so \( G(j, \chi) = 0 \) if and only if \( \gcd(j, k) > 1 \). And for some other characters this does not seem to be the case.

To make a longer story shorter, we cut to the chase. There are two types of Dirichlet characters: primitive and imprimitive. The definition of imprimitive is easier, the definition of primitive then being “not imprimitive”. Suppose that \( \psi \) is a character mod \( d \), we have \( d \mid k \), \( d < k \), and \( \chi_1 \) is the principal character mod \( k \). Then \( \psi\chi_1 \) is a character mod \( k \) and it is imprimitive. So, a primitive character mod \( k \) is one that cannot be “induced” in this way by a character of smaller modulus.

If \( k \) is a prime, then every non-principal character mod \( k \) is primitive, since \( k \) has only the proper divisor 1. Here are the 4 real characters mod 15 given by their values at 1, 2, 4, 7, 8, 11, 13, 14:

\[
1, 1, 1, 1, 1, 1, 1, 1
\]
\[
1, -1, 1, -1, -1, 1, -1, 1
\]
\[
1, -1, 1, -1, -1, 1, -1, 1
\]
\[
1, 1, 1, -1, 1, -1, -1, -1.
\]

Can you tell which are primitive and which are imprimitive? The key is to look at \( \chi(7) \) and \( \chi(11) \), the 4th and 6th entries. Since 7 is 1 mod 3 and a generator mod 5, if the value here is
1, then we see that chi is induced from either a character mod 1 or one mod 3, so the first two are not primitive. Since 11 is 1 mod 5 and a generator mod 3, we see that the first and third are not primitive. Only the 4th line is from a primitive character.

Here is a criterion for being imprimitive: there is a divisor \( d \mid k \) with \( d < k \) such that for each \( a \equiv 1 \) (mod \( d \)) and \( \gcd(a, k) = 1 \) we have \( \chi(a) = 1 \). The proof of this is fairly easy; one shows that if \( u \equiv v \) (mod \( d \)) and \( \gcd(uv, k) = 1 \) then \( \chi(u) = \chi(v) \), and so we are on the way to defining a character mod \( d \) that induces \( \chi \).

**Proposition 1.** Suppose that \( \chi \) is a character mod \( k \) and \( \gcd(n, k) > 1 \). If \( G(n, \chi) \neq 0 \), then \( \chi \) is not primitive.

*Proof.* Let \( d = k / \gcd(n, k) \) so that \( d \mid k \) and \( d < k \). Suppose that \( a \) is an integer coprime to \( k \) and satisfying \( a \equiv 1 \) (mod \( d \)). Note that as \( m \) runs over a complete residue system modulo \( m \), so does \( am \), and so

\[
G(n, \chi) = \sum_{m=0}^{k-1} \chi(am)e_k(nam) = \chi(a) \sum_{k=0}^{m-1} \chi(m)e_k(nam) = \chi(a)G(na, \chi).
\]

Seeing that \( k/d = \gcd(n, k) \mid n \), we write \( n = uk/d \). Also write \( a = dv + 1 \). Thus,

\[
a = (uk/d)(dv + 1) = ukv + uk/d = ukv + n \equiv n \pmod{k},
\]

which implies that \( G(na, \chi) = G(n, \chi) \). Thus, \( G(n, \chi) = \chi(a)G(n, \chi) \), and since \( G(n, \chi) \neq 0 \), we see that \( \chi(a) = 1 \). We have seen that this is the criterion for \( \chi \) to be imprimitive, and to be induced by a character mod \( d \). This completes the proof.

What’s interesting to us is the contrapositive of Proposition 1: *If \( \chi \) is a primitive character mod \( k \) and if \( \gcd(n, k) > 1 \), then \( G(n, \chi) = 0 \).*

We exploit this result in a couple of ways.

**Corollary 1.** If \( \chi \) is a primitive character mod \( k \) and \( n \) is any integer, then

\[
G(n, \chi) = \overline{\chi}(n)G(1, \chi).
\]

*Proof.* We have seen that this holds if \( \gcd(n, k) > 1 \), since then both sides are 0. So assume that \( \gcd(n, k) = 1 \). In this case, the relation to be proved is equivalent to showing that \( \chi(n)G(n, \chi) = G(1, \chi) \), and we shall show this latter relation holds regardless if \( \chi \) is primitive. As above, since \( n \) is coprime to \( k \), as \( m \) runs through a complete residue system mod \( k \), so does \( nm \), so that

\[
G(1, \chi) = \sum_{m=0}^{k-1} \chi(nm)e_k(nm) = \chi(n) \sum_{m=0}^{k-1} \chi(m)e_k(nm) = \chi(n)G(n, \chi).
\]

\[\square\]
Corollary 2. If $\chi$ is a primitive character mod $k$ then $|G(1, \chi)| = \sqrt{k}$.

Proof. We have

$$|G(1, \chi)|^2 = G(1, \chi)\overline{G(1, \chi)} = G(1, \chi) \sum_{m=0}^{k-1} \overline{\chi(m)} e_k(-m) = \sum_{m=0}^{k-1} G(m, \chi) e_k(-m),$$

since Corollary 1 implies that $\overline{\chi(m)} G(1, \chi) = G(m, \chi)$. Thus,

$$|G(1, \chi)|^2 = \sum_{m=0}^{k-1} \sum_{j=0}^{k-1} \chi(j) e_k(jm) e_k(-m) = \sum_{j=0}^{k-1} \chi(j) \sum_{m=0}^{k-1} e_k(m(j-1)).$$

We recognize the inner sum as 0 except in the case $j = 1$ when the inner sum is $k$. Thus, $|G(1, \chi)|^2 = \chi(1)k = k$, which was to be proved.

5 The Pólya–Vinogradov inequality

Suppose $\chi$ is a character mod $k$ and we wish to get a reasonable upper estimate for

$$|S(\chi, n)|,$$

where $S(\chi, n) = \sum_{m=0}^{n} \chi(m)$.

First note some simple cases: If $\chi$ is principal, then the sum is exactly the number of integers in $[0, n]$ that are coprime to $k$, and a simple upper estimate is $n$. But what if $\chi$ is not principal? Since $S(\chi, k) = 0$ in this case, we see that $S(\chi, n)$ is periodic with period $k$, so we may assume that $n \leq k$. We always have the trivial estimate that $|S(\chi, n)| \leq n$. The Pólya–Vinogradov inequality gives a nontrivial estimate for $|S(\chi, n)|$ if $\chi$ is not principal and $n$ is not too small compared with $k$.

First assume that $\chi$ is primitive and $k \geq 3$. (There are no primitive characters mod 2, and the only character mod 1 is principal.)

As a periodic function with period $k$, we have

$$\chi(m) = \frac{1}{k} \sum_{j=0}^{k-1} c_j e_k(jm),$$

where $c_j = \frac{1}{k} \sum_{l=0}^{k-1} \chi(l) e_k(-lj) = G(-lj, \chi)$.

Thus,

$$S(\chi, n) = \frac{1}{k} \sum_{m=0}^{n} \chi(m) = \frac{1}{k} \sum_{m=0}^{n} \sum_{j=0}^{k-1} c_j e_k(jm) = \frac{1}{k} \sum_{j=0}^{k-1} c_j \sum_{m=0}^{n} e_k(jm).$$
This is neat because the inner sum is now a geometric progression and the numbers $c_j$, being Gauss sums, have a known magnitude via Corollary 2. In particular, we have

$$|S(\chi, n)| \leq \frac{1}{\sqrt{k}} \sum_{\substack{j \leq k-1 \atop \gcd(j, k) = 1}} \left| \sum_{m=0}^{n} e_k(jm) \right|.$$  

We now figure out what this geometric progression sums to:

$$\sum_{m=0}^{n} e_k(jm) = \frac{e_k(j(n + 1) - 1)}{e_k(j) - 1} = \frac{e_k(j(n + 1)/2)}{e_k(j/2)} \cdot \frac{e_k(j(n + 1)/2) - e_k(-j(n + 1)/2)}{e_k(j/2) - e_k(-j/2)}.$$  

Since $|e_k(\theta)| = 1$ for $\theta$ real, and since $\sin(2\pi\theta/k) = (e_k(\theta) - e_k(-\theta))/2$, we have

$$\left| \sum_{m=1}^{n} e_k(jm) \right| = \left| \frac{\sin(2\pi j(n + 1)/(2k))}{\sin(2\pi j/(2k))} \right| \leq \frac{1}{|\sin(\pi j/k)|}.$$  

Thus,

$$|S(\chi, n)| \leq \frac{1}{\sqrt{k}} \sum_{\substack{j \leq k-1 \atop \gcd(j, k) = 1}} \frac{1}{|\sin(\pi j/k)|} \leq \frac{2}{\sqrt{k}} \sum_{1 \leq j \leq k/2} \frac{1}{|\sin(\pi j/k)|},$$

using the symmetry $\sin(\pi - \theta) = \sin(\theta)$. Note that $\sin \theta \geq 2\theta/\pi$ for $0 < \theta \leq \pi/2$ (just draw a line from $(0, 0)$ to $(\pi/2, 1)$ and see that $\sin \theta$ lies at or above this line). Thus, $1/|\sin(\pi j/k)| \leq (\pi/2)k/(\pi j) = \frac{1}{2}k/j$ for $1 \leq j \leq k/2$, and so

$$|S(\chi, n)| \leq \frac{1}{\sqrt{k}} \sum_{j \leq k/2} \frac{k}{j} = \sqrt{k} \sum_{j \leq k/2} \frac{1}{j}.$$  

We recognize this last sum as $\log(k/2) + \gamma + o(1)$ as $k \to \infty$ and to make this “numerical” (that is, not involving a limit at infinity), note that if $y \geq 5$, then

$$\sum_{j \leq y} \frac{1}{j} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \sum_{6 \leq j \leq y} \frac{1}{j} \leq \frac{137}{60} + \int_{5}^{y} \frac{dt}{t} = \frac{137}{60} - \log 5 + \log y.$$  

Applying this for $k \geq 10$, we have

$$\sum_{j \leq k/2} \frac{1}{j} < \frac{137}{60} - \log 5 + \log(k/2) < \log k,$$

since $137/60 < \log 10$. Hence for $k \geq 10$ and $\chi$ a primitive character mod $k$, we have

$$|S(\chi, n)| < \sqrt{k} \log k \quad (2)$$
for every integer $n$. One can check directly for the 16 primitive characters with modulus in
$[3, 9]$ and see that this result continues to be true there too. This then is the Pólya–Vinogradov
inequality.

The result can be rather easily extended to non-principal characters that are not primitive,
with a small sacrifice in the constant factor. (The constant factor is invisible in (2) since it is
“1”.) With considerably more work, the Pólya–Vinogradov inequality can be improved a little.
When $n$ is not too small, but smaller than $\sqrt{k}$, there is another inequality for $|S(\chi, n)|$, due to
Burgess, that does better. But perhaps it is time to leave this where it is.