DIRICHLET'S THEOREM ON PRIMES IN PROGRESSIONS, IV

K. Soundararajan

In the last article we define $\phi(q)$ Dirichlet characters $\pmod q$, and used these to isolate the $\phi(q)$ reduced residues $\pmod q$. Then the proof of Dirichlet's theorem boiled down to showing that as $s \to 1^+$ and $\chi \neq \chi_0$, $L(s,\chi)$ tends neither to zero nor infinity.

First we show that $L(s,\chi)$ does not go to infinity as $s \to 1^+$. To do this we introduce a technique known as partial summation which will be very useful in our subsequent work.

Lemma 1. *Let $a_1, a_2, \ldots$ be a sequence of complex numbers. Let f be some function from $\mathbb{R}$ to $\mathbb{C}$. Set $S(n) = \sum_{k=1}^n a_k$. Then*

$$\sum_{n=A+1}^{B} a_n f(n) = S(B)f(B) - S(A)f(A+1) - \sum_{n=A+1}^{B-1} S(n)(f(n+1)-f(n)).$$

Proof. Since $a_n = S(n) - S(n-1)$, our sum is

$$\begin{aligned}\sum_{n=A+1}^{B} (S(n)-S(n-1))f(n) &= \sum_{n=A+1}^{B} S(n)f(n) - \sum_{n=A+1}^{B} S(n-1)f(n) \\ &= \sum_{n=A+1}^{B} S(n)f(n) - \sum_{n=A}^{B-1} S(n)f(n+1) \\ &= S(B)f(B) - S(A)f(A+1) \\ &\qquad - \sum_{n=A+1}^{B-1} S(n)(f(n+1)-f(n)).\end{aligned}$$

This proves our Lemma.

With this in hand we can state a generalization of the alternating series test which will apply to our L-functions.

Proposition 2. *Let $f(n)$ be a sequence of positive real numbers which decreases monotonically to zero. Let a_n denote a sequence of complex numbers with the property that the*

Typeset by $\mathcal{AMS}$-TEX

partial sums $S(n) = \sum_{k=1}^n a_k$ *are bounded in size; that is* $|S(n)| \le S$ *for all* n*. Then the series* $\sum_{n=1}^\infty a_n f(n)$ *converges conditionally. In fact we have*

$$\Big| \sum_{n=A+1}^{B} a_n f(n) \Big| \le 2Sf(A+1).$$

If $a_n = (-1)^{n-1}$ then the partial sums $S(n)$ are either zero or one, and so Proposition 2 generalizes the usual alternating series test.

Proof. We only need to establish the bound in the Proposition; that estimate shows that the partial sums $\sum_{n=1}^N a_n f(n)$ form a Cauchy sequence which therefore converges. Applying Lemma 1 we see that

$$\sum_{n=A+1}^{B} a_n f(n) = S(B)f(B) - S(A)f(A+1) - \sum_{n=A+1}^{B-1} S(n)(f(n+1) - f(n)).$$

Using now that $|S(n)| \le S$, the triangle inequality, and that f is monotone decreasing, we obtain that the above is bounded in size by

$$Sf(B) + Sf(A+1) + S \sum_{n=A+1}^{B} (f(n) - f(n+1)) = 2Sf(A+1).$$

This proves our proposition.

Corollary 3. *Let* χ *be a non-principal character* $\pmod q$*. Then the series* $L(s,\chi) = \sum_{n=1}^\infty \chi(n) n^{-s}$ *converges (conditionally) for all* $s > 0$*. Moreover we have*

$$\Big| L(s,\chi) - \sum_{n=1}^{N} \frac{\chi(n)}{n^s} \Big| \le 2qN^{-s}.$$

Proof. We simply use Proposition 2 with $a_n = \chi(n)$ and $f(n) = n^{-s}$. The key point is that if χ is a non-principal character $\pmod q$, then the partial sums $S(n) = \sum_{k \le n} \chi(k)$ are always bounded in size by q. To see this, note that $\sum_{k=a+1}^{a+q} \chi(k) = 0$ since χ is non-principal, and so we may divide the numbers from 1 to n into intervals of length q with one smaller interval at the end. In each interval of length q the character values sum to zero, and the final interval has fewer than q terms and the character values there can sum at most to q. So $|S(n)| \le q$ always, and our Corollary follows from Proposition 2.

From Corollary 3 it is clear that if χ is non-principal then $L(s,\chi)$ converges to the bounded quantity $L(1,\chi)$ as $s \to 1^+$. In fact we see that, by taking $s = 1$ and $N = q$ in Corollary 3,

$$L(1,\chi) = \sum_{n \le q} \frac{\chi(n)}{n} + O(1) = O\Big(\sum_{n \le q} \frac{1}{n} \Big) + O(1) = O(\log q).$$

This proves the easy part of what we need, and it remains lastly to show that $L(s, \chi)$ does not tend to zero as $s \to 1^+$.

The partial summation idea can also be used to show that $L(s, \chi)$ is a nice smooth function (that is, infinitely differentiable) for $s > 0$. This takes a little thought, but for example you can check that the series

$$\sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}(-\log n)$$

converges for $s > 0$, and in fact it equals the derivative of $L(s, \chi)$ (as you may guess by differentiating term by term). In other words, $L(s, \chi)$ behaves very nicely in $s > 0$ and has a nice Taylor expansion around 1. For example, in the neighborhood of 1 we have $L(s, \chi) = L(1, \chi) + O((s-1))$ so that

$$\log |L(s, \chi)| = O(1),$$

if $L(1, \chi) \neq 0$, and for $s > 1$

$$\log |L(s, \chi)| \leq \log(s-1) + O(1)$$

in case $L(1, \chi) = 0$. We will use these below.

Theorem 4. *There is at most one non-principal character $\chi \pmod q$ for which $L(1, \chi)$ could be zero. This character, if it exists, must be real valued (that is, take values 0, ± 1).*

Proof. Recall that for $s > 1$ we have

$$\sum_{p \equiv a \pmod q} \frac{1}{p^s} = \frac{1}{\phi(q)} \sum_{\chi \pmod q} \overline{\chi(a)} \log L(s, \chi) + O(1).$$

Taking here $a = 1$ we obtain that

$$\sum_{p \equiv 1 \pmod q} \frac{1}{p^s} = \frac{1}{\phi(q)} \sum_{\chi \pmod q} \log L(s, \chi) + O(1).$$

We may as well take real parts on both sides above; after all the LHS is plainly real and so must the RHS. Noting that the real part of $\log z$ is $\log |z|$ we obtain

$$\sum_{p \equiv 1 \pmod q} \frac{1}{p^s} = \frac{1}{\phi(q)} \sum_{\chi \pmod q} \log |L(s, \chi)| + O(1). \tag{1}$$

The term $\chi = \chi_0$ contributes

$$\frac{1}{\phi(q)} \log L(s, \chi_0) = \frac{1}{\phi(q)} \log \frac{1}{s-1} + O(1),$$

to the RHS of (1). This goes to $+\infty$ as $s \to 1^+$. Note that the other terms appearing in the RHS of (1) either remain bounded (if $L(1,\chi) \neq 0$), or go to $-\infty$ (if $L(1,\chi) = 0$). Let's analyze this a little more carefully. Suppose there are two characters χ_1 and χ_2 (mod q) with $L(1,\chi_1) = L(1,\chi_2) = 0$. Then for s near 1 we would have $L(s,\chi_1) = L(1,\chi_1) + O(s-1) = O(s-1)$, and similarly $L(s,\chi_2) = O(s-1)$. Hence these two characters contribute to the RHS of (1) an amount less than

$$\frac{2}{\phi(q)} \log(s-1) + O(1).$$

The contribution of characters $\chi \neq \chi_0$, χ_1, or χ_2 is less than some constant. Therefore if there are two characters for which $L(1,\chi_1) = L(1,\chi_2) = 0$ then the RHS of (1) is less than

$$\frac{1}{\phi(q)} \log \frac{1}{s-1} + \frac{2}{\phi(q)} \log(s-1) + O(1) = \frac{1}{\phi(q)} \log(s-1) + O(1),$$

and this tends to $-\infty$ as $s \to 1^+$. But the LHS of (1) is a sum over primes which is obviously non-negative! Thus if there are two characters for which the L-function at 1 is zero, then we have reached a contradiction.

In other words, there is at most one character χ (mod q) for which $L(1,\chi)$ could be zero. Why must χ be real valued? If not then $\overline{\chi}$ would be a character distinct from χ. If $L(1,\chi) = 0$ then $L(1,\overline{\chi})$ being the complex conjugate of $L(1,\chi)$ would also be zero, and we would have two characters with the L-value at 1 being zero. We have already ruled that out.

We are now at the final, but difficult, stage of the proof. We must show that $L(1,\chi) \neq 0$ for a real valued character χ (mod q). The argument to establish this is quite subtle and relies upon the quantity

$$r_\chi(n) = \sum_{d|n} \chi(d).$$

We shall consider the behavior of

$$R(x) := \sum_{n \leq x} \frac{r_\chi(n)}{\sqrt{n}}$$

for large x. First we shall show that $R(x) \geq \log x + O(1)$ so that $R(x) \to \infty$ as $x \to \infty$. However, if $L(1,\chi) = 0$ we shall prove that $R(x)$ must be bounded (this is delicate!). This contradiction shows that $L(1,\chi) \neq 0$, and finally completes the proof of Dirichlet's theorem.

Lemma 5. *The function $r_\chi(n)$ is multiplicative; that is, if $(m,n) = 1$ then $r_\chi(m)r_\chi(n) = r_\chi(mn)$. For all natural numbers n we have $r_\chi(n) \geq 0$, and if $n = m^2$ is a perfect square then $r_\chi(n) \geq 1$.*

Proof. If m and n are coprime then we can associate to every divisor d of mn a unique pair of divisors (d_1, d_2) with $d_1|m$ and $d_2|n$ and $d = d_1 d_2$. From this correspondence we see that $r_\chi(n)$ is multiplicative. To check the second assertion it suffices to verify that $r_\chi(p^k) \geq 0$ for all prime powers p^k and that $r_\chi(p^{2k}) \geq 1$. Now

$$r_\chi(p^k) = \sum_{j=0}^{k} \chi(p^j) = \begin{cases} k+1 & \text{if } \chi(p) = 1 \\ 1 & \text{if } \chi(p) = 0 \\ (1+(-1)^k)/2 & \text{if } \chi(p) = -1. \end{cases}$$

Our result therefore follows.

Proposition 6. *We have*

$$R(x) \geq \frac{1}{2}\log x + O(1).$$

Proof. From Lemma 5 and the definition of $R(x)$ we obtain that

$$R(x) \geq \sum_{m \leq \sqrt{x}} \frac{r_\chi(m^2)}{m} \geq \sum_{m \leq \sqrt{x}} \frac{1}{m} \geq \int_1^{[\sqrt{x}]} \frac{dt}{t} = \frac{1}{2}\log x + O(1).$$

Now we must obtain an upper bound for $R(x)$, under the assumption that $L(1,\chi) = 0$. We may rewrite the definition of $r_\chi(n)$ as $r_{chi}(n) = \sum_{ab=n} \chi(a)$. Thus

$$R(x) = \sum_{\substack{a,b \leq x \\ ab \leq x}} \frac{\chi(a)}{\sqrt{ab}}$$

may be viewed as a sum over all lattice points (a,b) lying in the first quadrant and below the hyperbola $ab = x$. To estimate such sums Dirichlet devised an ingenious method called the *hyperbola method*. The idea is to choose to auxiliary parameters A and B with $AB = x$; in our case at hand we shall choose $A = B = \sqrt{x}$. Then we can split the lattice points (a,b) with $ab \leq x$ into two cases: terms of the first type with $a \leq A$ and $b \leq x/a$; terms of the second type with $b \leq B$ but with $A < a \leq x/b$. Call $R_1(x)$ the contribution of the first type of cases to $R(x)$, and $R_2(x)$ the contribution of the second type of cases.

To estimate these sums we will require the following Proposition, which will follow from partial summation, but whose proof we postpone for the moment.

Proposition 7. *For all $y \geq 1$ we have the asymptotic formula*

$$\sum_{n \leq y} \frac{1}{\sqrt{n}} = 2\sqrt{y} + C + O\Big(\frac{1}{\sqrt{y}}\Big),$$

where C is a constant.

Consider $R_1(x)$. From its definition it equals

$$\sum_{a \leq A} \frac{\chi(a)}{\sqrt{a}} \sum_{b \leq x/a} \frac{1}{\sqrt{b}},$$

and using Proposition 7 to evaluate the sum over b we get

$$\sum_{a\le A} \frac{\chi(a)}{\sqrt{a}}\Big(2\frac{\sqrt{x}}{\sqrt{a}} + C + O\Big(\frac{\sqrt{a}}{\sqrt{x}}\Big)\Big). \tag{2}$$

The first term above is

$$2\sqrt{x}\sum_{a\le A}\frac{\chi(a)}{a} = 2\sqrt{x}(L(1,\chi)+O(A^{-1})) = O(1),$$

where the first step follows from Corollary 3, and the second step follows from our assumption that $L(1,\chi)=0$, and our choice $A=\sqrt{x}$. The second term in (2) gives

$$C\sum_{a\le A}\frac{\chi(a)}{\sqrt{a}} = C(L(\tfrac{1}{2},\chi)+O(A^{-\frac{1}{2}})) = O(1),$$

upon using Corollary 3 and noting that $L(\frac{1}{2},\chi)$ is simply some constant independent of x. The last term in (2) gives

$$O\Big(\sum_{a\le A}\frac{1}{\sqrt{x}}\Big) = O(1).$$

Thus we conclude that $R_1(x) = O(1)$.

Let us now consider $R_2(x)$. This is

$$R_2(x) = \sum_{b\le B}\frac{1}{\sqrt{b}}\sum_{A<a\le x/b}\frac{\chi(a)}{\sqrt{a}}.$$

Consider the inner sum over a. This is part of the 'tail' of the series defining $L(\frac{1}{2},\chi)$. So certainly we would expect it to be small. Indeed we could use Proposition 2 (with $a_n=\chi(n)$ and $f(n)=1/\sqrt{n}$; or see Corollary 3) to see that

$$\Big|\sum_{A<a\le x/b}\frac{\chi(a)}{\sqrt{a}}\Big| \le 2q\frac{1}{\sqrt{A}} = O\Big(\frac{1}{\sqrt{A}}\Big).$$

Therefore, using Proposition 7,

$$R_2(x) = O\Big(\frac{1}{\sqrt{A}}\Big)\sum_{b\le B}\frac{1}{\sqrt{b}} = O\Big(\frac{\sqrt{B}}{\sqrt{A}}\Big) = O(1).$$

We deduce that $R(x) = R_1(x)+R_2(x) = O(1)$ which contradicts Proposition 6. Therefore our assumption that $L(1,\chi)=0$ must have been wrong, and we have proved Dirichlet's theorem!

It remains lastly to give a proof of Proposition 7. This is a simple application of the partial summation idea of Lemma 1.

Proof of Proposition 7. Take $a_n = 1$ for $n \geq 1$, and $f(x) = 1/\sqrt{x}$, and note that $S(x) = \sum_{k \leq x} a_k = [x]$. Then applying Lemma 1 with $A = 0$ and $B = [y]$ we obtain that

$$\sum_{n \leq y} \frac{1}{\sqrt{n}} = \frac{[y]}{\sqrt{y}} - \sum_{n=1}^{y-1} n\Big(\frac{1}{\sqrt{n+1}} - \frac{1}{\sqrt{n}}\Big). \tag{3}$$

The first term above is

$$\frac{y + O(1)}{\sqrt{y}} = \sqrt{y} + O\Big(\frac{1}{\sqrt{y}}\Big).$$

The second term in (3) we may write as

$$-\sum_{n=1}^{y-1} n \int_n^{n+1} \Big(\frac{1}{\sqrt{t}}\Big)' dt = \sum_{n=1}^{y-1} n \int_n^{n+1} \frac{1}{2t^{\frac{3}{2}}} dt = \sum_{n=1}^{[y]-1} \int_n^{n+1} \frac{[t]}{2t^{\frac{3}{2}}} dt = \int_1^{[y]} \frac{[t]}{2t^{\frac{3}{2}}} dt.$$

Since $[t] = t - \{t\}$ (where $\{t\}$ denotes the fractional part of t) the above is

$$\int_1^{[y]} \frac{t}{2t^{\frac{3}{2}}} dt - \int_1^{[y]} \frac{\{t\}}{2t^{\frac{3}{2}}} dt = \sqrt{[y]} - 1 - \int_1^{[y]} \frac{\{t\}}{2t^{\frac{3}{2}}} dt.$$

You can check easily that $\sqrt{[y]} = \sqrt{y} + O(1/\sqrt{y})$. The integrand in the second integral above is always less than $1/(2t^{\frac{3}{2}})$ in size; therefore the integral will converge when extended to infinity. In other words,

$$\begin{aligned} \int_1^{[y]} \frac{\{t\}}{2t^{\frac{3}{2}}} dt &= \int_1^{\infty} \frac{\{t\}}{2t^{\frac{3}{2}}} dt + O\Big(\int_{[y]}^{\infty} \frac{\{t\}}{t^{\frac{3}{2}}} dt\Big) = \int_1^{\infty} \frac{\{t\}}{2t^{\frac{3}{2}}} dt + O\Big(\int_{[y]}^{\infty} \frac{dt}{t^{\frac{3}{2}}}\Big) \\ &= \int_1^{\infty} \frac{\{t\}}{2t^{\frac{3}{2}}} dt + O\Big(\frac{1}{\sqrt{y}}\Big). \end{aligned}$$

Putting everything together we find that

$$\sum_{n \leq y} \frac{1}{\sqrt{n}} = 2\sqrt{y} - 1 - \int_1^{\infty} \frac{\{t\}}{2t^{\frac{3}{2}}} dt + O\Big(\frac{1}{\sqrt{y}}\Big),$$

and so the constant C in Proposition 7 equals

$$-1 - \int_1^{\infty} \frac{\{t\}}{2t^{\frac{3}{2}}} dt.$$