Math 259: Introduction to Analytic Number Theory

Exponential sums IV:
The Davenport-Erdős and Burgess bounds on short character sums

Finally, we consider upper bounds on exponential sum involving a character, again concentrating on the simplest such sum

$$S_\chi(N) := \sum_{n=1}^{N} \chi(n).$$

Here χ is a primitive character mod $q > 1$. We already encountered the Pólya-Vinogradov bound:

$$|S_\chi(N)| \ll q^{1/2} \log q$$

for all N. We motivate our further study of $S_\chi(N)$ with two paradigmatic applications.

The first application concerns an analogue of Lindelöf's conjecture "in the q-direction": fix $s = \sigma + it$ with $\sigma \in [1/2, 1]$, and ask how $|L(s,\chi)|$ can grow as we vary χ. (If $\sigma < 1/2$ then $1 - \sigma > 1/2$, so we first estimate $L(1-s, \overline{\chi})$ and then use the functional equation; if $\sigma > 1$ then the Dirichlet series converges, and $|L(s,\chi)| < \zeta(\sigma)$ is an upper bound independent of χ.) Partial summation gives $L(s,\chi) = s \int_1^\infty y^{-1-s} S_\chi(y)\, dy$. Thus for fixed s we have $L(s,\chi) \ll \int_1^\infty y^{-1-\sigma} |S_\chi(y)|\, dy$. Hence upper bounds on $|S_\chi(y)|$ directly yield upper bounds on $L(s,\chi)$. For instance, by combining the trivial estimate $|S_\chi(y)| \le y$ with the Pólya-Vinogradov bound we find

$$L(s,\chi) \ll \int_1^{q^{1/2}\log q} y^{-\sigma} dy \;+\; q^{1/2} \log q \int_{q^{1/2}\log q}^{\infty} y^{-1-\sigma} dy \ll (q^{1/2} \log q)^{1-\sigma},$$

except for $\sigma = 1$, for which we obtain $L(s,\chi) \ll \log q$ as we have already seen. In particular, $|L(s,\chi)| \ll q^{((1-\sigma)/2)+\epsilon}$ holds for all $\epsilon > 0$. Now for arbitrary y the Pólya-Vinogradov bound is almost as good as we can hope for (we shall make this precise soon), and is consistent with the typical behavior of a sum of about q random numbers of absolute value 1. But when y is significantly smaller than q, that is $y \ll q^\theta$ for $\theta < 1$, we expect that the exponent of $1/2$ in the Pólya-Vinogradov bound is too large, and indeed we know it is too large once $\theta \le 1/2$ because then the trivial bound $|S_\chi(y)| \le y$ is better. Nontrivial estimates on $|S_\chi(y)|$ for small y will yield improved upper bounds on $|L(s,\chi)|$, and if we could prove $|S_\chi(y)| \ll y^{1/2+o(1)}$ for all $y \gg q^\epsilon$ then it would follow that $L(s,\chi) \ll q^{o(1)}$, the q-analogue of Lindelöf's conjecture.

The second application is to the smallest quadratic nonresidue modulo a large prime p. Here we take $q = p$ and χ the Legendre symbol $\chi(n) = (n/p)$, and seek small solutions of $\chi(n) = -1$. Under the Extended Riemann Hypothesis, the smallest such n is $O(\log^2 p)$, but without unproved conjectures the best bounds known are of the form $p^{\theta+o(1)}$ for some positive θ. If there is no quadratic nonresidue in $[1,N]$ then $S_\chi(N)=N$, so $|S_\chi(N)|$ attains the trivial upper bound. Hence any nontrivial bound on $|S_\chi(N)|$ proves the existence of a quadratic nonresidue less than N. For example, Pólya-Vinogradov lets us take $N \gg p^{1/2}\log p$, giving $p^{\theta+o(1)}$ for $\theta=1/2$. Now it is true that an upper bound $2p^{1/2}$ on the least quadratic nonresidue is elementary,[1] but in fact Pólya-Vinogradov yields an upper bound $p^{\theta+o(1)}$ with $\theta=1/(2e^{1/2})$, as observed in [Davenport-Erdős 1952]. More generally:

Proposition 1. *Let χ be a character modulo a prime p. Suppose $|S_\chi(N)|<\epsilon N$ for some positive integer N and some $\epsilon>0$. Then there exists a positive integer $x<N^{\exp((\epsilon-1)/2)+o(1)}$ such that $\chi(x)\neq 1$.*

The $o(1)$ is effective; that is, for each $c>\exp((-\epsilon-1)/2)$ there exists an effectively bounded $N_0(c)$ such that if $N>N_0(c)$ then $|S_\chi(N)|<\epsilon N$ implies the existence of $x<N^c$ not in the kernel of χ.

Proof: Suppose $\chi(n)=1$ for all positive integers $x\leq X$. Then the same is true for every integer none of whose prime factors exceeds X. For each prime $l>X$ there are at most N/l integers $n\leq N$ for which $l|N$. Therefore the total number of integers $n\leq N$ divisible by such a prime is at most $N\sum_{X<l\leqslant N}1/l$. This number must exceed $(1-\epsilon)N/2$ for $|S_\chi(N)|$ to be as small as ϵN. But we already know that $\sum_{X<l\leqslant N}1/l=\log(\log(N)/\log(X))+o(1)$ as $X\to\infty$. Therefore $\log(N)/\log(X)$ must exceed $\exp((1-\epsilon)/2)-o(1)$, as claimed.

In fact it is possible to improve on both the trivial bound and the Pólya-Vinogradov bound on $S_\chi(N)$ when $N\ll q^\theta$ for certain $\theta<1$. As with exponential sums $\sum_{n=1}^N e(f(n))$, there are excellent bounds on average, and less good but still nontrivial bounds on individual sums. We address the average case first.

To get enough sums to average, we generalize $S_\chi(N)$ to what we shall call

$$S_\chi(n_0,n_0+N) := \sum_{n=1}^{N}\chi(n_0+n) = S_\chi(n_0+N)-S_\chi(n_0).$$

We fix the length N of the sum, and vary the starting point n_0 over all of $\mathbf{Z}/q\mathbf{Z}$. If q is much larger than N, We expect $S_\chi(n_0,n_0+N)$ to behave like the sum of about $(\phi(q)/q)N$ independent random numbers of absolute value 1, drawn uniformly from $\{1,-1\}$ if χ is real and from the m-th roots of unity for some $m\geq 3$ if χ is complex. By the Central Limit Theorem, for large N the distribution of these sums should approach a normal distribution with mean

[1] If $\chi(-1)=-1$ then $\chi(p-n^2)=-1$ for all $n\not\equiv 0 \bmod p$, and taking $n=\lfloor\sqrt{p}\rfloor$ makes $p-n^2<2p^{1/2}$. If $\chi(p)=+1$ it is enough to prove that not all nonzero $n\in(-p^{1/2},p^{1/2})$ are quadratic residues. If they were then the same would be true of all $m\equiv n/n' \bmod p$ with nonzero $n,n'\in(-p^{1/2},p^{1/2})$. But this is impossible because every $m\in(\mathbf{Z}/p\mathbf{Z})^*$ is of that form, as may be seen from a standard application of the "pigeonhole principle": there are more than p expressions $am+b$ for integers $a,b\in[0,p^{1/2})$, so two of them must coincide, and solving $am+b=a'm+b'$ we write m as a fraction of the desired form.

zero variance $(\phi(q)/q)N$ on $\mathbf{R}$ (for χ real) or $\mathbf{C}$ (for χ complex). That is, if χ is real we expect that

$$\frac{1}{q}\sum_{n_0=0}^{q-1} f\left(\frac{S_\chi(n_0,n_0+N)}{[(\phi(q)/q]\,N)^{1/2}}\right) \to \frac{1}{\sqrt{2\pi}}\int\limits_{x\in\mathbf{R}} f(x)\,e^{-x^2/2}\,dx, \tag{1}$$

$$\frac{1}{q}\sum_{n_0=0}^{q-1} F\left(\frac{S_\chi(n_0,n_0+N)}{[(\phi(q)/q]\,N)^{1/2}}\right) \to \frac{1}{2\pi}\iint\limits_{(x,y)\in\mathbf{R}^2} F(x+iy)\,e^{-(x^2+y^2)/2}\,dx\,dy \tag{2}$$

for all continuous functions $f:\mathbf{R}\to\mathbf{R}$ or $F:\mathbf{R}\to\mathbf{C}$ for which the integrals converge absolutely. Remarkably this can be proved for sequences of triples (q,χ,N) for which $N\to\infty$ and $\log(q)/\log(N)\to\infty$. We shall follow the proof under the hypothesis that q is prime. The same methods suffice, with some more gruntwork, to handle the case that q is composite but squarefree or nearly so (meaning that q factors as q_1q_2 with q_1 squarefree and $q_2=O(1)$). It is much harder to obtain such results when q has large or numerous repeated prime factors, but fortunately such q cannot arise as the modulus of a primitive real character, or more generally of a character taking values in the m-th roots of unity for fixed m.

Since q is a prime we write $q=p$; since $p\to\infty$, we can write (1,2) more simply as

$$\frac{1}{p}\sum_{n_0=0}^{p-1} f\left(\frac{S_\chi(n_0,n_0+N)}{N^{1/2}}\right) \to \frac{1}{\sqrt{2\pi}}\int\limits_{x\in\mathbf{R}} f(x)\,e^{-x^2/2}\,dx, \tag{3}$$

$$\frac{1}{p}\sum_{n_0=0}^{p-1} F\left(\frac{S_\chi(n_0,n_0+N)}{N^{1/2}}\right) \to \frac{1}{2\pi}\iint\limits_{(x,y)\in\mathbf{R}^2} F(x+iy)\,e^{-(x^2+y^2)/2}\,dx\,dy. \tag{4}$$

This is proved (as is done in one proof of the Central Limit Theorem itself) by proving that the *moments* of the numbers $N^{-1/2}S_\chi(n_0,n_0+N)$ converge to the corresponding moments of the normal distribution on $\mathbf{R}$ or $\mathbf{C}$. That is, one proves (3) for the functions $f(x)=x^r$ $(r=0,1,2,\ldots)$, and proves (4) for the functions $F(z)=z^r\overline{z}^{r'}$ $(r,r'=0,1,2,\ldots)$. This suffices because it is known that the normal distribution on any finite-dimensional real inner product space is characterized by its moments (and in the complex case the monomials $z^r\overline{z}^{r'}$ span all the polynomials in the real and imaginary parts of z). These moments are as follows: in the real case, the r-th moment is

$$\frac{1}{\sqrt{2\pi}}\int\limits_{x\in\mathbf{R}} x^r\,e^{-x^2/2}\,dx = \begin{cases} r!\,/\,(2^{r/2}(r/2)!), & \text{if } r \text{ is even;}\\ 0, & \text{if } r \text{ is odd,}\end{cases} \tag{5}$$

and in the complex case, the (r,r') moment is

$$\frac{1}{2\pi}\iint\limits_{(x,y)\in\mathbf{R}^2} (x+iy)^r(x-iy)^{r'}\,e^{-(x^2+y^2)/2}\,dx = \begin{cases} r!, & \text{if } r=r';\\ 0, & \text{if } r\neq r'\end{cases} \tag{6}$$

(see Exercise 3). The constant $r!\,/\,(2^{r/2}(r/2)!)$ in (5) is also the product of odd numbers from 1 to $r-1$, sometimes denoted by "$(r-1)!!$"; combinatorially it is the number of partitions of an r-element set into $r/2$ pairs.

The first few moments of the $N^{-1/2}S_\chi(n_0, n_0+N)$ are elementary. The zeroth moment is trivially 1; the first moment vanishes:

$$\sum_{n_0=0}^{p-1} S_\chi(n_0, n_0+N) = N \sum_{n=0}^{p-1} \chi(n) = 0.$$

For the second moment we have:[2]

Lemma 1. *Let χ be a nontrivial character modulo a prime p. Then for any integer $N \in [0,p]$ we have*

$$\sum_{n_0=0}^{p-1} |S_\chi(n_0, n_0+N)|^2 = pN - N^2. \tag{7}$$

Proof: As we did for continuous mean squares of exponential sums, we expand and interchange the order of summation:

$$\begin{aligned}
\sum_{n_0=0}^{p-1} |S_\chi(n_0, n_0+N)|^2 &= \sum_{n_0=0}^{p-1} \left| \sum_{n=1}^{N} \chi(n_0+n) \right|^2 \\
&= \sum_{n_0=0}^{p-1} \left(\sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \chi(n_0+n_1)\,\overline{\chi}(n_0+n_2) \right) \\
&= \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \left(\sum_{n_0=0}^{p-1} \chi(n_0+n_1)\,\overline{\chi}(n_0+n_2) \right) \\
&= \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \left(\sum_{x=0}^{p-1} \chi(x+n_1-n_2)\,\overline{\chi}(x) \right),
\end{aligned}$$

where the last step is the change of variable $x = n_0 + n_2$. For each of the N choices of (n_1, n_2) for which $n_1 = n_2$, the inner sum is $\sum_{x=0}^{p-1} |\chi(x)|^2 = p-1$. For the remaining $N^2 - N$ choices, we might expect the inner sum to consist of $p-2$ "random" numbers of absolute value 1, and thus to be roughly of order $p^{1/2}$. But in fact we have almost complete cancellation: the sum equals -1. To see this, omit the term $x=0$, for which $\overline{\chi}(x)$ vanishes, and write the remaining sum as

$$\sum_{x=1}^{p-1} \chi(x+n_1-n_2)\,\chi(x^{-1}) = \sum_{x=1}^{p-1} \chi(1+(n_1-n_2)x^{-1}).$$

[2] This is Lemma 1 of [Davenport-Erdős 1952]; the authors report at the end of the paper that it is also contained in "Vinogradov's *Osnovy teroii čisel*, p.109", published a year or so earlier.

Since $x \mapsto (n_1 - n_2)\, x^{-1}$ is a permutation of $(\mathbf{Z}/p\mathbf{Z})^*$, each $y \in \mathbf{Z}/p\mathbf{Z}$ occurs exactly once as $1 + (n_1 - n_2)x^{-1}$, except for $y = 1$ which does not occur at all. Hence the sum is $\left(\sum_{y=0}^{p-1} \chi(y)\right) - \chi(1) = -1$, and we have

$$\sum_{n_0=0}^{p-1} |S_\chi(n_0, n_0 + N)|^2 = N(p-1) + (N^2 - N)(-1) = pN - N^2,$$

as claimed.

As an immediate application we confirm that the Pólya-Vinogradov bound is within a factor $O(\log q)$ of the truth:

Corollary. *For every nontrivial character χ modulo a prime p, there exist $N \bmod p$ such that $|S_\chi(N)| \geq p^{1/2}/4 - O(p^{-3/2})$.*

Proof: Let $N = \lfloor p/2 \rfloor$. Then $pN - N^2 = p^2/4 - O(1)$. Therefore there exists some n_0 for which $|S_\chi(x_0, x_0 + N)|^2$ is at least as large as its average value $p/4 - O(1/p)$. Hence $|S_\chi(x_0, x_0 + N)| \geq p^{1/2}/2 - O(p^{-3/2})$. Since

$$S_\chi(x_0, x_0 + N) = S_\chi(x_0 + N) - S_\chi(x_0),$$

it follows that at least one of $S_\chi(x_0 + N)$ and $S_\chi(x_0)$ exceeds the claimed lower bound $(p^{1/2}/2 - O(p^{-3/2}))/2 = p^{1/2}/4 - O(p^{-3/2})$.

There is one more case where we can give an elementary nontrivial upper bound on one of our moments, namely the $(2,0)$ or $(0,2)$ moment in the complex case:

Lemma 2. *Let χ be a* complex *character modulo a prime p. Then for any integer $N \in [0, p]$ we have*

$$\left| \sum_{n_0=0}^{p-1} \left(S_\chi(n_0, n_0 + N)\right)^2 \right| \leq p^{1/2} N^2. \tag{8}$$

Proof: We begin as in the proof of Lemma 1 but without taking complex conjugates, finding that

$$\sum_{n_0=0}^{p-1} \left(S_\chi(n_0, n_0 + N)\right)^2 = \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \left(\sum_{x=0}^{p-1} \chi(x(x + n_1 - n_2)) \right).$$

Let $h = n_1 - n_2$. If $h = 0$ then the inner sum is $\sum_{x=0}^{p-1} \chi^2(x) = 0$, because by hypothesis χ^2 is nontrivial. Otherwise, we count for each $y \in (\mathbf{Z}/p\mathbf{Z})$ the number of representations of y as $x(x + h)$. Because p is odd (there being no complex characters mod 2), we may use the quadratic formula, finding that this number is $1 + \psi((h/2)^2 + y)$ where ψ is the quadratic character mod p. Therefore

$$\sum_{x=0}^{p-1} \chi(x(x+h)) = \sum_{y=0}^{p-1} \bigl(1 + \psi((h/2)^2 + y)\bigr)\, \chi(y) = \sum_{y=0}^{p-1} \psi((h/2)^2 + y)\, \chi(y),$$

since $\sum_{y=0}^{p-1} \chi(y) = 0$. Writing $y = -(h/2)^2 c$, we transform the last sum into

$$\chi(-(h/2)^2) \sum_{c \bmod p} \chi(c)\psi(1-c) = \chi(-(h/2)^2)\, J(\chi, \psi)$$

and since none of χ, ψ, or $\chi\psi$ is trivial this Jacobi sum has absolute value $p^{1/2}$. Summing over all n_1 and n_2 we thus get an upper bound of $N + p^{1/2}(N^2 - N)$ on the left-hand side of (8), whence the claimed inequality follows.

That is, the $(2,0)$ and $(0,2)$ moments of our numbers $N^{-1/2} S_\chi(n_0, n_0 + N)$ are $O(p^{-1/2}N)$, which is $o(1)$ as desired once $N = o(p^{1/2})$. Note that here the cross-terms $\sum_{x=0}^{p-1} \chi(n_0 + n_1)\chi(n_0 + n_2)$ do have the expected size $p^{1/2}$, though they still do not really behave like a sum of random numbers of absolute value 1 (such a sum rarely lies exactly on the circle of radius $p^{1/2}$ about the origin).

Beyond these quadratic moments, the main terms are still easy to handle but the cross-terms are no longer elementary. We proceed as before and show that the main terms agree (within negligible errors) with the corresponding moments of the normal distributions. We assume throughout that $N < p$.

In the real case, we have

$$\sum_{n_0=0}^{p-1} \left(S_\chi(n_0, n_0 + N)\right)^r = \sum_{1 \le n_1, \ldots, n_r \le N} \!\!\cdots\!\! \sum \left(\sum_{n_0=0}^{p-1} \chi\Bigl(\prod_{i=1}^{r} (n_0 + n_i)\Bigr) \right). \tag{9}$$

If the polynomial $\prod_{i=1}^{r} (x + n_i)$ is a perfect square then the inner sum is $p - O(r)$. For this to happen, r must be even and the n_i must match in pairs. We already noted that there are $r!/(2^{r/2}(r/2)!)$ ways to pair the indices $1, 2, \ldots, r$. Each of these accounts for $N^{r/2}$ choices of $(n_1, \ldots, n_r)$. This counts some r-tuples more than once, because the same number may equal n_i for $4, 6, 8, \ldots$ choices of i; but this overcounting affects only $O_r(N^{(r/2)-1})$ of the r-tuples. Hence the perfect squares in (9) sum to

$$p \left(\frac{r!}{2^{r/2}(r/2)!} + O_r(1/N) \right) N^{r/2}. \tag{10}$$

Dividing by $N^{r/2}$, we find that the squares' contribution to the r-th moment of the p real numbers $N^{-1/2} S_\chi(n_0, n_0 + N)$ is within $O_r(1/N)$ of the desired $r!/(2^{r/2}(r/2)!)$.

In the complex case we likewise expand $\bigl(S_\chi(n_0, n_0 + N)\bigr)^r \bigl(\overline{S_\chi(n_0, n_0 + N)}\bigr)^{r'}$ and sum over over $n_0 \bmod p$ to obtain

$$\sum_{\substack{1 \le n_1, \ldots, n_r \le N \\ 1 \le n'_1, \ldots, n'_{r'} \le N}} \!\!\cdots\!\! \sum \left(\sum_{n_0=0}^{p-1} \chi\Bigl(\prod_{i=1}^{r} (n_0 + n_i)\Bigr)\, \overline{\chi}\Bigl(\prod_{i'=1}^{r'} (n_0 + n'_i)\Bigr) \right). \tag{11}$$

Here the main terms are those for which $r = r'$ and the n'_i are some permutation of the n_i; this happens if and only if $r = r'$, in which case the number of such terms is $N^r r! - O_r(N^{r-1})$ (again the error is due to overcounting when the n_i are not distinct). These terms contribute $(1 + O_r(1/N))pN^r r!$ to the sum (11), and thus $r! + O_r(1/N)$ to the (r, r) moment of the p complex numbers $N^{-1/2} S_\chi(n_0, n_0 + N)$. There can be other ways to make the sum over n_0 be as large as $p - O(r + r')$; for instance if χ is cubic we may take $(r, r') = (3, 0)$ and $n_1 = n_2 = n_3$. But for a complex character the number of such alternatives is at most $O_{r,r'}(N^{(r+r'-1)/2})$ so their contribution is negligible as $N \to \infty$.

But the cross-terms are more troublesome. Already when we take $r = 3$ in (9) the typical inner sum is $\sum_{n_0=0}^{p-1} \chi\big((n_0 + n_1)(n_0 + n_2)(n_0 + n_3)\big)$. We already noted in the proof of Lemma 2 that when χ is the quadratic character mod p we can interpret $\chi(m)$ as the number of solutions mod p of $y^2 = m$, minus 1. Thus the inner sum is p less than the number of solutions mod p of

$$Y^2 = (X + n_1)(X + n_2)(X + n_3). \tag{12}$$

With three distinct factors on the right-hand side, (12) is an elliptic curve E. The elementary methods we have used thus far do not let us obtain a good upper bound on the size of such a character sum. Hasse [1936] developed enough of the theory of elliptic curves over finite fields to prove that the number of solutions of (12) is $p-a$ where a, the "trace" of E, satisfies[3] $|a| \leq 2p^{1/2}$. When two or more of the n_i coincide we easily obtain $\left|\sum_{n_0=0}^{p-1} \chi((n_0 + n_1)(n_0 + n_2)(n_0 + n_3))\right| \leq 1$. Summing over n_1, n_2, n_3 we thus find that the third moment of the real numbers $N^{-1/2} S_\chi(n_0, n_0 + N)$ is $O(p^{-1/2} N^3)$, which is $o(1)$ as desired once $N = o(p^{1/6})$.

Still in the real case, but taking r arbitrary, we need to bound the discrepancy between p and the number of solutions of

$$Y^2 = (X + n_1)(X + n_2) \cdots (X + n_r). \tag{13}$$

Now we have a hyperelliptic curve whose genus g can be as large as $\lfloor (r-1)/2 \rfloor$ (this upper bound on g is attained if and only if the n_i are distinct). Again a bound $O(q^{1/2})$ on the discrepancy is available, but is even harder: we need Weil's analogue of the Riemann Hypothesis for hyperelliptic curves of arbitrary genus. Weil proved [1948] that a smooth, projective curve of genus g over a finite field of q elements has $q + 1 - a$ rational points with $|a| \leq 2gq^{1/2}$. In our setting $q = p$ and $g < r/2$. The solutions of (13) may not correspond exactly to rational points on the associated smooth projective curve, because the model (13) of the curve is singular at infinity and also at any points where $X + n_i = 0$ for two or more i. Still the two counts differ by at most $(r/2) - g$, whence the Weil bound

$$\left| \sum_{n_0=0}^{p-1} \chi\big((n_0 + n_1)(n_0 + n_2) \cdots (n_0 + n_r)\big) \right| < rp^{1/2} \tag{14}$$

[3] Of course the inequality must be strict, but the proof applies also to an elliptic curve over a finite field of q elements for any prime power q, and when q is a square the values $a = \pm 2q^{1/2}$ can be attained. Note that we write $p - a$, not the usual formula $p + 1 - a$ for the number of rational points on E, because thiat formula includes the point at infinity $(X : Y : 1) = (0 : 1 : 0)$ of (12), which we did not count.

holds for all $(n_1, \ldots, n_r)$ that cannot be partitioned into $r/2$ pairs n_i, n_j with $i \neq j$ but $n_i = n_j$. We have seen already that the number of r-tuples that do have such a partition with each $n_i \in [1, N]$ is $\bigl(r!/(2^{r/2}(r/2)!) + O(1/N)\bigr)N^{r/2}$ if r is even, and zero otherwise. We conclude that

$$\frac{1}{p}\sum_{n_0=0}^{p-1}\left(\frac{S_\chi(n_0, n_0+N)}{N^{1/2}}\right)^r = \frac{1+O_r(1/N)}{\sqrt{2\pi}}\int\limits_{x\in\mathbf{R}} x^r e^{-x^2/2}dx + O(rp^{-1/2}N^r) \quad (15)$$

holds for every nonnegative integer r. This proves our claim that that when χ is real the distribution of $N^{-1/2}S_\chi(n_0, n_0+N)$ approaches the real normal distribution of mean 0 and variance 1 as (p, N) varies over a family with $N \to \infty$ and $\log p/\log N \to \infty$.

In the case that χ is complex new difficulty arises. Fix $n_1, \ldots, n_r$ and $n_1', \ldots, n_{r'}'$, and let D, D' be the polynomials

$$D(X) := \prod_{i=1}^{r}(X+n_i), \quad D'(X) := \prod_{i'=1}^{r'}(X+n_{i'}').$$

Assume that these polynomials are distinct (equivalently, that the $n_{i'}'$ are not a permutation of the n_i). For every character χ mod p set

$$\Phi_\chi := \sum_{n_0=0}^{p-1}\chi\Bigl(\prod_{i=1}^{r}(n_0+n_i)\Bigr)\,\overline{\chi}\Bigl(\prod_{i'=1}^{r'}(n_0+n_i')\Bigr) = \sum_{n_0=0}^{p-1}\chi(D(n_0))\,\overline{\chi}(D'(n_0)). \quad (16)$$

We assume that $D(X)/D'(X)$ is not of the form $R(X)^f$ for some rational function R and integer $f > 1$, else we may write $R = D_1/D_1'$ in lowest terms and write

$$\Phi_\chi = \sum_{n_0=0}^{p-1}\chi^f(D_1(n_0))\,\overline{\chi}^f(D_1'(n_0)) \;+\; O(r),$$

in which the sum is of the same form as (16) but with D_1/D_1' satisfying our condition on D/D'. (The error $O(r)$ arises because we may have removed as many as r/f factors $\chi^f(n_0+n_i)\overline{\chi}^f(n_0+n_{i'}')$ with $n_i = n_{i'}'$, and each removal introduces a term of norm 1 into the sum where in (16) there was zero.) We then want to prove that $\Phi_\chi \ll_r q^{1/2}$ for every nontrivial character χ.

Let d be the exponent of χ (that is, the least positive integer such that χ^d is the trivial character). Then we expect that Φ_χ will be related with the number of solutions mod p of

$$Y^d = D(X)\,(D'(X))^{d-1}. \quad (17)$$

Indeed this number is $\sum_{j=0}^{d-1}\Phi_{\chi^j}$; the $j = 0$ term is $p - O(1)$, and we expect the remaining terms to be $O_r(p^{1/2})$. The number of solutions of (17) is in turn within $O(r)$ of the number of rational points on the superelliptic curve[4] with

[4] A "superelliptic curve" has the form $Y^d = R(X)$ for some $d > 1$ and rational function R that is not a scalar multiple of an f-th power for any integer f such that $\gcd(f, d) > 1$; a "hyperelliptic curve" is a superelliptic curve with $d = 2$.

equation $Y^d = D(X)/D'(X)$, and Weil's theorem gives an upper bound on the discrepancy between this count and p. But once $d > 2$ we cannot recover Φ_χ from the number of rational points on the curve, because the other Φ_{χ^j} contribute to it as well.

To isolate the individual Φ_{χ^j} we need not just the count of d-th powers among the nonzero values of $D(X)(D'(X))^{d-1}$ but their full distribution among the d classes in $(\mathbf{Z}/p\mathbf{Z})^*/((\mathbf{Z}/p\mathbf{Z})^*)^d$. Equivalently, we need, for each $\zeta \in \mathbf{C}^*$ such that $\zeta^d = 1$, the number of $n_0 \bmod p$ such that $\chi(D(n_0))\,\overline{\chi}(D'(n_0)) = \zeta$. Call this number $F(\zeta)$. Then $\Phi_{\chi^j} = \sum_{\zeta^d=1} \zeta^j F(\zeta)$. In other words, we may regard the map $j \mapsto \Phi_{\chi^j}$ as the discrete Fourier transform of $\zeta \mapsto F(\zeta)$. But each $F(\zeta)$ is related with the number of points on some superelliptic curve over $\mathbf{Z}/p\mathbf{Z}$. For instance, $F(1)$ is within $O(1)$ of the number of solutions mod p of (17), divided by d. More generally, let us fix for each ζ some $c_\zeta \in (\mathbf{Z}/p\mathbf{Z})^*$ such that $\chi(c_\zeta) = \zeta$; then

$$F(\zeta) = \frac{1}{d}\#\{(n_0, y) \in (\mathbf{Z}/p\mathbf{Z})^2 : y \neq 0,\ c_\zeta y^d = D(n_0)\,(D'(n_0))^{d-1}\}. \tag{18}$$

Now the Weil bound applies to the right-hand side of (18). Unfortunately the genus of the superelliptic curve $c_\zeta Y^d = D(X)\,(D'(X))^{d-1}$ can be as large as a positive multiple of $(r + r')d$, so Weil only tells us that $F(\zeta) - p/d = O((r + r')p^{1/2})$, which yields $\Phi_{\chi^j} = \sum_{\zeta^d=1} \zeta^j F(\zeta) = O((r + r')dp^{1/2})$. This is good enough if d is bounded, but in general $(r + r')dp^{1/2}$ is much too large because d can be (and typically is) as large as $p - 1$.

Fortunately Weil's theory gives more information than just the size of $F(\zeta) - (p/d)$: it decomposes the difference (up to the usual $O(1)$ due to points on the curve where y is zero or infinite) as a sum of d contributions that exactly correspond with the Φ_{χ^j}, and bounds each of them by $O((r + r')p^{1/2})$, without a factor of d in the error estimate. (Again we see that the bound contains the same factor $p^{1/2}$ we expect from the behavior of sums of random numbers, but does not match exactly this behavior because $p^{-1/2}\Phi_{\chi^j}$ is bounded.) Therefore we finally obtain an estimate

$$\frac{1}{p}\sum_{n_0=0}^{p-1}\left(\frac{S_\chi(n_0, n_0+N)}{N^{1/2}}\right)^r \left(\frac{S_{\overline{\chi}}(n_0, n_0+N)}{N^{1/2}}\right)^{r'}$$

$$= \frac{1 + O_r(N^{-1/2})}{2\pi} \iint\limits_{(x,y)\in\mathbf{R}^2} (x+iy)^r\,(x-iy)^{r'}\,e^{-(x^2+y^2)/2}\,dx\,dy\ + O((r+r')p^{-1/2}N^r) \tag{19}$$

for all nonnegative integers r, r'. This proves the complex case of the Davenport-Erdős theorem: when χ is complex the distribution of $N^{-1/2}S_\chi(n_0, n_0 + N)$ approaches the complex normal distribution of mean 0 and variance 1 as (p, N) varies over a family with $N\to\infty$ and $\log p/\log N\to\infty$.

The Davenport-Erdős theorem applies equally when we generalize the sum $S_\chi(n_0, n_0 + N) = \sum_{n=1}^{N} \chi(n_0 + n)$ to $\sum_{n\in\mathcal{N}} \chi(n_0 + n)$ for any N-element subset of $\mathbf{Z}/p\mathbf{Z}$. Indeed, when $\mathcal{N}$ is fixed and n_0 varies, the moments of the resulting sums satisfy the same estimates that we showed in the special case $\mathcal{N} = \{1, 2, 3, \ldots, N\}$ (with error terms depending on p, N but not $\chi, \mathcal{N}$), and with the same proof. Thus the sums $N^{-1/2} \sum_{n \in \mathcal{N}} \chi(n_0 + n)$ approach the same real or complex normal distributions as $N \to \infty$ and $\log p / \log N \to \infty$.

This means that such techniques cannot be strong enough to produce nontrivial bounds on individual sums $S_\chi(n_0, n_0 + N)$: the bounds would then apply equally to all $\sum_{n \in \mathcal{N}} \chi(n_0 + n)$, but it is easy to find $\mathcal{N}$ that makes the sum as large as the trivial bound N for a single choice of n_0: simply make $\mathcal{N}$ an arbitrary subset of $-n_0 + \ker(\chi)$. (For instance, if χ is the quadratic chracater mod p, choose N numbers of the form $n = x^2 - n_0$.) Hence nontrivial bounds on individual sums $S_\chi(n_0, n_0 + N)$ must exploit the structure of $\mathcal{N}$ in the special case $\mathcal{N} = \{1, 2, 3, \ldots, N\}$.

We start from the fact that the set $\{1, 2, 3, \ldots, N\}$ contains many translates of $\{1, 2, 3, \ldots, H\}$ for $H = o(N)$. This means that if a single $|S_\chi(n_0, n_0 + N)|$ attains the trivial bound of N then each of $N - H + 1$ values of $|S_\chi(n_0', n_0' + H)|$ attains its trivial bound of H. More generally, if instead of $|S_\chi(n_0, n_0 + N)| = N$ we assume only that $|S_\chi(n_0, n_0 + N)|$ is unusually large then the same is true of $|S_\chi(n_0', n_0' + H)|$ for many choices of n_0'. Indeed we have

$$S_\chi(n_0, n_0 + N) = \frac{1}{H} \sum_{h=1}^{N} S_\chi(n_0 + h, n_0 + h + H) + O(H). \tag{20}$$

By Hölder's inequality[5] we have for $r = 1, 2, 3, \ldots$

$$\left| \sum_{h=1}^{N} S_\chi(n_0 + h, n_0 + h + H) \right| \le N^{1 - \frac{1}{2r}} \Big(\sum_{h=1}^{N} |S_\chi(n_0 + h, n_0 + h + H)|^{2r} \Big)^{1/2r}.$$

We can then use

$$\sum_{h=1}^{N} |S_\chi(n_0 + h, n_0 + h + H)|^{2r} \le \sum_{m=0}^{p-1} |S_\chi(m, m + H)|^{2r} \ll_r pH^r + p^{1/2} H^{2r}.$$

Unfortunately the cost of bounding the sum of $|S_\chi(m, m + H)|^{2r}$ over the N values $m = n_0 + h$ by the sum over all p choices is too high: there is no value of r,

[5] This inequality asserts that for $p, q > 1$ with $p^{-1} + q^{-1} = 1$ the upper bound

$$\sum_{i=1}^{n} x_j \overline{y_j} \le \Big(\sum_{i=1}^{n} |x_j|^p \Big)^{1/p} \Big(\sum_{i=1}^{n} |y_j|^q \Big)^{1/q}$$

holds for all $x_j, y_j \in \mathbf{C}$, with equality if and only if one of the vectors $(x_j), (y_j) \in \mathbf{C}^n$ is a nonnegative multiple of the other. The inequality asserts in effect that the l^p and l^q norms on $\mathbf{C}^n$ are each other's dual; the special case $p = q = 2$ is the Cauchy-Schwarz inequality. In the present application we have $y_j = 1$ for each j, a special case that is equivalent to the inequality of the means: for fixed $x_j \ge 0$, the p-th power mean $(n^{-1} \sum_{i=1}^{n} x_j^p)^{1/p}$ is an increasing function of p.

N, and $H < N$ that makes the resulting upper bound on $|S_\chi(n_0, n_0{+}N)|$ smaller than both the trivial bound N and the Pólya-Vinogradov bound $O(p^{1/2}\log p)$, even when the $O(H)$ error in (20) is ignored.

Burgess's key idea for going beyond both the trivial bound and Pólya-Vinogradov is the observation that $\{1, 2, 3, \ldots, N\}$ contains not just many translates of $\{1, 2, 3, \ldots, H\}$ but even more images of $\{1, 2, 3, \ldots, H\}$ under affine transformations — that is, arithmetic progressions of length H. If $|S_\chi(n_0, n_0 + N)| = N$ then $|\sum_{n=1}^{H} \chi(n_0' + dn)| = H$ for all n_0' and $d > 0$ such that $n_0' + d > n_0$ and $n_0' + Hd \le n_0 + N$. But χ is multiplicative, so

$$\sum_{n=1}^{H} \chi(n_0' + dn) = \chi(d) \sum_{n=1}^{H} \chi(d^{-1}(n_0' + dn)) = \chi(d)\, S_\chi(d^{-1}n_0, d^{-1}n_0 + H),$$

whence $\left|S_\chi(d^{-1}n_0, d^{-1}n_0 + H)\right| = H$ as well, where d^{-1} is the multiplicative inverse of $d \bmod p$. Proceeding as before, we try to bound $S_\chi(n_0, n_0 + N)$ by writing

$$\begin{aligned} S_\chi(n_0, n_0 + N) &= \frac{1}{H} \sum_{h=1}^{N} \chi(n_0 + h + dn) + O(Hd) \qquad (21) \\ &= \frac{\chi(d)}{H} \sum_{h=1}^{N} \sum_{n=1}^{H} S_\chi(d^{-1}(n_0 + h), d^{-1}(n_0 + h) + H) + O(Hd) \end{aligned}$$

for each $d < N/H$. Varying both d and h, we now expect to obtain an upper bound on $S_\chi(n_0, n_0{+}N)$ in terms of the sum of N^2/H powers $|S_\chi(n_0', n_0'{+}H)|^{2r}$ rather than only N, and thus to incur a lower penalty when replacing that sum by $\sum_{m=1}^{p} |S_\chi(m, m + H)|^{2r}$.

To see what would happen, suppose we could ignore the edge effects $O(Hd)$ in (21). We would then have

$$(?) \qquad S_\chi(n_0, n_0 + N) \ll \frac{1}{N} \sum_{d=1}^{N/H} \sum_{h=1}^{N} \left|S_\chi(d^{-1}(n_0 + h), d^{-1}(n_0 + h) + H)\right|,$$

and by Hölder the double sum is

$$\ll \left(\frac{N^2}{H}\right)^{1-\frac{1}{2r}} \Bigl(\sum_{d=1}^{N/H} \sum_{h=1}^{N} \left|S_\chi(d^{-1}(n_0 + h), d^{-1}(n_0 + h) + H)\right|^{2r}\Bigr)^{1/2r}.$$

We want to bound the sum over (d, h) by

$$\sum_{m=1}^{p-1} |S_\chi(m, m + H)|^{2r} \ll_r pH^r + p^{1/2}H^{2r}.$$

This would give

$$S_\chi(n_0, n_0 + N) \ll_r \frac{1}{N} \left(\frac{N^2}{H}\right)^{1-\frac{1}{2r}} (p^{1/2r}H^{1/2} + p^{1/4r}H),$$

and we would take $H \approx p^{1/2r}$ to balance $p^{1/2r}H^{1/2}$ with $p^{1/4r}H$, finally obtaining

$$(??) \qquad S_\chi(n_0, n_0+N) \ll_r N^{1-\frac{1}{r}} p^{\frac{r+1}{4r^2}}.$$

This would be a genuine improvement for suitable N and $r > 1$; for instance if $N \approx p^{1/2}$ we could take $r = 2$ to get $S_\chi(n_0, n_0+N) \ll p^{3/8}$, of lower order than both N and $p^{1/2}\log p$. For any N, the bound (??) for $r = 2$ and $n_0 = 0$ would give $S_\chi(N)/N^{1/2} \ll p^{3/16}$, and thus $L(1/2+it) \ll p^{3/16}\log p$. We note for future reference that if $N = p^\nu$ then (??) improves on both $|S_\chi(n_0, n_0+N)| < N$ and $|S_\chi(n_0, n_0+N)| \ll p^{1/2}\log p$ for $1/4 < \nu - (1/4r) \leq 1/2$; in particular for each $\nu > 1/4$ we would get a nontrivial bound by taking r large enough, and would thus prove the existence of a quadratic nonresidue mod p less than $p^{1/4\sqrt{e}+o(1)}$. For $r = 1$, (??) recovers the Pólya-Vinogradov bound without the $\log p$ factor (already hinting that we cannot expect to quite prove (??) as it stands), and as $r \to \infty$ (??) converges to the trivial bound N. Thus for finite integers $r > 1$ we may regard (??) as an interpolation between those two bounds.

But we have two difficulties to overcome before obtaining a bound such as (??). One is the "edge effects" error $O(Hd)$ that we ignored in (?): if we combine all these errors in the average over $d < N/H$, we get $O(N)$, which obviates any improvement over the trivial bound on $S_\chi(n_0, n_0+N)$. Fortunately these error terms are themselves averages over shorter character sums $S_\chi(n_0', n_0'+N')$, so we will be able to use induction on N to reduce the edge effects to a constant factor.

A more serious difficulty is that we cannot bound the sum over $d^{-1}(n_0+h)$ in (?) by a sum over m, because the same m may arise as $d^{-1}(n_0+h)$ in more than one way. There are several ways to overcome this diffculty, at the cost of a fractional power of $\log(p)$ which will not significantly affect the quality of our bound. When $n_0 = 0$ (the case relevant to bounds on $L(1/2+it, \chi)$), we can simply restrict the sum to prime d, because then the rational numbers h/d are all distinct, and remain distinct mod p provided their denominators do not exceed $p^{1/2}$. (Recall that $d \leq N/H$ and we shall take $H = p^{1/2r}$ while N will be no greater than $p^{1/2+(1/4r)}$.) This reduces the number of (d, h) pairs by a factor $O(\log p)$, and thus increases the bound on $S_\chi(0, N)$ by $O(\log p)^{1-(1/2r)}$.

For arbitrary n_0, we can argue instead as follows. Fix $D < N/H$, and vary d over all integers in $[1, D]$. For m mod p let c_m be the number of representations of m as $(n_0+h)/d$ with $1 \leq h \leq N$ and $1 \leq d \leq D$. Then

$$\frac{1}{N}\sum_{d=1}^{D}\sum_{h=1}^{N} \left|S_\chi(d^{-1}(n_0+h), d^{-1}(n_0+h)+H)\right| = \sum_{m=0}^{p-1} c_m \left|S_\chi(m, m+H)\right|$$

and Hölder bounds the sum over m by

$$\left(\sum_{m=0}^{p-1} c_m^{2r/(2r-1)}\right)^{1-(1/2r)} \left(\sum_{m=0}^{p-1} |S_\chi(m, m+H)|\right)^{1/2r}.$$

If all c_m were 0 or 1, the first factor would be $(ND)^{1-1/(2r)}$, which is what we used to obtain (??). Since some c_m may exceed 1, the actual bound is larger, but we shall show that in fact

$$\left(\sum_{m=0}^{p-1} c_m^{2r/(2r-1)}\right)^{1-(1/2r)} \ll \log(p)^{1/r}\,(ND)^{1-(2r)}),$$

and thus that only a fractional power of $\log(p)$ is required to fix (??).

Clearly $\sum_{m=0}^{p-1} c_m = ND$. We shall show that $\sum_{m=0}^{p-1} c_m^2 \ll (ND)^2/p + ND\log p$, from which the desired estimate will follow by another application of Hölder as long as $ND < p$. We can certainly assume $ND < p$, because $ND < N^2/H \approx p^{-1/2r}N^2$ which as already noted does not exceed p for any N that makes (??) better than Pólya-Vinogradov. Now $\sum_{m=0}^{p-1} c_m^2$ is the number of solutions of the congruence $(n_0+h)/d \equiv (n_0+h')/d' \bmod p$ in positive integers $h, h' \leq N$, $d, d' \leq D$. Equivalently, we must have

$$kp = d'(n_0+h) - d(n_0+h') = (d'-d)n_0 + d'h - dh' \tag{22}$$

for h, h', d, d' as above and $k \in \mathbf{Z}$. Fix d, d', and let $d_0 = \gcd(d, d')$. Then $|d'h - dh'| \leq ND$, so there are $O(1 + ND/(d_0 p))$ choices for k. Given k, if there is any solution (h_0, h_0') of $kp - (d'-d)n_0 = d'h - dh'$ then the general solution is $(h_0, h_0') + j(d/d_0, d'/d_0)$ for $j \in \mathbf{Z}$, so the number of solutions is $O(Nd_0/\max(d,d'))$. Thus the total number of solutions (k, h, h') of (22) is $O(Nd_0 + N^2D/p)/\max(d,d')$. It remains to sum this over $d, d' \leq D$. The sum of $N^2D/(p\max(d,d'))$ is $< 2(ND)^2/p$. Given $d_0 < D$, the sum of $1/\max(d,d')$ over $d, d' < D$ such that $\gcd(d,d') = d_0$ is at most $(1/d_0)\sum_{e,e'<D/d_0}\max(e,e') \ll \log(D/d_0)/d_0 < \log(p)/d_0$. Therefore our upper bound on $\sum_{m=0}^{p-1} c_m^2$ is

$$\frac{(ND)^2}{p} + \log p \sum_{d_0=1}^{D} N = ND\log p + \frac{(ND)^2}{p},$$

as claimed.

We can now prove:

Theorem. *For every positive integer r, nontrivial character χ modulo a prime p, and integers n_0 mod p and $n < N$, we have*

$$S_\chi(n_0, n_0+N) \ll_r N^{1-\frac{1}{r}} p^{\frac{r+1}{4r^2}} (\log p)^{1/r}. \tag{23}$$

Proof:

TO BE CONTINUED...

Exercises

1. Suppose χ is a cubic character modulo a prime p, that is, a nontrivial Dirichlet character such that χ^3 is trivial. Show that if $|S_\chi(N)| < \epsilon N$ then there exists $x < N^{\exp(2(\epsilon-1)/3)+o(1)}$ such that $\chi(x) \neq 1$.

[Note that $\chi(x) \neq 1$ if and only if x is not a cubic residue. See Theorem 2 of [Davenport-Erdős 1952] for a generalization to k-th power nonresidues; the power of N that occurs for $k > 3$ is smaller than the quadratic and cubic case suggests, due to overcounting of numbers divisible by more than one large prime.]

2. Take $N = \lfloor p/3 \rfloor$ and $p - \lfloor p/3 \rfloor$ instead of $N = \lfloor p/2 \rfloor$ in Lemma 1 to increase the constant 1/4 in the Corollary to that Lemma. (The improved constant will depend on whether χ is real or complex; in the real case you should be able to get $\sqrt{2}/3$.)

3. Prove the integral formulas (5) and (6). [For the former, change variables to obtain a multiple of $\Gamma((r+1)/2)$; for the latter, use polar coordinates.]

4. Show that the error $O_r(1/N)$ in (15) can be replaced by $O(r^2/N)$ with a universal implied constant. What is the corresponding result for (19)?

References

[Burgess 1957] Burgess, D.A.: The distribution of quadratic residues and non-residues, *Mathematika* **4** (1957), 106–112.

[Burgess 1962] Burgess, D.A.: On character sums and primitive roots, *Proc. London Math. Soc. (3)* **12** (1962), 179–192.

[Burgess 1962a] Burgess, D.A.: On character sums and L-series, *Proc. London Math. Soc. (3)* **12** (1962), 193–206.

[Davenport-Erdős 1952] Davenport, H., and Erdős, P.: The distribution of quadratic and higher residues, *Publicationen Mathematicae (Debrecen)* **2** (1952), 252–265.

[Hasse 1936] Hasse, H.: Zur Theorie der abstrakten elliptischen Funktionenkörper. III. Die Struktur des Meromorphismenrings. Die Riemannsche Vermutung. *J. reiner angew. Math.* **175** (1936), 193–208.

[Weil 1945] Weil, A.: Sur les courbes algébriques et les variétés qui s'en déduisent, *Actualités math. sci.* **1041** (Paris, 1945), Deuxième Partie, § IV.