CHARACTER SUMS

TAE KYU KIM

1. Introduction

In this expository paper, I introduce the character sum problem and give the elementary proof of the Polya-Vinogradov inequality, which gives a non-trivial bound for the sum. Afterwards, I briefly discuss the reason for which we might be interested in bounding the character sums.

2. Definitions

Definition 2.1. Let m be a positive integer. A *Dirichlet character modulo* m is a function $\chi : \mathbb{Z} \to \mathbb{C}$ satisfying the following properties:

(1) For any integers a and b, we have $\chi(ab) = \chi(a)\chi(b)$.
(2) χ is period with period m, i.e. $\chi(a+m) = \chi(a)$ for all integers a.
(3) $\chi(a) = 0$ if and only if $\gcd(a, m) > 1$.

Example. The *principal character* χ modulo m is defined to be

$$\chi(a) = \begin{cases} 1 & \gcd(a,m) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Example. Let $\zeta = e^{i\pi/3}$ be a primitive 6^{th} root of unity. Then let

$$\chi(a) = \begin{cases} 0 & a \equiv 0 \pmod 7, \\ 1 & a \equiv 1 \pmod 7, \\ \zeta^2 & a \equiv 2 \pmod 7, \\ \zeta & a \equiv 3 \pmod 7, \\ \zeta^4 & a \equiv 4 \pmod 7 \\ \zeta^5 & a \equiv 5 \pmod 7, \\ \zeta^3 = -1 & a \equiv 6 \pmod 7. \end{cases}$$

You can check that χ is indeed a character.

For complex characters, we also have a notion of a conjugate $\overline{\chi}$ where we take the output of $\overline{\chi}$ to be the complex conjugate of the output of χ for any input.

Definition 2.2. A character χ modulo q is said to be *induced* by another character χ_1 modulo q_1 if $q_1 | q$, and $\chi(n) = \chi_1(n)$ for $\gcd(n, q) = 1$ and $\chi(n) = 0$ otherwise.
A character χ is *primitive* if it is not induced by any other character.

Date: November 29, 2019.

Example. The following χ_1 modulo 5 is primitive

$$\chi_1(a) = \begin{cases} 0 & a \equiv 0 \pmod 5, \\ 1 & \text{otherwise.} \end{cases}$$

The following χ modulo 10 is induced by χ_1:

$$\chi(a) = \begin{cases} 0 & a \equiv 0 \pmod 5 \text{ ,or } a \text{ even}, \\ 1 & \text{otherwise.} \end{cases}$$

3. Problem Statement and Some Easy Bounds

We wish to study the upper bounds on the maximum sum attainable by summing $\chi(M)$ to $\chi(N)$. Formally, for any χ modulo q, we wish to find a function $f(q)$ such that for any integers $M < N$,

$$\left| \sum_{i=M}^{N} \chi(i) \right| \ll f(q).$$

We want to focus our attention to nonprincipal characters, as the sum of principal characters is unbounded (see above example).

Proposition 3.1. *Two easy bounds are*

$$\left| \sum_{i=M}^{N} \chi(i) \right| \leq q$$

and

$$\left| \sum_{i=M}^{N} \chi(i) \right| \leq \phi(q)$$

where ϕ denotes the Euler totient function.

Proof. We start by showing that

$$\sum_{a=0}^{q-1} \chi(a) = 0.$$

for all nonprincipal χ modulo q.

Since χ is nonprincipal, there is some b with $\gcd(b, q) = 1$ such that $\chi(b) \neq 1$. Suppose that $\sum_{a=0}^{q-1} \chi(a) = s$. Then, multiplying both sides by $\chi(b)$, we have

$$s\chi(b) = \chi(b) \sum_{a=0}^{q-1} \chi(a) = \sum_{a=0}^{q-1} \chi(ab).$$

As a runs from 0 to $q-1$, ab takes on each value modulo m exactly once. That is, multiplying the modular residues by b simply permutes the residues into a different order. Thus $\sum_{a=0}^{q-1} \chi(ab)$ is the same sum as $\sum_{a=0}^{q-1} \chi(a)$, but in a scrambled order. Thus we have $s\chi(b) = s$, or $s = 0$.

It's not very difficult to see that $|\chi(a)| = 0$ or 1 for all non-negative integers a because χ is multiplicative. Furthermore, as q consecutive terms always cancel out to 0, the sum can

accumulate to at most q.

There are actually only $\phi(q)$ modular residues of q that have nonzero character values (those that are relatively prime to q). Hence, the sum can be at most $\phi(q)$. ■

4. The Polya-Vinogradov Inequality

Theorem 4.1. *An improvement on the upper bounds is the* Polya-Vinogradov inequality *which says that*

$$\left|\sum_{n=M}^{N} \chi(n)\right| \leq 2\sqrt{q}\log q.$$

In order to prove this inequality, we will follow Schur's elementary proof as described by Davenport in his book [DM13]. We first show the following two lemmas:

Lemma 4.2. *Define the* Gaussian sum *to be*

$$\tau(\chi) = \sum_{m=1}^{q} \chi(m)e_q(m).$$

Then,

$$\chi(n) = \frac{1}{\tau(\overline{\chi})}\sum_{m=1}^{q} \overline{\chi}(m)e_q(mn)$$

for primitive χ and any integer n coprime to q.

Here, the function $e_q(m) = e^{\frac{2\pi i m}{q}}$.

Proof. If $\gcd(n,q) = 1$, then we can use the fact that $\chi(n) = \overline{\chi}(n^{-1})$ and that there exists h such that $m \equiv nh \pmod q$ to write

$$\begin{aligned}\chi(n)\tau(\overline{\chi}) &= \sum_{m=1}^{q} \overline{\chi}(m)\chi(n)e_q(m)\\ &= \sum_{h=1}^{q} \overline{\chi}(h)e_q(nh).\end{aligned}$$

This gives our desired expression for $\gcd(n,q) = 1$ and $\tau(\chi) \neq 0$.

Now, we let $\gcd(n,q) > 1$, so we can put

$$\frac{n}{q} = \frac{n_1}{q_1},$$

where $\gcd(n_1, q_1) = 1$ and $q_1 | q$ and $q_1 < q$. When $q_1 = 1$, n is a multiple of q, so the relation easily holds.

As $\chi(n) = 0$ when $\gcd(n,q) > 1$, we want to prove that

$$\chi(n)\tau(\overline{\chi}) = \sum_{h=1}^{q} \overline{\chi}(h)e(n_1 h/q_1) = 0,$$

where $e(n) = e^{2\pi i n}$ so that $e(n) = e(n+1)$ for all real n.

Now we write $q = q_1 q_2$ and put $h = uq_1 + v$ where

$$0 \leq u < q_2, \quad 1 \leq v \leq q_1.$$

Substituting in the expression for h gives us

$$\begin{aligned}
&\sum_{v=1}^{q_1}\sum_{u=0}^{q_2-1}\overline{\chi}(uq_1+v)e\left(n_1u+\frac{n_1v}{q_1}\right)\\
&=\sum_{v=1}^{q_1}\sum_{u=0}^{q_2-1}\overline{\chi}(uq_1+v)e\left(\frac{n_1v}{q_1}\right)\\
&=\sum_{v=1}^{q_1}\left[e\left(\frac{n_1v}{q_1}\right)\sum_{u=0}^{q_2-1}\overline{\chi}(uq_1+v)\right],
\end{aligned}$$

so it suffices to show that the inner sum is 0 for all v. We will denote this inner sum as $S(v)$, as a function of v.

The function $S(v)$ is periodic with period q_1 because the effect of replacing v by $v+q_1$ changes the range of u to $1 \le u \le q_2$, and the terms for $u = q_2$ and $u = 0$ are equivalent. Then, if c is any number satisfying $\gcd(c, q) = 1$ and $c \equiv 1 \pmod{q_1}$, then

$$\chi(c)S(v) = \sum_{u=0}^{q_2-1}\overline{\chi}(cuq_1+cv) = \sum_{u=0}^{q_2-1}\overline{\chi}(uq_1+cv) = S(v),$$

as $\chi(c) = \chi(c^{-1}) = \overline{\chi}(c)$.

Now, if we can find a c such that $\chi(c) \neq 1$, then we would know that $S(v) = 0$. To show that there exists some c with $\chi(c) \neq 1$, we use a characteristic property of primitive characters, namely that for $\gcd(n, q) = 1$, the function $\chi(n)$ is not periodic to any modulus q_1 that is a proper factor of q. This implies that there exists integers c_1, c_2 such that

$$\gcd(c_1, q) = \gcd(c_2, q) = 1, \quad c_1 \equiv c_2 \pmod{q_1}, \quad \chi(c_1) \neq \chi(c_2).$$

So there exists $c \equiv c_1c_2^{-1}$ which satisfies our constraints on c and has $\chi(c) \neq 1$. It follows that $S(v) = 0$, which completes our proof. ■

Lemma 4.3. *For a primitive character χ modulo q,*

$$|\tau(\chi)| = \sqrt{q}.$$

Proof. We can multiply the equation from Lemma 1.4 and the conjugate of that equation to get the square of the norm of the left side:

$$|\chi(n)|^2|\tau(\chi)|^2 = \sum_{h_1=1}^{q}\sum_{h_2=1}^{q}\overline{\chi}(h_1)\chi(h_2)e_q[n(h_1-h_2)].$$

Now sum for n over a complete set of residues $\pmod q$. The sum of the values of $|\chi(n)|^2$ is $\phi(q)$ (as each nonzero term is exactly 1), and the sum of the exponentials is 0 unless $h_1 \equiv h_2$. (To see this, consider some fixed h_1 and h_2 such that $h_1 \not\equiv h_2$. Summing the terms $\overline{\chi}(h_1)\chi(h_2)e_q[n(h_1-h_2)]$ over the residues $\pmod q$ gives all of the qth roots of unity, which gives 0 when summed up.) Hence,

$$\phi(q)|\tau(\chi)|^2 = q\sum_{h=1}^{q}|\chi(h)|^2 = q\phi(q),$$

which gives our desired expression. ■

The following is the proof of the Polya-Vinogradov inequality.

Proof. First let χ be primitive with $q > 1$. From lemma 4.2, we have

$$\chi(n) = \frac{1}{\tau(\overline{\chi})} \sum_{m=1}^{q} \overline{\chi}(m) e_q(mn).$$

We also saw from lemma 4.3 that $|\tau(\chi)| = \sqrt{q}$. Substituting this, we get

$$\begin{aligned} \left| \sum_{n=M}^{N} \chi(n) \right| &= \frac{1}{\sqrt{q}} \left| \sum_{n=M}^{N} \sum_{k=1}^{q} \overline{\chi}(k) e(kn/q) \right| \\ &= \frac{1}{\sqrt{q}} \left| \sum_{k=1}^{q} \overline{\chi}(k) \sum_{n=M}^{N} e(kn/q) \right|. \end{aligned}$$

The inner sum is a geometric series, and using the identity $e^{ix} - e^{-ix} = 2 \sin ix$ we can show that the expression is

$$\begin{aligned} &\frac{1}{\sqrt{q}} \left| \sum_{k=1}^{q} \overline{\chi}(k) e\left(\frac{(M + \frac{1}{2}N + \frac{1}{2})k}{q} \right) \frac{\sin \pi N k/q}{\sin \pi k/q} \right| \\ &\leq \frac{1}{\sqrt{q}} \sum_{k=1}^{q-1} \frac{1}{|\sin \pi k/q|}. \end{aligned}$$

This sum appears to be the Riemann sum for $\frac{1}{sin \pi x}$, and because the function is convex (the second derivative is positive), we can bound the sum using the integral:

$$\frac{1}{\sqrt{q}} \sum_{k=1}^{q-1} \frac{1}{|\sin \pi k/q|} \leq \sqrt{q} \int_{\frac{1}{2q}}^{1-(\frac{1}{2q})} \frac{1}{\sin \pi \beta} d\beta.$$

As $\sin \pi \beta$ is symmetric about $\frac{1}{2}$, we have

$$= 2\sqrt{q} \int_{\frac{1}{2q}}^{\frac{1}{2}} \frac{1}{\sin \pi \beta}.$$

Using that $\sin \pi \beta > 2\beta$ for $0 < \beta < \frac{1}{2}$, we get

$$< 2\sqrt{q} \int_{\frac{1}{2q}}^{\frac{1}{2}} \frac{1}{2\beta} = \sqrt{q} \log q.$$

Hence we have the inequality for primitive χ.

Now, assume that χ is induced by a character $\chi_1 \pmod{q'}$ with $q = q'r$. Then $\chi(n) = \chi_1(n)$ whenever $(n, r) = 1$ and 0 otherwise, so

$$\left| \sum_{n=M}^{N} \chi(n) \right| = \left| \sum_{\substack{n=M \\ \gcd(n,r)=1}}^{N} \chi_1(n) \right|.$$

We use the Mobius function $\mu * \mathbb{1} = I$ to represent the characteristic function of $\gcd(n, r)$. Then,

$$\begin{aligned}\left|\sum_{\substack{n=M\\ \gcd(n,r)=1}}^{N} \chi_1(n)\right| &= \left|\sum_{n=M}^{N} \sum_{d|\gcd(n,r)} \mu(d)\chi_1(n)\right| \\ &= \left|\sum_{d|r} \mu(d) \sum_{\substack{n=M\\ d|n}}^{N} \chi_1(n)\right| \\ &= \left|\sum_{d|r} \mu(d) \sum_{m=M/d}^{N/d} \chi_1(dm)\right| \\ &= \left|\sum_{d|r} \mu(d)\chi_1(d) \sum_{m=M}^{N/d} \chi_1(m)\right|.\end{aligned}$$

Now using the Polya-Vinogradov inequality to the primitive character χ_1, we see that the sum is

$$\sqrt{q_1}\log q_1 \left|\sum_{d|n} \mu(d)\chi_1(d)\right| \leq \sqrt{q_1}\log q_1 d(r),$$

where $d(r)$ is the number of divisors function (each term in the new sum is at most 1, and there are $d(r)$ terms). We can use the inequality $d(r) \leq 2\sqrt{r}$ because there are at most $\sqrt{r}$ divisors up to $\sqrt{r}$, and the total number of divisors is at most twice that. Hence, using $\sqrt{q} = \sqrt{q_1}\sqrt{r}$,

$$\left|\sum_{n=M}^{N} \chi(n)\right| \leq 2\sqrt{q}\log q.$$

■

The Polya-Vinogradov Inequality is impressive in that it has optimal bounds of $\sqrt{q}$ (plus some logarithm factors), which has not been broken even with stronger assumptions such as the Generalized Riemann Hypothesis.

5. Applications of the Bound

An application of the result is on the worst-case bounds of the first quadratic non-residue for a large prime q. A residue r is a quadratic residue of q if there exists a such that $r \equiv a^2 \pmod{q}$. We can create a Dirichlet character by assigning 0 when $\gcd(r, q) > 1$, 1 when r is a quadratic residue, -1 when it is a non-quadratic residue.

Because the partial sum of any character is bounded for any prime modulo q, that means that a non-quadratic residue must occur before that bound is reached. The same applies to quadratic residues, except that in that case it's a bit boring because 1 is a quadratic residue.

One of the reasons we care about this result about quadratic residues is that we want to observe the worst-case short-term behavior of number-theoretic sets. For random sets in probability, it turns out that short-term behavior can fluctuate greatly even when globally,

the set has a normal distribution (This is the central limit theorem). This means that we cannot say anything local from our global statistics. So, instead of looking at purely random sets, we look at sets with some structure, like quadratic residues with their multiplicative property. The nonzero values of quadratic residues $\{-1, 1\}$ are like coin flips, but unlike completely random sets, we can use their special property to better approximate their short-term behavior [Tao09].

In fact, using the property, one can "amplify" the bad behavior in one short intervals to other short intervals to influence the global behavior. Hence, we can actually obtain a better bound for the first non-quadratic residue. The Burgess bound does this, and it gives a bound on the order of $q^{\frac{1}{4}}$ plus some logarithm factors. Assuming the Generalized Riemann Hypothesis gives us a bound on the order of $(\log q)^2$ [jon14].

References

[DM13] H. Davenport and H.L. Montgomery. *Multiplicative Number Theory.* Graduate Texts in Mathematics. Springer New York, 2013.

[jon14] Character sums and pólya-vinogradov inequality, Oct 2014.

[Tao09] Terence Tao. The least quadratic nonresidue, and the square root barrier, Aug 2009.