ON SIEGEL EXCEPTIONAL ZEROS AND SIEGEL'S THEOREM

ZHENGHENG BAO AND TAN VO

ABSTRACT. In this note, we will review Siegel zeros, sketch the proof of Siegel's theorem which gives an estimation of the values of exceptional zeros, and summarize some results concerning their existence.

1. PRELIMINARIES

Let's first recall the definitions of Dirichlet character and Dirichlet L-functions.
The following definitions are taken from [1] chapter 4 and chapter 1.

Definition 1.1. [1] A Dirichlet character modulo q is a function $\chi : \mathbb{Z} \to \mathbb{C}$ satisfying:
1. χ is periodic with period q;
2. $\chi(n) \neq 0$ if and only if $(n, q) = 1$;
3. χ is totally multiplicative.

Definition 1.2. [1] A Dirichlet character χ modulo q is called principal, if $\chi(n) = 1$ when $(n, q) = 1$ and $\chi(n) = 0$ otherwise. We denote such a character by χ_0. Note that χ_0 still depends on q.

A Dirichlet character χ modulo q is called imprimitive, if there exists $d|q, d < q$ such that $\chi(m) = \chi(n)$ whenever $m \equiv n(\text{mod } d)$ and $(mn, q) = 1$. If χ is not imprimitive, then we say that χ is primitive.

Definition 1.3. [1] Let χ be a Dirichlet character modulo q. Its Dirichlet L-function is defined as

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$$

for $\Re(s) > 1$. It can be proved that $L(s, \chi)$ can be analytically continued to a meromorphic function on $\mathbb{C}$. Furthermore, its only pole is at $s = 1$ when χ is a principal character, and is entire when χ is non-principal.

Note that when $\chi(n) = 1$ for all $n \in \mathbb{N}$, in other words, a Dirichlet character modulo 1, then the L-function is the Riemann zeta function $\zeta(s)$.

Definition 1.4. [1] We call a Dirichlet character χ quadratic, if χ^2 is a primitive character. Equivalently, we can define it as a Dirichlet character that only takes real values, so it is also called a real character.

Notation 1.5. [1] For $s \in \mathbb{C}$, we denote the real part of s by σ, and the imaginary part of s by t. We also denote $\tau = |t| + 4$.

2. DEFINITION OF EXCEPTIONAL ZEROS

It has been proved that the Dirichlet L-function has no zeros in a classical zero-free region, except for a special case which we will discuss below. Such a zero is called an exceptional zero.

Theorem 2.1. *([1] p.360, theorem 11.3) There is an absolute constant $c > 0$ such that if χ is a Dirichlet character modulo q, then the region*

$$R_q = \{s : \sigma > 1 - \frac{c}{\log q\tau}\}$$

contains no zero of $L(s,\chi)$ unless χ is a quadratic character, in which case $L(s,\chi)$ has at most one, necessarily real, zero $\beta < 1$ in R_q.

Definition 2.2. Zeros of $L(s,\chi)$ located in the region R_q above are called Siegel zeros, or exceptional zeros. These two terms are used interchangeably.

Remark 2.3. Note that this absolute constant c does not depend on q or χ. For each single χ, since there is at most one exceptional zero, we can change the constant c so that the region doesn't contain any zero. However, this c is an absolute constant, so we are not necessarily able to change the constant c in this way to avoid exceptional zeros for all q and χ.

Also note that it has not been proved whether exceptional zeros exist or not, so these zeros are hypothetical zeros. Many people believe that the generalized Riemann hypothesis is true, in other words, all zeros of $L(s,\chi)$ have real part $1/2$, which means that exceptional zeros do not exist.

Now we sketch the proof of theorem 2.1, which is mainly based on [1] p.360–362, theorem 11.3.

Proof. Here we only sketch the proof when χ is a complex character, in other words not a real character.

Lemma 11.1 in [1] tells us that for $5/6 \le \sigma \le 2$, we have

$$-\frac{L'}{L}(s,\chi) = \frac{E_0(\chi)}{s-1} - \sum_{\rho} \frac{1}{s-\rho} + O(\log q\tau)$$

where the sum is over the zeros ρ of $L(s,\chi)$ in a certain region, and $E_0(\chi) = 1$ if χ is principal and 0 otherwise.

Using a trigonometric identity, lemma 11.2 in [1] shows that for $\sigma > 1$, we have

$$\Re\Big(-3\frac{L'}{L}(\sigma,\chi_0) - 4\frac{L'}{L}(\sigma+it,\chi) - \frac{L'}{L}(\sigma+2it,\chi^2)\Big) \ge 0.$$

Roughly speaking, if L has a zero at $s = \beta_0 + i\gamma_0$, by plugging in appropriate values of σ and t, the second term $-4\frac{L'}{L}(\sigma+it,\chi)$ gives a pole of residue -4. Observe that the only way to contribute a positive residue is in $\frac{E_0(\chi_0)}{s-1}$ and $\frac{E_0(\chi^2)}{s-1}$, therefore the first term contributes a residue at most 3. The third term cannot contribute a positive residue because χ^2 is non-principal since χ is complex, which is why we need this assumption. Hence we have reached our desired contradiction, as the error term is negligible in our desired classical zero-free region. $\square$

3. Estimations of exceptional zeros and $L(s,\chi)$

In this section, we sketch the proof of Siegel's theorem, which gives an upper bound of the value of $L(1,\chi)$ and the exceptional zeros, if they exist. In other words, they cannot stay too close to 1.

The following theorem allows us to estimate the value of the exceptional zero β_1, by estimating the value of $L(1,\chi)$.

Theorem 3.1. *([1] p.362–364, theorem 11.4) Let χ be a non-principal character modulo q, let c be the constant in theorem 2.1, and suppose that $\sigma \ge 1 - c/(2\log q\tau)$.*

If $L(s,\chi)$ has no exceptional zero, or if β_1 is an exceptional zero of $L(s,\chi)$ but $|s-\beta_1| \geq 1/\log q$, then

$$\frac{1}{L(s,\chi)} \ll \log q\tau.$$

Alternatively, if β_1 is an exceptional zero of $L(s,\chi)$ and $|s-\beta_1| \leq 1/\log q$, then

$$|s-\beta_1| \ll |L(s,\chi)| \ll |s-\beta_1|(\log q)^2.$$

Proof. See [1] p. 362–364, theorem 11.4. □

Remark 3.2. The first case is what we hope to happen: having no exceptional zeros, or the exceptional zeros are not too close to 1. If the second case happens, the above theorem tells us that $|s-\beta_1|$ and $|L(s,\chi)|$ are only two $\log$'s away from each other, so we only need to estimate $L(1,\chi)$, to estimate β_1.

Using Mellin transform, Page showed the following results.

Theorem 3.3. *([1] p. 370, theorem 11.11) If χ is a quadratic character modulo q, then $L(1,\chi) \gg q^{-1/2}$.*

Proof. See [2] p. 117, or [1] p. 370, theorem 11.11. □

On the other hand, Siegel showed the following theorems, which improves the previous theorem of $q^{-1/2}$ to an arbitrary negative power of q.

Theorem 3.4. *([1] p. 372–373, theorem 11.14) (Siegel) For each positive number ϵ there is a positive constant $C(\epsilon)$ such that if χ is a quadratic character modulo q, then*

$$L(1,\chi) > C(\epsilon)q^{-\epsilon}.$$

Before giving the proof of this theorem, which is mainly based on [1] p. 372–373, theorem 11.14, we need the following two lemmas, whose proofs can be found in [1] as cited below.

Lemma 3.5. *([1] p. 350, lemma 10.15) Let χ be a non-principal character modulo q, and suppose that $\delta > 0$ is fixed. Then*

$$L(s,\chi) \ll (1+(q\tau)^{1-\sigma}) \min(\frac{1}{|\sigma-1|}, \log q\tau)$$

uniformly for $\delta \leq \sigma \leq 2$.

Remark 3.6. Note that this lemma is different from theorem 3.1. Theorem 3.1 gives an estimation of $L(s,\chi)$ in terms of β_1, so that we can transform our problem to an estimation of $L(s,\chi)$, while this lemma gives an upper bound of $L(s,\chi)$ only in terms of s and χ, by estimating the L-function directly from its formula. Such an upper bound is helpful, because we will, as we did in theorem 2.1, pair $L(1,\chi)$ with some other L-functions, which we need an upper bound.

Lemma 3.7. *([1] p. 370, lemma 11.13) (Estermann) Suppose that $f(s)$ is analytic for $|s-2| \leq 3/2$, and that $|f(s)| \leq M$ for s in this disc. Suppose also that*

$$F(s) = \zeta(s)f(s) = \sum_{n=1}^{\infty} r(n)n^{-s}$$

for $\sigma > 1$, that $r(1) = 1$, and that $r(n) \geq 0$ for all n. If there is a $\sigma \in [19/20, 1)$ such that $f(\sigma) \geq 0$, then

$$f(1) \geq \frac{1}{4}(1-\sigma)M^{-3(1-\sigma)}.$$

Remark 3.8. On one hand, lemma 3.5 gives us an upper bound for $L(s,\chi)$. On the other hand, suppose there is a zero $\sigma \in [19/20, 1)$ for f, which we will later plug in some L-functions, lemma 3.7 gives a lower bound for f. Although these bounds don't contradict each other, in other words cannot help us rule out exceptional zeros, there will be q and $L(1,\chi)$ involved to help us estimate.

Usually, if we have a zero β_1 close to 1, and if the function $L(s,\chi)$ is nice enough, we would expect $L(1,\chi)$ cannot be too large. However, this lemma gives us a surprise: a close zero gives a large lower bound! Its proof is based on Landau's theorem ([1] theorem 1.7), which gives a singularity of a Dirichlet series of nonnegative coefficients, which are the $r(n)$ in this theorem.

Now we prove theorem 3.4, which is mainly following [1] p. 372, theorem 11.14. In the proof of theorem 2.1, we paired a character with its conjugate. This time, we still need to pair the character χ with another character, namely, a fixed character with an exceptional zero.

Proof. The case where there's no exceptional zero or all exceptional zeros are far from 1 is relatively easy, as well as generalizing the result from primitive characters to imprimitive characters. So we assume that χ is primitive and that there is a character χ_1 modulo q_1 such that $L(s,\chi_1)$ has a real zero $\beta_1 \geq 1 - \epsilon/4$.

Let χ be a primitive quadratic character (recall that only quadratic characters may have exceptional zeros), and $\chi \neq \chi_1$. Let $f(s) = L(s,\chi)L(s,\chi_1)L(s,\chi\chi_1)$ be the function f in lemma 3.7. Then

$$\begin{aligned}\log(\zeta(s)f(s)) &= \log(\zeta(s)) + \log(L(s,\chi)) + \log(L(s,\chi_1)) + \log(L(s,\chi\chi_1))\\ &= \sum_{n=2}^{\infty} \frac{\Lambda(n)}{\log n}(1+\chi(n)+\chi_1(n)+\chi\chi_1(n))n^{-s}\end{aligned}$$

Since $(1+\chi(n)+\chi_1(n)+\chi\chi_1(n)) = (1+\chi(n))(1+\chi_1(n))$, and $\chi(n), \chi_1(n) = 0, \pm 1$ because the characters are quadratic, so the coefficients of $\log(\zeta(s)f(s))$ are nonnegative. Let $g(s) = \log(\zeta(s)f(s))$. To see that the coefficients of $\zeta(s)f(s) = \exp g(s)$ in the Dirichlet series are nonnegative, let $g(s) = \sum_{n=1}^{\infty} a_n n^{-s}$. Then $\exp g(s) = \sum_{m=0}^{\infty} \frac{1}{m!}(\sum_{n=1}^{\infty} a_n n^{-s})^m$. By expanding the m-th power, we will get $\sum a_{n_1} \dots a_{n_m}(n_1 \dots n_m)^{-s}$ where the sum is over all m-tuples of positive integers $n_1, \dots, n_m$, which is still a Dirichlet series. This tells us that $\exp(g(s)) = \zeta(s)f(s)$ has nonnegative coefficients. The condition on the first coefficient of $\zeta(s)f(s)$ is not important, as we can scale f, because our estimation of f will contain a constant.

To find an upper bound M for f within the disk $|s-2| \leq 3/2$, we claim that we can take $M = C_3(qq_1)^{4/3}$ for some constant C_3. To see this, we apply lemma 3.5. Since $\sigma \geq 1/2$ in this disk, we know that $1 + (q\tau)^{1-\sigma} \ll q^{1/2}$ (note that τ is bounded in this disk) by lemma 3.5. So $L(1,\chi) \ll q^{1/2}\log q$, $L(1,\chi_1) \ll q_1^{1/2}\log q_1$, and $L(1,\chi\chi_1) \ll (qq_1)^{1/2}\log(qq_1)$. Multiply them together, we get our claim for M.

Now we have checked all hypotheses for lemma 3.7. Applying this lemma (with $\sigma = \beta_1$, and recall that we are considering the case where $\beta_1 \geq 1 - \epsilon/4$, and that q_1 is fixed), we get

$$f(1) \geq \frac{1}{4}(C_3(qq_1)^{4/3})^{-3(1-\beta_1)} \geq \frac{1}{4}(C_3(qq_1)^{4/3})^{-3\epsilon/4} \geq C_4(\epsilon)q^{-\epsilon}$$

for some constant $C_4(\epsilon)$. This gives a lower bound for $f(1)$.

To get an upper bound for $L(1,\chi_1)L(1,\chi\chi_1)$, by lemma 3.5 again, noticing that $\chi \neq \chi_1$ so $\chi\chi_1$ is non-principal, we know that

$$f(1) = L(1,\chi)L(1,\chi_1)L(1,\chi\chi_1) \ll L(1,\chi)\log(qq_1)^2$$

Combining these two inequalities, we get $L(1,\chi) \geq C_5(\epsilon)q^{-2\epsilon}$ for some constant $C_5(\epsilon)$. This is true for all $\chi \neq \chi_1$.

When $\chi = \chi_1$, since $L(1,\chi_1) \neq 0$, we can just pick a constant $C(\epsilon)$ such that $L(1,\chi_1) > C(\epsilon)q_1^{-2\epsilon}$. This finishes our proof. □

Remark 3.9. Note that the constant in theorem 3.3 is effective, meaning that there is a way to compute the constant by following the proof. However, the constant in Siegel's theorem 3.4 is not effective. This is because the constant came out when we concluded $L(1,\chi) \gg q^{-2\epsilon}$ from $C_4(\epsilon)q^{-\epsilon} \ll L(1,\chi)(\log qq_1)^2$, noticing that q_1 is fixed. However, q_1 is hypothetical and we have no way to compute its value, so there's no way to compute the constant in the theorem. There is no way to compute the constant for χ_1 itself either.

By theorem 3.1, the estimation of $L(1,\chi)$ easily gives the following bound on exceptional zeros.

Corollary 3.10. *([1] p. 373, corollary 11.15) For any $\epsilon > 0$ there is a positive number $C(\epsilon)$ such that if χ is a quadratic character modulo q and β is a real zero of $L(s,\chi)$, then $\beta < 1 - C(\epsilon)q^{-\epsilon}$.*

4. Rarity of Siegel zeros

The upper bound in the last section has many applications, including a better error term in the prime number theorem on arithmetic progressions. However, it would be even better if we can prove that exceptional zeros do not exist at all, which is believed by many people. Although this has not been proved yet, there are some partial results: if they exist, they occur very rarely. See the citations below for proofs.

Theorem 4.1. *([1] p. 368, corollary 11.8) For each positive integer q, there is at most one Dirichlet character χ modulo q such that $L(s,\chi)$ has an exceptional zero.*

Theorem 4.2. *([1] p. 368, corollary 11.9) For each positive number A there is a $c(A) > 0$ such that if $\{q_i\}$ is a strictly increasing sequence of natural numbers with the property that for each q_i, there is a primitive quadratic character $\chi_i \pmod{q_i}$ for which $L(s,\chi_i)$ has a zero β_i satisfying $\beta_i > 1 - c(A)/\log q_i$, then $q_{i+1} > q_i^A$.*

Remark 4.3. Note that $\tau = 4$ because exceptional zeros are necessarily real, so this expression is basically the same as the one in theorem 2.1 defining exceptional zeros, except that we may have a different constant. In this case, we know that the next possible modulus for χ is at least a certain power of the previous one, so the possible moduli are rare. Also, theorem 4.1 shows that each modulus gives at most one character with exceptional zeros, so exceptional zeros are rare.

References

[1] H. L. Montgomery and R. C. Vaughan, Multiplicative Number Theory. I. Classical theory, Cambridge Studies in Advanced Mathematics **97**, Cambridge University Press, Cambridge, 2007.

[2] A. Page, On the number of primes in an arithmetic progression. Proc. London Math. Soc. (2), **39**, 116–141 (1935)