MATH 730 ASSIGNMENT: A SUMMARY OF "ITERATION OF ANALYTIC FUNCTIONS" BY CARL LUDWIG SIEGEL

DASHENG WANG

1. BACKGROUND

Let

$$f(z) = \lambda z + \sum_{k=2}^{\infty} a_k z^k, \tag{1}$$

be a holomorphic map defined in some neighborhood U of the fixed point $z = 0$. It is of many interests to know whether or not this map is conjugate to a linear map $w \mapsto \lambda w$, i.e. whether or not we can find a holomorphic change of coordinate $z = \varphi(w)$ such that $\varphi \circ f \circ \varphi^{-1}(w) = \lambda w$. In other words, the following diagram commutes

$$\begin{array}{ccc} U & \xrightarrow{f} & f(U) \\ \varphi \uparrow & & \varphi \uparrow \\ \mathbb{C} & \xrightarrow{\lambda} & \mathbb{C} \end{array}$$

In 1884, Koenigs proved that if $|\lambda| \neq 0, 1$, then the map can be linearized, i.e., the formal solution

$$\varphi(w) = w + \sum_{k=2}^{\infty} c_k w^k \tag{2}$$

to the functional equation

$$f(\varphi(w)) = \varphi(\lambda w), \tag{3}$$

is convergent. We call the functioanl equation the Schröder equation and $\varphi(w)$ the Schröder series because Schröder first wrote the equation in this form in 1871. Leau in 1897 studied the case where the λ is a root of unity ([3]). In 1904, Böttcher treated the case where $\lambda = 0$ ([4]). The last case where $|\lambda| = 1$ and λ is not a root of unity is much harder. In 1938, Cremer proved that for λ satisfying the condition

$$\liminf_{n\to\infty} |\lambda^n - 1|^{1/n} = 0, \tag{4}$$

then $f(z)$ is not linearizable. It wasn't known whether $f(z)$ can be linearized until 1942, when Siegel gave a sufficient condition for $f(z)$ to be linearizable.

2. Main result of the paper

Write $\lambda = e^{2\pi i\xi}$ and let κ be a positive real number, then

Theorem. *If for every rational number p/q, there exists $\epsilon > 0$ such that*

$$\left|\xi - \frac{p}{q}\right| > \frac{\epsilon}{q^\kappa}, \tag{5}$$

then the Schröder series is convergent.

3. the ideas of the proof

The main part of the paper consists of the proofs of three lemmas: Lemma 1 gives a useful inequality for proving lemma 2 and 3, it answers the following kind of question: given the sum of several positive integers and some other constraints, what can we say about a certain product of these integers? The second lemma gives a bound for product of the form $\prod_n |\lambda^n - 1|^{-1}$. Its proof contains the main observation of the whole proof:

Let $\epsilon_n = |\lambda^n - 1|^{-1}$ $(n = 1, 2, \ldots)$, on account of (5), $\epsilon_n < (2n)^\nu$ for some positive value ν. Since

$$\lambda^m(\lambda^{n-m} - 1) = (\lambda^n - 1) - (\lambda^m - 1), \quad 0 < m < n, \tag{6}$$

we have

$$\epsilon_{n-m}^{-1} \leq \epsilon_n^{-1} + \epsilon_m^{-1}, \tag{7}$$

$$\min(\epsilon_n, \epsilon_m) \leq 2\epsilon_{n-m} < 2^{\nu+1}(n-m)^\nu. \tag{8}$$

Then the author introduces another useful set of parameters $\{\delta_n\}$: let $\delta_1 = 1$, and for every $k > 1$, let μ_k denote the maximum of all products $\delta_{l_1}, \delta_{l_2}, \ldots, \delta_{l_r}$ with $l_1 + l_2 + \cdots + l_r = k > l_1 \geq l_2 \geq \cdots \geq l_r \geq 1$ and let $\delta_k = \epsilon_{k-1}\mu_k$. The third lemma gives a bound for δ_k.

The proof of the theorem works roughly as follows:

By (1), (2) and (3),

$$\sum_{k=2}^{\infty} c_k \lambda(\lambda^{k-1} - 1) w^k = \sum_{l=2}^{\infty} a_l \left(w + \sum_{r=2}^{\infty} c_r w^r\right)^l. \tag{9}$$

The author considered an analogue of this equation

$$\sum_{k=2}^{\infty} \eta_k \gamma_k w^k = \sum_{l=2}^{\infty} \left(w + \sum_{r=2}^{\infty} \gamma_r w^r\right)^l, \tag{10}$$

where $\eta_2, \eta_3, \dots$ are positive parameters. Then the coefficients $\gamma_1 = 1, \gamma_2, \gamma_3, \dots$ are uniquely determined by the formula

$$\gamma_k = \eta_k^{-1} \sum \gamma_{l_1}\gamma_{l_2}\cdots\gamma_{l_r} \quad (k = 2, 3, \cdots), \tag{11}$$

where $l_1, \dots, l_r$ run over all positive integral solutions of $l_1 + \cdots + l_r = k$ $(r = 2, ..., k)$. Now another two parameters σ_k and τ_k are introduced: set $\eta_k = \epsilon_{k-1}^{-1}$ and let $\sigma_k = \gamma_k$; set $\eta_k = 1$ and let $\tau_k = \gamma_k$. Then by induction on k and use the second lemma, it is proved that

$$\sigma_k \leq \delta_k \tau_k. \tag{12}$$

Let $\psi = \sum_{k=1}^{\infty} \tau_k w^k$, then

$$\psi - w = (1 - w)^{-1}\psi^2 \tag{13}$$

hence

$$4\psi = 1 + w - (1 - 6w + w^2)^{\frac{1}{2}} \tag{14}$$

in order for it to have real solutions, determinant $(2|w| - 2)^2 - 16|w| > 0$, i.e. $|w| < 3 - 2\sqrt{2}$. By (9), (10) and (12),

$$|c_k| \leq \delta_k \tau_k \quad (k = 2, 3, \dots). \tag{15}$$

Then it follows from the third lemma that the Schröder series converges in the disk $|w| < (3 - 2\sqrt{2})2^{-5\nu-1}$.

4. Follow-up work

A natural question to ask is whether this condition can be refined. A condition given by Bryuno in 1965 involves continued fraction.

Definition. *Let $\xi \in (0, 1)$ be an irrational number, consider the continued fraction expansion*

$$\xi = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}} \tag{16}$$

where the a_i are uniquely defined strictly positive integers. The rational number

$$\frac{p_n}{q_n} = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots + \cfrac{1}{a_{n-1}}}} \tag{17}$$

is called the nth convergent to ξ.

Theorem. (Bryuno) *With* $\lambda = e^{2\pi i \xi}$ *and* $\{q_n\}$ *as above, if*

$$\sum_n \frac{\log(q_{n+1})}{q_n} < \infty, \tag{18}$$

then $f(z)$ *is linearizable.*

In 1988, Yoccoz proved that this is a best possible result, i.e.,

Theorem. (Yoccoz) *If the sum in* (18) *diverges,* $f(z)$ *is not linearizable.*

5. what I like about the paper

As Selberg once said, "The things that Siegel tended to do were usually things that seemed impossible. Also after they were done, they still seemed almost impossible". It is hard to think of anything when you are given an extremely complicated thing like (9). However, through his idea of introducing those parameters, the problem is transferred to something one can analyze. This ability to create something from seemingly nothing amazes me.

6. Improvements of the paper

6.1. **Lack of details.** In sereval parts of the paper, there are not obvious steps that are given without justifications, for example, in page 608, why are the two conditions in the theorem equvilent? Why does (7) imply (6)?...

6.2. **The writing style is not very heuristic.** For example, in the paragraph following the theorem, the author wrote "where λ and μ denote positive numbers depending only upon ω", it gives an impression that one can simutaneously control the two parameters to achieve the inequailty, but in fact μ is a given, the task is to find λ, so it is better to make the logic clearer by using "given..., if there exists..., then..." style.

Secondly, it would be better if the author could explain the motivation when introducing a new variable or parameter, so the reader can better understand the proof.

References

[1] C.L. Siegel, Iteration of Analytic Functions, Annals of Mathematics, Vol. 43, 1942

[2] J. Milnor, Dynamics in One Complex Variable, Princeton University Press, 2006

[3] L. Leau, Etude sur les equations fonctionelles aune ou plusiers variables, Ann. Fac. Sci. Toulouse 11, EI-EI10, 1897

[4] L. E. Bottcher, The principal laws of convergence of iterates and their application to analysis (Russian), Izv. Kazan. Fiz.-Mat. Obshch. 14, 155-234, 1904