Siegel Disk Notes and Write Up

David Blessing * and J.D. Mireles James †

Florida Atlantic University, Department of Mathematical Sciences

February 1, 2021

Abstract

Siegel disks are the first case where small divisors were overcome. This is largely due to the near complete removal of geometry from the problem. We look at parameterizing Siegel disks via Fourier series and set the problem up for a computer assisted proof.

Contents

*Email: `dblessing2014@fau.edu`

†J.M.J partially supported by NSF grant DMS - 1318172 Email: `jmirelesjames@fau.edu`

1 Introduction

1.1 Purpose

The plan is to study Siegel discs (based on [1]) in a fashion that will allow an understanding of the one dimensional KAM theorem and computer assisted proofs on Siegel disks.

This document will serve as:

- notes
- bibliography
- explaination of numerics
- general information on the problem

I intend to keep detailed notes of my understanding of the task at hand here, perhaps a section or two of this will be able to be re-purposed for an actual paper.
Additionally, this paper is intimately linked to the MatLab files on Siegel disks, changing those files will change this document.

1.2 Basic Setup

We start with an analytic map $f : \mathbb{C} \to \mathbb{C}$, defined by

$$\begin{aligned} f(z) &= az + N(z) \\ N(0) &= 0 \\ N'(0) &= 0 \end{aligned}$$

on some sufficiently small neighborhood of the origin. Do note that the above conditions imply $N = O(z^2)$.

We demand that $|a| = 1$ in order for the dynamics to be potentially conjugate to a rotation near the origin, i.e.

$$f \circ h(z) = h(az)$$

on some neighborhood of $0 \in \mathbb{C}$ for some analytic $h : \mathbb{C} \to \mathbb{C}$ with $h(0) = 0$, to be determined.

We will frequently write $f(z) = \sum_{n\in\mathbb{Z}} \hat{f}_n e_n$ where e_n are the basis functions of the space that f lives in. Then $\hat{f}_n = \int_{S^1} f\overline{e_n}$ For our purposes e_n will be $z^n = e^{2\pi i n\theta}$ on S^1 the unit circle in $\mathbb{C}$.

1.3 What is a Siegel Disk?

Definition (Siegel Disk)**.** *A connected component of the Fatou set where the dynamics are conjugate to an irrational rotation of the disk.*

Definition (Fatou Set)**.** *This is the compliment of the Julia Set, defined later. Roughly the Fatou set is a set where iterates of points behave similarly to their neighbors.*

Example (Behavior in Fatou Components)**.** *If f is a non-linear rational function on the extended complex plane, then for a periodic component U of the Fatou set, exactly one holds*

- *U contains an attracting periodic point.*
- *U is parabolic*
- *U is a Siegel Disc*
- *U is a Herman ring*

In case one in iterested:

Definition (Herman Ring)**.** *A component of the Fatou set where the rational function is conformally conjugate to an irrational rotation of the standard annulus.*

Definition (Confromal Mapping)**.** *Let $U \subseteq \mathbb{C}$ be open, $f : U \to \mathbb{C}$ that is holomorphic and it's derivative is everywhere non-zero on U. This leads to the property of local angle preservation.*

In the literature there are many mentions of the following:

Definition (Julia Set)**.** *The compliment of the Fatou set, that is values whose behavior is sensitive to small perturbations. Points near by each other can have dramatically different orbits under the mapping.*

In summary, the Fatou set is where the behavior of the function is "regular", whereas the Julia set is where the function is "chaotic". A Siegel Disk is a class of component of the Fatou set, where the function's behavior is analytically conjugate to an irrational rotation of the disk.

Remark. In [1], De la Llave requires h to be near the identity, that is $h'(0) = 1 = \hat{h_1}$. Computer numerics are better suited to working with smaller h_1 values, since small h_1 allows the control of the tail size. This is because an $\hat{h_1}$ near 1 leads to large tails of the Taylor series. Which does not play well with small divisors through the coarse lens of floating point numbers. So, our approach is fixes the radius of the pre-image to 1, then vary the $\hat{h_1} < 1$.

The system we consider is

$$f(z) = az + z^2 \qquad \text{with} \qquad a = e^{\varphi i}. \tag{1}$$

where $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden ratio or some other Diophantine irrational number. Since h is analytic in the disk, write $h(z) = \sum_{k\in\mathbb{N}} \hat{h_k} z^k$, $\hat{h_k} \in \mathbb{C}$, noting that $\hat{h_0} = 0$.

We manipulate the cohomology equation,

$$f \circ h(z) = h(az) \tag{2}$$

$$ah(z) + (h(z))^2 = h(az) \tag{3}$$

$$a\sum_{k\in\mathbb{N}} \hat{h_k} z^k + \left(\sum_{k\in\mathbb{N}} \hat{h_k} z^k\right)^2 = \sum_{k\in\mathbb{N}} \hat{h_k} (az)^k \tag{4}$$

We now equate powers of z, noting $\hat{h}_0 = 0$ and $\hat{h}_1$ is free to be assigned.

$$(a^k - a)h_k = \sum_{j=1}^{k-1} h_j h_{k-j} \tag{5}$$

$$h_k = \frac{\overline{a}\sum_{j=1}^{k-1} h_j h_{k-j}}{a^{k-1} - 1} \tag{6}$$

$$h_k = \frac{e^{-\varphi i}\sum_{j=1}^{k-1} h_j h_{k-j}}{e^{(k-1)\varphi i} - 1} \tag{7}$$

This is a formal solution since there is question of the convergence, namely concerning the small divisor $e^{(k-1)\varphi i} - 1$. The numerator will not effect divergence since $e^{-\varphi i}(h(z))^2$ is convergent, it is the composition of analytic functions, specifically h and $e^{-\varphi i} z^2$.

So, by equating coefficients we have a formal solution for the conjugacy equation. By changing the value of $\hat{h}_1$ we can change the image of $\mathbb{S}^1$, that is the topological circle $h(\mathbb{S}^1)$. We will be interested in finding how large $\hat{h}_1$ can be while still maintaining convergence on this boundary of the disk.

For the sake of numerical simplicity we now focus our attention on the boundary of the image. In an abuse of notation we will write h for the restriction $h|_{\mathbb{S}^1}$ throughout the sequel. We may consider our Taylor series as Fourier series as we are always considering the domain to be $\mathbb{S}^1$, so the composition $h\left(e^{i\theta}\right) = \sum_{k=1}^{N} \hat{h_k} e^{ik\theta}$ converts to Fourier series.

1.4 Function Space

We describe the function space that our approximation theory will take place in. We choose the Sobolev norm to assure a certain amount of regularity for the solutions. Do note that we really don't need the full generalization of the following because we are working only with analytic functions (that means they are locally expressible as a power series so they are infinitely differentiable and integrable). However, for my personal understanding we include some more generality.

The one dimensional Sobolev space $W^{k,p}(\mathbb{C})$ (there may be some incorrect translation from the real case to complex case) is a subspace of $L^p(\mathbb{C})$ such that functions and their weak derivatives up to order k have finite L^p norm.

Definition (Lesbegue Space L^p). *Let $1 \le p < \infty$. These are functions $f : S \to \mathbb{C}$ whose p-th power is absolutely Lebesgue integrable, identifying as expected functions that agree a.e. This is a vector space under pointwise addition and scalar multiplication. It becomes a normed space with the norm:*

$$\|f\|_p = \left(\int_S |f|^p \, \mathrm{d}\mu \right)^{1/p}$$

where (S, Σ, μ) is a measure space (we will use $S = [0, 1]$), and $f : S \to \mathbb{C}$ is a measurable function whose p-th power is absolutely integrable.

We will use Σ as the Borel sets on S under the standard topology, and μ to be the Lebesgue measure.

As we are working in Sobolev spaces, we need the notion of weak derivatives. Here is the vanilla definition of weak derivative,

Definition (Weak Derivative). *Let $u \in L^1([a, b])$. We say $v \in L^1([a, b])$ is the weak derivative of u if*

$$\int_a^b u(t)\varphi'(t)\, dt = -\int_a^b v(t)\varphi(t)\, dt$$

for all $\varphi \in C^\infty([a, b])$.

This formula comes from integration by parts. We can generalize this to $\mathbb{R}$ then to $\mathbb{R}^{2n}$ then to $\mathbb{C}^n$, we just look at $\mathbb{R}^n$ for simplicity, as the others are just special cases.

Definition (Weak Derivative on $U \subseteq \mathbb{R}^n$). *Let U be open, and $u, v \in L^1_{loc}(U)$ meaning u, v are integrable on every compact subset of U. We say that v is the α-th weak derivative of u, some $\alpha \in \mathbb{N}^n$, if*

$$\int_U u \frac{\partial^{|\alpha|}\varphi}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}} = (-1)^{|\alpha|} \int_U v\varphi$$

for all $\varphi \in C_c^\infty(U)$ that is smooth functions on U with compact support.

We now define the Sobolev space.

Definition (The Sobolev space $W^{k,p}(\mathbb{C})$ for $1 \le p < \infty$ and $k \in \mathbb{N}$). *Let $f \in L^p(\mathbb{C})$ such that $\|f\|_{k,p} < \infty$ where*

$$\|f\|_{k,p} = \left(\sum_{i=0}^k \|f^{(i)}\|_p^p \right)^{1/p} = \left(\sum_{i=0}^k \int_{\mathbb{S}^1} |f^{(i)}|^p \right)^{1/p}$$

and $f^{(i)}$ is the i-th weak derivative of f.

From now on we will only consider $p = 2$ because the function space $H^s = W^{k,2}$ is a Hilbert space, and we want this structure for the pen-and-paper analysis of computer assisted proofs.

Definition (Parseval's Identity). *Let H be a Hilbert space, with inner product $\langle \cdot, \cdot \rangle$, (e_n) an orthonormal basis for H, then,*

$$\sum_n |\langle x, e_n \rangle|^2 = \|x\|^2$$

The following paragraph needs some work, I am not quite sure how to set up the function spaces on $\mathbb{S}^1$.

We will work with functions from the circle $\mathbb{S}^1 \subseteq \mathbb{C} = \mathbb{R}^2$. Thought of as a smooth manifold with the usual smooth structure. For the purposes of integration, we will consider the chart with target $[0, 1)$ and coordinate function induced by the map $e^{2\pi i\theta} : [0, 1) \to \mathbb{S}^1$. The inner product on $H^k(\mathbb{S}^1)$ is $\langle f, g\rangle = \int_{\mathbb{S}^1} f\overline{g}$. Under this inner product, the functions $(e^{2\pi in\theta})$ for $n \in \mathbb{Z}$ form an orthonormal basis when we identify $H^k(\mathbb{S}^1)$ with the functions $f \in H^k([0, 1])$ such that $f(0) = f(1)$.

Putting this together,

$$\begin{aligned}
\|f\|_{k,2}^2 &= \sum_{j=0} \|f^{(j)}\|_2^2 = \sum_{j=0}^{k} \sum_{n\in\mathbb{Z}} |\langle f^{(j)}, e^{2\pi in\theta}\rangle|^2 = \sum_{n\in\mathbb{Z}} \sum_{j=0}^{k} |\langle f^{(j)}, e^{2\pi in\theta}\rangle|^2 \\
&= \sum_{n\in\mathbb{Z}} \sum_{j=0}^{k} \left| \int_{\mathbb{S}^1} f^{(j)}(\theta)\overline{e^{2\pi in\theta}} \right|^2 = \sum_{n\in\mathbb{Z}} \sum_{j=0}^{k} \left| \int_{\mathbb{S}^1} f^{(j)}(\theta)e^{-2\pi in\theta} \right|^2 = \sum_{n\in\mathbb{Z}} \sum_{j=0}^{k} \left| \hat{f_n^{(j)}} \right|^2
\end{aligned}$$

We need to figure out the relationship between $\hat{f}_n$ and $\hat{f_n^{(j)}}$. Assuming f has a Fourier series expansion,

$$f^{(j)}(\theta) = \frac{d^j}{d\theta^j} \sum_{n\in\mathbb{Z}} \hat{f}_n e^{2\pi in\theta} = \sum_{n\in\mathbb{Z}} \hat{f}_n \frac{d^j}{d\theta^j} e^{2\pi in\theta} = \sum_{n\in\mathbb{Z}} \hat{f}_n (2\pi in)^j e^{2\pi in\theta}$$

So, $\hat{f_n^{(j)}} = (2\pi in)^j \hat{f}_n$, therefore the norm may be expressed as:

$$\begin{aligned}
\|f\|_{k,2}^2 &= \sum_{n\in\mathbb{Z}} \sum_{j=0}^{k} \left| (2\pi in)^j \hat{f}_n \right|^2 = \sum_{n\in\mathbb{Z}} \sum_{j=0}^{k} |2\pi n|^{2j} \left| \hat{f}_n \right|^2 = \sum_{n\in\mathbb{Z}} \left(\left| \hat{f}_n \right|^2 \sum_{j=0}^{k} |2\pi n|^{2j} \right) \\
&= \sum_{n\in\mathbb{Z}} \left(1 + (2\pi n)^2 + \cdots + (2\pi n)^{2k} \right) \left| \hat{f}_n \right|^2
\end{aligned}$$

Definition (Equivalent Norms)**.** *Two norms are said to be equivalent if they generate the same topology. Two norms on a vector space V are equivalent if there exist $c, C > 0$ such that $\forall x \in V$, $c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1$.*

For tradition's sake, we use the equivalent norm (can be seen to be equivalent because it is the norm above with different coefficients on n^{2i}):

$$\|f\|_{k,2}^2 = \left(\sum_{n\in\mathbb{Z}} \left(1 + n^2\right)^k \left| \hat{f}_n \right|^2 \right)^{1/2}$$

as seen in [1] p. 28.

2 Overview of Process

1. Choose f and solve cohomological equation for $\hat{h}_k$
2. Choose $a \in \mathbb{R} \setminus \mathbb{Q}$
3. Plot some orbits?

4. Choose small $\hat{h}_1$ value
5. Compute Taylor coefficents via recursion
6. Check conjugacy error and tail size error
7. Enlarge $\hat{h}_1$ and compute new coefficents
8. Check error and tail size
9. Repeat previous two steps until unable to control error or decay
10. Check Sobolev Norm of resulting parameterizations

Remark. The error is only computed on the truncated series, hence, it may not accurately reflect the non-truncated error. Typically, we are expecting rapid convergence so this should not be a problem.

Remark. We need to set up a Ψ function that h would be an approximate zero of.

3 Validated Numerics

3.1 Multi-Dimensional Reference Material

We record Lemma 2.1 from [3], but there is need for some background definitions in order to understand the statement.

Definition. *Poly-disk*

$$D_r^m(z) = \{w \in \mathbb{C}^m : |w_i - z_i| \leq r, \forall 1 \leq i \leq m\}$$

These are the open balls generated by the ℓ^∞ norm on $\mathbb{C}^m$.

For $\ell < n$, let $\mathcal{X}_r^{\ell,n} = \mathcal{X}_r$ be the space of analytic maps from $D_r^\ell(0) \to \mathbb{C}^n$, mapping the origin to the origin, $Q(0) = 0$. Then, let $\mathcal{Y}_r \subset \mathcal{X}_r$ be the set of functions whose first derivatives vanish at the origin, $Df(0) = 0$ for $f \in \mathcal{X}_r$. Observe $\mathcal{X}_r$ and $\mathcal{Y}_r$ are Banach algebras under the supremum norm

$$\|Q\|_r = \sup_{|z_i|=r} |Q(z)|$$

Lemma 1 (Lemma 2.1)**.** *Let $\{\lambda_1, \dots, \lambda_\ell\} \subset \mathbb{C}$ be complex numbers and fix $\beta \in \mathbb{C}$. Let Λ be the matrix with diagonal entries λ_i. Let $\mathcal{X}_r = \mathcal{X}_r^{\ell,1}$ and $\mathcal{Y}_r$ be as above, so that if $q \in \mathcal{Y}_r$ then $q : D_r^\ell(0) \subset \mathbb{C}^\ell \to \mathbb{C}$, $q(0) = 0$, and $Dq(0) = 0$. Consider the bounded linear operator $\mathcal{L} : \mathcal{Y}_r \to \mathcal{Y}_r$ defined by*

$$\mathcal{L}(q) = \beta q - q \circ \Lambda. \tag{8}$$

Assume that for all multi-indicies α with $|\alpha| \geq 2$, and all $\tau > 0$, β and Λ satisfy the Diophantine condition

$$|\beta - \Lambda^\alpha|^{-1} \leq C_D |\alpha|^\tau. \tag{9}$$

Then for any $p \in \mathcal{Y}_r$ having $\|p\|_r < \infty$, and any $\delta > 0$, the equation

$$\mathcal{L}(q) = p, \tag{10}$$

has unique solution $q \in \mathcal{Y}_{re^{-\delta}}$. Furthermore there is a $\tilde{C}$ depending only on ℓ and C_D, so that

$$\begin{aligned} \|q\|_{re^{-\delta}} &\leq \tilde{C}|\delta|^{-\tau-\ell-1}\|p\|_r \\ \|q \circ \Lambda\|_{re^{-\delta}} &\leq \beta\|q\|_{re^{-\delta}} + \|p\|_r \leq 2\tilde{C}\delta^{-\tau-\ell-1}\|p\|_r. \end{aligned} \tag{11}$$

Proof. Let $p \in \mathcal{Y}_r$ with $\|p\|_r < \infty$, and write

$$p(\theta) = \sum_{|\alpha|\geq 2} p_\alpha \theta^\alpha$$

We desire a $q \in \mathcal{Y}_r$ with corresponding power series decomposition such that

$$\beta q(\theta) - q(\Lambda\theta) = p(\theta)$$

for all $\theta \in D_r^\ell(0) \subset \mathbb{C}^\ell$. Formally, we may solve by comparing like powers,

$$\begin{aligned}
\beta q(\theta) - q(\Lambda\theta) &= \beta \sum_{|\alpha|\geq 2} p_\alpha \theta^\alpha - \sum_{|\alpha|\geq 2} q_\alpha (\Lambda\theta)^\alpha \\
&= \sum_{|\alpha|\geq 2} (\beta - \Lambda^\alpha) q_\alpha \theta^\alpha
\end{aligned}$$

The exponent α on $\Lambda\theta$ is a multi-index so this is component wise possibly should be written $(\Lambda\theta)^{\otimes\alpha}$, the Hadamard or Schur exponent. Since Λ is diagonal, the last equality works out. And so, we have by hypothesis

$$\sum_{|\alpha|\geq 2} (\beta - \Lambda^\alpha) q_\alpha \theta^\alpha = \sum_{|\alpha|\geq 2} p_\alpha \theta^\alpha$$

As is standard in KAM, we match like powers and solve,

$$q_\alpha = \frac{p_\alpha}{\beta - \Lambda^\alpha}$$

We now show that this formal solution is indeed in $\mathcal{Y}_{re^{-\delta}}$. Observe that $q(0) = Dq(0) = 0$ since $p_\alpha = 0 \Leftrightarrow q_\alpha = 0$ because the Diophantine condition guarantees the denominator is never zero. Using the Diophantine condition and Cauchy estimates on p we may show that the norm of q is bounded,

$$\begin{aligned}
\|q\|_{re^{-\delta}} &= \left\| \sum_{|\alpha|\geq 2} q_\alpha \theta^\alpha \right\|_{re^{-\delta}} && \leq \sum_{|\alpha|\geq 2} |q_\alpha| r^{|\alpha|} e^{-|\alpha|\delta} \\
&= \sum_{|\alpha|\geq 2} \left| \frac{p_n}{\beta - \Lambda^\alpha} \right| r^{|\alpha|} e^{-|\alpha|\delta} && \leq \sum_{|\alpha|\geq 2} C_D |\alpha|^\tau |p_\alpha| r^{|\alpha|} e^{-|\alpha|\delta} \\
&\leq \sum_{|\alpha|\leq 2} C_D |\alpha|^\tau \frac{\|p\|_r}{r^{|\alpha|}} r^{|\alpha|} e^{-|\alpha|\delta} && = C_D \|p\|_r \sum_{|\alpha|\geq 2} |\alpha|^\tau e^{-|\alpha|\delta} \\
&\leq \tilde{C} |\delta|^{-\tau-\ell-1} \|p\|_r
\end{aligned}$$

The last inequality comes from bounding the sum by M_ℓ and letting $\tilde{C} = C_D M_\ell$, the bound is

$$\begin{aligned}
\sum_{|\alpha|\geq 2} |\alpha|^\tau e^{-|\alpha|\delta} &\leq \int_{\mathbb{R}^\ell_{x^i\geq 0}} |x|^\tau e^{-|x|\delta}\, dx && \leq \int_0^{\frac{\pi}{2}} \cdots \int_0^{\frac{\pi}{2}} \int_0^\infty r^\tau e^{-r\delta} r^{\ell-1} dr d\phi_1 \dots d\phi_\ell \\
&= \left(\frac{\pi}{2}\right)^\ell \int_0^\infty r^{\tau+\ell-1} e^{-r\delta}\, dr && = \left(\frac{\pi}{2}\right)^\ell \frac{\Gamma(\tau+\ell)}{\delta^{\tau+\ell}} \\
&= M_\ell \delta^{-\tau-\ell}
\end{aligned}$$

The integral is a table lookup $\int_0^\infty x^n e^{-cx}\, dx = \frac{\Gamma(n+1)}{c^{-n-1}}$. This is not quite where we want to end up, however, WLOG $\delta \geq 1$, so we can divide by it in the bound to get the exponent of $-\tau - \delta - 1$ not sure why we would do this. [ASK QUESTION ABOUT THIS]

From this computation we get the first estimate in the lemma as well as q is analytic on $D^\ell_{re^{-\delta}}(0)$, ie $q \in \mathcal{Y}_{re^{-\delta}}$.

We now show the second estimate. We have that

$$q \circ \Lambda = \beta q + p$$

by 8 and 10. Observe, that

$$\begin{aligned}
\|q \circ \Lambda\|_{re^{-\delta}} = \|\beta q + p\|_{re^{-\delta}} \qquad & \leq \beta \|q\|_{re^{-\delta}} + \|p\|_{re^{-\delta}} \\
\leq \tilde{C}|\delta|^{-\tau-\ell-1}\|p\|_r + \|p\|_r \qquad & \leq 2\tilde{C}|\delta|^{-\tau-\ell-1}\|p\|_r
\end{aligned}$$

Under the assumption that $1 \leq \tilde{C}|\delta|^{-\tau-\ell-1}$. □

Remark. The paper mentions that the exponent of δ can be $-\tau - 1$ if better bounds are used, I may have just used a slightly tighter estimate of the sum above. It is fine for the purposes we have in mind, the bounds will work for CAP either way. (I think)

3.2 Translation to 1-Dimension

The goal is to modify this for the case $\ell = n = 1$ and then convert it to be a tail control theorem. Here we will attempt to translate to dimension 1.

Lemma 2. *Let $\lambda, \beta \in \mathbb{C}$, and $\mathcal{X}_r = \mathcal{X}_r^{1,1}$ and $\mathcal{Y}_r$ be as above. Then, for $q \in \mathcal{Y}_r$, we have $q : D^1_r(0) \to \mathbb{C}$, $q(0) = 0$, and $Dq(0) = 0$. Consider the bounded linear operator $\mathcal{L} : \mathcal{Y}_r \to \mathcal{Y}_r$ defined by*

$$\mathcal{L}(q)(z) = \beta q(z) - q(\lambda z)$$

Assume that for all $\alpha \in \mathbb{Z}$ with $|\alpha| \geq 2$, and all $\tau > 0$, β and λ satisfy the Diophantine condition,

$$|\beta - \lambda^\alpha|^{-1} \leq C_D|\alpha|^\tau$$

Then, for any $p \in \mathcal{Y}_r$ with $\|p\|_r < \infty$, and any $\delta > 0$, the equation

$$\mathcal{L}(q) = p$$

has unique solution $q \in \mathcal{Y}_{re^{-\delta}}$. Furthermore, there is a $\tilde{C}$ depending only on C_D, such that

$$\begin{aligned}
&\|q\|_{re^{-\delta}} \leq \tilde{C}|\delta|^{-\tau-2}\|p\|_r \\
&\|q \circ \lambda\|_{re^{-\delta}} \leq \beta\|q\|_{re^{-\delta}} + \|p\|_r \leq 2\tilde{C}\delta^{-\tau-2}\|p\|_r
\end{aligned} \tag{12}$$

Well, that was fairly trivial.

3.3 Translation to Tail Control Theorem

3.4 Estimates and Bounds

I am finding the thought process of bounds to require more careful consideration. While I understand that most of the analysis world is done via inequalities, the specifics of what bound is useful for what task eludes me. More importantly, I would like know know more concretely what the named bounds involved in computer assisted proof are.

For Q analytic and bounded on $D_r^\ell(0)$ and $\delta > 0$, we have the following two inequalities

Definition (Cauchy Bounds)**.**

$$\|DQ\|_{re^{-\delta}} \leq \frac{C_*}{|\delta|}\|Q\|_r$$

Definition (Cauchy Estimates)**.**

$$|q_\alpha| \leq \frac{\|Q\|_r}{r^{|\alpha|}}$$

where C_* depends only on the dimension ℓ of the domain. It is not clear in [3] if C_* is able to be the same constant, however it is of little consequence for our purposes. Knowing that it is a constant times a function is what we are typically after.

The observation here is that Cauchy Bounds bound the norm of the derivative of a function. Whereas Cauchy Estimates bound the size of the Fourier coefficients, which are typically very small, hence they are estimated by their magnitude.

We now consider the notion of composition estimates for parameterizations. Assume $\Phi(K) = f \circ K - K \circ \Lambda \in \mathcal{Y}_r$ and $\epsilon = \|\Phi(K)\|_r$. Let $\rho_* > 0$ such that image$(K) \subset D^n_{\rho_*}(0)$. Then, by the triangle inequality and definition of ϵ,

Definition (Composition estimates for K)**.**

$$\|K \circ \Lambda\|_r \leq \|f \circ K\|_r + \|\Phi(K)\|_r \leq \|f\|_{\rho_*} + \epsilon$$

Moreover, letting $\epsilon' = \|\Phi(K')\|_r$, and assuming $K' = K + \Delta$,

$$\Phi(K + \Delta) = \Phi(K) + Df(K)\Delta + R_K(\Delta) - \Delta \circ \Lambda$$

where R_K is the remainder of the Taylor series for f at K. Rearranging, and assuming $\epsilon' < \epsilon$, we have

Definition (Composition estimates for Δ)**.**

$$\|\Delta \circ \Lambda\|_r \leq 2\epsilon + 2n\|Df\|_{\rho_*}\|\Delta\|_r + \|R_K(\Delta)\|_r$$

There is an explicit bound for R_K derived later, which is rather involved, and does not seem to fit well here. The composition estimates are used to bound the size of $K \circ \Lambda$ and $\Delta \circ \Lambda$ which both appear in $\Phi(K + \Delta)$.

4 MatLab Functions

Here we collect the functions we are going to use in the scripts below.

4.1 compute_coeff.m

```
function P_k = compute_coeff(k, P, a)
% A basic implementation of equation (6), the recursive definition of
% Fourier coefficents.
    P_k = 0;
    for j = 1:k-1
        P_k = P_k + P(k-j+1)*P(j+1);
    end % for loop
    P_k = P_k * conj(a) / (a^(k-1) - 1);
end % function
```

This function just implements the formal sum (7). I believe there is a vectorization that could be done to speed this up, perhaps changing the final multiplications and divisions would help as well.

4.2 evaluate_taylor.m

```
function Pz = evaluate_taylor(P,z)
% Evaluate the Taylor polynomial P at Z
    Pz= 0;
    for j = 1:length(P)
       Pz = Pz + P(j)*z^(j-1);
    end % for loop
end % function
```

Here, we evaluate a vector as if it was a Taylor polynomial at a given point z. This could probably also be vectorized for improved performance.

4.3 map_f.m

```
function newP = map_f(P, a)
% P is input complex taylor list
% a is the parameter
% newP is the image of P under f
    if size(P,1) == 1 % P is scalar
        newP = P*P + a*P;
    else % P is vector
        prod = conv(P,P);
        newP = prod(1:length(P)) + a*P;
    end % if
end % function
```

This function returns the image of a Taylor polynomial under f, truncated to be the same order as the input. The scalar case is listed because sometimes this is used on a point $z \in \mathbb{C}$ instead of a Taylor polynomial.

4.4 fcnPhi.m

```
function defect = fcnPhi(P,a)
% The objective function, we expect to find a zero.
    defect = map_f(P,a) - rotate_by_a(P,a);
end % function
```

Phi is the conjugacy equation expressed as a function.

4.5 rotate_by_a.m

```
function newP = rotate_by_a(P, a)
% Rotate a taylor series P by a
    newP = zeros(length(P),1);
    for j = 1:length(P)
        newP(j) = P(j)*a^(j-1);
    end % for loop
end % function
```

This function takes a Taylor polynomial $P(z)$ and returns $P(az)$. Definitely could be vectorized.

4.6 sobolevNorm.m

```
function sobolevNorm = sobolevNorm(P, s)
% Takes a Taylor series and returns the sobolev norm
% H^s from Rafael de la Llave p 28
if nargin == 1
    s = 1;
end % if
sobolevNorm = 0;
for k = 1:length(P)
    sobolevNorm = sobolevNorm + (1+abs(k-1)^2)^s*abs(
        P(k))^2;
end % for
sobolevNorm = sqrt(sobolevNorm);
end % function
```

As described in Section 1.4, we are computing the Sobolev Norm of the Taylor series when viewed as a Fourier series on the unit circle.

4.7 map_f_preimage.m

```
function [newP1, newP2] = map_f_preimage(z, a)
    dis = sqrt(a^2+4.*z);
    newP1 = (-a + dis)/2;
    newP2 = (-a - dis)/2;
end
```

Here we are computing the two preimages of f given an image z.

5 MatLab Script Files

5.1 Plotting the Julia Set for background

Listing 1: julia_color_script.m

```
% A script to plot the Julia set for z^2 + az
clear, close all
a = exp(1i*(1+sqrt(5))/2);
R = 1000;
maxIter = 1000;
xMin = -1.5;
xMax = 1.5;
yMin = -2;
yMax = 1;
step = .01;
xNum = length(xMin:step:xMax);
yNum = length(yMin:step:yMax);
juliaColor = zeros(xNum, yNum);
juliaSet = [];
x = 1;
offSet = 0;
for j = xMin:step:xMax
    y = 1;
    for k = yMin:step:yMax
        p0 = j + 1i*k;
        p = p0;
        iter = 0;
        for iter = 1:maxIter
            p = p^2 + a*p;
            if abs(p) > R
                break % Stop counting more iterations
            end % if
        end % for loop
        if (iter == maxIter)
            juliaColor(x,y) = 1 + offSet;
        else % not maximum iterations
            juliaColor(x,y) = iter;
        end % if
```

```
        if (y == 1) && (x == 1) && (iter > 0)
            offSet = iter - 1;
        elseif (y == 1) && (x == 1)
            offSet = 0;
        end % if
        y = y + 1;
    end % for loop
    x = x + 1;
end % for loop
contour(log(log(log((juliaColor-offSet)+1)+1))', 'Fill','on')
title('Julia Set')
xlab = xMin:.5:xMax;
xval = (xlab-xMin)/step + 1;
xticks(xval)
xticklabels(xlab)
ylab = yMin:.5:yMax;
yval = (ylab-yMin)/step + 1;
yticks(yval)
yticklabels(ylab)
```

Compute the Julia set of the system. Plot the time to escape via a color map.

5.2 Finding the Siegel Disk

Listing 2: expanding_script.m

```
%% Seigel Disc boundary approximation
% Clean up MatLab
clear, close all, hold on
%% Set up parameters
% Order of approximation
N = 2500;
% Scaling of preimage
r = 1;
% Number of points to plot on each circle
numPoints = 2500;
% Paramerter
% a = exp(1i*(1+sqrt(5))*pi);
% a = exp(1i*(1+sqrt(5))/2*pi);
a = exp(1i*(1+sqrt(5))/2);
% Number of iterations
maxIter = 200;
% Maximum error or norm of last coeff for next iteration
tolerance = 1e-10;
% Minimum step size for h_1 expressed as 10^-maxPower
maxPower = 15;
```

```
% Maximum s in sobolev H^s norm
maxSobolev = 10;
%% Record results for posterity
diary siegel_disk_output.txt
fprintf('f map parameter                  %.15f\n',a)
fprintf('Order of Taylor polynomials      %d\n', N)
fprintf('Radius of preimage               %d\n', r)
fprintf('Number of points plotted         %d\n', numPoints)
fprintf('Maximum defect                   %g\n', tolerance)
fprintf('Minimum step size of h_1         %g\n', 10^(-maxPower))
fprintf('Maximum number of iterations     %d\n\n', maxIter)
%% Set up loop variables
% Storage vector for P
P = zeros(N+1,maxIter);
% Storage for output points
Pz = zeros(numPoints+1,maxIter);
% Limit number of iterations
iter = 0;
% Starting h_1 value
h_1 = 0.1;
maxPower = 15;
% Record keeping
successes = [];
errorValues = [];
tailMagnitude = [];
h_1Values = [];
%% Set up output figure
figSuccesses = figure(1);
set(figSuccesses, 'Units', 'Normalized', 'OuterPosition', [.05 .4 .3 .5]);
title('Siegel Disks')
%% Main Computation Section
for j = 2:maxPower % j will control the attempted step size
    fprintf('Step size %g.\n',(10^(-j)))
    while iter <= maxIter % Repeatedly try larger h_1 values
        iter = iter + 1;
        h_1 = h_1 + (10^(-j)); % Enlarge h_1
        fprintf('h_1 value = %.15f\n', h_1)
        % Compute coeffients
        P(1,iter) = 0;
        P(2,iter) = h_1;
        for k = 2:N
            P(k+1,iter) = compute_coeff(k, P(:,iter), a);
```

```
        end % for loop
        % Look at the last coeff magnitude
        mag = abs(P(end,iter));
        % Compute the image of concentric circles under P
        for m = 1:numPoints+1
            Pz(m,iter) = ...
                evaluate_taylor(P(:,iter),(r)*exp(2*pi*1i*m/numPoints));
        end % for loop
        % Compute the norm of the image under the operator phi
        normF = norm(fcnPhi(P(:,iter), a),2);
        if normF > tolerance || mag > tolerance % Backtrack on failure
            fprintf('Attempt %d failed. \n\n', iter)
            % Zero out failed attempt
            P(:,iter) = zeros(N+1,1);
            Pz(:,iter) = zeros(numPoints+1, 1);
            % Reset P1 to previous value
            h_1 = h_1 - (10^(-j));
            break % Exit the while loop and decreas step size
        else % Report success and plot
            fprintf('Attempt %d successful. \n',iter)
            successes(length(successes)+1) = iter; %#ok<SAGROW>
            h_1Values(length(successes)) = h_1; %#ok<SAGROW>
            fprintf('    Norm of last coefficient: %1.2g\n', mag)
            tailMagnitude(length(successes)) = mag; %#ok<SAGROW>
            fprintf('    Norm of defect: %1.2g\n', normF)
            errorValues(length(successes)) = normF; %#ok<SAGROW>
            plot(Pz(:,iter));
            figure(1)
        end % if
    end % while loop
end % for loop
%% Report end of computation information.
figError = figure(2);
set(figError, 'Units', 'Normalized', 'OuterPosition', [.4 .2 .5 .7]);
subplot(2,2,1);
semilogy(h_1Values, tailMagnitude);
title('Log Tail magnitudes');
xlabel('h_1 value')
```

```
subplot(2,2,2);
semilogy(h_1Values, errorValues);
title('Log Conjugacy Error');
xlabel('h_1 value')
subplot(2,2,3);
plot(log(abs(P(:,successes(end)))));
title('Log Norm of Last Disk Coefficients');
subplot(2,2,4);
sobolevNorms = zeros(length(successes),maxSobolev);
for k = 1:length(successes)
    for s = 1:maxSobolev
        sobolevNorms(k, s) = sobolevNorm(P(:,successes(k)),s);
    end % for loop
end % for loop
mesh(1:size(sobolevNorms,2),1:size(sobolevNorms,1),sobolevNorms);
title('Sobolev Norm of Successes H^s');
xlabel('Sobolev Parameter s')
ylabel('Success number')
fprintf('Number of successes = %d\n', length(successes))
fprintf('Final h_1 value = %.15f\n', h_1)
% Comment to not overwrite the figures in the paper.
saveas(figSuccesses, 'siegel_disk.png')
saveas(figError, 'siegel_disk_defect.png')
diary off
```

Most of what this script does is set up parameters and storage for outputs. Then, given a $\hat{h}_1$ value, it tries to repeatedly increase $\hat{h}_1$ until either the tail is large or the conjugacy error is large.

5.3 Exploring the preimages of the siegel disk

Listing 3: post_expanding_script.m

```
%% Plot some pre-images of the largest siegel disk
orbit = Pz(:,iter-1);
figure(1)
hold on
for k = 1:50
    [newOrbit1, newOrbit2] = map_f_preimage(orbit, a);
    orbit = [newOrbit1; newOrbit2];
    plot((real(orbit)-xMin)/step,(imag(orbit)-yMin)/step,'.k','markersize',1);
    if k >= 1
       orbit = orbit(1:2:end);
    end
```

```
end
```

This script plots preimages of the largest Siegel disk. This allows for us to visualize the Julia Set via inverse iteration.

6 Numerical Results

6.1 Phase Space ℂ Output

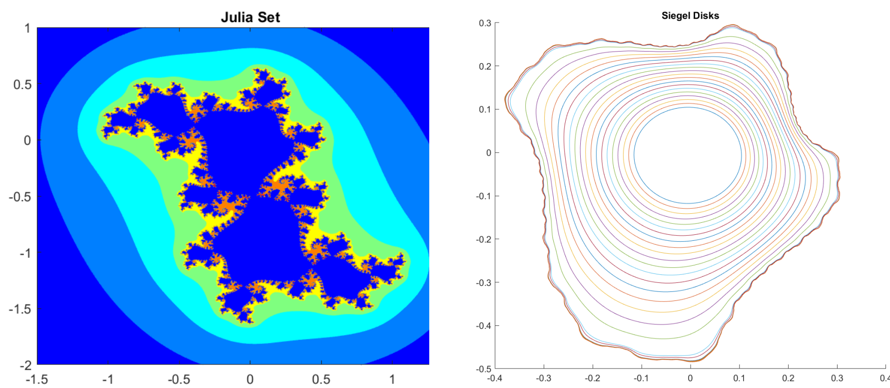
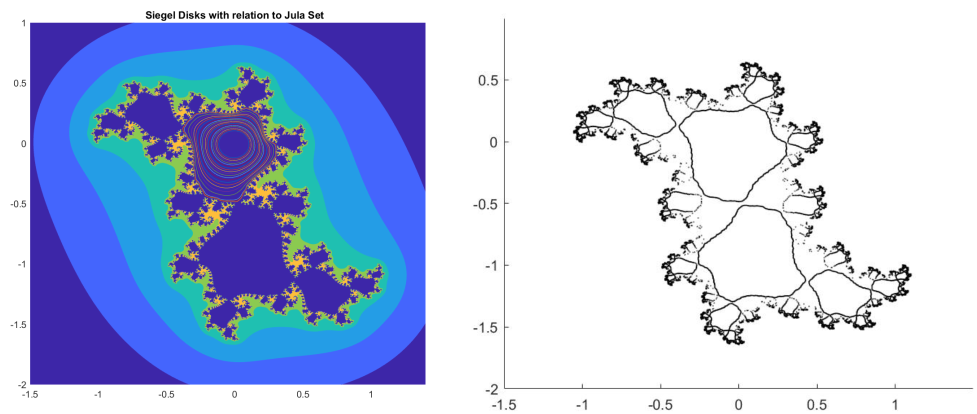


Figure 1: Visualization of output



6.2 Function Space Information

Number of successful attempts is 58, out of 72 iterations.
Final $\hat{h}_1$ value is 0.302326823002451.
Error size on final step 1e-17.
Magnitude of last coefficent on last step 1e-10.

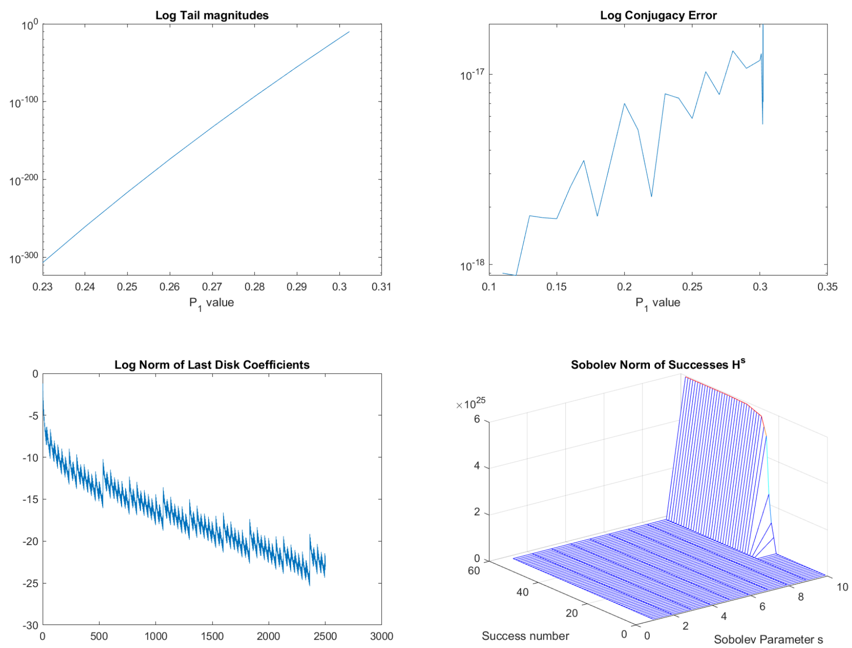


Figure 3: Log plot of tail magnitudes and conjugacy error.

7 Siegel Center Theorem

This section will be used to prove the Siegel Center theorem, in a couple different formats as a starting point to the preliminary presentation.

I'd like to know a few things about what is going on here...

- What does the theorem actually say?

  An analytic function, under certain assumptions, is analytically conjugate to its linear part on a small-enough neighborhood of the origin. More or less.
- What sort of problems would be solved by this?
- What sort of numbers are actually Diophantine?
- How does one practically implement the iterative process to compute h?
- How does one compute the Siegel radius?
- What is the iterative scheme exactly?

Theorem 1 (1 Dimensional Siegel Center Theorem). Let $U \subset \mathbb{C}$ be a neighborhood of $0 \in \mathbb{C}$, and $f : U \to \mathbb{C}$ be an analytic function of the form,

$$f(z) = az + \hat{f}(z) \tag{13}$$

where $\hat{f}(z)$ is $O(z^2)$.

Assume that

$$\left|(a^n - 1)^{-1}\right| \leq n^\nu K \tag{14}$$

for some $\nu \geq 1$ and $K > 0$. Assume as well that

$$\|\hat{f}\|_1 \leq \rho(\nu) K^2 \tag{15}$$

where ρ is an explicit function of ν.

There there exists an unique function

$$h(z) = z + \hat{h}(z) \tag{16}$$

with $\hat{h}$ analytic in a disk of radius $\sigma = 1 - 2\rho(\nu)$ such that

$$f \circ h(z) = h(az) \tag{17}$$

Moreover,

$$\|\hat{h}\|_\sigma \leq C \|f\|_1 \tag{18}$$

Questions

- What are $\|\cdot\|_1$ and $\|\cdot\|_\sigma$? I believe they are sup norms... just have to loop up what is used for analytic functions.
- What does it mean for a function $\rho(\nu)$ to be an explicit function of ν?

We should break down the proof as follows,

- Basic Outline
- Difficulties: small divisors, shrinking of domain, smoothness assumptions
- Technical Lemma
- Set up function spaces & Newton-like method
- Convergence bounds
- Induction bounds
- Put it all together

Proof. Sketch:

We can think of

$$f \circ h(z) = h(az)$$
$$0 = f \circ h(z) - h(az)$$

as an implicit equation in a space of functions, ie as a zero of

$$\mathcal{T}(f, h) = f \circ h - h \circ a$$

Note, $\mathcal{T}(a, Id) = 0$.

Consider a f fixed close to a, and an approximate solution h and say,

$$\mathcal{T}(f, h) = R \tag{19}$$

for some remainder R which we will think of a small.

We obtain a Δ that eliminates most of R so $\mathcal{T}(f, h+\Delta) \ll R$. This is a Newton-like method, we start by approximating the correction,

$$\mathcal{T}(f, h+\Delta) \approx \mathcal{T}(f, h) + D_2\mathcal{T}(f, h)\,\Delta$$

We want

$$\mathcal{T}(f, h+\Delta) = 0$$

Which can be approximated by using the approximation above,

$$R + D_2\mathcal{T}(f, h)\,\Delta = 0 \tag{20}$$

We can write the derivative,

$$\begin{aligned} D_2\mathcal{T}(f, h)\,\Delta &= \left.\frac{d}{dt}\right|_{t=0} F(f, h+t\Delta) \\ &= \left.\frac{d}{dt}\right|_{t=0} f\circ(h+t\Delta) - (h+t\Delta)\circ a \\ &= \left.\big(\big(f'\circ(h+t\Delta)\big)\,\Delta - \Delta\circ a\big)\right|_{t=0} \\ &= \big(f'\circ h\big)\,\Delta - \Delta\circ a \end{aligned}$$

So we can substitute this into 20,

$$\big(f'\circ h\big)\,\Delta - \Delta\circ a = -R$$

If we could reduce $f'\circ h = a + \hat{f}'\circ h$ was just a, this would reduce to that considered in Lemma 3 (see below).

Take derivatives wrt z of 19,

$$\big(f'\circ h\big)\,h' - a\,\big(h'\circ a\big) = R'$$

Moreover, instead of looking for Δ directly, we use the fact that h is analytic and near the identity, so we look equivalently for w defined by $\Delta = h'w$. The previous equation becomes,

$$\big(f'\circ h\big)\,h'w - \big(h'\circ a\big)\,(w\circ a) = -R$$

Substituting $(f'\circ h)\,h' = a\,(h'\circ a) + R'$ in,

$$a\,\big(h'\circ a\big)\,w - \big(h'\circ a\big)\,(w\circ a) = -R - R'w$$

We note that h' is of order 1, since Δ is small, w must be small. Also R is small, so $R'w$ must be small. So we ignore $R'w$. Factor and divide

$$aw - w\circ a = -\big(h'\circ a\big)^{-1} R \tag{21}$$

This is of the form from Lemma 3. Therefore we can solve it for w, and Δ in turn.

The proof finishes by the iteration

1. Take w solving 21.
2. Form $\Delta = h'w$

3. Then $h + \Delta$ should be a better solution to the problem.

It must be shown that the procedure improves the estimate, as well as the procedure can be repeated infinitely many times and converges to an analytic function. $\square$

Here is the technical lemma referenced in the sketch of proof. It was proved in Chapter 2 of KAM Tutorial in a more general format. I would like to prove it as stated here...

Lemma 3. *Assume that a is Diophantine (K, ν). Then if $\nu(0) = \nu_0 = 0$, we can find a solution of*

$$\varphi(az) - a\varphi(z) = \nu(z); \qquad \varphi(0) = 0$$

Moreover,

$$\|\varphi\|_{re^{-\delta}} \leq CK|\delta|^{-\tau}\|\nu\|_r$$

where τ is related to ν from the Diophantine condition.

Proof. $\square$

Here is the lemma showing the convergence of the iteration.

Lemma 4. *Let f be as in the 1-D Siegel Center Theorem, $h(z) = z + \hat{h}(z)$, $\hat{h}(z) = O(|z|^2)$, defined in a ball of radius $\frac{1}{2} < \sigma < 1$ satisfying*

$$\|\hat{h}'\|_\sigma \leq M \leq \frac{1}{2}$$

with

$$\sigma + M < 1$$
$$\|f \circ h - h \circ a\|_\sigma \leq \varepsilon$$

Assume furthermore that $\delta > 0$ is such that

$$KC\delta^{-\nu-1}\varepsilon + \sigma e^{-\delta} < \sigma$$

Then, the prescription above can be carried out and we have,

$$\|f \circ (h + \Delta) - (h + \Delta) \circ a\|_{\sigma e^{-\delta}} \leq KC\delta^{-\nu-1}\varepsilon^2 + \frac{1}{2}\|f\|_1(KC\delta^{-\nu-1}\varepsilon M)^2\delta^{-2}$$

8 Conclusions

8.1 General Observations

Remark. I have noticed that the large steps will fail, however this is not always due to there not being a disk there. Sometimes it s a matter of not having a high enough order.

Remark. It is also worth nothing that the iterative approach is unnecessary in the sense that each guess does not depend in any meaningful way on the previous guesses. Therefore, parallel computations could be employed to speed up the process.

8.2 Questions

- ~~Is there a norm that we would expect to blow up as we approach the boundary?~~
  Yes, the sobolev $H^s = W^{s,2}$ norms.
- ~~Rafael de la Lave fixes $\hat{h}_1 = 1$ that doesn't seem to lead to a convergent series, any idea why?~~
  It is better for pen and paper work.
- ~~Follow up, would the radius of convergence be extremely small?~~
  Yes, hence having to vary r to be smaller than 1. Note, outside the Siegel Disk this this series probably diverges.
- ~~Better way to visualize coeffs?~~
  Done.

8.3 Next Tasks

- ~~Try with different a, look for other diophantine numbers.~~ It works well
- ~~Make a Julia set plotter, pseudo-code available on wikipedia~~.
- ~~Combine Julia set plot with output~~ Clean it up some
- ~~Implement a pre-image plotter~~ not sure this does anything useful for us.
- ~~Look over Sobolev spaces and Parsevalś theorem on wikipedia and justify the use.~~
- Add sections about INTLAB as outlined in [4]. First, read this paper and learn some INTLAB. [2]
- Try with different f, for transcendental f this will require autodifferentiation.
- ~~Work on single dimensional version of Lemma 2.1~~ in paper, work on a tail only version $|\alpha| \geq n$ [3]
- Write section explaining background for above... did some... do more

References

[1] Rafael De la Llave. *A tutorial on KAM theory*. Jan. 2001. DOI: 10.1090/pspum/069/1858536.

[2] Gareth Hargreaves. "Interval Analysis in MATLAB". In: *Numerical Analysis Report* No. 416 (Feb. 2002).

[3] Jason D. Mireles James Rafael de la Llave. "Parameterization of invariant manifolds by reducibility for volume preserving and symplectic maps". In: *Discrete & Continuous Dynamical Systems - A* 32.1078-0947_2012_12_4321 (2012), p. 4321. ISSN: 1078-0947. DOI: 10.3934/dcds.2012.32.4321. URL: http://aimsciences.org//article/id/df5d5499-439d-4563-870f-329c75d16dfc.

[4] S.M. Rump. "INTLAB - INTerval LABoratory". In: *Developments in Reliable Computing*. Ed. by Tibor Csendes. http://www.ti3.tuhh.de/rump/. Dordrecht: Kluwer Academic Publishers, 1999, pp. 77–104.