UNIVERSITAT DE BARCELONA

Treball final de grau

GRAU DE MATEMÀTIQUES

Facultat de Matemàtiques i Informàtica
Universitat de Barcelona

SIEGEL'S LINEARIZATION THEOREM

Carles Raich Bros

Director:	**Dr. Alex Haro**
Realitzat a:	**Departament de Matemàtiques i Informàtica**
Barcelona,	**July 7, 2017**

Abstract

The purpose of this study is to give an insight to Siegel's linearization theorem, a result in discrete dynamics of one-dimensional holomorphic maps that claims the existence of a change of coordinates in a neighbourhood of a map's fixed point to its linear part, whenever the multiplier for such point satisfies the Diophantine condition. This overall approach aims to provide an understanding of the theorem and all it encompasses. It firstly puts forward necessary knowledge in Diophantine approximations as well as complex and functional analysis and introduces some background to Schröder's equation, the conjugacy problem in which the theorem originates. Once set, the theorem is proved in great detail and the dissertation concludes with a numerical exploration performed to visualize and ponder about the most relevant results aforementioned.

How one encounters reality is a choice.
Martin Heidegger

Contents

1 Introduction

In 1942, Carl Ludwig Siegel (1896 - 1981) gave an important result within the field of discrete dynamics of holomorphic functions, known as his linearization theorem. Such result shed some light on the discussion started years ago by Ernst Schröder (1841 - 1902) that seemed to have reached an end point when trying to describe the behaviour of an orbit near a fixed elliptic point.

While it did not give closure to the whole discussion, it outlined the path with which to properly tackle the issue, showing that the right approach on the matter was a refined argument containing a bit of number theory. This allowed, later on, other mathematicians to explore in greater depth the nature of elliptic fixed points, which turns out to be much more difficult to understand, yet mesmerizing.

In his proof, he also encountered and used traits quite common in one of the modern tools of the analysis of many dynamical systems: the KAM theory. A fully-fledged theory named after Andrey Nikolaevich Kolmogorov (1903 - 1987), Jürgen Kurt Moser (1928 - 1999) and Vladimir Igorevich Arnold (1937 - 2010) that roots, among others, in the studies and results of Siegel. Hence, his theorem entails much more than just a linearization result, thus being in fact the reason why it becomes the main theme of this graduate thesis.

In order to fully comprehend the depth of it, we first review in section 2 some crucial elements that will eventually play an important role: the Diophantine approximations and some elements of complex and functional analysis. Although the proof of theorem mainly relies on a KAM version of the Newton Method that requires very precise estimates, many fundamental results in complex and functional analysis lie underneath. And when dealing with the so-called *small divisors equation*, the notion of Diophantine number will prove to be essential so as to overcome one of the main mathematical challenges in the proof.

In section 3, we proceed to give some basic background in discrete dynamics so as to introduce the notion of conjugacy and the origin of Schröder's equation by which a general studying method for a holomorphic map's dynamic in a neighbourhood of its fixed points is intended. Once set and sorted, in section 4 we tackle right away the proof of the theorem for which we use all the previously acquired knowledge.

Finally, in section 5 we move on to some numerical explorations in order to visualize all the aforementioned by spotting the linearizations that Siegel's theorem guarantees for different multipliers. In fact, a further study is put forward so as to see how the approximated conjugacy computed and the perturbation to a linear map may affect its linearization, thus concluding the whole study revolving around the Siegel's linearization theorem and all it encompasses.

2 Preliminaries

Prior to embarking upon the history where Siegel's linearization theorem roots in, it is precise to set forth some elements from different mathematical fields: the Diophantine approximations and some results in complex and functional analysis. Despite the technicality of these preliminaries, the tools provided from the three areas will help framing the upcoming sections. Hence, this summary covers most of the relevant results, but just gives proof to some of them and references for the others.

2.1 Diophantine Analysis

Throughout this section we shall consider $\omega \in \mathbb{R}$-$\mathbb{Q}$, as irrationals that are not *too well approximated by rationals* play a central role when finding whether or not there exists an interesting change of coordinates near a fixed point of a one-dimensional holomorphic map. The main idea is to compare the distance between ω to any rational. As long as it can be somehow lower-bounded, good approximations by rationals are not to be expected and, as we shall see, this discussion translates straight into the field of Diophantine Analysis.

To this purpose, let us start off by giving the definition of a Diophantine number.

Definition 2.1.1. An irrational number ω is called Diophantine if and only if there exists $\varepsilon > 0$ and $\nu > 1$ such that

$$|q\omega - p| > \frac{\varepsilon}{q^{\nu}}$$

$\forall p, q \in \mathbb{Z}$, $q \neq 0$. We then say that ω is a Diophantine number of type (ε,ν).

The inequality given by the Diophantine condition is often rewritten as $|\omega - \frac{p}{q}| > \varepsilon q^{-\nu-1}$ where $k := \nu + 1 > 2$ is defined as the order of ω. According to this notation, we define

$$D_k = \{\omega \in \mathbb{R}\text{-}\mathbb{Q} \text{ ; } \omega \text{ is a Diophantine number of order } k\}$$

as the set of Diophantine numbers of order k and

$$D = \cup_{k>2} D_k$$

as the set of Diophantine numbers.

The natural question to be posed now is to what an extent are these numbers *common* or, in other words, how *many* there are. It is apparent that a positive answer to this question would prove the aforementioned definition to be key in this matter.

A first result by Joseph Liouville (1809 - 1882) in number theory proves that the set of algebraic irrationals happen to be Diophantine numbers. So it turns out that despite the apparently restrictive definition, D is not that rare after all.

Theorem 2.1.2. *(Liouville, 1844) Let ω be a real algebraic number of degree $n > 1$. Then there exists a positive constant C=C(ω) such that*

$$|\omega - \frac{p}{q}| > \frac{C}{q^n}$$

$\forall p, q \in \mathbb{Z}$, $q \neq 0$.

Proof. Given any $p, q \in \mathbb{Z}$, $q \neq 0$, if we denote by $P(x)$ the minimum polynomial of ω, the mean value theorem yields

$$P(\omega) - P(\frac{p}{q}) = P^{'}(\xi)(\omega - \frac{p}{q})$$

for some ξ lying in between ω and $\frac{p}{q}$.

Since P is irreducible of degree $n > 1$ and has integer coefficients, then $|P(\frac{p}{q})| \geq \frac{1}{q^n}$. Let us now assume that $|\omega - \frac{p}{q}| \leq 1 \implies |\xi| \leq 1 + |\omega|$. Then $|P^{'}(\xi)| < \frac{1}{C}$ for some $C > 0$ since $P^{'}(x)$ is continuous on the compact $\{x \in \mathbb{R}; |x| \leq 1 + |\omega|\}$. Hence

$$1 \geq |\omega - \frac{p}{q}| > C|P(\frac{p}{q})| \geq \frac{C}{q^n}$$

Notice that if $|\omega - \frac{p}{q}| > 1$, the result is apparent.

□

It is interesting to highlight that given an irrational ω outside the class D, as a consequence of Liouville's theorem, ω ought to be a transcendent number. These type of irrational numbers are often called *Liouville numbers* and are broadly studied. In fact, due to many of the applications of Liouville's theorem, the bound to the distance it provides was significantly enhanced by others afterwards. It is worth mentioning a theorem by Klaus Roth (1925 - 2015), previously conjectured by Siegel, since it is the best refined version of nowadays.

Theorem 2.1.3. *(Roth, 1955) Let ω be an irrational algebraic number and $\varepsilon > 0$, then there are finitely many coprime integers $p, q \neq 0$ such that*

$$|\omega - \frac{p}{q}| < \frac{1}{q^{2+\varepsilon}}$$

Going back to the set of Diophantine numbers, we now know that D is larger than first expected. However the following result tackles the issue directly.

Proposition 2.1.4. *Let $|\cdot|$ denote Lebesgue measure. The set D satisfies:*

1. *D can be expressed as the union of countably many nowhere dense subsets of $\mathbb{R}$, that is, sets whose closure has empty interior. In other words, it is a meager set.*
2. *$|D^c| = 0$.*

So D happens to have full Lebesgue measure, hence *almost every* real number is a Diophantine number. This feature clearly give us some idea as to how *big* the set of Diophantine numbers is, but on the other hand, it also turns out that D is a meagre set. This topological property conveys the idea of how *small* the set is, so as we can see there is this duality confronting two ideas related to the size of D, a rather vague term in mathematics. Let us prove now both properties.

Proof.

1. To prove meagerness we will show that D is in fact a subset of a meagre set. Let us define, given a $n \in \mathbb{N}$, the set $A_n := \{x \in \mathbb{R}\,;\, \forall \frac{p}{q} \in \mathbb{Q}, |x - \frac{p}{q}| \geq \frac{1}{nq^3}\}$.

   Notice that $\forall n \in \mathbb{N}$, the set A_n is closed since it is a countable intersection of preimages of closed sets by continuous functions. It also has an empty interior because it does not contain any rational number and, therefore, $A := \cup_{n\in\mathbb{N}} A_n$ is a meagre set. Since $D \subset A$, this yields the result.

2. As for its measure, we will prove the result on the unit interval $[0,1]$ since the same assertion can be shown considering any interval in $\{[n, n+1]\}_{n\in\mathbb{Z}}$. Let us fix $q \in \mathbb{N}$ and $k = 2 + \alpha > 2$ where $\alpha > 0$, we now define the set

$$A(\alpha, q) := \{x \in [0,1]\ ;\ |x - \frac{p}{q}| \leq \frac{1}{q^{2+\alpha}} \text{ for some p} \in \mathbb{N}\}$$

   which consists of about q interval of size $\frac{2}{q^{2+\alpha}} \implies |A(\alpha, q)| \leq \frac{2}{q^{\alpha+1}}$. Then the measure of all sets $\sum_{q\in\mathbb{N}} |A(\alpha, q)| < \infty$ since it is upper-bounded by the convergent harmonic series $2\sum_{q=1}^{\infty} \frac{1}{q^{\alpha+1}}$ and by the Borel-Cantelli lemma,

$$|A(\alpha)| := |\limsup_{q\to\infty} A(\alpha, q)| = 0 \implies |A| := |\cup_{n\in\mathbb{N}}\, A(\frac{1}{n})| = 0$$

   Now, since $(D^c \cap [0,1]) \subset A$, this yields the result.

□

2.2 Continued fractions

In section 2.1, we tackled the issue of lower-bounding the distance between an irrational ω and any irrational, thus providing a set of irrationals that are *not too well approximated by rationals*. However, there was no emphasis on how to actually approximate numbers by a rational expression. The answer to this are continued fractions: rational expressions that not only will allow us to redefine the set D, but they will also prove to come in handy when weakening the condition of Diophantine number. For this reason we shall briefly introduce them. In section 3.3.2 we will shed some light on the need of weakening at some point the Diophantine condition in relation to the study of a holomorphic map's fixed points.

Let us start off by giving some elemental definitions.

Definition 2.2.1. A finite or infinite expression of the form:

$$a_0 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \cdots}}} \tag{2.1}$$

where $a_n \in \mathbb{R}$ is called a continued fraction. We say it is finite if it terminates and infinite otherwise. The numbers a_n are called partial quotients. We often denote this expression as $[a_0; a_1, a_2, \dots]$.

Since we are interested in approximations by rationals, we will consider expressions such that their partial quotients are integers, also called simple continued fraction. So as to avoid expressions not well defined in this case, we shall allow a_0 to be 0 or negative and we will require the rest to be positive.

Definition 2.2.2. Given a continued fraction $[a_0; a_1, a_2, \dots]$, we define the sequence $(c_n)_{n\in\mathbb{N}}$ as

$$\forall n \in \mathbb{N}, \quad c_n := [a_0; a_1, \dots, a_n]$$

The terms c_n are called convergents.

Notice that if the continued fraction is simple, then each c_n is a rational number and we denote $c_n = \frac{p_n}{q_n}$. Thus, $(c_n)_{n\in\mathbb{N}}$ is a sequence of rational numbers that outlines the path to the sought approximation.

Definition 2.2.3. An infinite continued fraction $[a_0; a_1, a_2, \dots]$ is said to be convergent if and only if the sequence of its convergents $(c_n)_{n\in\mathbb{N}}$ converges. That is, the limit

$$\alpha := \lim_{n\to\infty} c_n$$

exists and $\alpha \in \mathbb{R}$. In this case, we say that $[a_0; a_1, a_2, \dots]$ is a continued fraction expansion of α.

Let us see now that provided a real number α, we are able to give an algorithm with which we can build a simple continued fraction that converges to that number.

Proposition 2.2.4. *(Algorithm) Let α be a real number and $(\alpha_n)_{n\in\mathbb{N}}$ the sequence defined as*

$$\alpha_0 := \alpha$$

and for $n \geq 1$, if $\alpha_{n-1} \in \mathbb{Z}$ we stop the sequence, otherwise

$$\alpha_n := \frac{1}{\alpha_{n-1} - \lfloor \alpha_{n-1} \rfloor}$$

where $\lfloor \cdot \rfloor$ denotes the floor function.

If we define $(a_n)_{n\in\mathbb{N}}$ as the sequence $a_n := [\alpha_n]$, then $[a_0; a_1, a_2, \cdots]$ is a simple continued fraction expansion of α.

Therefore the result claims that any real number can be approximated by a sequence of rationals numbers. Notice that this sequence might be finite if for some $n \in \mathbb{N}$, $\alpha_n \in \mathbb{Z}$. This leads to a more refined result that we now shall state. If a proof to any of the herein propositions and theorems is desired, refer to [9].

Theorem 2.2.5. *(Rational numbers) Any finite simple continued fraction represents a rational number. Conversely, any rational number con be expressed as a finite simple continued fraction. Furthermore, if we require $a_n > 1$ for all $n \geq 1$, then the representation is unique.*

Theorem 2.2.6. *(Irrational numbers) Any infinite simple continued fraction represents an irrational number. Conversely, any irrational number can be expressed as a infinite simple continued fraction. Furthermore, this representations is unique.*

So it turns out that for every irrational α there exists a unique infinite simple continued fraction that its convergents, which are rational numbers, converge to α. This is quite remarkable since the study of the convergents gives an interesting insight into the matter. But let us begin with some basic properties.

Proposition 2.2.7. *Let $[a_0; a_1, a_2, \dots]$ be a infinite simple continued fraction with convergents $c_n = [a_0; a_1, \dots, a_n] = \frac{p_n}{q_n}$ and α the irrational number represented, then*

1. *If we set $p_{-1} = 1$, $q_{-1} = 0$, $p_0 = a_0$ and $q_0 = 1$, the numbers p_n and q_n for $n \geq 1$ satisfy*

$$p_n = a_n p_{n-1} + p_{n-2}$$

$$q_n = a_n q_{n-1} + q_{n-2}$$

   which can also be rewritten as

$$\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix}$$

2. *The denominators are strictly increasing. That is: $q_1 < q_2 < q_3 < \dots$.*

3. *The even-indexed convergents form an increasing sequence whereas the odd-indexed a decreasing one such that*

$$c_0 < c_2 < \cdots < \alpha < \cdots < c_3 < c_1$$

4. *For $n \geq 0$, $|\alpha - \frac{p_n}{q_n}| < \frac{1}{q_n q_{n+1}}$.*

5. *The convergent c_n is the best possible approximation to α among all rationals numbers with the same or smaller denominator, that is, for any rational $\frac{a}{b}$, with $a \in \mathbb{Z}$, $b \in \mathbb{N}$ such that $1 \leq b \leq q_n$, then*

$$|\alpha - \frac{p_n}{q_n}| \leq |\alpha - \frac{a}{b}|$$

So as we can see, this very summary shows that convergents are a very powerful tool with which to tackle the approximation of an irrational number using rationals.

Now, in order to bridge simple continued fractions with Diophantine numbers, we set from now on $\lambda = e^{2\pi i \omega}$ on the unit circle $S^1 \subset \mathbb{C}$, being ω an irrational with simple continued fraction $[a_0; a_1, , a_2, \dots]$ and convergents $c_n = \frac{p_n}{q_n}$ and let us intend to study the orbit of λ under the rotation $z \mapsto \lambda z$, that is $\{\lambda, \lambda^2, \dots\}$, and its closeness to 1.

Definition 2.2.8. Given $m \in \mathbb{N}$, $m > 0$, the point λ^m is said to be a closest return to 1 if and only if

$$|\lambda^m - 1| < |\lambda^n - 1|$$

for every n such that $0 < n < m$.

The following result shows that closest returns are related to the simple continued fraction of ω from which we shall lead the way to D.

Proposition 2.2.9. *The point λ^m is a closest return to 1 $\iff$ m is one of the denominators of any convergent c_n. Moreover if $m = q_n$ with $n \geq 2$, then*

$$\frac{2}{q_{n+1}} < |\lambda^m - 1| < \frac{2\pi}{q_{n+1}} \tag{2.2}$$

It can be shown that if we require ω to be Diophantine of order k is equivalent to the requirement that for every $n \in \mathbb{N}$, $|\lambda^n - 1| > \varepsilon n^{k-1}$ for some $\varepsilon > 0$ depending on λ - in section 4.2 we shall see some hints to this. This combined with 2.2 are the essential ingredients in the next proposition.

Proposition 2.2.10. *Let ω be an irrational with simple continued fraction $[a_0; a_1, , a_2, \dots]$ and convergents $c_n = \frac{p_n}{q_n}$, then*

$$\omega \in D \iff \sup_{n\in\mathbb{N}} \frac{log(q_{n+1})}{log(q_n)} < \infty$$

Now if we get a little bit ahead of the course, when we look for conjugacies near elliptic fixed points, it will be precise to weaken the Diophantine condition and to do so in terms of the convergents' denominators q_n turns out to be much more refined. The following conditions are the ones we will eventually need.

Definition 2.2.11. Let ω be an irrational with simple continued fraction $[a_0; a_1, a_2, \dots]$ and convergents $c_n = \frac{p_n}{q_n}$, then

1. we say ω is a Brjuno number $\iff$ $\sum_{n=0}^{\infty} \frac{\log(q_{n+1})}{q_n} < \infty$ and we denote the set of Brjuno numbers as BR.
2. we say ω is a Pérez-Marco number $\iff$ $\sum_{n=0}^{\infty} \frac{\log(\log(q_{n+1}))}{q_n} < \infty$ and we denote the set of Pérez-Marco numbers as PM.

And of course it can be checked that $D \subset BR \subset PM$.

2.3 Complex Analysis

It is now time to put forward some fundamental results in complex analysis for the upcoming sections. Due to the fact that we will be mainly studying one-dimensional holomorphic maps in a neighbourhood of a fixed point, local properties of holomorphic maps will come in handy in different situations we will encounter.

We will denote by

- $B(r, p)$ the open disc and $\overline{B}(r, p)$ the closed disc of radius $r > 0$ centered in p. If p is the origin, we will simply omit it.
- $\mathcal{C}(\Omega)$ the set of continuous maps and $\mathcal{H}(\Omega)$ the set of holomorphic maps on the open subset $\Omega \subset \mathbb{C}$.

If need be, we might consider sets of continuous or holomorphic functions on compact sets $\Delta \subset \mathbb{C}$, meaning that for each function f in the set there exists an open set $\Omega = \Omega(f) \subset \mathbb{C}$ such that $\Delta \subset \Omega$ and $f \in \mathcal{C}(\Omega)$ or $f \in \mathcal{H}(\Omega)$ respectively. Being that set, let us define the most natural set of functions within the framework of this study. Take $r > 0$, we define

$$\Lambda_r := \mathcal{C}(\overline{B}(r)) \cap \mathcal{H}(B(r))$$

Remark 2.3.1. The idea behind the definition of such a set is that the center of the disc where the functions are defined corresponds to the fixed point of our dynamical system. In section 3.3, we will show that, without loss of generality, we can assume that the fixed point is none other than the origin, therefore the disc where the functions of the set Λ_r are defined has as center the origin, but it could be also considered $\Lambda_{r,p} = \mathcal{C}(\overline{B}(r,p)) \cap \mathcal{H}(B(r,p))$ with $p \in \mathbb{C}$. In fact, the following properties also hold for $\Lambda_{r,p}$ for all $p \in \mathbb{C}$.

Notice that the functions in Λ_r are defined on a compact subset of $\mathbb{C}$. This will allow us to obtain many properties, for instance, recall that if $K \subset \mathbb{C}$ is a compact set and $f \in \mathcal{C}(K)$, we can define the norm $|\cdot|$ as follows

$$|f| := \sup_{z \in K} |f(z)|$$

It is known as the uniform norm. Then, as we shall prove, the pair $(\Lambda_r, |\cdot|)$ is a Banach space. Let us recall first the maximum modulus principle.

Theorem 2.3.2. *(Maximum Modulus Principle) Let $\Omega \subset \mathbb{C}$ be a connected open subset and $f \in \mathcal{H}(\Omega)$, then*

1. *if f is not constant $\implies$ $|f|$ has no local maxima.*
2. *if Ω is bounded and $f \in \mathcal{C}(\partial\Omega)$ $\implies$ the maximum of $|f|$ lies in $\partial\Omega$.*

Hence, when we consider $f \in \Lambda_r$ and the uniform norm $|\cdot|$, we can rewrite the norm as $|f| = \sup_{|z|=r} |f(z)|$. Let us now prove the completeness of Λ_r as a vector space with norm $|\cdot|$.

Proposition 2.3.3. *Let $K \subset \mathbb{C}$ be a compact subset, then the pair $(\mathcal{C}(K), |\cdot|)$ is a Banach Space.*

Proof. It is apparent that $(\mathcal{C}(K), |\cdot|)$ is a normed vector space over the field $\mathbb{C}$. Thus we shall only prove the completeness. Let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathcal{C}(K)$. Then $\forall \varepsilon > 0$, $\exists n_0 \in \mathbb{N}$ such that $\forall n, m \geq n_0$ and $\forall z \in K$, $|f_n(z) - f_m(z)| < \varepsilon$.

Observe that for all $z \in K$, $(f_n(z))_{n\in\mathbb{N}} \subset \mathbb{C}$ is a Cauchy sequence and due to the completeness of $\mathbb{C}$, $\exists f(z) = \lim_{n\to\infty} f_n(z)$. Therefore, we choose some $z \in K$ and $n \in \mathbb{N}$ such that $n \geq n_0$, then $\lim_{m\to\infty} |f_n(z) - f_m(z)| = |f_n(z) - f(z)| < \varepsilon$. Therefore $\forall n \geq n_0$, $|f_n - f| < \varepsilon \implies f_n \rightrightarrows f$ on K. Since $f_n \in \mathcal{C}(K)$ for all n, then $f \in \mathcal{C}(K)$.

□

Observe that this proposition is not enough so as to ensure the completeness of Λ_r. We know that a Cauchy sequence $(f_n)_{n\in\mathbb{N}}$ has limit $f \in \mathcal{C}(\overline{B}(r))$, but we still do not know whether or not $f \in \mathcal{H}(B(r))$. The following theorem by Karl Theodor Wilhelm Weierstrass (1815 - 1897) will prove that the holomorphy of f on the disc $B(r)$ will be automatically inherited from the sequence $(f_n)_{n\in\mathbb{N}}$. Before proving that, we shall state a couple of results in order to prove the theorem.

Lemma 2.3.4. *Let $\emptyset \neq \Omega \subset \mathbb{C}$ be an open subset, then there exists a sequence of compact subsets $(K_n)_{n\in\mathbb{N}} \subset \mathbb{C}$ such that*

1. $\forall n \geq 1$, $K_n \subset \Omega$ *and* $K_n \subset int(K_{n+1})$.

2. $\cup_{n=1}^{\infty} K_n = \cup_{n=1}^{\infty} int(K_n) = \Omega$

This proposition allow us to translate the completeness of $(\mathcal{C}(K), |\cdot|)$ for any compact $K \subset \mathbb{C}$ into the completeness of $\mathcal{C}(\Omega)$ for any $\Omega \subset \mathbb{C}$ open subset with a specific metric d we will now define.

Theorem 2.3.5. *Let $f, g \in \mathcal{C}(\Omega)$ where $\Omega \subset \mathbb{C}$ is an open subset. If we define*

$$d(f,g) := \sum_{n=1}^{\infty} \frac{\rho_{K_n}(f,g)}{2^n}$$

where $(K_n)_{n\in\mathbb{N}}$ is a sequence of compacts as in the previous proposition and $\rho_{K_n} = \frac{|f-g|}{1+|f-g|}$. Then

1. *d is a distance*

2. *let $(f_n)_{n\in\mathbb{N}} \subset \mathcal{C}(\Omega)$ and $f \in \mathcal{C}(\Omega)$,*

   - *$f_n \to f$ with metric d $\iff$ $f_n \rightrightarrows f$ in K_n for all n $\iff$ $\forall K \subset \Omega$ compact subset, $f_n \rightrightarrows f$ in K.*
   - *$(f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $(\mathcal{C}(\Omega), d)$ $\iff$ $(f_n)_{n\in\mathbb{N}}$ is a uniformly Cauchy sequence in K_n for all n $\iff$ $\forall K \subset \Omega$ compact subset, $(f_n)_{n\in\mathbb{N}}$ is uniformly Cauchy in K.*

3. *$(\mathcal{C}(\Omega), d)$ is a complete metric space.*

If both proofs are desired, refer to [3]. We are now able to state and prove the theorem that will prove that Λ_r is a Banach space. In fact, this theorem will prove much more than that.

Theorem 2.3.6. *(Weierstrass Theorem) Let $\Omega \subset \mathbb{C}$ be an open subset, then*

1. *$\mathcal{H}(\Omega)$ is a closed subspace of $\mathcal{C}(\Omega)$.*

2. *for every $k \geq 1$, the map*

$$\begin{aligned} \Delta : \mathcal{H}(\Omega) &\longrightarrow \mathcal{H}(\Omega) \\ f &\mapsto f^{(k)} \end{aligned}$$

   is continuous.

Proof.

1. Let $(f_n)_{n\in\mathbb{N}} \subset \mathcal{H}(\Omega)$ and $f \in \mathcal{C}(\Omega)$ such that $f_n \to f$ in $\mathcal{C}(\Omega)$. We aim to prove that $f \in \mathcal{H}(\Omega)$ and we already know $f \in \mathcal{C}(\Omega)$, hence we will use the theorem of Giacinto Morera (1856 - 1909). Consider $\Delta \subset \Omega$ a closed triangle $\implies$ $\partial\Delta$ is a compact set and due to the fact that $f_n \to f$ in $\mathcal{C}(\Omega)$, we know that $f_n \rightrightarrows f$ in $\partial\Delta \implies \int_{\partial\Delta} f_n dz \to \int_{\partial\Delta} f dz$.

   By the homotopy version of the theorem of Augustin Louis Cauchy (1789 - 1857), since $\partial\Delta \backsim 0$, then $\int_{\partial\Delta} f_n dz = 0 \implies \int_{\partial\Delta} f dz = 0 \implies f \in \mathcal{H}(\Omega)$.

2. Let $(f_n)_{n\in\mathbb{N}} \subset \mathcal{H}(\Omega)$ such that $f_n \to f$ in $\mathcal{H}(\Omega)$ and take $k \geq 1$. Observe that $f_n^{(k)} \to f^{(k)}$ in $\mathcal{H}(\Omega) \iff \forall K \subset \Omega$ compact subset, $f_n^{(k)} \rightrightarrows f$ in $K \iff \forall a \in \Omega$, $\exists \overline{B}(R,a) \subset \Omega$ where $f_n^{(k)} \rightrightarrows f$ in $\overline{B}(R,a)$.

   Therefore, let us consider $a \in \Omega \implies \exists R > 0$ such that $\overline{B}(a,R) \subset \Omega$. We now define $r := \frac{R}{2}$. Notice that $\forall z \in B(r,a)$, $\overline{B}(r,z) \subset \overline{B}(R,a)$. If we apply Cauchy's inequality at z and $\overline{B}(r,z) \subset \Omega$, we obtain

$$|f_n^{(k)}(z) - f^{(k)}(z)| \leq \frac{k!}{r^k} \sup_{w\in\partial B(r,z)} |f_n(w) - f(w)|$$

   As $n \to \infty$, $|f_n(w) - f(w)| \to 0$.

$\square$

Corollary 2.3.7. *For all $r > 0$, $(\Lambda_r, |\cdot|)$ is a Banach space.*

Recall that due to Cauchy's theorem, every $f \in \Lambda_r$ can be expressed as

$$f(z) = \sum_{n=1}^{\infty} f_n z^n$$

where $f_n = \frac{f^{(n)}(0)}{n!} \in \mathbb{C}$ and that $\forall n \in \mathbb{N}$,

$$f^{(n)}(0) = \frac{n!}{2\pi i} \int_{|w|=r} \frac{f(w)}{w^{n+1}} dw$$

In fact, since power series are going to play an important role, let us shortly state some of its most important properties.

Theorem 2.3.8. *Consider $\sum_{n=0}^{\infty} c_n(z-a)^n$ where $c_n, a \in \mathbb{C}$. We define the radius of convergence as*

$$R := \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|c_n|}}$$

Then,

1. *the series converges absolutely if $|z-a| < R$ and diverges if $|z-a| > R$.*
2. *the series converges in $\mathcal{C}(B(R,a))$.*
3. *the function $f(z) = \sum_{n=0}^{\infty} c_n(z-a)^n$ defined on $B(R,a)$ is holomorphic.*
4. *the series $\sum_{n=1}^{\infty} nc_n(z-a)^{n-1}$ has the same radius of convergence.*

When it comes to the norm, we have the following result for power series.

Lemma 2.3.9. *Let $(f_n)_{n\in\mathbb{N}} \subset \mathbb{C}$ be a sequence such that $|f_n| < \varepsilon r^{-n}$ for every n where $r > 0$ is fixed. Then the map $f(z) = \sum_{n=0}^{\infty} f_n z^n$ is holomorphic on $B(r)$ and $|f| \leq \frac{\varepsilon r}{\rho}$ on $B(r-\rho)$ for all $r > \rho > 0$.*

Proof. Let us first compute the radius of convergence. Since

$$\limsup_{n\to\infty} \sqrt[n]{|f_n|} \leq \limsup_{n\to\infty} \varepsilon^{\frac{1}{n}} r^{-1} = r^{-1}$$

then we have that $R \geq r$. As for the norm

$$|f| \leq \varepsilon \sum_{n=0}^{\infty} r^{-n}(r-\rho)^n = \varepsilon \sum_{n=0}^{\infty} (1 - \frac{\rho}{r})^n = \frac{\varepsilon r}{\rho}$$

□

These lasts results lead as well to the following lemma that will also allow us to obtain useful estimates for the maps we deal with.

Lemma 2.3.10. *Let $f(z) = \sum_{n=1}^{\infty} f_n z^n \in \Lambda_r$ such that $|f| < \varepsilon$. Then $\forall n \in \mathbb{N}$, $|f_n| < \varepsilon r^{-n}$.*

Proof. Recall that $f_n = \frac{f^{(n)}(0)}{n!}$, then

$$|f_n| = |\frac{1}{2\pi i} \int_{|w|=r} \frac{f(w)}{w^{n+1}} dw| \leq \frac{1}{2\pi r^{n+1}} \int_{|w|=r} |f(w)| dw < \frac{\varepsilon}{r^n}$$

□

Thus far, the aforementioned results, which happened to be rather easy consequences of elementary complex analysis, have provided us with pretty much all the properties needed with regard to Λ_r. However, there are still two basic results that we ought to state before moving forward. It will be sometimes useful to consider the inverse of our map f, hence let us recall the complex version of the inverse function theorem.

Theorem 2.3.11. *(Inverse Function Theorem) Let $f \in \mathcal{H}(\Omega)$ where $\Omega \subset \mathbb{C}$ is a domain. Suppose there exists $z_0 \in \Omega$ such that $f'(z_0) \neq 0$. Then $\exists\, r > 0$ such that $B(r, z_0) \subset \Omega$ where f is one-to-one on, the image $U = f(B(r, z_0))$ is open and the inverse function*

$$f^{-1} : U \to B(r, z_0)$$

is analytic and satisfies

$$1 = f^{'}(z)(f^{-1})^{'}(f(z))$$

for all $z \in U$.

A complex version of the mean value theorem will also be needed when obtaining different types of estimates. Let us state it.

Theorem 2.3.12. *(Mean Value Theorem) Let $f \in \mathcal{H}(\Omega)$ where $\Omega \subset \mathbb{C}$ is an open convex set, then for every pair $a, b \in \Omega$, there exists $c \in L(a, b) \subset \Omega$ - where $L(a, b)$ denotes the segment between a and b - such that $|f(a) - f(b)| = |f'(c)||a - b|$.*

In particular, when it comes to $f \in \Lambda_r$ we can bound the distance between the image of two points by the radius r and the norm of $f^{'}$. Proofs to both theorems can be found at [2].

2.4 Functional Analysis

In the last section we prove that the space Λ_r alongside the uniform norm $|\cdot|$ is a Banach space that turns out to be an important space in the study of one-dimensional holomorphic map's fixed points. And as it could be expected, this Banach structure is to be related with some concepts in functional analysis.

As an outline, the first steps of Siegel's theorem proof will require to define and even compute some derivatives of an operator, thus revising some concepts with regard to Maurice René Fréchet (1878 - 1973) differential calculus - the generalization of derivative he provided in 1925 - is quite convenient.

Throughout this section, we shall consider the Banach spaces $(E, |\cdot|_E)$ and $(F, |\cdot|_F)$, both linear spaces over the field $\mathbb{C}$, a non-empty open subset $A \subset E$ and the map $f : A \subset E \longrightarrow F$, which we refer to it as an operator. Let us start off by recalling some basic definitions.

Definition 2.4.1. A map $L : E \longrightarrow F$ is a linear transformation $\iff L(\alpha x + \beta y) = \alpha L(x) + \beta L(y)$ for all $x, y \in E$ and all $\alpha, \beta \in \mathbb{C}$.

Even if E or F were just normed and not Banach spaces, there exists a characterization for a linear transformation's continuity that is quite simple.

Proposition 2.4.2. *Let $L : E \longrightarrow F$ be a linear transformation, the following are equivalent:*

1. *L is continuous at the origin.*
2. *L is bounded on $\{x \in E \ ; \ |x|_E < 1\}$.*
3. *$\exists M > 0$ such that $|L(x)|_F \leq |x|_E$ for all $x \in E$.*
4. *L is continuous in every point of E.*

As a consequence of this, if $L : E \longrightarrow F$ is a continuous linear transformation, one can define a norm for the transformation as

$$|L| := \inf\{M > 0 \ ; \ |T(x)|_F \leq M|x|_E, \ \forall x \in E\} = \sup\{|T(x)|_F \ ; \ |x|_E \leq 1, \ \forall x \in E\}$$

In fact, if we let $\mathfrak{L}(E, F)$ be the set of linear transformations from E to F that are continuous, we have the following theorem when E and F are Banach spaces.

Theorem 2.4.3. *If $(E, |\cdot|_E)$ and $(F, |\cdot|_F)$ are Banach spaces, then $(\mathfrak{L}(E, F), |\cdot|)$ is also a Banach space.*

The reason behind introducing linear transformations is that the most natural way, as Fréchet did, of generalizing the concept of derivative in a Banach space is by simply translating the usual definition in differential calculus into Banach spaces, thus the need of a linear transformation within this context. It therefore follows the next definition.

Definition 2.4.4. The operator f is differentiable at a point $a \in A \iff$ there exists a linear transformation $L : E \longrightarrow F$ satisfying that $\forall \varepsilon > 0$, $\exists \delta > 0$ such that for all $x \in A$ with $|x - a|_E < \delta$, $|f(x) - f(a) - L(x - a)|_F \leq \varepsilon |x - a|_E$. The operator f is said to be differentiable on A if and only if it is differentiable at each point of A.

So we have defined the concept of being differentiable in a Banach space, but we still have to figure out precisely what is the derivative of f at a point $a \in A$ as well as what the derivative of f on A. The sensible step now is to check whether or not the well-known results of differential calculus hold in this case.

Proposition 2.4.5. *If f is differentiable at a point $a \in A$, then the linear transformation is unique.*

Proof. Let $\varepsilon > 0$ and suppose there exists two linear transformations L_1 and L_2 such that

$$|f(x) - f(a) - L_1(x-a)|_F \leq \varepsilon |x-a|_E$$

for all $x \in E$ with $|x-a| < \delta_1$ and

$$|f(x) - f(a) - L_2(x-a)|_F \leq \varepsilon |x-a|_E$$

for all $x \in E$ with $|x-a| < \delta_2$.

Let now $\delta = \min\{\delta_1, \delta_2\}$, then for all $x \in A$ with $|x-a|_E < \delta$

$$\begin{aligned}
|(L_1 - L_2)(x-a)|_F &= |L_1(x-a) - L_2(x-a)|_F \\
&\leq |L_1(x-a) - f(x) + f(a)|_F + |f(x) - f(a) - L_2(x-a)|_F \\
&\leq 2\varepsilon |x-a|_E
\end{aligned}$$

Consider $y \in E$, $y \neq 0$ and observe that $\delta > \frac{\delta}{2} = |\frac{\delta y}{2|y|_E}|_E = |(\frac{\delta y}{2|y|_E} + a) - a|_E$, therefore

$$|(L_1 - L_2)(\frac{\delta y}{2|y|_E})|_F \leq 2\varepsilon |\frac{\delta y}{2|y|_E}|_E = \delta\varepsilon$$

which yields $|(L_1 - L_2)(y)|_F \leq 2\varepsilon |y|_E$ for all $y \in E$ since the inequality is also true for $y = 0$. Due to 2.4.2, $L_1 - L_2 \in \mathfrak{L}(E, F)$ and $|L_1 - L_s| \leq 2\varepsilon$.

Since $\varepsilon > 0$ is arbitrary, we conclude $|L_1 - L_2| = 0 \implies L_1 = L_2$.

□

It is now apparent that the uniqueness of the linear transformation is what allow us to define the derivative of f at the point $a \in A$, pretty much as in the usual differential calculus.

Definition 2.4.6. Let f be differentiable at a point $a \in A$, then the unique linear transformation L is called the Fréchet derivative and it is usually denoted as $f^{'}(a)$.

In order to attain the idea of an operator's derivative on an open subset, the following theorem provides a characterization of the derivative at a point in terms of the continuity of the operator at the same point that will lead to the desired definition.

Theorem 2.4.7. *Let f be differentiable at a point $a \in A$, then $f^{'}(a) \in \mathfrak{L}(E, F) \iff f$ is continuous at a.*

We are now set to give the definition of derivative.

Definition 2.4.8. Let f be differentiable and continuous on A. It is said to be continuously differentiable on $A \iff$ the mapping

$$A \longrightarrow \mathfrak{L}(E,F)$$
$$a \mapsto f^{'}(a)$$

is continuous. Such mapping is denoted by $f^{'}$ and is called the derivative of f on A.

It is to be expected that this definition matches well with the usual definition when taking E and F as the usual $\mathbb{R}^n$ or $\mathbb{C}^n$ spaces. Indeed it does, however some proofs are required since the results are not that apparent at first and hence we recommend [4] for a detailed explanation. Let us now set out the elementary rules of differentiation.

Let $\alpha \in \mathbb{C}$ and $h : E \longrightarrow F$. We define for all $x \in E$

$$(f+h)(x) := f(x) + h(x)$$

and

$$(\alpha f)(x) := \alpha f(x)$$

The following proposition claims the result we are expecting.

Proposition 2.4.9. *If f and h are differentiable at a point $a \in A$, then so is $f + h$ and αf. What is more, it is satisfied that*

$$(f+h)^{'}(a) = f^{'}(a) + h^{'}(a)$$

and

$$(\alpha f)^{'}(a) = \alpha f^{'}(a)$$

If we instead take $h : F \longrightarrow G$ where $(G, |\cdot|_G)$ is a Banach space so that the composition $h \circ f$ makes sense, the next result gives us the version of the chain rule of the classical differential calculus.

Proposition 2.4.10. *Let $B \subset F$ be an open subset. If f is continuous and differentiable at $a \in A$ and h is continuous and differentiable at $b \in B$ where $b = f(a)$, then $h \circ f$ is continuous and differentiable at a and*

$$(h \circ f)^{'}(a) = h^{'}(b) f^{'}(a)$$

Thus far, this sums up the most elementary results of differential operators, nevertheless there is still one concept that we ought to bring up in regard to the operator we deal with later on: partial derivatives. The aforementioned definition of derivative can be broadened for operators with more than one variable. Here we shall focus on operators with two variables since it is the case that concerns us. To this purpose, let us assume from now one that $(E_1, |\cdot|_{E_1})$ and $(E_2, |\cdot|_{E_2})$ are Banach spaces.

Proposition 2.4.11. *The pair $(E_1 \times E_2, |\cdot|_{E_1 \times E_2})$ where*

$$E_1 \times E_2 := \{(a,b)\ ;\ a \in E_1,\ b \in E_2\}$$

$$\begin{aligned} |\cdot|_{E_1 \times E_2} : E_1 \times E_2 &\longrightarrow \mathbb{R}_+ \\ (a,b) &\mapsto \max\{|a|_{E_1}, |b|_{E_2}\} \end{aligned}$$

is a Banach space with respect to the operations

$$(a,b)+(c,d):=(a+c,b+d)$$

$$\alpha(a,b):=(\alpha a,\alpha b)$$

for all $\alpha\in\mathbb{C}$ *and all* $(a,b),(c,d)\in E_1\times E_2$.

Remark 2.4.12. There are other norms with which we can give $E_1\times E_2$ the structure of Banach space. For instance

$$\begin{aligned}|\cdot|_{E_1\times E_2}:E_1\times E_2&\longrightarrow\mathbb{R}_+\\(a,b)&\mapsto|a|_{E_1}+|b|_{E_2}\end{aligned}$$

or more generally

$$\begin{aligned}|\cdot|_{E_1\times E_2}:E_1\times E_2&\longrightarrow\mathbb{R}_+\\(a,b)&\mapsto\sqrt[p]{|a|^p_{E_1}+|b|^p_{E_2}}\end{aligned}$$

where $1\leq p<\infty$.

So let us now consider $A\subset E_1\times E_2$ a non-empty open subset and the operator

$$f:A\subset E_1\times E_2\longrightarrow F$$

and see how we can adapt the previous definition of differentiability to each of the variables of f. To this end, for a point $(a_1,a_2)\in A$, we define the following auxiliary mappings

$$\begin{aligned}h_1:A_1&\longrightarrow F\\a&\mapsto h_1(a):=f(a,a_2)\end{aligned}$$

$$\begin{aligned}h_2:A_2&\longrightarrow F\\a&\mapsto h_2(a):=f(a_1,a)\end{aligned}$$

where

$$A_1:=\{a\in E_1\ ;\ (a,a_2)\in A\}$$

$$A_2:=\{a\in E_2\ ;\ (a_1,a)\in A\}$$

for which the definition of being differentiable at the points a_1 and a_2 respectively is applicable. Hence the next definition.

Definition 2.4.13. The operator f is differentiable with respect to the first variable at a point $(a_1,a_2)\in A$ if and only if g_1 is differentiable at a_1 and we write $f_1'(a_1,a_2)=g'(a_1)$, which is called the partial derivative of f with respect to the first variable at the point (a_1,a_2). The operator is said to be differentiable with respect to the first variable on A if and only if it is differentiable with respect to the first variable at each point of A. It is analogous for the second variable.

Observe that $f_1^{'}(a_1, a_2)$ is the unique linear transformation of E_1 into F that satisfies the condition of differentiability for the mapping g_1. What is more, the mapping $a \mapsto (a, a_2)$ is a continuous function, therefore if f is continuous on A, so is g_1 on A_1 and $f_1^{'}(a_1, a_2) \in \mathfrak{L}(E_1, F) \iff g_1$ is continuous at $a_1 \in A_1$. Once more, it is completely analogous for the second variable.

It is important to notice that, since $(E_1 \times E_2, |\cdot|_{E_1 \times E_2})$ has the structure of a normed linear space, the differentiability of f at a point $a \in A$ is studied as the aforementioned definition, however the way it is linked with the partial derivatives is yet unknown. The following theorem shows that partial derivatives and their continuity are necessary conditions for differentiability and continuous differentiability respectively.

Theorem 2.4.14. *Let f be differentiable at a point $a = (a_1, a_2) \in A$, then f is differentiable with respect to both variables at (a_1, a_2) and $\forall (x_1, x_2) \in E_1 \times E_2$*

$$f^{'}(a_1, a_2)(x_1, x_2) = f_1^{'}(a_1, a_2)(x_1) + f_2^{'}(a_1, a_2)(x_2)$$

Moreover, if f is continuously differentiable on A, then the mappings

$$\begin{aligned} f_1^{'} : A &\longrightarrow \mathfrak{L}(E_1, F) \\ (x_1, x_2) &\mapsto f_1^{'}(x_1, x_2) \end{aligned}$$

$$\begin{aligned} f_2^{'} : A &\longrightarrow \mathfrak{L}(E_2, F) \\ (x_1, x_2) &\mapsto f_2^{'}(x_1, x_2) \end{aligned}$$

are continuous on A.

As for the converse, the continuity of f is needed.

Theorem 2.4.15. *Let f be continuous on A and suppose it is differentiable with respect to both variables on A. If $f_1^{'}$ and $f_2^{'}$ are continuous mappings of A into $\mathfrak{L}(E_1, F)$and $\mathfrak{L}(E_2, F)$ respectively, then f is continuously differentiable on A.*

All the proofs of the theorems that have been simply stated can be found at [4].

3 Dynamics of one-dimensional holomorphic maps

The main theme concerning this dissertation is a theorem that embodies many features purely about the dynamics of one-dimensional holomorphic maps. Due to the nature of its context, an introduction to its history and to the most basic notions of discrete dynamics is rather adequate.

We shall consider throughout this section an open subset $\Omega \subset \mathbb{C}$ and $f : \Omega \mapsto \Omega$ a holomorphic function for which we will intend to study its dynamics. So as to tackle that, we begin with the study of the orbits, in particular the local behaviour near the map's fixed points. Afterwards, we will focus on the existence of a change of coordinates in a fixed point's neighbourhood, a problem that still is a subject of nowadays.

Recall that given a point $p \in \Omega$, we define $O(p) := \{p, f(p), f^2(p), \dots\}$ as the orbit of this point. We say that

1. a point $p \in \Omega$ is a fixed point $\iff$ $f(p) = p$, that is, $O(p) = \{p\}$.
2. a point $p \in \Omega$ is a periodic point of period n $\iff$ $f^n(p) = p$, that is, $O(p) = \{p, f(p), \dots, f^{n-1}(p)\}$. The least positive n for which this happens is called the prime period of p.
3. a point $p \in \Omega$ is eventually periodic of period n if p is not periodic but there exists $m > 0$ such that $f^{l+n}(p) = f^l(p)$ for all $l \geq m$.

The general understanding of the orbits of all points in Ω is essential to comprehend to some extent the behaviour of f. It is difficult to study them all, but it turns out that the presence of fixed points in the dynamic are likely to have a strong influence over some of the surrounding points. That is why the study of fixed points is particularly relevant.

What is more, since we are only requiring f to be holomorphic, all the results we obtain within the next propositions and theorems for fixed points can also be applied to periodic - and even eventually periodic - points of period $n > 1$ taking f^n instead of f, for which the periodic points are merely fixed points and its holomorphy is inherited.

In the upcoming subsections we will discuss the notions of stability of fixed points and provide a very easy characterization. We will then study some elementary examples and dig into the field of conjugacies so as to obtain a rather general method of study.

3.1 Stability of fixed points

We consider again the map f and we assume that $\exists\, p \in \Omega$ such that $f(p) = p$.

Definition 3.1.1. We say that a point $z \in \Omega$ is forward asymptotic to p $\iff$ $\lim_{n\to\infty} f^n(z) = p$. We denote by $W^s(p)$ the set consisting of all points in Ω forward asymptotic to p. We call it the stable set of p.

The previous definition is also applied for periodic points of period m by substituting f^n for f^{mn} in the above limit. Also, if f^{-1} exists, we may also consider points backward asymptotic to p by letting $n \to -\infty$. In this case, it is defined $W^u(p)$ as the unstable set of p. It can even be considered this definition for points p that are not periodic by requiring that $|f^n(z) - f^n(p)| \to 0$ as $n \to \infty$. However, when it comes to fixed points, ensuring that the stable or unstable set is not empty is mostly feasible. That is the reason behind this definition and also what motivates the following one.

Definition 3.1.2. We say that p is a sink if and only if there exists $\varepsilon > 0$ such that $B(\varepsilon, p) \subset W^s(p)$. Sometimes the set $B(\varepsilon, p)$ is referred to it as the local stable set, denoted by $W^s_{loc}(p)$. If $W^s(p) = \mathbb{C}$, then p is called a global sink. On the other hand, we say that p is a source if and only if there exists $\varepsilon > 0$ such that for every $z \in B(\varepsilon, p)$, $z \neq p$, there exists $n > 0$ such that $f^n(z) \notin B(\varepsilon, p)$. This neighbourhood is often called the local unstable set and is denoted by $W^u_{loc}(p)$. This time, if $W^u(p) = \mathbb{C} - \{p\}$, then p is called a global source.

It is important to notice that, even though the definition is local, if it is a common thing for a fixed point to be either a sink or a source, then its influence in the dynamics is undeniable. The following proposition proves so. But first, we need another definition.

Definition 3.1.3. The derivative at the the fixed point $f'(p)$ is called the multiplier of p. If its modulus is different to one, we say that p is a hyperbolic fixed point. Otherwise we write $f'(p)$ as $e^{2\pi i\omega}$ and if $\omega \in \mathbb{Q}$, we say it is parabolic and if not, we say it is elliptic.

Proposition 3.1.4. *If p is a hyperbolic fixed point and*

1. $|f'(p)| < 1$, *then p is a sink.*
2. $|f'(p)| > 1$, *then p is a source.*

Proof.

1. Let us rewrite the condition $|f^{'}(p)| < 1$ as $|f^{'}(p)| < \rho < 1$, then by the mean value theorem on some neighbourhood $B(\varepsilon, p)$, $|f(z) - f(p)| < \rho|z - p|$. From the last inequality we have $|f(z) - p| < |z - p|$ since $\rho < 1$, therefore it can be inductively proved on n that $|f^n(z) - p| < \rho^n|z - p|$. Consequently the iterates of f^n converge uniformly to p on $B(\varepsilon, p)$.
2. It suffices to consider f^{-1}, which exists as a consequence of the inverse function theorem on some neighbourhood of p, and apply the first case since $|(f^{-1})^{'}(p)| = \frac{1}{|f^{'}(p)|} < 1$.

□

Remark 3.1.5. Suppose p is a periodic point of period n and let us rewrite $O(p) = \{p, f(p), \ldots, f^{n-1}(p)\} = \{p, p_1, \ldots .p_{n-1}\}$. If we now consider f^n, it is important to observe that its multiplier at $z = p$

$$(f^n)'(p) = f^{'}(f^{n-1}(p))f^{'}(f^{n-2}(p)) \ldots f^{'}(p) = f^{'}(p_{n-1})f^{'}(p_{n-2}) \ldots f^{'}(p)$$

is the same as the multiplier at every other point in the orbit. Therefore, this allow us to translate the idea of *sink* and *source* into periodic points, thus obtaining a result for every type of periodic point of the map f.

3.2 Linear and power maps

Now let us analyze a couple of one-dimensional holomorphic maps that set out the most common behaviours within complex dynamical systems. The general understanding of these maps is going to be key later on to understand other dynamical systems and that is the reason why we ought to examine them carefully. On a side note, throughout this section we shall use the polar form for complex number due to its suitability for this case.

- Linear maps

Consider $\lambda, \beta \in \mathbb{C}$, $\lambda \neq 0$, the maps we are dealing with are of the form

$$z \mapsto \lambda z + \beta$$

However, unless $\lambda = 1$, we can assume without loss of generality that $\beta = 0$ and we shall explain in the next section why. If $\lambda = 1$, the map $z \mapsto z + \beta$ is merely a translation on the real axis direction, so let us suppose from now on that $\lambda \neq 1$ and $\beta = 0$.

It is apparent that if we use the polar forms $z = re^{2\pi i\theta}$ and $\lambda = \lambda_0 e^{2\pi i\alpha}$, the mapping

$$re^{2\pi i\theta} \mapsto (\lambda_0 r)e^{2\pi i(\theta+\alpha)}$$

is a combination of a rotation of angle α and a homothetic transformation with ratio λ_0. And by induction, it is easy to see that

$$O(z) = \{re^{2\pi i\theta}, (\lambda_0 r)e^{2\pi i(\theta+\alpha)}, \dots, (\lambda_0^n r)e^{2\pi i(\theta+n\alpha)}, \dots\}$$

being the origin the only fixed point unless $\lambda = 1$ and $\alpha = 0$ for which the map is $Id(z) = z$.

Consider now $|\lambda| \neq 1$. If $\lambda_0 < 1$, the origin is a global sink because for every $r \geq 0$, $\lambda_0^n r \to 0$ when $n \to \infty$. When $\lambda_0 > 1$, then the origin is a global source since for all $r > 0$, $\lambda_0^n r \to \infty$ when $n \to \infty$. Had we considered the maps on the Riemann sphere - named after Georg Friedrich Bernhard Riemann (1826 - 1866) -, then we could say that ∞ is also a fixed point that plays the opposite role of the origin in each case. As for α, the relevant role it plays when $|\lambda| \neq 1$ is when it equals to 0, for which the origin is a star stable or unstable node. When $\alpha > 0$, there is a rotation within each orbit - a spiral pattern indeed -, but no other significant change.

If $\lambda_0 = 1$, the previous proposition says nothing about the origin as a sink or a source, but it is easy to check that the orbit for $z = re^{2\pi i\theta}$ stays in $|z| = r \implies W^s(0) = \emptyset = W^u(0) \implies$ the origin is neither a sink nor a source. When it comes to α, there is a substantial difference when the angle is a rational or an irrational multiple of 2π. If $\alpha = \frac{p}{q} \in \mathbb{Q}$, then all points are periodic of period q since $re^{2\pi i(\theta+q\alpha)} = re^{2\pi i\theta}$. When α is irrational, we have the following theorem due to Carl Gustav Jacob Jacobi (1804 - 1851).

Theorem 3.2.1. *Consider the rotation $R_\alpha(z) = e^{2\pi i\alpha}z$ on $|z| = 1$ where α is an irrational angle. Then all orbits are dense.*

Proof. Set $z_0 = e^{2\pi i\theta}$. Since $\alpha \notin \mathbb{Q} \implies \forall n \neq m$, $R_\alpha^n(z_0) \neq R_\alpha^m(z_0)$, otherwise there would exist a $k \in \mathbb{Z}$ such that $\theta + n\alpha = \theta + m\alpha + k \implies \alpha = \frac{k}{n-m} \in \mathbb{Q}$!

Since $|z| = 1$ is a compact, the sequence $\{R_\alpha^n(z_0)\}_{n\in\mathbb{N}}$ has a limit point, say $R_\alpha^m(z_0)$, which implies that $\forall\varepsilon > 0$, there exists $n > m$ such that $|R_\alpha^m(z_0) - R_\alpha^n(z_0)| < \varepsilon$. Since R_α is an isometry, then by fixing $\varepsilon > 0$, there exists $k > 0$ such that $|R_\alpha^k(z_0) - z_0| < \varepsilon$.

Take $N \geq \frac{2\pi}{\varepsilon}$. Then the set $A = \{z_0, R_\alpha^k(z_0), \ldots, R_\alpha^{Nk}(z_0)\}$ yields the result since $N\varepsilon \geq 2\pi$, which is the length of $|z| = 1$ and the distance between a pair of consecutive points in A is less than ε.

$\square$

Observe that this theorem can be applied to every circle with center in the origin in the complex plane. This concludes the study of linear maps that, as we can see, it is quite simple and easy to understand.

- Power maps

Consider $d \in \mathbb{N}$, $d > 1$, the maps we now focus on are

$$z \mapsto z^d$$

Pretty much like the linear maps, by using the polar forms $z = re^{i\theta}$ we can rewrite it as

$$re^{i\theta} \mapsto r^d e^{id\theta}$$

Therefore by induction, the orbits are

$$O(z) = \{re^{i\theta}, r^d e^{id\theta}, \ldots, r^{d^n} e^{id^n\theta}, \ldots\}$$

and it is apparent that

1. If $r > 1$, $r^{d^n} e^{id^n\theta} \to \infty$ when $n \to \infty$.
2. If $r = 1$, $\forall n \in \mathbb{N}$ $|r^{d^n} e^{id^n\theta}| = 1$.
3. If $r < 1$, $r^{d^n} e^{id^n\theta} \to 0$ when $n \to \infty$.

Observe that if we were considering the power map on the Riemann sphere, once again the infinity would play the role of a fixed point being a sink in this case whose stable set is $W^s(\infty) = \{z \in \mathbb{C}\ ;\ |z| > 1\}$. When it comes to the origin, it is also a sink whose stable set is $W^s(0) = B(1)$.

However, the dynamics on the unit circle remain a bit unexplained. They are ruled by the angle $d^n\theta$, so the object of study can be translated into the study of the mapping

$$\theta \mapsto d\theta \quad (mod\ 2\pi)$$

which is also known as expanding map. Let us recall a result for such mappings.

Proposition 3.2.2. *Let $m \in \mathbb{N}$, $m \geq 2$ and consider the surjective continuous mapping*

$$\begin{aligned} E_m : \mathbb{R}/2\pi\mathbb{Z} &\longrightarrow \mathbb{R}/2\pi\mathbb{Z} \\ x &\mapsto mx \end{aligned}$$

Then

1. *$Per_p(E_m) := \{x \in \mathbb{R}/2\pi\mathbb{Z}\ ;\ E_m^p(x) = x\} = \{\frac{k}{m^p-1}\ ;\ k \in \{0, \ldots, m^p - 2\}\}$ and $\#Per_p(E_m) = m^p - 1$.*

2. $\overline{Per(E_m)} = \overline{\cup_{p\in\mathbb{N}} Per_p(E_m)} = \mathbb{R}/2\pi\mathbb{Z}$ *and* $|Per(E_m)| = 0$ *where* $|\cdot|$ *is the Lebesgue measure.*

3. *For each* m *there exists a dense orbit.*

4. $\forall x \in \mathbb{R}/2\pi\mathbb{Z}$ *and* $\forall \delta > 0$ *there exist* $n \in \mathbb{N}$ *and* $y \in B(\delta, x)$ *such that the distance between* x *and* $E_m^n(y)$ *with the usual distance on the circle* $\mathbb{R}/2\pi\mathbb{Z}$ *is larger or equal to* $\frac{1}{2}$.

Remark 3.2.3. Robert Luke Devaney (1948 - still alive) introduced in 1989 in his text *Introduction to Chaotic Dynamical Systems* one of the several definitions of chaotic dynamical system. He established three conditions for which a continuous map $f : X \longrightarrow X$, where X is a metric space, would be chaotic:

1. the set $Per(f) = \cup_{n\in\mathbb{N}} Per_p(f)$ is dense in X.

2. f is topologically transitive.

3. f exhibits a sensitive dependence on initial conditions.

As for the second condition, we observe that it suffices in the one-dimensional case to prove the existence of a dense orbit. When it comes to the third, the definition for this sensitivity is exactly the same as the property 4 of the expanding maps, where the amount $d = \frac{1}{2}$ is previously fixed and can be varied. Therefore, in the Devaney sense, the expanding maps are in fact chaotic maps.

So now that we know how both system behave, in the next section we shall provide the tool with which we can apply our knowledge on these maps to a wide variety of other dynamical systems.

3.3 Conjugacy

The conjugacy is the tool that will help us bridge two dynamical systems and allow us to transfer the results we know for one system to the other, thus being a rather wise way with which to address many problems.

Definition 3.3.1. Let $X, Y \subset \mathbb{C}$ be two open subsets and $f : X \longrightarrow X$, $g : Y \longrightarrow Y$ two holomorphic maps. We say that f and g are topologically conjugated if and only if there exists a homeomorphism $h : X \longrightarrow Y$ such that the diagram

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "h"'] & X \arrow[d, "h"'] \\ Y \arrow[r, "g"] & Y \end{tikzcd}$$

commutates. In other words, $\forall x \in X$

$$h(f(x)) = g(h(x))$$

Remark 3.3.2. A conjugacy is at least bijective, therefore the equation

$$h(f(x)) = g(h(x)) \quad \forall x \in X$$

can also be written in terms of the inverse h^{-1}

$$f(h^{-1}(y)) = h^{-1}(g(y)) \quad \forall y \in Y$$

Hence the choice between both expressions is irrelevant.

The fact that h is at least a homeomorphism between X and Y guarantees the preservation of the most important topological properties from both spaces, in particular, the convergence of sequences. This leads to the preservation of the dynamical system's properties that the following theorem sums up.

Theorem 3.3.3. *Let $X, Y \subset \mathbb{C}$ be two open subsets and $f : X \longrightarrow X$, $g : Y \longrightarrow Y$ holomorphic maps topologically conjugated by $h : X \longrightarrow Y$. Then,*

1. *for all $n \in \mathbb{N}$, $h \circ f^n = g^n \circ h$.*
2. *p is a periodic point of $f \iff h(p)$ is a periodic point of g. What is more, their prime period are the same.*
3. *if the stable set of a periodic point p of f is $W^s(p)$, then the stable set of the periodic point $h(p)$ of g is $h(W^s(p))$.*
4. *the periodic points of f are dense in $X \iff$ the periodic points of g are dense in Y.*

Proof. Since h is a homeomorphism, then h^{-1} is also a homeomorphism, therefore it suffices to prove just one of the implications and not both in every property.

1. Trivial by induction.
2. It is apparent from the previous equation $h \circ f^n = g^n \circ h$.
3. Suppose p is a periodic point of f of period k and $z \in W^u(p)$. Then for each $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that if $n \geq n_0$, then $|f^{kn}(z) - p| < \varepsilon$. Due to the continuity of h, given $\overline{\varepsilon} > 0$, there exists a $\delta > 0$ such that if $|w - p| < \delta$, then $|h(w) - h(p)| < \overline{\varepsilon}$. By taking $\varepsilon = \delta$, there exists $m_0 \in \mathbb{N}$ such that for $n \geq m_0$, $|h(f^{kn}(z)) - h(p)| = |g^{kn}(h(z)) - h(p)| < \overline{\varepsilon}$.
4. The fact that h is a homeomorphism that maps fixed points to fixed points yields the result.

□

Remark 3.3.4. Observe that if there exists a dense orbit in X and f is topologically conjugates to g, then there would also exist a dense orbit in g since the homeomorphism maps orbits to orbits and the density is preserved. Therefore, for the one-dimensional case the topological transitivity is also preserved by conjugacy. What is more, in 1992, a proposition appeared in the issue of *American Mathematical monthly* that claimed for one-dimensional continuous maps $f : X \longrightarrow X$ that as long as X is infinite, then the sensitive dependence on initial conditions is a redundant property if the set of periodic points of f are dense and f is topologically transitive whereas if X is finite, then it consists only of the orbit of a single periodic point. Therefore, in this case we also have chaos' preservation by conjugacy. For further details, refer to [5].

So qualitatively speaking, if a conjugation exists between two dynamical systems, then they behave the same way. In fact, recall that in the previous section we supposed without loss of generality that $\beta = 0$ in the linear map if $\lambda \neq 1$. The reason to this is that

$$h(z) = z + \frac{\beta}{\lambda - 1}$$

is a homeomorphism that conjugates the linear map to its linear part in the whole plane, therefore it only sufficed to study linear maps of the form $z \mapsto \lambda z$ with $\lambda \neq 0, 1$. On a side note, notice that the conjugacy h is defined on $\mathbb{C}$, which is not a common case since most of the times the conjugacy is locally defined and thus all the aforementioned results only hold in the domain of the conjugacy. What is more, h also happens to be more than just a homeomorphism, it is a holomorphic bijection and that in fact carries along a stronger bond between maps. Let us recall the following theorem:

Theorem 3.3.5. *Let $\Omega \subset \mathbb{C}$ be an open subset and $h : \Omega \longrightarrow h(\Omega)$ a bijective holomorphic map. Then*

1. *$h^{'}(z)$ has no zeros in Ω.*
2. *$h^{-1} : h(\Omega) \longrightarrow \Omega$ is holomorphic on $h(\Omega)$ and $(h^{-1})^{'}(h(z)) = \frac{1}{h^{'}(z)}$ for all $z \in \Omega$.*

In particular, h is a conformal map, that is, it locally preserves angles.

As we can see, conjugacies that happen to be holomorphic are much more interesting than just the topological ones, that is why we have the following definition.

Definition 3.3.6. Let $h : X \longrightarrow Y$ be a topological conjugacy between maps $f : X \longrightarrow X$, $g : Y \longrightarrow Y$ where $X, Y \subset \mathbb{C}$ are two open subsets. If $h \in \mathcal{H}(X)$, then we say the conjugacy is holomorphic or conformal or that there exists a change of coordinates between both maps.

Being that set, it is now time to set forth the central ideas in order to assemble a powerful tool to study in greater depth the nature of a one-dimensional analytic complex map's fixed points. The course with which to tackle our goal will mainly rely on establish conditions under which we can guarantee the existence of a conjugacy in a fixed point's neighbourhood to, preferably, one of the maps studied in section 3.2. Thus we can locally obtain a great deal of information with regard to the map's dynamic behaviour. That is the reason why conjugacies, holomorphic in particular, to linear maps are our main concern.

Let us start with an important remark to bear in mind.

Remark 3.3.7. Let $\Omega \subseteq \mathbb{C}$ be an open subset and $f : \Omega \to \mathbb{C}$ a holomorphic map with a fixed point $p \in \Omega$. We can assume without loss of generality that Ω is a neighbourhood of the origin and $p = 0$. This assumption is due to the commutativity of the diagram:

$$\begin{array}{ccc}\Omega & \xrightarrow{f} & \mathbb{C} \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle h} \\ h(\Omega) & \xrightarrow{\widetilde{f}} & \mathbb{C}\end{array}$$

where $h(z) = z - p$ and $\widetilde{f}(z) = f(z+p) - p$ are holomorphic maps. That is, we can always consider the conjugacy given by h and then study the translated map $\widetilde{f}$.

As a consequence of that and due to the fact that holomorphic maps are analytic, given a holomorphic map f with a fixed point, we will assume from now on that

$$f(z) = \sum_{n=1}^{\infty} f_n z^n \tag{3.1}$$

where $f_n \in \mathbb{C}\ \forall n \in \mathbb{N}$ in a certain neighbourhood of the origin. Notice that the term f_1 is in fact the multiplier of $z = 0$, which may provide us with some information in terms of stability as seen in section 3.1. Because of its relevance, we shall refer to it as λ and later on we will have to distinguish two cases according to its module.

Back to the conjugacy discussion, it was Ernst Schröder (1841 - 1902) who introduced in 1871 the idea of conformal conjugation when studying the iteration of rational complex functions and finding effective methods for computing iterates. It arose consequently the problem of finding out whether or not there existed a $\rho > 0$ and a change of coordinates h for a given holomorphic map f with a fixed point such that for $|w| < \rho$ the equation of conjugacy

$$f(h(w)) = h(\lambda w) \tag{3.2}$$

was satisfied. This equation and some of its variants are often named after him. For this reason, it is convenient to give the following definition.

Definition 3.3.8. We say that a holomorphic map f with a fixed point is linearizable at the origin if and only if Schröder's equation has a solution h.

Finding out whether or not a map is linearizable at the origin will mainly rely on the nature of its multiplier as we shall see. The first results for Shcröder's equation came at the end of the 19th century for the hyperbolic cases, then followed by some results in the parabolic case at early 20th century and eventually the elliptic case some years later.

3.3.1 Hyperbolic case

When $|\lambda| \neq 1$, the origin is either a sink or a source, so its stability as a fixed point is known. The fact that there are some fixed traits about the origin might not suggest much at first, but in most cases seems to come across as something compatible with the existence of a solution to Schröder's equation. In this subsection, we gather the most basic results about local conjugacies in a neighbourhood of a hyperbolic fixed point that roughly speaking ensure us for almost all cases a solution.

In 1884, it was proved by Gabriel Koenigs (1858 - 1931) that if $z = 0$ is not a critical point of f, then the map is locally conjugated to its linear part. In other words, he showed that f is linearizable.

Theorem 3.3.9. *(Koenigs, 1884) Let $f(z) = \lambda z + \sum_{n=2}^{\infty} f_n z^n$ be a holomorphic map in a neighbourhood Ω of the origin such that $|\lambda| \neq 0, 1$. Then there exists $\rho > 0$ and a change of coordinates $z = h(w)$, such that $h(0) = 0$ and*

$$f(h(w)) = h(g(w))$$

where g is the linear map $w \mapsto \lambda w$ and $|w| < \rho$. Moreover, h is unique up to multiplication by a nonzero constant.

Therefore, the theorem claims that for a wide range of λ values, the linearization is indeed possible. What is more, it provides a solution to the equation, which happens to be unique up to multiplication by a nonzero constant. In the light of this theorem, it may seem that pretty much all work is covered, but we shall see that the remaining cases are going to raise some interesting questions in regard to the dynamics.

In section 4, the proof of Siegel's theorem is also valid to prove the existence of this very conjugacy, but we are still going to give an easier and shorter proof for Koenigs' theorem.

Proof. Existence. Let us assume that $0 < |\lambda| < 1$. We choose a $0 < \delta < 1$ such that $\delta^2 < |\lambda| < \delta$ and write $f(z) = \lambda z + z^2 r(z)$. We can find $\varepsilon > 0$ such that $|\lambda| + M\varepsilon < \delta$ where $M = \max_{z \in \overline{\mathfrak{B}}(\varepsilon)} |r(z)|$ - observe that $M < \frac{\delta - |\lambda|}{\varepsilon}$, therefore by taking $\varepsilon > 0$ small enough, this condition is satisfied. It follows then that on $\overline{\mathfrak{B}}(\varepsilon)$

1. $|f(z) - \lambda z| = |z^2 r(z)| \leq |z|^2 M$
2. $|f(z)| \leq |z|(|\lambda| + |r(z)|) < |z|\delta \leq \varepsilon\delta \implies |f^n(z)| < |z|\delta^n$ for all $n \in \mathbb{N}$.

We now define the sequence $\{h_n\}_{n \in \mathbb{N}}$ as $h_n = \frac{f^n}{\lambda^k} \in \Lambda_\varepsilon$ and we have on $B(\varepsilon)$

$$\begin{aligned} |h_{n+1}(z) - h_n| &= \frac{1}{|\lambda|^{n+1}} |f(f^n(z) - \lambda f^n(z)| \leq \frac{M}{|\lambda|^{n+1}} |f^n(z)|^2 \\ &\leq \frac{M}{|\lambda|} (\frac{\delta^2}{|\lambda|})^n |z|^2 \leq \frac{M\varepsilon^2}{|\lambda|} (\frac{\delta^2}{|\lambda|})^n \end{aligned}$$

Since $\frac{\delta^2}{|\lambda|} < 1$, then the series $\sum_{n=0}^{\infty} (\frac{\delta^2}{|\lambda|})^n$ converges $\implies \sum_{n=1}^{\infty} (h_{n+1} - h_n)$ converges absolutely and uniformly to $h(z) - h_1(z)$ in Λ_r. Observe now that $\forall n \in \mathbb{N}$, $h_n^{'}(0) = 1$, therefore $h^{'}(0) = 1$. If needed, we can shrink ε so as to obtain that $h(z)$ is a change of coordinates. And finally,

$$h(f(z)) = \lim_{n \to \infty} \frac{f^n(f(z))}{\lambda^n} = \lambda \lim_{n \to \infty} \frac{f^{n+1}}{\lambda^{n+1}} = \lambda h(z) = g(h(z))$$

If $|\lambda| > 1$, it suffices to apply the same argument to f^{-1}.

Uniqueness. Let h and t be two change of coordinates such that satisfy equation 3.2. Then for $|w| < \rho$, we have that $t \circ g \circ t^{-1} = f = h \circ g \circ h^{-1}$, which implies the commutativity between maps g and $\Upsilon \coloneqq t^{-1} \circ h$. Notice that we may write $\Upsilon(w) = \sum_{n=1}^{\infty} \upsilon_n w^n$ because it is an holomorphic map and, therefore, we can compare coefficients of the resulting power series from the commutativity. We obtain:

$$\Upsilon(g(w)) = \sum_{n=1}^{\infty} \upsilon_n \lambda^n w^n = \sum_{n=1}^{\infty} \upsilon_n \lambda w^n = g(\Upsilon(w))$$

So for all $n \geq 1$, $\upsilon_n \lambda^n = \upsilon_n \lambda$. Since λ is neither a roof of unity nor zero, for all $n \geq 2$, $\upsilon_n = 0$. This yields

$$\upsilon_1 w = \Upsilon(w) = t^{-1}(h(w))$$

□

Of course the same result cannot be expected when $z = 0$ is a critical point. In fact, not even a conjugacy to a linear map is to be established. In 1904, Lucjan Böttcher (1872 - 1937) stated a result for this case.

Theorem 3.3.10. *(Böttcher, 1904) Let $f(z) = \sum_{n=k}^{\infty} f_n z^n$ be a holomorphic map in a neighbourhood Ω of the origin such that $k \geq 2$ and $f_k \neq 0$. Then there exists $\rho > 0$ and a change of coordinates $z = h(w)$, such that $h(0) = 0$ and*

$$f(h(w)) = h(g(w))$$

where g is the k-th power map $w \mapsto w^k$ and $|w| < \rho$. Moreover, h is unique up to multiplication by an $(k-1)-th$ root of unity.

Hence, the dynamic near a critical point is much richer than the other hyperbolic cases, but still feasible enough to study in depth despite not being linearizable. If a proof is desired, refer to [9].

So when it comes to $|\lambda| \neq 1$, we may conclude that the dynamics near a fixed point of a holomorphic map turn out to be quite elementary. But on the obverse side of the coin, the remaining case will prove to be quite difficult to deal with, thus being the reason why the first results came some years afterwards.

3.3.2 Parabolic and elliptic case

When it comes to $|\lambda| = 1$, it is convenient to think of λ in the unit circle $S^1 \subset \mathbb{C}$, thus referring to it as $e^{2\pi i \omega}$ with $\omega \in \mathbb{R}$. As we shall see, the discussion will mainly focus on the nature of ω rather than λ itself due to the effect it has on the dynamics's behaviour as it was apparent, for instance, in the linear maps studied in section 3.2. This tightly relates to the difference in nomenclature for this case, however, other terms are broadly used and we shall introduce them as well.

Definition 3.3.11. If ω is a rational number, we say that the fixed parabolic point is rationally indifferent. We say it is irrationally indifferent otherwise.

It was in the studies of Léopold Leau (1868 - 1943) and Pierre Joseph Louis Fatou (1878 - 1929) in 1897 - 1919/1920 respectively that some light was shed for the parabolic case $\lambda = 1$. In fact, this led to the conclusion that f is not linearizable if the origin happens to be a rationally indifferent fixed point, thus providing a dense subset in S^1 for which Schröder's equation has no solution - needless to say the identity map is not taken into account. However, the two french mathematicians went up a notch and, in fact, they were able to describe the dynamic for these cases in a rather visual way.

So as to get there, let us consider first the case $\lambda = 1$ and write

$$f(z) = z + f_{k+1} z^{k+1} + \dots$$

where $f_{k+1} \neq 0$ and the integer $k+1 \geq 2$ is called the multiplicity of the fixed point. Since $f'(0) = 1$ and the origin is a fixed point, we can consider a neighbourhood of the origin N such that it is mapped diffeomorphically onto some other neighbourhood of the origin M. It turns out that both french mathematicians noticed that in $N \cap M$ there exist some special subsets for which the dynamic is quite mesmerizing. That is the reason of the following definition.

Definition 3.3.12. A connected open subset U such that $\overline{U} \subset N \cap \ M$ is an attracting petal for f at the origin if and only if

1. $f(\overline{U}) \subset U \cup \{0\}$
2. $\cap_{n \geq 0} f^n(\overline{U}) = \{0\}$

If U satisfies the same conditions but with f^{-1} instead, then U is a repelling petal for f at the origin.

We are now set to state the theorem that describes in detail the behaviour of the map f in relation with the attracting an repelling petals.

Theorem 3.3.13. *(Leau-Fatou Flower) Let $f(z) = z + \sum_{n=k+1}^{\infty} f_n z^n$ be a holomorphic map in a neighbourhood Ω of the origin where $k+1 \geq 2$ is the multiplicity of the origin. Then there exist k disjoint attracting petals $\{U_i\}_{i=1}^k$ and k disjoint repelling petals $\{V_i\}_{i=1}^k$ such that*

1. *If we define $U = \cup_{i=1}^k U_i$ and $V = \cup_{i=1}^k V_i$, then $F := U \cup V \cup \{0\}$ is a neighbourhood of the origin.*
2. *The petals alternate with each other, that is, U_i only intersects with V_{i-1} and V_i, where we identify V_0 as V_k.*

So as we can see, the origin is surrounded by overlapped domains - hence the flower name - that alternate utterly different behaviours for orbits starting out in each of them. As for the attracting petals, the orbit will never escape the petal and eventually will tend to zero and when it comes to the repelling ones, the orbit will leave the petal at some point although it might come back. Therefore, it is apparent that no periodic orbits other than the fixed point exist in the flower F. What is more, Schröder's equation definitely has no solution for such case. The question now is whether or not this applies to other rationally indifferent origins.

To that purpose, let us consider $f(z) = \lambda z + \sum_{n=k+1}^{\infty} f_n z^n$ holomorphic in a neighbourhood Ω of the origin with λ a primitive $n-th$ root of unity. Then, there exists a collection of petals for f^n. As a consequence, it can be shown that f has also a collection of petals that in number ought to be a multiple of the number of petals for f^n, that is, the multiplicity $m+1$ of the origin in f^n satisfies that $m+1 \equiv 1 \ (mod\, n)$. Therefore, for every irrationally indifferent origin, the Schröder's equation has no solution. For further results, refer to [9] and [10].

Nevertheless, we can still wonder if, in general, a conjugacy is to be found for this cases. It was not until 1978, that Leopoldo Camacho (1943 - still alive) proved that for $\lambda = 1$, f is topologically conjugated to the map

$$g(z) = z - z^{k+1}$$

where $k+1$ is the multiplicity of the origin. In fact, this result can be refined up to a holomorphic conjugacy by adding a term of the form βz^{2k+1} to the map g, where β is a holomorphic invariant. Once again, refer to [10] for more information.

So it seems there is only one small left case to address: the elliptic fixed points. The history behind the other cases legitimates the believe this one would have a rather accessible outcome, but it turned out to be thus far the toughest of them all, yet the most well-rounded. In 1912, Edward Kasner (1878 - 1955) conjectured that every holomorphic map near an irrationally indifferent fixed point was linearizable, but in 1917 George Adam Pfeiffer (1889 - 1943) provided him with a counterexample that proved Kasner wrong. Two years later, Gaston Maurice Julia (1893 - 1978) claimed that for rational functions of degree two or more, Schröder's equation had no solution, but then again it would happen to be incorrect.

The first one to state a remarkable result on the matter was Hubert Cremer (1897 - 1983) in 1927. He provided a characterization of the multiplier for which any rational function of degree two or more would not be linearizable. The most relevant aspect of his result was that the set of multipliers in the unit circle that satisfy such condition is quite *large* and although he did not prove that the result held for all multipliers, he indeed settled a good starting point. Before getting to his result, let us first give some notions in regard to the condition he imposed.

Definition 3.3.14. Let $\lambda = e^{2\pi i\omega}$ with ω irrational and let $d \geq 2$, we say that ω satisfies the Cremer condition of degree d if and only if

$$\limsup_{n\to\infty} \frac{\log(\log(\frac{1}{|\lambda^n - 1|}))}{n} > \log(d)$$

Remark 3.3.15. This condition is often presented in terms of the multiplier, but it also is equivalent to other expressions. If we use the simple continued fraction $[a_0; a_1, a_2, \dots]$ and convergents $c_n = \frac{p_n}{q_n}$ for ω, then the Cremer condition of degree d is satisfied if and only if

$$\limsup_{n\to\infty} \frac{\log(\log(q_{n+1}))}{q_n} > \log(d)$$

The Cremer condition is in fact much more common than what might seem. It can be shown that the set of real numbers satisfying the condition for every d is generic, which means it contains a countable intersection of dense open subsets of $\mathbb{R}$ and is necessarily dense and uncountably infinite.

Theorem 3.3.16. *(Cremer, 1927) Let f(z) be a rational function of degree $d \geq 2$ whose origin is fixed and its multiplier satisfies the Cremer condition of degree d, then any neighbourhood of the origin contains infinitely many periodic orbits.*

It is apparent that if Schröder's equation had a solution, it would not be compatible with the infinitely many periodic orbits each neighbourhood contains, hence no local linearization is possible. This property is often referred to is as the small cycle property and it is advisable to bear it in mind since we will eventually come back to it. If wished, the reader may find a proof at [9].

So back then it would seem that every elliptic fixed point was not likely to be linearizable. But the question as whether this was true for all irrationals remained open for quite some years. Eventually Carl Ludwig Siegel (1896 - 1981) set out the main ideas to tackle this situation and it turned out that how *well* the irrational ω was approximated by rationals played an essential role in the matter. Refer to section 2.2 for more details with regard to these approximations.

The Diophantine condition on ω, which is a translation of being badly approximated by rationals, allowed Siegel to carry on with an iterative process that led to the construction of a change of coordinates for the map f to its linear part, thus solving Schröder's equation for the first time in the elliptic case. Not only that but since the set of Diophantine numbers has full Lebesgue measure, his result holds for almost every multiplier on the unit circle. Quite surprising, taking into account that Cremer provided a set of counterexamples for a generic multiplier in the unit circle.

Let us recall first the definition of the Diophantine condition and then proceed to his theorem.

Definition 3.3.17. An irrational number ω is called Diophantine if and ony if there exists $\varepsilon > 0$ and $\nu > 1$ such that

$$|q\omega - p| > \frac{\varepsilon}{q^{\nu}}$$

$\forall p, q \in \mathbb{Z}$, $q \neq 0$. We then say that ω is a Diophantine number of type (ε,ν).

Remark 3.3.18. Just like Cremer's condition, there are other expressions with which one can rewrite the Diophantine condition on ω. Let us give, in particular, the one using the convergent's denominators q_n of the simple continued fraction for ω. We have that ω is Diophantine if and only if

$$\sup_{n\in\mathbb{N}} \frac{\log(q_{n+1})}{\log(q_n)} < \infty$$

Theorem 3.3.19. *(Siegel, 1942) Let $f(z) = \lambda z + \sum_{n=2}^{\infty} f_n z^n$ be a holomorphic map in a neighbourhood Ω of the origin and $\lambda = e^{2\pi i\omega}$ where ω is a Diophantine number. Then there exists $\rho > 0$ and a change of coordinates $z = h(w) = w + \sum_{n=2}^{\infty} h_n w^n$ such that*

$$f(h(w)) = h(g(w))$$

where g is the linear map $w \mapsto \lambda w$ and $|w| < \rho$.

So compared to the other cases, when the origin is an elliptic fixed point clearly two possible situations arise for Schröder's equation and, because of that, it is often used the following definition.

Definition 3.3.20. Suppose the origin is an elliptic fixed point of a holomorphic map f. If Schröder's equation can be solved, we say that the origin is a Siegel point, whereas if f is not linearizable, we say that the origin is a Cremer point.

Siegel's theorem came as a major breakthrough in the study of the elliptic case and in fact for many reasons. The condition he provided holds for every type oh holomorphic function - even rationals of degree $d \geq 2$ - and the set of multipliers for which it is satisfied has full Lebesgue measure, which despite the contrast with the behaviour of a generic set, in applied dynamics often comes across as a better case. He also brought a new perspective to the argument that afterwards proved to be more suitable for addressing the problem, that is, the approximation by rationals; having control over the orbit of $\lambda \mapsto \lambda^2 \mapsto \cdots \mapsto \lambda^n \mapsto \dots$ and how close might get to 1 was key in his theorem and, as we shall see, this line of discussion allowed others to enhance the result.

Moreover, within his resolution of the problem, he encountered and solved many problems that would shortly after become central in the KAM theory - one of the modern approaches of study of the dynamical systems named after Andrey Nikolaevich Kolmogorov (1903 - 1987), Jürgen Kurt Moser (1928 - 1999) and Vladimir Igorevich Arnold (1937 - 2010). He therefore began to outline the path for much more than just the study of conformal conjugacies and because of that, in the next section, we shall focus exclusively on proving such theorem and going through all its details and particularities.

Later on, Alexander Dmitrievich Brjuno (1940 - still alive) and Jean-Christophe Yoccoz (1957 - 2016) proved a much sharper version of Siegel's theorem that took the argument up a notch. In 1965, Brjuno gave a weaker condition than the Diophantine on ω with which he could still ensure a solution for Schröder's equation.

Definition 3.3.21. Let $[a_0; a_1, a_2, \dots]$ be the simple continued fraction of ω and $c_n = \frac{p_n}{q_n}$ its convergents. We say that ω satisfies the Brjuno condition if and only if

$$\sum_{n=0}^{\infty} \frac{\log(q_{n+1})}{q_n} < \infty$$

In particular, if ω is a Diophantine number, it satisfies the Brjuno condition.

Theorem 3.3.22. *(Brjuno, 1965) Let $f(z) = \lambda z + \sum_{n=2}^{\infty} f_n z^n$ be a holomorphic map in a neighbourhood Ω of the origin and $\lambda = e^{2\pi i \omega}$ where ω is a Brjuno number. Then there exists $\rho > 0$ and a change of coordinates $z = h(w) = w + \sum_{n=2}^{\infty} h_n w^n$ such that*

$$f(h(w)) = h(g(w))$$

where g is the linear map $w \mapsto \lambda w$ and $|w| < \rho$.

Notice that, even though Brjuno's case entails more than Siegel's, it seems as if not much more has been said on the subject. The remarkable thing about his theorem is that Brjuno gave in fact the optimal condition on ω for f to be linearizable. This was proved by Yoccoz in 1988 by giving, pretty much in the same line as Cremer, a set of examples that are not linearizable. If a proof to both theorems is desired, refer to [10].

Theorem 3.3.23. *(Yoccoz, 1988) Let $f(z) = e^{2\pi i \omega} z + z^2$ where ω is not a Brjuno number, that is, $\sum_{n=0}^{\infty} \frac{\log(q_{n+1})}{q_n} = \infty$. Then the origin has the property of small cycles and, therefore, it is a Cremer point.*

The property of small cycle appears once again in the context of non-linearizable functions with an elliptic fixed origin. For this reason, the natural question to be posed is whether or not this property is necessary to determine if the origin is a Cremer point. In 1990, the answer to this was provided by Ricardo Pérez-Marco (1967 - still alive) who introduced a weaker condition on ω and characterized the multipliers for which this property appears.

Definition 3.3.24. Let $[a_0; a_1, a_2, \dots]$ be the simple continued fraction of ω and $c_n = \frac{p_n}{q_n}$ its convergents. We say that ω satisfies the Pérez-Marco condition if and only if

$$\sum_{n=0}^{\infty} \frac{\log(\log(q_{n+1}))}{q_n} < \infty$$

In particular, if ω is a Brjuno number, it satisfies the Pérez-Marco condition.

Theorem 3.3.25. *(Pérez-Marco, 1990) Let ω be an irrational number and let $c_n = \frac{p_n}{q_n}$ be the convergents of its simple continued fraction. Suppose that*

$$\sum_{n=0}^{\infty} \frac{\log(q_{n+1})}{q_n} = \infty$$

Then

1. *If ω is not a Pérez-Marco number, there exists a non-linearizable map f whose origin has a neighbourhood where it has no other periodic orbit other than itself. In other words, there exists a map f with a Cremer origin but without the small cycles property.*

2. *If ω is a Pérez-Marco number, every non-linearizable map f has the property of small cycles.*

Observe that the first result proves that a characterization of Cremer's points in terms of the small cycles property is not possible. The second one represents a criterion for finding out whether or not the Schröder's equation can be solved when ω is a Pérez-Marco but not a Brjuno number. More results and a proof to Pérez-Marco's theorem can be found at [9].

Further progress has been achieved in the past years on the study of Schröder's equation in the elliptic case, specially alongside the line of Pérez-Marco's discussion, but not as fundamental as the previous theorems. Recall that even though there are still cases to be covered, the subject remains open for just a subset of zero Lebesgue measure. Such cases mainly include questions about rationals functions for which is unknown, for instance, whether or not they might have a Cremer point without the small cycles property or a Siegel point but with a multiplier not satisfying the Brjuno condition. In spite of that, Pérez-Marco theorem gives a well-rounded closure to this section.

It remains one final observation. Throughout this section the holomorphic map f has been considered on an open subset of the complex plane, but it could have been considered on the Riemann sphere as well where the infinity point could have played the role of a fixed point. For such case, there are also results in conformal conjugacy and if further information is desired, refer to [8].

4 Siegel's linearization theorem

We now focus on the theorem that thus far has proved to be crucial in this study of elliptic fixed points, as shown in section 3.3.2. Bearing in mind the notation used in section 3.3, let us recall first what the theorem claims.

Theorem 4.0.1. *(Siegel, 1942) Let $f(z) = \lambda z + \sum_{n=2}^{\infty} f_n z^n$ be a holomorphic map in a neighbourhood Ω of the origin and $\lambda = e^{2\pi i \omega}$ where ω is a Diophantine number. Then there exists $r > 0$ and a change of coordinates $z = h(w) = w + \sum_{n=2}^{\infty} h_n w^n$ such that*

$$f(h(w)) = h(g(w)) \tag{4.1}$$

where g is the linear map $w \mapsto \lambda w$ and $|w| < r$.

Roughly speaking, the proof of this theorem revolves around solving for h the equation 4.1, that is, Schröder's equation. As earlier mentioned in section 3.3.1, for nonzero $|\lambda| \neq 1$, the equation was already solved by Koenings in a rather uncomplicated manner. However many refinements are required in the Siegel's case due to the nature of the multiplier and, specifically, when solving the *small divisors equation.* The silver lining is that this proof can also be used for Koenigs' case and thus providing us with a wider result. In order to prove step by step the theorem, we shall first introduce some common elements in KAM theory in the upcoming sections.

4.1 The KAM method

To address Schröder's equation in a straightforward way, we will intend to reduce it to an implicit function problem, for we will give a generalization of the Newton Method that shall lead the way to the desired change of coordinates. This technique is also known as the KAM method, since it is broadly applied in KAM theory due to its suitability for many non hyperbolic problems.

The first step is to consider the two functional variable operator $\Gamma(\varphi, \psi) := \psi^{-1} \circ \varphi \circ \psi$ so we can rewrite the conjugacy equation as

$$\Gamma(f, h) = g \tag{4.2}$$

The second step is to linearize the operator near the solution (g, Id) as in the Newton Method. Recall that in our case f is a fixed map *close* to g - we shall discuss this closeness later on - and our goal is to find the unknown map h. Assuming a linear structure on the neighbourhood of (g, Id), we will first seek an approximate solution $h_0 = Id + \sigma_0$, hence we write

$$\begin{aligned}\Gamma(f, h_0) = g &+ \frac{\partial \Gamma(g, Id)}{\partial f}(f - g) + \frac{\partial \Gamma(g, Id)}{\partial h}(h_0 - Id) \\ &+ O\Big((f - g)(h_0 - Id), (f - g)^2, (h_0 - Id)^2\Big)\end{aligned}$$

and we then drop the high order terms so as to replace the linear part in equation 4.2. We obtain:

$$\frac{\partial \Gamma(g, Id)}{\partial h}\sigma_0 = -\frac{\partial \Gamma(g, Id)}{\partial f}\mu_0 \tag{4.3}$$

where $\mu_0 = f - g$.

The third step is to solve equation 4.3 for σ_0, provided that is possible. Then it is all just a matter of iterating this process and proving it actually converges. That is, we take $f_1 := \Gamma(f, Id + \sigma_0)$, which will be closer to g than f, and we look for an approximate solution $h_1 = Id + \sigma_1$, hence we define $\mu_1 := f_1 - g$ and solve the equation

$$\frac{\partial\Gamma(g, Id)}{\partial h}\sigma_1 = -\frac{\partial\Gamma(g, Id)}{\partial f}\mu_1$$

for σ_1. Therefore at step n, if we define

$$\eta_n := h_0 \circ h_1 \circ \cdots \circ h_{n-1}$$

we will have

$$f_{n+1} = h_n^{-1} \circ f_n \circ h_n = (Id + \sigma_n)^{-1} \circ f_n \circ (Id + \sigma_n) = \eta_{n+1}^{-1} \circ f \circ \eta_{n+1}$$

and once the sequences $\{f_n\}_{n\in\mathbb{N}}$ and $\{\eta_n\}_{n\in\mathbb{N}}$ are constructed, the only thing that will be left to prove is their converge to g and h respectively. In section 4.3 we shall discuss this procedure in more detail.

Thus far, that is the architecture of the KAM method applied to the Siegel theorem. However, as neat as it is, there are many intrinsic difficulties within this method with which we shall deal carefully, one of them being the iteration process itself and another one being solving equation 4.3. But let us first start off with the linear approximation to the operator in a neighbourhood of (g, Id). It is convenient to notice that a feature of the operator is the group property, that is

- $\Gamma(\varphi, Id) = Id \circ \varphi \circ Id = \varphi$
- $\Gamma(\varphi, \psi \circ \phi) = (\psi \circ \phi)^{-1} \circ \varphi \circ (\psi \circ \phi) = \phi^{-1} \circ (\psi^{-1} \circ \varphi \circ \psi) \circ \phi = \Gamma(\Gamma(\varphi, \psi), \phi)$

Since $\Gamma(\cdot, Id) = Id$, the partial derivative with respect to f is simply Id. As for the one with respect to h, we shall compute it as

$$\frac{\partial\Gamma(g, Id)}{\partial h}\sigma_0 = \lim_{t\to 0}\frac{1}{t}\big(\Gamma(g, Id + t\sigma_0) - \Gamma(g, Id)\big)$$

If $|t|$ is small enough, then by the inverse function theorem there exists $(Id + t\sigma_0)^{-1}$, hence the previous limit is well defined. On the same note, we can discard the high order terms in t and take $Id - t\sigma_0$ as the inverse. This leads to

$$\begin{aligned}\frac{\partial\Gamma(g, Id)}{\partial h}\sigma_0 &= \lim_{t\to 0}\frac{1}{t}\big((Id - t\sigma_0) \circ g \circ (Id + t\sigma_0) - g\big)\\ &= \lim_{t\to 0}\frac{1}{t}\big(g + t(g \circ \sigma_0) - t(\sigma_0 \circ g) - t^2(\sigma_0 \circ g \circ \sigma_0) - g\big)\\ &= g \circ \sigma_0 - \sigma_0 \circ g\end{aligned}$$

Therefore, the equation 4.3 turns into

$$\sigma_0 \circ g - g \circ \sigma_0 = \mu_0 \tag{4.4}$$

which is often referred to it as *small divisors equation* due to the adversity they represent when solving it. Since this equation will have to be solved in each step of the iteration, it is remarkably important to understand the way it behaves. As we will see in the following section, this is where the Diophantine condition on ω will play an important role.

4.2 Small divisors equation

It is quite common in KAM theory to encounter equations such as the functional one previously introduced, whose solutions and estimates for the solving maps are essential in order to proceed. That is the reason we ought to take some time to carefully study the small divisors equation 4.4, which happens to be one of the first of its kind, solved by Siegel in 1942 and that preceded many more within the advent of KAM theory.

The resolution mainly struggles with small denominators for which the Diophantine condition on ω provides us with some useful bounds that will guarantee the existence of a holomorphic solving function. However, to obtain an estimate for σ_0 in its domain of holomorphy will not be possible and as a consequence the domain will need to be shrunk a little bit to achieve so. In this sense, the loss within the search of estimates will place a burden in the iteration process that, if it is controlled wisely, it can be overcome.

For now, let us focus merely on how to solve equation 4.4 and we shall deal with the rest later on. If we write $\sigma_0(z) = \sum_{n=2}^{\infty} \sigma_n z^n$, the equation can be seen in terms of power series as

$$\sum_{n=2}^{\infty} f_n z^n = \mu_0 = f - g = \sigma_0 \circ g - g \circ \sigma_0 = \sum_{n=2}^{\infty} (\lambda^n - \lambda)\sigma_n z^n \tag{4.5}$$

so the coefficients for σ_0 ought to be

$$\sigma_n = \frac{f_n}{\lambda^n - \lambda} \quad n \geq 2 \tag{4.6}$$

It now becomes apparent the problem raised by the multiplier. If $\lambda = e^{2\pi i\omega}$ with a rational $\omega = \frac{p}{q}$, then all coefficients with $n = q, 2q, 3q, \ldots$ cannot be defined. Whereas if ω is instead irrational, we may assure that $\lambda^n \neq 1$ for all n, but since it could be arbitrarily close to 1, it would not make much of a difference. That is why the condition on ω of not being *too well approximated by rationals* is crucial to prove that the power series expression for σ_0 converges.

Proposition 4.2.1. *There exists $\kappa \in \mathbb{N}$ such that $\forall n \in \mathbb{N}$, $n \neq 0$,*

$$|\lambda^n - 1| \geq \frac{n^{-\nu}}{\kappa} \tag{4.7}$$

where $\nu + 1$ is ω's order as a Diophantine number.

Proof. Observe that for all $n \in \mathbb{N}$,

$$|\lambda^n - 1| = |1 - e^{2\pi i n\omega}| = 2|\sin(\pi n\omega)|$$

Since $\sin(x) \geq \frac{2x}{\pi}$ if $x \leq \frac{\pi}{2}$, let us consider $m \in \mathbb{N}$ the closest integer to $n\omega$ and then we obtain

$$2|\sin(\pi n\omega)| = 2|\sin(|\pi n\omega - \pi m|)| \geq 4|n\omega - m|$$

because $|n\omega - m| \leq \frac{1}{2}$. Since ω is a Diophantine number, there exists $\varepsilon > 0$ and $\nu > 0$ such that

$$|n\omega - m| \geq \frac{\varepsilon}{n^{\nu}} \implies |\lambda^n - 1| \geq \frac{n^{-\nu}}{\kappa}$$

where $\kappa = 4^{-1}\varepsilon^{-1}$.

$\square$

Remark 4.2.2. If we considered Koenigs' case - recall section 3.3.1 -, it would be easy to check out that the multiplier satisfies the following

$$|\lambda^n - 1| \geq \frac{n^{-\nu}}{\kappa|\lambda|}$$

for some $\nu \in \mathbb{N}$ and $\kappa \in \mathbb{N}$. The fact that the term $|\lambda^n - 1|$ is bounded from below yields to the result. By fixing $\nu \in \mathbb{N}$ as small as desired, we can always adjust κ to obtain the inequality.

We are now set to study in greater detail the coefficients for σ_0 map given the bound 4.7. The following argument is going to be applied often throughout the proof of the theorem, hence we shall state it as a proposition.

Proposition 4.2.3. *Let λ be the multiplier in Siegel's theorem - or even in Koenigs' case - and let $\Psi(z) = \sum_{n=0}^{\infty} \psi_n z^n$ be a holomorphic map on $B(\delta)$ and continuous on $\overline{B}(\delta)$. We now define*

$$\Upsilon(z) := \sum_{n=2}^{\infty} \upsilon_n z^n, \quad \upsilon_n := \frac{\psi_n}{\lambda^n - \lambda}$$

If $|\Psi| < \rho$ on $\overline{B}(\delta)$, then

1. *Υ is holomorphic on $B(\delta)$.*
2. *There exists $\xi = \xi(\nu) > 0$ such that*

$$|\Upsilon| < \rho\kappa\xi\chi^{-\lceil\nu\rceil - 1}$$

on $\overline{B}(\delta(1-\chi))$, where $0 < \chi < 1$ and $\lceil\cdot\rceil$ denotes the ceiling function.

Proof. Due to 2.3.10, the hypothesis on Ψ allow us to bound $\forall n \in \mathbb{N}$ the coefficients $|\psi_n| < \rho\delta^{-n}$. Therefore, for $n \geq 2$ we have

$$|\upsilon_n| = \frac{|\psi_n|}{|\lambda||\lambda^{n-1} - 1|} < \rho\delta^{-n}(n-1)^{\nu}\kappa \leq \rho\delta^{-n}n^{\nu}\kappa$$

and since

$$\limsup_{n\to\infty} \sqrt[n]{|\upsilon_n|} \leq \limsup_{n\to\infty} \frac{\sqrt[n]{\rho\kappa}\,\sqrt[n]{n^{\nu}}}{\delta} = \frac{1}{\delta}$$

the radius of convergence for the power series of Υ is

$$R = \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|\upsilon_n|}} \geq \delta$$

and we can conclude that Υ is holomorphic on $B(\delta)$.

We now aim to find a bound for Υ map. Observe that, since $|\Upsilon(z)| < \rho\kappa \sum_{n=2}^{\infty} n^{\nu}\delta^{-n}|z|^n$, if we take $0 < \chi < 1$ and consider the map on $\overline{B}(\delta(1-\chi))$, we can rewrite

$$|\Upsilon(z)| < \rho\kappa \sum_{n=2}^{\infty} n^{\nu}(1-\chi)^n \leq \rho\kappa \sum_{n=1}^{\infty} n^{\lceil\nu\rceil}(1-\chi)^n$$

Let us now introduce a set of functions that were first considered in correspondence of Gottfried Wilhelm Leibniz (1646 - 1716) with Johann Bernoulli (1667 - 1748) in 1969: the polylogarithm functions. One of their main features is that they happen to satisfy quite some functional equations and are closely related to many zeta functions among others, however we are just going to use one of their properties in order to specifically find the value of ξ.

Take $s \in \mathbb{R}$, we define the power series

$$Li_s(z) := \sum_{n=1}^{\infty} \frac{z^n}{n^s}$$

as the polylogarithm function of order s. Notice that its radius of convergence is $R = 1$, hence $Li_s(z)$ is defined for $|z| < 1$. For non-positive integer orders $-k$, it is possible to express the series as

$$\frac{1}{(1-z)^{k+1}} \sum_{n=0}^{k-1} A(k,n) z^{k-n}$$

where $A(k,n)$ denotes the Eulerian number

$$A(k,n) := \sum_{i=0}^{n+1} (-1)^i \begin{pmatrix} k+1 \\ i \end{pmatrix} (n+1-i)^k$$

which is the number of permutations of the numbers $\{1, \ldots, k\}$ where exactly n elements are greater than their previous element. Since these numbers satisfy that

$$\sum_{n=0}^{k-1} A(k,n) = k!$$

for $k \geq 1$, by taking $k = \lceil \nu \rceil$ we obtain

$$|Li_{-\lceil \nu \rceil}| = |\sum_{n=1}^{\infty} z^n n^{-\lceil \nu \rceil}| \leq \frac{\lceil \nu \rceil !}{|1-z|^{\lceil \nu \rceil + 1}}$$

which yields

$$|\Upsilon(z)| < \rho\kappa \frac{\lceil \nu \rceil !}{\chi^{\lceil \nu \rceil + 1}}$$

on $\overline{B}(\delta(1-\chi))$.

$\square$

It follows from this proposition that $\sigma_0(z) = \sum_{n=2}^{\infty} \sigma_n z^n$ is indeed solution to the small divisors equations whose holomorphy and estimates will depend on $\mu_0 = f - g$, the *closeness* between the map f and its linearized part g. An estimate on σ_0' is going to be needed as well in the iteration process, therefore we shall discuss the following for $\mu_0'(z) = \sum_{n=2}^{\infty} n f_n z^{n-1}$.

Since μ_0' vanishes at zero, we can choose $\delta > 0$ such that $|\mu_0'| < \rho$ on $B(\delta)$ where $\rho > 0$ shall be specified later on. By applying on $B(\delta)$ the mean value theorem, we obtain that $|\mu_0| < \rho\delta$ on $B(\delta)$. Hence,

1. The map $\sigma_0(z)$ is holomorphic on $B(\delta)$.

2. Let $0 < \chi < 1$, then on $\overline{B}(\delta(1-\chi))$

$$|\sigma_0| < \rho\delta\kappa\frac{\lceil\nu\rceil!}{\chi^{\lceil\nu\rceil+1}}$$

As for σ_0', the holomorphy is inherited from σ_0 but when it comes to the estimate, it is required to work a little bit more. Following the same idea behind the proof, on $B(\delta)$ we have that $|z\mu_0'(z)| = |\sum_{n=2}^{\infty} nf_n z^n| < \rho\delta$, hence $|nf_n| < \rho\delta^{-n+1}$ and on $\overline{B}(\delta(1-\chi))$

$$\begin{aligned}|\sigma_0'(z)| &= |\sum_{n=1}^{\infty}\frac{(n+1)f_{n+1}}{\lambda^{n+1}-\lambda}z^n| < \kappa\rho\sum_{n=1}^{\infty}(1-\chi)^n(n+1)^{\lceil\nu\rceil}\\ &\leq \rho\kappa\frac{\lceil\nu\rceil!}{\chi^{\lceil\nu\rceil+1}(1-\chi)} - \rho\kappa < \rho\kappa\frac{\lceil\nu\rceil!}{\chi^{\lceil\nu\rceil+1}(1-\chi)}\end{aligned}$$

where it has been used that

$$\frac{\lceil\nu\rceil!}{\chi^{\lceil\nu\rceil+1}} \geq 1-\chi+\sum_{n=1}^{\infty}(1-\chi)^{n+1}(n+1)^{\lceil\nu\rceil}$$

Therefore, we conclude that for $0 < \chi < 1$, on $\overline{B}(\delta(1-\chi))$ we have

$$|\sigma_0| < \rho\delta\kappa\frac{\lceil\nu\rceil!}{\chi^{\lceil\nu\rceil+1}} \tag{4.8}$$

$$|\sigma_0'| < \rho\kappa\frac{\lceil\nu\rceil!}{\chi^{\lceil\nu\rceil+1}(1-\chi)} \tag{4.9}$$

We are now set to begin with the iteration process. It is important to highlight that the existence of a solution to the small divisors equation relied on the bound for $|\lambda^n - 1|$ and the map μ_0 rather than the equation itself. Hence, in the iteration process is going to be remarkably important to find out the domain of holomorphy of the new map f_n as well as giving some estimates to $\mu_n = f_n - g$. All the above is discussed in the following section.

4.3 Iteration process

Now that we know that equation 4.4 is solved by the holomorphic map

$$\sigma_0(z) = \sum_{n=2}^{\infty}\frac{f_n}{\lambda^n-\lambda}z^n$$

in $B(\delta)$, we may move forward to the first step of the iteration process. We set $f_1 = \Gamma(f, h_0) = \Gamma(f, Id+\sigma_0)$, which is closer to the linear map g, and the idea is to repeat the same step as we did in section 4.1. From now on, we will consider that $|\lambda| \leq 1$, thus covering Siegel's theorem as well as one part of Koenigs'. For $|\lambda| > 1$, it will suffice to consider from the very beginning f^{-1} instead, which can be defined since the multiplier is not zero.

The first issue we run into is that we need to know more precisely where f_1 is defined. It is apparent that is holomorphic in a neighbourhood of the origin, which is a fixed point, but we still need to know where the composition $h_0^{-1} \circ f \circ h_0$ makes sense. We will also need estimates for the *closeness* between f_1 and g, more specific of its derivative μ_1'. Recall that for σ_0, the estimates were only obtained in a disk of radius smaller than δ, so it is to be expected that for f_1 the domain of definition is smaller.

Let us tackle the composition $h_0^{-1} \circ f \circ h_0$ for which upper and lower bounding h will be essential so as to determine the domain of f_1. Since $h_0 = Id + \sigma_0$, we choose $0 < \chi < 1$ and consider $|z| < \delta(1 - \chi)$. Then on $B(\delta(1 - \chi))$

$$|h_0(z)| = |z + \sigma_0(z)| \leq |z| + |\sigma_0(z)|$$

$$|h_0(z)| = |z - (h_0(z) - z)| \geq |z| - |\sigma_0(z)|$$

and recall that

$$|\sigma_0| < \rho\delta\kappa \frac{\lceil\nu\rceil!}{\chi^{\lceil\nu\rceil+1}}$$

on $\overline{B}(\delta(1 - \chi))$, where $\rho > 0$ has not yet been specified. In order to obtain rather neat bounds for h_0, it would be convenient to get rid of some of the terms in the expression that estimates σ_0, so for now let us set $\rho > 0$ such that

$$\rho\kappa\lceil\nu\rceil! < \chi^{\lceil\nu\rceil+1}$$

However, we therefore obtain on $B(\delta(1 - \chi))$ that $|\sigma_0| < \delta$, which implies that $|h_0| < \delta(2 - \chi)$ and even though the domain for f might be bigger than a disk of radius δ, we cannot assure that the composition $f(h_0(z))$ takes place in $B(\delta(1 - \chi))$. So let us set instead $\rho > 0$ such that

$$\rho\kappa\lceil\nu\rceil! < \chi^{\lceil\nu\rceil+2} \tag{4.10}$$

for which we obtain $|\sigma_0| < \delta\chi$ and $|h_0(z)| < \delta$ on $B(\delta(1 - \chi))$.

Let us now focus on the last part of the composition since this will probably be determining. We would want indeed the biggest domain possible, so let us allow $|z| = \delta(1-\chi)$ and observe that, if we require $0 < \chi < \frac{1}{2}$, then $|z| - \delta(1-2\chi) = \delta\chi > |\sigma_0|$, hence $|h_0| > \delta(1 - 2\chi)$. Since $h_0(0) = 0 + \sigma_0(0) = 0$, we may conclude that $B(\delta(1 - 2\chi)) \subset h_0(B(\delta(1 - \chi)))$ where the factor 2 cannot be shrunk. Therefore we aim to find a radius $0 < r \leq \delta(1 - \chi)$ such that for $|z| < r$, $|f(h_0(z))| < \delta(1 - 2\chi)$ since $f(h_0(0)) = 0$ and h_0^{-1} is defined on $B(\delta(1 - 2\chi))$.

If we write $r = \delta(1-\alpha\chi)$ where $\alpha \in \mathbb{N}$ and assume $|z| < r$, then $|f(z)| = |\lambda z+\mu_0(z)| \leq |z| + |\mu_0| < r + \rho\delta$ because we chose $\delta > 0$ such that $|\mu_0'| < \rho$ on $B(\delta)$. Since we need $r + \rho\delta = \delta(1 - \alpha\chi + \rho) \leq \delta(1 - 2\chi) \implies \rho \leq \chi(\alpha - 2)$, we can get a specific value for α. From inequality 4.10, we obtain that $\rho < \chi$ because $\kappa, \lceil\nu\rceil! \in \mathbb{N}$ and $0 < \chi < \frac{1}{2}$, hence every $\alpha \geq 3$ suits our need. Notice that the bigger α is, the smaller χ is required to be, therefore we take $\alpha = 3$ being that the smallest possible value and set $0 < \chi < \frac{1}{3}$.

All there is left to do now is look for a radius r that for $|z| < r$, $|h_0(z)| < \delta(1-3\chi)$. Let us set $r = \delta(1-\beta\chi)$ where $\beta \in \mathbb{N}$ and suppose $|z| < r$. We obtain $|h_0(z)| \leq \delta(1-\beta\chi)+\delta\rho$ and we need this estimate to be smaller than $\delta(1 - 3\chi) \implies \rho \leq (\beta - 3)$. Following the above argument, we take $\beta = 4$ and $0 < \chi < \frac{1}{4}$.

So it turns out that for $\rho\kappa\lceil\nu\rceil! < \chi^{\lceil\nu\rceil+2}$ and $0 < \chi < \frac{1}{4}$, f_1 is defined on $B(\delta(1 - 4\chi))$. As we can see, there has been a considerable loss of the domain and this is something that will require to be dealt carefully in the iteration process since we do not want to run out of domain for the limit function.

When it comes to estimating $|\mu_1'|$, let us consider on $B(\delta(1-4\chi))$ with $0<\chi<4$

$$f_1(z) = \lambda z + \mu_1(z)$$

$$f(z) = \lambda z + \mu_0(z)$$

and we may rewrite

$$h_0 \circ f_1 = f \circ h_0$$

as

$$\lambda z + \mu_1(z) + \sigma_0(\lambda z + \mu_1(z)) = \lambda(z + \sigma_0(z)) + \mu_0(z + \sigma_0(z))$$

Since $\lambda \sigma_0(z) = \sigma_0(\lambda z) - \mu_0(z)$, we get to

$$\mu_1(z) = \sigma_0(\lambda z) - \sigma_0(\lambda z + \mu_1(z)) + \mu_0(z + \sigma_0(z)) - \mu_0(z)$$

This last expression is quite useful because by the mean value theorem we may bound the following

$$|\sigma_0(\lambda z) - \sigma_0(\lambda z + \mu_1(z))| \leq \sup_{|z|<\delta(1-4\chi)} |\sigma_0'(z)| \sup_{|z|<\delta(1-4\chi)} |\mu_1(z)| < \frac{\chi}{1-\chi} \sup_{|z|<\delta(1-4\chi)} |\mu_1| < \frac{|\mu_1|}{3}$$

$$|\mu_0(z + \sigma_0(z)) - \mu_0(z)| \leq \sup_{|z|<\delta(1-4\chi)} |\mu_0'(z)| \sup_{|z|<\delta(1-4\chi)} |\sigma_0(z)| < \rho^2 \kappa \delta \frac{\lceil \nu \rceil !}{\chi^{\lceil \nu \rceil + 1}}$$

which leads to

$$\begin{aligned} \sup_{|z|<\delta(1-4\chi)} |\mu_1(z)| &\leq \sup_{|z|<\delta(1-4\chi)} |\sigma_0(\lambda z) - \sigma_0(\lambda z + \mu_1(z))| + \sup_{|z|<\delta(1-4\chi)} |\mu_0(z + \sigma_0(z)) - \mu_0(z)| \\ &< \frac{|\mu_1|}{3} + \rho^2 \kappa \delta \frac{\lceil \nu \rceil !}{\chi^{\lceil \nu \rceil + 1}} \end{aligned}$$

and we obtain

$$|\mu_1| < \rho^2 \kappa \delta \frac{3 \lceil \nu \rceil !}{2 \chi^{\lceil \nu \rceil + 1}}$$

on $B(\delta(1-4\chi))$. So if we shrink a little bit the domain, say by $\delta\chi$ with $0 < \chi < \frac{1}{5}$, then the Cauchy estimates imply that on $B(\delta(1-5\chi))$ we have

$$|\mu_1'| < \rho^2 \kappa \frac{3 \lceil \nu \rceil !}{2 \chi^{\lceil \nu \rceil + 2}}$$

We therefore have been able to find a domain where the composition $h_0^{-1} \circ f \circ h_0$ made sense and also bound $|\mu_1|$ on $B(r_1)$ with $r_1 := \delta(1-5\chi)$ and $0 < \chi < \frac{1}{5}$. As a consequence of that, we know that the small divisors equation for this step

$$\mu_1 = g \circ \sigma_1 - \sigma_1 \circ g$$

can be solved by writing σ_1 in terms of power series, that the series is holomorphic on $B(r_1)$ and that by setting $0 < \chi_1 < 1$, we can bound on $B(r_1(1-\chi_1))$

$$|\sigma_1| < \rho_1 \kappa r_1 \frac{\lceil \nu \rceil !}{\chi_1^{\lceil \nu \rceil + 1}}$$

$$|\sigma_1'| < \rho_1 \kappa \frac{\lceil \nu \rceil !}{\chi_1^{\lceil \nu \rceil + 1}(1-\chi_1)}$$

where $\rho_1 := \rho^2 \kappa \frac{3\lceil \nu \rceil!}{2\chi^{\lceil \nu \rceil + 2}}$. Set $0 < \chi_1 < \frac{1}{5}$ and under the assumption that $\rho_1 \kappa \lceil \nu \rceil! < \chi_1^{\lceil \nu \rceil + 2}$, we can bound on $B(r_1(1 - 5\chi_1))$

$$|\mu_2'| < \rho_2 := \rho_1^2 \kappa \frac{3\lceil \nu \rceil!}{2\chi_1^{\lceil \nu \rceil + 2}}$$

So as long as we set a sequence of $\{\chi_n\}_{n \in \mathbb{N}}$ such that for all n, $0 < \chi_n < \frac{1}{5}$ and, alongside the sequence $\{\rho_n\}_{n \in \mathbb{N}}$, satisfies that $\rho_n \kappa \lceil \nu \rceil! < \chi_n^{\lceil \nu \rceil + 2}$, then the process can be iterated infinitely. However, it is important to choose a sequence wisely because we do not want the radius to converge to zero. The following proposition provides an answer to this choice, but it requires a stricter condition on ρ than 4.10.

Proposition 4.3.1. *Set $0 < \chi = \frac{1}{10} < \frac{1}{5}$ and $\rho > 0$ such that*

$$\frac{3}{2} \rho \kappa \lceil \nu \rceil! < (\frac{\chi}{2})^{\lceil \nu \rceil + 2}$$

If we define the sequences $\{\chi_n\}_{n \in \mathbb{N}}$ and $\{r_n\}_{n \in \mathbb{N}}$ as

$$\chi_0 := \chi, \quad \chi_{n+1} := \frac{\chi_n}{2}$$

$$r_0 := \delta, \quad r_{n+1} := r_n(1 - 5\chi_n)$$

then the process can be iterated infinitely and $r_n \not\to 0$. In particular, at the $n-th$ step, the map f_n is defined on $B(r_n(1 - 4\chi_n))$ and

$$|\mu_{n+1}'| < \rho_{n+1}$$

on $B(r_{n+1})$ where the sequence $\{\rho_n\}_{n \in \mathbb{N}}$ is defined as

$$\rho_0 := \rho$$

$$\rho_{n+1} := \rho_n^2 \kappa \frac{3\lceil \nu \rceil!}{2\chi_n^{\lceil \nu \rceil + 2}}$$

as expected.

Proof. Observe that the key to proceed with the iteration is to require that

$$0 < \chi_n < \frac{1}{5}$$

and

$$\rho_n \kappa \lceil \nu \rceil! < \chi_n^{\lceil \nu \rceil + 2}$$

for every n.

The first condition is apparent because $\chi_n = \frac{\chi_0}{2^n} < \frac{1}{5}$ due to the choice of χ. As for the second one, let us prove it inductively on n.

For $n = 0$,

$$\frac{3}{2} \rho \kappa \lceil \nu \rceil! < (\frac{\chi}{2})^{\lceil \nu \rceil + 2} \implies \rho_0 \kappa \lceil \nu \rceil! < \frac{3}{2} \rho_0 \kappa \lceil \nu \rceil! 2^{\lceil \nu \rceil + 2} < \chi_0^{\lceil \nu \rceil + 2}$$

Let us suppose that for $n = k \geq 1$ the result holds, thus

$$\frac{3}{2}\rho_n\kappa\lceil\nu\rceil! < (\frac{\chi_n}{2})^{\lceil\nu\rceil+2}$$

and for $n = k+1$, due to the induction hypothesis, we obtain

$$\frac{3}{2}\kappa\lceil\nu\rceil!\rho_k^2\frac{3}{2}\kappa\lceil\nu\rceil! < (\frac{\chi}{2})^{\lceil\nu\rceil+2}(\frac{\chi}{2})^{\lceil\nu\rceil+2}$$

$$\begin{aligned}\implies \rho_{n+1}\kappa\lceil\nu\rceil! \leq& \frac{3}{2}\rho_{n+1}\kappa\lceil\nu\rceil!2^{\lceil\nu\rceil+2}\\ =&\frac{3}{2}(\frac{\frac{3}{2}\rho_n^2\kappa\lceil\nu\rceil!}{\chi_k^{\lceil\nu\rceil+2}})\kappa\lceil\nu\rceil!2^{\lceil\nu\rceil+2}\\ <&(\frac{\chi_k}{2})^{\lceil\nu\rceil+2} = \chi_{k+1}^{\lceil\nu\rceil+2}\end{aligned}$$

Therefore at the $n-th$ step, we know that on $B(r_n)$ the map σ_n is holomorphic and solves the small divisors equation. What is more, on $B(r_n(1-\chi_n))$

$$|\sigma_n| < \rho_n\kappa r_n\frac{\lceil\nu\rceil!}{\chi_n^{\lceil\nu\rceil+1}}$$

$$|\sigma_n'| < \rho_n\kappa\frac{\lceil\nu\rceil!}{\chi_n^{\lceil\nu\rceil+1}(1-\chi_n)}$$

and since $\rho_n\kappa\lceil\nu\rceil! < \chi_n^{\lceil\nu\rceil+2}$, we know that f_{n+1} is defined on $B(r_n(1-4\chi_n))$ and

$$|\mu_{n+1}'| < \rho_{n+1} = \rho_n^2\kappa\frac{3\lceil\nu\rceil!}{2\chi_n^{\lceil\nu\rceil+2}}$$

on $B(r_n(1-5\chi_n)) = B(r_{n+1})$.

So the last thing left to check is that the radius $r := \lim_{n\to\infty} r_n = \delta\lim_{n\to\infty}\prod_{i=0}^{n-1}(1-5\chi_i)$ does not tend to 0. It is apparent that the sequence $\tau_n := \prod_{i=0}^{n-1}(1-5\chi_i)$ is strictly decreasing and $\tau_n > 0$, therefore it converges to τ. However, in the interest to prove afterwards the convergence, instead of showing $\tau \neq 0$, we shall look for a rather interesting lower bound.

First of all, let us set $\chi_0 = \frac{1}{10} \implies 5\chi_0 = \frac{1}{2}$. Therefore, for all i

$$0 > -5\chi_i = \frac{-5\chi_0}{2^i} = \frac{-1}{2^{i+1}} \geq \frac{-1}{2}$$

We define the real map $s(x) := \log(1+x)-2x$ for $x \in I = [\frac{-1}{2}, 0]$ and since $s'(x)\frac{1}{1+x}-2 < 0$ for $x < 0$, $s'(0) = 0$ and $s(0) = 0$, we conclude that $\log(1+x) \geq 2x$ in I and hence for all i

$$\log(1-5\chi_i) \geq -\frac{1}{2^i}$$

so we obtain

$$\sum_{i=0}^{\infty}\log(1-5\chi_i) = \lim_{N\to\infty}\sum_{i=0}^{N}\log(1-5\chi_i) = \log(\lim_{N\to\infty}\prod_{i=0}^{N}(1-5\chi_i)) = \log(\prod_{i=0}^{\infty}(1-5\chi_i))$$

and that leads to

$$\log(\prod_{i=0}^{\infty}(1-5\chi_i)) \geq -\sum_{i=0}^{\infty}\frac{1}{2^i} = -2$$

So now we can lower bound the radius r since $\tau \geq e^{-2} > \frac{1}{10} \implies r > \frac{\delta}{10}$.

□

Now that we know that $\overline{B}(\frac{\delta}{10}) \subset B(r)$ and that the sequences $\{f_n\}_{n\in\mathbb{N}}, \{\eta_n\}_{n\in\mathbb{N}} \subset \Lambda_{\frac{\delta}{10}}$, the convergence can be proved easily. The fact that for every n,

$$\frac{3}{2}\rho_n\kappa\lceil\nu\rceil! < (\frac{\chi_n}{2})^{\lceil\nu\rceil+2} \implies \rho_n \to 0$$

and therefore

$$|f_n - g| = |\mu_n| < \rho_n\frac{\delta}{10} \to 0$$

on $\overline{B}(\frac{\delta}{10})$ and

$$f_n \to g$$

$$\eta_n \to h$$

in $\Lambda_{\frac{\delta}{10}}$, where h satisfies $g = h^{-1} \circ f \circ h$ and hence it is the conjugation we were seeking and f is indeed linearizable.

5 Numerical study

Bearing in mind the results assuring the existence of a map's linearization, we may now proceed to numerically study such conjugacy. To this purpose, we choose the following family of one-dimensional holomorphic maps to study

$$f_{\lambda,\varepsilon} = \lambda z + \varepsilon z^2$$

where $\varepsilon \in (0, 1]$ is a perturbation of the linear map and $\lambda \in \mathbb{C}$ is the multiplier of the origin, which shall either be $0 < |\lambda| < 1$ or $\lambda = e^{2\pi i\omega}$ with a Diophantine number ω, just for simplicity. Notice that $z = -\frac{\lambda}{\varepsilon}$ is also another fixed point, however we shall focus merely on the origin.

For every pair of values (λ, ε) provided, there is a change of coordinates $h(z)$ such that

$$f(h(z)) = h(\lambda z)$$

By writing it as

$$h(z) = z + \sum_{n=2}^{\infty} h_n z^n$$

we obtain the following algorithm to compute the coefficients h_n as

$$h_n = \frac{\varepsilon\beta_{n-1}}{\lambda^n - \lambda} \qquad \beta_{n-1} = \sum_{k=1}^{n-1} h_k h_{n-k}$$

where $h_1 = 1$ and $n \geq 2$.

Once the coefficients are computed up to a certain order, say N, an estimate of the radius of convergence of the power series of $h(z)$ can be obtained by the largest non-zero coefficient h_n as

$$R_c \approx \frac{1}{\sqrt[n]{|h_n|}}$$

as well as an approximation of the radius where the conjugacy holds by checking the largest $R_d > 0$ for which

$$|f(h(z) - h(\lambda z)| < tol$$

if $|z| \leq R_d$, where tol is a given a tolerance.

Let us start off by displaying some examples. We set $N = 100$, $\varepsilon = 0.5$, a tolerance of 10^{-8} and we take three different multipliers: $\lambda_1 = 0.2$, $\lambda_2 = -0.3 + i0.1$ and $\lambda_3 = e^{2\pi i\sqrt{2}}$. For each case, we compute seven orbits of $f_{\lambda_i,0.5}(z)$ and its linear map $\lambda_i z$ in a neighbourhood of the origin so as to show their behaviour locally. Then, we compute the series and both R_c and R_d. Within the area of conjugacy, we plot seven orbits of $f_{\lambda_i,0.5}(h(z))$ and $h(\lambda_i z)$ with the same initial condition for each orbit in order to visually compare the outcome.

Each point in a orbit is connected to the next one with a dashed line and each orbit has a colour assigned. The red circle displays the radius R_d.

1. $\lambda_1 = 0.2$

$$R_c \approx 0.101 \quad R_d \approx 0.087$$

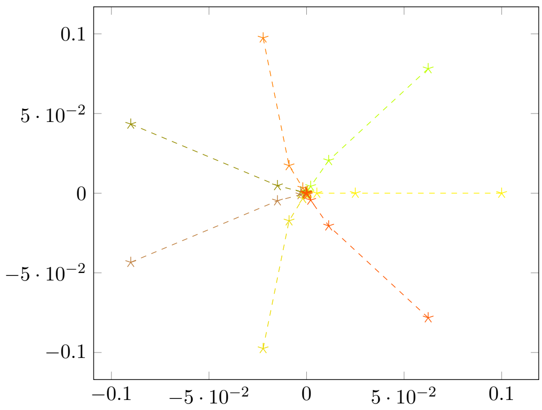

Figure 1: Seven orbits of $f_{\lambda_1,0.5}(z)$

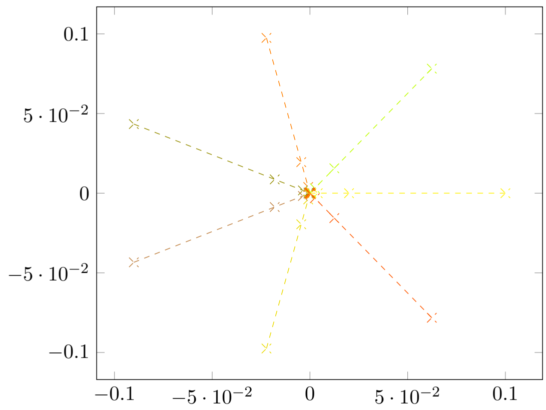

Figure 2: Seven orbits of $\lambda_1 z$

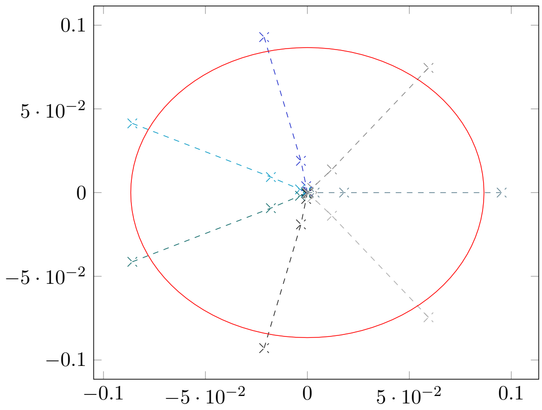

Figure 3: Seven orbits of $h(\lambda_1 z)$

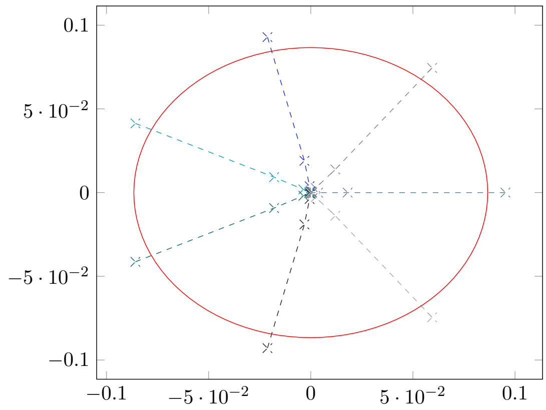

Figure 4: Seven orbits of $f_{\lambda_1,0.5}(h(z))$

2. $\lambda_2 = -0.3 + i0.1$

$$R_c \approx 0.180 \quad R_d \approx 0.153$$

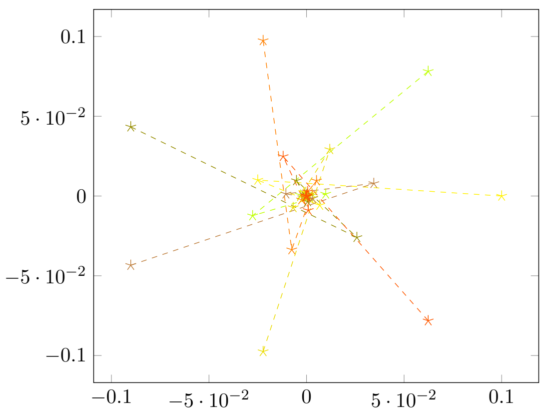

Figure 5: Seven orbits of $f_{\lambda_2,0.5}(z)$

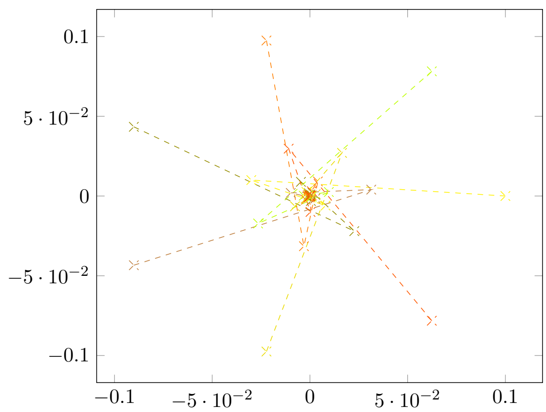

Figure 6: Seven orbits of $\lambda_2 z$

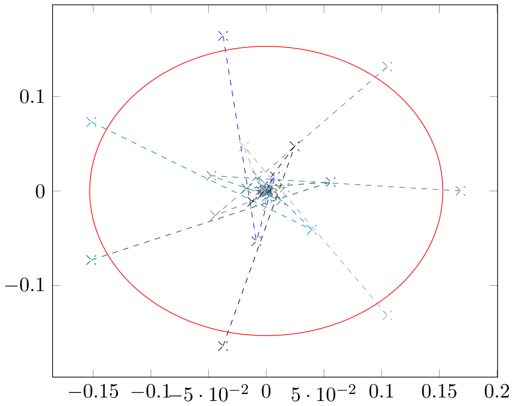

Figure 7: Seven orbits of $h(\lambda_2 z)$

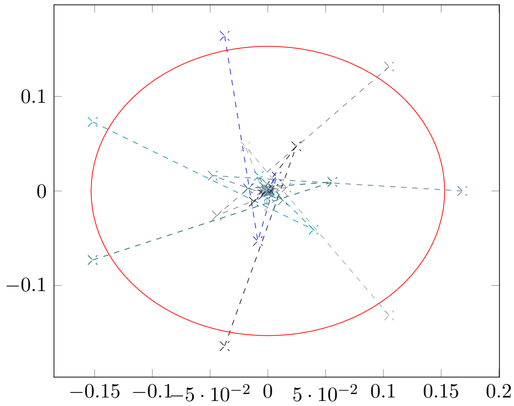

Figure 8: Seven orbits of $f_{\lambda_2,0.5}(h(z))$

3. $\lambda_3 = e^{2\pi i\sqrt{2}}$

$$R_c \approx 0.698 \quad R_d \approx 0.581$$

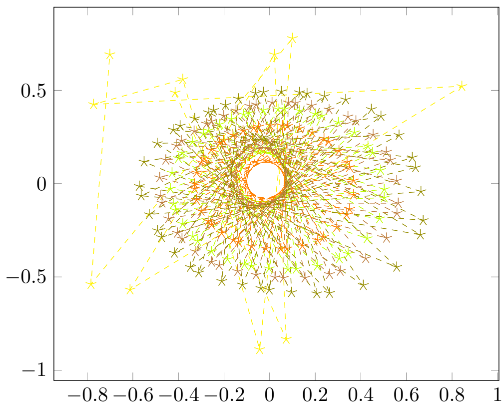

Figure 9: Seven orbits of $f_{\lambda_3,0.5}(z)$

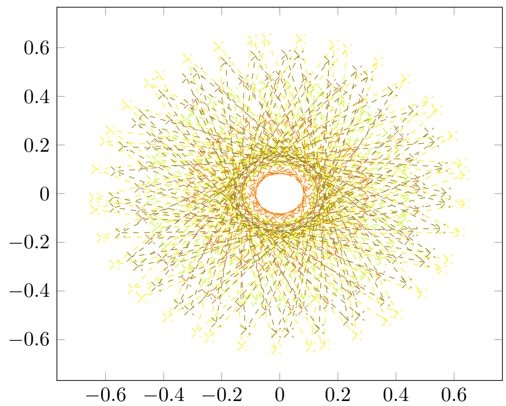

Figure 10: Seven orbits of $\lambda_3 z$

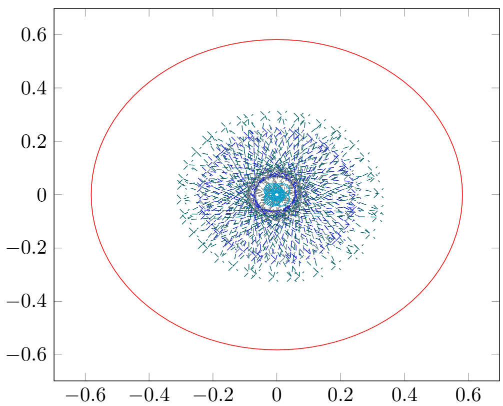

Figure 11: Seven orbits of $h(\lambda_3 z)$

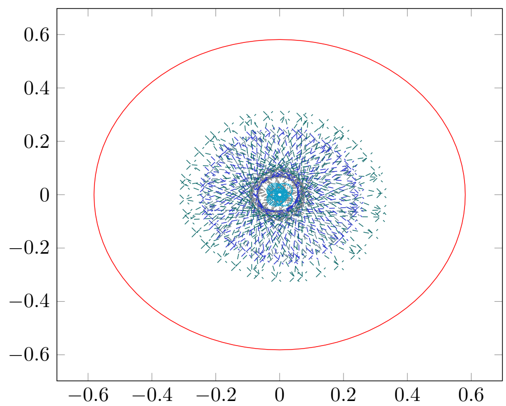

Figure 12: Seven orbits of $f_{\lambda_3,0.5}(h(z))$

As we can see, both hyperbolic cases show a clear correspondence between maps within the computed radius of convergence, which also happens to be quite close to the approximated radius of convergence of the series. Since $|\lambda| < 1$, the origin is a sink and therefore each map conveys a rather ease layout. However, when it comes to the Diophantine case, even though the ratio between both radius is similar, it is not that certain whether the conjugacy holds since, for instance, plotting orbits with initial conditions near the radius R_d has had many numerical difficulties.

So as to overcome this issue when dealing with Diophantine numbers, since $f(h(z))$ is expected to behave similarly to an irrational rotation and these provide a dense orbit for each initial condition, within the domain of conjugacy we shall expect the same. Therefore, by taking initial conditions on one of the axis and exploring how the orbits grow as their initial condition moves away from the origin, we can numerically determine where the end of the conjugacy lies.

For instance, when $\lambda = e^{2\pi i\sqrt{2}}$ and $\varepsilon = 0,5$ the aforementioned exploration leads to the following figure, which displays three orbits within the domain on conjugacy - in red, blue and green - and a fourth one in grey located on the boundary.

$$R_d \approx 0.581$$

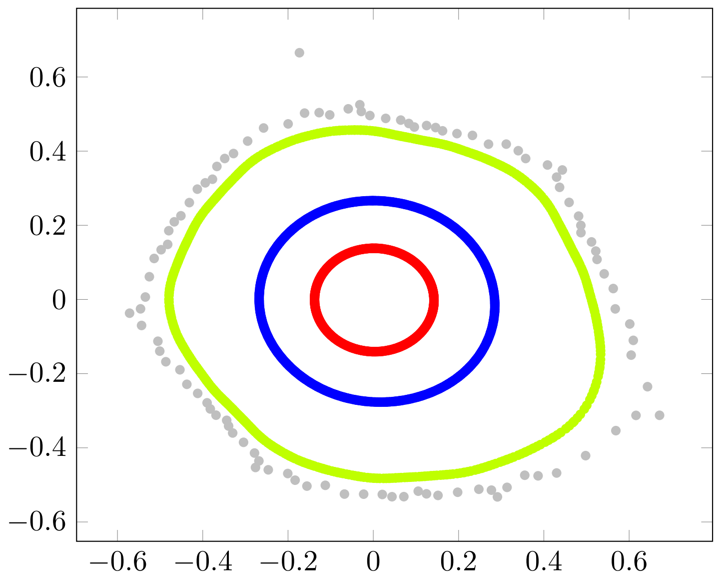

Figure 13: Four orbits - $\lambda = e^{2\pi i\sqrt{2}}$ and $\varepsilon = \frac{1}{2}$

The correspondence between the numeric boundary shown in the graphic and the one provided by the computed radius R_d is apparent as well as the quasi-periodic orbits that fill an entire somewhat circled area as expected.

Since this is a more appropriate way of studying the Diophantine case, let us now set $\lambda = e^{2\pi i\sqrt{5}}$ and explore a little bit further by varying the value of the perturbation ε. Let us plot the previous orbits for four different values.

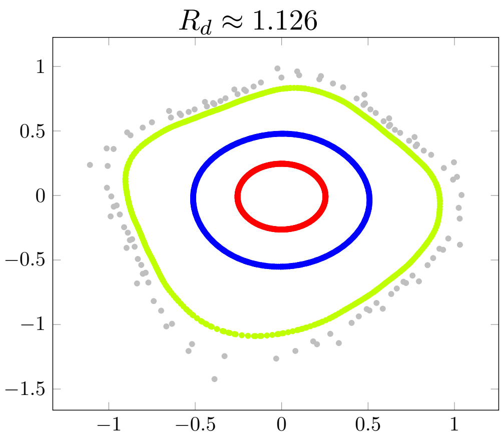

Figure 14: Four orbits - $\varepsilon = \frac{1}{4}$

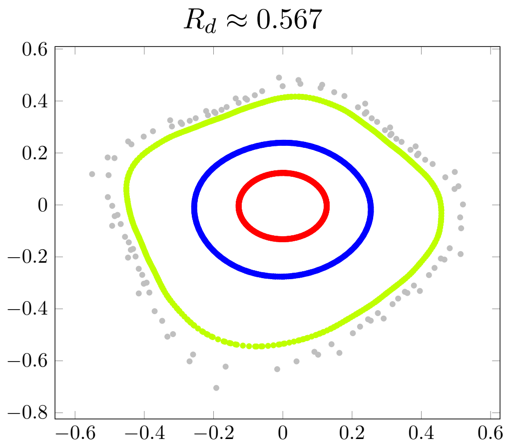

Figure 15: Four orbits - $\varepsilon = \frac{1}{2}$

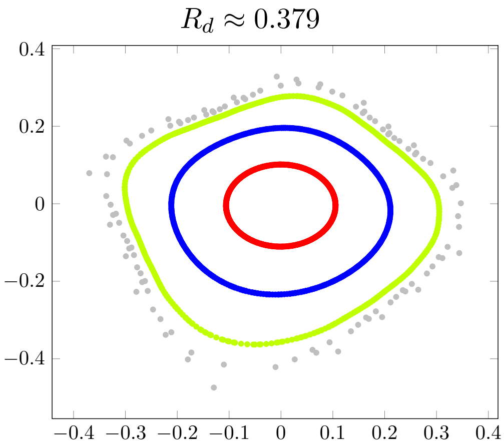

Figure 16: Four orbits - $\varepsilon = \frac{3}{4}$

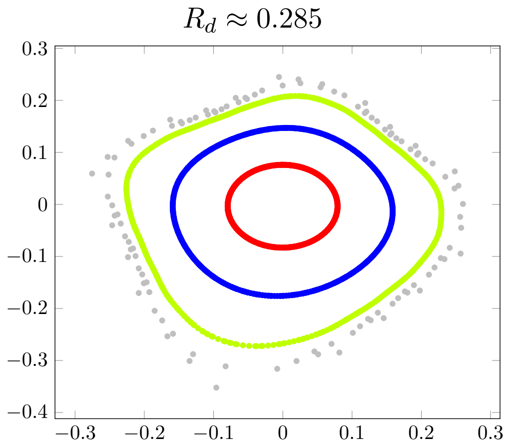

Figure 17: Four orbits - $\varepsilon = 1$

As we can see, even as the map is less closer to the linear map λz, there is a persistent neighbourhood where these quasi-periodic orbits remain. In all four cases, the correspondence between R_d and the domain shown by the graphic still holds, however there is an apparent shrinking of the radius according to the perturbation.

In fact, many numerical studies on the border of the domain of conjugacy can be considered, but thus far, we have been able to spot the linearization in all cases where we knew it existed and we have also computed an approximation of the domain, which accomplishes the main goals in this numerical study. One last exploration we shall consider is the study of the growth of both radius varying ε as well as the number of coefficients so as to get an idea on how it might affect the conjugacy and its domain.

When it comes to the number of coefficients, there is barely a time cost when computing N=100, which seems at first a fair number. However the operations needed afterwards spend significantly more time as N increases, so setting a maximum of N=250, we compute both radius for different numbers of coefficients, starting off with 10 coefficients and increasing N in five. We set $\varepsilon = \frac{1}{2}$. The outcome follows.

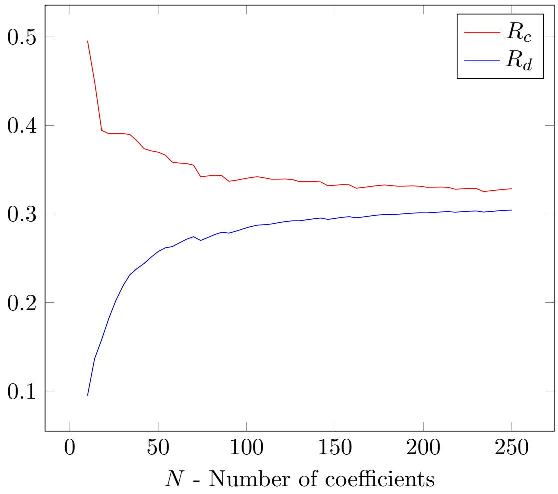

Figure 18: Growth of R_c and R_d varying the number of coefficients N

The radius of convergence of the series clearly becomes smaller as more coefficients are computed, thus showing that the error was increasing its value. It seems as if the value of the radius tends to range in a small interval near 0.3, which is achieved around the 100 coefficients. Exactly the same pattern but increasing instead, R_d grows as more coefficients are computed and it also seems to tend to range within the same interval. This behaviour was to be expected because the domain of conjugacy cannot be larger than the radius where the series converge and although we might not extrapolate how both radius behave as more coefficients are computed, since they both met within the interval they range in at N=100, it seems indeed as a fair number of coefficients to be computed.

As for the graphic varying ε, we set N=100 and starting off with $\varepsilon = 0.01$ and increasing it by 0.01, we compute both radius until the value $\varepsilon = 1$ is achieved.

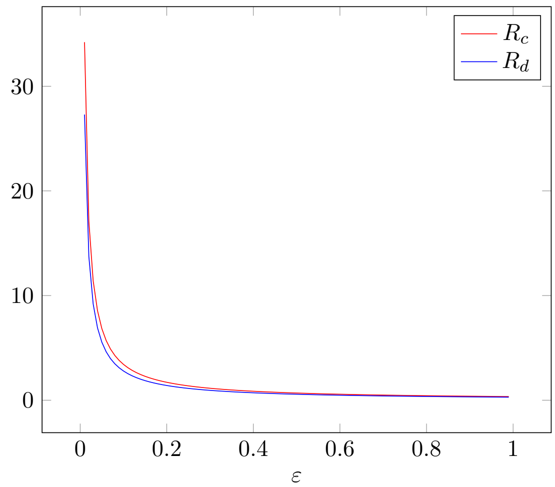

Figure 19: Growth of R_c and R_d varying the perturbation value ε

The graphic shows that throughout the unit interval where ε ranges, the correlation between both radius is very tight, an outcome to be expected from the previous discussion on the number of coefficients. As for their growth, it is apparent that the closer to the rotation, the bigger the domain is, whereas when the perturbation becomes bigger, the linearization takes places in a much smaller place, in fact, the loss of domain is quite remarkable.

This numerical exploration has therefore led to different, yet wonted, results on the matter. First of all, all linearizations were indeed spotted by using a rather uncomplicated algorithm and a suitable number of coefficients. When dealing with the Diophantine case, looking for quasi-periodic orbits instead of plotting some orbits with random initial conditions proved to be more useful in order to determine an area where the conjugacy held. It is important to notice that an analogous study can also be applied for the hyperbolic case if an alternative method to explore the conjugacy is desired. Second, it seems as if both radius tend to the same or at least a similar value as the number of computed coefficients grow, being 100 coefficients a fair amount. The boundary of such domain clearly arises as a very interesting object of study in case of a further analysis on the matter. And finally, it is also conveyed the idea that the bigger the perturbation is, the smaller the linearization seems to become.

6 Conclusions

It is now time to conclude this dissertation by providing an overview of it and going through it as a whole. There are a three fundamental aspects to ponder over due to the nature of a graduate thesis that we shall now analyze.

First of all, it is the achievements accomplished and the enhancements to be considered. The main aim was to provide, as well as personally gaining, an insight to Siegel's linearization theorem and all it encompasses. Within a reasonable amount of theory, the dissertation briefly goes through all the fields that provide the reader with the tools and then sets out a background for the theorem, which is straightforwardly attained. However, even though an overview of the conjugacy issue near an elliptic fixed point is put forward, it is undeniable that since many questions still remain, a deeper inquiry about the matter could have been made, especially in terms of continued fractions and unsolved characterizations.

When it comes to the proof, it carefully goes through most of its details, despite there are refinements on the estimates that could be improved. In fact, the KAM method is one of the standard procedures used to prove the result, but other options can be considered, as well as bounds and operators. It requires a great deal of experience to precisely know which one should be chosen in order to suit and emphasize the author's needs conveniently. As for the numerical study, all graphics meet and broadly satisfy what was meant to be expected, although arose many other subjects of study such as the border of the domain of conjugacy, among others.

Secondly, since this is a graduate thesis, there is an extricable link to analyze between the degree and the herein dissertation. Subjects as *Models Matemàtics i Sistemes Dinàmics* and *Sistemes Dinàmics* have both predisposed a certain, yet crucial, facility when dealing with discrete dynamics concepts, more precisely in complex dynamics and numerical explorations correspondingly. It also needs to be pointed out that *Anàlisi Complexa* and *Funcions de Variable Complexa* have contributed significantly to the easiness of handling holomorphic functions and most of its properties throughout this dissertation, which played an essential role. Finally, it is worth mentioning *Anàlisi Real i Funcional* for which notions related to Banach spaces as well as operators became less of a burden.

And last but not least, the consistency of it as a whole. While it is true that some improvements and deeper insights on different matters could have been made as it has been already mentioned, this thesis gives a fulfilling overview of the subject in terms of providing the reader with the necessary knowledge, setting out a background for Schröder's equation, carefully dealing with every step of the proof and concluding the study with some numerical explorations to display some of the acquired results. Thus becoming a well-rounded dissertation and an easy introduction of the subject for undergraduates students.

References

[1] J. Steuding, *Diophantine Analysis*, Chapman and Hall/CRC, 2005, 36–51 and 142–145.

[2] T.W. Gamelin, *Complex Analysis*, Springer, 2001, 33–53.

[3] J.B. Conway, *Functions of one complex variable*, Springer, 1978, 142–154.

[4] A.L. Brown and A. Page, *Elements of functional analysis*, Van Nostrand Reinhold, 1970, 263–300.

[5] R.A. Holmgren, *A first course in discrete dynamical systems*, Springer, 1996, 80–91.

[6] R.L. Devaney, *An introduction to chaotic Dynamical Systems*, Westview Press, 2003, 17–59, 260–267 and 300–310.

[7] L. Carleson and T.W. Gamelin, *Complex Dynamics*, Springer, 1993, 27–46.

[8] A. Morosawa, Y. Nishimura, M. Taniguchi and T. Ueda, *Holomorphic Dynamics*, Cambridge University Press, 2000, 12–20.

[9] J.W. Milnor, *Dynamics in one complex variable*, Princeton University Press, 2006, 76–139.

[10] M. Abate, E. Bedford, M. Brunella, T. Dinh, D. Schleicher and N. Sibony, *Holomorphic Dynamical Systems*, Lecture notes in mathematics (Springer-Verlag), 1998, 4–31.

[11] A. Katok and B. Hasselblatt, *Introduction to the modern theory of dynamical systems*, Cambridge University Press, 1995, 90–99.