Chapter 3

Around the boundary of complex dynamics

Roland K. W. Roeder[1]

Dedicated to Emile and Eli

Contents

Preface

These notes were written for the 2015 "Thematic Program on Boundaries and Dynamics" held at Notre Dame University. They are intended for an advanced undergraduate student who is majoring in mathematics. In an ideal world, a student reading these notes will have already taken undergraduate level courses in complex variables, real analysis, and topology. As the world is far from ideal, we will also review the needed material.

There are many fantastic places to learn complex dynamics, including the books by Beardon [3], Carleson and Gamelin [10], Devaney [11, 12], Milnor [37], and Steinmetz [45], as well as the Orsay notes [13] by Douady and Hubbard, the surveys by Blanchard [5] and Lyubich [31, 33], and the invitation to transcendental dynamics by Shen and Rempe-Gillen [43]. The books by Devaney and the article by Shen and Rempe-Gillen are especially accessible to undergraduates. We will take a complementary approach, following a somewhat different path through some of the same

[1]rroeder@math.iupui.edu
IUPUI Department of Mathematical Sciences, LD Building, Room 270, 402 North Blackford Street, Indianapolis, IN 46202, United States.
This work was partially supported by NSF grant DMS-1348589.

material as presented in these sources. We will also present modern connections at the boundary between complex dynamics and other areas.

None of the results presented here are new. In fact, I learned most of them from the aforementioned textbooks and from courses and informal discussions with John Hubbard and Mikhail Lyubich.

Our approach is both informal and naive. We make no effort to provide a comprehensive or historically complete introduction to the subject. Many important results will be omitted. Rather, we will simply have fun doing mathematics.

Lecture 1. Warm up

Let's start at the very beginning:

1.1 Complex numbers. Recall that a complex number has the form $z = x + iy$, where $x, y \in \mathbb{R}$ and i satisfies $i^2 = -1$. One adds, subtracts, multiplies, and divides complex numbers using the following rules:

$$(a+bi) \pm (c+di) = (a \pm c) + (b \pm d)i,$$
$$(a+bi)(c+di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i, \quad \text{and}$$
$$\frac{a+bi}{c+di} = \frac{(a+bi)}{(c+di)}\frac{(c-di)}{(c-di)} = \frac{(ac+bd)+(bc-ad)i}{c^2+d^2}.$$

The set of complex numbers forms a field $\mathbb{C}$ under the operations of addition and multiplication.

The *real part* of $z = x + iy$ is $\mathrm{Re}(z) = x$ and the *imaginary part* of $z = x + iy$ is $\mathrm{Im}(z) = y$. One typically depicts a complex number in the *complex plane* using the horizontal axis to measure the real part and the vertical axis to measure the imaginary part; see Figure 3.1. One can also take the real or imaginary part of more complicated expressions. For example, $\mathrm{Re}(z^2) = x^2 - y^2$ and $\mathrm{Im}(z^2) = 2xy$.

The *complex conjugate* of $z = x + iy$ is $\overline{z} = x - iy$ and the *modulus* of z is $|z| = \sqrt{x^2 + y^2} = \sqrt{z\overline{z}}$. In the complex plane, $\overline{z}$ is obtained by reflecting z across the real axis and $|z|$ is the distance from z to the origin $0 = 0 + 0i$. The *argument* of $z \neq 0$ is the angle counterclockwise from the positive real axis to z.

A helpful tool is the

Triangle inequality. *For every $z, w \in \mathbb{C}$ we have*

$$|z| - |w| \leq |z + w| \leq |z| + |w|.$$

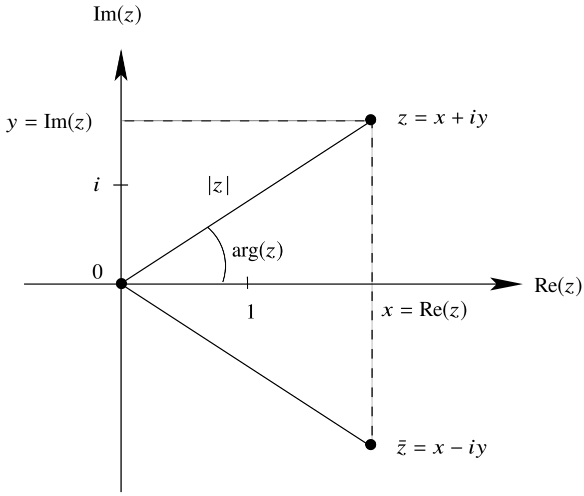


Figure 3.1. The complex plane.

A *complex polynomial* $p(z)$ of degree d is an expression of the form

$$p(z) = a_d z^d + a_{d-1} z^{d-1} + \cdots + a_1 z + a_0,$$

where $a_d, \ldots, a_0$ are some given complex numbers with $a_d \neq 0$. Historically, complex numbers were introduced so that the following theorem holds:

Fundamental theorem of algebra. *A polynomial $p(z)$ of degree d has d complex zeros $z_1, \ldots, z_d$, counted with multiplicity.*

In other words, a complex polynomial $p(z)$ can be factored over the complex numbers as

$$p(z) = c(z - z_1)(z - z_2) \cdots (z - z_d), \tag{1.1}$$

where $c \neq 0$ and some of the roots z_j may be repeated. (The number of times z_j is repeated in (1.1) is the *multiplicity* of z_j as a root of p.)

Multiplying and dividing complex numbers is often simpler in *polar form*. Euler's formula states

$$\mathrm{e}^{i\theta} = \cos\theta + i \sin\theta \qquad \text{for any } \theta \in \mathbb{R}.$$

We can therefore represent any complex number $z = x + iy$ by $z = r\mathrm{e}^{i\theta}$ where $r = |z|$ and $\theta = \arg(z)$. Suppose $z = r\mathrm{e}^{i\theta}$ and $w = s\mathrm{e}^{i\phi}$ and $n \in \mathbb{N}$. The simple formulae

$$zw = rs\mathrm{e}^{i(\theta+\phi)}, \qquad z^n = r^n \mathrm{e}^{in\theta}, \quad \text{and} \quad \frac{z}{w} = \frac{r}{s}\mathrm{e}^{i(\theta-\phi)} \tag{1.2}$$

follow from the rules of exponentiation. Multiplication and taking powers of complex numbers in polar form are depicted geometrically in Figure 3.2.

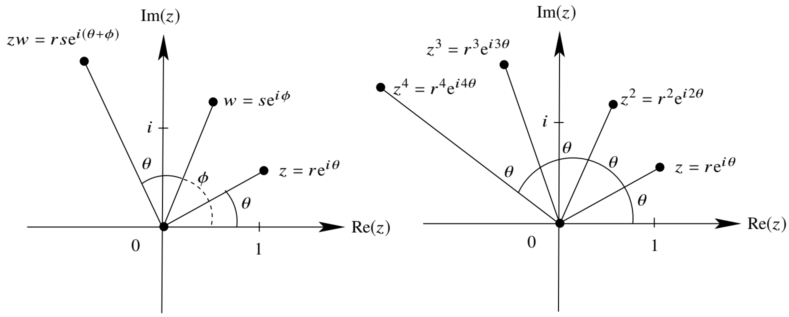


Figure 3.2. Multiplication and taking powers in polar form.

1.2 Iterating linear maps. A *linear map* $L : \mathbb{C} \to \mathbb{C}$ is a mapping of the form $L(z) = az$, where $a \in \mathbb{C} \setminus \{0\}$. Suppose we take some *initial condition* $z_0 \in \mathbb{C}$ and repeatedly apply L:

$$z_0 \longrightarrow L(z_0) \longrightarrow L(L(z_0)) \longrightarrow L(L(L(z_0))) \longrightarrow \cdots . \tag{1.3}$$

For any natural number $n \geq 1$ let $L^{\circ n} : \mathbb{C} \to \mathbb{C}$ denote the composition of L with itself n times. We will often also use the notation

$$z_n := L^{\circ n}(z_0).$$

The sequence $\{z_n\}_{n=0}^{\infty} \equiv \{L^{\circ n}(z_0)\}_{n=0}^{\infty}$ is called the sequence of iterates of z_0 under L. It is also called the *orbit* of z_0 under L.

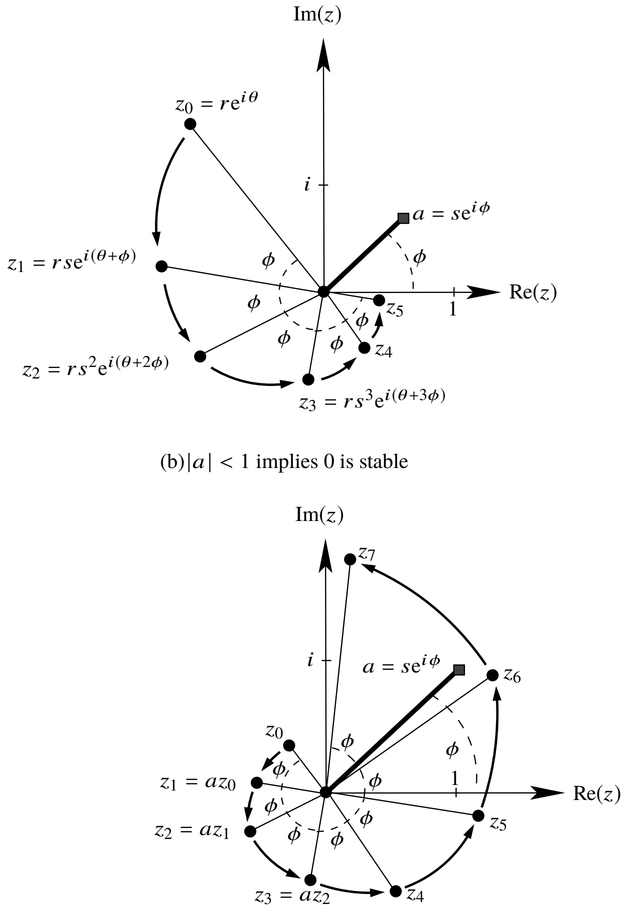
Remark. The notion of *linear* used above is from your course on linear algebra: a linear map must satisfy $L(z + w) = L(z) + L(w)$ for all $z, w \in \mathbb{C}$ and $L(cz) = cL(z)$ for all $z, c \in \mathbb{C}$. For this reason, mappings of the form $z \mapsto az + b$ are *not considered* linear. Instead, they are called *affine*. (See Exercise 1.1.)

The number a is called a *parameter* of the system. We think of it as describing the overall state of the system (think, for example, temperature or barometric pressure) that is fixed for all iterates n. One can change the parameter to see how it affects the behavior of sequences of iterates (for example, if the temperature is higher, does the orbit move farther in each step?).

Our rules for products and powers in polar form (1.2) allow us to understand the sequence of iterates (1.3). Suppose $z_0 = re^{i\theta}$ and $a = se^{i\theta}$ with $r, s > 0$. Then, the behavior of the iterates depends on $s = |a|$, as shown in Figure 3.3.

Remark. For a linear map $L(z) = az$ with $|a| \neq 1$, the orbits $\{z_n\}$ and $\{w_n\}$ for any two nonzero initial conditions z_0 and w_0 have the same dynamical behavior. If $|a| < 1$ then

$$\lim_{n\to\infty} z_n = 0 = \lim_{n\to\infty} w_n,$$



(b)$|a| < 1$ implies 0 is stable

(a)$|a| > 1$ implies 0 is unstable

Figure 3.3. Iterating the linear map $L(z) = az$. Above: $|a| < 1$ implies orbits spiral into 0. Below: $|a| > 0$ implies orbits spiral away from 0. Not shown: $|a| = 1$ implies orbits rotate around 0 at constant modulus.

and if $|a| > 1$ then

$$\lim_{n\to\infty} z_n = \infty = \lim_{n\to\infty} w_n.$$

This is atypical for dynamical systems—the long-term behavior of the orbit usually depends greatly on the initial condition. For example, we will soon see that when iterating the quadratic mapping $p(z) = z^2 + \frac{i}{4}$ there are many initial conditions whose orbits remain bounded and many whose orbit escapes to ∞. There will also be many initial conditions whose orbits have completely different behavior! Linear maps are just too simple to have interesting dynamical properties.

Exercise 1.1. An affine mapping $A : \mathbb{C} \to \mathbb{C}$ is a mapping given by $A(z) = az + b$, where $a, b \in \mathbb{C}$ and $a \neq 0$. Show that iteration of affine mappings produces no dynamical behavior that was not seen when iterating linear mappings.

1.3 Iterating quadratic polynomials. Matters become far more interesting if one iterates quadratic mappings $p_c : \mathbb{C} \to \mathbb{C}$ given by $p_c(z) = z^2 + c$. Here, c is a parameter, which we sometimes include in the notation by means of a subscript, writing $p_c(z)$, and sometimes omit, writing simply $p(z)$.

Remark. Like in Exercise 1.1, one can show that quadratic mappings of the form $p_c(z) = z^2 + c$ actually capture all of the types of dynamical behavior that can arise when iterating a more general quadratic mapping $q(z) = az^2 + bz + c$.

Applying the mapping p_c can be understood geometrically in two steps: one first squares z using the geometric interpretation provided in polar coordinates (1.2). One then translates (shifts) the result by c. This two-step process is illustrated in Figure 3.4.

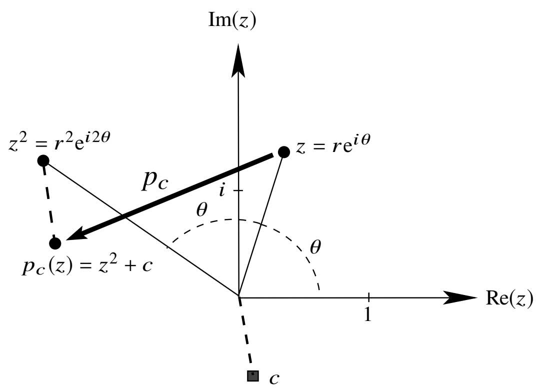


Figure 3.4. Geometric interpretation of applying $p_c(z) = z^2 + c$.

Remark. Solving the exercises in this subsection may require some of the basic complex analysis from the following subsection. They are presented here for better flow of the material.

Example 1.2 (Exploring the dynamics of $p_c : \mathbb{C} \to \mathbb{C}$ for $c = \frac{i}{4}$.). In Figure 3.5 we show the first few iterates under $p(z) = z^2 + \frac{i}{4}$ of two different orbits: $\{z_n\}$ of initial condition $z_0 = i$ and $\{w_n\}$ of initial condition $w_0 = 1.1\,i$. Note that orbit $\{z_n\}$ seems to converge to a point $z \approx -0.05 + 0.228i$ while orbit $\{w_n\}$ seems to escape to ∞.

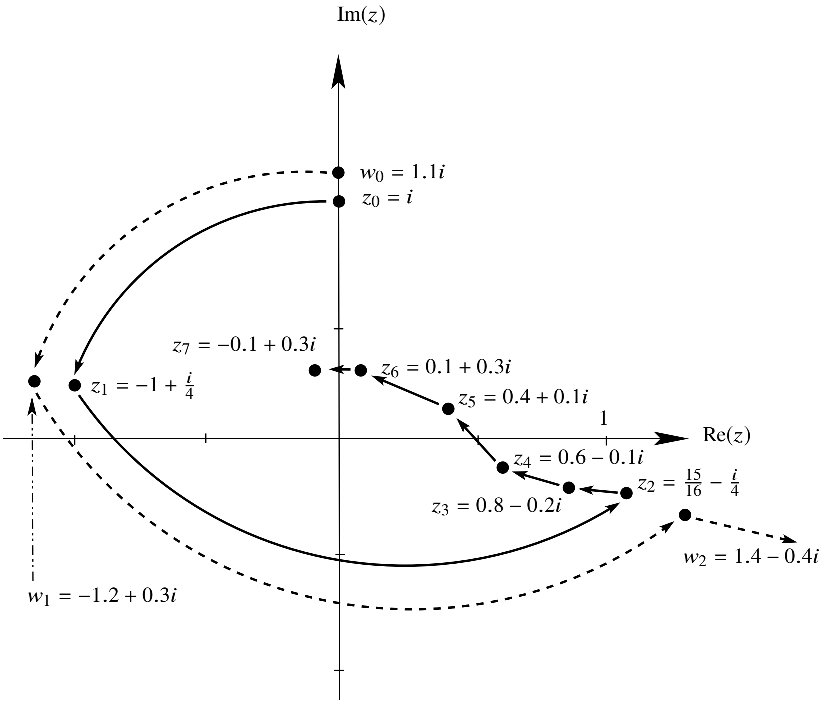


Figure 3.5. Orbits $\{z_n\}$ for initial condition $z_0 = i$ and $\{w_n\}$ for $w_0 = 1.1i$ under $p(z) = z^2 + \frac{i}{4}$.

Exercise 1.3. Use the quadratic formula to prove that there exists $z_\bullet \in \mathbb{C}$ that is close to $-0.05 + 0.228i$ and satisfies

$$p(z_\bullet) = z_\bullet.$$

Such a point is called a *fixed point* for $p(z)$ because if you use $z_\bullet$ as the initial condition the orbit is a constant sequence $\{z_\bullet, z_\bullet, z_\bullet, \ldots\}$.

Show that there is a second fixed point z_* for $p(z)$ with $z_* \approx 1.05 - 0.228i$.

Compute $|p'(z_\bullet)|$ and $|p'(z_*)|$, where $p'(z) = 2z$ is the derivative of $p(z) = z^2 + \frac{i}{4}$. Use the behavior of linear maps, as shown in Figure 3.3, to make a prediction about the behavior of orbits for $p(z)$ near each of these fixed points.

Exercise 1.4. Let $z_\bullet$ be the fixed point for $p(z)$ discovered in Exercise 1.3. Prove that for any point z_0 sufficiently close to $z_\bullet$, the orbit $\{z_n\}$ under $p(z) = z^2 + \frac{i}{4}$ converges to $z_\bullet$ (i.e., prove that there exists $\delta > 0$ such that for any z_0 satisfying $|z_0 - z_\bullet| < \delta$ and any $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$ we have $|z_n - z_\bullet| < \epsilon$).

Why does your proof fail if you replace the fixed point $z_\bullet$ with z_*?

Now, prove that the orbit of $z_0 = i$ converges to $z_\bullet$.

Exercise 1.5. Prove that there exists $r > 0$ such that for any initial condition z_0 with $|z_0| > r$, the orbit $\{z_n\}$ of z_0 under $p(z) = z^2 + \frac{i}{4}$ escapes to ∞ (i.e., prove that there exists $r > 0$ such that for any z_0 satisfying $|z_0| > r$ and any $R > 0$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$ we have $|z_n| > R$).

Now prove that the orbit of $w_0 = 1.1i$ escapes to ∞.

Example 1.6 (Exploring the dynamics of $p_c : \mathbb{C} \to \mathbb{C}$ for $c = -1$.). In Figure 3.6 we show the first few iterates under $p(z) = z^2 - 1$ of two different orbits: $\{z_n\}$ of initial condition $z_0 \approx 0.08 + 0.66i$ and $\{w_n\}$ of initial condition $w_0 = \frac{\sqrt{2}}{2}(1+i)$. Orbit $\{z_n\}$ seems to converge to a periodic behavior ("periodic orbit") while $\{w_n\}$ seems to escape to ∞.

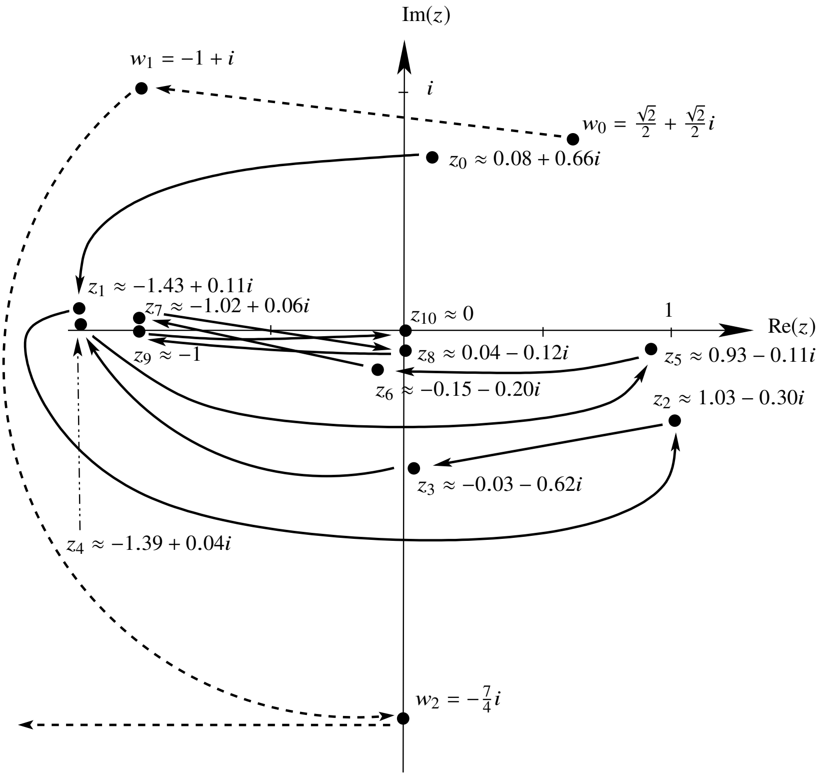


Figure 3.6. Orbits $\{z_n\}$ of $z_0 \approx 0.08 + 0.66i$ and $\{w_n\}$ of $w_0 = \frac{\sqrt{2}}{2}(1+i)$ under the quadratic polynomial $p_{-1}(z) = z^2 - 1$.

In fact, the periodic orbit that $\{z_n\}$ seems to converge to is easy to find for this mapping. If we use initial condition $u_0 = 0$ we have

$$u_1 = p_{-1}(u_0) = 0^2 - 1 = -1.$$

Then,

$$u_2 = p(u_1) = p(-1) = (-1)^2 - 1 = 0 = u_0.$$

We conclude that the orbit of $u_0 = 0$ is periodic with period 2:

$$\begin{tikzcd} 0 \arrow[r, "p_{-1}"] & -1. \arrow[l, bend left=60, "p_{-1}"] \end{tikzcd}$$

(Subsequently, this periodic orbit will be denoted $0 \leftrightarrow 1$.)

The following two exercises are in the context of Example 1.6.

Exercise 1.7. Make precise the statement that if z_0 is an initial condition sufficiently close to 0, then its orbit "converges to the periodic orbit $0 \leftrightarrow 1$." Prove the statement.

Now, suppose $z_0 \approx 0.08 + 0.66i$ and prove that its orbit converges to the periodic orbit $0 \leftrightarrow 1$.

Exercise 1.8. Find an initial condition $z_0 \in \mathbb{C}$ such that for any $\epsilon > 0$ there are

(1) infinitely many initial conditions w_0 with $|w_0 - z_0| < \epsilon$ having orbit $\{w_n\}$ under p_{-1} that remains bounded, and

(2) infinitely many initial conditions u_0 with $|u_0 - z_0| < \epsilon$ having orbit $\{u_n\}$ under p_{-1} that escapes to ∞.

Hint: Work within $\mathbb{R}$ and consider the graph of $p(x) = x^2 - 1$.

Example 1.9 (Exploring the dynamics of $p_c : \mathbb{C} \to \mathbb{C}$ for $c = \frac{1}{2}$.). As in the previous two examples, we will try a couple of arbitrary initial conditions. Figure 3.7 shows the orbits of initial conditions $z_0 = 0$ and $w_0 \approx 0.4 + 0.6i$ under $p(z) = z^2 + \frac{1}{2}$. Both orbits seem to escape to ∞.

Exercise 1.10. Prove that for any real initial condition $z_0 \in \mathbb{R}$ the orbit $\{z_n\}$ under $p(z) = z^2 + \frac{1}{2}$ escapes to ∞.

Exercise 1.11. Determine whether there is any initial condition z_0 for which the orbit under $p_{1/2}$ remains bounded.

Exercise 1.12. Repeat the type of exploration done in Examples 1.2–1.9 for

$$c = 0, \quad c = -2, \quad c = i, \quad \text{and} \quad c = -0.1 + 0.75i.$$

Try other values of c.

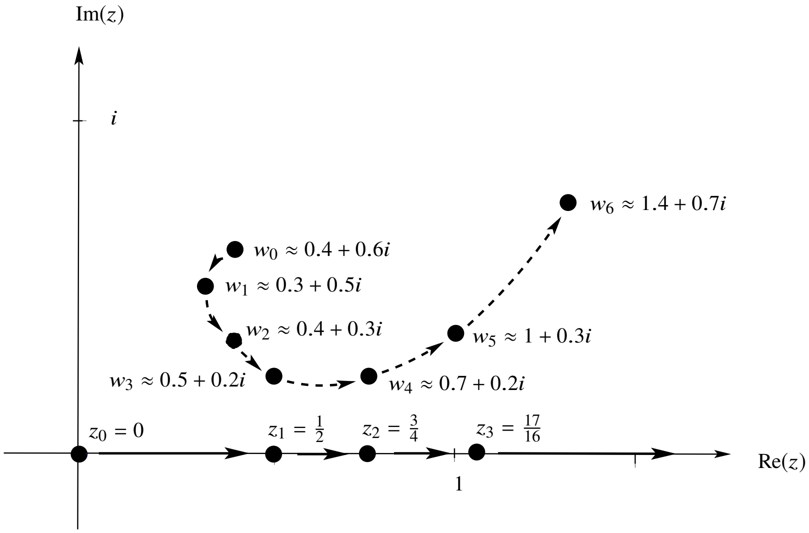


Figure 3.7. Orbits $\{z_n\}$ of $z_0 = 0$ and $\{w_n\}$ of $w_0 \approx 0.4 + 0.6i$ under the quadratic polynomial $p_{1/2}(z) = z^2 + \frac{1}{2}$.

1.4 Questions. During our explorations we've discovered several questions. Some of them were answered in the exercises, but several of them are still open:

(1) Does every quadratic map have some initial condition z_0 whose orbit escapes to ∞?

(2) Does every quadratic map have some periodic orbit

$$z_0 \mapsto z_1 \mapsto z_2 \mapsto \cdots \mapsto z_n \mapsto z_0$$

which attracts the orbits of nearby initial conditions? $\Big($Perhaps we didn't look hard enough for one when $c = \frac{1}{2}$?$\Big)$

(3) Can a map $p_c(z)$ have more than one such attracting periodic orbit?

(4) For any $m \geq 1$, does there exist a parameter c such that $p_c(z)$ has an attracting periodic orbit of period m?

Exercise 1.13. Answer question (1) by showing that for any c there is a radius $R(c)$ such that for any initial condition z_0 with $|z_0| > R(c)$, the orbit $\{z_n\}$ escapes to ∞.

Generalize your result to prove that for any polynomial $q(z)$ of degree at least 2, there is some $R > 0$ so that any initial condition z_0 with $|z_0| > R$ has orbit $\{z_n\}$ that escapes to ∞.

1.5 Crash course in complex analysis. In order to answer the questions posed in the previous subsection and explore the material more deeply, we will need some basic tools from complex analysis. We have slightly adapted the following results from the textbook by Saff and Snider [42]. We present at most sketches of the proofs and leave many of the details to the reader.

This subsection is rather terse. The reader may want to initially skim over it and then move forwards to see how the material is used in the later lectures.

We begin with some topological properties of $\mathbb{C}$. The *open disk* of radius $r > 0$ centered at z_0 is $D(z_0, r) := \{z \in \mathbb{C} : |z - z_0| < r\}$.

Definition 1.14. A set $S \subset \mathbb{C}$ is *open* if for every $z \in S$ there exists $r > 0$ such that $D(z, r) \subset S$. A set $S \subset \mathbb{C}$ is *closed* if its complement $\mathbb{C} \setminus S$ is open.

Exercise 1.15. Prove that for any $z_0 \in \mathbb{C}$ and any $r > 0$, the "open disk" $D(z_0, r)$ is *actually* open. Then prove that the set

$$\overline{D(z_0, r)} := \{z \in \mathbb{C} : |z - z_0| \le r\}$$

is closed. It is called the *closed disk* of radius r centered at z_0.

Definition 1.16. The *boundary* of $S \subset \mathbb{C}$ is

$$\partial S := \{z \in \mathbb{C} : D(z, r) \text{ contains points in } S \text{ and in } \mathbb{C} \setminus S \text{ for every } r > 0\}.$$

Definition 1.17. A set $S \subset \mathbb{C}$ is *disconnected* if there exist open sets U and V with

(i) $S \subset U \cup V$;

(ii) $S \cap U \neq \emptyset$ and $S \cap V \neq \emptyset$; and

(iii) $U \cap V = \emptyset$.

A set $S \subset \mathbb{C}$ is *connected* if it is not disconnected.

An open connected $U \subset \mathbb{C}$ is called a *domain*. Any set denoted U in this subsection will be assumed to be a domain. If $z_0 \in U$, a *neighborhood* of z_0 will be another domain $V \subset U$ with $z_0 \in V$. (A round disk $D(z_0, r)$ for some $r > 0$ sufficiently small will always suffice.)

Definition 1.18. A *contour* $\gamma \subset U$ is a piecewise smooth function $\gamma : [0, 1] \to U$. (Here, the notation implicitly identifies the function $\gamma : [0, 1] \to U$ with its image $\gamma[0, 1] \equiv \gamma \subset U$.)

A contour γ is *closed* if $\gamma(0) = \gamma(1)$. A closed contour γ is *simple* if $\gamma(s) \neq \gamma(t)$ for $t \neq s$ unless $t = 0$ and $s = 1$ or vice versa. (Informally, a simple closed contour is a loop that does not cross itself.)

A simple closed contour is *positively oriented* if as you follow the contour, the region it encloses is on your left. (Informally, this means that it goes counterclockwise.)

Remark. An open set S is connected if and only if for every two points $z, w \in S$ there is a contour $\gamma \subset S$ with $\gamma(0) = z$ and $\gamma(1) = w$.

Definition 1.19. A domain $U \subset \mathbb{C}$ is *simply connected* if any closed contour $\gamma \subset U$ can be continuously deformed within U to some point $z_0 \in U$.

We refer the reader to [42, Section 4.4, Definition 5] for the formal definition of *continuously deformed*. In these notes, we will need only that the disk $D(z_0, r)$ is simply connected. It follows from the fact that any closed contour $\gamma \subset D(z_0, r)$ can be affinely scaled within $D(z_0, r)$ down to the center z_0.

Remark. You have seen Definition 1.19 in your multivariable calculus class, where it was used in the statement of Green's theorem.

Definition 1.20. A set $K \subset \mathbb{C}$ is *compact* if for any collection $\{W_\lambda\}_{\lambda \in \Lambda}$ of open sets with

$$K \subset \bigcup_{\lambda \in \Lambda} W_\lambda,$$

there are a finite number of sets $W_{\lambda_1}, \ldots, W_{\lambda_n}$ so that

$$K \subset W_{\lambda_1} \cup \cdots \cup W_{\lambda_n}.$$

Heine–Borel theorem. *A set $S \subset \mathbb{C}$ is compact if and only if it is closed and bounded.*

Exercise 1.21. Suppose $K_1 \supset K_2 \supset K_3 \supset \cdots$ is a nested sequence of nonempty connected compact sets in $\mathbb{C}$. Prove that $\bigcap_{n \geq 1} K_n$ is nonempty and connected.

We are now ready to start doing complex calculus. The notion of a limit is defined in exactly the same way as in calculus, except that modulus $|\cdot|$ takes the place of absolute value.

Definition 1.22. Let $z_0 \in U$ and let $f : U \setminus \{z_0\} \to \mathbb{C}$ be a function. We say that $\lim_{z \to z_0} f(z) = L$ for some $L \in \mathbb{C}$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that $0 < |z - z_0| < \delta$ implies $|f(z) - L| < \epsilon$.

If we write

$$f(z) = u(x, y) + iv(x, y)$$

with $u : \mathbb{R}^2 \to \mathbb{R}$ and $v : \mathbb{R}^2 \to \mathbb{R}$, then $\lim_{z \to z_0} f(z) = L$ if and only if

$$\lim_{(x,y) \to (x_0, y_0)} u(x, y) = \mathrm{Re}(L) \quad \text{and} \quad \lim_{(x,y) \to (x_0, y_0)} v(x, y) = \mathrm{Im}(L).$$

(The limits on the right-hand side are taken as in the sense of your multivariable calculus class.)

Definition 1.23. The function $f : U \to \mathbb{C}$ is *continuous* if for every $z_0 \in U$ we have $\lim_{z \to z_0} f(z) = f(z_0)$.

Definition 1.24. The function $f : U \to \mathbb{C}$ is *differentiable* at $z_0 \in U$ if

$$f'(z_0) := \lim_{h \to 0} \frac{f(z_0 + h) - f(z_0)}{h}$$

exists.

Remark. The usual rules for differentiating sums, products, and quotients, as well as the chain rule, hold for complex derivatives. They are proved in the same way as in your calculus class.

Remark. It is crucial in Definition 1.24 that one allows h to approach 0 from any direction and that the resulting limit is independent of that direction.

Now for the most important definition in this whole set of notes:

Definition 1.25. The function $f : U \to \mathbb{C}$ is *analytic* (or *holomorphic*) if it is differentiable at every $z_0 \in U$.

We will see that analytic functions have marvelous properties! It will be the reason why studying the iteration of analytic functions is so fruitful.

Exercise 1.26. Show that $f(z) = z$ is analytic on all of $\mathbb{C}$ and that $g(z) = \overline{z}$ is not analytic in a neighborhood of any point of $\mathbb{C}$. (In fact, it is "anti-analytic.")

Exercise 1.27. Show that any complex polynomial

$$p(z) = a_d z^d + a_{d-1} z^{d-1} + \cdots + a_1 z + a_0$$

gives an analytic function $p : \mathbb{C} \to \mathbb{C}$.

Definition 1.28. Suppose U and V are domains. A mapping $f : U \to V$ is called *conformal* if it is analytic and has an analytic inverse $f^{-1} : V \to U$.

Cauchy–Riemann equations. *Let $f : U \to \mathbb{C}$ be given by*

$$f(z) = u(x, y) + iv(x, y)$$

with $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$, and $\frac{\partial v}{\partial y}$ continuous on U. Then

$$f \text{ is analytic on } U \quad \Leftrightarrow \quad \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \quad \text{for all } (x, y) \in U.$$

Inverse function theorem. *Suppose $f : U \to \mathbb{C}$ is analytic and $f'(z_0) \neq 0$. Then, there is an open neighborhood V of $f(z_0)$ in $\mathbb{C}$ and an analytic function $g : V \to U$ such that $g(f(z_0)) = z_0$ and for all $w \in V$ we have $f(g(w)) = w$ and all $z \in g(V)$ we have $g(f(z)) = z$. Moreover,*

$$g'(f(z_0)) = \frac{1}{f'(z_0)}.$$

Exercise 1.29. Show that $f(z) = z^2 - 1$ satisfies the hypotheses of the inverse function theorem for any $z \neq 0$. Use the quadratic equation to explicitly find the function $g(z)$ whose existence is asserted by the inverse function theorem. What goes wrong with g at $-1 = f(0)$?

Exponential and logarithm. According to Euler's formula, if $z = x + iy$ with $x, y \in \mathbb{R}$ then

$$\mathrm{e}^z = \mathrm{e}^{x+iy} = \mathrm{e}^x \left(\cos y + i \sin y\right),$$

which can be verified to be analytic on all of $\mathbb{C}$ by using the Cauchy–Riemann equations. It satisfies $(\mathrm{e}^z)' = \mathrm{e}^z$, which is never 0.

Let $S := \{z \in \mathbb{C} \, : \, -\pi < \mathrm{Im}(z) < \pi\}$ and $\mathbb{C}^\dagger := \mathbb{C} \setminus (-\infty, 0]$. Then, the exponential function maps the strip S bijectively onto $\mathbb{C}^\dagger$. Therefore, it has an inverse function

$$\mathrm{Log}(z) : \mathbb{C}^\dagger \to S,$$

which is analytic by the inverse function theorem. (This function is called the *principal branch of the logarithm*. One can define other branches that are analytic on domains other than $\mathbb{C}^\dagger$; see [42, Section 3.3].)

Definition 1.30. Suppose $f : U \to \mathbb{C}$ is an analytic function. A point $z \in U$ with $f'(z) = 0$ is called a *critical point* of f. A point $w \in \mathbb{C}$ with $w = f(z)$ for some critical point z is called a *critical value*.

The neighborhood V provided by the inverse function theorem could be very small. When combined with the monodromy theorem [1, pp. 295–297], one can control the size of the domain, so long as it is simply connected:

Simply connected inverse function theorem. *Suppose $f : U \to \mathbb{C}$ is an analytic function and $V \subset f(U)$ is a simply connected domain that doesn't contain any of the critical values of f.*

Given any $w_\bullet \in V$ and any $z_\bullet \in f^{-1}(w_\bullet)$, there is a unique analytic function $g : V \to \mathbb{C}$ with $g(w_\bullet) = z_\bullet$, $f(g(w)) = w$ for all $w \in V$, and $g(f(z)) = z$ for all $z \in g(V)$.

Remark. Our name for the previous result is not standard. Use it with caution!

If $f : U \to \mathbb{C}$ is continuous and $\gamma \subset U$ is a contour, then the integral

$$\int_\gamma f(z)\, dz$$

is defined in terms of a suitable complex version of Riemann sums; see [42, Section 4.2]. For our purposes, we can take as a definition

$$\int_\gamma f(z)\, dz := \int_0^1 f(\gamma(t))\gamma'(t)\, dt,$$

which is stated as [42, Theorem 4, Section 4.2].

Exercise 1.31. Let γ be the positively oriented unit circle in $\mathbb{C}$. Show that

$$\int_\gamma \frac{dz}{z} = 2\pi i, \tag{1.4}$$

which is perhaps "the most important contour integral."

Cauchy's theorem. *If $f : U \to \mathbb{C}$ is analytic and U is simply connected, then for any closed contour $\gamma \subset U$ we have*

$$\int_\gamma f(z)\, dz = 0.$$

Sketch of proof. The following is "cribbed" directly from [42, pp. 192–193]. Write

$$f(z) = u(x, y) + iv(x, y) \qquad \text{and} \qquad \gamma(t) = (x(t), y(t)).$$

Then,

$$\begin{aligned}
\int_\gamma f(z)\, dz &= \int_0^1 f(\gamma(t))\gamma'(t)\, dt \\
&= \int_0^1 \Big(u(x(t), y(t)) + iv(x(t), y(t))\Big)\left(\frac{dx}{dt} + i\frac{dy}{dt}\right) dt \\
&= \int_0^1 \left(u(x(t), y(t))\frac{dx}{dt} - v(x(t), y(t))\frac{dy}{dt}\right) dt \\
&\quad + i\int_0^1 \left(v(x(t), y(t))\frac{dx}{dt} + u(x(t), y(t))\frac{dy}{dt}\right) dt.
\end{aligned}$$

The real and imaginary parts above are just the parameterized versions of the real contour integrals

$$\int_\gamma u(x, y)\, dx - v(x, y)\, dy \qquad \text{and} \qquad \int_\gamma v(x, y)\, dx + u(x, y)\, dy$$

considered in a multivariable calculus class. Since U is simply connected, Green's theorem [46] gives

$$\int_\gamma u(x,y)\,dx - v(x,y)\,dy = \int\int_D \left(-\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) dx\,dy \quad \text{and}$$
$$\int_\gamma v(x,y)\,dx + u(x,y)\,dy = \int\int_D \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) dx\,dy.$$

Since f is analytic, the Cauchy–Riemann equations imply that both integrands are 0. Thus, $\int_\gamma f(z)\,dz = 0$. □

Remark. In the proof we have used the additional assumption that the partial derivatives $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$, and $\frac{\partial v}{\partial y}$ are all continuous functions of (x,y). This was needed in order for us to apply Green's theorem. This hypothesis is not needed, but the general proof of Cauchy's theorem is more complicated; see, for example, [1, Section 4.4].

There is also an amazing "converse" to Cauchy's theorem:

Morera's theorem. *If $f : U \to \mathbb{C}$ is continuous and if*

$$\int_\gamma f(z)\,dz = 0$$

for any closed contour $\gamma \subset U$, then f is analytic in U.

Cauchy integral formula. *Let γ be a simple closed positively oriented contour. If f is analytic in some simply connected domain U containing γ, and z_0 is any point inside γ, then*

$$f(z_0) = \frac{1}{2\pi i}\int_\gamma \frac{f(z)}{z - z_0}\,dz.$$

Sketch of proof. Refer to Figure 3.8 throughout the proof. For any $\epsilon > 0$ we can apply Cauchy's theorem to the contour η proving that

$$\int_\gamma \frac{f(z)}{z - z_0}\,dz = \int_{\gamma'} \frac{f(z)}{z - z_0}\,dz,$$

where γ' is the positively oriented circle $|z - z_0| = \epsilon$. Since $f(z)$ is analytic it is continuous, implying that if we choose $\epsilon > 0$ sufficiently small, $f(z) \approx f(z_0)$ on γ'. Then,

$$\int_{\gamma'} \frac{f(z)}{z - z_0}\,dz \approx f(z_0)\int_{\gamma'} \frac{1}{z - z_0}\,dz = 2\pi i f(z_0),$$

with the last equality coming from (1.4). □

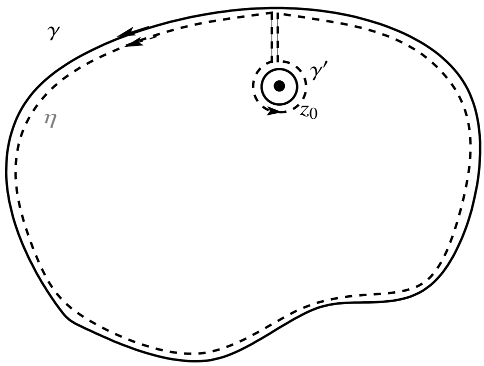


Figure 3.8. Illustration of the proof of the Cauchy integral formula.

Exercise 1.32. Use the fact that if $|f(z) - g(z)| < \epsilon$ for all z on a contour γ then

$$\left| \int_\gamma f(z)\, dz - \int_\gamma g(z)\, dz \right| < \epsilon \text{ length}(\gamma),$$

to make rigorous the estimates $\approx$ in the proof of the Cauchy integral formula.

Let us write the Cauchy integral formula slightly differently:

$$f(z) = \frac{1}{2\pi i} \int_\gamma \frac{f(\zeta)}{\zeta - z} d\zeta, \tag{1.5}$$

where z is any point inside γ. (This makes it more clear that we think of z as an independent variable.) By differentiating under the integral sign (after checking that it's allowed) we obtain the next result.

Cauchy integral formula for higher derivatives. *Let γ be a simple closed positively oriented contour. If f is analytic in some simply connected domain U containing γ and z is any point inside γ, then*

$$f^{(n)}(z) = \frac{n!}{2\pi i} \int_\gamma \frac{f(\zeta)}{(\zeta - z)^{n+1}}\, dz. \tag{1.6}$$

In particular, an analytic function is infinitely differentiable!

Cauchy estimates 1.33. *Suppose $f(z)$ is analytic on a domain containing the disk $D(z_0, r)$ and suppose $|f(z)| < M$ on the boundary $\partial D(z_0, r) = \{z \in \mathbb{C} : |z - z_0| = r\}$. Then, for any $n \in \mathbb{N}$ we have*

$$\left| f^{(n)}(z_0) \right| \leq \frac{n!M}{r^n}.$$

Exercise 1.34. Prove the Cauchy estimates, supposing (1.6).

Suppose $\overline{D(z_0, r)} \subset U$ and $f : U \to \mathbb{C}$ is analytic. If we parameterize $\partial D(0, r)$ by $\gamma(t) = z_0 + r\mathrm{e}^{it}$, then the Cauchy integral formula becomes

$$f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + r\mathrm{e}^{it})\, dt.$$

From this, one sees that it is impossible to have $|f(z_0)| \geq |f(z_0 + r\mathrm{e}^{it})|$ for all $t \in [0, 2\pi]$ without the inequality actually being an equality for all t. From this, it is straightforward to prove the following result.

Maximum modulus principle. *Suppose $f(z)$ is analytic in a domain U and $|f(z)|$ achieves its maximum at a point $z_0 \in U$. Then $f(z)$ is constant on U.*

If, moreover, $\overline{U}$ is compact and f extends continuously to $\overline{U}$, then f achieves its maximum modulus on the boundary of U.

Meanwhile, by using the geometric series to write

$$\frac{1}{\zeta - z} = \frac{1}{\zeta} \cdot \frac{1}{1 - \frac{z}{\zeta}} = \frac{1}{\zeta} \sum_{n=0}^{\infty} \left(\frac{z}{\zeta}\right)^n,$$

for any $\left|\frac{z}{\zeta}\right| < 1$, the Cauchy integral formula (1.5) implies

Existence of power series. *Let f be analytic on a domain U and suppose the disk $D(z_0, r)$ is contained in U. Then, we can write $f(z)$ as a power series*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

that converges on $D(z_0, r)$.

The *multiplicity* of a zero z_0 for an analytic function $f(z)$ is defined as the order of the smallest nonzero term in the power series expansion of $f(z)$ around z_0.

Argument principle. *Suppose $f : U \to \mathbb{C}$ is analytic and $\gamma \subset U$ is a positively oriented simple closed contour such that all points inside γ are in U. Then, the number of zeros of f (counted with multiplicities) is equal to the change in* $\arg(f(z))$ *as z traverses γ once in the counterclockwise direction.*

Definition 1.35. Let $f_n : U \to \mathbb{C}$ be a sequence of functions and $f : U \to \mathbb{C}$ be another function. Let $K \subset U$ be a compact set. The sequence $\{f_n\}$ converges to f *uniformly on K* if for every $\epsilon > 0$ there is a $\delta > 0$ such that for every $z \in K$ $|f_n(z) - f(z)| < \epsilon$.

Note that the order of quantifiers in Definition 1.35 is crucial. If δ was allowed to depend on z, we would have the weaker notion of *pointwise convergence*.

Uniform limits theorem. *Suppose $f_n : U \to \mathbb{C}$ is a sequence of analytic functions and $f : U \to \mathbb{C}$ is another (potentially nonanalytic) function. If for any compact $K \subset U$ we have that $\{f_n\}$ converges uniformly to f on K, then $f : U \to \mathbb{C}$ is also analytic.*

Moreover, for any $k \geq 1$, the kth derivatives $f_n^{(k)}(z)$ converge uniformly to $f^{(k)}(z)$ on any compact $K \subset U$.

Sketch of the proof. By restricting to a smaller domain, we can suppose U is simply connected. For any contour $\gamma \subset U$, Cauchy's theorem gives $\int_\gamma f_n(z)\,dz = 0$. Since the convergence is uniform on the compact set $\gamma \subset U$, we have

$$\int_\gamma f(z)\,dz = \int_\gamma \lim_{n\to\infty} f_n(z)\,dz = \lim_{n\to\infty} \int_\gamma f_n(z)\,dz = 0.$$

Thus, Morera's theorem gives that $f(z)$ is analytic.

Convergence of the derivatives follows from the Cauchy integral formula for higher derivatives. □

The following exercises illustrate the power of the uniform limits theorem.

Exercise 1.36. Suppose that for some $R > 0$ the power series

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n \tag{1.7}$$

converges for each $z \in D(z_0, R)$. Prove that for any $0 < r < R$ the power series converges uniformly on the closed disk $\overline{D(z_0, r)}$. Use Exercise 1.27 and the uniform limits theorem to conclude that power series (1.7) defines an analytic function $f : D(z_0, R) \to \mathbb{C}$.

Exercise 1.37. Suppose we have a sequence of polynomials $p_n : [0, 1] \to \mathbb{R}$ and that $p_n(x)$ converges uniformly on $[0, 1]$ to some function $f : [0, 1] \to \mathbb{R}$. Does f even have to be differentiable?

We close this section with the following famous result:

Schwarz lemma. *Let $\mathbb{D} := D(0, 1)$ be the unit disk and suppose $f : \mathbb{D} \to \mathbb{D}$ is analytic with $f(0) = 0$. Then*

(a) $|f'(0)| \leq 1$, *and*

(b) $|f'(0)| = 1$ *if and only if* $f(z) = e^{i\theta} z$ *for some* $\theta \in \mathbb{R}$.

Sketch of the proof. By the existence of power series theorem we can write f as a power series converging on $\mathbb{D}$:

$$f(z) = a_1 z + a_2 z^2 + a_3 z^3 + \cdots,$$

where the constant term is 0 because $f(0) = 0$. Therefore,

$$F(z) := \frac{f(z)}{z} = a_1 + a_2 z + a_3 z^3 + \cdots$$

is also analytic on $\mathbb{D}$, by Exercise 1.36. Applying the maximum modulus principle to $F(z)$ we see that for any $0 < r < 1$ and any ζ satisfying $|\zeta| < r$,

$$|F(\zeta)| \leq \frac{\max_{\{|z|=r\}} |f(z)|}{r} \leq \frac{1}{r}.$$

Since this holds for any $0 \leq r \leq 1$, we find that $|F(\zeta)| \leq 1$ for any $\zeta \in \mathbb{D}$. Part (a) follows because $F(0) = f'(0)$.

If $|f'(0)| = 1$, then $|F(0)| = 1$, implying that F attains its maximum at a point of $\mathbb{D}$. The maximum modulus principle implies that $F(z)$ is constant, i.e., $F(z) = c$ for some c with $|c| = 1$. Any such c is of the form $\mathrm{e}^{i\theta}$ for some $\theta \in \mathbb{R}$, so by the definition of F, we have $f(z) = \mathrm{e}^{i\theta} z$ for all $z \in \mathbb{D}$. □

Remark. There was nothing special about radius 1. If $f : D(0, r) \to D(0, r)$ for some $r > 0$ and $f(0) = 0$, then (a) and (b) still hold.

Lecture 2. Mandelbrot set from the inside out

We will work our way to the famous Mandelbrot set from an unusual perspective.

2.1 Attracting periodic orbits. In Section 1.3 we saw that the quadratic maps $p_c(z) = z^2 + c$ for $c = \frac{i}{4}, -1$, and $-0.1 + 0.75i$ seemed to have attracting periodic orbits of periods 1, 2, and 3, respectively. In this subsection we will make that notion precise and prove two results about attracting periodic orbits. We will also see that the set of initial conditions whose orbits converge to an attracting periodic orbit can be phenomenally complicated.

While we are primarily interested in iterating quadratic polynomials $p_c(z) = z^2 + c$, it will also be helpful to consider iteration of higher-degree polynomials $q(z)$.

Definition 2.1. A sequence

$$z_0 \xrightarrow{q} z_1 \xrightarrow{q} z_2 \xrightarrow{q} \cdots \xrightarrow{q} z_m = z_0$$

is called a *periodic orbit of period m* for q if $z_n \neq z_0$ for each $1 \leq n \leq m-1$. The members of such a periodic orbit for q are called *periodic points* of period m for q. A periodic point of period 1 is called a *fixed point* of q.

If z_0 is a periodic point of period m for q, then it is a fixed point for the polynomial $s(z) = q^{\circ m}(z)$. Meanwhile, if z_0 is a fixed point for $s(z)$, then it is a periodic point of period j for q, where j divides m. Thus, we can often reduce the study of periodic points to that of fixed points.

Definition 2.2. A fixed point z_* of q is called *attracting* if there is some $r > 0$ such that such $q\big(D(z_*, r)\big) \subset D(z_*, r)$ and for any initial condition $z_0 \in D(z_*, r)$ the orbit $\{z_n\}$ under q satisfies $\lim z_n = z_*$.

A periodic orbit $z_0 \to z_1 \to \cdots \to z_m = z_0$ is *attracting* if for each $n = 0, \ldots, m-1$ the point z_n is an attracting fixed point for $s(z) = q^{\circ m}(z)$.

Definition 2.3. The *multiplier* of a periodic orbit $z_0 \to z_1 \to \cdots \to z_m = z_0$ is

$$\lambda = q'(z_0) \cdot q'(z_1) \cdot \cdots \cdot q'(z_{m-1}).$$

Note that if $s(z) = q^{\circ m}(z)$, then the chain rule gives that

$$s'(z_j) = q'(z_0) \cdot q'(z_1) \cdot \cdots \cdot q'(z_{m-1}) = \lambda \qquad \text{for each } 0 \leq j \leq m-1.$$

Thus the multiplier of the periodic orbit $z_0 \to z_1 \to \cdots \to z_m = z_0$ under q is the same as the multiplier of each point z_j, when considered as a fixed point of $s(z)$.

The next lemma tells us that the same criterion we had in Section 1.3 for 0 being attracting under a linear map applies to fixed points of nonlinear maps.

Attracting periodic orbit lemma. *A periodic orbit*

$$z_0 \to z_1 \to \cdots \to z_m = z_0$$

of q is attracting if and only if its multiplier satisfies $|\lambda| < 1$.

Proof. Replacing q by a suitable iterate we can suppose the periodic orbit is a fixed point z_* of q. If $z_* \neq 0$ then we can consider the new polynomial $q(z + z_*) - z_*$ for which 0 replaces z_* as the fixed point of interest. (We call this a shift of coordinates.)

Suppose 0 is an attracting fixed point for q. Then, there exists $r > 0$ so that $q\big(D(0, r)\big) \subset D(0, r)$ and so that the orbit $\{z_n\}$ of any initial condition $z_0 \in D(0, r)$ satisfies $\lim_{n\to\infty} z_n = 0$. Since $q(0) = 0$, the Schwarz lemma implies that $|q'(0)| \leq 1$. If $|q'(0)| = 1$, then the Schwarz lemma implies that q is a rigid rotation $z \mapsto \mathrm{e}^{i\theta} z$. This would violate that the orbit of any initial condition $z_0 \in D(0, r)$ converges to 0. Therefore, $|q'(0)| < 1$.

Now, suppose 0 is a fixed point for q with multiplier $\lambda = q'(0)$ of modulus less than 1. We will consider the case $\lambda \neq 0$, leaving the case $\lambda = 0$ as Exercise 2.6, below. We have

$$q(z) = \lambda z + a_2 z^2 + \cdots + a_d z^d = \lambda \left(1 + \frac{a_2}{\lambda} z + \cdots + \frac{a_d}{\lambda} z^{d-1}\right) z.$$

Since $\lim_{z \to 0}(1 + \frac{a_2}{\lambda} z + \cdots + \frac{a_d}{\lambda} z^{d-1}) = 1$ and $|\lambda| < 1$ there exists $\epsilon > 0$ so that if $|z| < \epsilon$ then

$$\left|1 + \frac{a_2}{\lambda} z + \cdots + \frac{a_d}{\lambda} z^{d-1}\right| < 1 + \frac{1 - |\lambda|}{2|\lambda|}.$$

Thus, for any $|z| < \epsilon$ we have

$$|q(z)| = \left|\lambda \left(1 + \frac{a_2}{\lambda} z + \cdots + \frac{a_d}{\lambda} z^{d-1}\right)\right| |z| \leq \frac{1 + |\lambda|}{2} |z|. \tag{2.1}$$

In particular, $q\left(D(0, r)\right) \subset D(0, r)$ and (2.1) implies that for any $z_0 \in D(0, r)$ the orbit satisfies $|z_n| \leq \left(\frac{1+|\lambda|}{2}\right)^n r \to 0$. We conclude that 0 is an attracting fixed point for q. □

Exercise 2.4. Use the attracting periodic orbit lemma to verify that

(a) $z_* = \frac{1}{2} - \frac{\sqrt{1-i}}{2}$ is an attracting fixed point for $p(z) = z^2 + \frac{i}{4}$;

(b) $0 \leftrightarrow 1$ is an attracting periodic orbit of period 2 for $p(z) = z^2 - 1$; and

(c) if c satisfies $c^3 + 2c^2 + c + 1 = 0$, then $0 \to c \to c^2 + c \to 0$ is an attracting periodic orbit of period 3. (One of the solutions for c is the parameter $c \approx -0.12 + 0.75i$ studied in Exercise 1.12.)

Exercise 2.5. Verify that there exists $r > 0$ such that for any initial condition $z_0 \in \mathbb{R}$ with $|z_0| < r$ the orbit under $q(z) = z - z^3$ converges to 0. Why is 0 not attracting as a complex fixed point?

Exercise 2.6. Prove that if z_* is a fixed point for a polynomial q having multiplier $\lambda = 0$, then z_* is attracting.

Definition 2.7. Suppose $\mathcal{O} = z_0 \to z_1 \to \cdots \to z_m = z_0$ is an attracting periodic orbit. The *basin of attraction* $\mathcal{A}(\mathcal{O})$ is

$$\mathcal{A}(\mathcal{O}) := \{z \in \mathbb{C} \, : \, s^{\circ n}(z) \to z_j \text{ as } n \to \infty \text{ for some } 0 \leq j \leq m - 1\},$$

where $s(z) = q^{\circ m}(z)$. The *immediate basin* $\mathcal{A}_0(\mathcal{O})$ is the union of the connected components of $\mathcal{A}(\mathcal{O})$ containing the points $z_0, \ldots, z_{m-1}$.

Computer generated images of the basins of attraction for the attracting periodic orbits discussed in Exercise 2.4 are shown in Figures 3.9–3.11. Notice the remarkable complexity of the boundaries of the basins of attraction, something we would never have guessed during our experimentation in Section 1.3.

Remark on computer graphics. We used Fractalstream [19] to create Figures 3.9–3.12, 3.16–3.20, 3.22, and 3.23. Other useful programs include Dynamics Explorer [16] and the Boston University Java Applets [9].

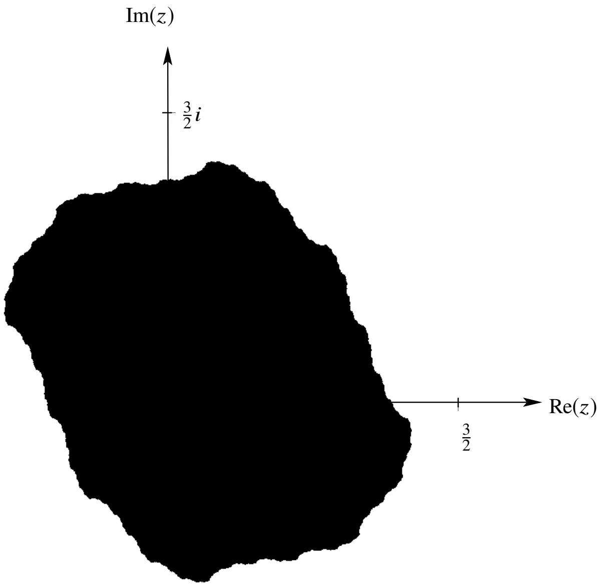


Figure 3.9. Basin of attraction of the fixed point $z_* = \frac{1}{2} - \frac{\sqrt{1-i}}{2}$ for $p(z) = z^2 + \frac{i}{4}$.

It is natural to ask how complicated the dynamics for iteration of q can be near an attracting fixed point. The answer is provided by Kœnig's theorem and Böttcher's theorem.

Kœnig's theorem. *Suppose $z_\bullet$ is an attracting fixed point for q with multiplier $\lambda \neq 0$. Then, there exists a neighborhood U of $z_\bullet$ and a conformal map*

$$\phi : U \to \phi(U) \subset \mathbb{C}$$

so that for any $w \in \phi(U)$ we have

$$\phi \circ q \circ \phi^{-1}(w) = \lambda w. \tag{2.2}$$

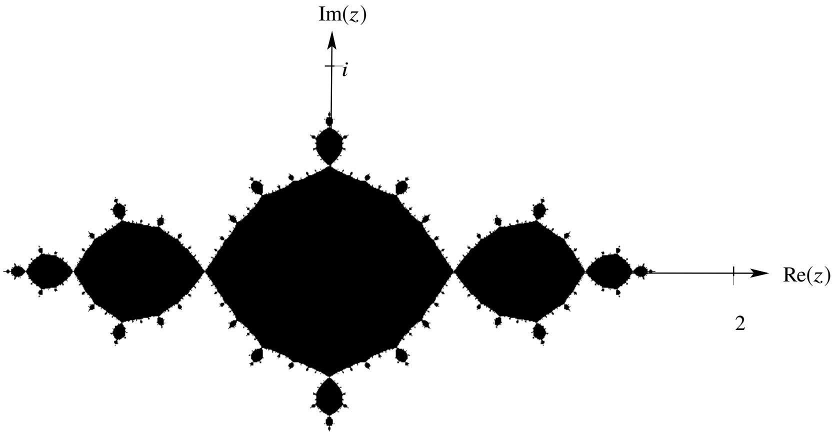


Figure 3.10. Basin of attraction of the period-2 cycle $0 \leftrightarrow -1$ for $p(z) = z^2 - 1$.

In other words, Theorem 2.1 gives that there is a neighborhood U of $z_\bullet$ in which there is a coordinate system $w = \phi(z)$ in which the nonlinear mapping q becomes linear! This explains why the same geometric spirals shown on the top of Figure 3.3 for the linear map appear sufficiently close to an attracting fixed point $z_\bullet$ for a nonlinear map. This is illustrated in Figure 3.12.

Proof. Shifting the coordinates if necessary, we can suppose $z_\bullet = 0$. The attracting periodic orbit lemma gives that the multiplier of 0 satisfies $|\lambda| < 1$. Therefore, as in the second half of the proof of the attracting periodic orbit lemma, we can find some $r > 0$ and $|\lambda| < a < 1$ so that

$$\text{for any} \quad z \in D(0,r) \quad \text{we have} \quad |z_n| \leq a^n r, \tag{2.3}$$

where $z_n := q^{\circ n}(z)$.

Since $q(0) = 0$ we have

$$q(z) = \lambda z + s(z), \tag{2.4}$$

where $s(z) = a_2 z^2 + a_3 z^3 + \cdots + a_d z^d$. In particular, there exists $b > 0$ so that

$$|s(z)| \leq b|z|^2, \tag{2.5}$$

for all $z \in D(0,r)$.

Let

$$\phi_n : D(0,r) \to \mathbb{C} \qquad \text{be given by} \qquad \phi_n(z) := \frac{z_n}{\lambda^n},$$

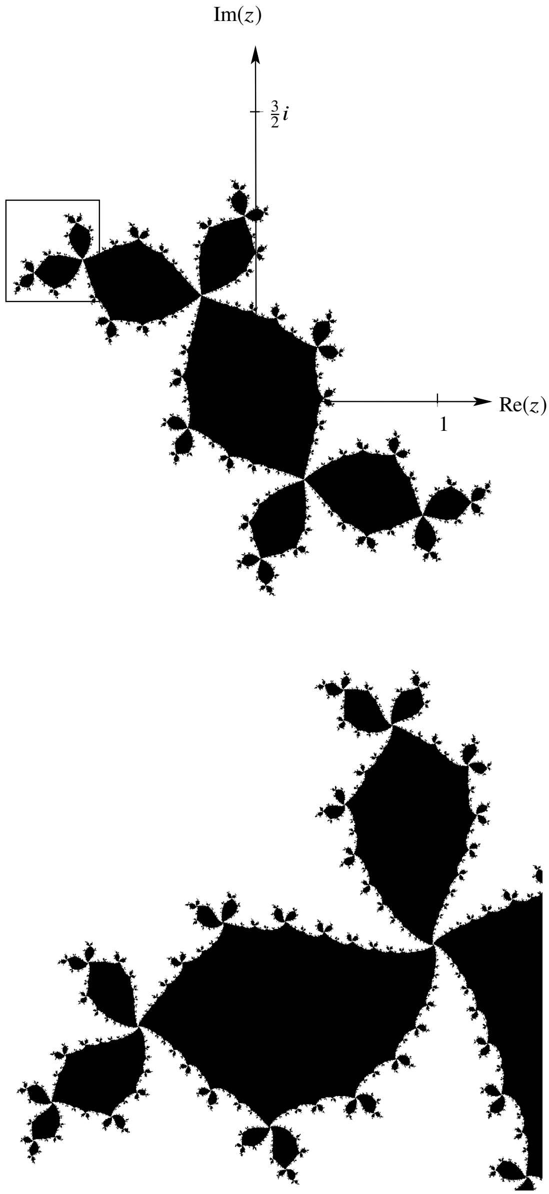


Figure 3.11. Top: Basin of attraction of the attracting period-3 cycle $0 \to c \to c^2 + c$ for $c \approx -0.12 + 0.75i$. Bottom: Zoomed-in view of the boxed region from the left.

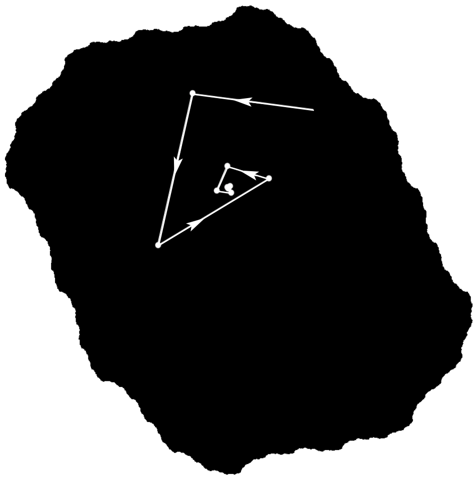


Figure 3.12. An orbit converging to the attracting fixed point for $p(z) = z^2 + \frac{i}{4}$. Here, $\lambda = 1 - \sqrt{1 - 2\,i} \approx 0.8e^{1.9i}$.

which satisfies $\phi_n(0) = 0$, since 0 is a fixed point. Notice that

$$\phi_n(q(z)) = \frac{z_{n+1}}{\lambda^n} = \lambda \cdot \frac{z_{n+1}}{\lambda^{n+1}} = \lambda\phi_{n+1}(z). \tag{2.6}$$

Suppose we can prove that ϕ_n converges uniformly on $D(0, r)$ to some function $\phi : D(0, r) \to \mathbb{C}$. Then, ϕ will be analytic by the uniform limits theorem. Meanwhile, the left and right sides of (2.6) converge to $\phi(q(z))$ and $\lambda\phi(z)$, respectively, implying

$$\phi(q(z)) = \lambda\phi(z). \tag{2.7}$$

Since $\phi_n(0) = 0$ for each ϕ_n we will also have $\phi(0) = 0$.

To see that ϕ_n converges uniformly on $D(0, r)$, let us rewrite it as

$$\phi_n(z) = \frac{z_n}{\lambda^n} = z_0 \cdot \frac{z_1}{\lambda z_0} \cdot \frac{z_2}{\lambda z_1} \cdot \frac{z_3}{\lambda z_2} \cdot \cdots \cdot \frac{z_n}{\lambda z_{n-1}}.$$

By (2.4), the general term of the product becomes

$$\frac{z_k}{\lambda z_{k-1}} = \frac{q(z_{k-1})}{\lambda z_{k-1}} = \frac{\lambda z_{k-1} + s(z_{k-1})}{\lambda z_{k-1}} = 1 + \frac{s(z_{k-1})}{\lambda z_{k-1}}.$$

By the estimates (2.5) and (2.3) on $|z_n|$ we have

$$\left|\frac{s(z_{k-1})}{\lambda z_{k-1}}\right| \leq b\frac{|z_k|}{\lambda} \leq b\frac{a^k r}{\lambda}. \tag{2.8}$$

We will now make r smaller, if necessary, to ensure that the right-hand side of (2.8) is less than $\frac{1}{2}$.

To show that the ϕ_n converge uniformly on $D(0,r)$, it is sufficient to show that the infinite product

$$\frac{z_1}{\lambda z_0} \cdot \frac{z_2}{\lambda z_1} \cdot \frac{z_3}{\lambda z_2} \cdot \cdots \cdot \frac{z_n}{\lambda z_{n-1}} \cdots$$

does. Such a product converges if and only if logarithms of the finite partial products converge, i.e., if and only if the infinite sum

$$\operatorname{Log} \phi(z) = \sum_{k=1}^{\infty} \operatorname{Log}\left(1 + \frac{s(z_{k-1})}{\lambda z_{k-1}}\right) \tag{2.9}$$

converges. (We can take the logarithms on the right-hand side of (2.9) because our bound of (2.8) by $\frac{1}{2}$ implies that $1 + \frac{s(z_{k-1})}{\lambda z_{k-1}} \in \mathbb{C} \setminus (-\infty, 0]$.) Using the estimate

$$|\operatorname{Log}(1+w)| \leq 2|w| \qquad \text{for any} \quad |w| < \tfrac{1}{2}$$

and (2.8) we see that the kth term is geometrically small:

$$\left|\operatorname{Log}\left(1 + \frac{s(z_{k-1})}{\lambda z_{k-1}}\right)\right| \leq 2b\frac{a^k r}{\lambda}.$$

This proves convergence of (2.9)

It remains to show that ϕ is conformal when restricted to a small enough neighborhood $U \subset D(0,r)$ of 0. By the chain rule, each ϕ_n satisfies $\phi_n'(0) = 1$. Since the ϕ_n converge uniformly to ϕ in a neighborhood of 0, the Cauchy integral formula (1.6) implies that $\phi_n'(0) \to \phi'(0)$. Thus $\phi'(0) = 1$.

By the inverse function theorem, there is a neighborhood V of $0 = \phi(0)$ and an analytic function $g : V \to D(0,r)$ so that $\phi(g(w)) = w$ for every $w \in V$. If we let $U = g(V)$, then $\phi : U \to V$ is conformal.

To obtain (2.2), precompose (2.7) with $\phi^{-1} = g$ on V. □

Extended exercise 2.8. Adapt the proof of Kœnig's theorem to prove the following:

Böttcher's theorem. *Suppose $p(z)$ has a fixed point $z_\bullet$ of multiplier $\lambda = 0$ and thus is of the form*

$$p(z) = z_\bullet + a_k(z - z_\bullet)^k + a_{k+1}(z - z_\bullet)^{k+1} + \cdots + a_d(z - z_\bullet)^d$$

for some $2 \leq k < d$*. Then, there exists a neighborhood* U *of* $z_\bullet$ *and a conformal map*

$$\phi : U \to \phi(U) \subset \mathbb{C}$$

so that for any $w \in \phi(U)$ *we have* $\phi \circ p \circ \phi^{-1}(w) = w^k$.

2.2 First exploration of the parameter space: The set M_0. Let us try to understand the space of parameters $c \in \mathbb{C}$ for the quadratic polynomial maps $p_c(z) = z^2 + c$. Consider

$$M_0 := \{c \in \mathbb{C} \,:\, p_c(z) \text{ has an attracting periodic orbit}\}.$$

We have already seen in Section 1.3 that $c = \frac{i}{4}$ and $c = -1$ are in M_0 and that $c = \frac{1}{2}$ is probably not in M_0. We will now use the attracting periodic orbit lemma to find some regions that are in M_0.

The fixed points of $p_c(z) = z^2 + c$ are

$$z_* = \frac{1}{2} + \frac{\sqrt{1-4c}}{2} \qquad \text{and} \qquad z_\bullet = \frac{1}{2} - \frac{\sqrt{1-4c}}{2}$$

and, since $p_c'(z) = 2z$, their multipliers are

$$\lambda_* = 1 + \sqrt{1-4c} \qquad \text{and} \qquad \lambda_\bullet = 1 - \sqrt{1-4c}.$$

If $|\lambda_*| = 1$, then

$$1 + \sqrt{1-4c} = \mathrm{e}^{i\theta}$$

for some $\theta \in \mathbb{R}$. Solving for c, we find

$$c = \frac{\mathrm{e}^{i\theta}}{2} - \frac{\mathrm{e}^{i2\theta}}{4}.$$

The resulting curve C is a “cardioid,” shown in Figure 3.13.

In each of the two regions of $\mathbb{C} \setminus C$ we choose the points $c = 0$ and $c = 1$, which result in $\lambda_* = 2$ and $\lambda_* = 1 + \sqrt{3}i$, respectively. Thus, neither of the regions from $\mathbb{C} \setminus C$ corresponds to parameters c for which z_* is an attracting fixed point. Thus, we conclude that the smallest $|\lambda_*|$ can be is 1, occurring exactly on the cardioid C.

Doing the same computations with the multiplier $\lambda_\bullet$ of the second fixed point $z_\bullet$, we also find that $|\lambda_\bullet| = 1$ if and only if c is on the cardioid C. However, at $c = 0$ we have $\lambda_\bullet = 0$ and at $c = 1$ we have $\lambda_\bullet = 1 - \sqrt{3}i$, which is of modulus greater than 1. Therefore, according to the attracting periodic orbit lemma, fixed point $z_\bullet$ is attracting if and only if c is inside the cardioid C. We summarize the past three paragraphs with the following lemma.

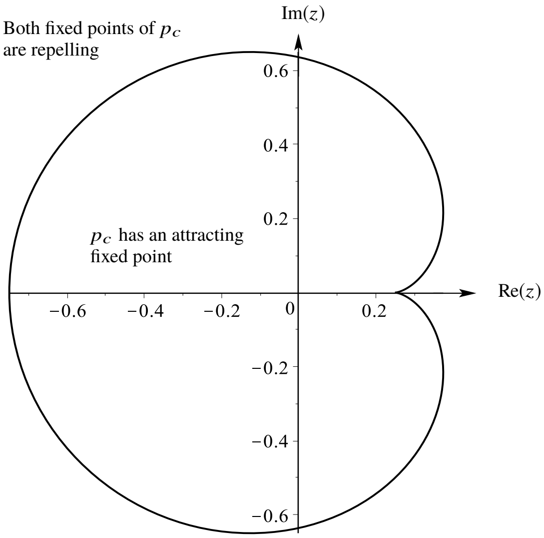


Figure 3.13. $p_c(z) = z^2 + c$ has an attracting fixed point if and only if c lies inside the cardioid $c = \frac{e^{i\theta}}{2} - \frac{e^{i2\theta}}{4}$, where $0 \le \theta \le 2\pi$, depicted here.

Lemma 2.9. *$p_c(z) = z^2 + c$ has an attracting fixed point if and only if c lies inside the cardioid curve $C := \left\{ c = \frac{e^{i\theta}}{2} - \frac{e^{i2\theta}}{4} \ : \ 0 \le \theta \le 2\pi \right\}$.*

To find periodic orbits of period 2, we solve $p_c^{\circ 2}(z) = \left(z^2 + c\right)^2 + c = z$. In addition to the two fixed points z_* and $z_\bullet$, we find

$$z_0 = -\frac{1}{2} + \frac{\sqrt{-3 - 4\,c}}{2} \qquad \text{and} \qquad z_1 = -\frac{1}{2} - \frac{\sqrt{-3 - 4\,c}}{2}.$$

One can check that $p_c(z_0) = z_1$ and $p_c(z_1) = z_0$. These points are equal if $c = -\frac{3}{4}$; otherwise, they are indeed a periodic orbit of period 2.

The multiplier of this periodic orbit is

$$\lambda = \left(-1 + \sqrt{-3 - 4\,c}\right)\left(-1 - \sqrt{-3 - 4\,c}\right) = 4 + 4c,$$

which has modulus 1 if and only if $|c + 1| = \frac{1}{4}$. Since $\lambda = 0$ for $c = -1$ (inside the circle) and $\lambda = 4$ for $c = 0$ (outside the circle) we find the following lemma.

Lemma 2.10. *$p_c(z) = z^2 + c$ has a periodic orbit of period 2 if and only if c lies inside the circle $|c + 1| = \frac{1}{4}$.*

In Figure 3.14 we show the regions of M_0 that we have discovered.

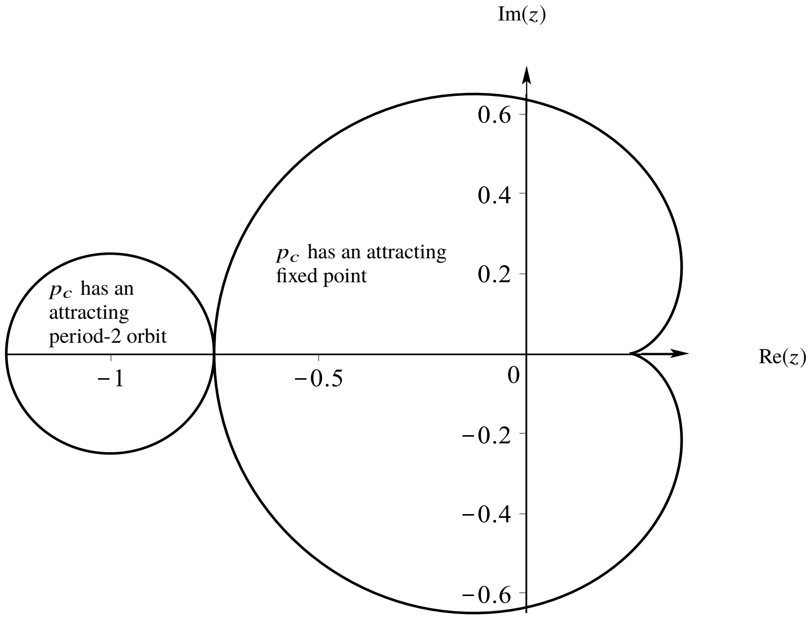


Figure 3.14. The regions in the parameter plane where $p_c(z) = z^2 + c$ has an attracting fixed point and where p_c has an attracting periodic orbit of period 2. Combined, they form a subset of M_0.

Exercise 2.11. If possible, determine the region of parameters c for which $p_c(z) = z^2 + c$ has an attracting periodic orbit of period 3.

As n increases, this approach becomes impossible. We need a different approach, which requires a deeper study of attracting periodic orbits.

2.3 Second exploration of the parameter space: The Mandelbrot set M.

Fatou–Julia lemma. *Let q be a polynomial of degree $d \geq 2$. Then, the immediate basin of attraction for any attracting periodic orbit contains at least one critical point of q. In particular, since q has $d-1$ critical points (counted with multiplicity), q can have no more than $d-1$ distinct attracting periodic orbits.*

The following proof is illustrated in Figure 3.15.

Proof. Replacing q by an iterate, we can suppose that the attracting periodic orbit is a fixed point $z_\bullet$ of q. Performing a shift of coordinates, we suppose $z_\bullet = 0$.

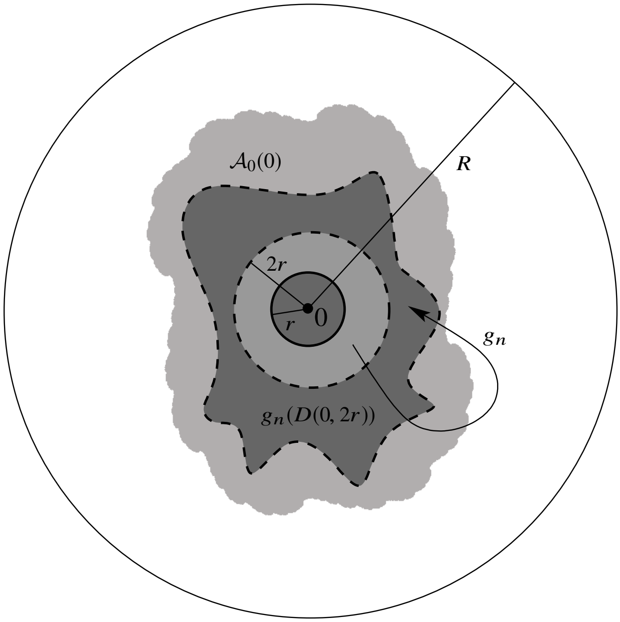


Figure 3.15. Illustration of the proof of the Fatou–Julia lemma.

If 0 has multiplier $\lambda = 0$, then 0 is already a critical point in the immediate basin $\mathcal{A}_0(0)$. We therefore suppose 0 has multiplier $\lambda \neq 0$. By the attracting periodic orbit lemma, $|\lambda| < 1$.

Suppose for contradiction that there is no critical point for q in $\mathcal{A}_0(0)$.

According to Exercise 1.13 there is some $R > 0$ so that any initial condition z_0 with $|z_0| > R$ has orbit $\{z_n\}$ that escapes to ∞. In particular,

$$\mathcal{A}_0(0) \subset D(0, R). \tag{2.10}$$

We claim that $q(\mathcal{A}_0(0)) = \mathcal{A}_0(0)$. Since $\mathcal{A}(0)$ is forward invariant, $q(\mathcal{A}_0(0)) \subset \mathcal{A}(0)$. Because $\mathcal{A}_0(0)$ is connected, so is $q(\mathcal{A}_0(0))$, which is therefore contained in one of the connected components of $\mathcal{A}(0)$. Since $0 = q(0) \in q(\mathcal{A}_0(0))$, we have $q(\mathcal{A}_0(0)) \subset \mathcal{A}_0(0)$.

Conversely, suppose $z_* \in \mathcal{A}_0(0)$. Let γ be a simple contour in $\mathcal{A}_0(0)$ connecting z_* to 0 and avoiding any critical values of q. (By hypothesis, such critical values would be images of critical points that are not in $\mathcal{A}_0(0)$.) Then, $q^{-1}(\gamma)$ is a union of several simple contours. Since $q^{-1}(0) = 0$, one of them is a simple contour ending at 0. The other end is a point $z_\#$, which is therefore in $\mathcal{A}_0(0)$. By construction, $q(z_\#) = z_*$.

To simplify notation, let $f := q|_{\mathcal{A}_0(0)} : \mathcal{A}_0(0) \to \mathcal{A}_0(0)$, which satisfies

(1) $f(\mathcal{A}_0(0)) = \mathcal{A}_0(0)$ and

(2) f has no critical points.

These properties persist under iteration, giving that $f^n(\mathcal{A}_0(0)) = \mathcal{A}_0(0)$ and f^n has no critical points for every $n \geq 1$. (The latter uses the chain rule.)

Let $r > 0$ be chosen sufficiently small so that $D(0, 2r) \subset \mathcal{A}_0(0)$. Since $D(0, 2r)$ is simply connected, the simply connected inverse function theorem gives for each $n \geq 1$ an analytic function

$$g_n : D(0, 2r) \to \mathcal{A}_0(0) \subset D(0, R)$$

with $g_n(0) = 0$ and $f^{\circ n}(g_n(w)) = w$ for all $w \in (0, 2r)$. Its derivative satisfies

$$g_n'(0) = \frac{1}{(f^n)'(0)} = \frac{1}{\lambda^n}, \tag{2.11}$$

which can be made arbitrarily large by choosing n sufficiently large, since $|\lambda| < 1$.

Meanwhile, we can apply the Cauchy estimates 1.33 to the closed disk $\overline{D(0, r)} \subset D(0, 2r)$. They assert that

$$|g_n'(0)| \leq \frac{R}{r},$$

where R is the bound on the radius of $\mathcal{A}_0(0)$ given in (2.10). This is a contradiction to (2.11). We conclude that the immediate basin $\mathcal{A}_0(0)$ contains a critical point of q. $\square$

Exercise 2.12. Use the Fatou–Julia lemma and the result of Exercise 1.10 to (finally) prove that $p(z) = z^2 + \frac{1}{2}$ does not have any attracting periodic orbit. This answers Section 1.4, question (2) in the negative.

Remark. The Fatou–Julia lemma also answers Section 1.4, question (3) by telling us that a quadratic polynomial can have at most one attracting periodic orbit.

In Section 2.2 we were interested in the set

$$M_0 := \{c \in \mathbb{C} : p_c(z) \text{ has an attracting periodic orbit}\}.$$

Using the attracting periodic orbit lemma to find regions in the complex plane for which $p_c(z) = z^2 + c$ has an attracting periodic point of period n became hopeless once n is large. The results for $n = 1$ and 2 are shown in Figure 3.14.

If we decide to lose control over what period the attracting periodic point has, the Fatou–Julia lemma gives us some very interesting information:

Corollary (Consequence of Fatou–Julia lemma). *If $p_c(z)$ has an attracting periodic orbit, then the orbit $\{p_c^{\circ n}(0)\}$ of the critical point* 0 *remains bounded.*

This motivates the definition of another set:

Definition 2.13. The *Mandelbrot set* is

$$M := \{c \in \mathbb{C} : p_c^{\circ n}(0) \text{ remains bounded for all } n \geq 0\}. \tag{2.12}$$

A computer image of the Mandelbrot set is depicted in Figure 3.16. One sees small “dots” at the left end and top and bottom of the figure. They are in M, but it is not at all clear if they are connected to the main cardioid and period-2 disk of M that are shown in Figure 3.14. If one looks closer, one sees many more such “dots.” In Lecture 3 we will use a smart coloring of $\mathbb{C} \setminus M$ to better understand this issue. We will then state a theorem of Douady and Hubbard, which clears up this mystery.

The Mandelbrot set was initially discovered around 1980, but the historical details are a bit controversial. We refer the reader to [37, Appendix G] for an unbiased account. (The reader who seeks out controversy may enjoy [23].)

The corollary to the Fatou–Julia lemma implies that $M_0 \subset M$. In other words, the Mandelbrot set is an “outer approximation” to our set M_0. The reader should compare Figure 3.16 with Figure 3.14 for a better appreciation of how much progress we’ve made!

Exercise 2.14. Prove that $M_0 \neq M$ by exhibiting a parameter c for which $p_c^{\circ n}(0)$ remains bounded but with p_c having no attracting periodic orbit.

Density of hyperbolicity conjecture. $\overline{M_0} = M$.

Although this conjecture is currently unsolved, the corresponding result for real polynomials $p_c(x) = x^2 + c$ with $x, c \in \mathbb{R}$ was proved by Lyubich [32] and by Graczyk–Świątek [21]. Both proofs use complex techniques to solve the real problem.

We have approached the definition of M from an unusual perspective, i.e., “from the inside out.” In the next section we will use the fixed point at ∞ for p_c to study M again, but “from the outside in.” It is the more traditional way of introducing M.

Lecture 3. Complex dynamics from the outside in

Definition 3.1. The *filled Julia set* of $p_c(z) = z^2 + c$ is

$$K_c := \{z \in \mathbb{C} : p_c^{\circ n}(z) \text{ remains bounded for all } n \geq 0\}.$$

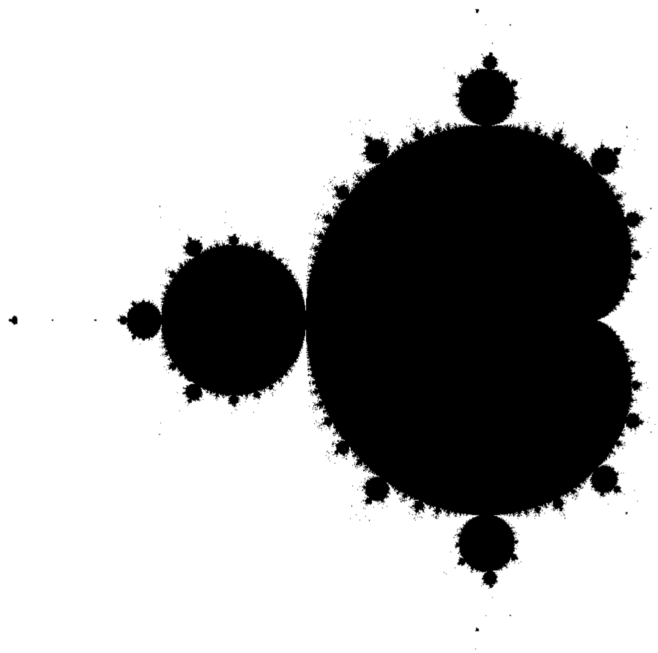


Figure 3.16. The Mandelbrot set M, shown in black. The region displayed is approximately $-2.4 \leq \mathrm{Re}(z) \leq 1$ and $-1.6 \leq \mathrm{Im}(z) \leq 1.6$.

If p_c has an attracting periodic orbit $\mathcal{O}$, then the basin of attraction $\mathcal{A}(\mathcal{O})$ is contained in K_c. However, K_c is defined for any $c \in \mathbb{C}$, even if p_c has no attracting periodic orbit in $\mathbb{C}$.

There is a natural way to extend p_c as a function:

$$p_c : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}.$$

(More formally, the space $\mathbb{C} \cup \{\infty\}$ is called the Riemann sphere; see [42, Section 1.7].) This extension satisfies $p_c(\infty) = \infty$ and, by your solution to Exercise 1.13, ∞ always has a nonempty basin of attraction:

$$\mathcal{A}(\infty) := \{z \in \mathbb{C} \,:\, p_c^{\circ n}(z) \to \infty\} = \mathbb{C} \setminus K_c.$$

Thus, ∞ is an attracting fixed point of p_c for any parameter $c \in \mathbb{C}$. In this way, the definition of K_c is always related to the basin of attraction for an attracting fixed point, even if p_c has no attracting periodic point in $\mathbb{C}$. A detailed study of $\mathcal{A}(\infty)$ will help us to prove nice theorems later in this subsection.

Remark. Note that the Fatou–Julia lemma still applies to the extended function $p_c : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\}$. If you follow through the details of how this extension is done, you find that ∞ is a critical point of p_c for every c.

Definition 3.2. The *Julia set* of $p_c(z) = z^2 + c$ is $J_c := \partial K_c$, the boundary of K_c.

Exercise 3.3. Check that for any $c \in \mathbb{C}$ the sets K_c and J_c are *totally invariant* meaning that $z \in K_c \Leftrightarrow p_c(z) \in K_c$ and $z \in J_c \Leftrightarrow p_c(z) \in J_c$.

Exercise 3.4. Use the Cauchy estimates and the invariance of J_c to prove that any repelling periodic point for p_c is in J_c.

Before drawing some computer images of Julia sets, it will be helpful to study $\mathcal{A}(\infty)$ a bit more.

Definition 3.5. A *harmonic function* $h : \mathbb{C} \to \mathbb{R}$ is a function with continuous second partial derivatives $h(x + iy) \equiv h(x, y)$ satisfying

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} = 0.$$

One can use the Cauchy–Riemann equations to verify that the real or imaginary part of an analytic function is harmonic and also that any harmonic function can be written (locally) as the real or imaginary part of some analytic function. Thus, there is a close parallel between the theory of analytic functions and of harmonic functions. We will need only two facts which follow directly from their analytic counterparts:

Maximum principle. *Suppose $h(z)$ is harmonic in a domain U and $h(z)$ achieves its maximum or minimum at a point $z_0 \in U$. Then $h(z)$ is constant on U.*

If, moreover, $\overline{U}$ is compact and h extends continuously to $\overline{U}$, then h achieves its maximum and minimum on the boundary of U.

Uniform limits of harmonic functions. *Suppose $h_k : U \to \mathbb{R}$ is a sequence of harmonic functions and $h : U \to \mathbb{R}$ is some other function. If for any compact $K \subset U$ we have that $\{h_k\}$ converges uniformly to h on K, then h is harmonic on U.*

Moreover, any (repeated) partial derivative of h_k converges uniformly to the corresponding partial derivative of h on any compact $K \subset U$.

Lemma 3.6. *The following limit exists:*

$$G_c(z) := \lim_{n\to\infty} \frac{1}{2^n} \log_+ \left|p_c^{\circ n}(z)\right|, \quad \textit{where} \quad \log_+(x) = \max(\log(x), 0),$$

for any parameter $c \in C$ and any $z \in \mathbb{C}$. For each c the resulting function $G_c : \mathbb{C} \to \mathbb{R}$ is called the Green function *associated to p_c. It satisfies* *(i)* G_c *is continuous on* $\mathbb{C}$ *and harmonic on* $\mathcal{A}(\infty)$*;*

(ii) $G_c(p_c(z)) = 2G_c(z)$*;*

(iii) $G(z) \approx \log|z|$ *for* $|z|$ *sufficiently large; and*

(iv) $G(z) = 0$ *iff* $z \in K_c$.

The Green function G_c is interpreted as describing the rate at which the orbit of initial condition $z_0 = z$ escapes to infinity under iteration of p_c. (The proof of Lemma 3.6 is quite similar to the proofs of Kœnig's theorem and Böttcher's theorem, so we will omit it.)

It is customary, when drawing filled Julia sets on the computer, to color $\mathcal{A}(\infty) = \mathbb{C} \setminus K_c$ according to the values of $G_c(z)$. This is especially helpful for parameters c at which p_c has no attracting periodic orbit. Using how the colors cycle one can "view" where K_c should be. In Figure 3.17 we show the filled Julia sets for four different values of c. (Among these is $c = \frac{1}{2}$, from Example 1.9. We can now see where the bounded orbits are.)

For the parameter values $c = \frac{i}{4}$, -1, and $c \approx -0.12 + 0.75i$, the filled Julia set is the closure of the basin of attracting periodic orbit. Thus, Figures 3.9–3.11 also depict the filled Julia sets for these parameter values.

Remark. Like the ancient people who named the constellations, people doing complex dynamics also have active imaginations. They have named the filled Julia set for $c = -1$ the "basilica" and the filled Julia set for $c \approx -0.12 + 0.75i$ "Douady's rabbit."

The Green function also helps us to make better computer pictures of the Mandelbrot set. The value $G_c(0)$ expresses the rate at which the critical point 0 of p_c escapes to ∞ under iteration of p_c. Thus, points c with larger values of $G_c(0)$ should be farther away from M. Therefore, it is customary to color $\mathbb{C} \setminus M$ according to the values of $G_c(0)$, as in Figure 3.18. It is interesting to compare Figures 3.18 and 3.16. It now looks more plausible that the black "dots" in Figure 3.16 might be connected to the "main part" of M.

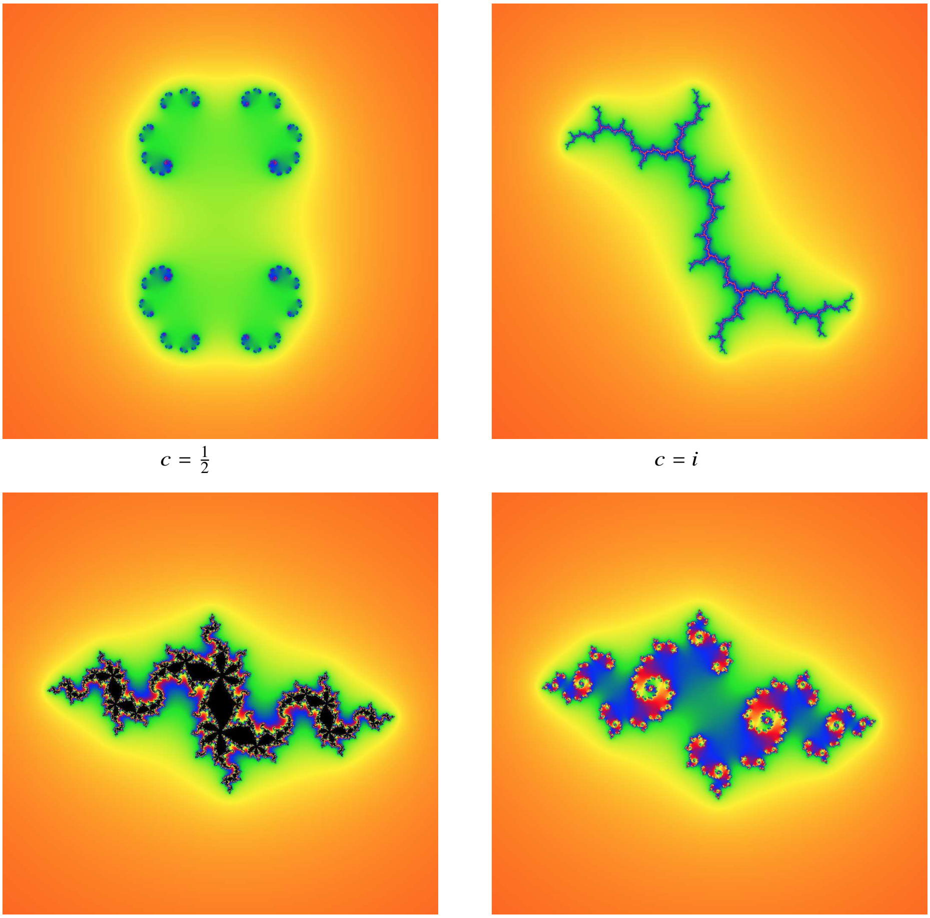
The Green function is not only useful for making nice pictures. It also plays a key role in the proof of the following result.

Topological characterization of the Mandelbrot set. *K_c is connected if and only if $c \in M$.*

We illustrate this theorem with Figure 3.19. The reader may also enjoy comparing the parameter values shown in Figure 3.18 with their filled Julia sets shown in previous figures.

According to definition (2.12) of M, this statement is equivalent to

Topological characterization of the Mandelbrot set$'$. *K_c is connected if and only if the orbit $\{p_c^{\circ n}(0)\}$ of the critical point 0 of p_c remains bounded.*



$c = \frac{1}{2}$ $c = i$

$c \approx -0.92 + 0.25i$ $c = -0.8 + 0.3i$

Figure 3.17. Filled Julia sets for four values of c. The basin of attraction for ∞ is colored according to the value of $G_c(z)$.

Although the Mandelbrot set was not defined at the time of Fatou and Julia's work (they lived from 1878–1929 and 1893–1978, respectively), the proof of the topological characterization of the Mandelbrot set$'$ is due to them.

Sketch of the proof. We will consistently identify $\mathbb{C}$ with $\mathbb{R}^2$ when taking partial derivatives and gradients of $G_c : \mathbb{C} \to \mathbb{R}$. We claim that $G_c(z)$ has a critical point at $z_0 \in \mathcal{A}(\infty)$ if and only if $p^{\circ n}(z_0) = 0$ for some $n \geq 0$. Consider the finite

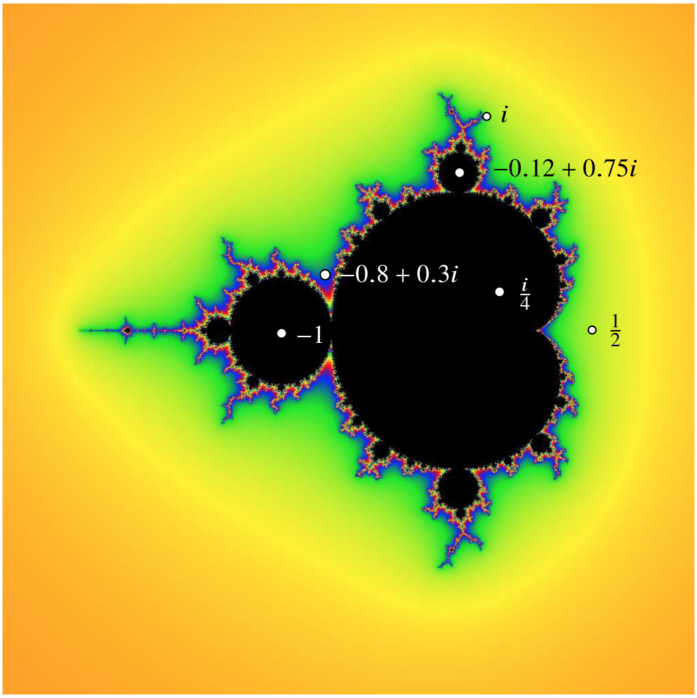


Figure 3.18. Mandelbrot set with the approximate locations of parameters $c = \frac{i}{4}, -1, -0.12+0.75i$, i, $-0.8+0.3i$, and $\frac{1}{2}$ labeled.

approximations

$$G_{c,n}(z) := \frac{1}{2^n} \log_+ |p_c^{\circ n}(z)|,$$

which one can check converge uniformly to $G_c(z)$ on any compact subset of $\mathbb{C}$. For points $z \in \mathcal{A}(\infty)$ we can drop the subscript + and use that $\log|z|$ is differentiable on $\mathbb{C} \setminus \{0\}$. Combined with the chain rule, we see that $\frac{\partial}{\partial x} G_{c,n}(z) = \frac{\partial}{\partial y} G_{c,n}(z) = 0$ if and only if $(p^{\circ n})'(z) = 0$. This holds if and only if $p^{\circ m}(z) = 0$ for some $0 \leq m \leq n-1$. The claim then follows from the uniform limits of harmonic functions theorem.

Suppose that the critical point 0 has bounded orbit under p_c. Then, according to the previous paragraph, G_c has no critical points in $\mathcal{A}(\infty)$. For any $t > 0$ let

$$L_t := \{z \in \mathbb{C} : G_c(z) \leq t\}.$$

By definition, if $t < s$ then $L_t \subset L_s$. For each $t > 0$, L_t is closed and bounded since $G_c : \mathbb{C} \to \mathbb{R}$ is continuous and $G_c(z) \to \infty$ as $|z| \to \infty$, respectively. Therefore, by the Heine–Borel theorem, L_t is compact. Since $K_c \neq \emptyset$ and $G_c(z) = 0$ on K_c, each L_t is nonempty.

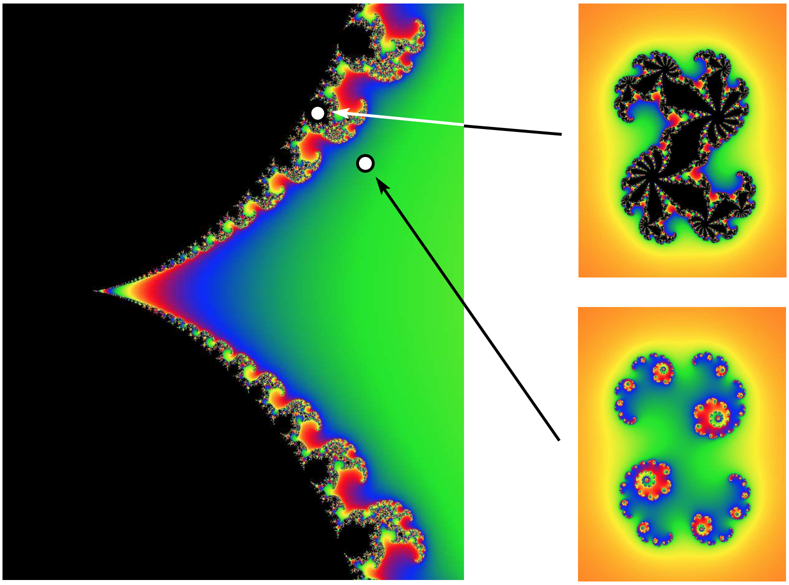


Figure 3.19. Left: Zoomed-in view of the Mandelbrot set near the cusp at $c = \frac{1}{4}$. Right: Two filled Julia sets corresponding to points inside M and outside M.

Since $z \in K_c$ if and only if $G_c(z) = 0$, we can write K_c as a nested intersection of nonempty compact sets:

$$K_c = \bigcap_{n \geq 1} L_{1/n}.$$

If we can show that L_t is connected for each $t > 0$, then Exercise 1.21 will imply that K_c is connected.

Since $G_c(z) \approx \log |z|$ for $|z|$ sufficiently large, there exists $t_0 > 0$ sufficiently large so that L_{t_0} is connected (it is almost a closed disk of radius $\log t_0$). We will show that for any $0 < t_1 < t_0$ the sets L_{t_1} and L_{t_0} are homeomorphic (i.e., there is a continuous bijection with continuous inverse from L_{t_1} to L_{t_0}). Since L_{t_0} is connected, this will imply that L_{t_1} is also connected.

The following is a standard construction from Morse theory; see [36, Theorem 3.1]. Because G_c is harmonic and has no critical points on $\mathcal{A}(\infty)$, $-\nabla G_c$ is a nonvanishing smooth vector field on $\mathcal{A}(\infty)$. It is a relatively standard smoothing construction to define a new vector field $\boldsymbol{V} : \mathbb{R}^2 \to \mathbb{R}^2$ that is smooth on all of $\mathbb{C} \equiv \mathbb{R}^2$ and equals $\frac{-\nabla G_c}{\|\nabla G_c\|^2}$ for $z \in \mathbb{C} \setminus L_{t_1/2}$.

For any $t \in [0, \infty)$ let $\Phi_t : \mathbb{R}^2 \to \mathbb{R}^2$ denote the flow obtained by integrating $\boldsymbol{V}$.

According to the existence and uniqueness theorem for ordinary differential equations (see, e.g., [38]), $\Phi_t : \mathbb{R}^2 \to \mathbb{R}^2$ is a homeomorphism for each $t \in [0, \infty)$. (We're using that $\boldsymbol{V}$ "points inwards" from ∞ so that the solutions exist for all time.)

For any $z_0 \in \mathbb{C} \setminus L_{t_1/2}$ and any $0 \leq t \leq G_c(z_0) - t_1/2$ we have

$$\begin{aligned}\frac{d}{dt} G_c(\Phi_t(z_0)) &= \nabla G_c(\Phi_t(z_0)) \cdot \frac{d}{dt}\Phi_t(z_0) = \nabla G_c(\Phi_t(z_0)) \cdot \boldsymbol{V}(\Phi_t(z_0)) \\ &= \nabla G_c(\Phi_t(z_0)) \cdot \frac{-\nabla G_c(\Phi_t(z_0))}{\|\nabla G_c(\Phi_t(z_0))\|^2} = -1.\end{aligned}$$

In particular, $\Phi_{t_0-t_1}(L_{t_0}) = L_{t_1}$, implying that L_{t_0} is homeomorphic to L_{t_1}.

Now suppose that 0 has unbounded orbit under p_c. In this case, 0 and all of its iterated preimages under p_c are critical points of G_c. Since p_c has a simple critical point at 0, one can check that these critical points of G_c are all "simple" in that the Hessian matrix of second derivatives has nonzero determinant. Moreover, by the maximum principle, they cannot be local minima or local maxima. They are therefore saddle points. From the property $G_c(p(z)) = 2G_c(z)$, the saddle point at $z = 0$ is the one with the largest value of G_c.

There are two paths along which we can start at 0 and walk uphill in the steepest way possible—call them γ_1 and γ_2. Since 0 is the highest critical point, they lead all the way from 0 out to ∞. Together with 0, these two paths divide $\mathbb{C}$ into two domains U_1 and U_2. Meanwhile, there are two directions that one can walk downhill from a saddle point. Walking the fastest way downhill leads to two paths η_1 and η_2 which lead to points in U_1 and in U_2 along which $G_c(z) < G_c(0)$.

To make this idea rigorous, one considers the flow associated to the vector field $-\nabla G_c$. The saddle point 0 becomes a saddle-type fixed point for the flow with the paths γ_1 and γ_2 being the stable manifold of this fixed point. The paths η_1 and η_2 are the unstable manifolds of this fixed point. (See again [38].)

The union $\gamma_1 \cup \gamma_2 \cup \{0\}$ divides the complex plane into two domains U_1 and U_2 with $\eta_1 \subset U_1$ and $\eta_2 \subset U_2$. We claim that both of these domains contain points of K_c. Suppose for contradiction that one of them (say U_1) does not. Then, $U_1 \subset \mathcal{A}(\infty)$ and hence G_c would be harmonic on U_1. However, $G_c(z) \sim \log |z|$ for $|z|$ large and $G_c(z) > G_c(0)$ for points $z \in \gamma_1 \cup \gamma_2$. Since $G_c(z) < G_c(0)$ for points on η_1, this would violate the maximum principle. □

Remark. A stronger statement actually holds: if K_c is disconnected, then it is a Cantor set. (See [17] for the definition of a Cantor set.) In particular, it is *totally disconnected*: for any $z, w \in K_c$ there exist open sets $U, V \subset \mathbb{C}$ such that $K_c \subset U \cup V$, $z \in U$, $w \in V$, and $U \cap V = \emptyset$. This follows from the fact that once G_c has the critical point $0 \in \mathcal{A}(\infty)$ then it actually has infinitely many critical points in $\mathcal{A}(\infty)$. These additional critical points of G_c are the iterated preimages of 0 under p_c.

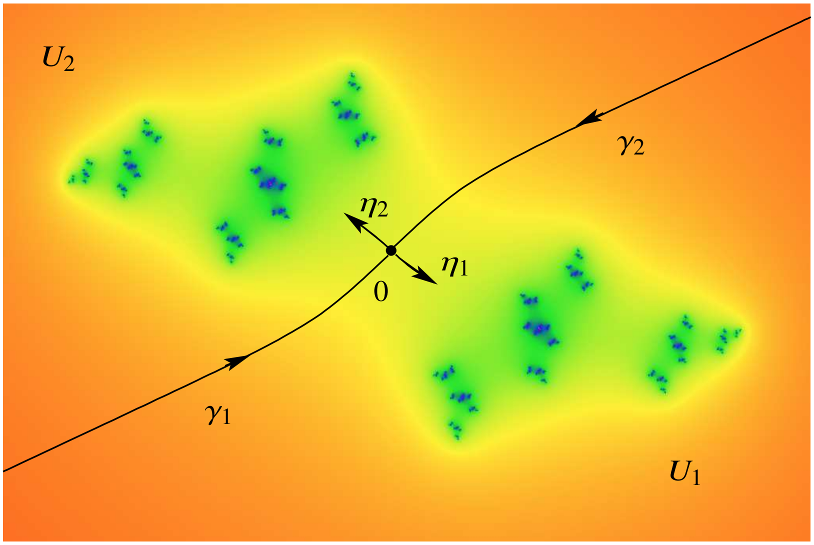


Figure 3.20. Stable and unstable trajectories of $-\nabla G_c$ for the critical point 0.

For a somewhat different proof from the one presented above, including a proof of this stronger statement, see [11, 12].

Exercise 3.7. Prove that if $c \neq 0$ then $\log|z^2+c|$ has a saddle-type critical point at $z=0$.
Hint: Write $z = x+iy$ and $c = a+ib$ and use that

$$\log|z^2+c| = \frac{1}{2}\log\left(z^2+c\right)\left(\overline{z^2+c}\right).$$

We will now state (without proofs) several interesting properties of the Mandelbrot set:

Theorem (Douady–Hubbard [14]). *The Mandelbrot set M is connected.*

(Nessim Sibony gave an alternate proof around the same time.) This theorem clears up the mystery about the black "dots" in Figure 3.16.

The following very challenging extended exercise leads the reader through a proof that M is connected, that is related to the coloring of $\mathbb{C} \setminus M$ according to the value of $G_c(0)$. (It will be somewhat more convenient to consider $G_c(c) = G_c(p_c(0)) = 2G_c(0)$.)

Extended exercise 3.8. Let $H : \mathbb{C} \to \mathbb{R}$ be given by $H(c) = G_c(c)$. Prove that

(1) H is continuous;

(2) H is harmonic on $\mathbb{C} \setminus M$;

(3) H is identically 0 on M;

(4) $\lim_{|c|\to\infty} H(c) = \infty$; and

(5) H has no critical points in $\mathbb{C} \setminus M$.

(Step 5 is the hardest part.) Use these facts to adapt the proof of the topological characterization of the Mandelbrot set to prove that M is connected.

Hausdorff dimension extends the classical notion of dimension from lines and planes to more general metric spaces. As the formal definition is a bit complicated, we instead illustrate the notion with a few examples. A line has Hausdorff dimension equal to 1 and the plane has Hausdorff dimension equal to 2. A contour has Hausdorff dimension equal to 1 because, if you zoom in sufficiently far near any of the smooth points, the contour appears more and more like a straight line. However, sets of a "fractal nature" can have noninteger Hausdorff dimension. One example is the Koch curve, which is a simple closed curve in the plane that is obtained as the limit of the iterative process shown in Figure 3.21. No matter how far you zoom in, the Koch curve looks the same as a larger copy of itself, and not like a line! This results in the Koch curve having Hausdorff dimension equal to $\log(4)/\log(3) \approx 1.26$. We refer the reader to [17] for a gentle introduction to Hausdorff dimension.

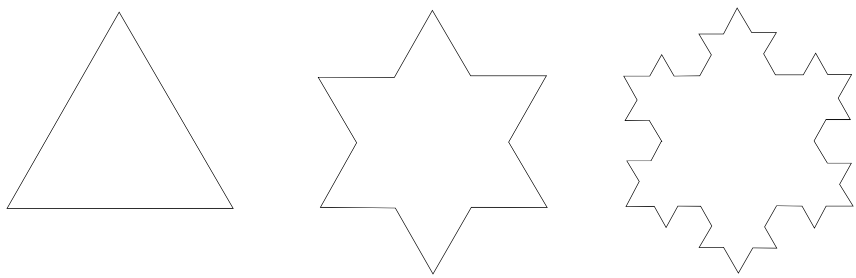


Figure 3.21. The Koch curve.

If $S \subset \mathbb{C}$ contains an open subset of $\mathbb{C}$, then it is easy to see that it has Hausdorff dimension equal to 2. It is much harder to imagine a subset of $\mathbb{C}$ that contains no such open set having Hausdorff dimension 2. Therefore, the following theorem shows that the boundary ∂M of the Mandelbrot set M has amazing complexity. It also shows that for many parameters c from ∂M, the Julia set J_c has amazing complexity.

Theorem (Shishikura [44]). *The boundary of the Mandelbrot set ∂M has Hausdorff dimension equal to 2. Moreover, for a dense set of parameters c from the boundary of M, the Julia set J_c has Hausdorff dimension equal to 2.*

Another interesting property of the Mandelbrot set is the appearance of "small copies" within itself. (Some of these were the "dots" from Figure 3.16.) Figure 3.22 shows a zoomed-in view of M, where several small copies of M are visible. These copies are explained by the renormalization theory [15, 34].

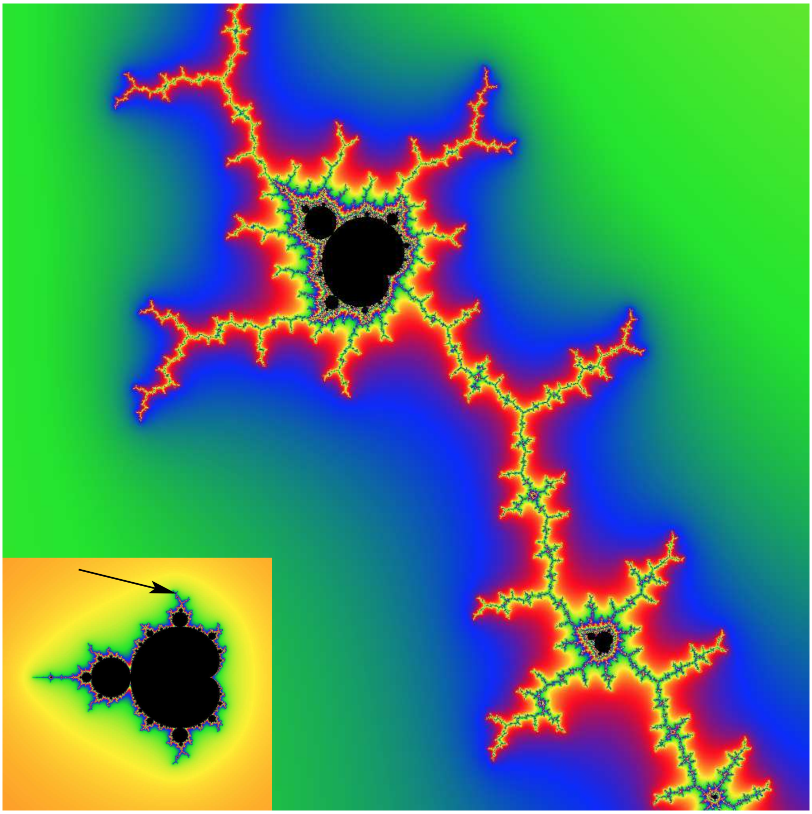


Figure 3.22. Zoomed-in view of part of the Mandelbrot set showing two smaller copies. The approximate location where we have zoomed in is marked by the tip of the arrow in the inset figure.

It would be remiss not to include one of the most famous conjectures about the Mandelbrot set. We first need a definition.

Definition 3.9. A topological space X is *locally connected* if for every point $x \in X$ and any open set $V \subset X$ that contains x there is another connected open set U with $x \in U \subset V$.

MLC conjecture. *The Mandelbrot set M is locally connected.*

According to the Orsay notes [13] of Douady and Hubbard, if this were the case, then one could have a very nice combinatorial description of M. Given a proposed way that p_c acts on the Julia set J_c (described by means of the so-called Hubbard tree), one can use this combinatorial description of M to find the desired value of c.

To better appreciate the difficulty in proving the MLC conjecture, we include one more zoomed-in image of the Mandelbrot set in Figure 3.23.

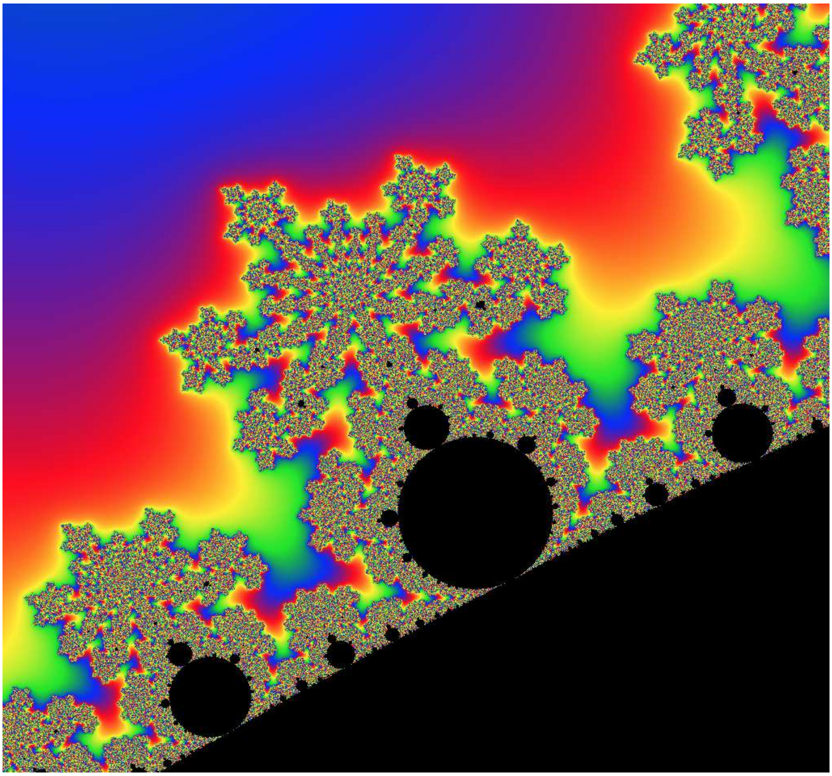


Figure 3.23. Another zoomed-in view of part of the Mandelbrot set.

Let us finish the section, and our discussion of iterating quadratic polynomials, by returning to mathematics that can be done by undergraduates. The reader is now ready to answer Section 1.4, question (4):

Extended exercise 3.10. Prove that for every $m \geq 1$, there exists a parameter $c \in \mathbb{C}$ such that $p_c(z)$ has an attracting periodic orbit of period exactly m.
Hint: Prove that there is a parameter c such that $p_c^{\circ m}(0) = 0$ and $p_c^{\circ j}(0) \neq 0$ for each $0 \leq j < m$.

Lecture 4. Complex dynamics and astrophysics

Most of the results discussed in Lectures 1–3 of this chapter are now quite classical. Let us finish our lectures with a beautiful and quite modern application of the Fatou–Julia lemma to a problem in astrophysics [26, 24]. We also mention that there are connections between complex dynamics and the Ising model from statistical physics (see [8, 7] and the references therein) and the study of droplets in a Coulomb gas [30, 29].

4.1 Gravitational lensing. Einstein's theory of general relativity predicts that if a point mass is placed directly between an observer and a light source, then the observer will see a ring of light, called an "Einstein ring." The Hubble Space Telescope has sufficient power to see these rings—one such image is shown in Figure 3.24. If the point mass is moved slightly, the observer will see two different images of the same light source. With more complicated distributions of mass, like n point masses, the observer can see more complicated images, resulting from a single point light source. Such an image is shown in Figure 3.25. (Thanks to NASA for these images and their interpretations.)

There are many excellent surveys on gravitational lensing that are written for the mathematically inclined reader, including [25, 39, 47], as well as the book [40]. We will be far more brief, with the goal of this lecture being to explain how Rhie [41] and Khavinson–Neumann [24] answered the following question:

What is the maximum number of images that a single light source can have when lensed by n point masses?

We will tell some of the history of how this problem was solved and then focus on the role played by the Fatou–Julia lemma.

Suppose that n point masses lie on a plane that is nearly perpendicular to the line of sight between the observer and the light source and that they lie relatively close to the line of sight. If we describe their positions relative to the line of sight to the light source by complex numbers z_j and their normalized masses by $\sigma_j > 0$ for $1 \leq j \leq n$, then the images of the light source seen by the observer are given by solutions z to the *lens equation:*

$$z = \sum_{j=1}^{n} \frac{\sigma_j}{\overline{z} - \overline{z_j}}. \tag{4.1}$$

The "mysterious" appearance of complex conjugates on the right-hand side of this equation makes it difficult to study. It will be explained in Section 4.3, where we derive (4.1) from.

Exercise 4.1. Verify that (4.1) gives a full circle of solutions (Einstein ring) when there is just one mass at $z_1 = 0$. Then, verify that when $z_1 \neq 0$ there are two solutions. Can you find a configuration of two masses so that (4.1) has five solutions?

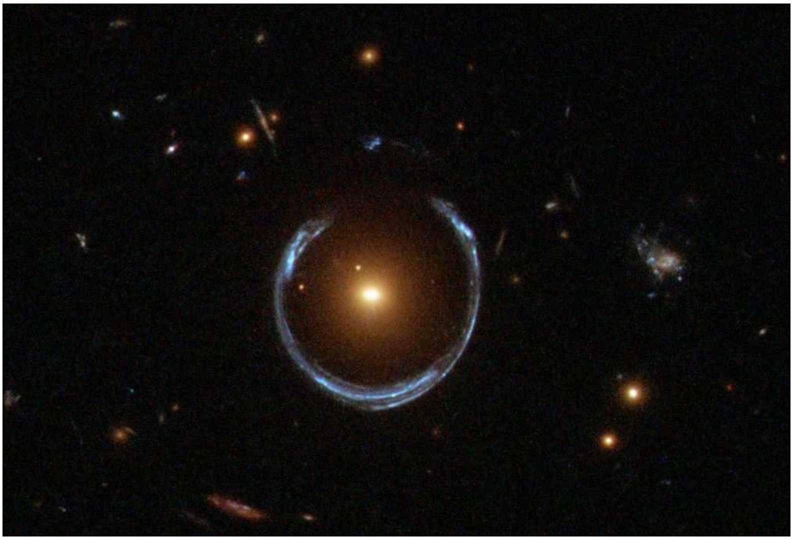


Figure 3.24. An Einstein ring. For more information, see http://apod.nasa.gov/apod/ap111221.html.

Remark. Techniques from complex analysis extend nicely to lensing by mass distributions more complicated than finitely many points, including elliptical [18] and spiral [4] galaxies.

The right-hand side of (4.1) is of the form $\overline{r(z)}$, where $r(z)$ is a rational function $r(z) = \frac{p(z)}{q(z)}$ of degree n. (The degree of a rational function is the maximum of the degrees of its numerator and denominator.) Thus, our physical question becomes the problem of bounding the number of solutions to an equation of the form

$$z = \overline{r(z)} \tag{4.2}$$

in terms of $n = \deg(r(z))$. Sadly, the fundamental theorem of algebra cannot be applied to

$$z\overline{q(z)} - \overline{p(z)} = 0 \tag{4.3}$$

because the resulting equation is a polynomial in *both* z and $\overline{z}$. If one writes $z = x + iy$ with $x, y \in \mathbb{R}$, one can change (4.3) to a system of two real polynomial equations

$$a(x, y) := \mathrm{Re}\left(z\,\overline{q(z)} - \overline{p(z)}\right) = 0 \quad \text{and}$$
$$b(x, y) := \mathrm{Im}\left(z\,\overline{q(z)} - \overline{p(z)}\right) = 0,$$

each of which has degree $n + 1$. So long as there are no curves of common zeros for $a(x, y)$ and $b(x, y)$, Bezout's theorem (see, e.g., [27]) gives a bound on the number of solutions by $(n + 1)^2$.

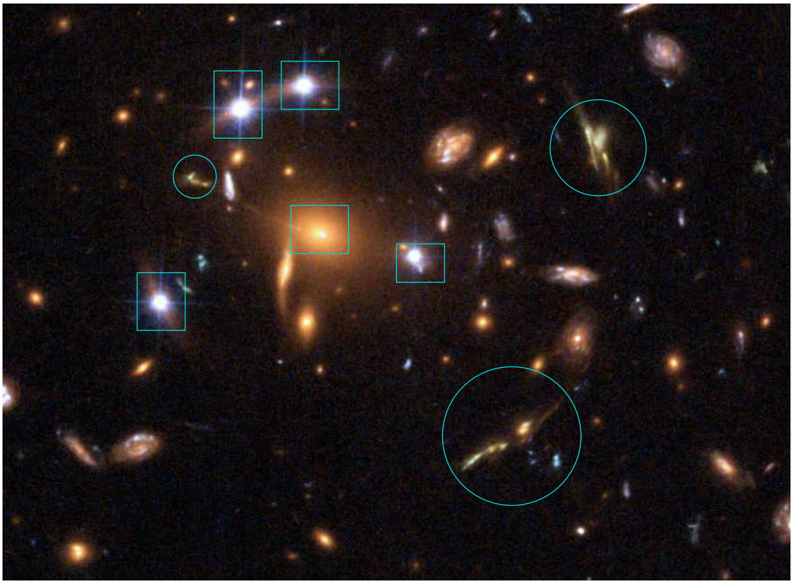


Figure 3.25. Five images of the same quasar (boxed) and three images of the same galaxy (circled). The middle image of the quasar (boxed) is behind the small galaxy that does the lensing. For more information, see http://www.nasa.gov/multimedia/imagegallery/image_feature_575.html.

In 1997, Mao, Petters, and Witt [35] exhibited configurations of n point masses at the vertices of a regular polygon in such a way that $3n + 1$ solutions were found. They conjectured a linear bound for the number of solutions to (4.1). For large n this would be significantly better than the bound given by Bezout's theorem.

In 2003, Rhie [41] showed that if one takes the configuration of masses considered by Mao, Petters, and Witt and places a sufficiently small mass centered at the origin, then one finds $5n - 5$ solutions to (4.1). (We refer the reader also to [6, Section 5] for another exposition on Rhie's examples.)

In order to address a problem on harmonic mappings $\mathbb{C} \to \mathbb{C}$ posed by Wilmshurst in [49], in 2003 Khavinson and Świątek studied the number of solutions to $z = \overline{p(z)}$ where $p(z)$ is a complex polynomial. They proved

Theorem (Khavinson–Świątek [26]). *Let $p(z)$ be a complex polynomial of degree $n \geq 2$. Then, $z = \overline{p(z)}$ has at most $3n - 2$ solutions.*

Khavinson and Neumann later adapted the techniques from [26] to prove

Theorem (Khavinson–Neumann [24]). *Let $r(z)$ be a rational function of degree $n \geq 2$. Then, $z = \overline{r(z)}$ has at most $5n - 5$ solutions.*

Apparently, Khavinson and Neumann solved this problem because of its mathematical interest and only later were informed that they had actually completed the solution to our main question of this lecture:
When lensed by n point masses, a single light source can have at most $5n - 5$ *images.*

Remark. Geyer [20] used a powerful theorem of Thurston to show that for every $n \geq 2$ there is a polynomial $p(z)$ for which $z = \overline{p(z)}$ has $3n - 2$ solutions, thus showing that the Khavinson–Świątek theorem above is sharp. It would be interesting to see an "elementary" proof.

4.2 Sketching the proof of the $5n - 5$ bound. We provide a brief sketch of the proof of the Khavinson–Neumann upper bound in the special case that

$$r(z) = \sum_{j=1}^{n} \frac{\sigma_j}{z - z_j}, \tag{4.4}$$

with each $\sigma_j > 0$. It is the case arising in the lens equation (4.1). The locations $\{z_1, \ldots, z_n\}$ of the masses are called *poles* of $r(z)$. They satisfy $\lim_{z \to z_j} |r(z)| = \infty$ for any $1 \leq j \leq n$.

The function

$$f : \mathbb{C} \setminus \{z_1, \ldots, z_n\} \to \mathbb{C} \quad \text{given by} \quad f(z) = z - \overline{r(z)} \tag{4.5}$$

is an example of a *harmonic mapping with poles* since its real and imaginary parts are harmonic. It is orientation preserving near a point $z_\bullet$ with $|r'(z_\bullet)| < 1$ and orientation reversing (like a reflection $z \mapsto \overline{z}$) near points with $|r'(z_\bullet)| > 1$. A zero $z_\bullet$ of f is *simple* if $|r'(z_\bullet)| \neq 1$ and a simple zero is called *sense preserving* if $|r'(z_\bullet)| < 1$ and *sense reversing* if $|r'(z_\bullet)| > 1$.

Step 1: Reduction to simple zeros. Suppose $r(z)$ is of the form (4.4) and $f(z) = z - \overline{r(z)}$ has k zeros, some of which are not simple. Then, one can show that there is an arbitrarily small perturbation of the locations of the masses so that the resulting rational function $s(z)$ produces $g(z) = z - \overline{s(z)}$ having at least as many zeros as $f(z)$ all of which are simple.

Therefore, it suffices to consider rational functions $r(z)$ of the form (4.4) such that each zero of $f(z) = z - \overline{r(z)}$ is simple.

Step 2: Argument principle for harmonic mappings. Suffridge and Thompson [48] have adapted the argument principle to harmonic mappings with poles $f : \mathbb{C} \setminus \{z_1, \ldots, z_k\} \to \mathbb{C}$.

Since $r(z)$ has the form (4.4), $\lim_{z \to \infty} |r(z)| = 0$. Therefore, we can choose $R > 0$ sufficiently large so that all of the poles of $r(z)$ lie in $D(0, R)$ and the change of argument for $f(z)$ while traversing $\gamma = \partial D(0, R)$ counterclockwise is 1. This variant of the argument principle then gives

$$(m_+ - m_-) + n = 1, \tag{4.6}$$

where m_+ is the number of sense-preserving zeros, m_- is the number of sense-reversing zeros, and n is the number of poles. (We are using that all of the zeros and poles are simple so that they do not need to be counted with multiplicities.)

Step 3: Fatou–Julia lemma bound on m_+. Zeros of $f(z)$ correspond to fixed points for the anti-analytic mapping

$$z \mapsto \overline{r(z)}.$$

Moreover, sense-preserving zeros correspond to attracting fixed points (those with $|r'(z_\bullet)| < 1$).

Since the coefficients of $r(z)$ are real, taking the second iterate yields

$$Q(z) = \overline{r\left(\overline{r(z)}\right)} = r(r(z)),$$

which is an analytic rational mapping of degree n^2. Such a mapping has $2n^2 - 2$ critical points and an adaptation of the Fatou–Julia lemma implies that each attracting fixed point of Q attracts a critical point.

However, the chain rule gives that critical points of $Q(z) = r(r(z))$ are the critical points of $r(z)$ and their inverse images under $r(z)$. Since a generic point has n inverse images under r, this can be used to show that each attracting fixed point of $Q(z)$ actually attracts $n + 1$ critical points of Q. Therefore, $Q(z)$ has at most $2n - 2$ attracting fixed points.

Since any sense-preserving zero for $f(z)$ is an attracting fixed point for Q, we conclude that $m_+ \leq 2n - 2$.

Step 4: Completing the proof. Since $m_+ \leq 2n - 2$, equation (4.6) implies $m_- \leq 3n - 3$. Therefore, the total number of zeros is

$$m_+ + m_- \leq 5n - 5.$$

The reader is encouraged to see [24] for the full details, including how to prove the bound for general rational functions $r(z)$.

4.3 Derivation of the lens equation. This derivation is a synthesis of ideas from [2] and [40, Section 3.1] that was written jointly with Bleher, Homma, and Ji when preparing [6]. Since it was not included in the published version of [6], we present it here.

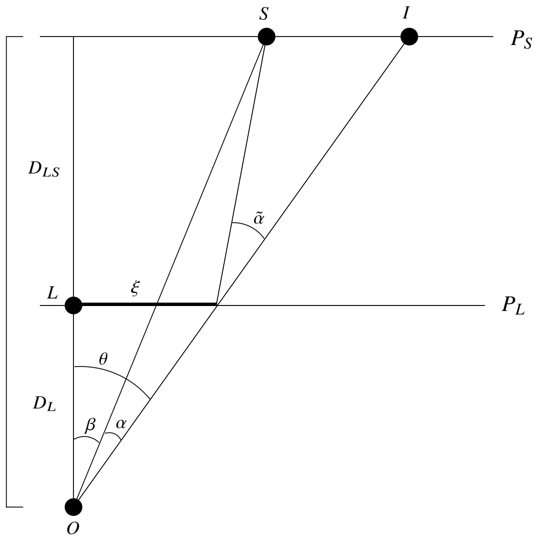


Figure 3.26. S is the light source, I is an image, O is the observer, L is a point mass, P_L is the lens plane, P_S is the source plane.

We will first derive the lens equation for one point mass using Figure 3.26, and then adapt it to N point masses. Suppose the observer is located at point O, the light source at a point S, and a mass M at point L. Also, suppose P_L is the plane perpendicular to OL that contains L, and P_S is the plane perpendicular to OL that contains S. Due to the point mass, an image, I, will be created at angle α with respect to S.

Einstein derived, using general relativity, that the bending angle is

$$\tilde{\alpha} = \frac{4GM}{c^2 \xi}, \tag{4.7}$$

where G is the universal gravitational constant and c is the speed of light; see [2].

The observer O describes the location S of the light source using an angle β and the perceived location I using another angle θ (see Figure 3.26). By a small angle approximation, $\xi = D_L \theta$, which we substitute into (4.7) obtaining

$$\tilde{\alpha}(\theta) = \frac{4GM}{c^2 D_L \theta}. \tag{4.8}$$

A small angle approximation also gives that $D_{SI} = D_{LS}\tilde{\alpha}(\theta) = D_S \alpha(\theta)$. Substituting this into $\beta = \theta - \alpha(\theta)$ gives

$$\beta = \theta - \frac{D_{LS}}{D_S D_L} \cdot \frac{4GM}{c^2 \theta}. \tag{4.9}$$

For $\beta \neq 0$, exactly two images are produced. When $\beta = 0$, the system is rotationally symmetric about OL, thereby producing an Einstein ring, whose angular radius is given by equation (4.9).

In order to describe systems of two or more masses, we need to describe locations in the source plane P_S and the lens plane P_L using two-dimensional vectors of angles (polar and azimuthal angles) as observed from O. Complex numbers will be a good way to do this:

$$\alpha = \alpha^{(1)} + i\alpha^{(2)}, \ \tilde{\alpha} = \tilde{\alpha}^{(1)} + i\tilde{\alpha}^{(2)}, \ \beta = \beta^{(1)} + i\beta^{(2)}, \ \text{and } \theta = \theta^{(1)} + i\theta^{(2)}.$$

When there is only one mass, the whole configuration must still lie in one plane, as in Figure 3.26. In particular, all four complex numbers have the same argument, forcing us to replace the θ on the right-hand side of (4.8) with $\overline{\theta}$:

$$\tilde{\alpha}(\theta) = \frac{4GM}{c^2 D_L \overline{\theta}} \qquad \text{and hence} \qquad \beta = \theta - \frac{D_{LS}}{D_S D_L} \cdot \frac{4GM}{c^2 \overline{\theta}}. \tag{4.10}$$

This is why the complex conjugate arises in the lens equation (4.1).

We now generalize to n point masses. Let L be the center of mass of the n masses, and redefine S_L as the plane that is perpendicular to OL and contains L. We assume that the distance between L and the individual point masses is extremely small with respect to the pairwise distances between O, P_L, and P_S. Now consider the projection of the n point masses onto S_L. We continue to let $\beta \in \mathbb{C}$ describe the location of the center of mass and we describe the location of the jth point mass by $\epsilon_j = \epsilon_j^{(1)} + i\epsilon_j^{(2)} \in \mathbb{C}$. It has mass M_j.

In general, the bending angle is expressed as an integral expressed linearly in terms of the mass distribution; see [40, Equation 3.57]. In particular, with point masses, the bending angle decomposes to a sum of bending angles, one for each point mass. Each is computed as in the one-mass system:

$$\tilde{\alpha}_j = \frac{4GM_j}{c^2 D_L \overline{\theta_j}}, \qquad \text{where} \qquad \theta_j = \theta - \epsilon_j.$$

We obtain

$$\beta = \theta - \sum_{j=1}^{n} \alpha_j = \theta - \frac{D_{LS}}{D_S D_L} \sum_{j=1}^{n} \frac{4GM_j}{c^2 \left(\overline{\theta} - \overline{\epsilon_j}\right)}.$$

Letting

$$w = \beta, \quad z = \theta, \quad z_j = \epsilon_j, \quad \text{and} \quad \sigma_j = \frac{D_{LS}}{D_S D_L} \cdot \frac{4GM_j}{c^2}$$

gives

$$w = z - \sum_{j=1}^{n} \frac{\sigma_j}{\overline{z} - \overline{z_j}}. \tag{4.11}$$

Equation (4.11) requires the assumption that the center of mass is the origin, i.e., $\sum \sigma_j z_j = 0$. A translation by w allows us to fix the position of the light source at the origin and vary the location of the center of mass. This simplifies equation (4.11) to equation (4.1).

4.4 Wilmshurst's conjecture. Let us finish our notes with an open problem that can be explored by undergraduates. In [49], Wilmshurst considered equations of the form

$$p(z) = \overline{q(z)}, \tag{4.12}$$

where $p(z)$ and $q(z)$ are polynomials of degree n and m, respectively. By conjugating the equation, if necessary, one may suppose $n \geq m$. If $m = n$, then one can have infinitely many solutions (e.g., $p(z) = z^n = q(z)$), but once $n > m$ Wilmshurst showed that there are finitely many solutions. He conjectured that the number of solutions to (4.12) is at most $3n - 2 + m(m - 1)$.

Unfortunately, this conjecture is false! Counterexamples were found when $m = n - 3$ by Lee, Lerario, and Lundberg [28]. They propose the following conjecture:

Conjecture (Lee, Lerario, Lundberg). *If* $\deg(p(z)) = n$, $\deg(q(z)) = m$, *and* $n > m$, *then the number of solutions to* $p(z) = \overline{q(z)}$ *is bounded by* $2m(n - 1) + n$.

Note that this conjectured bound is not intended to be sharp. For example, Wilmshurst proved his conjecture in the case that $m = n - 1$, providing a stronger bound in that case [49].

This problem was further studied using certified numerics by Hauenstein, Lerario, Lundberg, and Mehta [22]. Their work provides further evidence for this conjecture.

Question. Can techniques from complex dynamics be used to prove this conjecture?

Acknowledgments

I am grateful to the University of Notre Dame for their hospitality during the thematic program "Boundaries and Dynamics." Ivan Chio, Youkow Homma, Lyndon Ji, Scott Kaschner, Dmitry Khavinson, Seung-Yeop Lee, Rodrigo Pérez, and Mitsuhiro Shishikura provided many helpful comments. All of the computer-drawn images of basins of attraction, filled Julia sets, and the Mandelbrot set were created using the Fractalstream software [19] that was written by Matthew Noonan.

Bibliography

[1] Lars V. Ahlfors. *Complex analysis*. Third edition, McGraw-Hill, New York (1978). An introduction to the theory of analytic functions of one complex variable, International Series in Pure and Applied Mathematics. Zbl 0395.30001.

[2] H. Asada. Perturbation theory of N point-mass gravitational lens systems without symmetry. *Monthly Notices of the Royal Astronomical Society*, **394**(2): 818–830 (2009). doi 10.1111/j.1365-2966.2008.14367.x.

[3] Alan F. Beardon. *Iteration of rational functions*. Graduate Texts in Mathematics, volume 132. Springer, New York (1991). Complex analytic dynamical systems. Zbl 0742.30002.

[4] Steven R. Bell, Brett Ernst, Sean Fancher, Charles R. Keeton, Abi Komanduru, and Erik Lundberg. Spiral galaxy lensing: a model with twist. *Math. Phys. Anal. Geom.*, **17**(3-4): 305–322 (2014). Zbl 1308.85002.

[5] Paul Blanchard. Complex analytic dynamics on the Riemann sphere. *Bull. Amer. Math. Soc. (N.S.)*, **11**(1): 85–141 (1984). Zbl 0558.58017.

[6] Pavel M. Bleher, Youkow Homma, Lyndon L. Ji, and Roland K. W. Roeder. Counting zeros of harmonic rational functions and its application to gravitational lensing. *Int. Math. Res. Not. IMRN*, **2014**(8): 2245–2264 (2014). Zbl 1329.85001.

[7] Pavel Bleher, Mikhail Lyubich, and Roland Roeder. Lee–Yang–Fisher zeros for DHL and 2D rational dynamics, II. Global pluripotential interpretation. Preprint available at arXiv:1107.5764.

[8] Pavel Bleher, Mikhail Lyubich, and Roland Roeder. Lee–Yang zeros for DHL and 2D rational dynamics, I. Foliation of the physical cylinder. Preprint available at arXiv:1009.4691.

[9] *Boston University Java Applets for Dynamical Systems*. http://math.bu.edu/DYSYS/applets/.

[10] Lennart Carleson and Theodore W. Gamelin. *Complex dynamics*. Universitext: Tracts in Mathematics. Springer, New York (1993). Zbl 0782.30022.

[11] Robert L. Devaney. *A first course in chaotic dynamical systems*. Addison-Wesley Studies in Nonlinearity. Addison-Wesley, Advanced Book Program, Reading, MA (1992). Theory and experiment. With a separately available computer disk. Zbl 0768.58001.

[12] Robert L. Devaney. *An introduction to chaotic dynamical systems*. Studies in Nonlinearity. Westview Press, Boulder, CO (2003). Reprint of the second (1989) edition. Zbl 1025.37001.

[13] A. Douady and J. H. Hubbard. *Étude dynamique des polynômes complexes. Partie I*. Publications Mathématiques d'Orsay [Mathematical Publications of Orsay], volume 84. Université de Paris-Sud, Département de Mathématiques, Orsay (1984). Zbl 0571.30026.

[14] Adrien Douady and John Hamal Hubbard. Itération des polynômes quadratiques complexes. *C. R. Acad. Sci. Paris Sér. I Math.*, **294**(3): 123–126 (1982). Zbl 0483.30014.

[15] Adrien Douady and John Hamal Hubbard. On the dynamics of polynomial-like mappings. *Ann. Sci. École Norm. Sup. (4)*, **18**(2): 287–343 (1985). Zbl 0587.30028.

[16] *Dynamics Explorer software.* Written by Brian and Suzanne Boyd. http://sourceforge.net/projects/detool/.

[17] Gerald A. Edgar. *Measure, topology, and fractal geometry*. Undergraduate Texts in Mathematics. Springer, New York (1990). Zbl 0727.28003.

[18] C. D. Fassnacht, C. R. Keeton, and D. Khavinson. Gravitational lensing by elliptical galaxies, and the Schwarz function. *Analysis and Mathematical Physics*, Trends Math., pp. 115–129. Birkhäuser, Basel (2009). Zbl 1297.85003.

[19] *Fractalstream dynamical systems software.* Written by Matthew Noonan. https://code.google.com/p/fractalstream/downloads/list.

[20] Lukas Geyer. Sharp bounds for the valence of certain harmonic polynomials. *Proc. Amer. Math. Soc.*, **136**(2): 549–555 (2008). Zbl 1133.26009.

[21] Jacek Graczyk and Grzegorz Świątek. Generic hyperbolicity in the logistic family. *Ann. of Math. (2)*, **146**(1): 1–52 (1997). Zbl 0936.37015.

[22] Jonathan Hauenstein, Antonio Lerario, Erik Lundberg, and Dhagash Mehta. Experiments on the zeros of harmonic polynomials using certified counting. *Exp. Math.*, **24**(2): 133–141 (2015). See also http://arxiv.org/pdf/1406.5523.pdf. Zbl 06465603.

[23] John Horgan. Mandelbrot set-to: Did the father of fractals "discover" his namesake set? *Scientific American*, **262**: 30–34 (April 1990).

[24] Dmitry Khavinson and Genevra Neumann. On the number of zeros of certain rational harmonic functions. *Proc. Amer. Math. Soc.*, **134**(4): 1077–1085 (2006). (Electronic.) Zbl 1090.30008.

[25] Dmitry Khavinson and Genevra Neumann. From the fundamental theorem of algebra to astrophysics: a "harmonious" path. *Notices Amer. Math. Soc.*, **55**(6): 666–675 (2008). Zbl 1194.30001.

[26] Dmitry Khavinson and Grzegorz Świątek. On the number of zeros of certain harmonic polynomials. *Proc. Amer. Math. Soc.*, **131**(2): 409–414 (2003). Zbl 1034.30003.

[27] Frances Kirwan. *Complex algebraic curves*. London Mathematical Society Student Texts, volume 23. Cambridge University Press, Cambridge (1992). Zbl 0744.14018.

[28] Seung-Yeop Lee, Antonio Lerario, and Erik Lundberg. Remarks on Wilmshurst's theorem. *Indiana Univ. Math. J.*, **64**(4): 1153–1167 (2015). See also http://arxiv.org/pdf/1308.6474.pdf. Zbl 1332.31001.

[29] Seung-Yeop Lee and Nikolai Makarov. Sharpness of connectivity bounds for quadrature domains. Preprint available at arXiv:1411.3415.

[30] Seung-Yeop Lee and Nikolai Makarov. Topology of quadrature domains. *J. Am. Math. Soc.*, **29**(2): 333–369 (2016). (See also arXiv:1009.4691.) Zbl 06552399.

[31] Mikhail Lyubich. Dynamics of rational transformations: Topological picture. *Uspekhi Mat. Nauk*, **41**(4(250)): 35–95, 239 (1986). Zbl 0619.30033.

[32] Mikhail Lyubich. Dynamics of quadratic polynomials. I, II. *Acta Math.*, **178**(2): 185–247, 247–297 (1997). Zbl 0908.58053.

[33] Mikhail Lyubich. Six lectures on real and complex dynamics (based on European lectures given in Barcelona, Copenhagen and St. Petersburg in May–June 1999). See http://www.math.sunysb.edu/~mlyubich/papers/lectures.ps.

[34] Mikhail Lyubich. Baby Mandelbrot sets, renormalization and MLC. *Gaz. Math.*, **113**: 45–50 (2007). Zbl 1171.37300.

[35] S. Mao, A. O. Petters, and H. J. Witt. Properties of point mass lenses on a regular polygon and the problem of maximum number of images. *The Eighth Marcel Grossmann Meeting, Part A, B (Jerusalem, 1997)*, pp. 1494–1496. World Scientific Publishing, River Edge, NJ (1999).

[36] J. Milnor. *Morse theory*. Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, volume 51. Princeton University Press, Princeton, NJ (1963). Zbl 0108.10401.

[37] John Milnor. *Dynamics in one complex variable*. Annals of Mathematics Studies, volume 160, third edition. Princeton University Press, Princeton, NJ (2006). Zbl 1085.30002.

[38] Lawrence Perko. *Differential equations and dynamical systems*, volume 7. Springer Science & Business Media (2001). Zbl 0973.34001.

[39] A. O. Petters. Gravity's action on light. *Notices Amer. Math. Soc.*, **57**(11): 1392–1409 (2010). Zbl 1221.83013.

[40] Arlie O. Petters, Harold Levine, and Joachim Wambsganss. *Singularity theory and gravitational lensing*. Progress in Mathematical Physics, volume 21. Birkhäuser Boston, Boston, MA (2001). With a foreword by David Spergel. Zbl 0979.83001.

[41] S. H. Rhie. n-point gravitational lenses with $5(n-1)$ images. Preprint available at arXiv:astro-ph/0305166.

[42] E. B. Saff and A. D. Snider. *Fundamentals of complex analysis for mathematics, science, and engineering*, third edition. Prentice Hall (2003).

[43] Zhaiming Shen and Lasse Rempe-Gillen. The exponential map is chaotic: An invitation to transcendental dynamics. *American Mathematical Monthly*, **122**(10):919–940 (2015). Zbl 06593895. See also http://arxiv.org/pdf/1408.1129v2.pdf.

[44] Mitsuhiro Shishikura. The Hausdorff dimension of the boundary of the Mandelbrot set and Julia sets. *Ann. of Math. (2)*, **147**(2): 225–267 (1998). Zbl 0922.58047.

[45] Norbert Steinmetz. *Rational iteration*. de Gruyter Studies in Mathematics, volume 16. Walter de Gruyter, Berlin (1993). Complex analytic dynamical systems. Zbl 0773.58010.

[46] James Stewart. *Calculus*, seventh edition. Cengage Learning (2012).

[47] Norbert Straumann. Complex formulation of lensing theory and applications. *Helvetica Physica Acta*, **70**(6): 894–908 (1997). Zbl 0890.53071.

[48] T. J. Suffridge and J. W. Thompson. Local behavior of harmonic mappings. *Complex Variables Theory Appl.*, **41**(1): 63–80 (2000). Zbl 1020.30015.

[49] A. S. Wilmshurst. The valence of harmonic polynomials. *Proc. Amer. Math. Soc.*, **126**(7): 2077–2081 (1998). Zbl 0891.30009.