Global convergence behavior of Newton's method

Kiyoko NISHIZAWA, Masayo FUJIMURA

Department of mathematics, Faculty of Science and Technology,

Sophia University

1 Introduction

The subject treated here is an attempt to understand the efficiency of algorithms for solving a non-linear equation. Among others, Newton's method has played a central role in root-finding algorithms for a polynomial and amounts to the dynamical properties of a certain rational function (Newton map) under iterations. The global study of the algorithm forces the introduction of topology and geometry into this subject.

We write $N: P_d \times \overline{\mathbb{C}} \to \overline{\mathbb{C}}$, where P_d is the space of polynomials of degree $\leq d$ and $\overline{\mathbb{C}}$ is the Riemann sphere $\mathbb{C} \cup \infty$. Then $N(p, z) = N_p(z) = z - p(z)/p'(z)$ is rational over $\overline{\mathbb{C}}$ in $p \in P_d$ and $z \in \overline{\mathbb{C}}$; that is, N can be formed from the complex rational operations $(+, -, \times, \div)$ from the coefficients of p and z.

If z is sufficiently close to a root η of p, then the sequence defined by

$$z_1 = N_p(z), \ z_2 = N_p^2(z) = N_p(z_1), \cdots, z_k = N_f^k(z) = N_p(z_{k-1})$$

converges to η as k tends to ∞ . However, as is well known there is an open set U in $P_d \times \overline{\mathbf{C}}$ such that this convergence will not happen for (f, z) in U. Consequently, for Newton's iterative scheme, two distinctly different types of behavior have been observed. In the first case, this algorithm succeeds for an open dense set of starting points. The set of exceptional points (Julia set) is closed, nowhere dense and has two dimensional measure zero. The second case exhibits an open set of initial conditions where this algorithm fails. The failure is due to the existence of an attracting periodic cycle of a Newton map.

By Smale([10]), it was conjectured that no such algorithm could be generally convergent. C.McMullen ([5]) answered the question by showing that there is no generally convergent purely iterative algorithm, rational over \overline{C} , for finding roots of polynomials of degree ≥ 4 . Here "purely iterative" means that the algorithm can be expressed as a discrete dynamical system on \overline{C} parameterized by the polynomial.

However it was shown by M.Shub and S.Smale([9]) that if one adds the operation of complex conjugation, then there do exist generally convergent purely iterative algorithms for finding zeros of polynomials.

In this paper we shall analyze the global behavior of Newton's map from the viewpoint of complex dynamics of rational functions. In Section 3 as Theorem 3.1, we give a complete criterion for a rational function to be a Newton's method as applied to a polynomial map. In Section 4, we study how one can guarantee success of Newton's method, by measuring the width of basins of roots. In Section 5, Newton map is studied for the generic cubic polynomials. After a change of variables, these polynomials parameterized by a single complex parameter. The Newton map has a single critical point other than its fixed points at the roots of the polynomial. One observe the variety of behavior of the orbit of the free critical point depending on parameters. The Julia set, points where Newton's method fail to converge, is also pictured.

2 Preliminaries from Complex Dynamical Systems

In this section some minimal knowledge and notations from iteration theory for rational functions are given. For examples, to study the behavior of orbits, the structure of invariant sets (Julia sets etc.) for the iteration of rational functions on the Riemann Sphere, and how they change when the functions are varied.

Let $\overline{\mathbf{C}} = \mathbf{C} \cup \{\infty\}$ be the Riemann sphere. We consider a rational function $f : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$ as a dynamical system (complex dynamical system) on $\overline{\mathbf{C}}$. The *n*-th iteration f^n $(n \in \mathbf{Z})$ is defined by:

$$f^0 = id, f^{n+1} = f^n \circ f, \text{ and } f^{-n} = (f^n)^{-1} \ (n \ge 0).$$

The orbit of a point $z \in \overline{\mathbb{C}}$ is the sequence $\{f^n(z); n \ge 0\}$. The degree of $f(z) = \frac{P(z)}{Q(z)}$, denoted by deg(f), is equal to the maximum of the degree of P(z) and Q(z), i.e. deg(f) = max{deg(P), deg(Q)}, where we assume that P(z) and Q(z) have no common roots. This is equal to degree of mapping of f. In fact, for any $z \in \overline{\mathbb{C}}$, the inverse image $f^{-1}(z)$ consists of exactly d(= deg(f)) solutions, counted with multiplicity. From now on, we assume $d \ge 2$, since for $d \le 1$ there are only simple dynamical systems. At first we note that

• a rational function of degree d has precisely d + 1 fixed points.

conjugacy

global conjugacy A rational function f is (analytic) conjugate to a rational function g iff there exist a Möbius transformation $A(z) = \frac{az+b}{cz+d} \in PSL(2, \mathbb{C})$, which satisfies the following commutative diagram :

$$\begin{array}{cccc} \overline{\mathbf{C}} & \xrightarrow{f} & \overline{\mathbf{C}} \\ A & & \downarrow A \\ \overline{\mathbf{C}} & \xrightarrow{g} & \overline{\mathbf{C}} \end{array}$$

By this transformation A(z), the orbit of a point z_0 under f corresponds to that of the point $A(z_0)$ under $A \circ f \circ A^{-1}$. Therefore, if necessary, we consider $A \circ f \circ A^{-1}$ instead of f, after such a conjugation. We note that if f and g are (global)conjugate, then deg(f) = deg(g).

local conjugacy Another important property of conjugacy is concerning about fixed

points. Let f and g are rational function with fixed point z_0 and w_0 respectively. If there exists a biholomorphic function $\varphi: U(z_0) \to V(w_0)$ such that $\varphi \circ f = g \circ \varphi$, then it is said that f is (locally) conjugate to g by φ in U. Namely the following diagram commutes:

$$\begin{array}{cccc} U & \stackrel{f}{\longrightarrow} & U \\ \varphi & & & \downarrow \varphi \\ V & \stackrel{g}{\longrightarrow} & V \end{array}$$

period point

Let x_0 be a periodic point of period s, i.e. $f^s(x_0) = x_0$ for a rational function f. If $x_0 \neq \infty$ then we define **eigen value** of x_0 as follows:

$$\lambda = \left(f^s\right)'\left(x_0\right)$$

A periodic point x_0 is said to be

attracting	if	$0 < \lambda < 1$
super-attracting	if	$\lambda = 0$
neutral (indifferent)	if	$ \lambda = 1$
repelling	if	$ \lambda > 1.$

Julia set

We fix a metric ρ , on $\overline{\mathbb{C}}$. A family \mathcal{F} of continuous mappings from an open set U on $\overline{\mathbb{C}}$ to $\overline{\mathbb{C}}$ is said to be equicontinuous if for any $\varepsilon > 0$ and $x \in U$ there exists $\delta > 0$ such that $\rho(f(x), f(y)) < \varepsilon$ for any $f \in \mathcal{F}$, whenever $\rho(x, y) < \delta$ and $y \in U$. According to Ascoli-Arzela's theorem, it is equivalent to the condition that \mathcal{F} is a normal family (it is a family of which any sequence of functions contains a subsequence that converges uniformly on any compact set.)

Now for any rational function f, z is said to be **normal** with respect to f if there exists a neighborhood of $z \in \overline{\mathbb{C}}$ such that $\{f^n|_U ; n \ge 0\}$ is a equicontinuous. The Julia set of f is defined as

 $J_f = \{z \in \overline{\mathbf{C}} ; z \text{ is not normal with respect to } f\}$

and $F_f = \overline{\mathbf{C}} - J_f$ is called the **Fatou set**.

Critical Point

A point z is called a critical point if f is not injective on any neighborhood of z. When $z \neq \infty$ and $f(z) \neq \infty$, z is a critical point if and only if f'(z) = 0. And multiplicity as the number of the solutions of a equation f'(z) = 0 is called the multiplicity of z (as the critical point).

By an easy calculation, we can see that a rational function with degree d has exactly 2(d-1) critical points, counted with multiplicity. These critical points play very important rolls when we characterize the behavior of dynamical systems.

basin

Let α be a (super-)attracting fixed point for a rational function f(z). The set

$$\{z ; f^n(z) \longrightarrow \alpha , (n \to \infty)\}$$

is called the attracting basin of α . The immediate basin of α , denoted by $B(\alpha)$, is the connected component of attracting basin which contains α .

index of a fixed point

For f(z), we define the multiplicity of a fixed point z_0 as follows.

- If $f'(z_0) \neq 1$, multiplicity of z_0 is 1.
- If $f'(z_0) = 1$, then the Taylor expansion of f at z_0 is

$$f(z) = z_0 + (z - z_0) + a(z - z_0)^m + \cdots, \ a \neq 0.$$

In this case, multiplicity is defined by m.

We define the holomorphic index ι of a fixed point z_0 for f as follows([6]):

$$\iota(f; z_0) = \frac{1}{2\pi i} \int_{|z-z_0|=\delta} \frac{1}{z-f(z)} dz = \operatorname{Res}\left(\frac{1}{z-f(z)}; z_0\right).$$

Theorem 2.1 (Milnor) ([6]) For a rational function $f(z) \neq z$, we have

$$\sum_{f(z)=z}\iota(f\,;\,z)=1.$$

If $z_0 = f(z_0), f'(z_0) = \lambda \neq 1$ then $\iota(f; z_0) = \frac{1}{1 - \lambda}$.

Since a repelling or a certain neutral periodic points belong to the Julia set, this theorem yields the following:

Corollary 2.2 The Julia set for a rational function of degree two or more is always non empty.

3 Characterization of rational functions to be a Newton map

For a polynomial p(z), we define Newton map as follows.

$$N_p(z) = z - \frac{p(z)}{p'(z)}$$

It is clear that if p(z) has n distinct roots then N_p is a rational function of degree n.

The following facts concerning to an immediate basin $B(\alpha)$ of Newton map N_p are known, where α is a root of p.

- 1. An immediate basin $B(\alpha)$ is simply connected. ([7],[8])
- 2. ∞ lies on the boundary of $B(\alpha)$ for every root α of p(z). ([4])
- 3. If the local degree of $N|_{B(\alpha)}$ is s, then $B(\alpha)$ approaches ∞ in s-1 different directions.
 - ([7])

We shall characterize rational functions that are conjugate to a Newton's map as applied to a polynomial map. Namely, there is a complete criterion for a rational function to be a Newton's method.

Lemma 3.1 Let p be a polynomial of degree d, and N_p Newton map for p.

- 1. The set of fixed points of $N_p(z)$ is $\{\infty\} \cup p^{(-1)}(0)$.
- 2. If α is a root of p(z) with multiplicity m then $N_p'(\alpha) = \frac{m-1}{m}$. Hence a root of p is attracting fixed point of N_p . Especially if α is simple root of p(z) (i.e. m = 1) then α is super-attracting fixed point of N_p .
- 3. ∞ is the unique repelling fixed point of N_p , and its eigen value is $\frac{d}{d-1}$.

Theorem 3.2 The next two statements are equivalent for a rational function f of degree d.

1. f has distinct d fixed points, z_1, z_2, \cdots, z_d , whose eigen values are given as

$$f'(z_i) = \frac{m_i - 1}{m_i}, \ m_i \in \mathbf{N}, \ i = 1, \ \cdots, \ d$$

2. There exists a polynomial p for which Newton map N_p is conjugate to f.

Proof of Theorem

It is well known and easy to check that 2. \Rightarrow 1. Now we shall only to prove that 1. \Rightarrow 2. Note that f(z) has precisely d+1 fixed points. Let ζ be a fixed point $\neq z_i$ $(i = 1, \dots, d)$. The multiplicity of the fixed point ζ is 1, because that multiplicity of each fixed point z_i $(1 \le i \le d)$ is 1.

Put
$$k = \sum_{i=1}^{d} m_i$$
. We obtain that $f'(\zeta) = \frac{k}{k-1}$ from the equation in Theorem 2.1:

$$\sum_{z=f(z)} \iota(f;z) = \sum_{i=1}^{d} \frac{1}{1 - \frac{m_i - 1}{m_i}} + \frac{1}{1 - f'(\zeta)} = 1.$$

Hence it turns out that ζ is a repelling fixed point.

If $\zeta \neq \infty$ then by change of coordinate, ζ is transformed to ∞ , and f to a conjugate rational function \tilde{f} . We denote \tilde{f} by same f for notational simplicity.

Hence we can write $f(z) = \frac{q(z)}{r(z)}$ where q(z) and r(z) are polynomial of deg q = dand deg r < d. Let

$$f(z) - z = \frac{q(z) - z \cdot r(z)}{r(z)}.$$

The polynomial $q(z) - z \cdot r(z)$ of degree less than d has fixed points z_1, \dots, z_d . Therefore

$$q(z) - z \cdot r(z) \equiv 0$$
 or $q(z) - z \cdot r(z) = c \cdot \prod_{i=1}^{d} (z - z_i).$

From $f(z) \not\equiv z$, the first case does not occur. Then $q(z) - z \cdot r(z) = c \cdot \prod(z - z_i) = c \cdot s(z)$. Hence

$$f(z)-z=\frac{c\cdot\prod(z-z_i)}{r(z)}=\frac{c\cdot s(z)}{r(z)}.$$

For each $i \ (1 \leq i \leq d)$,

$$f'(z_i) = \left(z + \frac{c \cdot s(z)}{r(z)}\right)' \Big|_{z=z_i} = 1 - \frac{1}{m_i}$$

Hence

$$\left. \left(\frac{c \cdot s(z)}{r(z)} \right)' \right|_{z=z_i} = \frac{c \cdot s'(z_i)}{r(z_i)} = -\frac{1}{m_i}$$

So we have that $r(z_i) = -c \cdot m_i \cdot s'(z_i)$.

Let

$$t(z) = r(z) - \sum_{i=1}^{d} m_i \cdot \prod_{i \neq j} (z - z_j).$$

Then t(z) is a polynomial of degree d-1, satisfying that $t(z_i) = 0$ for each $i (1 \le i \le d)$,. Therefore t(z) must be constant 0.

Hence

$$r(z) = \sum_{i=1}^d m_i \cdot \prod_{i \neq j} (z - z_j).$$

Put $P_0(z) = c \cdot \prod_{i=1}^d (z - z_j)^{m_i}$. Then we have

$$\begin{split} f(z) &= z - \frac{c \cdot \prod_{i=1}^{d} (z - z_j)}{\sum_{i=1}^{d} m_i \cdot \prod_{i \neq j} (z - z_j)} \cdot \frac{\prod_{i=1}^{d} (z - z_i)^{m_i - 1}}{\prod_{i=1}^{d} (z - z_i)^{m_i - 1}} \\ &= z - \frac{P_0(z)}{P'_0(z)} = N_{P_0}(z). \end{split}$$

This theorem covers the result by J.Head as a corollary:

Corollary 3.3 (Head) ([3]) Any rational function f of degree d having d distinct supper-attracting fixed points is conjugate to the N_P for a polynomial P of degree d. Moreover if ∞ is not supper-attracting for f, but fixed point, then $f = N_P$ for some polynomial P of degree d.

4 Sutherland's estimate for the basins of Newton's method

The problem of devising optimal methods for finding numerically approximate zeros of a polynomial has been of interest and is now far from solved. In order to have an estimate on the complexity of a root-finding algorithm, we need a compactness condition under a suitable norm on the space of polynomials. This can be done by placing conditions either of the location of the roots or of the coefficients. Hence we consider hereafter a polynomial in the family $\mathcal{P}_d(1)$:

$$\mathcal{P}_d(1) = \{ p(z) = z^d + a_{d-1} z^{d-1} + \dots + a_0, \quad |a_i| \le 1 \ (i = 0, 1, \dots, d-1) \}$$

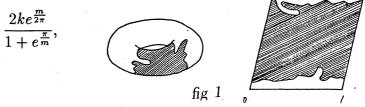
Moreover if $a_{d-1} = 0$ then p(z) is called **centered polynomial**. It is easy to see that the roots of a polynomial in $\mathcal{P}_d(1)$ lie within the disk with center 0 and radius 2. It is possible to transform an arbitrary polynomial into an element of $\mathcal{P}_d(1)$: for $p(z) = \sum_{j=0}^d c_j z^j$, let $C = \max\left\{\frac{c_j}{c_d}\right\}$, then $q(z) = \frac{p(Cz)}{C^d} \in \mathcal{P}_d(1)$. Note that the Newton map N_p induced by p

is conjugate to N_q induced by q. Therefore, withoutloss of generality, we can treat only centered polynomials $p \in \mathcal{P}_d(1)$, after such conjugacy.

Newton's method converges very fast in a neighborhood of a simple root, but can fail in for some initial point outside this neighborhood. Let $B(\alpha)$ be an immediate basin of attracting fixed point α (a root of p) of N_p . S.Sutherland ([11]) attempted how one can gurantee convergence of Newton's method, by estimating the "width" of $B(\alpha)$. But unfortunately there are ambigous parts and errors in his paper([11]). In this section we shall correct his results. Especially Proposition 3.5 in ([11]) is reformed as Proposition 4.2

Definition 4.1 Any annulus can be mapped by an analytic diffeomorphism onto a unique "standard annulus" whose inner boundary is the unit circle and with outer boundary the circle of radius $e^{2\pi m}$ for some $m \in \mathbb{R}^+$. In this case, the modulus of the annulus is said to be m.

Proposition 4.2 Let T be a torus, isomorphic to $C/(Z \oplus Z\tau)$, and A an annulus with modulus(A) = m, contained in T (see fig 1). Then the distance between the boundary curves of A is at least



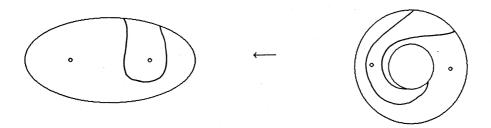
where $k = \min\{1, \Im(\tau)\}.$

Proof of Proposition Consider an open ellipse whose major axis is the interval $\left(-\frac{r+1/r}{2}, \frac{r+1/r}{2}\right)$ and minor axis is $\left(-\frac{r-1/r}{2}i, \frac{r-1/r}{2}i\right)$. Remove two points -1, 1 from the ellipse and denote by E the resulting set. Let Γ be the set of curves in E which join the boundary of the ellipse passing through the interval (-1, 1).

The map $z \mapsto \frac{z+1/z}{2}$ is two to one. And a punched annulus

$$PA = \{z \, ; \, \frac{1}{r} < |z| < r\} - \{-1, 1\}$$

is mapped to E. Let Γ' be the set of curves in PA joining inner boundary and outer boundary.



Then the extremal lengths are calculated as follows:

$$\lambda(\Gamma') = rac{1}{\pi}\log r \;, \;\; ext{and} \;\; \lambda(\Gamma) = rac{2}{\pi}\log r \;.$$

We may assume that the "narrow part" of this embedded annulua is located at the center of T. Let δ be the width of the narrow part of E. Scale the ellipse by $\frac{\delta}{2}$ and embed it in T so that the interval $\left[-\frac{\delta}{2}, \frac{\delta}{2}\right]$ corresponds to the narrow part.

Let

$$r = \frac{k}{\delta} + \sqrt{\left(\frac{k}{\delta}\right)^2 - 1},$$

where $k = \min\{1, \Im(\tau)\}.$

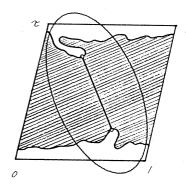
Let Γ_A be the family of closed curves in A. Then $\lambda(\Gamma_A) = \frac{1}{m}$.

Since each curve in Γ_A contains a curves in Γ , we have

$$\frac{1}{m} \ge \lambda(\Gamma) = \frac{2\log r}{\pi} = \frac{2\log\left(\left(\frac{k}{\delta}\right) + \sqrt{\left(\frac{k}{\delta}\right)^2 - 1}\right)}{\pi}$$

Solving for δ , we obtain

$$\delta \ge \frac{2k \exp(\frac{\pi}{2m})}{1 + \exp(\frac{\pi}{m})}.$$

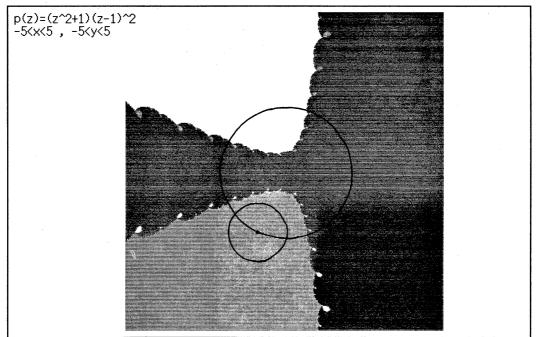


By estimating the "width" of immediate basin of a root of $p(z) \in \mathcal{P}_d(1)$, we have the following revised results, essentially due to S.Sutherland ([11]).

Theorem 4.3 (Sutherland) Let α be a root of multiplicity m, with $N|_{B(\alpha)}$ of degree s + 1. Then there are points $t_1, \dots t_s$ of magnitude $2 + 2\sqrt{2}$ for which a disk of radius r_i centered at t_i lies entirely within $B(\alpha)$. These radii satisfy $\sum_{i=1}^{s} r_i \geq \frac{(2+2\sqrt{d})\pi}{12d(1+\sqrt{\frac{2m}{2m-1}})}$.

Corollary 4.4 Let p(z) be a centered polynomial in $\mathcal{P}_d(1)$, and $|z| \ge 2 + 2\sqrt{d}$. Then the probability that $N^n(z)$ will converge to a root of p is at least $\frac{1}{29\pi d}$.

Corollary 4.5 Let p(z) be a centered polynomial in $\mathcal{P}_d(1)$. Let t_1, \dots, t_n be points equally spaced around the circle of radius $2 + 2\sqrt{d}$, where $n \ge 29\pi d(d-1)$. Then for each root α_i of p(z), at least one of the points t_j lies in $B(\alpha_i)$.



5 Computer experiments with Newton's method

In this section we study Newton's method for cubic polynomials using computer graphics. We forcus on the Fatou set of N_p containing basins of an attracting periodic point of period two or more. If such periodic points exist, they do not correspond to roots of p, and Newton's method with any initial point in such a basin fails to converge. The following theorem plays central role:

Theorem 5.1 (Fatou) ([2]) The immediate basin of an attracting periodic point contains at least one critical point.

Thus we only follow the orbits of free critical points of N_p .

We consider hereafter one-parameter family, called Milnor model, { $p_{\mu} = z^3 - 2z + \mu$ ($\mu \in \mathbb{C}$)}. Note that $p_{\mu} \notin \mathcal{P}_d(1)$. $N_{p_{\mu}}$ has only one free critical point 0. Note that the Newton map for an arbitrary cubic polynomial with more than one root can be conjugated to $N_{p_{\mu}}$ for some μ .

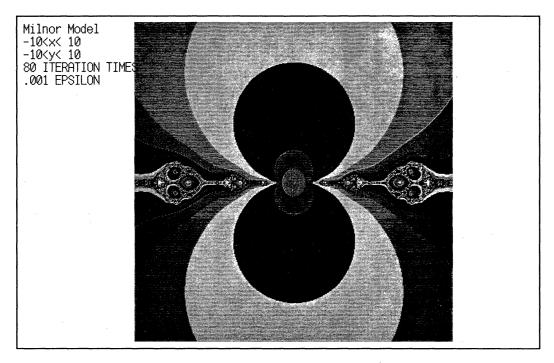
In the next figures 5.1 and 5.2 (enlargement), we plot a white point at a parameter whose associated Newton map fails to converge if the free critical point 0 is a starting point.

We observed in these parameter plane that there are copies of Mandelbrot set \mathcal{M} (see figure 5.3) :

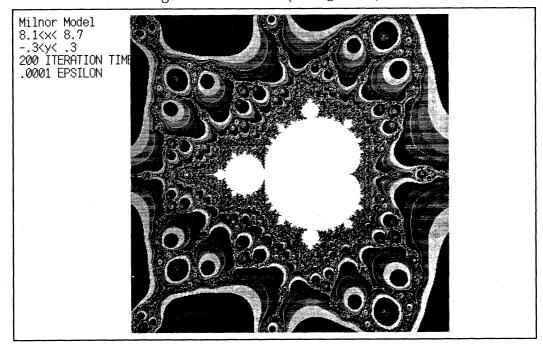
$$\mathcal{M} = \{ c \in \mathbf{C} : z_0 = 0, z_{n+1} = z_n^2 + c \not\to \infty \}$$

The figure 5.4 is filled-in Julia set of quadratic map $z^2 + c$ where c is in Mandelbrot set. The figure 5.5 is Fatou set for a value μ chosen from white region in Figure 5.2.

To the question why we obtain images of the Mandelbrot set or of a filled-in Julia set of quadratics, A. Douady and J. H. Hubbard ([1]) answered by using the theory of polynomial-like mapping. Roughly speaking, Newton map $N_{p_{\mu}}$ locally may behave as a quadratic polynomial.



figures 5.1 and 5.2 (enlargement)



Handelbrot set. -2 < Re Z < 1.25 -1.25 < Im Z < 1.25

figure 5.3

Mandelbrot set \mathcal{M}

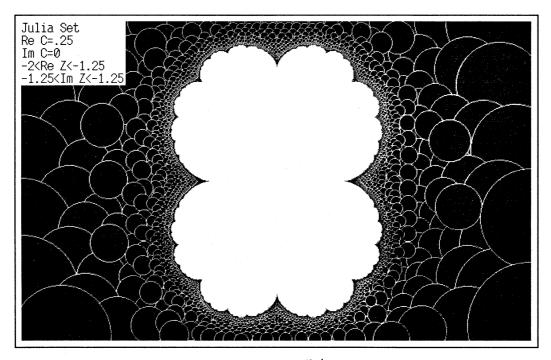


figure 5.4

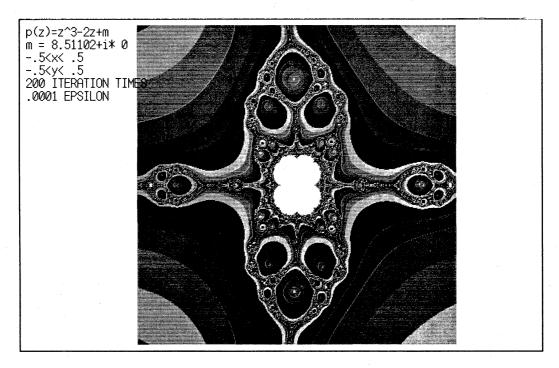


figure 5.5

References

- A.Douady & J.Hubbard. On the Dynamics of Polynomial-like mappings, Ann. Sci. Ec. Norm. Sup., 4e série, t.18(1985), 285-343
- [2] P.Fatou. Sur les Equations fonctionnelles, Bull. Soc. Math. France, 47(1919), pp. 161-271, 48(1920), pp. 33-94, pp. 208-304.
- [3] J.Head. The combinatorics of Newton's method for cubic polynomials, Ph.D. thesis, Cornell Univ. Ithaca, U.S.A., 1987
- [4] A.Manning. How to be Sure of Solving a Complex Polynomial using Newton's Method, preprint, Math.Inst., Univ. Warwick, 1986
- [5] C.McMullen. famillies of Rational Maps and Iterative Root Finding Algorithms, Ann. of Math., 125(1987), pp. 467-493
- [6] J.Milnor. Dynamics in one complex variables: Introductory lectures, Preprint # 1990/5, SUNY Stony Brook, 1990
- [7] F.Przytycki. Remarks on the Simple Connectedness of basins of Sinks for Iterations of Rational Maps, Dynamical Systems and Ergodic theory, PWN-Polish Scientific Pub., 1989, pp. 229-235
- [8] M.Shishikura. The connectivity of the Julia set of rational maps and Fixed points, IHES/M/90/37, I.H.E.S., Bures-sur-Yvette, France, 1990
- [9] M.Shub & S.Smale. On the Existence of Generally Convergent Algorithms, J. of complexity, 2(1986), pp.2-11.
- [10] S.Smale. On the efficiency of algorithms of analysis, Bull. A.M.S., 13(1985), pp. 87-121
- [11] S.Sutherland. Finding Roots of Comple Polynomials with Newton's Method, Preprint # 1990/7, SUNY Stony Brook, 1990