Finding Roots of Complex Polynomials with Newton's Method

SCOTT SUTHERLAND

Institute for Mathematical Sciences
SUNY Stony Brook

A mildly revised version of the author's doctoral dissertation
Submitted to Boston University Graduate School, May 1989
Revised September 1989

Abstract

Newton's method is one of the most widely known numerical algorithms for finding the roots of smooth functions of one variable. However, it is also considered to be "unpredictable" in that many initial values for the method do not converge to a root of the function. We will show that, for polynomial functions, if the initial values are chosen properly convergence to the roots can be guaranteed.

More specifically, Newton's method for a polynomial yields a rational function of the Riemann sphere to itself, and we can apply techniques of complex dynamics to analyze its behavior. One of the results is that the immediate basins of the roots are large and have a definite "width" which we can bound from below. For a root of a polynomial of degree d, this width is of order $1/d$. We present an algorithm for choosing starting values for Newton's method which will find all of the roots (to within ϵ) in presumably no more than order $d^3 \log(\frac{1}{\epsilon})$ evaluations.

We also present some experimental results concerning the relationship between the "Newton flow" for a polynomial, the relaxed Newton's method, and Newton's method. Some conjectures are made about how one might find polynomials which cause problems for Newton's method and related methods.

Acknowledgments

I would like to express my gratitude to my thesis adviser, Paul Blanchard. He has been extraordinarily helpful in my research into Newton's method, as well as introducing me to the subject of Dynamical Systems, in particular Complex Dynamics. Much of what I know about dynamics was taught to me by him. Finally, he served as Director of Graduate Studies for much of my tenure as a graduate student. Only he can fully appreciate the implications of that.

Unbounded thanks go to Dick Hall for his constant willingness to discuss mathematics (or anything else). His advice to me has been invaluable; I view him as a great friend and colleague.

I would like to thank my friends and long-term office mates, Greg Buck and Robert Winters, for the friendship they have provided and the many discussions we have had over the years.

I would also like to acknowledge the assistance I have received from helpful discussions with Bob Devaney, Mitsuhiro Shishikura, Adrien Douady, Joel Friedman, Anthony Manning, Mike Shub, Steven Smale, Mike Hurley, Bruce Peckham, and Alec Norton. These mathematicians may not realize how much they have assisted and encouraged me, but without their help, I probably would still be working on this.

My wife Beth has provided support, love, and friendship. Without her assistance, I doubt I could have accomplished anything.

Table of Contents

Addendum to
Finding Roots of Complex Polynomials with Newton's Method

We can now improve the major result in chapter 3 and simultaneously simplify the proof. This simplification makes the "tedious, but necessary, calculations" (section 3.7) no longer necessary (although still tedious). The main change is that we estimate the width of $\mathbf{B}(\alpha)$ near a circle of moderate radius directly, instead of obtaining estimates near infinity and pulling them in.

Note that if $R \geq 2\left(\frac{d+1}{d-1}\right)$, then the image of the circle of radius R lies outside of $\mathbf{D}_2$. Denote the annulus bounded by this circle and its image by A_R. By Proposition 1.9, we can find an annulus A around ∞ which will be mapped univalently onto A_R by some iterate of N. Furthermore, $\mathbf{B}(\alpha) \cap A_R$ is isomorphic to an annulus (in the torus obtained by identifying the boundary curves of A_R under the action of N). This annulus has the same modulus as the sector $\mathcal{V}$, namely $\frac{\pi}{\log(M'(\xi))}$.

In order to estimate the width of the "thinnest part" of $\mathbf{B}(\alpha)$ lying in A_R, we apply the argument of section 3.5 (almost). First, note that $\log(A_R)$ is a "rectangle" with 3 straight sides. The "wiggly side" corresponds to the inner boundary of A_R. As a result of Lemma 1.7, the inner boundary of A_R lies inside the disk of radius $R - \frac{R-2}{d}$, and so the "wiggly rectangle" contains a straight-sided rectangle of height 2π and width $\log\left(\frac{Rd}{Rd-R+2}\right)$. We now use the construction in section 3.5 to conclude that the thinnest part of $\log\left(\mathbf{B}(\alpha) \cap A_R\right)$ is at least

$$\frac{2\pi\sqrt{M'(\xi)}}{1+M'(\xi)} \log \frac{Rd}{Rd-R+2}.$$

Now apply the same argument as in Theorem 3.6 to obtain

THEOREM 1. *If $R \geq 2\left(\frac{d+1}{d-1}\right)$, there is a point $t_\xi \in \mathbf{B}(\alpha)$, with $|t_\xi| = R$, which is the center of a disk of radius at least*

$$R \tanh\left(\frac{\pi\sqrt{M'(\xi)}}{1+M'(\xi)} \log \frac{Rd}{Rd-R+2}\right) \geq \frac{2\pi(R-2)}{3d\left(1+\sqrt{M'(\xi)}\right)}$$

contained entirely in $\mathbf{B}(\alpha)$. (The point ξ is the appropriate fixed point of the map M.)

The rest of the argument is unchanged except for the constants, so we obtain our major result:

Theorem 2. *Choose $R \geq 2\left(\frac{d+1}{d-1}\right)$ and let α be a root of multiplicity m, with $N|_{\mathbf{B}(\alpha)}$ of degree s. Then there are points $t_1, \ldots, t_s$ of magnitude R for which a disk of radius r_i centered at t_i lies entirely within $\mathbf{B}(\alpha)$. These radii satisfy*

$$\sum_{i=1}^{s} r_i \geq \frac{2\pi(R-2)}{3d\left(1+\sqrt{\frac{2m}{2m-1}}\right)}.$$

Note the lower bound on R decreases with d, so we may always use $R = 4$. Using this, we have the immediate result that the immediate basin of any root α takes up at least $\frac{1}{6\sqrt{3}}$ of the circumference of the circle of radius 4. We restate two immediate consequences here:

Corollary 3. *Let $p(z)$ be a centered polynomial in $\mathcal{P}_d(1)$, and $|z| \geq 4$. Then the probability that $N^n(z)$ will converge to a root of p is at least $\frac{1}{6d\sqrt{3}}$.*

Corollary 4. *Let $p(z)$ be a centered polynomial in $\mathcal{P}_d(1)$. Let $t_1, \ldots, t_n$ be points equally spaced around the circle of radius 4, where $n \geq 11d(d-1)$. Then for each root α_i of $p(z)$, at least one of the points t_j lies in $\mathbf{B}(\alpha_\mathbf{i})$.*

1. Introduction

Newton's method is one of the simplest, most widely known methods for approximating the roots of smooth functions of one variable. It is often taught in first semester calculus courses as a nice application of the derivative.

Briefly stated, the method (for a function with real coefficients) is this: Suppose we want to find a root of the equation

$$p(x) = 0.$$

Make an initial guess x_0, and then approximate the function $p(x)$ by its tangent line at x_0. Let x_1 be the point where the tangent line meets the x–axis. If x_1 is not a sufficiently good approximation to a root, we repeat the process with x_1 as the next guess. See Figure 1.1.

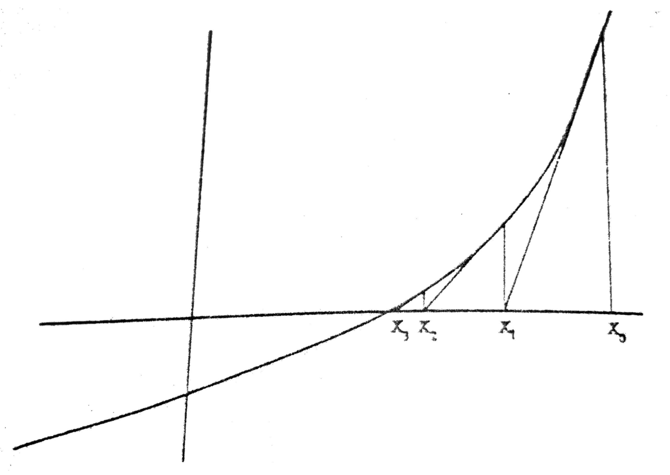

FIGURE 1.1: Newton's method for a real polynomial.

This geometric idea yields the iteration scheme

$$x_{i+1} = x_i - \frac{p(x_i)}{p'(x_i)}.$$

We can perform the same iteration where p is a complex function, and our initial value z_0 is also complex. In fact, we shall see that it is often preferable to think of the process as occurring in $\mathbf{C}$, even for a real function. We can think of Newton's method as iterating the map

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

If p is a polynomial (a restriction we will make from here on), the resulting Newton map N_p is a rational map of the Riemann Sphere $\overline{\mathbf{C}} = \mathbf{C} \cup \{\infty\}$ to itself. This fact enables us to analyse Newton's method using techniques of complex dynamics.

The main theme of this paper is “How can we be assured of finding all the roots of a polynomial with Newton’s method?”

1.1 Some History

Newton’s method is almost certainly due in part to Newton himself, although in a rather different form. According to [Cj], Newton first explained his method for approximating the real roots of a polynomial in *De analysi per aequationes numero terminoriaum infinitas* in 1669, although it was not published until 1704. He also gave essentially the same explanation in *Methodus fluxionum et serierum infinitarum*, which was scheduled to be published in 1671, but not printed until 1736. The first printed account of the technique actually appeared in Wallis’ *Algebra* in 1685.

Newton found an approximate root of the cubic

$$y^3 - 2y - 5 = 0$$

by setting $y = 2 + p$ to obtain

$$p^3 + 6p^2 + 10p - 1 = 0.$$

He ignores the higher powers of p to get the next approximation $p = 0.1 + q$. Making this substitution gives

$$q^3 + 6.3q^2 + 11.23q + 0.061,$$

or $q = -0.0054 - r$. He repeats this process once more to obtain $r = -0.00004854 + s$, which yields an approximate root

$$y = 2 + 0.1 - 0.0054 - 0.00004854 = 2.09455147.$$

Newton was apparently aware that the process might fail to converge, but he fails to go any deeper into that question. He gives no other examples of the method for finding roots. He does, however, exploit the method to find series solutions for implicitly defined functions, and to compute the infinite series for $\sin x$ and $\cos x$ by inverting the series for $\arcsin x$ and $\arccos x$! (See [E].)

Note that Newton derived a new equation from each successive approximation, instead of iterating the “Newton function” as is done today. This iteration of $x - \frac{p(x)}{p'(x)}$ was a modification made by Joseph Raphson in 1690. Raphson’s reformulation greatly simplifies the computational process; it is for this reason that Newton’s method is often called as “Newton–Raphson iteration”.

Later (1740), Thomas Simpson discussed the extension of the Newton–Raphson method to irrational and transcendental equations. He did not mention Newton or Raphson, however, and says his procedure was “a new method”.

Joseph Fourier is generally credited with providing sufficient (but not necessary) conditions for convergence of the iterates to a root in 1831. However, his results were anticipated by J. R. Mourraille in 1768. Both give the geometric condition that convergence is assured if the initial approximation x_0 is chosen so that the polynomial is concave up on the interval between the root and x_0.

In 1879, Arthur Cayley introduced what he calls the "Newton–Fourier Imaginary Problem". Here, "throwing aside the restrictions as to reality", he suggests applying Newton-Raphson iteration to a complex rational function $p(z)$, and investigating what the fate is of an arbitrary point in $\mathbf{C}$. He stated in [Ca] and [Ca1] that "the solution is easy and elegant in the case of a quadratic equation", but gave few details. He also said that the solution for the cubic "appears to present considerable difficulty".

In [Ca3] and [Ca4], Cayley gave a more detailed explanation of the solution of the quadratic case, which in modern language is essentially this: For a quadratic polynomial $p(z)$, the Newton map $N_p(z)$ is conjugate to the map $z \mapsto z^2$ via a Möbius transformation. That is, we can make a change of variables so that iterating N_p behaves just like iterating z^2. This change of variables sends one root to 0 and the other to ∞, and maps the perpendicular bisector of the segment joining the roots to the unit circle. The points on the bisector never converge to a root, and all other points converge to the root which lies nearest. See Figure 1.2. More detailed explanations of Cayley's result appear in [PSH], [PR] and [S2].

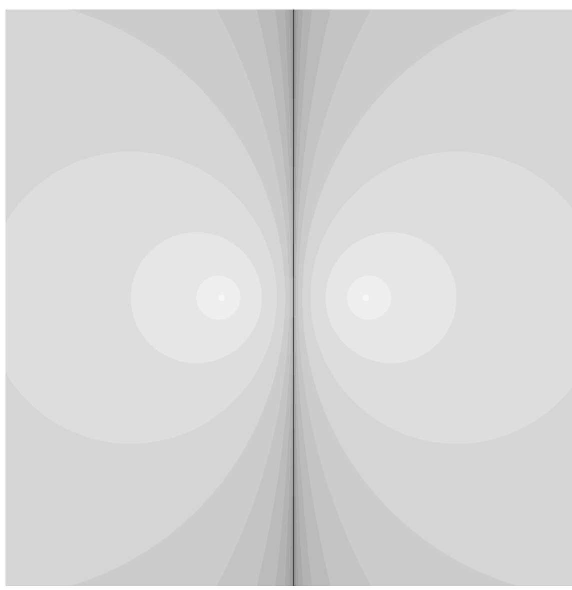

FIGURE 1.2: The convergence of Newton's method for $p(z) = z^2 - 1$. The gray level indicates the speed of convergence to a root, with lighter colors converging in fewer iterations. The black line in the middle is the imaginary axis, which does not converge to a root.

Until recently, little was known about the global convergence of Newton's method. It was known that quadratic convergence occurs in a neighborhood of a simple root, and some of the behavior for special cases was understood. In the middle of this century, Barna [Ba] showed that for a polynomial with all roots real, Newton's method converges to a zero on all of $\mathbf{R}$ except a Cantor set.

Due to the lack of assured convergence, the general feeling among the users of the method was that the method works well if one has a good idea of where a root might be. If not, just pick a point at random, and if that fails, pick another. This is still common practice today.

In an attempt to overcome the difficulties with Newton's method, a large number of variants of the method have been introduced. For example, one technique is to use a slower, but "surer" method to get near a root, and then apply Newton's method. Another common modification is the *relaxed Newton's method*, which consists of iterating

$$N_{h,p}(z) = z - h\frac{p(z)}{p'(z)}.$$

For $h \neq 1$, the relaxed Newton's method does not have quadratic convergence in a neighborhood of a simple root, although it may near a multiple root (for appropriate h). Smale [S1], [S3] has studied the efficiency of such methods extensively, and Kim [K1] has computed the topological complexity of such algorithms. Flexor and Sentenac [FS] have recently shown that there is a choice of $h \in \mathbf{C}$ (depending on p) such that $N_{h,p}$ converges almost everywhere. Kim [K2] has also examined the behavior of the relaxed Newton's method where the parameter h is variable. Because of its intimate connection with the regular Newton's method, we will return to a discussion of the relaxed Newton's method in Chapter 4.

Recently, there has been a great resurgence of interest in the behavior of numerical algorithms, especially in viewing them as dynamical systems (See [Sh], [S2], [SS1], [SS2], [S3], [S4], and others). In particular, N_p is a rational map of $\overline{\mathbf{C}}$, and the theory of complex dynamical systems, originated by Fatou ([F1], etc.) and Julia ([J]), can be applied. Cayley's work on Newton's method may have been one of the original motivations for the development of complex dynamics by Fatou and Julia.

Iterating N_p is a special case of a *purely iterative algorithm*, as introduced in [S2]. It is not *generally convergent* for polynomials of degree $d > 2$. That is, there are open sets of polynomials for which an open set of initial conditions fail to converge. We give examples of such polynomials in a later section. McMullen has shown that there is no generally convergent purely iterative algorithm for polynomials of degree $d \geq 4$, as well as constructing a modification of Newton's method which is generally convergent for cubics; see [Mc1] and [Mc2]. In [DM], Doyle and McMullen give an algorithm for solving a quintic polynomial by iteration; this algorithm uses a "tower of purely iterative algorithms".

Newton's method fails on an open set when there is an attracting periodic orbit. Curry, Garnett, and Sullivan [CGS] have several computer studies of this case for a family of cubic polynomials. Here, one sees the appearance of Mandelbrot sets in the parameter plane; this is explained by the theory of polynomial–like mappings as developed in [DH]. (See [Bl], [Ma], [DH1], [DH2], [PR], and others for a discussion of the Mandelbrot Set.) Hurley [Hu] has shown that there can be as many as $d-2$ such periodic attractors. See also [HM], [Hu2], and [SU].

Let p be a polynomial in the family $\mathcal{P}_d(1)$ (defined below) and define $\mathcal{B}_p$ to be the set of points which tend to a root of p under iteration of N_p. In [S2], Smale asks (Problem 7A) for a lower bound on the area of $\mathcal{B}_p \cap \mathbf{D}_2$. Friedman [Fr] gives a bound which tends to 0 as $d \to \infty$. It is not known if there is a nonzero lower bound for this area which is independant of d.

In [M], Manning gives an algorithm for finding the roots of a polynomial with Newton's method. He does this by estimating the size of the immediate basin of attraction for an "exposed root" α as it crosses the region between the circle of radius d and its preimage. A root is *exposed* if it lies on the convex hull of the roots, and the interior angle of the hull is not too great. There are always at least 2 exposed roots; to obtain the interior roots, the polynomial is deflated (that is, the roots already determined are divided out) and the process is repeated. Several of the techniques in our Chapter 3 are quite similar to those in [M].

1.2 Basic Complex Dynamics

In this section, we review some basic facts from complex dynamics. A good reference for this material is [Bl].

The field of complex dynamics began with work done in the early part of this century by G. Julia [J] and P. Fatou [F1-4]. Its main concern is understanding the behavior of points under iteration of an analytic function $f : \overline{\mathbf{C}} \to \overline{\mathbf{C}}$, where $\overline{\mathbf{C}}$ is the one-point compactification of $\mathbf{C}$ obtained by adding a point at ∞. The Riemann surface $\overline{\mathbf{C}}$ is usually referred to as the *Riemann Sphere*.

We denote the n–fold composition of f with itself by f^n. The *orbit* of a point z is the sequence $\{f^n(z)\}_{n=0}^{\infty}$. A point z is said to be a *periodic point* if $f^n(z) = z$; such a point is called a *fixed point* if $n = 1$.

By the chain rule, the derivative along a periodic orbit of least period n is $(f^n)'(z) = \prod_{i=0}^{n-1} f'(f^i(z))$. This gives the total amount of local expansion along the orbit. A point z (resp. orbit) is said to be *attracting* if the modulus of its derivative is less than 1; if the modulus is greater than 1, the point (resp. orbit) is called *repelling*. A point (resp. orbit) with derivative 0 is called *superattracting*; in a neighborhood of a superattracting orbit, f contracts dramatically.

DEFINITION. Let α be an attracting fixed point for the map $N(z)$. Then the set

$$\{z : \; N^n(z) \to \alpha \text{ as } n \to \infty\}$$

is the *basin of attraction* of α. The connected component of this set which contains α is called its *immediate basin*, which we shall denote $\mathbf{B}(\alpha)$. The basin of attraction of a periodic orbit can be defined similarly.

Definition. The ***Julia Set of*** f, denoted J_f, is the set of points $z \in \overline{\mathbf{C}}$ for which the family of iterates $\{f^n\}$ is not normal in any neighborhood of z.

The Julia set is often referred to as the chaotic set; it is where all the complicated dynamics occur.

Proposition 1.1.

(i) $J_f = \overline{\{\text{repelling periodic points of } f\}}$.

(ii) *J_f is a closed, perfect, non-empty set without interior.*

(iii) *If A is the basin of attraction of an attracting orbit, then $J_f = \mathrm{Fr}(A)$.*

Theorem 1.2. *The immediate basin of a periodic attractor of a rational map f must contain at least one critical value of f. Moreover, if every critical point of f lies in some basin of a periodic attractor, then these basins have full measure in $\overline{\mathbf{C}}$.*

1.3 Newton's Method: Mathematical Preliminaries

Here we give the preliminary facts about Newton's method that we will need in the sequel.

For a polynomial $p(z)$, we define the Newton map to be

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

We will usually omit the subscript when there is no confusion about which polynomial we are working with. The following proposition lists some of the basic properties of N, which are easily verified.

Proposition 1.3.

(i) *If $p(z)$ has n distinct roots, then $N_p(z)$ is a degree n rational map .*

(ii) *If $q(z) = p(az+b)$, then N_p is conjugate to N_q by the map $z \mapsto az+b$.*

(iii) *The fixed points of N_p are exactly the roots $\alpha_1, \ldots, \alpha_d$ of p and ∞.*

(iv) *The derivative at z is given by* $N_p'(z) = \dfrac{p(z)p''(z)}{[p'(z)]^2}$.

(v) *Infinity is the only fixed repeller of N_p, with derivative $\frac{d}{d-1}$.*

(vi) *If α is a root with multiplicity m, then $N_p'(\alpha) = \frac{m-1}{m}$. Thus a simple root is a superattracting fixed point.*

Definition. We say that $p(z) \in \mathcal{P}_d(1)$ if

$$p(z) = z^d + a_{d-1}z^{d-1} + a_{d-2}z^{d-2} + \ldots + a_0$$

where $|a_i| \leq 1$. Such a polynomial is said to be ***centered*** if $a_{d-1} = 0$.

So that we have some initial idea of where the roots lie, we shall make the restriction that $p(z)$ is a centered polynomial in $\mathcal{P}_d(1)$. This normalization can be achieved by precomposing $p(z)$ with an affine map, which by (ii) has no qualitative effect on the behavior of N.

Proposition 1.4. *If $p(z) = \prod_{i=1}^{d}(z-\alpha_i)$ is a centered polynomial in $\mathcal{P}_d(1)$, then all of the roots α_i lie in the disk of radius 2 about the origin. Furthermore, the sum of the roots is 0.*

Proof. The fact that $|\alpha_i| < 2$ follows from the fact that $|z^d| > |p(z) - z^d|$ for $|z| > 2$. Since the coefficient of z^{d-1} is equal to the sum of the roots, this yields the second fact. $\square$

Now we give a geometric interpretation of Newton's method in the complex plane. Note that

$$\frac{p'(z)}{p(z)} = \frac{\sum_{j=1}^{d}\prod_{i\neq j}(z-\alpha_i)}{\prod_{i=1}^{d}(z-\alpha_i)} = \sum_{i=1}^{d}\frac{1}{z-\alpha_i}.$$

Thus

$$N(z) = z - \frac{1}{\sum_{i=1}^{d}\frac{1}{z-\alpha_i}}.$$

This gives us the size of the "step" that is taken by one iteration of the map as a function of the vectors to the roots. We shall use this fact to understand the behavior of N away from the roots of p. First, we state Lucas' Theorem, which tells us that the poles of N (which are the critical points of p) must lie inside the convex hull of the roots. We shall denote the convex hull of the roots by $\langle\{\alpha_i\}\rangle$.

Theorem 1.5. (Lucas, 1874) *Let $p(z)$ be polynomial with coefficients in $\mathbf{C}$. Then the zeros of p' lie inside the convex hull of the roots of p.*

Proof. Note that

$$p'(z) = p(z)\sum_{i=1}^{d}\frac{1}{z-\alpha_i}.$$

Suppose that $p'(\beta) = 0$ for some β outside $\langle\{\alpha_i\}\rangle$. But then all the vectors from β to the α_i lie in a half-plane through β, and so their inverses $\frac{1}{\alpha_i-\beta}$ also lie in a (possibly different) half-plane. But then these vectors cannot sum to zero, giving a contradiction. $\square$

The next lemma tells us that N moves points outside the convex hull of the roots toward it. The proof is similar to that of Lucas' Theorem (Theorem 1.5).

Lemma 1.6. *If z is outside $\langle\{\alpha_i\}\rangle$, then $N(z)$ lies on a ray emanating from z and passing through $\langle\{\alpha_i\}\rangle$.*

Proof. We shall show that the vector $N(z) - z$ lies in the wedge based at z and containing $\langle\{\alpha_i\}\rangle$. Note that each of the vectors $\alpha_i - z$ lie in this wedge, and so their inverses lie in a second wedge with the same angle. Since this angle is less than $\frac{\pi}{2}$, we have $\sum_{i=1}^{d} \frac{1}{z-\alpha_i}$ also lying in this second wedge. But this means that its inverse

$$\frac{1}{\sum\limits_{i=1}^{d} \frac{1}{\alpha_i - z}} = N(z) - z$$

lies in the original wedge. $\square$

For a point z of very large modulus, $N(z)$ is essentially the same as $(1 - \frac{1}{d})z$, since the difference between $z - \alpha_i$ and z is negligible. The following lemma gives explicit bounds on this difference.

Lemma 1.7. *If $|z| > 2$, then $N(z) \in \mathbf{D}_{\frac{2}{d}}(\frac{d-1}{d}z)$.*

Proof. For ease of calculation, we change coordinates by an affine isometry so that $z = 0$, and that the α_i lie in a disk of radius 2 about some point c on the real line, with $c > 2$. Let $\widehat{N}$ denote the map in this new coordinate system, and $\hat{\alpha}_i$ be the corresponding roots. Since $\hat{\alpha}_i \in \mathbf{D}_2(c)$,

$$\frac{1}{\hat{\alpha}_i} \in \mathbf{D}_{\frac{2}{c^2-4}}(\frac{c}{c^2-4})$$

Summing the d terms,

$$\sum \frac{1}{\hat{\alpha}_i} \in \mathbf{D}_{\frac{2d}{c^2-4}}(\frac{cd}{c^2-4})$$

and inverting,

$$\widehat{N}(0) = \frac{1}{\sum_{i=1}^{d} \frac{1}{\hat{\alpha}_i}} \in \mathbf{D}_{\frac{2}{d}}(\frac{c}{d})$$

Changing back to our original coordinates gives the desired result. $\square$

To make the estimate in the previous lemma, we approximated $\langle\{\alpha_i\}\rangle$ by $\mathbf{D}_2$. A similar result, which identifies a half-plane containing $N(z)$, can be obtained for all z outside of $\langle\{\alpha_i\}\rangle$. The argument is almost identical to that in Lemma 1.7, but uses a half-plane instead of $\mathbf{D}_2$. We omit the proof here (see [Fr] or [M]).

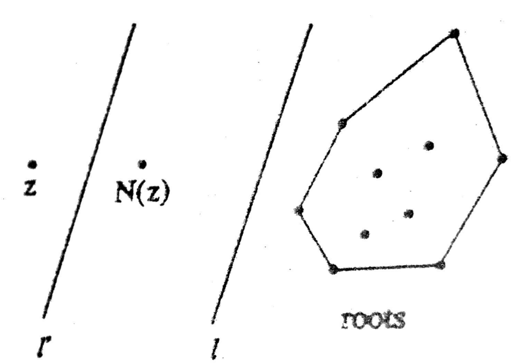

Figure 1.3: A pictorial version of Theorem 1.8.

THEOREM 1.8. *Let ℓ be a line separating z from the roots, and ℓ' be a line parallel to ℓ lying $1/d$ of the way between z and ℓ. Then ℓ' separates z from $N(z)$. (see Figure 1.3).*

PROPOSITION 1.9. ([M]) *There is an open set U in $\overline{\mathbf{C}}$, homeomorphic to a disk, such that $N|_U$ is a diffeomorphism from U onto $\overline{\mathbf{C}} - \langle\{\alpha_i\}\rangle$. For $z \notin \langle\{\alpha_i\}\rangle$, we have $(N|_U)^{-n}(z) \to \infty$ as $n \to \infty$.*

PROOF. The inverse images of ∞ are itself and the roots of p' that are not also roots of p. By Theorem 1.5, these lie inside $\langle\{\alpha_i\}\rangle$. Also, the critical points of N are the simple roots of p and the roots of p''; these must also lie in the convex hull by the same theorem.

The quotient space of $\overline{\mathbf{C}}$ by $\langle\{\alpha_i\}\rangle$ is homeomorphic to the sphere. Denote the image of $\langle\{\alpha_i\}\rangle$ under this identification by H. We define a map $\widehat{N}$ from this space to itself by

$$\widehat{N}(\hat{z}) = \begin{cases} H & \text{if } \hat{z} = H; \\ H & \text{if the straight line from } z \text{ to } N(z) \text{ crosses } \langle\{\alpha_i\}\rangle; \\ N(z) & \text{otherwise.} \end{cases}$$

The map $\widehat{N}$ is continuous because all points outside H move toward it (Lemma 1.6). Infinity is a regular point of the mapping $\widehat{N}$, and $\widehat{N}^{-1}(\infty) = \{\infty\}$. The map $\widehat{N}$ has local degree $+1$ everywhere except on $\widehat{N}^{-1}(H)$. Thus every point of the quotient except H has exactly one inverse image under $\widehat{N}$, and so N^{-1} is well defined except on $\langle\{\alpha_i\}\rangle$. Set $U = N^{-1}\left(\overline{\mathbf{C}} - \langle\{\alpha_i\}\rangle\right)$.

The last statement follows from Theorem 1.8. ☐

Our main analysis will concentrate on the set $\mathbf{B}(\alpha)$. We summarize some important results about this set in the following.

THEOREM 1.10.

(i) The set $\mathbf{B}(\alpha)$ is simply connected.

(ii) Infinity lies on the frontier of $\mathbf{B}(\alpha)$ for every root α of p.

(iii) If the local degree of N restricted to $\mathbf{B}(\alpha)$ is s, then $\mathbf{B}(\alpha)$ approaches ∞ in $s-1$ different directions.

We shall omit the proof of this theorem here, but note that parts (i) and (iii) have been proven by F. Przytycki in [Pz] using classical arguments. M. Shishikura has also shown (i) using holomorphic surgery. Property (ii) is an immediate consequence of a theorem of Fatou [F3] which states that if a map has a single repelling fixed point with all other fixed points being sinks, then that fixed point lies on the boundary of the basins of attraction of each of the sinks. All three of these arguments essentially come down to showing that the family $\{N^{-n}\}$ of iterates of one branch of the inverse is a normal family, and that the limit function must be constant.

Computer studies of the case with the degree of $N|_{\mathbf{B}(\alpha)}$ is greater than 1 can be seen in the next chapter, especially Figure 2.5 and Figure 2.6.

2. A Newton's Method Picture Book

In this section, we give the results of several of our computer studies of the convergence of Newton's method for various polynomials. This should help develop some intuition for the result of Chapter 3.

The pictures in this section were computed in the following way: Take an $n \times n$ grid of points over some region in the complex plane. Typically, we take $n = 1000$ and the region to be the square with lower left corner $-2 - 2i$ and upper right corner $2 + 2i$. Using each point in the grid as a starting value, apply Newton's method, iterating at most m times. If the orbit of the point comes within ϵ of a root α, color that point according to the number of iterations required. We color those points that converge in 1 iteration white, and those that took m tries black, with levels of grey in between (see Figure 2.1). Typically, $m = 30$ and $\epsilon = 10^{-8}$. Refer to the appendix for the exact values used in each picture.

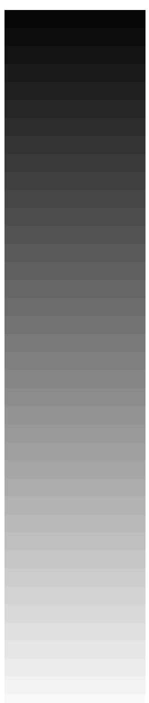

FIGURE 2.1: The gray levels used.

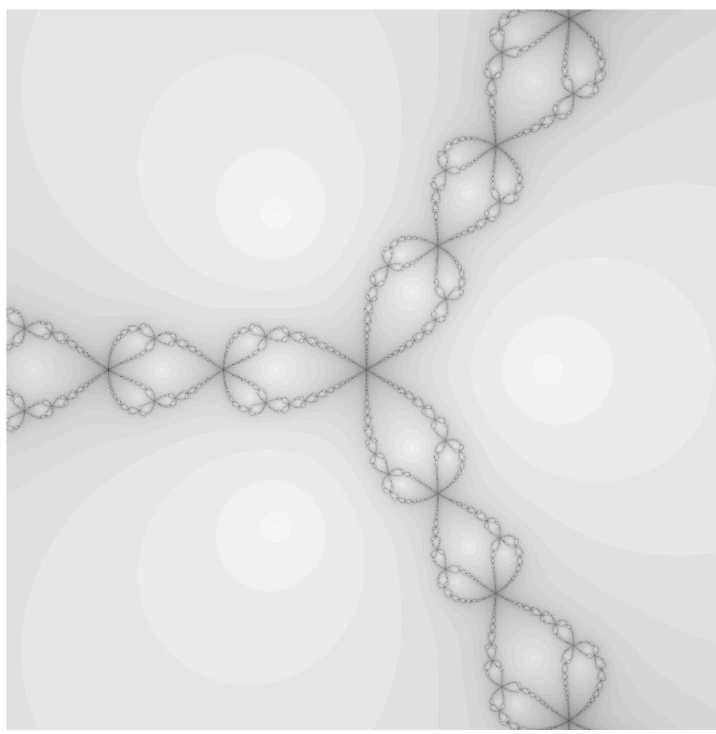

FIGURE 2.2: The convergence speeds for N_p where $p(z) = z^3 - 1$.

Pictured in Figure 2.2 is a study of the convergence of N_p for $p(z) = z^3 - 1$. The 3 roots of unity are in the white regions, and 0 is a pole. The Julia set of N lies inside the dark gray areas.

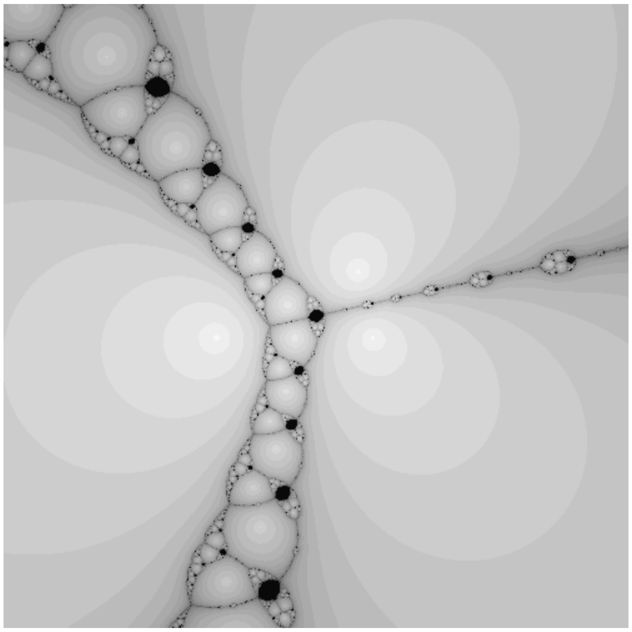

FIGURE 2.3: Newton's method for a cubic polynomial which has an attracting period 2 orbit.

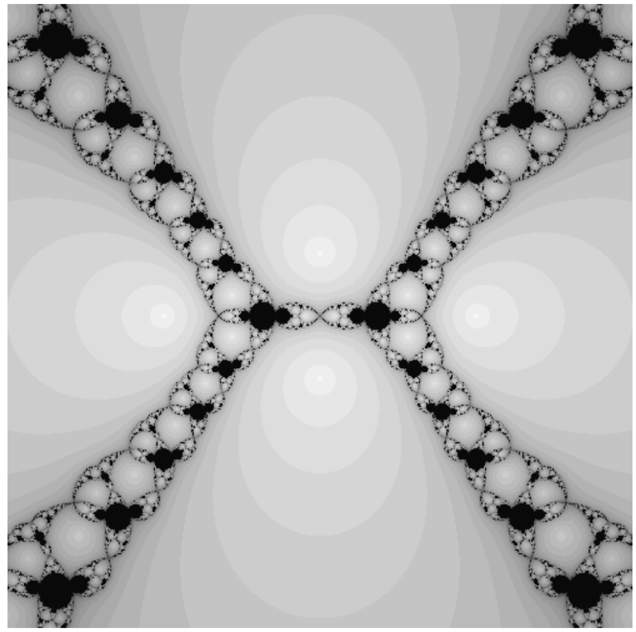

FIGURE 2.4: Newton's method for a quartic polynomial with two attracting period 2 orbits.

Figure 2.3 shows the behavior of Newton's method for a cubic where there is an attracting period 2 orbit. One point on this orbit lies inside the large black blob in the center, and the other lies in the tiny black dot between the two closest roots. All the other black regions are preimages of these two.

Figure 2.4 shows Newton's method for the polynomial $(z^2 - 1)(z^2 + 0.16)$. In this case there are two attracting period two orbits: these lie on the real line and are approximately $\{.3192, -.2599\}$ and $\{-.3192, .2599\}$. Four of the large black disks in the center of the picture contain these orbits; all of the other black disks are preimages of these. For both Figure 2.3 and Figure 2.4, these attracting orbits persist for small perturbations of the polynomials. In Figure 2.3, N^2 is polynomial-like of degree 2 near the attracting orbit (see [DH]) – all of the behavior of the family $z \mapsto z^2 + c$ occur for various perturbations of this map. The map in Figure 2.4 is polynomial–like of degree 3.

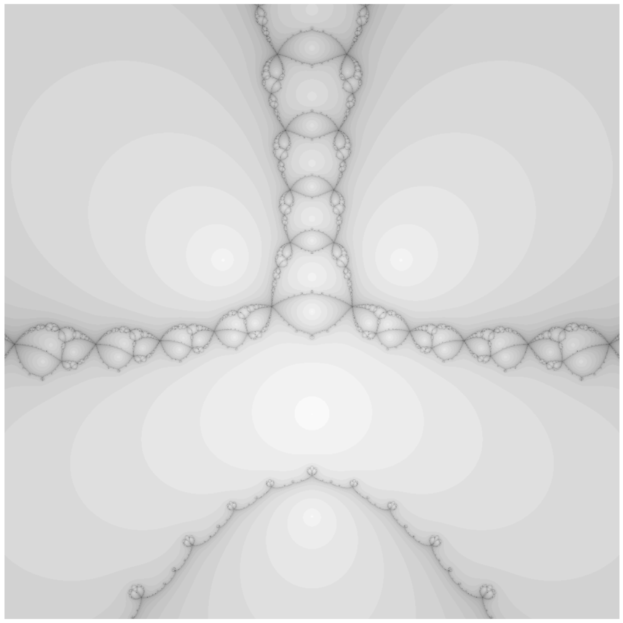

FIGURE 2.5: Sylvester the cat.

In Figure 2.5, the central root of the polynomial has an extra critical point in its immediate basin. This increases the local degree of N and so creates an extra approach to ∞. Note that near the edge of the picture, the sum of the "widths" of these two approaches is approximately the same as the width for the roots with only one approach to ∞. Figure 2.6 shows a nearby polynomial, where we have adjusted things so that the central root has 3 approaches to ∞. Note that although the width on one side has widened, the other side has closed off somewhat.

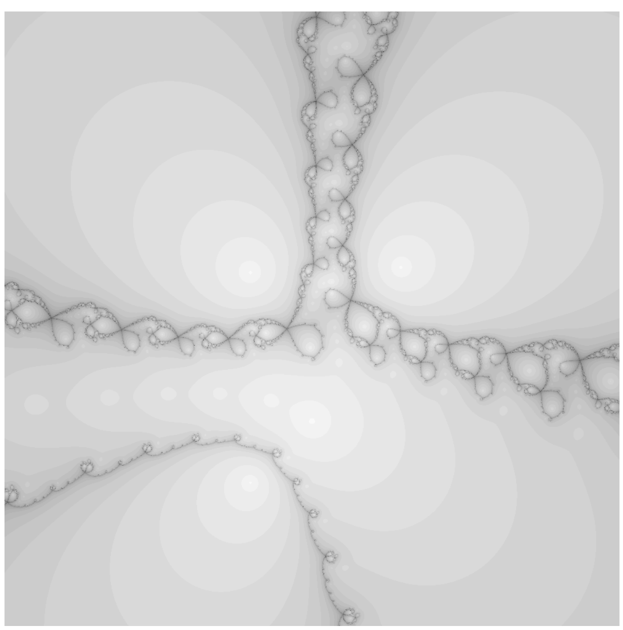

FIGURE 2.6: Sylvester squints.

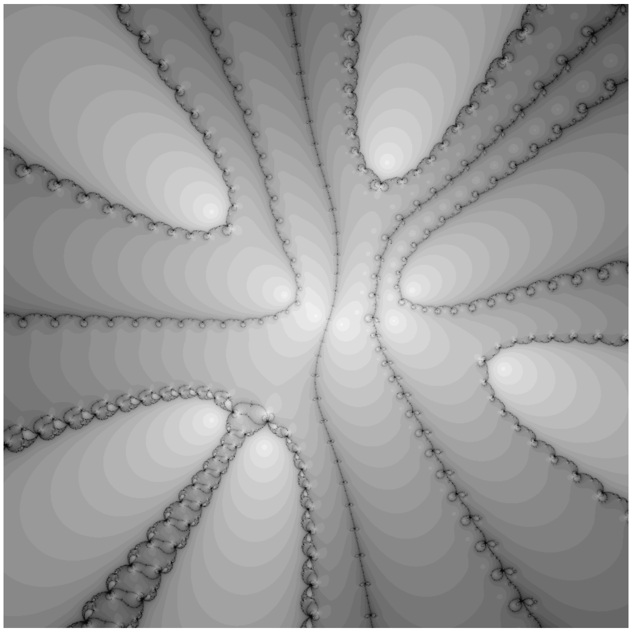

FIGURE 2.7: A degree 10 polynomial.

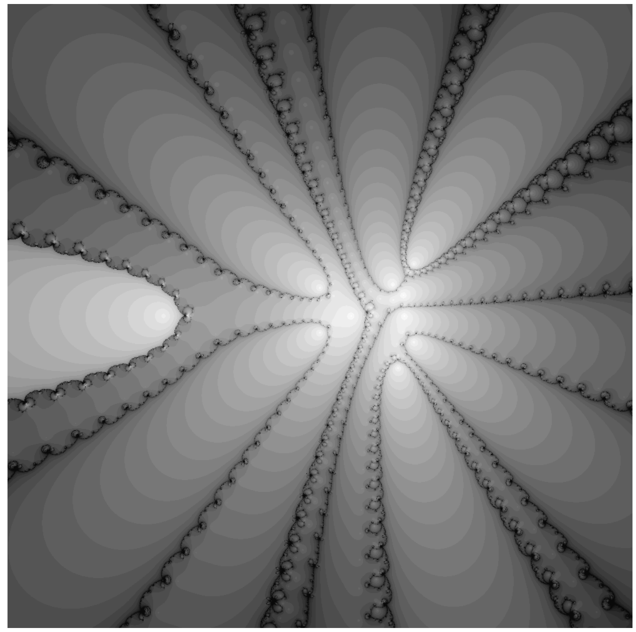

FIGURE 2.8: Another degree 10 polynomial.

Figure 2.7 and Figure 2.8 show Newton's method on 2 typical degree 10 polynomials. Again, note how the immediate basins distribute themselves somewhat equally as they tend toward the edge of the pictures. We will make this statement more precise in the next chapter.

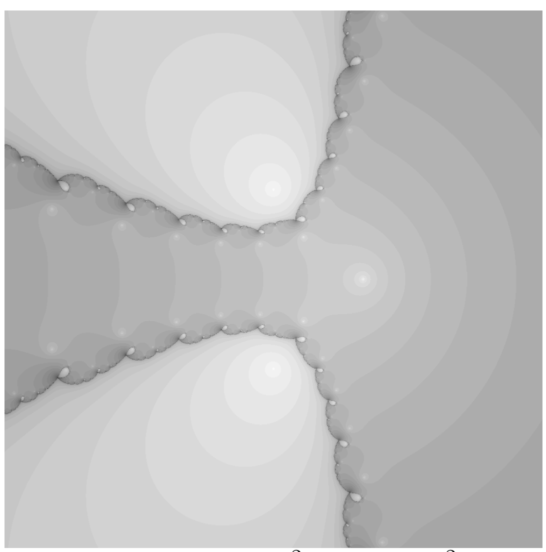

FIGURE 2.9: Newton's method for $(z^2 + 1)(z - 1)^2$, which has a double root at 1.

Figure 2.9 shows Newton's method for a quartic polynomial with a double root at 1. Note that the immediate basin of 1 is nearly twice as wide as that for the other 2 roots, although the convergence is much slower, since $N'(1) = \frac{1}{2}$.

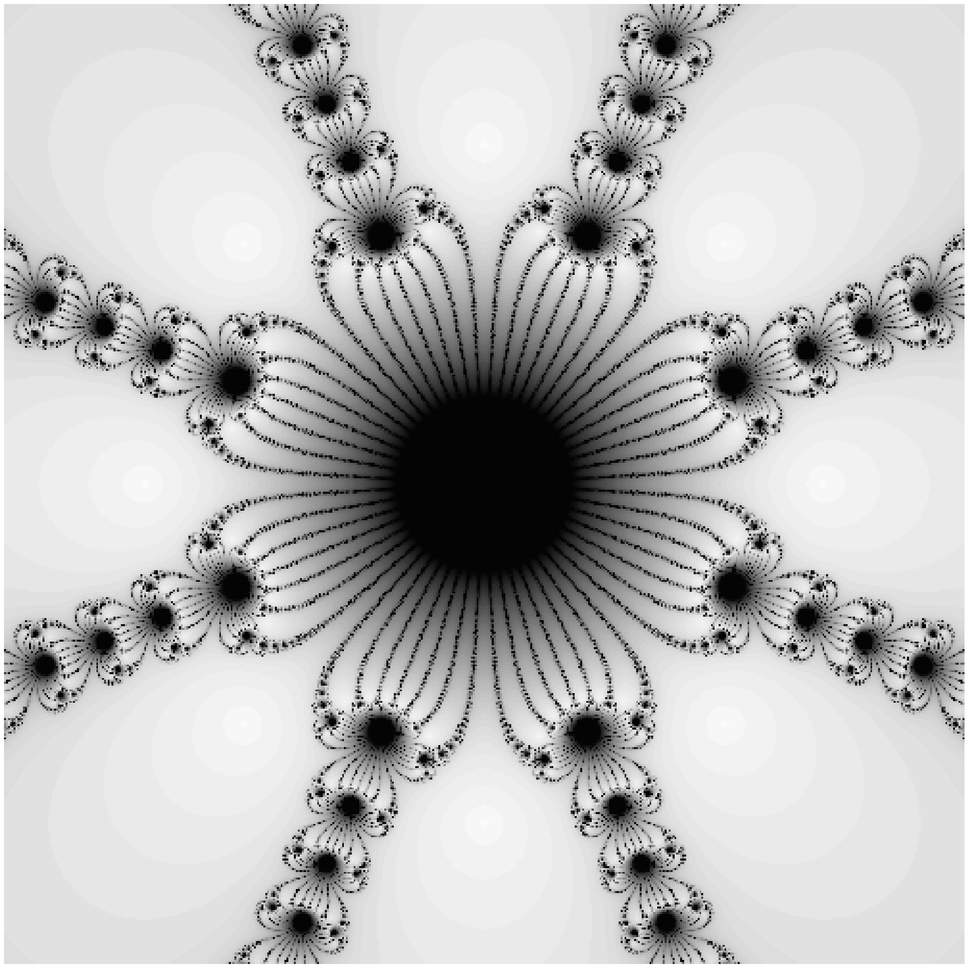

FIGURE 2.10: Newton's method for $z^8 - 1$. The large black area at the center is due to the fact that convergence is very slow near 0, not because of an attracting periodic orbit. Zero is the only pole of N_p, and lies on the boundary of the immediate basins of all eight roots.

3. Estimating the Width of the Immediate Basins

We have seen several examples in which the immediate basin of a root tends toward a given "width" as distance from the root tends toward ∞. In this chapter, we explicitly estimate this width. Throughout this section, we assume that $p(z)$ is a centered polynomial in $\mathcal{P}_d(1)$ of degree $d \geq 4$. (We could take $d \geq 2$, but a few of the estimates give better bounds for $d \geq 4$.)

3.1 Building a Model for N

We shall attempt to understand $\mathbf{B}(\alpha)$ by conjugating $N|_{\mathbf{B}(\alpha)}$ to another map M from the open unit disk to itself. Let $h : \mathbf{D} \to \mathbf{B}(\alpha)$ be the Riemann map which sends 0 to the root α. Since h is an analytic diffeomorphism, we can define M by the diagram

$$\begin{array}{ccc} \mathbf{B}(\alpha) & \xrightarrow{N} & \mathbf{B}(\alpha) \\ \uparrow h & & \uparrow h \\ \mathbf{D} & \xrightarrow{M} & \mathbf{D} \end{array}$$

The form of M depends on the number of critical points of N that lie in $\mathbf{B}(\alpha)$. If there are s such critical points, then M must be an analytic, degree $s+1$ mapping of $\mathbf{D}$ to itself. By the following proposition, it must be a finite Blashke product. (Adapted from [Bu], pp. 197–198)

PROPOSITION 3.1. *If f is an analytic, degree $s+1$ map of $\mathbf{D}$ to itself, then there exist $\mu_0, \ldots \mu_s \in \mathbf{D}$ and $\theta \in \mathbf{R}$ such that*

$$f(z) = e^{i\theta} \prod_{j=0}^{s} \frac{z - \mu_j}{1 - \overline{\mu_j} z}$$

for all $z \in \mathbf{D}$.

PROOF. *(i)* ([Ra]). First, we show that $\lim_{r \to 1} |f(re^{i\phi})| = 1$ for $\phi \in \mathbf{R}$. Suppose not. Then we can find a sequence $\{z_n\} \in \mathbf{D}$ converging to the boundary, but for which $\{f(z_n)\}$ converges to some interior point w, and with $f(z_n) \neq w$ for all n. Let $a_0, \ldots, a_k$ be distinct preimages of w, with multiplicities $m_0, \ldots, m_k$. Choose $\epsilon > 0$ and choose disjoint neighborhoods U_j of a_j which satify

(a) $\overline{U_j} \subset \mathbf{D}$

(b) each point $z \neq w$ in $\mathbf{D}_\epsilon(w)$ has exactly m_j distinct preimages.

Now choose $R < 1$ so that

$$\overline{U_0} \cup \ldots \cup \overline{U_k} \subset \mathbf{D}_R(0).$$

For n large enough, we have $|z_n| > R$, $f(z_n) \neq w$, and

$$f(z_n) \in f(U_0) \cap \ldots \cap f(U_k) \cap \mathbf{D}_\epsilon(w).$$

Thus the point $f(z_n)$ has $m_0 + \ldots + m_k = s + 1$ distinct preimages in $U_0 \cup \ldots \cup U_k$, as well as the preimage z_n which lies outside of $\mathbf{D}_R$. This means that the point z_n has $s+2$ distinct preimages, contradicting the fact that f is of degree $s+1$.

(ii) ([F4]). Let $\mu_0, \ldots, \mu_s$ be the zeros of f in $\mathbf{D}$, repeated according to multiplicity, and let

$$g(z) = \prod_{j=0}^{s} \frac{z - \mu_j}{1 - \overline{\mu_j} z}.$$

The functions f/g and g/f are both holomorphic functions in $\mathbf{D}$, and by *(i)*, we have

$$\lim_{r \to 1} \left| \frac{f(re^{i\phi})}{g(re^{i\phi})} \right| = \lim_{r \to 1} |f(re^{i\phi})| = 1.$$

Applying the Maximum Modulus Principle to f/g and g/f, we have

$$\left| \frac{f}{g} \right| = \left| \frac{g}{f} \right| = 1$$

on all of $\mathbf{D}$. Therefore,

$$f(z) = e^{i\theta} \prod_{j=0}^{s} \frac{z - \mu_j}{1 - \overline{\mu_j} z}$$

as desired. □

Since the roots of p are fixed points of N, we must have $M(0) = 0$. This gives

$$M(z) = z\mathrm{e}^{i\theta} \prod_{j=1}^{s} \frac{z - \mu_j}{1 - \overline{\mu_j} z}$$

where N has s critical points in $\mathbf{B}(\alpha)$. For a simple root α of p, $N'(\alpha) = 0$, and so we may take $\mu_s = 0$. In this case we may also conjugate away the $\mathrm{e}^{i\theta}$ term. The easiest case to keep in mind is a simple root with no "free critical points" in $\mathbf{B}(\alpha)$; here $M(z) = z^2$.

In all cases, M has an attracting fixed point at 0, and s repelling fixed points $\xi_1 \ldots \xi_s$ which lie on the unit circle. Because M preserves the unit disk, the derivative at each of these fixed points must be real and positive.

3.2 Manning's Construction

Let ξ be a repelling fixed point of M. (When $M(z) = z^2$, $\xi = 1$.) In general, $N'(\infty) \neq M'(\xi)$, so we cannot hope to analytically extend the conjugacy h to a neighborhood of the repeller. However, following Manning [M], we *can* estimate the angle $\mathbf{B}(\alpha)$ makes at ∞ by first linearizing N and M near ∞ and ξ, and then taking the $\frac{\log(\frac{d-1}{d})}{\log M'(\xi)}$ power to conjugate the linear repulsions.

There is a unique analytic map L_M tangent to the identity at ξ which conjugates M near ξ to multiplication by its derivative near 0. (For z^2, this map is $\log(z)$.) Similarly, there is a unique map L_N which conjugates N to its derivative at ∞ with $L_N'(\infty) = 1$. Finally, the map

$$R = z^{\frac{\log(\frac{d-1}{d})}{\log M'(\xi)}}$$

defined on the left half plane sends 0 to ∞ and conjugates multiplication by $M'(\xi)$ at 0 to multiplication by $\frac{d-1}{d}$ at ∞. We follow R with multiplication by a complex constant ρ, to be specified shortly. This can be represented in the following diagram, or refer to Figure 3.1.

$$\begin{array}{ccccccccc}
\overline{\mathbf{C}}, \infty & \xleftarrow{R} & \overline{\mathbf{C}}, 0 & \xleftarrow{L_M} & \mathbf{D}, \xi & \xrightarrow{h} & \mathbf{B}(\alpha), \infty & \xrightarrow{L_N} & \overline{\mathbf{C}}, \infty \\
\downarrow \times \frac{d-1}{d} & & \downarrow \times M'(\xi) & & \downarrow M & & \downarrow N & & \downarrow \times \frac{d-1}{d} \\
\overline{\mathbf{C}}, \infty & \xleftarrow{R} & \overline{\mathbf{C}}, 0 & \xleftarrow{L_M} & \mathbf{D}, \xi & \xrightarrow{h} & \mathbf{B}(\alpha), \infty & \xrightarrow{L_N} & \overline{\mathbf{C}}, \infty
\end{array}$$

Let $\mathcal{V}$ be the image under L_N of the neighborhood of ∞ where it is defined, let S be the composition $L_N h L_M^{-1} R^{-1}$, and let $\mathcal{W} = S^{-1}(\mathcal{V})$.

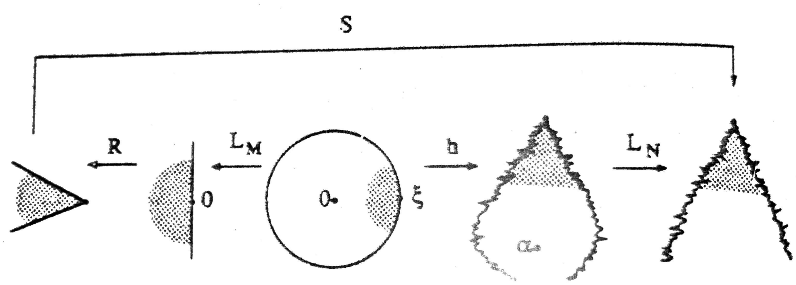

FIGURE 3.1: The various maps and regions involved in the construction of S.

Note that $\mathcal{W}$ is an open wedge which subtends an angle at ∞ of

$$\pi \frac{\log(\frac{d}{d-1})}{\log(M'(\xi))}.$$

The map S is an analytic, univalent function which commutes with multiplication by $\frac{d-1}{d}$. At this point, we could estimate the derivative of S at some point $w \in \mathcal{W}$, and apply the Koebe 1/4 Theorem (see page 29) to obtain a lower bound on the width of $\mathcal{V}$ near $S(w)$. In fact, this would give us estimates at all points of the form of the form $\left(\frac{d-1}{d}\right)^j z$ lying in $\mathcal{V}$. The problem with this approach is that this estimate only gives us the width at a discrete set of points, instead of through the whole of $\mathcal{V}$. In the next few sections, we shall present an alternative approach which was suggested to us by A. Douady.

3.3 The Opening Modulus of a Sector

We now have a univalent map S between the two sectors $\mathcal{W}$ and $\mathcal{V}$, each of which is invariant under multiplication by $\frac{d-1}{d}$. In order to estimate the "width" of $\mathcal{V}$, we will now introduce the notion of the opening modulus of a sector. Although this can be done in the more general situation of any holomorphic map with a repelling fixed point (see [BD]), we will only discuss the linear case.

Let p be a repelling fixed point of the linear map $z \mapsto cz$, and let Δ be a disk centered at p. In our case, $p = \infty$, and $c = \frac{d-1}{d}$. If we define the equivalence relation

$$z \sim c^n z,$$

then

$$T = (\Delta - \{p\}) \, / \sim$$

is a Riemann surface of genus 1, isomorphic to the torus

$$\mathbf{C} / \left(\mathbf{Z} \log c \oplus \mathbf{Z} 2\pi i\right).$$

Let ϖ be the projection

$$\varpi : \Delta - \{p\} \to T.$$

If $\mathcal{V}$ is a sector which is invariant under multiplication by c, then $\varpi(\mathcal{V})$ will be an annulus $A_{\mathcal{V}}$ in the torus T. We define the *opening modulus* of the sector $\mathcal{V}$ (relative to c) to be the modulus of the annulus $A_{\mathcal{V}}$ (see below).

Any annulus A 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 $m \in \mathbf{R}^+$. In this case, the modulus of A is said to be m. Using this definition, the modulus of an annulus A is clearly a conformal invariant. That is, if f is a univalent conformal mapping of an annulus A onto an annulus A', then A and A' must have the same modulus. It also follows that the opening modulus of a sector $\mathcal{V}$ is a conformal invariant.

Lemma 3.2. *The opening modulus of the sector $\mathcal{W}$ (and hence of $\mathcal{V}$) is*

$$\frac{\pi}{\log(M'(\xi))}.$$

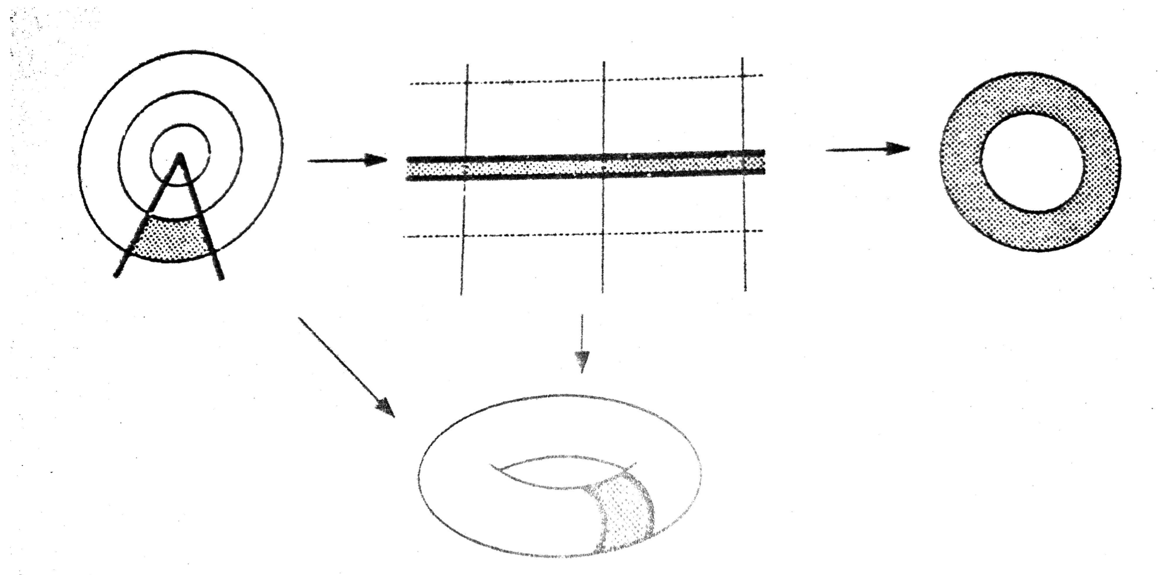

FIGURE 3.2: Calculating the modulus of the sector.

PROOF. For notational convenience, we set

$$\theta = \pi \frac{\log(\frac{d}{d-1})}{\log(M'(\xi))},$$

the angle $\mathcal{W}$ subtends at ∞. As we mentioned earlier, our repelling fixed point p is ∞. We take Δ to be the exterior of the disk of radius $\frac{d}{d-1}$, and set $\widehat{A}$ to be the part of $\mathcal{W}$ for which $\frac{d}{d-1} < |z| < 1$.

For any complex constant k, the map $\phi_k(z) = \log(k/z)$ is analytic on $\Delta - \{\infty\}$, and its image is (part of) the universal cover of the torus T. Choose a lift of the torus T so that $\varpi = \pi \circ \phi_k$, and choose k so that $\widetilde{A} = \phi_k(\widehat{A}) \cap \widetilde{T}$ has its upper right corner at the origin. Note that $\widetilde{A}$ is a rectangle of width $\log(\frac{d-1}{d})$ and height θ. Then $\zeta = \exp(\frac{2\pi i}{\log(\frac{d-1}{d})} z)$ maps $\widetilde{A}$ onto the "standard annulus"

$$1 < |z| < \exp\left(\frac{2\pi\theta}{\log(\frac{d-1}{d})}\right),$$

so the modulus of A is

$$\frac{\theta}{\log(\frac{d-1}{d})} = \frac{\pi}{\log(M'(\xi))}.$$

□

The modulus of an annulus (or a sector) is, in some sense, a measure of its "width". That is to say, the larger the modulus, the thicker the annulus. To make this statement more precise, we shall have to recall some basic facts about extremal length in the next section.

3.4 An Extremely Short Course on Extremal Length

We will only need a few basic facts about extremal length, so we present them here. This topic usually comes up in a discussion of quasiconformal mappings (which we will not be discussing), although it is also useful for conformal maps. A standard reference on the subject is [A].

Let Γ be a family of curves in the plane. The extremal length of Γ is a measure of the average minimal length of the curves in Γ.

We call a function ρ ***allowable*** if it satisfies:

1. $\rho \geq 0$ and is measurable.
2. $0 < \iint \rho^2 dx\, dy < \infty$, where the integral is taken over the whole plane.

For such a ρ, and a curve $\gamma \in \Gamma$, we define

$$L_\gamma(\rho) = \begin{cases} \int_\gamma \rho\, |dz|, & \text{if } \rho \text{ is measurable on } \gamma \text{ (as a function of arc length);} \\ \infty, & \text{otherwise.} \end{cases}$$

Set

$$L(\rho) = \inf_{\gamma \in \Gamma} L_\gamma(\rho).$$

Definition. The ***extremal length*** of Γ is the quantity

$$\lambda(\Gamma) = \sup_\rho \frac{(L(\rho))^2}{\iint \rho^2 dx\, dy}$$

for all allowable ρ.

Definition. If Γ_1 and Γ_2 are families of curves such that every $\gamma_2 \in \Gamma_2$ contains a $\gamma_1 \in \Gamma_1$, we say that $\Gamma_1 < \Gamma_2$. (There are "fewer" curves in Γ_2, and they are longer.)

Remark. Note that $\Gamma_2 \subset \Gamma_1 \Rightarrow \Gamma_1 < \Gamma_2$!

We shall need the following two standard facts:

Theorem 3.3. *If* $\Gamma_1 < \Gamma_2$, *then* $\lambda(\Gamma_1) < \lambda(\Gamma_2)$.

Theorem 3.4. $\lambda(\Gamma)$ *is a conformal invariant.*

Lastly, we give (without proof) a few examples of families of curves and their corresponding extremal length.

Example. Let R be a rectangle of height h and width w, and let Γ be the set of curves joining the top to the bottom. Then $\lambda(\Gamma) = h/w$.

EXAMPLE. Let Γ_1 be the set of curves which join the boundary arcs of the annulus $r_1 < |z| < r_2$. Then

$$\lambda(\Gamma_1) = \frac{1}{2\pi} \log \frac{r_2}{r_1}.$$

Note that this is the same as the modulus of the annulus. In fact, this is commonly used as a definition of the modulus.

EXAMPLE. Let Γ_2 be the set of closed curves in the annulus $r_1 < |z| < r_2$ that have nontrivial homotopy, and Γ_1 be as before. Then

$$\lambda(\Gamma_2) = \frac{1}{\lambda(\Gamma_1)} \ .$$

3.5 The Larger the Modulus, the Thicker the Annulus

In this section we make more explicit the relationship between the modulus and the thickness of an annulus. Mitsuhiro Shishikura very kindly pointed out to us how this could be accomplished. We shall prove the following:

PROPOSITION 3.5. *Let T be a torus isomorphic to $\mathbf{C}/(\mathbf{Z} \oplus \mathbf{Z}\tau)$, and let A be a nontrivial annulus contained in T with* modulus$(A) = m$. *Then the distance between the boundary curves of A is at least*

$$\frac{2\pi k e^{\frac{\pi}{2m}}}{1 + e^{\frac{\pi}{m}}}$$

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

PROOF. Consider an open ellipse with its center at 0, whose major axis is the interval $[-\frac{r+1/r}{2}, \frac{r+1/r}{2}]$ and with its minor axis being $[-\frac{r-1/r}{2}i, \frac{r-1/r}{2}i]$. Remove the intervals $(-\frac{r+1/r}{2}, -1]$ and $[1, \frac{r+1/r}{2})$ from that ellipse, and denote the resulting set E. Let Γ be the set of curves in E which join the upper and lower portions of the boundary of the ellipse.

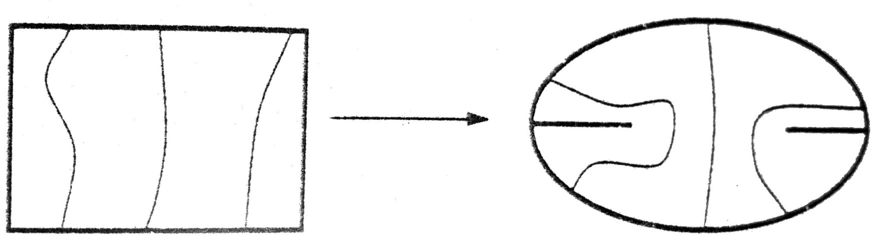

FIGURE 3.3: The rectangle mapped to the ellipse E, and some curves in Γ.

Then $\lambda(\Gamma) = \frac{\log r}{2\pi}$. This is because the map $\cosh(z)$ maps the open rectangle $(-\log r, \log r) \times (0, \pi\mathrm{i})$ univalently onto E and the set of curves joining top to bottom in the rectangle onto Γ.

We may assume that A lies in T so that (in the cover) it is homotopic to the real axis. Suppose that A has a "narrow part" of width δ. Then scale the ellipse E by $\delta/2$ and embed it in T so that the interval $[-\frac{\delta}{2}, \frac{\delta}{2}]$ corresponds to the narrow part. (See Figure 3.4.)

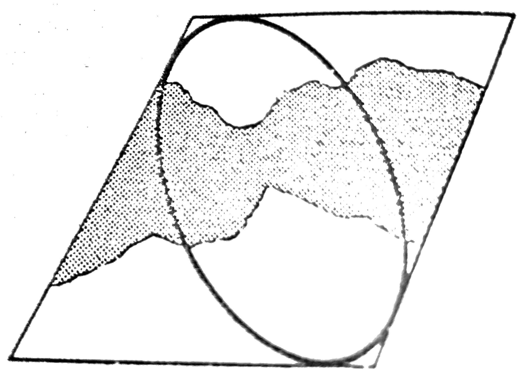

FIGURE 3.4: The torus T, with the embedded ellipse E and the annulus A (shaded).

Let $k = \min\{1, \mathrm{Im}(\tau)\}$. Then if we choose

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

the ellipse E will be injectively embedded.

Let Γ_A be the family of nontrivial closed curves in A, and recall that $\lambda(\Gamma_A) = 1/m$, where m is the modulus of A. Also, notice that every curve in Γ_A contains a curve from Γ, so $\Gamma < \Gamma_A$. We can apply Theorem 3.3 to obtain

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

Now solve for δ,

$$\delta \geq \frac{2\pi k \mathrm{e}^{\frac{\pi}{2m}}}{1 + \mathrm{e}^{\frac{\pi}{m}}}.$$

$\square$

3.6 The Width near Infinity

We can now combine the results of Lemma 3.2 and Proposition 3.5 to put a lower bound on the size of the sector $\mathcal{V}$ near ∞.

THEOREM 3.6. *For any R sufficiently large, there is a point $z_R \in \mathcal{V}$, with $|z_R| = R$, which is the center of a disk of radius at least*

$$R \tanh\left(\frac{\pi\sqrt{M'(\xi)}}{1+M'(\xi)} \log\left(\frac{d}{d-1}\right)\right) \geq \frac{2\pi R}{3d\left(1+\sqrt{M'(\xi)}\right)}$$

contained entirely in $\mathcal{V}$.

PROOF. Recall from Lemma 3.2 that the modulus of $\mathcal{V}$ is $\frac{\pi}{\log(M'(\xi))}$. This is also the modulus of the corresponding annulus $A_{\mathcal{V}}$ lying in the torus $\mathbf{C}/(\mathbf{Z}\log\frac{d}{d-1} \oplus \mathbf{Z}2\pi i)$. The exponential function maps the lift of the annulus back onto $\mathcal{V}$, so we can use that to pull our estimates over.

By Proposition 3.5, we know that the "thinnest part" of the annulus is at least

$$\delta = \frac{2\pi\sqrt{M'(\xi)}}{1+M'(\xi)} \log\frac{d}{d-1}$$

wide. (We had to rescale the result by a factor $\log\frac{d}{d-1}$ to account for the different sized torus.) This means there is a curve running through the annulus on which we can center disks of radius $\frac{\delta}{2}$ which stay in the annulus. Index these disks by their centers w. The exponential of the disk centered at w is a bean-shaped blob, which contains a disk of radius

$$|\mathrm{e}^w|(\mathrm{e}^{\frac{\delta}{2}} - \mathrm{e}^{-\frac{\delta}{2}}) = |e^w| \sinh\frac{\delta}{2}$$

and centered at $e^w \cosh\frac{\delta}{2}$. For each R, we set $z_R = e^w \cosh(\frac{\delta}{2})$; if there is more than one w that works, pick one arbitrarily. Substituting $\dfrac{z_R}{\cosh(\delta/2)}$ for e^w tells us that the radius of the disk centered at z_R is $|z_R|\tanh(\delta/2)$.

Notice that, since ξ is a repelling fixed point, $M'(\xi) > 1$, so

$$\tanh\left(\frac{\pi\sqrt{M'(\xi)}}{1+M'(\xi)} \log(\frac{d}{d-1})\right) \geq \tanh\left(\frac{\pi}{1+\sqrt{M'(\xi)}}\frac{1}{d}\right).$$

Then, for $x \leq 1$ we have

$$\tanh(x) \geq x - \frac{x^3}{3} > \frac{2x}{3},$$

which gives the final result. $\square$

What we have really done here is find the width of a "channel" through $\mathcal{V}$ which starts at infinity and heads off toward 0. Since the map $L_N : \mathbf{B}(\alpha) \to \mathcal{V}$ is analytic and tangent to the identity at ∞, we actually have an estimate for the width of $\mathbf{B}(\alpha)$ near ∞, because we can choose our point z_R as large as we like.

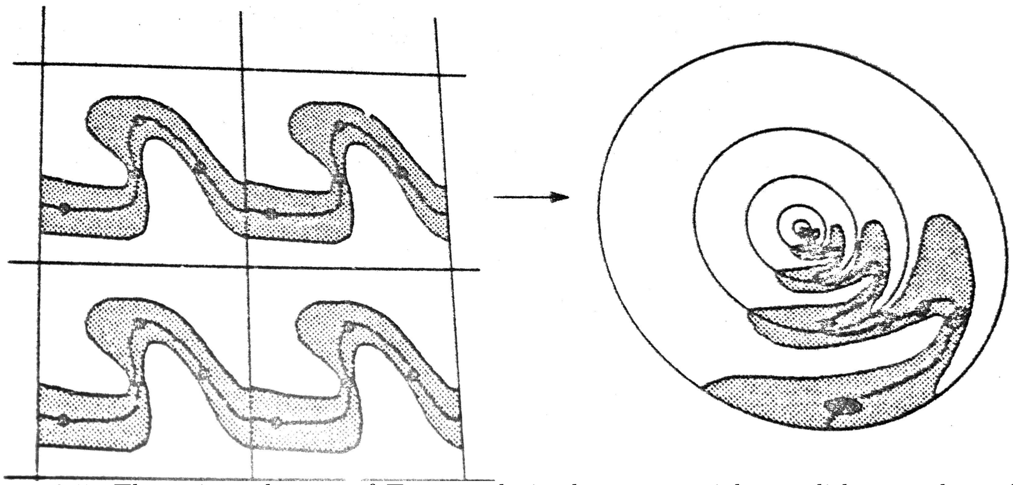

FIGURE 3.5: The universal cover of T mapped via the exponential to a disk around ∞. Some of the lifts of the annulus $A_{\mathcal{V}}$ are shown, mapping onto the sector $\mathcal{V}$. Also shown is the curve running through $A_{\mathcal{V}}$ and several of the disks of radius $\delta/2$ (in dark gray).

3.7 Some Tedious, But Necessary, Calculations

We have found a disk that lies entirely within $\mathbf{B}(\alpha)$, but it is rather far away. The following lemmas provide a lower bound on how much N contracts this disk as we iterate. The first lemma is essentially the same as Lemma 3.1 in [M], although the constants have been changed to better suit our circumstances. Lemma 3.9 is also similar to Lemma 3.2 in [M], although the proof is somewhat different. All of the proofs are rather tedious calculations, so be forewarned. Now is a good time for a cup of coffee.

LEMMA 3.7. *If* $|z| > 2$ *then*

$$\left|\frac{zN'(z)}{N(z)}\right| \geq \left(1 - \frac{8}{3|z|^2}\right)^2.$$

PROOF.

$$\left|\frac{zN'(z)}{N(z)}\right| = \left|z \cdot \frac{[p(z)][p''(z)]}{[p'(z)]^2} \cdot \frac{[p'(z)]}{[zp'(z) - p(z)]}\right|$$

$$= \left|\frac{z \cdot [z^d(1 + a_{d-2}z^{-2} + \cdots)][d(d-1)z^{d-2}(1 + \frac{(d-2)(d-3)}{d(d-1)}a_{d-2}z^{-2} + \cdots)]}{[dz^{d-1}(1 + \frac{d-2}{d}a_{d-2}z^{-2} + \cdots)] \cdot [(d-1)z^d(1 + \frac{d-3}{d-1}a_{d-2}z^{-2} + \cdots)]}\right|$$

$$= \left|\frac{(1 + a_{d-2}z^{-2} + \cdots)(1 + \frac{(d-2)(d-3)}{d(d-1)}a_{d-2}z^{-2} + \cdots)}{(1 + \frac{d-2}{d}a_{d-2}z^{-2} + \cdots)(1 + \frac{d-3}{d-1}a_{d-2}z^{-2} + \cdots)}\right|$$

set $x = |1/z|$

$$\geq \frac{(1 - x^2 - x^3 - \cdots)^2}{(1 + x^2 + x^3 + \cdots)^2}$$

$$= \frac{(\frac{-1}{1-x} + x + 2)^2}{(\frac{1}{1-x} - x)^2}$$

$$= \frac{(\frac{1-x-x^2}{1-x})^2}{(\frac{1-x+x^2}{1-x})^2}$$

$$= \left(1 - \frac{2x^2}{1-x+x^2}\right)^2$$

Since $\frac{2}{1-x+x^2} \leq \frac{8}{3}$ for $0 \leq x \leq \frac{1}{2}$,

$$\geq \left(1 - \frac{8x^2}{3}\right)^2$$

$$\geq \left(1 - \frac{8}{3|z|^2}\right)^2.$$

□

The following lemma is a souped-up version of Lemma 1.7, and gives some estimates on how quickly points move in toward the roots.

Lemma 3.8. *For z large, choose $k > 2$ and let n be the largest integer such that $|N^n(z)| > k$. Then*

$$|N^i(z)| - 2 \geq (\frac{d}{d-1})^{n-i}(k-2), \qquad i = 0, \ldots, n$$

Proof. From Lemma 1.7, we obtain

$$|N(z)| \leq \frac{\frac{d-1}{d}|z| + 2}{d}$$
$$= \tfrac{d-1}{d}(|z| - 2) + 2$$

and so

$$\frac{|N(z)| - 2}{|z| - 2} \leq \frac{d-1}{d}.$$

Applying this $n - i$ times, we get

$$\frac{|N^n(z)| - 2}{|N^i(z)| - 2} = \frac{|N^n(z)| - 2}{|N^{n-1}(z)| - 2} \cdot \frac{|N^{n-1}(z)| - 2}{|N^{n-2}(z)| - 2} \cdots \frac{|N(z)^(i+1)| - 2}{|N^i(z)| - 2}$$
$$\leq (\frac{d-1}{d})^{n-i}.$$

Rewriting,

$$|N^i(z)| - 2 \geq (\tfrac{d}{d-1})^{n-i}\,(|N^n(z)| - 2)$$
$$\geq (\tfrac{d}{d-1})^{n-i}(k-2).$$

□

For a final technical lemma, we combine Lemma 3.8 and Lemma 3.7 to get a bound on the nonlinearity of n iterations of N.

LEMMA 3.9. *Let* $|z| > 2 + 2\sqrt{d}$, $d \geq 4$, *and let* n *be the largest integer such that* $|N^n(z)| > 2 + 2\sqrt{d}$. *Then*

$$\left| \frac{z(N^{n+1})'(z)}{N^{n+1}(z)} \right| \geq 1/2.$$

PROOF. For notational convenience, let $k = 2 + 2\sqrt{d}$. We have

$$\left| \frac{z(N^{n+1})'(z)}{N^{n+1}(z)} \right| = \prod_{i=0}^{n} \left| \frac{N'(N^i(z))N^i(z)}{N^{i+1}(z)} \right|$$

Applying Lemma 3.7,

$$\geq \prod_{i=0}^{n} \left(1 - \frac{8}{3|N^i(z)|^2}\right)^2$$

$$= \exp\left(2\sum_{i=0}^{n} \log(1 - \frac{8}{3|N^i(z)|^2})\right).$$

Note that $\log(1 - ax) \geq \frac{x}{c}\log(1 - ac)$ when $0 \leq x \leq c \leq \frac{1}{a}$. We take $a = 8/3$, $c = 1/k^2$ and $x = 1/|N^i(z)|^2$ to obtain

$$\left| \frac{z(N^{n+1})'(z)}{N^{n+1}(z)} \right| \geq \exp\left(2k^2 \log(1 - \frac{8}{3k^2}) \sum_{i=0}^{n} \frac{1}{|N^i(z)|^2}\right)$$

Applying the previous lemma,

$$\geq \exp\left(2k^2 \log(1 - \frac{8}{3k^2}) \sum_{i=0}^{n} \frac{1}{(2 + (\frac{d}{d-1})^{n-i}(k-2))^2}\right)$$

$$\geq \exp\left(2k^2 \log(1 - \frac{8}{3k^2}) \int_0^{\infty} \frac{dx}{\left(2 + (k-2)(\frac{d}{d-1})^x\right)^2}\right)$$

Now we need to calculate the integral in the exponent. To make things a bit less messy, we set $r = \frac{d}{d-1}$ and $a = k - 2$. Then

$$\int \frac{dx}{(2 + ar^x)^2} = \frac{1}{\log r}\left(\frac{1}{4}\log\frac{r^x}{2 + ar^x} + \frac{1}{2(2 + ar^x)}\right).$$

So the value of the improper integral is

$$\frac{1}{\log r}\left(\frac{1}{4}\log\frac{2 + a}{a} - \frac{1}{2(2 + a)}\right).$$

If we rewrite r and a in terms of d, (remember $a = k - 2 = 2\sqrt{d}$), we obtain

$$\frac{1}{8} - \frac{1}{6\sqrt{d}} + \mathcal{O}\left(\frac{1}{d}\right).$$

Thus

$$\left|\frac{z(N^{n+1})'(z)}{N^{n+1}(z)}\right| \geq \exp\left(\frac{k^2}{4}\log(1-\frac{8}{3k^2})\right) \geq 1/2,$$

where we have used $k = 2 + 2\sqrt{d}$ and $d \geq 4$. □

REMARK. Notice that for any $c > 0$, we could take $k = 2 + c\sqrt{d}$ and still obtain a bound on the nonlinearity of N. This is because the integral in the proof is bounded below by $\frac{1}{2c^2}$. The resulting bound on the nonlinearity tends to 0 as $c \to 0$ and to 1 as $c \to \infty$. Specifically, $c = 1$ gives a bound of 0.0518, which seems quite low considering the numerical experiments. It is probably an artifact of the proof.

3.8 Finally, a Result!

We are now ready to estimate the size of the part of $\mathbf{B}(\alpha)$ corresponding to the fixed point ξ of M at a reasonable distance from the roots.

We will need to use the Koebe $\frac{1}{4}$ Theorem in this proof, so we state it here. A proof of this can be found in many univalent function theory texts, for example [Du].

THE KOEBE 1/4 THEOREM. *Let $f : \mathbf{D}_R(c) \to \mathbf{C}$ be an analytic, univalent mapping. Then the disk of radius*

$$\frac{R}{4}|f'(c)|$$

is contained in the image of $\mathbf{D}_R(c)$.

THEOREM 3.10. *Let ξ be a repelling fixed point of M. Then $\mathbf{B}(\alpha)$ contains a disk centered at t_ξ of radius*

$$\frac{(2+2\sqrt{d})\pi}{12d\left(1+\sqrt{M'(\xi)}\right)},$$

where $|t_\xi| = 2 + 2\sqrt{d}$.

PROOF. By Theorem 3.6, for every R large enough, we obtain a certain sized disk lying inside $\mathbf{B}(\alpha)$ which is centered at z_R. Let z_0 be one of these z_R which satisfies $|N^n(z_0)| = 2 + 2\sqrt{d}$ for some n. As long as we stay outside of the convex hull of the roots, $N(z)$ is a univalent function (by Proposition 1.9), and so we can apply the Koebe 1/4 theorem to obtain a disk centered at t_ξ of radius at least

$$\frac{|z_0|\pi}{6d\left(1+\sqrt{M'(\xi)}\right)}|(N^n)'(z_0)| = \frac{|t_\xi|\pi}{6d\left(1+\sqrt{M'(\xi)}\right)}\left|\frac{z_0(N^n)'(z_0)}{N^n(z_0)}\right|$$

As a result of Lemma 3.9, the radius is at least

$$\frac{(2+2\sqrt{d})\pi}{12d\left(1+\sqrt{M'(\xi)}\right)}$$

$\square$

In the simplest case (a simple root with no "free critical points"), we know that $M(z) = z^2$ and $\xi = 1$, so we have immediately the following estimate:

COROLLARY 3.11. *If α is a simple root of the polynomial $p(z)$ and $p''(z) \neq 0$ for all $z \in \mathbf{B}(\alpha)$, then $\mathbf{B}(\alpha)$ contains a disk of radius at least*

$$\frac{(2+2\sqrt{d})\pi}{28d},$$

centered at a point t of norm $2+2\sqrt{d}$.

This says that the width of the basin of a simple root takes up at least $1/56d$ of the circle of radius $2+2\sqrt{d}$. We would like an estimate that covers *all* cases, however. The next lemma gives us the tool we need.

LEMMA 3.12. *Let α be a root of multiplicity $m < d$ be such that $N|_{\mathbf{B}(\alpha)}$ is degree $s+1$. Let $\xi_1, \ldots, \xi_s$ be the repelling fixed points of the corresponding Blaschke product M. Then*

$$\sum_{i=1}^{s} \frac{1}{1+\sqrt{M'(\xi_i)}} \geq \frac{1}{1+\sqrt{\frac{2m}{2m-1}}}.$$

PROOF. First, note that $m < d$ implies $s \geq 1$. For if $s = 0$, $N|_{\mathbf{B}(\alpha)}$ is conjugate to a linear contraction on $\mathbf{C}$. This either forces N to have countably many fixed points on the boundary of $\mathbf{B}(\alpha)$, or that $\mathbf{B}(\alpha) = \mathbf{C}$. Both cases occur only for $m = d$.

Since $s > 1$, N is conjugate to a degree $s+1$ Blaschke product $M : \mathbf{D} \to \mathbf{D}$, which extends to a rational map of $\overline{\mathbf{C}}$. This map has fixed points at 0, ∞, and $\xi_1, \ldots, \xi_s$, where $|\xi_i| = 1$.

We now use a formula of Fatou [F1] which relates the derivatives at the fixed points of a rational map.

$$\sum_{\text{fixed points}} \frac{1}{M'(z_i)-1} = -1$$

Since $M'(0) = M'(\infty) = \frac{m-1}{m}$, we have

$$\sum_{i=0}^{s} \frac{1}{M'(\xi_i)-1} = 2m-1.$$

We know the ξ_i are repelling fixed points, so the maximum value of $\sum \frac{1}{1+\sqrt{M'(\xi_i)}}$ occurs when $M'(\xi_i)$ is the same for all i, and the minimum when the derivative at one fixed point is small and the derivatives at the others are allowed to get arbitrarily large. Solving for each of these cases yields the desired inequality. $\square$

As an immediate consequence, we obtain our main result:

THEOREM 3.13. *Let α be a root of multiplicity m, with $N|_{\mathbf{B}(\alpha)}$ of degree s. Then there are points $t_1, \ldots, t_s$ of magnitude $2+2\sqrt{d}$ for which a disk of radius r_i centered at t_i lies entirely within $\mathbf{B}(\alpha)$. These radii satisfy*

$$\sum_{i=1}^{s} r_i \geq \frac{(2+2\sqrt{d})\pi}{12d\left(1+\sqrt{\frac{2m}{2m-1}}\right)}$$

PROOF. We just apply Theorem 3.10 to each of the s fixed points of the map M, and then use Lemma 3.12 to put them together. □

This theorem says, roughly, that the basin of any given root α takes up at least $\frac{1}{56d}$ of the circumference of the circle of radius $2+2\sqrt{d}$, and even more if α is a multiple root. That is,

COROLLARY 3.14. *Let $p(z)$ be a centered polynomial in $\mathcal{P}_d(1)$, and $|z| \geq 2+2\sqrt{d}$. Then the probability that $N^n(z)$ will converge to a root of p is at least $\frac{1}{56d}$.*

If we want to use Newton's method to find all of the roots of a polynomial without deflating, then for each root α_i, we need some way to pick at least one initial condition lying in $\mathbf{B}(\alpha_\mathbf{i})$. Since we have a lower bound on the width of this set when it crosses a certain circle, we can find it by dividing that circle up into subintervals of that size and put a point in each subinterval.

COROLLARY 3.15. *Let $p(z)$ be a centered polynomial in $\mathcal{P}_d(1)$. Let $t_1, \ldots, t_n$ be points equally spaced around the circle of radius $2+2\sqrt{d}$, where $n \geq 56d(d-1)$. Then for each root α_i of $p(z)$, at least one of the points t_j lies in $\mathbf{B}(\alpha_\mathbf{i})$.*

PROOF. First, consider polynomials p with d distinct roots. The Newton method map N has $2d-2$ critical points, but d of them are coincident with the d roots. This leaves $d-2$ free. The smallest width basin can occur when all $d-2$ free critical points iterate to the same root α. In this case, $\mathbf{B}(\alpha)$ has $d-1$ canals reaching to ∞, whose total width at radius $2+2\sqrt{d}$ is at least $\dfrac{(2+2\sqrt{d})\pi}{28d}$. Thus there is one which has a width of at least a factor of $\frac{1}{d-1}$ of the total.

For a polynomial with multiple root of multiplicity m, the degree of the corresponding Newton map N is $d-m+1$ (or less if there are other multiple roots). Since the multiple root must also have a critical point of N in its basin (although not necessarily coincident with the root), we have $d-m-1$ "free" critical points. Placing them all in the basin of a particular root gives a larger width than in the simple root case, and this concludes the proof. □

REMARK. Note that placing $56d$ points around the circle of radius $2+\sqrt{d}$ guarantees that at least 2 of them lie in the basins of 2 different roots. This is because there are only $d-2$ "free critical points" and so at least 2 roots must have immediate basins with only 1 approach to ∞. By the same reasoning, at least half of the roots have at most 2 approaches to infinity, so $112d$ points are sufficient.

4. Experiments and Conjectures

In this chapter we present the results of some computer experiments related to Newton's method and some conjectures arising from them.

4.1 Root Finding Algorithms

Corollary 3.15 tells us that, for a centered polynomial in $\mathcal{P}_d(1)$, we need no more than $56d(d-1)$ points placed around the circle of radius $2+2\sqrt{d}$ to be sure that we have one point in the immediate basin of each root. This information is not quite sufficient to state a scheme for using Newton's method which will guarantee convergence to all the roots— we also need an upper bound on the iterations required. However, we can do so if we make the assumption that an orbit which converges to a root does so at least linearly, decreasing the distance to the root by a factor of $\frac{d-1}{d}$ with each iteration. Unless multiple roots are ruled out, this is the best we can hope for, since the Newton's method for the polynomial $p(z)=z^d$ has exactly this rate of convergence on the whole of $\overline{\mathbf{C}}$. Making this assumption, we give the following:

Conjectured Algorithm 4.1. *Let $p(z)$ be a centered polynomial in $\mathcal{P}_d(1)$. Assuming at least linear convergence, all the roots can be found by taking $56d(d-1)$ points equally spaced around the circle of radius $2+2\sqrt{d}$, using each one as a starting value for Newton's method. If after $d\log\left(\frac{4+2\sqrt{d}}{\epsilon}\right)$ iterations of N a particular starting value is not within ϵ of a root, it should be abandoned in favor of the next point.*

Remark. This algorithm takes at most $\mathcal{O}\left(d^3\log\frac{\sqrt{d}}{\epsilon}\right)$ evaluations of N to find an approximate zero of $p(z)$. This is for a worst-case scenario. Denote the points on the circle by $t_1,\ldots,t_n$ where $\arg(t_1) < \arg(t_2) < \ldots < \arg(t_n)$. We should choose the initial points for N in the order

$$t_1, t_{\frac{n}{2}}, t_{\frac{n}{4}}, t_{\frac{3n}{4}}, \ldots,$$

which keeps the points tested "balanced" around the circle. Using this ordering instead of $t_1, t_2, t_3, \ldots$ should, in general, decrease the number of points required.

A common technique in root finding is to *deflate* the polynomial after finding some approximate roots. That is, a new polynomial of lower degree is obtained from the original by dividing out by the approximate roots. The trouble with this technique is that it can introduce significant numerical errors, because the deflated polynomial is only an approximate factor of the original.

If deflation is not viewed to be a problem, we can modify Conjectured Algorithm 4.1 to reduce the number of evaluations required. Again, we make the same assumption of at least linear convergence.

Conjectured Algorithm 4.2. *Let $p(z)$ be a centered polynomial in $\mathcal{P}_d(1)$ of degree d. Place $112d$ points equally spaced around the circle of radius $2+2\sqrt{d}$, and use each one as a starting value for Newton's method, iterating each at most $d\log\left(\frac{4+2\sqrt{d}}{\epsilon}\right)$ times. If all of the roots have not been found after testing all $112d$ points, the polynomial should be deflated by those roots that have been found, and the process repeated.*

This algorithm uses $112d$ points to find at least half of the roots, and then an additional $56d$ (or fewer) points to find half of those remaining, and so on. This means we do at most $\mathcal{O}\left(d^2\log(d)\log\left(\frac{\sqrt{d}}{\epsilon}\right)\right)$ evaluations to find all of the roots.

Remark. Several numerical experiments seem to imply that the circle used can be of constant radius (independent of d). Radius 3 seems to be adequate. Making the appropriate modifications to Conjectured Algorithm 4.1 and Conjectured Algorithm 4.2 yield algorithms with requiring $\mathcal{O}\left(d^3\log\left(\frac{1}{\epsilon}\right)\right)$ and $\mathcal{O}\left(d^2\log(d)\log\left(\frac{1}{\epsilon}\right)\right)$ evaluations, respectively.

It also seems quite likely that the number of points required $(56d(d-1))$ is also too large. Experiments indicate that the number required is actually between $d(d-1)$ and $2d(d-1)$. Although this does not decrease the order of the number of computations required, it does have a significant impact in practice.

4.2 The Newton Flow

For a complex polynomial $p(z)$, we define the Newton flow to be the ordinary differential equation

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

Notice that the Newton map N_p is an Euler approximation to this O.D.E. with step size 1. Or, more generally, the relaxed Newton's method

$$N_{h,p} = z - h\frac{p(z)}{p'(z)}$$

is an Euler approximation with step size h.

It is often more convenient to work with the desingularized flow

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

which has the same solution curves as the original. We will refer to this latter flow as $\mathcal{N}_p$. This flow has been studied by several people in recent years. See [S1], [Sh], [JJT], [S2], and [STW]. We present some of the elementary properties in Proposition 4.3. We omit the proof, since these are readily verified.

PROPOSITION 4.3.

(i) The attractors of $\mathcal{N}_p$ are sinks located at the zeros of $p(z)$.

(ii) ∞ is the only source.

(iii) The only other rest points are the zeros β_i of p'. For a simple zero of p', these are hyperbolic saddles. Multiple zeros of p' correspond to degenerate (multipronged) saddles.

(iv) Solution curves of $\mathcal{N}_p$ are mapped by p to straight lines emanating from the origin.

PROPOSITION 4.4. *Let $\mathcal{C}_r$ be a circle of radius r centered at 0, and let p be a centered polynomial in $\mathcal{P}_d(1)$ with d simple roots. Then the portion of $\mathcal{C}_r$ that lies in the basin of any given root of p under $\mathcal{N}_p$ approaches $\frac{1}{d}$ as $d \to \infty$.*

Before we can prove this, we need a combinatorial lemma.

LEMMA 4.5. *Divide an oriented circle into d segments, and into each segment place the numbers $1, 2, \ldots, n$ in increasing order, where $n \geq d - 1$. If pairs of the same symbol are joined by non-intersecting arcs in the disk until no more pairs can be joined, the disk will be divided into d regions. Each region will have all n symbols occurring exactly once on the boundary of each region. (Count the paired symbols only once).*

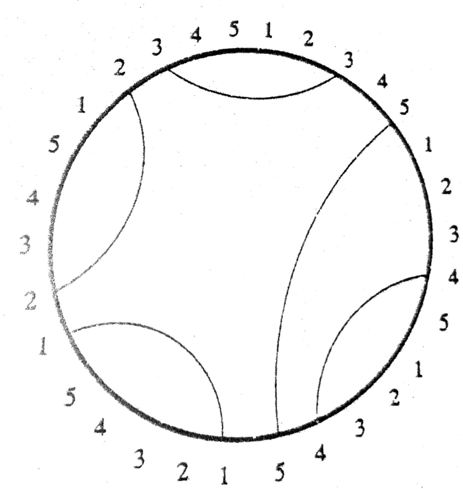

FIGURE 4.1: A circle with $d = 6$ segments, each containing $n = 5$ symbols.

PROOF. We use strong induction on d. For $d = 1$, the disk already contains 1 region, with all symbols only once on the boundary.

Now assume the lemma is true for all $m < d$. Connecting a pair of symbols divides the disk into 2 regions. After identifying the paired symbols, we have 2 disks, one containing k blocks of n symbols, and the other containing $d - k$ such blocks. We may not reuse the symbol just paired, but since the hypothesis was assumed true for all pairings, this causes no problem. After all the pairings are made, we have

$$1 + (k - 1) + (d - k - 1) = d - 1$$

arcs, giving us d regions with exactly n symbols on the boundary of each. $\square$

Proof of Proposition 4.4. By Proposition 4.3, we know that p maps the stable manifolds of the saddles to rays of constant argument which terminate at the critical values of p. The entire stable manifold of any given saddle must all be mapped onto the same ray, namely that one with the same argument as the corresponding critical value. For $|z|$ large, $p(z)$ is arbitrarily close to z^d (in the spherical metric). If we label the regions between the rays with the symbols $1, 2, \ldots, d-1$ on some circle of large radius, we can pull these back by $z^{1/d}$ to induce a labeling of the inverse circle. This labeling has d blocks of $d-1$ symbols, as in Lemma 4.5. Since $|z|$ was chosen very large, the stable manifolds of the saddles are close to some inverse image of the rays, and so form a pairing of the symbols. Therefore, each region contains $1/d$ of the total circumference. $\square$

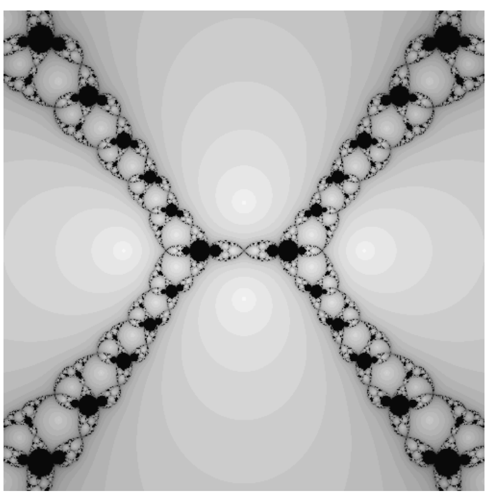

Figure 4.2: Newton's method for the polynomial $(z^2-1)(z^2+.16)$, which has 2 attracting periodic orbits. This is the same picture as Figure 2.4.

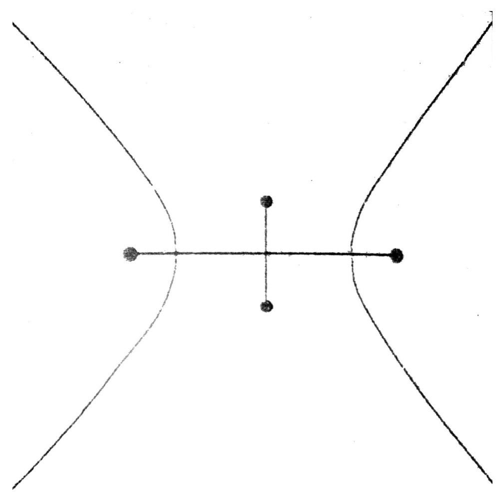

Figure 4.3: The stable manifolds of the sinks and saddles for the flow $\mathcal{N}_p$, where $p(z) = (z^2-1)(z^2+.16)$. The sinks are marked with black dots.

We have done a number of computer experiments which seem to indicate a close relationship between the Newton flow $\mathcal{N}_p$ and the gross structure of the Newton map N_p. Roughly stated, the Julia set for the map lies "near" the stable manifolds of the saddles of the flow. A proof of this statement, coupled with Proposition 4.4, would give an alternate version of our Corollary 3.15. This same technique should also give estimates for the behavior of the relaxed Newton's method $N_{h,p}$.

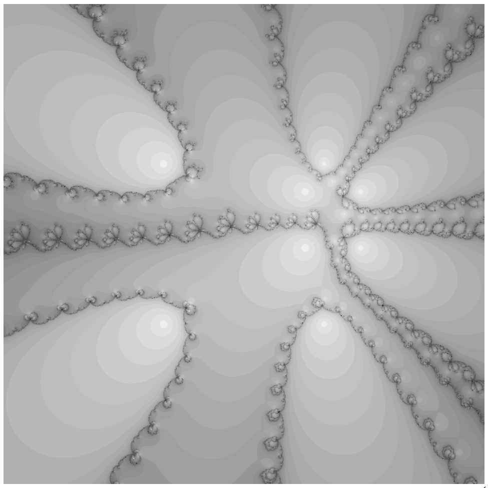

FIGURE 4.4: Newton's Method applied to $p(z) =$ $(z^4+4)(z-1.4535-.6535i)$ $(z-1.4535+.0465i)(z-.7465-.6535i)(z-.7465+0.0465i)$.

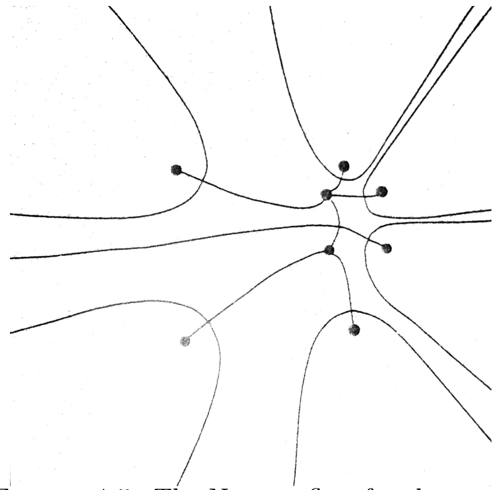

FIGURE 4.5: The Newton flow for the same polynomial.

4.3 Finding Bad Polynomials

In [S2] (problem 6), Smale asks for ways to find polynomials $p(z)$ for which $N_p(z)$ has attracting periodic sinks. It is our belief that such polynomials occur near those whose Newton flow $\mathcal{N}_p$ has saddle connections.

More precisely, let $p_0(z)$ be a polynomial for which $\mathcal{N}_{p_0}$ has saddle connections, and let $\mathcal{F}(p_0)$ be the family of polynomials which can be obtained from p_0 by a continuous deformation of the roots which does not break the saddle connections in the corresponding flow. Then it is our conjecture that there is a $p_\omega \in \mathcal{F}(p_0)$ for which the Newton map N_{p_ω} has an attracting periodic orbit. Furthermore, we also conjecture that if N_p has a periodic attractor, then there is a deformation of p to a polynomial p_ω for which N_{p_ω} has saddle connections; this deformation can be made so that each Newton map N_{p_t} has a periodic attractor.

4.4 A Family of Fourth Degree Newton's Methods

In [DH] and [CGS], the parameter space for Newton's method on a family of cubic polynomials is studied. For certain ranges of parameter values, copies of the Mandelbrot set appear in the pictures. This behavior is explained in [DH] by the theory of polynomial-like mappings.

Definition. Let U and U' be open sets isomorphic to $\mathbf{D}$ with U' relatively compact in U. A map $f : U' \to U$ is said to be *polynomial-like of degree d* if it is a proper holomorphic map whose local degree on U' is d.

Let K_f be the set of $z \in U'$ such that $f^n(z)$ is defined and belongs to U' for all $n \geq 0$. If $f : U' \to U$ and $g : V' \to V$ are two polynomial-like mappings, we say that they are *quasi-conformally equivalent* if there is a quasi-conformal homeomorphism

$$\phi : U_1 \to V_1,$$

where U_1 and V_1 are neighborhoods of K_f and K_g, satisfying

$$g \circ \phi = \phi \circ f \text{ on } f^{-1}(U_1).$$

The Straightening Theorem. ([DH]) *Let f be a polynomial-like mapping of degree d with K_f connected. Then f is quasi-conformally equivalent to a polynomial P of degree d.*

Let M_f be a connected region in parameter space for which the corresponding map f is polynomial-like of degree d with connected K_f. Then, under certain conditions, M_f is a quasi-conformal image of a cover of the connectedness locus for degree d polynomials. (See [DH] for the conditions).

When there is a periodic attracting orbit for Newton's method, it is often true that N_p is polynomial-like on some set U containing the attracting orbit. The figures in this section are studies of the parameter space of Newton's method for the family of quartic polynomials

$$p_c(z) = (z^2 - 1)(z - c)(z - \overline{c}).$$

This family has two free critical points, and thus may have at most two attracting periodic orbits. The free critical points are either real or complex conjugate; when they are complex conjugate, so are their orbits. But when they are real, their orbits may have different behaviors. Denote the free critical points γ_1 and γ_2. By Theorem 1.2, if there is an attracting periodic orbit, it must attract at least one of points γ_i.We follow both critical orbits for a maximum of m iterations of N_{p_c}. We then color the point c the grey tone corresponding to $n_1 + n_2$, where n_i is the number of iterations it takes for $N^n_{p_c}(\gamma_i)$ to get within ϵ of a root of p_c (or m if it never does). Thus, the dark regions in the figures correspond to those c

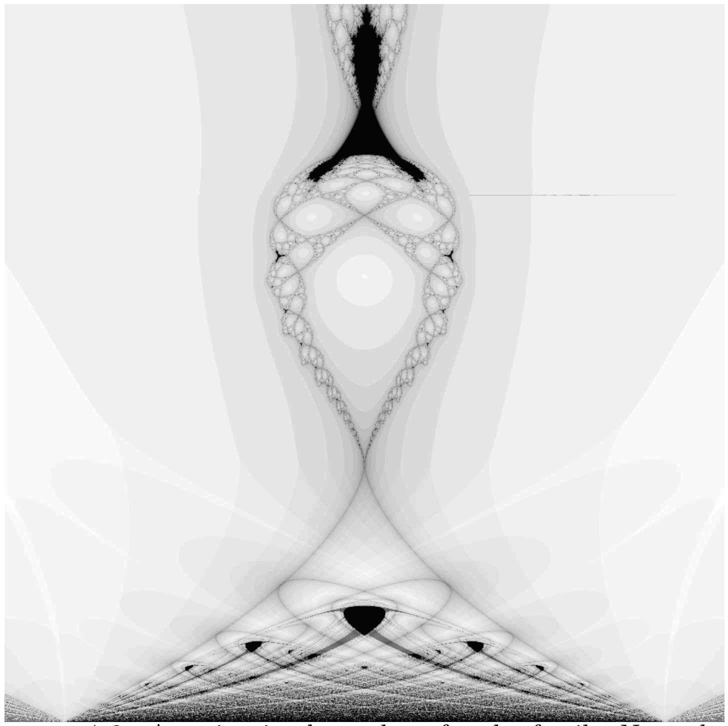

FIGURE 4.6: A region in the c-plane for the family N_{p_c}, where $p_c(z) = (z^2 - 1)(z - c)(z - \overline{c})$.

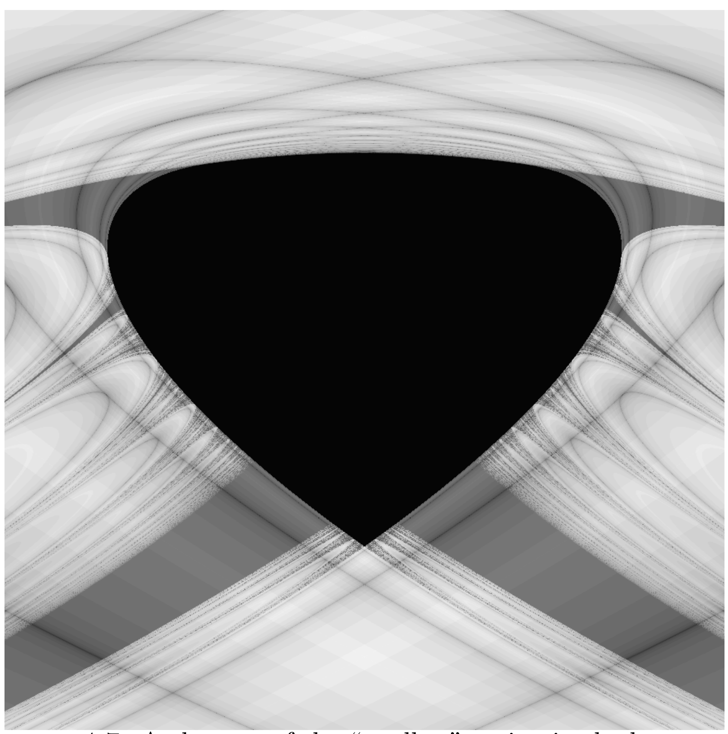

FIGURE 4.7: A close-up of the "swallow" region in the lower part of Figure 4.6.

values for which N_{p_c} has an attracting periodic orbit; the black regions are those where both critical points are attracted to a periodic orbit.

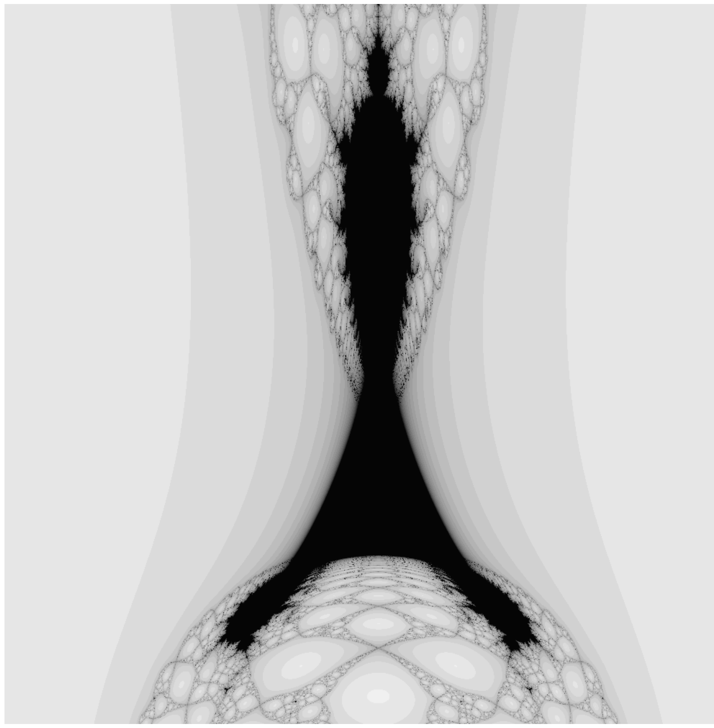

FIGURE 4.8: A close-up of the "tricorn" region on the head of the snake in Figure 4.6.

Figure 4.7 and Figure 4.8 show close-ups of Figure 4.6. In both these regions, N_{p_c} is cubic-like near the orbits of the critical points. Not surprisingly, we see regions which look like the slices of the cubic connectedness locus studied by John Milnor in [M]. Milnor refers to these regions as a swallows (Figure 4.7) and tricorns (Figure 4.8). In Milnor's models, these configurations occur when a disk is mapped over itself by a composition of two quadratic mappings

$$Q_a(z) = z^2 + a \quad \text{and} \quad Q_b(z) = z^2 + b.$$

The swallow occurs when a and b are both real; the tricorn when $b = \overline{a}$. This is precisely the behavior we have for the maps N_{p_c}.

Newton's method for the polynomial $(z^2-1)(z^2+.16)$ occurs at the center of the swallow in Figure 4.7; a polynomial with the same dynamics occurs in the tricorn in Figure 4.8. See Figure 2.4 or Figure 4.2 for pictures of this polynomial.

Appendix: Parameters for the Pictures

Here we give the values used to compute the pictures. The parameter "window" gives the pair of complex numbers which are the lower left and upper right corners of the picture. We stop iterating Newton's method when either the iterate comes within ϵ of a root, or when the number of iterations exceeds the value of "max its". These pictures were all computed using *Citool* [BSV], a program developed at Boston University for creating and maintaining such pictures.

Figure 1.2

title : Newton's method for a quadratic polynomial
polynomial: $z^2 - 1$
slices : 1001
window : $-4 - 4i, 4 + 4i$
max its : 30
ϵ : 0.000001

Figure 2.2

title : Newton's method for the cube roots of 1
polynomial: $z^3 - 1$
slices : 1000
window : $-2 - 2i, 2 + 2i$
max its : 50
ϵ : 0.00000001

Figure 2.3

title : Degree 3 N with a period 2 attractor
polynomial: $(z + 0.635445 + 0.140996i)(z - 0.364555 + 0.140996i)$
$(z - 0.27089 - 0.281992i)$
slices : 500
window : $-2 - 2i, 2 + 2i$
max its : 30
ϵ : .000001

Figure 2.4

title	: Degree 4 N with 2 period 2 attractors
polynomial	: $(z^2-1)(z^2+0.16)$
slices	: 500
window	: $-2-2i, 2+2i$
max its	: 30
ϵ	: .000001

Figure 2.5

title	: Sylvester the cat
polynomial	: $(z^3-i)(z+2i)$
slices	: 1000
window	: $-3-3i, 3+3i$
max its	: 40
ϵ	: .000001

Figure 2.6

title	: Sylvester Squints
polynomial	: $(z+i)(z+0.6+1.6i)(z+0.6-0.45i)(z-0.8660254-0.5i)$
slices	: 1000
window	: $-3-3i, 3+3i$
max its	: 40
ϵ	: .000001

Figure 2.7

title	: N for a "random" degree 10 polynomial
polynomial	: $z(z+1-i)(z+1+i)(z+0.494781+1.25887i)(z-.7465+.0465i)$ $(z-1.79749+0.507307i)(z-0.920668-0.244259i)(z+.3-.2i)$ $(z-0.682672-1.47182i)(z-0.244259+0.0814196i)$
slices	: 1000
window	: $-3-3i, 3+3i$
max its	: 30
ϵ	: .000001

Figure 2.8

title : N for another "random" degree 10 polynomial
polynomial: $(z+1.5)(z-.27i)(z+.16i)(z-.3)(z-.8-.2i)(z-.7-.3i)$
$(z-.75+.5i)(z-.81)(z-.9-.5i)(z-9+.25i)$
slices : 1000
window : $-3-3i, 3+3i$
max its : 30
ϵ : .000001

Figure 2.9

title : N for a polynomial with a double root
polynomial: $(z^2+1)(z-1)^2$
slices : 1000
window : $-3-3i, 3+3i$
max its : 40
ϵ : .000001

Figure 2.10

title : Newton's method for the eighth roots of 1
polynomial: z^8-1
slices : 500
window : $-2-2i, 2+2i$
max its : 60
ϵ : .000001

Figure 4.2

title : Degree 4 N with two period 2 attractors
polynomial: $(z^2-1)(z^2+0.16)$
slices : 500
window : $-2-2i, 2+2i$
max its : 30
ϵ : .000001

Figure 4.3

title : Newton Flow for Figure 4.2
polynomial: $(z^2-1)(z^2+0.16)$
window : $-2-2i, 2+2i$

Figure 4.4

title	: N for a degree 8 polynomial
polynomial:	$(z^4+4)(z-1.4535-.6535i)(z-1.4535+.0465i)$ $(z-.7465-.6535i)(z-.7465+0.0465i)$
slices	: 500
window	: $-3-3i, 3+3i$
max its	: 30
ϵ	: .000001

Figure 4.5

title	: Newton Flow for Figure 4.5
polynomial:	$(z^4+4)(z-1.4535-.6535i)(z-1.4535+.0465i)$ $(z-.7465-.6535i)(z-.7465+0.0465i)$
window	: $-3-3i, 3+3i$

Figure 4.6

title	: A snake with a hat; parameter plane for N_{p_c}
$p_c(z)$	: $(z^2-1)(z-c)(z-\overline{c})$
slices	: 1000
window	: $-1.2, 1.2+2.8i$
max its	: 50
ϵ	: .000001

Figure 4.7

title	: a swallow in the snake; parameter plane for N_{p_c}
$p_c(z)$	: $(z^2-1)(z-c)(z-\overline{c})$
slices	: 1000
window	: $-0.08114+0.282i, 0.08114+0.493526i$
max its	: 50
ϵ	: .000001

Figure 4.8

title	: the hat of the snake; parameter plane for N_{p_c}
$p_c(z)$	: $(z^2-1)(z-c)(z-\overline{c})$
slices	: 1000
window	: $-0.430188+2.0295i, 0.420188+2.82014i$
max its	: 50
ϵ	: .000001

References

[A] L. Ahlfors: *Lectures on Quasiconformal Mappings*, Wadsworth, Monterey, CA, 1987.

[Ba] B. Barna: Über die Divergenzpunkte des Newtonschen Verfahrens zur Bestimmung von Wurzeln Algebraischer Gleichungen, *Publ. Math. Debrecen* **4** (1956), 384–397.

[Bl] P. Blanchard: Complex Analytic Dynamics on the Riemann Sphere, *Bull. Amer. Math. Soc.* **11** (1984), 85-141.

[BSV] P. Blanchard, S. Sutherland, G. Vegter: *Citool*, computer software, Boston Univ. Math. Dept., 1986.

[BD] B. Branner and A. Douady: Surgery on Complex Polynomials, Preprint, Matematisk Institut, Danmarks Tekniske Højskole, April 1987.

[Bu] R. Burckel: *An Introduction to Classical Complex Analysis, Vol. 1*, Academic Press, New York, 1979.

[Cj] F. Cajori: Historical Note on the Newton-Raphson Method of Approximation, *Amer. Math. Monthly* **18** (1911), 29–33.

[Cj1] F. Cajori: *A History of Mathematics, (Third Edition)*, Chelsea, New York, 1980.

[Ca] A. Cayley: The Newton–Fourier Imaginary Problem, *Amer. J. Math* **II** (1879), 97.

[Ca1] A. Cayley: On the Newton–Fourier Imaginary Problem, *Proc. Cambridge Phil. Soc.* **3** (1880), 231-232.

[Ca2] A. Cayley: Application of the Newton–Fourier Method to an Imaginary Root of an Equation, *Quart. J. of Pure and Applied Math.* **XVI** (1879), 179–185.

[Ca3] A. Cayley: Sur les Racines D'une Équation Algébrique, *Comptes Rendus Acad. Sci. Paris* **t.CX** (Janvier–juin 1890), 174–176, 215–218.

[CGS] J. Curry, L. Garnett, and D. Sullivan: On the Iteration of a Rational Function: Computer Experiments with Newton's Method, *Comm. Math. Phys.* **91** (1983), 267–277.

[DH] A. Douady and J. Hubbard: On the Dynamics of Polynomial–like mappings, *Ann. Sci. Ecole Norm. Sup., 4e série* **t.18** (1985), 287–343.

[DH1] A. Douady and J. Hubbard: Itération des Polynômes quadratiques complexes, *Comptes Rendus Acad. Sci. Paris* **294**, Serie I

[DH2] A. Douady, and J. Hubbard: *Etude Dynamique des Polynômes Complexes I*, Publications Mathématiques d'Orsay **84–02** (1984).

[DH3] A. Douady, and J. Hubbard: *Etude Dynamique des Polynômes Complexes II*, Publications Mathématiques d'Orsay **85–04** (1985).

[DM] P. Doyle and C. McMullen: Solving the Quintic by Iteration, Preprint, Princeton University, 1988.

[Du] P. Duren: *Univalent Functions*, Springer–Verlag, New York, 1983.

[E] C. H. Edwards: *The Historical Development of the Calculus*, Springer–Verlag, New York, 1979.

[FS] M. Flexor and P. Sentenac: Algorithmes de Newton Généralises, Preprint, Univ. Paris Sud, Orsay France, 1988.

[F1] P. Fatou: Sur les équations fonctionnelles, *Bull. Soc. Math. France* **47** (1919) 161–271.

[F2] P. Fatou: Sur les équations fonctionnelles, *Bull. Soc. Math. France* **48** (1920) 33–94.

[F3] P. Fatou: Sur les équations fonctionnelles, *Bull. Soc. Math. France* **48** (1920) 208–314.

[F4] P. Fatou: Sur les fonctions holomorphes et bornées à l'intérieur d'un cercle, *Bull. Soc. Math. France* **51** (1923), 191–202.

[Fr] J. Friedman, On the Convergence of Newton's Method, *Proc. 27th Annual Symp. on Foundations of Comp. Sci.* (1986).

[HM] M. Hurley and C. Martin: Newton's Algorithm and Chaotic Dynamical Systems, *SIAM J. Math. Anal.* **16** (1984), 11–20.

[Hu] M. Hurley: Multiple Attractors in Newton's Method, *Ergod. Thy. and Dynam. Sys.* **6** (1986).

[Hu1] M. Hurley: Improbability of Nonconvergence in a Cubic Root Finding Method, *J. Math. Anal. and Applications*, 1987.

[JJT] H. T. H. Jongen, P. Jonker, and F. Twilt: The Continuous Desingularized Newton's Method for Meromorphic Functions, Preprint, Dept. of Applied Math., Twente University of Technology, the Netherlands.

[J] G. Julia: Iteration des Applications Fonctionelles, *J. Math. Pures Appl.* **4** (1918), 47-245.

[K1] M.-H. Kim: Topological complexity of a root finding algorithm, Preprint, Bellcore, Morristown, NJ, 1989.

[K2] M.-H. Kim: On approximate zeros and rootfinding algorithms for a complex polynomial, *Math. of Computation* **51** (1988), 707-719.

[M] A. Manning: How to be Sure of Solving a Complex Polynomial using Newton's Method, Preprint, Mathematics Insistitute, University of Warwick, Nov. 1986

[Ma] B. Mandelbrot: *The Fractal Geometry of Nature*, Freeman & Co., San Francisco, 1982.

[Mr] M. Marden: *The Geometry of Polynomials*, Amer. Math. Soc, Providence, R.I., 1966.

[Mc1] C. McMullen: Families of Rational Maps and Iterative Root Finding Algorithms, *Annals of Math.* **125** (1987), 467–493.

[Mc2] C. McMullen: Braiding of the Attractor and the Failure of Iterative Algorithms, *Inv. Math.*, to appear.

[Mi] J. Milnor: Remarks on Iterated Cubic Maps, Preprint, Institute for Advanced Study, August 1987.

[PR] H.-O. Peitgen and P. Richter: *The Beauty of Fractals*, Springer-Verlag, Heidelberg, 1986.

[PS] H.-O. Peitgen and D. Saupe, eds.: *The Science of Fractal Images*, Springer-Verlag, New York, 1988.

[PSH] H.-O. Peitgen, D. Saupe, and F. v. Haessler: Cayley's Problem and Julia Sets, *Math. Intelligencer* **6** (1984), 11–20.

[Pz] F. Przytycki: Remarks on the Simple Connectedness of Basins of Sinks for Iterations of Rational Maps, Preprint, Polish Academy of Sciences, Warsaw, 1987.

[Ra] T. Radó: Zur Theorie der Mehrdeutigen Konformen Abbildungen, *Acta Litt. ac Scient. Univ. Hung. (Szeged)* **1** (1922/23), 55–64.

[SU] D. Saari and J. Urenko: Newton's Method, Circle Maps, and Chaotic Motion, *Amer. Math. Monthly* **91** (1984), 3–17.

[Sh] M. Shub: The Geometry and Topology of Dynamical Systems and Algorithms for Numerical Problems, notes from lectures given at D.D.4 Peking University, Beijing, China, Aug.–Sept. 1983.

[SS1] M. Shub and S. Smale: Computational Complexity: on the Geometry of Polynomials and a Theory of Cost I, *Ann. Sci. Ecole Norm. Sup.* **18**(1985), 107–142.

[SS2] M. Shub and S. Smale: Computational Complexity: on the Geometry of Polynomials and a Theory of Cost II, *SIAM J. Comp.* **15**(1986), 145–161.

[STW] M. Shub, D. Tischler, and R. Williams: The Newtonian Graph of a Complex Polynomial, *SIAM J. Math. Anal.* **19** (1988), 246–256.

[S1] S. Smale: The Fundamental Theorem of Algebra and Complexity Theory, *Bull. Amer. Math. Soc.* **4** (1981), 1–36.

[S2] S. Smale: On the Efficiency of Algorithms of Analysis, *Bull. Amer. Math. Soc.* **13** (1985), 87–121.

[S3] S. Smale: Algorithms for Solving Equations, *Proceedings of a Conference in Honor of Gail S. Young*, Springer, New York, 1986.

[S4] S. Smale: Newtons's Method Estimates from Data at One Point, *Proceedings of the International Congress of Mathematicians*, 1986.