CONSTRUCTING RATIONAL ITERATING FUNCTIONS WITHOUT ATTRACTIVE CYCLES

a thesis submitted

by

William F. Smyth

to

The School of Graduate Studies

University of Ottawa

in partial fulfillment of the requirements

for the degree of

Master of Science

in the subject of

Mathematics

April 1983

Revised June 1983

CONTENTS

1 Abstract

The main results related to the global convergence properties of iterations of rational functions are surveyed, especially those related to the existence and properties of attractive cycles.

Suppose integers $n > 1$ and $\sigma > 2$ are given, together with n distinct points $z_1, \cdots, z_n$, in the complex plane. Define $C_M = C_M(\sigma; z_1, \cdots, z_n)$ to be the class of rational functions $f_{p,q}(z) = g_p(z)/h_q(z)$ (where g and h are polynomials of degree $p > 1$ and $q > 1$, respectively) such that (1) $p + q + 1 = M$, (2) f when iterated converges with order σ at each z_i, $i = 1, \cdots, n$. Then if $M < \sigma n$, C_M is null; if $M = \sigma n$, C_M contains exactly σn elements. For every $M > \sigma n$, all the elements of C_M may be constructed by expressing, for each choice of p and q which satisfies $p + q + 1 = M$, the coefficients of g_p and h_q in terms of $M - \sigma n$ arbitrarily chosen values. In fact, these coefficients are expressed in terms of generalized Newton sums $S_n^{j,k} = S_n^{j,k}(z_1, \cdots, z_n)$, $1 < j < n$, $k > n$, which it is shown may be calculated by recursion from the normal Newton sums $S_n^{j,n}$. Hence, given a polynomial $\phi_n(z)$ with n distinct (unknown) zeros $z_1, \cdots, z_n$, it is possible to construct all $f_{p,q}(z)$ which converge to the z_i with order σ: in the case $\sigma = 2$, the choice $p = n$, $q = n - 1$, yields the Newton-Raphson iteration $f_{n,n-1} \varepsilon C_{2n}$; the Schröder and König iterations are shown to be elements of $C_{2(2\sigma-3)(n-1)+2}$ and $C_{2(\sigma-1)(n-1)+2}$, respectively. It is shown by examples that there exist cases in which $f_{n,n-1}$ has an undesirable property (attractive cycles) not shared by other iterating functions in the same class C_{2n}.

2 Introduction

The advent of the digital computer since 1945 has greatly stimulated interest in numerical procedures, or algorithms, of all kinds [K1-K3], including in particular the computational algorithms required for engineering or scientific calculations [D1-D2] together with their mathematical foundations [H1-H2, H4-H5, B1]. Of major importance among these computational algorithms are iterations [T1].

An iteration consists of a more-or-less simple procedure, usually expressable as a function evaluation, applied repeatedly to a given set of initial values. The purpose of iteration is usually to determine the solutions of some given equation. Normally, then, the function used by the iteration will be chosen so that, when applied over and over to the initial values, it results in convergence to the solutions of the given equation. Regrettably, however, convergence to the desired solutions does not as a rule take place from any set of initial values: it is usually necessary to choose initial values which are in some sense "sufficiently close" to the solution, a process which may well be difficult and which often depends heavily on the particular characteristics of the equations to be solved.

There is motivation, then, in order to avoid such special cases and special methods, for studying the global convergence properties of iterations: to see what may be said in general about the convergence of iterations, and, if possible, to use this information to devise general methods for the construction of iterating functions with desirable convergence properties from any, or almost any, set of initial values. The foundations of such a study were in fact laid, some 25 years before the invention of the electronic digital computer, by two French mathematicians, working independently, G. Julia [J1] and P. Fatou [F1-F3]. Useful summaries of their work may be found in [C1] and [M1]. Their results were both interesting, from a theoretical point of view, and encouraging, from a practical point of view. They found that, for iterations of rational functions, convergence was assured in cases of practical interest from initial values which might lie "almost anywhere", in domains of convergence of the complex plane. The main difficulty turned out to be that, depending upon the convergence domain of the initial value, convergence might take place, not to solutions of the given problem, but to other sets of points, known as attractive cycles. From a practical point of view, then, it became important to try to develop general methods for the construction of iterating functions not giving rise to attractive cycles, but in the intervening 60 years there seem to have been very few papers [B2, S1] which bear directly on this problem.

In the present work, Chapters 3-5 review the theoretical results related to the global convergence properties of iterations of rational functions, with particular emphasis on the role of attractive cycles. Chapter 6 then develops new results which show how to arrange all the rational iterating functions convergent at n given points into equivalence classes based on the degree of the function and the degree of the iteration's convergence. In particular, given a polynomial $\phi_n(z)$ whose (distinct) zeros are to be found, a method is described for the construction of all rational iterating functions converging to the zeros with given order of convergence σ. It is shown by example that there exist cases in which one iteration in an equivalence class has an undesirable property (convergence to attractive cycles) not shared by other iterating functions in the same class. The chapter closes with a brief discussion of the significance of these results, especially of possible future avenues of research.

3 Meromorphic Functions

This chapter reviews definitions and fundamental results related to meromorphic functions, of which rational functions are of course a special case. These results will be required in Chapter 5 to establish the main properties of iterations of rational functions. The selection and use of the results has been strongly influenced by a recent paper of Burckel [B12] in which equivalent properties are developed for iterations of holomorphic functions.

Definition 3.1 A domain is an open arcwise-connected set.

Definition 3.2 A function $f(z)$ is said to be holomorphic at a point z of the finite plane $\pi - \{\infty\}$ iff it is single-valued and differentiable at z. It is holomorphic at $z = \infty$ if $f(1/z)$ is holomorphic at $z = 0$. It is holomorphic in a domain D iff it is holomorphic at every point $z \in D$.

Any polynomial $p(z)$ is holomorphic in the finite plane. A rational function $f(z) = p(z)/q(z)$ is holomorphic at ∞ provided the degree of the polynomial p does not exceed that of the polynomial q. The function $f(z) = z + 1/z$ is holomorphic everywhere in the extended plane π except at 0 and ∞.

Definition 3.3 A function is said to be transcendental entire (or, briefly, entire) iff it is holomorphic in the finite plane and not a polynomial.

Common examples of entire functions are $\sin z$, $\cos z$, e^z.

Definition 3.4 A function is said to have a singularity at w iff it is not holomorphic at w.

Definition 3.5 A singularity w of f(z) is said to be a pole of f(z) of order m iff there exists a least integer m > 0 such that $(z - w)^m f(z)$ is holomorphic at w.

Definition 3.6 A singularity is said to be an essential singularity iff it is not a pole.

Definition 3.7 A function is said to be meromorphic in a domain D iff it has no singularities other than poles in D.

We recall [S3, 147] that a function is meromorphic in the extended plane if and only if it is rational. The six standard trigonometric functions and e^z are meromorphic in the finite plane.

In order to discuss certain properties of meromorphic functions, especially sequences or families of such functions, it is convenient to be able to speak of neighbourhoods of ∞ and convergence to ∞. For this reason, the Riemann sphere σ and the corresponding chordal metric χ are introduced as in [H7, 38-44] and [O1] for both the domain and the range of meromorphic functions. This leads to the following definitions [H8, 241-243]:

Definition 3.8 A function f(z) defined in a domain D is said to be spherically continuous at $w \in D$ iff for each $\varepsilon > 0$ there exists δ

= δ(w, ε) such that

$$\left.\begin{array}{l} z \in D \\ \chi(z,w)<\delta \end{array}\right\} \Longrightarrow \chi[f(z),\ f(w)] < \varepsilon.$$

f(z) is <u>spherically continuous in D</u> iff it is spherically continuous at every point of D.

<u>Definition 3.9</u> A family F of functions defined in a domain D is said to be <u>spherically equicontinuous at w ε D</u> iff for each ε > 0 there exists δ = δ(w, ε) such that for every f ε F

$$\left.\begin{array}{l} z \in D \\ \chi(z,w)<\delta \end{array}\right\} \Longrightarrow \chi[f(z),\ f(w)] < \varepsilon.$$

F is <u>spherically equicontinous in D</u> iff it is spherically equicontinuous at every point of D.

<u>Definition 3.10</u> A sequence of functions $f_n(z)$ defined on a set S is said to converge <u>spherically uniformly on S</u> iff, given any ε > 0, there exists an N = N(ε) such that for all integers m, n > N, $\chi[f_m(z),\ f_n(z)] < \varepsilon$ for every z ε S.

Using these definitions and the properties of the chordal metric χ, it is a relatively straightforward matter [H8, 242] to show that equicontinuity in the complex plane implies spherical equicontinuity, and that uniform convergence implies spherical uniform convergence. In order to state the main results, two more definitions are necessary.

Definition 3.11 A set S is said to be compact iff every infinite subset of S has a limit point in S.

Definition 3.12 A family F of functions meromorphic in a domain D is said to be normal in D iff every infinite sequence chosen from F contains a subsequence which converges spherically uniformly on every compact subset of D. F is said to be normal at a point iff it is normal in some neighbourhood of the point.

The term "normal" is due to Montel [M2]. To fix the idea of normality, consider two simple sequences [C1]

$$\{z/2^n\} = \{z/2,\ z/4,\ z/8,\ \cdots\}$$

$$\{z^{2^n}\} = \{z^2,\ z^4,\ z^8,\ \cdots\}.$$

Both of these sequences are normal in the open unit disk $|z| < 1$, since each of them, and therefore every subsequence of each, converges to zero uniformly, hence spherically uniformly, on every compact subset. The sequence $\{z/2^n\}$ is in fact normal in any open disk $|z| < R$, while the sequence $\{z^{2^n}\}$ is normal in no domain which contains any point of the unit circle $|z| = 1$. On the other hand, $\{z/2^n\}$ is normal in no neighbourhood of ∞, while $\{z^{2^n}\}$ is normal in every domain $|z| > R > 1$. The combined family

$$\{z/2,\ z^2,\ z/4,\ z^4,\ \cdots\}$$

is also normal in the open unit disk, but not in any domain which properly contains it. The combined family is however normal in rings defined by $1 < R_1 < |z| < R_2 < \infty$.

The results required for the development of Chapter 5 may now be stated:

Theorem 3.13 (Uniqueness Theorem) If f and g are meromorphic in a domain D, and if f = g in some non-empty open subset of D, then f = g throughout D [S3, 152].

Theorem 3.14 (Open Map Theorem) The image of a domain under any non-constant meromorphic function is also a domain [S3, 162].

Theorem 3.15 (Weierstrass) A sequence of functions meromorphic in a domain D, which converges spherically uniformly on every compact subset of D, converges to a meromorphic function [H8, 243].

Theorem 3.16 A family of functions meromorphic in a domain D is normal in D iff it is equicontinuous in D [H8, 244].

Theorem 3.17 (Montel) A family of functions meromorphic in a domain D is normal in D if every function in the family omits the same three fixed values in D [H8, 248-249].

Theorem 3.18 If f is meromorphic in a domain D, and $f'(z) \neq 0$ for all $z \in D$, then the inverse function f_{-1} is meromorphic in the domain f(D) [53, 163].

4 Iteration

The study of iteration as a technique for the solution of equations goes back at least to Newton, whose famous iterating function was, according to [K4, 381], first published in 1685 and then improved by Raphson five years later. Other early iterating functions include Halley's, dated by Traub [T1] to 1694, and Euler's (mid eighteenth century).

This chapter begins with a fairly broad, though by no means universal, definition of iteration, illustrates the definition by several examples, and defines various other ideas of importance for the analysis of iterations. It is convenient at this stage to define iteration in terms of an arbitrary complete normed linear space (or Banach space) which gives rise in the usual way to a Banach algebra [H4, 67-71]. The norm of a point or element z of the space will as is customary be denoted $\|z\|$. As indicated in Chapter 3, in order to make use of Theorems 3.13 - 3.18, later work with rational functions will require further specialization: the Banach space will be assumed to be either the complex plane π or the Riemann sphere σ, with corresponding norms $|\ |$ and χ, respectively.

<u>Definition 4.1</u> Let X^m be a Banach space of dimension $m > 0$, and suppose $f: X^m \to X^m$ is defined on a point set $S \subseteq X^m$. Then an <u>iteration</u> — more precisely, the <u>sequence of iterates</u> of f at a point $z \in S$ — is given by

$$F(z) = \{f_n(z)\} = \{f_n(z) | n = 0, 1, \cdots \},$$

where $f_0(z) = z$ and $f_{n+1}(z) = f(f_n(z))$, $n = 0, 1, \cdots$, provided $f_n(z) \in S$. To represent an iteration everywhere in S, the shorter form $F = \{f_n\}$ may also be used.

$f(z)$ is called an immediate consequent of z (of course unique for single-valued f) and $f_{-1}(z)$ denotes a particular immediate antecedent of z (that is, a point w, if it exists, such that $f(w) = z$). More generally, $f_n(z)$ and $f_{-n}(z)$ denote an nth consequent and a particular nth antecedent, respectively, of z. The notation $\{f_{-n}(z)\}$ is used to denote an inverse iteration based on the repeated application of a particular f_{-1}, and $[f_{-n}(z)]$ will denote the set (perhaps empty, perhaps uncountable) of all nth antecedents of z. If it is supposed (see below, Axiom 4.8(d)) that every antecedent of $z \in S$ is also a point of S, then

$$[f_{-(n+1)}(z)] = [f_{-1}([f_{-n}(z)])].$$

(In the terminology of general iteration theory [K5, T4], F(z) is called the splinter (f-splinter) of z, and F(z) together with all its antecedents is called the orbit (f-orbit) of z. A basic result of this theory is that the points of an orbit form an equivalence class under the relation $x \sim_f y$ (x and y have a common consequent).)

Definition 4.2 An accumulation point α of F(z), $z \in S$, is a point in whose every neighbourhood there exists an infinite number of elements of F(z). α is then said to be the (unique) limit point of some subsequence $\{f_{n_k}(z)\} \subseteq F(z)$.

In connection with this definition, it should be remarked that the sequence F may give rise to limit points even when it covers only a finite number of points. For if F is an <u>infinite</u> sequence whose elements take only a finite number of values, then there exist least integers $r > 0$ and $s > 0$ such that $f_r(f_s(z)) = f_s(z)$. In this case

$$F(z) = \{z,\ f(z),\ \cdots,\ f_s(z),\ \cdots,\ f_{r-1}(f_s(z)),\ f_s(z),\ \cdots\}$$

and setting $\alpha = f_s(z)$, it follows that the distinct points of F(z) are $\{z,\ f(z),\ \cdots,\ f_{s-1}(z),\ \alpha,\ f(\alpha),\ \cdots,\ f_{r-1}(\alpha)\}$. If F(z) were thought of merely as a set of points, it would thus have no limit point; in accordance with Definition 4.2, however, it possesses limit points $\alpha,\ \cdots,\ f_{r-1}(\alpha)$.

From a practical point of view, the most interesting instance of limit points arises when, from a particular initial value z, the entire sequence of iterates converges to a single limit point α. Usually, therefore, the iterating function f is chosen in such a way that when the entire sequence F converges to a limit point α, then $\phi(\alpha) = 0$ — that is, a solution of a given equation is determined.

<u>Example 4.3</u> Suppose solutions are required to a given equation $\phi(z) = 0$, where the variable z represents a point in the complex plane π. Then, referring to Definition 4.1, $\pi = X^2$ may be identified as the Banach space of dimension $m = 2$. Consider the seventeenth-century iterating functions mentioned at the beginning of this chapter. If ϕ' and ϕ'' denote the first and second derivatives,

respectively, of ϕ at z, then these iterations may be written:

$$f(z) = z - \phi/\phi' \qquad \text{(Newton-Raphson)}$$

$$f(z) = z - \phi/[\phi' - \phi\phi''/(2\phi')] \qquad \text{(Halley's).}$$

From certain initial values z, the sequence of iterates generated by these functions would, hopefully, determine limit points α such that $\phi(\alpha) = 0$. Note that, in each case, when $\phi(\alpha) = 0$, and $\phi'(\alpha) \neq 0$, it follows that $f(\alpha) = \alpha$; that is, solutions of the given equation $\phi(z) = 0$ are (provided they are not also solutions of $\phi'(z) = 0$) points $z = \alpha$ at which the sequence of iterates $F(\alpha)$ reduces to $\{\alpha, \alpha, \cdots\}$.

Example 4.4 Definition 4.1 does not restrict f to be single-valued. It therefore includes iterations such as Laguerre's [D1, 72-74] for determining the roots of a polynomial $\phi(z)$ of degree k:

$$f(z) = z - k\phi/(\phi' \pm \sqrt{H}),$$

where again the Banach space $X^2 = \pi$ and $H(z) = (k - 1)^2(\phi')^2 - k(k - 1)\phi\phi''$. Euler's iteration, mentioned at the beginning of this chapter, is the special case k = 2 of Laguerre's. Note that as in 4.3 when $\phi(\alpha) = 0$ and $\phi'(\alpha) \neq 0$, $f(\alpha) = \alpha$.

Example 4.5 The Newton-Bairstow method [H6, 110-113] converges to real values u and v which define a quadratic factor $z^2 - uz - v$.

extracted from a given polynomial

$$\phi(z) = a_0 z^k + a_1 z^{k-1} + \cdots + a_k$$

with real coefficients and degree $k > 2$. First the quantities b_h and c_h are introduced:

$$b_h(u, v) = \sum_{j=0}^{h} a_j u^{2S_{hj}} (1 + v/u^2)^{R_{hj}}$$

$$c_h(u, v) = \sum_{j=0}^{h} b_j u^{2S_{hj}} (1 + v/u^2)^{R_{hj}},$$

$h = 0, \cdots, k$, where $S_{hj} = (h - j)/2$ and R_{hj} is the integral part of S_{hj}. Then the Newton-Bairstow method is defined by

$$f(u, v) = [u + \delta(u, v), v + \varepsilon(u, v)],$$

where

$$\delta(u, v) = (b_k c_{k-3} - b_{k-1} c_{k-2})/(c_{k-2}^2 - c_{k-1} c_{k-3})$$

$$\varepsilon(u, v) = (b_k c_{k-1} - b_k c_{k-2})/(c_{k-2}^2 - c_{k-1} c_{k-3}).$$

As in the above examples, this is an iteration in the sense of Definition 4.1, and $m = 2$; but now, instead of the complex plane, the Banach space is R^2, the Cartesian product of the reals; and instead of $|\ |$, the norm is the Euclidean norm.

Example 4.6 Given a set of m real simultaneous equations $Ax = b$ and an initial estimate $x_0 = (x_0^1, x_0^2, \cdots, x_0^m)$ of the solution vector, a better estimate may be achieved by means of a vector iteration [F4, 117-127]

$$x_{n+1} = (I - A)x_n + b.$$

In terms of Definition 4.1, this iteration is equivalent to

$$f(x) = (I - A)x + b,$$

where the Banach space of the points $x = (x^1, \cdots, x^m)$ is R^m and an appropriate norm might be $\|x\| = \sum_{j=1}^{m} |x^j|$.

In terms of the components x^j, $j = 1, \cdots, m$, this iteration could also be expressed by m simultaneous "iterative" relations in R^1:

$$x_{n+1}^j = x_n^j - \sum_{k=1}^{m} a^{jk}\, x_n^k + b_j, \quad j = 1, \cdots, m,$$

where the a^{jk} represent the elements of the $m \times m$ matrix A; but these relations would not constitute iterations in the sense of Definition 4.1, because at each stage x_{n+1}^j is computed not only from x_n^j, but also from $x_n^1, \cdots, x_n^{j-1}, x_n^{j+1}, \cdots, x_n^m$.

Example 4.7 The above examples have shown that the definition of iteration given above is sufficiently general to include a variety of different cases. That it does not, however, include all instances of what is commonly called "iteration" is shown by the simple example of the "iterated sine" [T3]. Given an initial real value x, this iteration may be described by

$$g_0(x) = x, \; g_{n+1}(x) = \sin\,[g_n(x)],$$

$$f_{n+1}(x) = n[g_n(x)]^2$$

giving rise to the sequence

$$\{\sin^2 x, \; 2\sin^2(\sin x), \; \cdots\}.$$

In order to include this sequence in the definition of iteration, it would be necessary both to permit the iterating function f to be a function of n as well as z, and to introduce an "auxiliary" iterating function g. This is an example of a non-stationary iteration with memory: by virtue of its non-stationary character, it would fall outside of Traub's classification [T1], but it would be included in the very general definition of iteration given by Brent et al. [B9]. The Brent definition however covers only a real, one-dimensional iterate.

On the other hand, it has been remarked [H11] that the iterated sine may be imbedded in an R^3 iteration which does satisfy Definition 4.1. Given a point $z = (z^1, z^2, z^3) \in R^3$, define f as follows:

$$f(z^1, z^2, z^3) = (\sin z^1, \; z^1 \cdot z^1 \cdot z^3, \; z^3 + 1).$$

Then from an initial value $z_0 = (x, 0, 0)$, the sequence $F(z_0)$ yields the iterated sine as the middle component of $f_n(z_0)$, $n = 2, 3, \cdots$.

In order now to state and prove some very elementary properties of iterations, certain restrictions on the iterating function f and the point set S are introduced. It shall from now on be required that the following axioms be satisfied:

<u>Axioms 4.8</u> For every $z \in S$,

(a) $f(z)$ is single-valued
(b) $f(z)$ is continuous
(c) $f(z) \in S$
(d) $[f_{-1}(z)] \subseteq S$.

Based on these axioms (compare with [S2]), some of the fundamental properties of a sequence of iterates may be established.

<u>Lemma 4.9</u> For every $z \in S$ and integers $s > t > 0$,

(a). $f_s(f_t(z)) = f_{s+t}(z) = f_t(f_s(z))$

(b) $f_s(f_{-t}(z)) = f_{s-t}(z) \in [f_{-t}(f_s(z))]$

(c) $[f_t(f_{-s}(z))] = [f_{t-s}(z)] \subseteq [f_{-s}(f_t(z))]$

(d) $[f_{-s}([f_{-t}(z)])] = [f_{-(s+t)}(z)] = [f_{-t}([f_{-s}(z)])]$.

<u>Proof</u> To prove (a), note first that by Axiom 4.8(c) every consequent of z exists. Then by Definition 4.1 the first part of (a) is true for $s = 0$. By Definition 4.1 and Axiom 4.8(a) the first part of (a) is also true for $s = 1$. Then suppose for $s > 1$, $f_{s-1}(f_t(z)) = f_{s-1+t}(z)$. Applying f to each side of this equality yields, by Definition 4.1 and Axiom 4.8(a), $f_s(f_t(z)) = f_{s+t}(z)$. Hence by induction the first part of (a) is true. The same argument with the roles of s and t reversed shows that the second part of (a) is also true. Then (a) is true.

To prove (b), note that by Axiom 4.8(d) every antecedent of z exists. Then for <u>every</u> $w = f_{-t}(z)$, it follows that $f_t(w) = z$. Since $s - t > 0$, this means by (a) that

$$f_s(f_{-t}(z)) = f_s(w) = f_{s-t}(f_t(w)) = f_{s-t}(z),$$

so that the first part of (b) is true. To prove the second part of (b), observe that by (a)

$$f_t(f_{s-t}(z)) = f_s(z),$$

which makes clear that $f_{s-t}(z)$ is a t^{th} antecedent of $f_s(z)$. Hence $f_{s-t}(z) \in [f_{-t}(f_s(z))]$, and (b) is true.

Similar arguments may be employed to establish (c) and (d). ∎

Lemma 4.10 For every integer $r > 0$, f_r satisfies Axioms 4.8.

Proof By Definition 4.1 the result is trivially true for $r = 1$. Suppose it is true for a specific value $r > 1$. Then by Axiom 4.8(c), $f_r(z) \in S$, from which it follows by Axiom 4.8(c) and Definition 4.1 that $f_{r+1}(z) = f(f_r(z)) \in S$, so that f_{r+1} satisfies 4.8(c). Similar arguments establish that f_{r+1} satisfies every Axiom 4.8. Then the Lemma follows by induction. ∎

Lemma 4.11 If for some $w \in S$, $\alpha = \lim f_{n_k}(w) \in S$, then for every integer

$$r > 0,\ f_r(\alpha) = \lim f_{n_k+r}(w).$$

Proof The result will be proved for $r = 1$. By induction it will hold then for every integer $r > 0$.

By Axiom 4.8(c) $f(\alpha) \in S$ and by Axiom 4.8(b) f is continuous at α. Then given any real number $\varepsilon > 0$, a real number δ may always be found such that, for $z \in S$,

$$|z - \alpha| < \delta \Longrightarrow |f(z) - f(\alpha)| < \varepsilon,$$

Since $\alpha = \lim f_{n_k}(w)$, it follows that all but a finite number of terms of the subsequence $\{f_{n_k}(w)\}$ are in the neighbourhood $M(\alpha, \delta)$. Then for these terms (that is, for $k > K(\varepsilon)$),

$$|f(f_{n_k}(w)) - f(\alpha)| < \varepsilon.$$

Now suppose the subsequence $\{f_{n_k+1}(w)\}$ has an accumulation point $\beta \neq f(\alpha)$. But then, if a positive ε is chosen such that $\varepsilon < |\beta - f(\alpha)|$, this inequality ensures that all but a finite number of terms $f_{n_k+1}(w)$ lie outside some neighbourhood of β, which can therefore not be an accumulation point of $\{f_{n_k+1}(w)\}$. It may be concluded then that $f(\alpha)$ is the only accumulation point of $\{f_{n_k+1}(w)\}$. ∎

If as in [S2] an Axiom 4.8(e) is added, stating that every accumulation point of $F(z)$ is a point of S, it may also, as in Lemma 1 of [S2], be shown that under the hypothesis of Lemma 4.11 there exists a particular $f_{-r}(\alpha)$ such that $f_{-r}(\alpha) = \lim f_{n_k-r}(w)$.

Lemma 4.12 If for some $w \in S$ and some least integer $r > 0$, $\alpha = \lim f_{nr}(w) \in S$, then for every integer $0 \leq s < r$,

(a) $f_s(\alpha) = \lim f_{nr+s}(w)$

(b) $f_r(f_s(\alpha)) = f_s(\alpha)$.

Proof Result (a) is a straightforward application of Lemma 4.11. Then, applying Lemma 4.11 to (a), (b) follows:

$$\begin{aligned} f_r(f_s(\alpha)) &= \lim f_{nr+s+r}(w) \\ &= \lim f_{(n+1)r+s}(w) \\ &= \lim f_{nr+s}(w) \\ &= f_s(\alpha). \end{aligned}$$ ∎

This lemma provides motivation for the following:

Definition 4.13 A point α is said to be a fixpoint of order r of F iff there exists a least integer $r > 0$ such that $f_r(\alpha) = \alpha$. The set of (distinct) points $\{\alpha, f(\alpha), \cdots, f_{r-1}(\alpha)\}$ is called a cycle of order r.

Definition 4.13 opens the door to the real difficulties of iteration theory. The main reason for this is that the behaviour of F(z) for points z in a neighbourhood of a fixpoint α seems to be connected in a very fundamental way with the global properties of F. In order to formulate definitions of this behaviour (especially in terms of attractive and repulsive fixpoints) which will be appropriate to the development of Chapter 5, it becomes convenient at this point to identify the Banach space with the Riemann sphere σ (or its projection onto the extended complex plane π). Definition 4.14 is therefore stated in terms of σ, and subsequent discussion will be confined to σ or π.

Definition 4.14 A fixpoint α of order r of F is said to be attractive

<u>(repulsive)</u> iff there exists a neighbourhood $M(\alpha)$ in which $\{f_{nr}\}(\{f_{-nr}\})$ converges spherically uniformly to α. α is said to be <u>indifferent</u> if no such $M(\alpha)$ exists. Points from which convergence to α takes place are called <u>attractive</u> <u>(repulsive)</u> points.

From Definition 4.14 it may be seen that convergence of $F(z)$ to an attractive fixpoint α is assured, provided that the initial value z is chosen sufficiently close to α. Iterating functions f are therefore normally constructed so as to have the property that every solution of a given equation $\phi(z) = 0$ is not just a fixpoint, but in fact an attractive fixpoint (usually of order 1), of the iteration sequence F. Note that in the repulsive case, $\{f_{-nr}\}$ is formed by iterating a branch f_{-r} for which $f_{-r}(\alpha) = \alpha$.

Definition 4.14 will not as far as is known be found in the literature. It has been introduced because it is a somewhat weaker definition than those which are generally employed, because it maintains a certain symmetry between attractive and repulsive fixpoints, and because it seems that the standard results for iterations of rational functions can still be developed without recourse to any stronger definition.

For example, Kuczma [K5, 18] defines attractive/repulsive fixpoints $\alpha \in \pi - \{\infty\}$ of order r in terms of ordinary convergence/divergence, but then also defines <u>strongly</u> attractive/repulsive fixpoints on the basis of the conditions that for every $z \neq \alpha$ in some neighbourhood $M(\alpha)$

$$|f_r(z) - \alpha| < \theta|z - \alpha| \quad \text{(attractive)}$$

$$|f_r(z) - \alpha| > |z - \alpha|/\theta \quad \text{(repulsive)}$$

for some fixed $0 < \theta < 1$. In this terminology, it would be important to specify conditions on f under which every attractive/repulsive fixpoint is strongly attractive/repulsive. For functions f holomorphic in $M(\alpha)$, Kuczma's strongly attractive/repulsive fixpoints are equivalent, as shown below, to the attractive/repulsive fixpoints of Cremer [C1] which are defined to depend on the sign of $|f_r'(\alpha)| - 1$. Targonski [T4, 32-33] gives much the same definition of attractive/repulsive as Kuczma, but does not mention strongly attractive/repulsive: he contents himself with showing that if $|f_r'(\alpha)| < 1$, then α is attractive. Targonski however quotes a remark of Kuczma that there is not "even a satisfactory definition of a repulsive fixpoint". Indeed, it is perhaps not too much to say that the definitions neither of repulsive nor of attractive fixpoints are satisfactory. The present attempt has been suggested in part by [C1], in which the proofs in one or two cases require that branches of the inverse iteration should converge to repulsive fixpoints of F; in part by [L2], in which the existence of attractive/repulsive regions in neighbourhoods of certain kinds of indifferent fixpoints is discussed.

To establish the relationship among Definition 4.14 and the definitions of Cremer and Kuczma, consider an iterating function f which is holomorphic in $M(\alpha)$. By Theorem 3.14 f_r is also holomorphic in some neighbourhood $M'(\alpha)$ and may therefore be represented by a Taylor's series there:

$$f_r(z) = f_r(\alpha) + f_r'(\alpha)\,(z - \alpha) + \cdots,$$

or $(z \neq \alpha)$

$$[f_r(z) - f_r(\alpha)]/(z - \alpha) = f_r'(\alpha) + (z - \alpha)\ [\cdots].$$

By bringing z close to α, the lefthand side of this equation may be brought as close as may be desired to $f_r'(\alpha)$. From this fact it follows that for a holomorphic function f, the condition $|f_r'(\alpha)| < 1$ (> 1) is equivalent to Kuczma's condition that α be a strongly attractive (repulsive) fixpoint. Further, if Kuczma's condition is satisfied, then f_r is a contraction (expansion) mapping in M'(α), from which the uniform convergence of $\{f_{nr}\}$ ($\{f_{-nr}\}$) follows directly. Then either Kuczma's or Cremer's definition implies Definition 4.14.

Definition 4.15 An indifferent fixpoint α is called a centre iff there exists a neighbourhood of α which contains neither attractive nor repulsive points of F.

This definition isolates a special kind of indifferent fixpoint which appears to arise only in special circumstances. For example, if in the complex plane π

$$f(z) = \omega_p z^{\omega} q,$$

where p, q > 1 are integers such that pq = r, and ω_i stands for the i^{th} root of unity, then $f_r(z) = z$ everywhere in π -- that is, every point of π is a centre of order less than or equal to r. Centres are discussed in [R1, F2]: they assume importance in the consideration of circumstances in which F does not give rise to constant limit functions [S2].

<u>Definition 4.16</u> A sequence F of iterates in the complex plane is said to have <u>order of convergence</u> σ at an attractive fixpoint α of order r if σ is the least non-negative integer such that, in every neighbourhood $M(\alpha)$, there exists $z \neq \alpha$ such that

$$|f_r(z) - \alpha| > |z - \alpha|^{\sigma+1}.$$

If f is assumed to be differentiable in a neighbourhood of α, and if α is assumed to be a fixpoint in the sense of Cremer, then the condition of Definition 4.16 may be replaced by the condition that

$$f_r^{(k)}(\alpha) = 0,\ k = 1, \cdots, \sigma; \quad |f_r^{(\sigma+1)}(\alpha)| > 1.$$

Then, since

$$f_r'(z) = f'[f_{r-1}(z)] \cdot f'[f_{r-2}(z)] \cdot \ldots \cdot f'(z),$$

it follows that in the special case $z = \alpha$,

$$f_r'(\alpha) = f_r'[f(\alpha)] = \cdots = f_r'[f_{r-1}(\alpha)],$$

from which it is seen that the fixpoints of a given cycle are all either attractive, or repulsive, or indifferent. Hence one may speak of attractive, repulsive, and indifferent cycles. Moreover, by the symmetry of $f_r'(\alpha)$, hence of $f_r^{(\sigma)}(\alpha)$,

$$f_r^{(\sigma)}(\alpha) = f_r^{(\sigma)}[f(\alpha)] = \cdots = f_r^{(\sigma)}[f_{r-1}(\alpha)],$$

so that one may speak also of the order of convergence of a cycle.

Example 4.17 From Example 4.3, for the Newton-Raphson iteration,

$$f'(z) = \phi(z)\phi''(z)/[\phi'(z)]^2,$$

an expression which is zero at a simple root $z = \alpha$ of the given equation $\phi(z) = 0$. Hence such a root α is an attractive fixpoint with order of convergence $\sigma \geq 2$.

As indicated above, the first problem in the construction of iterating functions is to ensure that f(z) has attractive fixpoints at points $z = \alpha$ which are solutions of the given problem. But there are other important factors as well. Among these, there has traditionally been heavy emphasis on the order of convergence σ, which should of course be as large as possible, and on the amount of work to be done at each stage of the iteration to evaluate the function f, an amount which should be small. Usually the work increases with σ. As remarked earlier, one important factor that has traditionally been neglected is the tendency of iterations to converge from certain initial values to attractive cycles which are not solutions of the given problem.

Example 4.18 Consider the polynomial $\phi(z) = z^4 - 6z^2 - 11$, which has zeros $z = \pm(3 \pm 2\sqrt{5})^{1/2}$. The Newton-Raphson iteration corresponding to $\phi(z) = 0$ is

$$f(z) = (3z^4 - 6z^2 + 11)/(4z^3 - 12z),$$

which does in fact have attractive fixpoints of order 1 at the four zeros of ϕ. But observe that $f(\pm 1) = \mp 1$, so that $\{+1, -1\}$ is a cycle of order 2. Moreover, this cycle is attractive, because $\phi''(z) = 12z^2 - 12$, which vanishes for $z = \pm 1$, thereby causing the expression for $f'(z)$ (see 4.17) to vanish also. In other words, for starting values sufficiently close to ± 1, the iteration converges, not to zeros of $\phi(z)$, but to ± 1.

This example is due to Turing and Brooker [B2]. A recent paper by Ascher [A1] gives examples of classes of problems giving rise to unwanted repulsive cycles in the Newton-Raphson iteration, a problem which in practice is less serious than that of attractive cycles. As in [S1], Chapter 6 will describe classes of iterating functions which have the property that, while some functions in the class (particularly the Newton-Raphson iteration) give rise to attractive cycles, others in the same class may not. In fact, for the Newton-Raphson iteration f applied to a polynomial ϕ of degree k, the expression given in 4.17 shows that f' may have $2k - 2$ zeros (those of ϕ together with those of ϕ''). Of these, the $k - 2$ zeros of ϕ'' may (or may not) give rise to attractive cycles of aggregate order $r < k - 2$ and order of convergence $\sigma > 2$. This example anticipates the result, proved in the next chapter, that the number of attractive cycles is bounded by the number of zeros of f'. Note that if ϕ is a quadratic function with distinct roots, f gives rise to no extra attractive cycles.

To conclude this chapter, brief reference is made to some of the work which has been done to generate classes of iterating functions with desirable properties. Such efforts date back at least to Schröder, who in 1870 published results due to himself and Eggers [S4] which yielded iterating functions of arbitrarily high order of convergence σ for solution of a given equation $\phi(z) = 0$ with simple roots:

$$f^{[\sigma]}(z) = z - b_1(z)\phi(z) + \cdots + (-1)^{\sigma-1}b_{\sigma-1}(z)[\phi(z)]^{\sigma-1},$$

where $b_1(z) = 1/\phi'(z)$, $b_j(z) = b'_{j-1}(z)/j\phi'(z)$, $j = 2, \cdots, \sigma - 1$. Similar iterating functions were constructed, also in the nineteenth century, by König [H1, 103-105]:

$$f^{[\sigma]}(z) = z + (\sigma - 1)[1/\phi(z)]^{(\sigma-2)}/[1/\phi(z)]^{(\sigma-1)}.$$

Both the Schröder and König iterations reduce to the Newton-Raphson iteration in the case $\sigma = 2$. The iterations of Schröder have recently [H3] been related to modern iterations due to Traub [T2] and Jenkins [J2], and these latter iterations have themselves been generalized by Ford [F5]. Schröder's iterations have also recently been generalized to apply to multiple roots by

Pomentale [P1]. The Newton-Raphson iteration and the iteration of Halley (Example 4.3) also emerge as special cases of a finite class of iterations due to Farmer and Loizou [F6], later developed independently, using other methods, by Nourein [N1] as an infinite class of iterations of arbitrarily high order of convergence. Hansen and Patrick [H10] exhibit a class of iterating functions dependent on a parameter s,

$$f^{[s]}(z) = z - (s + 1)\phi/\{s\phi' \pm [(\phi')^2 - (s + 1)\phi\phi'']^{1/2}\},$$

which reduces to Halley's iteration ($s \to -1$), Laguerre's iteration ($s = \frac{1}{k-1}$), Euler's iteration ($s = 1$), and Ostrowski's square root iteration [O2] ($s = 0$). Bairstow's method (Example 4.5) has been displayed by Brodlie [B10] as a member of a class of iterations. In the present work, Chapter 6 shows how to construct all rational iterating functions which converge with given order of convergence to the (simple) roots of a given polynomial: here again the Newton-Raphson, Schröder, and König iterations emerge as special cases.

5 Iteration of Rational Functions

In this chapter the results of Chapters 3 and 4 are used to develop the global properties of iterations of rational functions on the Riemann sphere σ, which now becomes the Banach space of Chapter 4. As is customary, the rational functions will be assumed to be irreducible (numerator and denominator have no common factor), so that there is no possibility of the indeterminate form 0/0 arising.

Let N denote the set of points at which the sequence F of iterates of a rational function f is both infinite and normal. Let E denote the set of points at which F is an infinite sequence but not normal. Then

Theorem 5.1 N + E = σ.

Proof An immediate consequence of the remark made after Definition 3.7 that a function is meromorphic in the extended plane if and only if it is rational [S3, 147]. ∎

This result shows that, for rational f, F is always an infinite sequence.

Theorem 5.2 N is open (E is closed).

Proof Consider any point z ε N. By the definition of normality (3.12), F is normal in some neighbourhood of z. Hence z is an interior point. ∎

Theorem 5.3 Let S denote (a) N (b) E. In each case Axioms 4.8, hence

Lemmas 4.9-4.12, hold.

Proof Since f is rational, it is single-valued and continuous everywhere in σ. Hence Axioms 4.8(a) and 4.8(b) are satisfied for both S = N and S = E.

Before considering 4.8(c) - (d), note that for every point $z \in \sigma$, f(z) exists since f is rational, and $f_{-1}(z)$ exists by the fundamental theorem of algebra (every polynomial has a zero).

Now consider any point $z \in N$. Then $F(z) = \{z, f(z), \cdots\}$ is normal. But then so are $F(f(z)) = \{f(z), f_2(z), \cdots\}$ and $F(f_{-1}(z)) = \{(f_{-1}(z), z, f(z), \cdots\}$, since by Definition 3.12 the addition or deletion of only a finite number of terms of F(z) cannot affect the normality property. Hence Axioms 4.8(c) and 4.8(d) are satisfied for S = N. Then no point of E may have an antecedent or consequent in N. Hence the axioms hold also for S = E.

∎

Theorem 5.4 The sequence of iterates of each neighbourhood of a point of E covers all points in σ except at most two exceptional points.

Proof This is a restatement of Montel's Theorem (3.17). ∎

Theorem 5.5 If E contains an interior point, then $N = \phi$.

Proof Suppose there exists a point $z \in E$ such that a neighbourhood

$M(z) \subset E$. By Theorem 5.4 $F(M(z))$ covers all points except at most two exceptional points, and all of these covered points must then by Axiom 4.8(c) be points of E. The exceptional points, if they exist, cannot by Theorem 5.2 be points of N. Hence $N = \phi$. ∎

The case $N = \phi$ may occur, as is shown by the example of Lattes [L1],

$$f(z) = (z^2 + 1)^2 / 4z(z^2 - 1),$$

which is the Newton-Raphson iteration applied to a branch of the function $g(z) = (3z^4 - 6z^2 - 1)^{1/3}$. The sequence of iterates of f(z) is known to be normal nowhere in σ [C1]. The following theorem describes what is, in practice, the more common case.

<u>Theorem 5.6</u> If $N \neq \phi$, E is nowhere dense and $E = \bar{N} - N$.

<u>Proof</u> By hypothesis, N exists. Hence by Theorem 5.2 N consists only of interior points, and by Theorem 5.5 E contains no interior points. Since by Theorem 5.1 $E + N = \sigma$, E is then precisely the boundary set of N. ∎

Another way of stating Theorem 5.6 is that if N exists, then $\bar{N} = \sigma$. Further information about E is provided by the next four theorems.

<u>Theorem 5.7</u> E contains no exceptional points.

<u>Proof</u> Suppose $z \in E$. Then by Axiom 4.8(d) every antecedent $f_{-1}(z) \in E$.

Then certainly $F(f_{-1}(z))$ covers z, which cannot by Theorem 5.4 be exceptional. ∎

<u>Theorem 5.8</u> For $z \in E$, every accumulation point of $F(z)$ is a point of E.

<u>Proof</u> Suppose there exists an accumulation point $\alpha \in N$ of $F(z)$. Then by Axiom 4.8(c) every neighbourhood of α contains points of E. Hence F is normal in no neighbourhood of α, contrary to Definition 3.12. ∎

<u>Theorem 5.9</u> E is a perfect set.

<u>Proof</u> Recall that E is perfect if it is closed (every limit point is a point of E) and dense-in-itself (every point is a limit point of E). By Theorem 5.2 only the second property remains to be proved. Further, it may be assumed that $E \neq \phi$, for otherwise E would be perfect.

Suppose therefore that E is not dense-in-itself. Then E contains an isolated point α. By Theorem 5.4 the sequence of iterates of each neighbourhood of α covers every point except at most two, which by Theorem 5.7 are not points of E. Then every point of E, including α itself, is a consequent of α. But this implies that E must break down into a finite cycle of order r:

$$E = \{\alpha, f(\alpha), \cdots, f_{r-1}(\alpha)\}.$$

Consider then the sequence of iterates $\{f_r\}$. This is normal wherever

F is normal. Suppose that $\{f_r\}$ is normal also at α. Then by Axiom 4.8(a) it is normal at every point of E, and so $\{f_r\}$ is normal in every domain D containing the points of E. By Theorem 3.14 the application of $\{f_r\}$ to a domain D containing the points of E will cover neighbourhoods $M(\alpha)$, $M(f(\alpha))$, $\cdots$, $M(f_{r-1}(\alpha))$, and it follows that every point covered by F in some neighbourhood $M'(\alpha)$ will be covered by $\{f_r\}$ in D. Hence $\{f_r\}$ omits at most two points in D, contrary to Theorem 3.17. Then $\{f_r\}$ is not normal at any point of E. Hence by Theorems 5.4 and 5.7 $\{f_r\}$ applied to every neighbourhood $M(\alpha)$ cannot omit any points of E. But it does omit $f(\alpha)$, $\cdots$, $f_{r-1}(\alpha)$. This is possible only if $r = 1$ and $E = \{\alpha\}$.

Consider then F(D), where $D = N = \sigma - \{\alpha\}$. Since $F(M(\alpha) - \{\alpha\})$ omits only α and at most two exceptional points β and γ, F(D) also omits at most α, β, γ. But $f(\beta) = \beta$ or γ and $f(\gamma) = \beta$ or γ, since otherwise these points could not be exceptional. Hence F(D) omits only α, in contradiction to Theorem 3.17. Then E contains no isolated point α and is therefore dense-in-itself. This proves the theorem. ∎

Theorem 5.10 E contains all the repulsive fixpoints and all the indifferent fixpoints which are not centres.

Proof Suppose α is a fixpoint of order r which is neither attractive nor a centre. Then by Definitions 4.14 and 4.15 there exists a repulsive point in every neighbourhood of α. Choose such a point w, and observe that by Definition 4.14 every neighbourhood of α contains an antecedent $f_{-nr}(w)$ of w.

Suppose $\alpha \in N$. Then by Theorem 3.16 $\{f_{nr}\}$ is equicontinuous at α: given $\varepsilon > 0$, it is possible to ensure that

$$\chi[f_{nr}(z), f_{nr}(\alpha)] = \chi[f_{nr}(z), \alpha] < \varepsilon$$

for all n simply by choosing $\chi(z, \alpha) < \delta$. But if $\varepsilon < \chi(w, \alpha)$, this condition cannot be satisfied for any choice of δ, no matter how small. Hence $\{f_{nr}\}$ is not equicontinuous at α and therefore, by Theorem 3.16, $\alpha \notin N$. ∎

The following four theorems provide further information about N.

Theorem 5.11 N contains all the attractive fixpoints and all the centres.

Proof Suppose α is an attractive fixpoint of order r. Then there exists a neighbourhood $P(\alpha)$ such that for any $\varepsilon > 0$, $\chi[f_{nr}(z), \alpha] < \varepsilon$ for all $n > N(\varepsilon)$ and every $z \in P(\alpha)$. Then consider $M(\alpha) = f_{(N+1)r}(P(\alpha))$. By Theorem 3.14 $M(\alpha)$ is a domain containing α and $z \in M(\alpha)$ implies $\chi[f_r(z), \alpha)] < \varepsilon$. By choosing ε small enough it may be ensured, using the spherical continuity of f, that $f(M(\alpha)), \cdots, f_{r-1}(M(\alpha))$ are disjoint and in fact that every consequent of a point of $M(\alpha)$ lies arbitrarily near one of the r fixpoints $\alpha, f(\alpha), \cdots, f_{r-1}(\alpha)$. Then F omits more than three points in $M(\alpha)$, and by Theorem 3.17 F is normal in $M(\alpha)$. Hence $\alpha \in N$.

Suppose α is a centre of order r. Then by Definition 4.15 there exists a neighbourhood of $M(\alpha)$ which contains no repulsive points. Then there exist no points in $M(\alpha) - \{\alpha\}$ whose antecedents accumulate at α. Then by Theorem 5.4 $\alpha \notin E$. ∎

Theorem 5.12 If α is an attractive fixpoint of order r in a domain $D \subseteq N$, then $\{f_{nr}\}$ converges to α everywhere in D.

Proof By Definition 4.14 there exists $M(\alpha) \subset D$ in which $\{f_{nr}\}$ converges uniformly to a limit function $g(z) = \lim f_{nr}(z) = \alpha$. By Definition 3.12 $\{f_{nr}\}$ contains a subsequence $\{f_{n_k(r)}\}$ which converges uniformly on every compact subset of D. By Theorem 3.15 the limit function h of $\{f_{n_k(r)}\}$ is a meromorphic function. But for $z \in M(\alpha)$, it must be true that $h(z) = g(z) = \alpha$, so that by Theorem 3.13 $h(z) = \alpha$ everywhere in D. Then by Lemma 4.11 $\alpha = \lim f_{n_k(r)+r}$ everywhere in D.

Suppose then that there exists $\beta \neq \alpha$ and a subsequence $\{f_{m_j(r)}\} \subset \{f_{nr}\}$ such that $\beta = \lim f_{m_j(r)}(w)$ for some point $w \in D$. Then given an arbitrary $\varepsilon > 0$ an integer $J(\varepsilon)$ may be found such that for every $j > J(\varepsilon)$, $\chi[f_{m_j(r)}(w), \beta] < \varepsilon$. Choose $\varepsilon < \chi(\alpha, \beta)$ and recall that by equicontinuity at α (Theorem 3.16) $\delta = \delta(\varepsilon)$ may be found such that for every $z \in M'(\alpha, \delta)$ and every integer $n > 0$,

$$\chi[f_{nr}(z), f_{nr}(\alpha)] = \chi[f_{nr}(z), \alpha] < \varepsilon.$$

By the convergence of $f_{n_k(r)}(w)$ to α, an integer $K = K(\delta)$ may be found such that $f_{n_K(r)}(w) \in M'(\alpha, \delta)$. Then

$$\chi[f_{nr}[f_{n_K(r)}(w)], \alpha] = \chi[f_{n_K(r)+nr}(w), \alpha] < \varepsilon$$

for every integer $n > 0$. Since $\{f_{n_K(r)+nr}\}$ includes all but a finite number of terms of $\{f_{m_j(r)}\}$, this statement conflicts with the convergence of $\{f_{m_j(r)}\}$ to β. This proves the theorem. ∎

As a consequence of Theorem 5.12, whenever an attractive fixpoint $\alpha \in D \subseteq N$, D is called the <u>immediate domain of convergence</u> of α.

<u>Theorem 5.13</u> If D is the immediate domain of convergence of a fixpoint of order r, then f_r: $D \to D$.

<u>Proof</u> This is an immediate consequence of the Open Map Theorem (3.14). ∎

<u>Theorem 5.14</u> The union of the immediate domains of convergence of the fixpoints of an attractive cycle contains at least one point w such that $f'(w) = 0$.

<u>Proof</u> Let $\alpha \in D$ be an attractive fixpoint of order r. Denote by f_{-r} a branch of the inverse function such that $f_{-r}(\alpha) = \alpha$. Suppose that $z \in D$ implies $f_r'(z) \neq 0$. Then by Theorem 3.18 f_{-r} is meromorphic in $f_r(D)$, which by Theorem 3.14 is a domain and by Theorem 5.13 contained in D. In fact the sequence $\{f_{-nr}\}$ is even normal in $f_r(D)$,

since it omits more than three points there (for example, all points of E). Hence by Theorems 5.10 and 5.11 α is either a centre or an attractive fixpoint of $\{f_{-nr}\}$. But since α is an attractive fixpoint of F, there exists a neighbourhood $M(\alpha)$ containing only attractive points of $\{f_{nr}\}$, so that by Definitions 4.14 and 4.15 α can be neither a centre nor an attractive fixpoint of $\{f_{-nr}\}$. Hence there exists $w \in D$ for which $f_r'(w) = 0$. But as noted in connection with Definition 4.16, this means that

$$f_r'(w) = f'[f_{r-1}(w)] \cdot f'[f_{r-2}(w)] \cdot \ldots \cdot f'(w) = 0,$$

from which it follows that $f_r'(w) = 0$ if and only if at least one of $f'(w)$, $f'(f(w))$, $\cdots$, $f'(f_{r-1}(w))$ is zero. Since by Theorem 3.14 every point $f_s(w)$, $s = 0, \cdots, r-1$, is contained in the immediate domain of convergence of $f_s(\alpha)$, the result is proved. ∎

Since for a rational function f, f' is also rational and therefore possesses a finite number of zeros, it follows that the number of attractive cycles of a rational iterating function is finite. More precisely, if $f = g_p/h_q$, where g, h are polynomials of degree p, q respectively, then

$$f' = (g_p' h_q - g_p h_q')/h_q^2,$$

so that F gives rise to at most $p + q - 1$ attractive cycles.

Theorems 5.1 - 5.14 constitute the basis of the global theory of rational iterating functions) and their convergence properties. To complete the theory, it would of course be desirable to show that in cases of practical interest ($E \neq \sigma$, $E \neq \phi$) the sequence of iterates converges everywhere in N to a constant limit function. Although this result has been proved for holomorphic functions (see the elegant summary by Burckel [B12]), it has not been proved in general, despite considerable work in the area by Baker [B3-B8] and a recent attempt, unfortunately incorrect, by Smyth [S2]. In the context of our present discussion, proof of a constant limit theorem amounts to showing that Definitions 4.14 and 4.15 account for all possible cases of convergence; in particular, not only that non-constant limit functions may not exist, but also that other forms of constant limit function are excluded (for example, centres), in cases of practical interest. There is reason to conjecture, if not believe, that both centres and non-constant limit functions are associated only with the case $E = \phi$ [S2, B12]. For example, the iterating function $f(z) = \omega_p z^{\omega_q}$ discussed in connection with Definition 4.15 gives rise to centres $f_r(z) = z$ and a non-constant limit function $g(z) = z$ at every point $z \in \sigma = N$. The equation $f_r(z) = z$ was studied by Babbage [B11], who recognized that if $f(z)$ was a solution of it, then so was $f^*(z) = h_{-1}\{f[h(z)]\}$ for any homeomorphism h. No non-trivial example of a rational or even meromorphic function which gives rise to a non-constant limit function has been found. An example due to Schröder [S5] has been shown to be incorrect [S2]. A recent article by Rice et al. [R2] also bears on this question.

The development given in Chapters 3-5 has been strongly influenced by [C1] and, especially, by [B12]. A number of new proofs have however been used, with the intention of avoiding any dependence on the special properties of rational functions. A paper by Rådström [R1] separates meromorphic functions f into four classes based on the properties of the set U of essential singularities of f and their antecedents; this classification can be used to prove the results of this chapter, slightly modified, for general meromorphic functions as well. An excellent summary by Brolin [B13] has been brought to the author's attention too late for incorporation in this work.

6 Rational Iterating Functions Without Attractive Cycles

The preceding three chapters have set forth the grounds for supposing that, in cases of practical interest, iterations of rational functions converge almost everywhere in the complex plane to attractive cycles. Examples 4.17 and 4.18 have shown however that these attractive cycles may sometimes be undesirable: they may be points which are not solutions of the problem which the iteration has been designed to solve.

It is therefore an ultimate aim of this research to find a means of constructing rational iterating functions which do not give rise to attractive cycles or superfluous attractive fixpoints, and which, therefore, converge globally — that is, almost everywhere in the complex plane — to solutions of the given problem $\phi(z) = 0$. In this chapter a step is taken toward this goal by showing how to construct all the rational iterating functions $f = f_{p,q}$ which converge with given order of convergence σ to n given distinct points z_i, $i = 1, \cdots, n$; in particular, how to construct all the $f_{p,q}$ which converge with order σ to the distinct (but unknown) zeros z_i of a given polynomial $\phi_n(z)$. Denote by

$$C_M = C_M(\sigma;\ z_1, \cdots, z_n)$$

the class of rational functions $f_{p,q}$ such that (1) $p + q + 1 = M$ (2) the sequence of iterates of $f_{p,q}$ converges with order σ at each z_i; then the construction technique will open up the new possibility of inspecting the iterating functions of each C_M to determine which give rise to attractive cycles, and which do not — or, indeed, to determine which may or may not have other desirable or undesirable properties. No techniques are proposed here for inspection of the elements of C_M; but it is shown, by an example derived from Brooker's, that, even though the Newton-Raphson iteration $f_{n,n-1} \in C_{2n}$ gives rise to attractive cycles, there may nevertheless exist other functions in the same class which have no such cycles.

6.1 Rational Iterating Functions with n Given Fixpoints

Consider

$$f_{p,q}(z) = g_p(z)/h_q(z) = \sum_{0 \leq j \leq p} \alpha_j z^j / \sum_{0 \leq k \leq q} \beta_k z^k,$$

where $p \geq 0$, $q \geq 0$, $\alpha_p \neq 0$, $\beta_q \neq 0$, and the polynomials g and h have no common factor. In this section the coefficients (α_j, β_k) will be chosen to satisfy n given conditions

(1) $$f_{p,q}(z_i) = z_i, \quad i = 1, \cdots, n,$$

where $i \neq i' \implies z_i \neq z_{i'}$. That is, (α_j, β_k) are to be found such that

(2) $$g(z_i) - z_i h(z_i) = 0;$$

in other words, such that

$$(3) \qquad \alpha_0 + (\alpha_1 - \beta_0)z_i + \cdots + (\alpha_p - \beta_{p-1})z_i^p + (0 - \beta_p)z_i^{p+1}$$

$$+ \cdots + (0 - \beta_q)z_i^{q+1} = 0$$

for each z_i, where if $q < p$, then $\beta_k \equiv 0$, $k = q + 1, \cdots, p$.

Setting

$$(4) \qquad \gamma_j^{(1)} \equiv \alpha_j - \beta_{j-1}, \quad j = 0, \cdots, m,$$

where $m \equiv \max(p, q + 1)$, and defining $\beta_{-1} \equiv 0$, $\alpha_j \equiv 0$ for $j = p + 1, \cdots, m$, (3) may be rewritten

$$\begin{pmatrix} 1 & z_1 & \cdots & z_1^m \\ \vdots & \vdots & & \vdots \\ 1 & z_n & \cdots & z_n^m \end{pmatrix} \cdot \begin{pmatrix} \gamma_0^{(1)} \\ \vdots \\ \gamma_m^{(1)} \end{pmatrix} = 0.$$

Introducing the notation

$$Z_n^{r,s} \equiv \begin{pmatrix} z_1^r & z_1^{r+1} \cdots & z_1^s \\ \vdots & \vdots & \vdots \\ z_n^r & z_n^{r+1} \cdots & z_n^s \end{pmatrix}, \quad \Gamma_{r,s} \equiv \begin{pmatrix} \gamma_r^{(1)} \\ \vdots \\ \gamma_s^{(1)} \end{pmatrix},$$

for integers r and s, $0 \le r \le s$, this becomes

(5) $$\sum Z_n^{0,m} \Gamma_{0,m} = 0,$$

a set of n homogeneous linear equations in m + 1 variables $\gamma_j^{(1)}$. Observe that for m + 1 = n, $\det(Z_{m+1}^{0,m})$ is simply the Vandermonde determinant

$$D_n = \prod_{1 \le i' < i \le n} (z_i - z_{i'}) \ne 0.$$

Accordingly, for m < n, the only solution of (5) is $\Gamma_{0,m} = 0$, which implies $\alpha_j = \beta_{j-1}$, j = 0, ⋯, m, hence that $f_{p,q}(z) = z$, identically.

However, if m = n, the rank of $Z_n^{0,m}$ is also n, and, therefore, (5) possesses a nontrivial solution in the $\gamma_j^{(1)}$ which is unique but for multiplication by a constant factor. Thus the minimum value of m leading to a nontrivial iterating function $f_{p,q}$ is m = n. Rewriting (5) for m = n produces

$$Z_n^{0,n-1} \Gamma_{0,n-1} = -Z_n^{n,n} \gamma_n^{(1)},$$

which yields the (unique) solution

$$\gamma_j^{(1)} = -(D_n^{j+1,n}/D_n)\gamma_n^{(1)}, \quad j = 0, \cdots, n-1,$$

where in general $D_n^{j,k}$ is the determinant obtained by replacing the jth column of D_n with $Z_n^{k,k}$. In Chapter 7 Theorem 7.2 shows that the quantities

$$S_n^{j,n} = (-1)^{n-j} D_n^{j,n}/D_n, \quad 1 \leq j \leq n;$$

are the familiar Newton sums

$$S_n^{1,n} = z_1 z_2 \cdots z_n, \cdots, S_n^{n-1,n} = \sum_{i \neq i'} z_i z_{i'}, \quad S_n^{n,n} = \sum_{1 \leq i \leq n} z_i.$$

Then for the case $m = n$,

$$\gamma_j^{(1)} = (-1)^{n-j} S_n^{j+1,n} \gamma_n^{(1)}, \quad j = 0, \cdots, n-1, \tag{6}$$

and the $\gamma_j^{(1)}$ are always symmetric functions of the z_i.

More generally, for $m > n$ conditions (1),

$$Z_n^{0,n-1} \Gamma_{0,n-1} = -Z_n^{n,m} \Gamma_{n,m}, \tag{7}$$

and it follows that each element of the solution $\Gamma_{0,n-1}$ may be expressed as a linear combination of $m - n + 1$ arbitrarily chosen values $\gamma_n^{(1)}, \cdots, \gamma_m^{(1)}$. If corresponding to each $K = n, \cdots, m$, these $m - n + 1$ values are chosen so that $\gamma_K^{(1)} \neq 0$, $\gamma_k^{(1)} = 0$ for $k \neq K$, $m - n + 1$ linearly independent particular solutions are derived,

$$\gamma_{jK}^{(1)} = -(D_n^{j+1,K}/D_n)\gamma_K^{(1)}, \quad j = 0, \cdots, n-1,$$

each dependent on a single arbitrarily chosen value $\gamma_K^{(1)}$. A general solution of (7) may then be reconstituted by forming a linear combination of these m − n + 1 particular solutions. Accordingly, summing the $\gamma_{jK}^{(1)}$ over K and making use of the generalized Newton sums $S_n^{j,k}$ defined in Chapter 7, the general solution of the system (5) may be written

$$(8) \qquad \gamma_j^{(1)} = (-1)^{n-j} \sum_{n \le K \le m} S_n^{j+1,K} \gamma_K^{(1)}, \quad j = 0, \cdots, n-1,$$

where the $\gamma_K^{(1)}$ are arbitrary complex numbers, not all zero. The quantities $S_n^{j+1,K}$ are by definition symmetric functions of the z_i, and may be calculated as indicated by Theorem 7.4 and Algorithm 7.5. Hence

<u>Theorem 6.1</u> In order that a rational function $f_{p,q}(z)$ satisfy the n conditions (1), it is necessary that $m \equiv \max(p, q+1) \ge n$. $f_{p,q}(z)$ satisfying (1) may be constructed by arbitrarily choosing m − n + 1 "combined coefficients" $\gamma_j^{(1)}$ as defined in (4), and then using (8) or an analogous system to find the remaining n combined coefficients, which are then uniquely determined as symmetric functions of the given points z_i, $i = 1, \cdots, n$.

A way has thus been found to generate all rational functions which as iterating functions have fixpoints at n given distinct points.

It should be remarked at this point that, despite Theorem 6.1, the construction technique may yield rational functions which are not irreducible, in contradiction to the assumption stated at the beginning of Chapter 5. The construction process described here is a special case of rational interpolation, which is succinctly described in [G1]: "the basic problem of rational interpolation is that of finding a rational fraction which assumes given values f_0, f_1, ···, f_n at distinct points x_0, x_1, ···, x_n." It is possible however that for specific choices of p, q and f_i, x_i, $i = 0, \cdots, n$, there may be no irreducible rational function which meets the conditions. Hildebrand [H12, 499] gives the example

$$n = 4 \qquad x = (0,\ 1,\ 2,\ 3)$$

$$p = 2 \qquad f = (2,\ 3/2,\ 4/5,\ 1/2)$$

$$q = 1$$

which yields the reducible solution

$$f_{p,q}(x) = (4 - x)(2 - x) \ /\ 2(2 - x).$$

6.2 Rational Iterating Functions with n Given Attractive Fixpoints

The conditions now to be satisfied may be expressed as follows:

(9) $$f_{p,q}(z_i) = z_i, \quad f'_{p,q}(z_i) = 0,$$

for every $i = 1, \cdots, n$. These are 2n independent simultaneous linear constraints on the p + q + 2 coefficients (α_j, β_k), which may be re-expressed in the form of (2):

$$g(z_i) - z_i h(z_i) = 0, \tag{10a}$$

$$g'(z_i) - z_i h'(z_i) = 0. \tag{10b}$$

From (10b), after some manipulation, a set of equations analogous to (3) is obtained:

$$\sum_{0 \le j \le m-1} [(j + 1)\alpha_{j+1} - j\,\beta_j] z_i^j = 0, \quad i = 1, \cdots, n. \tag{11}$$

It is convenient to define

$$\gamma_j^{(2)} = (j + 1)\alpha_{j+1} - j\,\beta_j, \quad j = 0, \cdots, m - 1, \tag{12}$$

analogous to (4). Then (11) simplifies to

$$\sum_{0 \le j \le m-1} z_i^j \gamma_j^{(2)} = 0, \quad i = 1, \cdots, n, \tag{13}$$

a set of n homogeneous linear equations in m unknowns $\gamma_j^{(2)}$ which is identical in form to (5). Proceeding as in the previous section, observe that for $m < n + 1$ there are only trivial solutions of (13), and that for $m = n + 1$ there is a unique solution

$$\gamma_j^{(2)} = (-1)^{n-j} S_n^{j+1,n} \gamma_n^{(2)}, \quad j = 0, \cdots, n - 1, \tag{14}$$

analogous to (6). Then for $m > n + 1$, it may be shown as before that general solutions $\gamma_j^{(2)}$, $j = 0, \cdots, n - 1$, of (10b) may be written

$$(15) \qquad \gamma_j^{(2)} = (-1)^{n-j} \sum_{n<K<m-1} S_n^{j+1,K} \gamma_K^{(2)}, \quad j = 0, \cdots, n-1,$$

identical in form to (8). The solutions (8) and (15) taken together are 2n linear homogeneous equations in $p + q + 2$ variables (α_j, β_k) equivalent to the original 2n equations (9); they will possess a solution if $p + q + 1 > 2n$, a condition in some respects stricter than $m > n + 1$, but which allows the equations (8) and (15) to be solved in the important "Newton-Raphson" case, characterized by $n = m = p = q + 1$. Then

<u>Theorem 6.2</u> In order that a rational function $f_{p,q}(z)$ satisfy the 2n conditions (9), it is necessary that $p + q + 1 > 2n$. $f_{p,q}(z)$ satisfying (9) may be constructed by arbitrarily choosing $(p + q + 2) - 2n$ of the coefficients (α_j, β_k), and then using (4), (8), (12), and (15) to find the remaining 2n coefficients, which are then uniquely determined as symmetric functions of the given points z_i, $i = 1, \cdots, n$.

A way has thus been found to generate all rational functions which as iterating functions have attractive fixpoints with order of convergence 2 at n given distinct points.

Once again it should be noted that in the case $p + q + 1 = 2n$, it may occur that there exists no <u>irreducible</u> rational function which satisfies the constraints. However, the cases $p = 0$ and $q = 0$ will certainly always yield an irreducible result. Further, in the Newton-Raphson case ($n = p = q + 1$), if there exists a polynomial ϕ such that

$$f_{p,q} = (z\phi' - \phi)/\phi'$$

it follows that the $f_{p,q}$ can be reducible only if the z_i are not distinct (one of them is a multiple root of ϕ). The problem of multiple fixpoints is discussed in Section 6.5 and the application to polynomials in Section 6.4.

6.3 Rational Iterating Functions of Given Order σ

Evidently, the above results may be generalized. Given n distinct points $z_1, \cdots, z_n$, and any integer $\sigma > 1$, it is possible to construct rational iterating functions $f_{p,q}$ with order of convergence σ at each of the given points by finding solutions to the σn linear homogeneous equations

$$(16) \qquad f_{p,q}(z_i) = z_i,$$

$$f_{p,q}^{(s-1)}(z_i) = 0, \quad s = 2, \cdots, \sigma,$$

where $i = 1, \cdots, n$. Introducing the notation $g^{(0)} \equiv g$, $h^{(0)} \equiv h$, these equations may be briefly re-expressed in the form of (10):

$$(17) \qquad g^{(s-1)}(z_i) - z_i h^{(s-1)}(z_i) = 0,$$

for $s = 1, \cdots, \sigma$, $i = 1, \cdots, n$. In accordance with (12), define

$$(18)\quad \gamma_j^{(s)} = (j + s - 1) \cdots (j + 1)\alpha_{j+s-1} - (j + s - 2) \cdots (j)\beta_{j+s-2},$$

for $s = 2, \cdots, \sigma$, $j = 0, \cdots, m - s + 1$, and then rewrite (17) in the form

$$(19)\qquad \sum_{0 \leq j \leq m-s+1} z_i^j \gamma_j^{(s)} = 0,$$

for $s = 1, \cdots, \sigma$, $i = 1, \cdots, n$, in order, as before, to derive general solutions

$$(20)\qquad \gamma_j^{(s)} = (-1)^{n-j} \sum_{n \leq K \leq m-s+1} S_n^{j+1,K} \gamma_K^{(s)},$$

$s = 1, \cdots, \sigma$, $j = 0, \cdots, n - 1$, valid provided $p + q + 1 \geq \sigma n$. Hence

Theorem 6.3 In order that a rational function $f_{p,q}(z)$ satisfy the σn conditions (16), it is necessary that $p + q + 1 \geq \sigma n$. $f_{p,q}(z)$ satisfying (16) may be constructed by arbitrarily choosing $(p + q + 2) - \sigma n$ of the coefficients (α_j, β_k), and then using (18) and (20) to find the remaining σn coefficients, which are then uniquely determined as symmetric functions of the given points z_i, $i = 1, \cdots, n$.

In terms of the iteration classes C_M defined at the beginning of this chapter, Theorem 6.3 means that C_M is null whenever $M < \sigma n$. Moreover, for $M =$

σn, each of the σn possible choices of p and q yields a single unique iterating function $f_{p,q}$; in other words, $C_{\sigma n}$ contains exactly σn members. For $M > \sigma n$, each choice of p and q yields an infinite subclass of iterating functions with coefficients dependent on the particular choice of the excess $(p + q + 2) - \sigma n$ coefficients. The case $M = \sigma n$ is, therefore, of particular interest, not only because it is a minimum (hence minimizing the time required for function evaluations, as well as the number of possible attractive cycles) but also because $C_{\sigma n}$ is finite.

In this context, recall that the Schröder and König iteration classes mentioned in Chapter 4 contain one iteration for each order of convergence σ. After some computation, the Schröder iterations are found to have denominators of degree $(2\sigma - 3)(n - 1)$ and numerators of degree one greater (they are Newton-Raphson-like in form). Hence in the new terminology Schröder's $f^{[\sigma]}$ is one of the infinitely many iterations

$$f_{(2\sigma-3)(n-1)+1,(2\sigma-3)(n-1)} \in C_{2(2\sigma-3)(n-1)+2},$$

and it is clear that $f^{[\sigma]}$ depends upon an implicit arbitrary choice of $2(2\sigma - 3)(n - 1) + 2 - \sigma n$ or $(\sigma - 2)(3n - 4)$ coefficients. As expected, in the case $\sigma = 2$, $f^{[2]} = f_{n,n-1}$, the Newton-Raphson iteration.

Similarly, the König iteration $f^{[\sigma]}$, also Newton-Raphson-like, is one of the infinitely many iterations

$$f_{(\sigma-1)(n-1)+1,(\sigma-1)(n-1)} \in C_{2(\sigma-1)(n-1)+2}$$

and depends upon an implicit arbitrary choice of $(\sigma - 2)(n - 2)$ coefficients.

For $\sigma > 2$, then, both these iteration classes are probably less than optimal.

6.4 Rational Iterating Functions for Zeros of a Polynomial $\phi_n(z)$

As mentioned earlier, it is shown in Chapter 7 that the quantities $S_n^{j,k} = S_n^{j,k}(z_1, \cdots, z_n)$ used to compute the coefficients of $f_{p,q}$ may themselves be computed from Newton sums. Since the Newton sums may in their turn be computed from the coefficients only of a polynomial [H9, 453], without direct knowledge of the roots, it is natural to seek to apply the method to construct rational iterating functions convergent to the unknown (but distinct) zeros of a given polynomial $\phi_n(z)$. As an illustration a polynomial suggested by Brooker's example [B2] is used:

$$\phi_4(z) = z^4 - 4z^2 - 5.$$

This polynomial has the property that the Newton-Raphson iteration applied to it gives rise to an attractive cycle. It has zeros $\pm\sqrt{5}$ and $\pm i$, and the Newton sums may be read off from the coefficients:

$$S_4^{1,4} = -5,\ S_4^{2,4} = 0,\ S_4^{3,4} = -4,\ S_4^{4,4} = 0.$$

With these values as the only input data, and making use of Theorem 6.3

together with Algorithm 7.5 and the other results of Chapter 7, it is possible to generate all the rational iterating functions $f_{p,q}$ which converge with order $\sigma = 2$ to the zeros of ϕ_4 and which belong to the minimum class $C_8(2; +\sqrt{5}, -\sqrt{5}, +i, -i)$. Table 6.1 displays the results: observe that while $f_{4,3}$, for example, gives rise to the attractive cycle $(+\sqrt{5/7}, -\sqrt{5/7})$, the iterations $f_{2,5}$, $f_{6,1}$, $f_{1,6}$, and $f_{7,0}$ converge to some zero of ϕ_4 from each zero of $f'_{p,q}$.

TABLE 6.1

Iterating functions for $\phi_4(z) = z^4 - 4z^2 - 5$ which belong to minimum class $C_8(2; +\sqrt{5}, -\sqrt{5}, +i, -i)$

m	p	q	iterating function $f_{p,q}$	additional critical points (cps)*	fixpoints or cycles reached from cps
4	4	3	$(3z^4 - 4z^2 + 5)/(4z^3 - 8z)$	$\pm \sqrt{2/3}$	$(+\sqrt{5/7}, -\sqrt{5/7})$
5	3	4	$(8z^3 + 20z)/(z^4 + 4z^2 + 15)$	$\pm \sqrt{15/2}i$	$(+\sqrt{7}i, -\sqrt{7}i)$
5	5	2	$(z^5 - 12z^3 - 25z)/(-8z^2 - 20)$	$\pm (5/\sqrt{6})i$	∞
6	2	5	$(92z^2 - 40)/(3z^5 - 20z^3 + 109z)$	$\pm \sqrt{218/207}i$	$\pm i$
6	6	1	$(z^6 - 6z^4 - 33z^2 + 10)/(-36z)$	$\pm \sqrt{2/5}i$	$\pm i$
7	1	6	$-180z/(2z^6 - 21z^4 + 42z^2 - 115)$	$\pm \sqrt{23/10}$	$\pm \sqrt{5}$
7	7	0	$(2z^7 - 21z^5 + 42z^3 + 245z)/180$	$\pm \sqrt{7/2}$	$\pm \sqrt{5}$
8	0	7	$-900/(23z^7 - 174z^5 + 33z^3 + 1130z)$	$\pm \sqrt{226/161}$	$(+0.9285, -0.9285)$

---

* The roots $\pm\sqrt{5}$, $\pm i$ are <u>critical points</u> of each $f^{-1}_{p,q}$; that is, points at which $f'_{p,q} = 0$. Thus, in each case, two other critical points remain to be specified, and these may possibly belong to the immediate domain of convergence of an attractive cycle, if such exist, or of a root. Note that "convergence" to ∞ is possible.

Table 6.2 displays 5 of the 12 iterating functions in $C_{12}(3; +\sqrt{5}, -\sqrt{5}, +i, -i)$; that is, the minimal class of rational iterating functions which converge with order $\sigma = 3$ to the zeros of ϕ_4.

TABLE 6.2

Iterating functions for $\phi_4(z)$ which belong to $C_{12}(3; +\sqrt{5}, -\sqrt{5}, +i, -i)$

m	p	q	iterating function $f_{p,q}$
6	6	5	$(49z^6 - 12z^4 + 675z^2 - 200)/(87z^5 - 204z^3 + 645z)$
7	5	6	$(129z^5 + 204z^3 + 435z)/(8z^6 + 135z^4 + 12z^2 + 245)$
7	7	4	$(8z^7 - 231z^5 - 252z^3 - 805z)/(-273z^4 + 84z^2 - 435)$
8	4	7	$(1443z^4 - 8220z^2 + 3225)/(49z^7 - 576z^5 + 3363z^3 - 8900z)$
8	8	3	$(29z^8 - 280z^6 - 1050z^4 + 4760z^2 - 2775)/(-2016z^3 + 5760z)$

The determination of the coefficients (α_j, β_k) for this example involved the solution of the simultaneous equations (20) in accordance with Theorem 6.3. This solution was facilitated by the fact that these equations are sparse in the (α_j, β_k); and an estimate of the maximum density of the matrix representing the equations may, therefore, be of general interest. (20) represents σn equations in $p + q + 2$ variables. The left-hand side of (20) specifies at most $2\sigma n$ nonzero coefficients (α_j, β_k), hence represents at most that many matrix positions. The right-hand side specifies at most $\Sigma_{1<s<\sigma}(m - s + 1 - n + 1)$ nonzero coefficients. Maximum nonzero entries in the coefficient matrix are, therefore,

$$2\sigma n + \Sigma\ (m - s - n + 2) = (\sigma/2)(2n + 2m - \sigma + 3),$$

out of a total $\sigma n(p + q + 2)$ positions. The maximum ratio of non-zero positions in the matrix representation of (20) is, therefore,

$$(2n + 2m - \sigma + 3)/2n(p + q + 2).$$

Letting m take its maximum value σn, and considering the "minimal" case $p + q + 1 = n$, this ratio becomes after some manipulation

$$1/n + (2n - \sigma + 1)/2n(\sigma n + 1),$$

which, since $\sigma > 2$, is less than $3/2n$. The system (20) will thus, in general, be fairly sparse.

6.5 Extension of Results to Multiple Fixpoints

In the preceding development it has most obviously been the definition of S_n^{j,k^1} in terms of an expression with denominator D_n (vanishing whenever $z_i = z_{i'}$, $i \neq i'$) which has precluded consideration of multiple fixpoints. In this section it will be found that the vanishing of D_n is not in itself a fundamental problem, but that other difficulties, similar to those encountered when applying the Newton-Raphson iteration to a function with multiple zeros, inhibit full extension of the results to the multiple fixpoint case.

Begin by considering (2), which may be interpreted as expressing the condition on $g(z) - zh(z)$, a polynomial of degree m, that it have $n < m$ roots z_i;

that is, that

$$\Big[\sum_{0\leq j\leq m} \gamma_j^{(1)} z^j\Big] = g(z) - zh(z)$$

$$= c(z) \prod_{1\leq i\leq n} (z - z_i) = \Big[c(z) \sum_{0\leq i\leq n} (-1)^{n-i} S_n^{i+1,n} z^i\Big],$$

where the $S_n^{i+1,n}$ are as before simply the Newton sums, $S_n^{n+1,n} \equiv 1$, and

$$c(z) = c_n + c_{n+1}z + \cdots + c_{m-1}z^{m-n-1} + c_m z^{m-n}, \qquad c_m \neq 0,$$

is a polynomial with arbitrary coefficients introduced so that the bracketed polynomials should be of equal degree m. Recall, however, that some of the left-hand coefficients — to be precise, $\gamma_K^{(1)}$, K = n, $\cdots$, m — may also be arbitrarily chosen; and it follows then that it is possible to either express the $\gamma_K^{(1)}$ in terms of the c_K, K = n, $\cdots$, m, or vice versa, simply by equating the coefficients of like powers of z:

$$\text{(A)} \qquad \begin{cases} \gamma_0^{(1)} = (-1)^n S_n^{1,n} c_n \\ \gamma_1^{(1)} = (-1)^{n-1} S_n^{2,n} c_n + (-1)^n S_n^{1,n} c_{n+1} \\ \quad \vdots \\ \gamma_{n-1}^{(1)} = \sum\limits_{0\leq K\leq \min(n-1,m-n)} (-1)^{K+1} S_n^{n-K,n} c_{n+K} \end{cases}$$

$$
(B)\quad \begin{cases}
\gamma_n^{(1)} = \sum_{0 \le K \le \min(n,m-n)} (-1)^K S_n^{n-K+1,n} c_{n+K} \\
\vdots \\
\gamma_{m-1}^{(1)} = S_n^{n+1,n} c_{m-n-1} + (-1) S_n^{n,n} c_{m-n}, \\
\gamma_m^{(1)} = S_n^{n+1,n} c_{m-n}.
\end{cases}
$$

(B) may first of all be solved for the c_K in terms of the $\gamma_K^{(1)}$, $K = n, \cdots, m$; this may clearly be accomplished by a process of back substitution. Then, by substituting the resulting expressions for the c_K into (A), it will be possible to write down solutions which express the $\gamma_j^{(1)}$, $j = 0, \cdots, n - 1$ in terms of the $\gamma_K^{(1)}$, $K = n, \cdots, m$, and the Newton sums, just as in the preceding sections, but this time avoiding the use of Vandermonde determinants. It is not necessary actually to carry out this complex calculation, however, since it is already known, from our previous work, what the result will be. For distinct z_i, it will be found, as before, that

$$
(8)\qquad \gamma_j^{(1)} = (-1)^{n-j} \sum_{n \le K \le m} S_n^{j+1,K} \gamma_K^{(1)}, \quad j = 0, \cdots, n - 1,
$$

but with the important difference now that the values $S_n^{j+1,K}$ are no longer defined in terms of Vandermonde determinants, but may now be regarded as defined by relation (10) of Chapter 7, which states that for $k \ge n \ge 1$ and $1 \le j \le n$

$$
S_n^{j,k} = \sum_{0 \le j' \le j-1} (-1)^{j'} S_n^{n,k-j'-1} S_n^{j-j',n}.
$$

This relation has the special case

$$S_n^{n,k} = \sum_{0 \leq j' \leq n-1} (-1)^{j'} S_n^{n,k-j'-1} S_n^{n-j',n},$$

which may be used, along with Algorithm 7.5, to recursively determine the $S_n^{n,k}$, hence the $S_n^{j,k}$, from the Newton sums and $S_n^{n,n-1} \equiv 1$. Based on this second definition of the $S_n^{j+1,K}$, these calculations may now equally be carried out for the case of multiple z_i; and they will, by the continuity of $\gamma_j^{(1)}(z_1, \cdots, z_n)$, yield the same result (8). Consequently,

<u>Theorem 6.4</u> Theorem 6.1 holds also in the case that the n given fixpoints z_i are not necessarily distinct.

Difficulties arise, however, with the extension of Theorems 6.2 and 6.3 to the case of multiple fixpoints. For suppose z_i is a fixpoint of exact multiplicity $a > 1$. Then

$$g(z) - zh(z) = (z - z_i)^a A(z),$$

for some polynomial $A(z)$ such that $A(z_i) \neq 0$. Since

$$g'(z) - zh'(z) = a(z - z_i)^{a-1} A(z) + (z - z_i)^a A'(z) + h(z)$$

and

$$f'(z) = [g'(z) - f(z)h'(z)]/h(z),$$

it may be concluded that if $(z - z_i)^{a-1}$ is not a factor of h(z),

$$f'(z_i) = \lim_{z \to z_i} [a(z - z_i)^{a-1}A(z) + (z - z_i)^a A'(z) + h(z) - (f(z) - z)h'(z)] \;/\; h(z)$$

$$= 1,$$

so that z_i cannot be an attractive, but is rather an indifferent, fixpoint of f. In this case, then, (10) and (17) are both internally inconsistent, and a meaningful solution cannot be expected.

Suppose then that $(z - z_i)^{a-1}$ is a factor of h(z), and that a > 1. Then $h(z) = (z - z_i)^{a-1}B(z)$, for some polynomial B(z), and

$$f(z) = (z - z_i)^{a-1}[zB(z) + (z - z_i)A(z)] \;/\; (z - z_i)^{a-1}B(z),$$

which, since the construction process makes no provision for cancellation of identical factors in numerator and denominator, is of the form 0/0 at $z = z_i$. Furthermore, f'(z) is of the same form 0/0 at $z = z_i$, and it appears therefore that in this case also multiple fixpoints introduce internal inconsistencies into (10) and (17).

It becomes clear then that for multiple fixpoints an improved construction process is required which will somehow cancel the terms $(z - z_i)^{a-1}$ to yield

$$\hat{f}(z) = z + (z - z_i)A(z)/B(z) \text{ and } \hat{f}'(z)|_{z=z_i} = 1 + A(z_i)/B(z_i),$$

effectively reducing the problem to the case of distinct fixpoints, and permitting solution of (10) and (17) as before.

It is nevertheless interesting to see what happens when the equations (10) or (17) are applied to a particular example of multiple fixpoints. The results of such an application are displayed in Table 6.3. Provided both $g_p(z)$ and $h_q(z)$ are of degree sufficiently large to contain a factor $(z - z_i)^{a-1}$, it seems that the iterating function $f_{p,q}$ may be used to solve the given problem, but with a lower order of convergence at multiple fixpoints z_i. The case in which $(z - z_i)^{a-1}$ does not divide h(z) apparently does not occur. The repeated result $\hat{f}'(1) = 1/2$ suggests a possible generalization of the well-known fact [H9, 53] that for the Newton-Raphson iteration, at a fixpoint z_i of multiplicity a,

$$f'(z_i) = 1 - 1/a.$$

6.6 Discussion

In this chapter a method of classifying all rational iterating functions -- especially those with a prescribed order of convergence — has been developed. In the context of the general theory of the global convergence

TABLE 6.3

Iterating functions for $\phi_4(z) = z^4 - 2z^3 + 2z^2 - 2z + 1$ with roots $\pm i$, 1, 1

m	p	q	constructed function $f_{p,q}$	corresponding iterating function $\hat{f}_{p,q}$	$\hat{f}(1)$	$\hat{f}'(1)$	$\hat{f}(\pm i)$	$\hat{f}'(\pm i)$
4	4	3	$(3z^4 - 4z^3 + 2z^2 - 1)/(4z^3 - 6z^2 + 4z - 2)$	$(3z^3 - z^2 + z + 1)/(4z^2 - 2z + 2)$	1	1/2	$\pm i$	0
5	3	4	$(-4z^3 + 14z^2 - 12z + 2)/(3z^4 - 8z^3 + 6z^2 + 4z - 5)$	$(-4z^2 + 10z - 2)/(3z^3 - 5z^2 + z + 5)$	1	1/2	$\pm i$	0
5	5	2	$(2z^5 - 5z^4 + 8z^3 - 14z^2 + 10z - 1)/(2z^2 - 8z + 6)$	$(2z^4 - 3z^3 + 5z^2 - 9z + 1)/(2z - 6)$	1	1/2	$\pm i$	0
6	2	5	$(14z^2 - 16z + 2)/(2z^5 - z^4 - 4z^3 + 6z^2 + 6z - 9)$	$(14z - 2)/(2z^4 + z^3 - 3z^2 + 3z + 9)$	1	1/2	$\pm i$	0
6	6	1	$(z^6 - 3z^4 + 4z^3 - 13z^2 + 12z - 1)/(-8z + 8)$	$(z^5 + z^4 - 2z^3 + 2z^2 - 11z + 1)/(-8)$	1	1/2	$\pm i$	0
7	1	6	$(-8z + 8)/(7z^6 - 20z^5 + 31z^4 - 44z^3 + 45z^2 - 32z + 13)$	$(-8)/(7z^5 - 13z^4 + 18z^3 - 26z^2 + 19z - 13)$	1	1/2	$\pm i$	0
7	7	0	$(z - 1)(z^2 + 1)\phi_4(z)/0$					
8	0	7	$0/(8z^3 - 15z^2 + 22z - 29)\phi_4(z)$					

properties of iterations of rational functions, this method has been used to generate classes of rational iterating functions which converge to the distinct zeros of a given polynomial. The method then becomes a tool for the empirical investigation of iterations which do or do not give rise to unwanted attractive cycles.

There are several interesting areas for future research in connection with these results:

(1) improvement of the construction method to eliminate iterations giving rise to unwanted attractive cycles;

(2) improvement of the construction method to recognize and converge to multiple zeros (perhaps achieved by the approach of Howland [H11]);

(3) improvement of the construction method to eliminate altogether the possibility of reducible rational functions;

(4) proof of, or a counterexample to, the conjecture that the number of attractive fixpoints (not cycles) is at most equal to the number of critical points of the inverse iteration;

(5) statement and proof of a constant limit theorem for rational iterating functions as discussed at the end of Chapter 5;

(6) extension of the theoretical results of Chapter 5 to meromorphic functions.

7 Appendix -- Generalized Newton Sums

This chapter is essentially an appendix to Chapter 6 in which a collection of results related to Vandermonde determinants and generalized Newton sums are proved. These results generally depend upon elementary, but still rather difficult, manipulative mathematics.

Given integers $n \geq 1$, j, and k, and distinct (real or complex) values z_i, $i = 1, \cdots, n$, quantities $S_n^{j,k} = S_n^{j,k}(z_1, \cdots, z_n)$ are defined as follows:

$$S_n^{j,k} = (-1)^{n-j} D_n^{j,k}/D_n \quad \text{for } 1 \leq j \leq n \text{ and } k \geq 0;$$

$$= 0 \text{ otherwise,}$$

where $D_n = D_n(z_1, \cdots, z_n)$ is the Vandermonde determinant of order n,

$$\begin{vmatrix} 1 & z_1 & \cdots & z_1^{n-1} \\ 1 & z_2 & \cdots & z_2^{n-1} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 1 & z_n & \cdots & z_n^{n-1} \end{vmatrix},$$

and $D_n^{j,k}$ is the determinant obtained by replacing the jth column of D_n, denoted (z_i^{j-1}), by the new column (z_i^k). Then, in particular, $D_1 = 1$ and $S_1^{1,k} = D_1^{1,k} = z_1^k$. For $n \geq 2$, recall that

(1) $$D_n = \Pi_n D_{n-1},$$

where $\Pi_n = \Pi_{1 \leq i \leq n-1}(z_n - z_i)$, from which it follows that

(2) $$D_n = \prod_{1 \leq i' < i \leq n} (z_i - z_{i'}).$$

Note also that when $k \leq n - 1$,

$$D_n^{j,k} = D_n \quad \text{for } j = k + 1;$$

$$= 0 \text{ otherwise.}$$

Theorem 7.1 For $k \geq n \geq 2$,

$$S_n^{j,k} = \sum_{1 \leq j' \leq k-n+1} S_{n-1}^{j,k-j'} z_n^{j'} + \sum_{0 \leq j' \leq k-n+1} S_{n-1}^{j-1,k-j'-1} z_n^{j'}.$$

Proof The result is trivially true for $j < 1$ and for $j > n$. For $j = n$,

$$D_n^{n,k} = \begin{vmatrix} 1 & z_1 & \cdots & z_1^{n-2} & z_1^k \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ 1 & z_n & \cdots & z_n^{n-2} & z_n^k \end{vmatrix}.$$

Subtracting the nth row of this determinant from each of the others, and then extracting the common factor Π_n, it is found that $D_n^{n,k}$ is equal to $(-1)^{n-1}(-1)^{n-1}\Pi_n$ times an n - 1 by n - 1 determinant whose h^{th} column, $1 \leq h \leq n-2$, may be represented

$$\sum_{1 \leq j' \leq h} z_n^{h-j'} z_i^{j'-1},$$

and whose (n - 1)th column is

$$\sum_{1 \leq j' \leq k} z_n^{k-j'} z_i^{j'-1}.$$

Then, subtracting z_n^{k-n+2} times the (n - 2)th column from the (n - 1)th column, and afterwards successively subtracting z_n times the (h - 1)th column from the h^{th} column, $h = n-2, \cdots, 2$,

$$D_n^{n,k} = \Pi_n \begin{vmatrix} 1 & z_1 & \cdots & z_1^{n-3} & \sum_{0 \leq j' \leq k-n+1} z_n^{j'} z_1^{k-j'-1} \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ 1 & z_{n-1} & \cdots & z_{n-1}^{n-3} & \sum_{0 \leq j' \leq k-n+1} z_n^{j'} z_{n-1}^{k-j'-1} \end{vmatrix}$$

$$= \Pi_n \sum_{0 \leq j' \leq k-n+1} D_{n-1}^{n-1,k-j'-1} z_n^{j'},$$

from which, in view of (1) and the definition of $S_n^{j,k}$, the result follows immediately. Similar manipulations establish the theorem in the case $1 \leq j < n$. ∎

Some special cases of Theorem 7.1 may be noted. For $j = 1$,

(4) $$S_n^{1,k} = \sum_{1 \leq j' \leq k-n+1} S_{n-1}^{1,k-j'} z_n^{j'},$$

and for j = n,

(5) $$S_n^{n,k} = \sum_{1 \leq j' \leq k-n+1} S_{n-1}^{n-1,k-j'-1} z_n^{j'}.$$

For k = n,

(6) $$S_n^{j,n} = (S_{n-1}^{j,n-1} + S_{n-1}^{j-1,n-2}) z_n + S_{n-1}^{j-1,n-1}$$

whence

(7) $$S_n^{1,n} = S_{n-1}^{1,n-1} z_n,$$

(8) $$S_n^{n,n} = S_{n-1}^{n-1,n-1} + z_n,$$

and the following theorem may easily be proved by induction.

<u>Theorem 7.2</u> For $1 \leq j \leq n$ the quantities $S_n^{j,n}(z_1, \cdots, z_n)$ are the Newton sums for a polynomial of degree n with distinct zeros z_i, i = 1, ⋯, n; that is, $S_n^{j,n}$ is the sum of all distinct products of the z_i taken n − j + 1 at a time.*

Theorem 7.2 establishes then the important fact that given any polynomial

*Using Theorem 7.1, it may also be shown more generally that for $1 \leq j \leq n$ and $k > n$, $S_n^{j,k}$ is the sum of all distinct products $z_{b'}^{a'} z_b^{b'} \cdots$, where $\max(a', b', \cdots) = k - n + 1$ and $a' + b' + \cdots = k - j + 1$.

$f_n(z)$ of degree n, corresponding quantities $S_n^{j,n}$ may be computed only from the coefficients of f_n. Theorem 7.3 shows that in fact every $S_n^{j,k}$ may be computed from the $S_n^{j,n}$, hence also from the coefficients of f_n.

<u>Theorem 7.3</u> For $k > n > 1$ and $1 < j < n$,

$$S_n^{j,k} + S_n^{j-1,k-1} = S_n^{n,k-1} S_n^{j,n}.$$

<u>Proof</u> Using the definition of $S_n^{j,k}$, this result may easily be verified for $n = 1$. Suppose it is known that

$$(9) \qquad S_{n-1}^{j-1,k} + S_{n-1}^{j-2,k-1} = S_{n-1}^{n-1,k-1} S_{n-1}^{j-1,n-1}$$

for all $k > n - 1 > 1$ and $1 < j < n$. Then by Theorem 7.1

$$S_n^{j,k} + S_n^{j-1,k-1} = \sum_{1<j'<k-n+1} S_{n-1}^{j,k-j'} z_n^{j'} + \sum_{0<j'<k-n+1} S_{n-1}^{j-1,k-j'-1} z_n^{j'}$$

$$+ \sum_{1<j'<k-n} S_{n-1}^{j-1,k-j'-1} z_n^{j'} + \sum_{0<j'<k-n} S_{n-1}^{j-2,k-j'-2} z_n^{j'}$$

$$= \sum_{1<j'<k-n+1} \{S_{n-1}^{j,k-j'} + S_{n-1}^{j-1,k-j'-1}\} z_n^{j'}$$

$$+ \sum_{0<j'<k-n} \{S_{n-1}^{j,k-j'-1} + S_{n-1}^{j-2,k-j'-2}\} z_n^{j'}$$

$$= \sum_{1 \leq j' \leq k-n+1} S_{n-1}^{n-1,k-j'-1} S_{n-1}^{j,n-1} z_n^{j'}$$

$$+ \sum_{0 \leq j' \leq k-n} S_{n-1}^{n-1,k-j'-2} S_{n-1}^{j-1,n-1} z_n^{j'},$$

by the inductive hypothesis, if $j \neq n$; then by (5) and (6),

$$S_n^{j,k} + S_n^{j-1,k-1} = \sum_{0 \leq j' \leq k-n} S_{n-1}^{n-1,k-j'-2} z_n^{j'} \{z_n S_{n-1}^{j,n-1} + S_{n-1}^{j-1,n-1}\}$$

$$= S_n^{j,n} S_n^{n,k-1}.$$

If $j = n$, then by (5) $S_n^{n,k} = \sum_{1 \leq j' \leq k-n+1} S_{n-1}^{n-1,k-j'-1} z_n^{j'} + S_{n-1}^{n-1,k-1}$

and

$$S_n^{n,k} + S_n^{n-1,k-1} = \sum_{1 \leq j' \leq k-n+1} S_{n-1}^{n-1,k-j'-1} z_n^{j'} + \sum_{0 \leq j' \leq k-n} S_{n-1}^{n-1,k-j'-2} z_n^{j'}$$

$$= \sum_{0 \leq j' \leq k-n} S_{n-1}^{n-1,k-j'-1} \{z_n + S_{n-1}^{n-1,n-1}\} = S_n^{n,k-1} S_n^{n,n},$$

by application of (5) and (8). This completes the proof. ∎

An immediate consequence of Theorem 7.3 is

<u>Theorem 7.4</u> For $k \geq n \geq 1$ and $1 \leq j \leq n$,

(10) $$S_n^{j,k} = \sum_{0<j'<j-1} (-1)^{j'} S_n^{n,k-j'-1} S_n^{j-j',n}.$$

Theorem 7.4 shows that the determination of the $S_n^{j,k}$ may be reduced to the determination of the Newton sums $S_n^{1,n}$, $\cdots$, $S_n^{j,n}$ together with the quantities $S_n^{n,k-1}$, $\cdots$, $S_n^{n,n+1}$. But these latter quantities also may be calculated from the Newton sums, since by rewriting (10) for the special case $j = n$,

$$S_n^{n,k} = \sum_{0<j'<n-1} (-1)^{j'} S_n^{n,k-j'-1} S_n^{n-j',n},$$

then successively setting $k = n + 1, n + 2, \cdots$, it is possible to recursively determine all the $S_n^{n,k}$ in terms of the Newton sums and $S_n^{n,n-1} = 1$. This is done by the following

Algorithm 7.5 Given Newton sums $S_n^{1,n}$, $\cdots$, $S_n^{n,n}$ stored in SN(1), $\cdots$, SN(n), respectively, to calculate the values $S_n^{n,n-1} = 1.0$, $S_n^{n,n} =$ SN(n), $S_n^{n,n+1}$, $\cdots$, $S_n^{n,n+K}$, and store them in SK(1), $\cdots$, SK(K+2).

(1) SK(1) $\leftarrow$ 1.0, SK(2) $\leftarrow$ SN(n), k $\leftarrow$ n + 1.

(2) j' $\leftarrow$ 0, SIGN $\leftarrow$ 1.0, $\alpha \leftarrow$ k − n + 2, SK(α) $\leftarrow$ 0.0.

(3) SK(α) $\leftarrow$ SK(α) + SIGN*SN(n−j')*SK(α − j' − 1).

(4) If k − j' $\neq$ n and j' $\neq$ n − 1, SIGN $\leftarrow$ − SIGN, j' $\leftarrow$ j' + 1, go to (3).

(5) If $k \neq K$, $k \leftarrow k + 1$, go to (2).

(6) Exit.

It has recently been suggested [H11] that the use of Waring-Lagrange polynomials may significantly reduce the complexity of the proofs in this chapter. This possibility is currently under investigation.

8 Acknowledgement

I take pleasure in acknowledging the support of Professor James L. Howland, University of Ottawa, over a period of 18 years!

9 References

[A1] Ascher, Marcia, Cycling in the Newton-Raphson iteration, Internat. J. Math. Ed. Sci. Tech. 5-2 (1974) 229-235.

[B1] Borodin, A. & Munro, I., The computational complexity of algebraic and numeric problems, American Elsevier (1975).

[B2] Brooker, R.A., The solution of algebraic equations on the EDSAC, Proc. Cambridge Philos. Soc. 48 (1952) 255-270.

[B3] Baker, I.N., Fixpoints and iterates of entire functions, Math. Z. 71 (1959) 146-153.

[B4] ————, The existence of fixpoints of entire functions, Math. Z. 73 (1960) 280-284.

[B5] ————, Multiply connected domains of normality in iteration theory, Math. Z. 81 (1963) 206-214.

[B6] ————, Sets of non-normality in iteration theory, J. London Math. Soc. 40 (1965) 499-502.

[B7] ————, Repulsive fixpoints of entire functions, Math. Z. 104 (1968) 252-256.

[B8] ————, Limit functions and sets of non-normality in iteration theory, Ann. Acad. Sci. Fenn. Ser. A-I 467 (1970) 1-11.

[B9] Brent, R., Winograd, S. & Wolfe, P., Optimal iterative processes for root-finding, Num. math. 20-5 (1973) 327-341.

[B10] Brodlie, Kenneth W., On Bairstow's method for the solution of polynomial equations, Math. Comp. 29-131 (1975) 816-826.

[B11] Babbage, Charles, An essay toward the calculus of functions, Trans. Royal Soc. London (1815) 389-423.

[B12] Burckel, R.B., Iterating analytic self-maps of discs, Amer. Math. Monthly 88-6 (1981) 396-407.

[B13] Brolin, M., Invariant sets under iterations of rational functions, Arkiv för Mathematik 6 (1965) 103-144.

[C1] Cremer, Hubert, Über die Iteration rationaler Funktionen, Jahresbericht Deutschen Math. - Vereinigung 33-1 Abt. Heft 9/12 (1925) 185-210.

[D1] Durand, E., Solutions numériques des équations algébriques I, Masson (1960).

[D2] , Solutions numériques des équations algébriques II, Masson (1961).

[F1] Fatou, P., Sur les équations fonctionnelles I, Bull. Soc. math. France 47 (1919) 161-271.

[F2] ———, Sur les équations fonctionnelles II, Bull. Soc. Math. France 48 (1920) 33-94.

[F3] ———, Sur les équations fonctionnelles III, Bull. Soc. Math. France 48 (1920) 208-314.

[F4] Faddeeva, V.N., Computational methods of linear algebra, Dover (1959).

[F5] Ford, J.A., A generalization of the Jenkins-Traub method, Math. Comp. 31-137 (1977) 193-203.

[F6] Farmer, M.R. & Loizou, G., A class of iteration functions for improving, simultaneously, approximations to the zeros of a polynomial, BIT 15-3 (1975) 250-258.

[G1] Graves-Morris, P.R. & Hopkins, T.R., Reliable rational interpolation, Numerische Math. 36 (1981) 111-128.

[H1] Householder, Alston S., Principles of numerical analysis, McGraw-Hill (1953).

[H2] ———, The theory of matrices in numerical analysis, Blaisdell (1964).

[H3] ________, Schroder and Trudi: a historical excursion, SIAM Review 16-3 (1974) 344-348.

[H4] Henrici, Peter, Applied and computational complex analysis I, Wiley & Sons (1974).

[H5] ________, Applied and computational complex analysis II, Wiley & Sons (1977).

[H6] ________, Elements of numerical analysis, Wiley & Sons (1964).

[H7] Hille, Einar, Analytic function theory I, Blaisdell (1959).

[H8] ________, Analytic function theory II, Blaisdell (1962).

[H9] Hall, H.S. & Knight, S.R., Higher algebra, MacMillan (1950).

[H10] Hansen, Eldon & Patrick, Merrell, A family of root finding methods, Num. Math. 27-3 (1976-1977) 257-269.

[H11] Howland, James L., private communications (1983).

[H12] Hildebrand, F.B., Introduction to numerical analysis, 2nd ed., McGraw-Hill (1974).

[J1] Julia, Gaston, Mémoire sur l'itération des fonctions rationnelles, J.

Math. Pures Appl. 8-1 (1918) 47-245.

[J2] Jenkins, M.A. & Traub, J.F., A three-stage algorithm for real polynomials using quadratic iteration, SIAM J. Num. Anal. 7 (1970) 545-566.

[K1] Knuth, Donald E., The art of computer programming I: fundamental algorithms, 2nd ed., Addison-Wesley (1973).

[K2] , The art of computer programming II: seminumerical algorithms, Addison-Wesley (1969).

[K3] , The art of computer programming III: sorting and searching, Addison-Wesley (1973).

[K4] Klein, Morris, Mathematical thought from ancient to modern times, Oxford University Press (1972).

[K5] Kuczma, Marek, Functional equations in a single variable, Polish Scientific Publishers (1968).

[L1] Lattes, S., Sur l'itération des substitutions rationnelles et les fonctions de Poincaré, C.R. Acad. Sci. Fr. 166 (1918) 26-28.

[L2] Leau, L., Etude sur les équations fonctionnelles, Annales Fac. Sciences Toulouse série 1, 11 (1897) 22, 29-35.

[M1] Montel, Paul, L'itération, Univ. Nac. La Plata Publ. Fac. Ci. Fisicomat. 3 (1940) 201-211.

[M2] ______, Lecons sur les familles normales de fonctions analytiques et leurs applications, Gauthier-Villars (1927).

[N1] Nourein, Abdel Wahab, Root determination by use of Padé approximants, BIT 16-3 (1976) 291-297.

[O1] Ostrowski, A.M., Über Folgen analytischer Funktionen und einige Verschärfungen des Picardschen Satzes, Math. Z. 24 (1925) 215-258.

[O2] Ostrowski, A.M., Solution of equations and systems of equations, Academic Press (1973).

[P1] Pomentale, T., A class of iterative methods for holomorphic functions, Num. Math. 18-3 (1971) 193-203.

[R1] Rådström, H., On the iteration of analytic functions, Math. Scand. I (1953) 85-92.

[R2] Rice, R.E., Schweizer, B. & Sklar, A., When is $f(f(z)) = az^2 + bz + c$?, Amer. Math. Monthly 87-1 (1980) 252-263.

[S1] Smyth, W.F., The construction of rational iterating functions, Math. Comp. 32-143 (1978) 811-827.

[S2] ———, Constant limit of a sequence of iterates, SIAM J. Math. Anal. 10-3 (1979) 463-471.

[S3] Saks, Stanislaw & Zygmund, Antoni, Analytic functions, 3rd ed., American Elsevier (1971).

[S4] Schröder, Ernst, Über unendlich viele Algorithmen zur Auflösung der Gleichungen, Math. Ann. 23 (1870) 317-365.

[S5] ———, Über iterierte Funktionen, Math. Ann. 3 (1871) 296-322.

[T1] Traub, J.F., Iterative methods for the solution of equations, Prentice-Hall (1964).

[T2] ———, A class of globally convergent iteration functions for the solution of polynomial equations, Math. Comp. 20 (1966) 113-138.

[T3] Temple University Problem Solving Group, The iterated sine, Amer. Math. Monthly 82-1 (1975) 82-83.

[T4] Targonski, György, Topics in iteration theory (Studia Mathematica No. 6), Vandenhoeck & Ruprecht, Göttingen (1981).

June 3, 1983
M/UCS/WFS.1/A/WFS1.0/dm