MATHEMATICS OF COMPUTATION
Volume 70, Number 235, Pages 1305–1310
S 0025-5718(00)01242-4
Article electronically published on March 6, 2000

ON ITERATES OF MÖBIUS TRANSFORMATIONS ON FIELDS

SAM NORTHSHIELD

ABSTRACT. Let p be a quadratic polynomial over a splitting field K, and S be the set of zeros of p. We define an associative and commutative binary relation on $G \equiv K \cup \{\infty\} - S$ so that every Möbius transformation with fixed point set S is of the form x "plus" c for some c. This permits an easy proof of Aitken acceleration as well as generalizations of known results concerning Newton's method, the secant method, Halley's method, and higher order methods. If K is equipped with a norm, then we give necessary and sufficient conditions for the iterates of a Möbius transformation m to converge (necessarily to one of its fixed points) in the norm topology. Finally, we show that if the fixed points of m are distinct and the iterates of m converge, then Newton's method converges with order 2, and higher order generalizations converge accordingly.

Consider the Fibonacci sequence $F_1 = F_2 = 1$ and, for $n \geq 1$, $F_{n+2} = F_{n+1} + F_n$. It is known that the ratios $r_n \equiv F_{n+1}/F_n$ converge (and in fact are continued fraction convergents) to the "golden ratio" $\frac{1+\sqrt{5}}{2}$ (see, for example, [1], [2]). If $m(x) = 1 + 1/x$, then the sequence (r_n) satisfies the recursion $r_{n+1} = m(r_n)$ and so, letting n approach ∞, the iterates of m converge to a fixed point of m. We associate to m its *characteristic polynomial* $\theta(x) \equiv x^2 - x - 1$ (the monic polynomial whose zeros are the fixed points of m). Hence the iterates of m (starting with 1) converge to a zero of θ. Iteration by Newton's method (applied to θ and starting with 1) also converges to this zero and, in fact, gives the 1st, 2nd, 4th, 8th, 16th, ... iterates of m (see [2]). Iteration by the secant method (applied to θ and starting with 1,2) gives the 1st, 2nd, 3rd, 5th, 8th, 13th, ... iterates of m (see [2]).

This paper grew out of an attempt to understand these and other (known [1], [2], [3], [4], [5]) phenomena. We shall first generalize some of the results of [2], [3], [4], [5] to the case where m is a Möbius transformation (i.e., function of the form $x \longmapsto \frac{ax+b}{cx+d}$) where a, b, c, d, x are elements of an arbitrarily chosen field K. We shall generalize results of [2], [4] to our case as well as introduce generalizations of Newton's method. We shall also derive a generalization of *Aitken acceleration* (a main result of [3], [5]). Our proofs are different (and perhaps simpler) than the the extant proofs.

Next, we shall assume that K is equipped with a norm or absolute value and that K contains the fixed points of m. We give necessary and sufficient conditions for the iterates of m (with a given starting point) to converge to a given fixed

Received by the editor October 3, 1997 and, in revised form, January 18, 1999 and July 30, 1999.

2000 *Mathematics Subject Classification.* Primary 41A20; Secondary 65B99, 11B39, 12J20.

Key words and phrases. Fibonacci numbers, Newton's method, secant method, Aitken acceleration, Möbius transformation, field, order of convergence.

This paper was written while the author was on sabbatical leave at the University of Minnesota. The author appreciated the hospitality of the Mathematics Department during that time.

point of m. Given convergence of the iterates of m, we show that Newton's method converges quadratically and that its generalizations converge with correspondingly high order.

Let $\theta(x) \equiv x^2 - ax - b$. We shall assume that K contains the two zeros (ξ_1 and ξ_2) of θ; we allow for the possibility that $\xi_1 = \xi_2$. For any c, it is easily seen that θ is the characteristic polynomial of the Möbius transformations

$$m(x) = \frac{cx+b}{x-a+c} \quad \text{and} \quad m^{-1}(x) = \frac{(a-c)x+b}{x-c}. \tag{1}$$

As in the real case, we introduce two conditions regarding ∞: $m(\infty) = c$ and $m(a-c) = \infty$. Let $G = K \cup \{\infty\} - \{\xi_1, \xi_2\}$. Given any r_0, we can form a sequence (r_n) in *both* directions:

$$r_{n+1} = m(r_n), \qquad r_{n-1} = m^{-1}(r_n). \tag{2}$$

Interestingly, the numbers r_n are ratios of "generalized Fibonacci numbers". Given initial values $G_0 = 1$ and $G_1 = r_0$, define

$$G_{n+2} = \frac{cG_{n+1} + bG_n}{G_{n+1} + (c-a)G_n}.$$

Note that when $c = a$, $G_{n+2} = aG_{n+1} + bG_n$. In any case, $r_0 = G_1/G_0$ and, if $r_n = G_{n+1}/G_n$, then $r_{n+1} = m(r_n) = m(G_{n+1}/G_n) = G_{n+2}/G_{n+1}$ so that $r_n = G_{n+1}/G_n$ for all n.

Given $x, y \in G$, let

$$x \oplus y = \frac{xy+b}{x+y-a}.$$

Here the conventions regarding ∞ are $x \oplus (a-x) = \infty$ and $x \oplus \infty = x$. Although it is clear that the binary relation $\oplus$ is commutative, it is perhaps less clear that it is associative. In the real case, it is a challenging problem to show *geometrically* that this is so. (The connection to geometry in this case is that the line through $(x, \theta(x))$ and $(y, \theta(y))$ has the x-intercept $x \oplus y$).

Theorem 1. *The relation $\oplus$ is associative.*

Proof. Given any two-by-two matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, define an associated Möbius transformation $\Phi_A(x) = \frac{ax+b}{cx+d}$. It is well known (and easy to verify) that the composition of such functions corresponds to matrix multiplication; that is, $\Phi_A \circ \Phi_B = \Phi_{AB}$.

Let $M = \left(\begin{smallmatrix} 0 & b \\ 1 & -a \end{smallmatrix}\right)$. Note that $\Phi_{M+xI}(y) = x \oplus y$. Hence

$$(x \oplus y) \oplus z = \Phi_{M+zI}(x \oplus y) = \Phi_{M+zI}(\Phi_{M+xI}(y)) = \Phi_{(M+zI)(M+xI)}(y).$$

Since $M + zI$ and $M + xI$ commute, we have $(x \oplus y) \oplus z = (z \oplus y) \oplus x$. $\square$

We remark that by this theorem $(G, \oplus)$ is an abelian group with identity ∞ such that $a - x$ is the inverse of x. By the proof above, $(M + xI)(M + yI)$ is a scalar multiple of $M + (x \oplus y)I$, and so the map $x \longmapsto M + xI$ is a projective representation of G into $GL_2(K)$.

We let $x^{\oplus n}$ denote the n-fold "sum" of x (i.e., $x^{\oplus 1} = x$ and $x^{\oplus(n+1)} = x \oplus x^{\oplus n}$). This definition extends to any integer n by $x^{\oplus 0} = \infty$ and $x^{\oplus(-n)} = (a-x)^{\oplus n}$. Note that, by (1),

$$m(x) = x \oplus c \quad \text{and} \quad m^{-1}(x) = x \oplus (a-c).$$

Therefore, with r_k defined above,

$$r_{n+k} = r_k \oplus c^{\oplus n} \tag{3}$$

for any n and k. By associativity and commutativity,

$$\frac{r_{l-i}r_i + b}{r_{l-i} + r_i - a} = r_{l-i} \oplus r_i = r_k \oplus r_k \oplus c^{\oplus(l-2k)} = r_k \oplus r_{l-k}.$$

Hence, for any i and j

$$\frac{r_{l-i}r_i + b}{r_{l-i} + r_i - a} = \frac{r_{l-j}r_j + b}{r_{l-j} + r_j - a}.$$

Since they are equal, they are also equal to the ratio of differences (i.e., if $A/B = C/D$, then $A/B = (A-C)/(B-D)$) and we have

Theorem 2. *For all i, j and k such that the denominator of the fraction below is nonzero,*

$$\frac{r_{l-i}r_i - r_{l-j}r_j}{r_{l-i} + r_i - r_{l-j} - r_j} = r_k \oplus r_{l-k}.$$

Note that if $r_1 = c$ (equivalently, $r_0 = \infty$), then $r_n = c^{\oplus n}$. We shall use this in the next four results. The reader is invited to extend those results to the case when $r_1 \neq c$. The following is a generalization of the Aitken acceleration formula (see [3], [5])

Corollary 3. *If $r_1 = c$, then for all n and l,*

$$\frac{r_{n+l}r_{n-l} - r_n^2}{r_{n+l} - 2r_n + r_{n-l}} = r_{2n}.$$

Proof. Replace i, j, k, and l in Theorem 2 by $n-l, n, l$, and $2n$, respectively. □

When $K = \mathbb{R}$, Newton's method to approximate the zeros of θ is, given a starting point t_0, to define a sequence inductively

$$t_{n+1} = t_n - \frac{\theta(t_n)}{\theta'(t_n)},$$

which converges (in many cases) to a zero of θ. In our case, this boils down to

$$t_{n+1} = \frac{t_n^2 + b}{2t_n - a} = t_n \oplus t_n.$$

We take this to be the definition of Newton's method in the general case.

If we take $t_0 = c$, then a simple induction argument shows that $t_n = c^{\oplus 2^n}$. If (r_n) is defined as in (2) above, then $r_n = r_0 \oplus c^{\oplus n}$ and we have

Theorem 4. *If $t_0 = c$, then $t_n = (a - r_0) \oplus r_{2^n}$.*

One may generalize further. Let $g^{(n)}(x) = x^{\oplus n}$. For example,

$$g^{(3)}(x) = \frac{x^3 + 3bx - ab}{3x^2 - 3ax + b + a^2}, \qquad g^{(4)}(x) = \frac{x^4 + 6bx^2 - 4abx + b(a^2+b)}{4x^3 - 6ax^2 + 4(a^2+b)x + a^3 - 2ab}.$$

Iteration of $g^{(3)}$ is Halley's method applied to θ. The rational functions $g^{(n)}$ appear, with different notation, in [4]. In that paper, a larger family of iterative procedures is introduced; our family is that of [4] when the parameter d introduced there is 1. This will be clear from a closed form for $g^{(n)}$ in terms of the numbers u_n defined by $u_0 = 0$, $u_1 = 1$, and $u_{n+2} = au_{n+1} + bu_n$. Note that these numbers are a special case of the "generalized Fibonacci numbers" G_n introduced earlier.

We define polynomials P_n and Q_n to be the unique polynomials satisfying $x^{\oplus n} = P_n(x)/Q_n(x)$, where P_n is monic and of minimal degree.

Hence $P_0(x) = 1$ and $Q_0(x) = 0$. Letting M be the matrix in the proof of Theorem 1, note that

$$\begin{pmatrix} P_{n+1}(x) \\ Q_{n+1}(x) \end{pmatrix} = (M + xI) \begin{pmatrix} P_n(x) \\ Q_n(x) \end{pmatrix}.$$

Let $\binom{v_n}{w_n} = (-1)^n M^n \binom{1}{0}$. Then

$$\begin{pmatrix} P_n(x) \\ Q_n(x) \end{pmatrix} = (M + xI)^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \sum_{k=0}^{n} \binom{n}{k} (-1)^k x^{n-k} \begin{pmatrix} v_k \\ w_k \end{pmatrix}.$$

Since M satisfies its characteristic equation, $M^2 + aM - bI = 0$ and thus (v_n) and (w_n) satisfy the difference equation $x_{n+2} = ax_{n+1} + bx_n$. Since $v_0 = 1$, $w_0 = 0 = v_1$, and $w_1 = -1$, we may write $v_n = bu_{n-1}$ and $w_n = -u_n$. Hence,

Proposition 5. *$x^{\oplus n} = P_n(x)/Q_n(x)$, where $P_n(x) = b\sum_{k=0}^{n} \binom{n}{k} x^{n-k}(-1)^k u_{k-1}$ and $Q_n(x) = -\sum_{k=0}^{n} \binom{n}{k} x^{n-k}(-1)^k u_k$.*

Iterates of $g^{(k)}$ give an exponential subsequence of iterates of m. Let $t_{n+1}^{(k)} = g^{(k)}(t_n^{(k)})$. As for Theorem 4,

Theorem 6. *If $t_0 = c$, then $t_n^{(k)} = (a - r_0) \oplus r_{k^n}$.*

The secant method is, given two starting points s_0 and s_1, to construct a sequence defined by,

$$s_{n+1} = s_n - \frac{\theta(s_n)(s_n - s_{n-1})}{\theta(s_n) - \theta(s_{n-1})},$$

which, in our case, boils down to

$$s_{n+1} = \frac{s_n s_{n-1} + b}{s_n + s_{n-1} - a} = s_n \oplus s_{n-1}.$$

As above, we take this to be the definition of the secant method in the general case.

The Fibonacci sequence (F_n) defined at the beginning shows up in a perhaps surprising way (see [2]).

Theorem 7. $s_n = s_0^{\oplus F_{n-1}} \oplus s_1^{\oplus F_n}$.

Proof. Taking $F_{-1} = 1$ and $F_0 = 0$, the theorem clearly holds for $n = 0, 1$. Supposing it holds for all $k \leq n$, $s_{n+1} = s_n \oplus s_{n-1} = s_0^{\oplus F_{n-1}} \oplus s_1^{\oplus F_n} \oplus s_0^{\oplus F_{n-2}} \oplus s_1^{\oplus F_{n-1}} = s_0^{\oplus F_n} \oplus s_1^{\oplus F_{n+1}}$. By induction, the theorem is proven. □

To discuss convergence, we assume that K has a topology defined by a norm (or absolute value) $|\cdot|$. That is, for all $x, y \in K$,

a) $|x| = 0$ if and only if $x = 0$,

b) $|x + y| \leq |x| + |y|$, and

c) $|xy| = |x||y|$.

If $r_n \to \xi$, then, by (2) and the definition of absolute value, $\xi = m(\xi)$ and so ξ must be a zero of θ. We still assume then that the zeros (ξ_1 and ξ_2) of θ are in K.

We now define a function on G. Let $f(\infty) = 1$ and, for $x \in K - \{\xi_1, \xi_2\}$,

$$f(x) = \left|\frac{x - \xi_1}{x - \xi_2}\right|.$$

Lemma 8. *For all $x, y \in K - \{\xi_1, \xi_2\}$,*

$$f(x \oplus y) = f(x)f(y).$$

Proof. Since $x^2 - ax - b = (x - \xi_1)(x - \xi_2)$, we have $\xi_1 + \xi_2 = a$ and $\xi_1\xi_2 = -b$. Hence, $z = x \oplus y = \frac{xy - \xi_1\xi_2}{x + y - \xi_1 - \xi_2}$, which implies

$$\frac{z - \xi_1}{z - \xi_2} = \frac{xy - (x + y)\xi_1 + \xi_1^2}{xy - (x + y)\xi_2 + \xi_2^2}.$$

Taking the absolute value of both sides,

$$f(z) = \left|\frac{(x - \xi_1)(y - \xi_1)}{(x - \xi_2)(y - \xi_2)}\right| = f(x)f(y).$$

□

We remark that f is a group homomorphism from $(G, \oplus)$ into the group of positive real numbers under multiplication. If $K = \mathbb{R}$, then G is an example of a disconnected Lie group and f is two-to-one.

We are now able to say some things about the convergence of (r_n).

Theorem 9. *Let $m(z) = z \oplus c$ and m_n be the n-th iterate of m.*

a) If $|c - \xi_1| > |c - \xi_2|$, then, for $z \neq \xi_1$, $m_n(z)$ converges to ξ_2 in the norm topology.

b) If $|c - \xi_1| = |c - \xi_2|$ but $\xi_1 \neq \xi_2$, then, for all $z \notin \{\xi_1, \xi_2\}$, $m_n(z)$ does not converge.

c) If ξ is the only zero of θ, then, for all $z \neq \xi$, $m_n(z)$ converges to ξ if and only if K is Archimedean (i.e., $\lim_{n\to\infty} |\tilde{n}| = \infty$ where $\tilde{n}$ denotes the n-fold sum of the unit in K).

Proof. a) If $|c - \xi_1| > |c - \xi_2|$, then $f(c) > 1$ and, by Lemma 8 and induction, $f(m_n(z)) = f(z)f(c)^n$. Unless $f(z) = 0$ (equivalently, $z = \xi_1$), $f(m_n(z)) \to \infty$. Since f is bounded outside any neighborhood of ξ_2 (triangle inequality), the result follows.

b) If $|c - \xi_1| = |c - \xi_2|$, then $f(c) = 1$ and so, for any $z \notin \{\xi_1, \xi_2\}$, $f(m_n(z))$ is nonzero and independent of n. Since $m_n(z)$ can converge only to ξ_1 or ξ_2 (in which case $f(m_n(z))$ would converge to 0 or ∞), the result follows.

c) If ξ is the only zero of θ, then $x \oplus y = \frac{xy - \xi^2}{x + y - 2\xi}$ and a simple calculation gives

$$\frac{1}{x \oplus y - \xi} = \frac{1}{x - \xi} + \frac{1}{y - \xi}.$$

Hence, by induction,

$$\left|\frac{1}{m_n(x) - \xi} - \frac{1}{x - \xi}\right| = \frac{|\tilde{n}|}{|c - \xi|},$$

and the result follows. □

We say x_n converges to x with order k if $\frac{|x_{n+1}-x|}{|x_n-x|^k}$ converges to a nonzero constant. For example, we shall see that Newton's method converges with order two and Halley's method converges with order three.

Let $t_n^{(k)}$ be defined as above.

Theorem 10. *If θ has distinct zeros and $r_n \to \xi$, then $t_n^{(k)} \to \xi$ with order k.*

Proof. Suppose $r_n \to \xi$. We write $f(x) \asymp g(x)$ if $lim_{x\to\xi} \frac{f(x)}{g(x)}$ exists and is nonzero.

Suppose y depends on x and $y \to \xi$ as $x \to \xi$. Using the fact that $\xi^2 = a\xi + b$,

$$(x-\xi)(y-\xi) = xy + b - (x+y-a)\xi = (x+y-a)(x \oplus y - \xi).$$

Since the zeros of θ are assumed distinct, $2\xi \neq \xi_1 + \xi_2 = a$ and so $|x+y-a| \asymp 1$. Hence $|x-\xi||y-\xi| \asymp |x \oplus y - \xi|$. If $|y-\xi| \asymp |x-\xi|^k$, then $|x \oplus y - \xi| \asymp |x-\xi|^{k+1}$ and so, by induction,

$$|x^{\oplus k} - \xi| \asymp |x-\xi|^k.$$

The result follows. □

References

1. M. Filaseta, *Newton's method and simple continued fractions*, Fibonacci Quart. **24** (1986), 41-46. MR **88c:**40003
2. J. Gill and G. Miller, *Newton's method and ratios of Fibonacci numbers*, Fibonacci Quart. **19** (1981), 1–4. MR **82h:**65090
3. J.H. McCabe and G.M. Phillips, *Aitken sequences and generalized Fibonacci numbers*, Math. Comp. **45** (172) (1985), 553-558. MR **87b:**41015
4. J.B. Muskat, *Generalized Fibonacci and Lucas sequences and rootfinding methods*, Math. Comp. **61** (203) (1993), 365-372. MR **93k:**65004
5. G.M. Phillips, *Aitken sequences and Fibonacci numbers*, Amer. Math. Monthly **91** (1984), 354-357. MR **85h:**65013

Plattsburgh State University, Plattsburgh, New York 12901

E-mail address: samuel.northshield@plattsburgh.edu