ON HALLEY'S ITERATION METHOD

WALTER GANDER
Seminar für angewandte Mathematik ETH, CH-8092 Zürich, Switzerland

1. Introduction. Solving a nonlinear equation is a problem that has occupied mathematicians for many centuries and many numerical analysts of the present generation first learned how to do it from P. Henrici's textbook [7] which has now been replaced by [8].

Nearly 300 years ago in 1694 Edmund Halley published a paper [6] in Latin where he presents a new method to compute roots of a polynomial. Halley is well known for first computing the orbit of the Halley comet, which he observed in 1682 and which will soon visit us again in 1986.

Halley generalized an iteration formula due to Lagney for computing the cubic root of a number and obtained an iteration to compute roots of a polynomial. Halley did not use calculus to derive his formula. If we try to generalize and translate his derivation into modern mathematical notation we can argue as follows:

Let f be a C^2 function and x_k an approximation of a zero s of f. We replace the equation $f(x) = 0$ by $T(h) = 0$, where T is the Taylor polynomial of degree 2 of f

$$T(h) = f(x_k) + f'(x_k)h + \frac{f''(x_k)}{2}h^2 \tag{1}$$

and $h = x - x_k$. Now if x_k is a good approximation of s, then neglecting h^2 we get from

$$T(h) = 0 \tag{2}$$

the *Newton correction*

$$h = -f(x_k)/f'(x_k). \tag{3}$$

Since we neglected h^2, the denominator in (3) is wrong: equation (3) should be

$$h = -f(x_k)\Big/\left(f'(x_k) + \frac{f''(x_k)}{2}h\right). \tag{4}$$

Therefore, replacing h in the denominator of (4) by the Newton correction (3), we obtain the *Halley correction*

$$h = \frac{-f(x_k)}{f'(x_k) - \dfrac{f''(x_k)f(x_k)}{2f'(x_k)}}. \tag{5}$$

For the following discussion we rearrange Halley's iteration formula (5) to

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \cdot \frac{1}{1 - \dfrac{t(x_k)}{2}}, \tag{6a}$$

where

$$t(x) = \frac{f''(x)f(x)}{f'(x)^2}. \tag{6b}$$

In his paper [6], Halley calls the iteration formula (6a) the *rational formula*. He also considers the *Euler correction*

$$h = \frac{-f'(x_k) \pm \sqrt{f'(x_k)^2 - 2f(x_k)f''(x_k)}}{f''(x_k)} \tag{7}$$

that is obtained by solving the quadratic equation (2) and which he calls the *irrational formula*. If we rearrange (7) and compute the correction closest to zero, we get Euler's iteration

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \cdot \frac{2}{1+\sqrt{1-2t(x_k)}}, \tag{8}$$

where t is defined by (6b).

In the following we shall show how Halley's method and also many other third order iteration formulas may be derived algebraically using an elementary technique. For geometric interpretations of Halley's method we refer to [1], [2].

2. Algebraic interpretation. Halley's method (6) belongs to the class of *one point iteration methods without memory* [12]

$$x_{k+1} = F(x_k). \tag{9}$$

We consider the special case where the iteration function F has the form

$$F(x) = x - \frac{f(x)}{f'(x)} G(x). \tag{10}$$

For the following discussion we assume that s is a simple zero of f and that f and G have a sufficient number of continuous derivatives in a neighborhood of s. It is well known [11] that the iteration (9) is of second order if

$$F'(s) = 0, F''(s) \neq 0. \tag{11}$$

Let $u(x) = f(x)/f'(x)$. Then we have $u'(x) = 1 - t(x)$ where t is defined by (6b). It follows that

$$u(s) = 0 \text{ and } u'(s) = 1. \tag{12}$$

Differentiating (10) we get

$$F'(x) = 1 - u'(x)G(x) - u(x)G'(x). \tag{13}$$

And using (12) we get $F'(s) = 1 - G(s)$.

LEMMA 1. *The iteration* (9) *with F defined by* (10) *is of second order if and only if* $G(s) = 1$.

We note that for the special case $G(x) \equiv 1$ we have Newton's method for the equation $f(x) = 0$.

The iteration (9) is of third order if

$$F'(s) = F''(s) = 0, F'''(s) \neq 0. \tag{14}$$

Differentiating (13) we get

$$F''(x) = -u''(x)G(x) - 2u'(x)G'(x) - u(x)G''(x).$$

Since

$$u''(x) = -t'(x) = -\frac{f''(x)}{f'(x)} + f(x)\left(\frac{2f''(x)^2}{f'(x)^3} - \frac{f'''(x)}{f'(x)^2}\right),$$

it follows that

$$u''(s) = -\frac{f''(s)}{f'(s)} = -t'(s) \tag{15}$$

and

$$F''(s) = \frac{f''(s)}{f'(s)} G(s) - 2G'(s).$$

LEMMA 2. *The iteration* (9) *with F defined by* (10) *is of third order if and only if $G(s) = 1$ and $G'(s) = (1/2)f''(s)/f'(s)$.*

The assumptions of Lemma 2 are not helpful for choosing a function G since they need the knowledge of the zero s. However, if we choose

$$G(x) = H(t(x)) \tag{16}$$

with some function H, then we have

$$G(s) = H(0)$$

and

$$G'(s) = H'(0)t'(s).$$

Therefore the assumptions of Lemma 2 are simply $H(0) = 1$ and $H'(0) = 1/2$. We get the following theorem.

THEOREM. *Let s be a simple zero of f and H any function with $H(0) = 1$, $H'(0) = 1/2$ and $|H''(0)| < \infty$. The iteration $x_{n+1} = F(x_n)$ with*

$$F(x) = x - \frac{f(x)}{f'(x)}H(t(x))$$

where

$$t(x) = \frac{f(x)f''(x)}{f'(x)^2}$$

is of third order.

Many well-known third order iterative methods are special cases of this theorem:

(1) Halley's method

$$H(t) = \left(1 - \tfrac{1}{2}t\right)^{-1} = 1 + \frac{1}{2}t + \frac{1}{4}t^2 + \cdots .$$

(2) Euler's formula

$$H(t) = 2\left(1 + \sqrt{1 - 2t}\right)^{-1} = 1 + \frac{1}{2}t + \frac{1}{2}t^2 + \cdots .$$

(3) Hansen-Patrick family [**5**]

$$H(t) = (a + 1)\left(a + \sqrt{1 - (a + 1)t}\right)^{-1} = 1 + \frac{1}{2}t + \frac{a + 3}{8}t^2 + \cdots .$$

(4) Ostrowski's square root iteration [**10**]

$$H(t) = (1 - t)^{-0.5} = 1 + \frac{1}{2}t + \frac{3}{8}t^2 + \cdots .$$

(5) Quadratic inverse interpolation

$$H(t) = 1 + \frac{1}{2}t.$$

Not all the third order iteration methods are special cases of this theorem. It is possible to describe all methods that do not explicitly depend on s by using the Schröder iteration functions [**9**, p. 531]. One can show [**4**] that all third order methods are given by the iteration function (10) with

$$G(x) = H(t(x)) + f(x)^2 b(x),$$

where b is an arbitrary function which is bounded for $x \to s$.

References

1. G. Alefeld, On the convergence of Halley's method, this MONTHLY, 88 (1981) 530–536.

2. G. H. Brown, Jr., On Halley's variation of Newton's method, this MONTHLY, 84 (1977) 726–728.

3. M. Davis and B. Dawson, On global convergence of Halley's iteration formula, Numer. Math., 24 (1975) 133–135.

4. H. Ehrmann, Konstruktion und Durchführung von Iterationsverfahren höherer Ordnung, Arch. Rational Mech. Anal., 4 (1959) 65–88.

5. E. Hansen and M. Patrick, A family of root finding methods, Numer. Math., 27 (1977) 257–269.

6. E. Halley, Methodus Nova, Accurata & Facilis Inveniendi Radices Aequationum Quarumcumque Generaliter, Sine Praevia Reductione, Philos. Trans. Roy. Soc. London, 18 (1694) 136–148.

7. P. Henrici, Elements of Numerical Analysis, Wiley, New York, 1964.

8. ______, Essentials of Numerical Analysis, Wiley, New York, 1982.

9. ______, Applied and Computational Complex Analysis, Wiley, New York, 1974.

10. A. M. Ostrowski, Solution of Equations and Systems of Equations, Academic Press, New York, 1973.

11. J. Stoer and R. Bulirsch, Introduction to Numerical Analysis, Springer-Verlag, New York, 1980.

12. J. F. Traub, Iterative Methods for Solution of Equations, Prentice-Hall, 1964.

A CONVERSE FOR THE CAYLEY-HAMILTON THEOREM

CARMEN CHICONE, N.J. KALTON*, AND IRA J. PAPICK

Department of Mathematics, University of Missouri-Columbia, Columbia, MO 65211

The Cayley-Hamilton theorem asserts that every $n \times n$ matrix A satisfies its characteristic polynomial, $\det(\lambda I - A)$. This note deals with the problem of characterizing those polynomials for which the Cayley-Hamilton theorem holds. Informally stated our result is that the polynomials which a square matrix satisfies are precisely the multiples (in a ring of polynomials) of the characteristic polynomial.

Some notation is necessary to make a precise statement of our theorem. If X denotes the $n \times n$ matrix of indeterminants (x_{ij}), then it is apparent that $\det(X)$ is a polynomial in n^2-variables. Moreover, if $F(x_{ij})$ is any polynomial in n^2-variables with coefficients in a commutative ring R with identity, then by $F(X)$ we shall mean $F(x_{ij})$. We now restate our problem: Characterize those polynomials $F(X)$ having the property that every $n \times n$ matrix A with entries in R satisfies the polynomial $F(\lambda I - A)$. We shall call such polynomials $F(X)$ Cayley-Hamilton polynomials, and we may now state a precise converse of the Cayley-Hamilton theorem.

THEOREM 1. *Let R be an infinite (commutative) integral domain, and let $F(X)$ be a polynomial in n^2-variables with coefficients in R. Then $F(X)$ is a Cayley-Hamilton polynomial if and only if $F(X) = \det(X)G(X)$, where $G(X)$ is a polynomial in n^2-variables with coefficients in R.*

Before proceeding we make two observations. The "if" direction of the theorem is clear. For if A is an $n \times n$ matrix with entries in R, and if $F(X) = \det(X)G(X)$, then $F(\lambda I - A) = \det(\lambda I - A)G(\lambda I - A)$. Hence, by the Cayley-Hamilton theorem, A satisfies the polynomial $F(\lambda I - A)$. Secondly, the theorem is false without some assumptions on R. For example if $R = Z_2$, the field with two elements, define $F(X) = (x_{12} + x_{21})x_{11}x_{22}$. Then given

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

with $a_{ij} \in Z_2$, we see that

$$F(\lambda I - A) = (a_{12} + a_{21})(\lambda - a_{11})(\lambda - a_{22}).$$

Thus $F(\lambda I - A) \equiv 0$ unless A is either upper or lower triangular and not diagonal, and in this case

$$F(\lambda I - A) = (\lambda - a_{11})(\lambda - a_{22}) = \det(\lambda I - A).$$

*Supported in part by NSF grant MCS-8001852.