Innotec Lectures
Masaryk University Brno

NONLINEAR EQUATIONS

Prof. Walter Gander
ETH Zurich

Decenber 2, 2008

1 The Condition of Zeros of Polynomials

Jim Wilkinson, the famous British numerical analyst, discovered that the zeros of polynomial are very sensitive to small changes of the coefficients of the polynomial. Wilkinson was testing a new computer. As a test example he constructed from the zeros $x_i = 1, 2, \ldots, 20$ by expanding the product the polynomial

$$P_{20}(x) = \prod_{i=1}^{20}(x - i) = x^{20} - 210x^{19} + 20615x^{18} - \cdots + 20!. \tag{1}$$

Then he used a numerical method to compute the zeros and he was astonished to observe that his program computed some complex zeros and would not reproduce the zeros $x_i = 1, 2, \ldots, 20$. After having checked very carefully that there was no programming error and that also the hardware was working correctly he then tried to understand the results by a backward error analysis and found out that the zeros were *ill-conditioned*.

Let z be a simple zero of

$$P_n(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0.$$

Let ε be a small parameter and let g be another polynomial of degree n. Consider the zero $z(\varepsilon)$ (corresponding to z, i.e.,$z(0) = z$) of the perturbed polynomial h:

$$h(x) := P_n(x) + \varepsilon g(x).$$

We expand $z(\varepsilon)$ in a series

$$z(\varepsilon) = z + \sum_{k=1}^{\infty} p_k \varepsilon^k.$$

The coefficient $p_1 = z'(0)$ can be computed by differentiating

$$P_n(z(\varepsilon)) + \varepsilon g(z(\varepsilon)) = 0$$

with respect to ε. We obtain

$$P_n'(z(\varepsilon))z'(\varepsilon) + g(z(\varepsilon)) + \varepsilon g'(z(\varepsilon))z'(\varepsilon) = 0$$

and

$$z'(\varepsilon) = -\frac{g(z(\varepsilon))}{P_n'(z(\varepsilon)) + \varepsilon g'(z(\varepsilon))}.$$

Thus for $\varepsilon = 0$

$$z'(0) = p_1 = -\frac{g(z)}{P_n'(z)}. \tag{2}$$

We now apply Equation (2) to Wilkinson's polynomial (1). Wilkinson perturbed only the coefficient $a_{19} = -210$ by 2^{-23} (which was the machine precision for single precision on some early computers). This modification corresponds to the choices of $g(x) = x^{19}$ and $\varepsilon = -2^{-23}$ and we obtain for the zero $z_r = r$ the perturbation

$$\delta z_r \approx 2^{-23}\frac{r^{19}}{|P_{20}'(r)|} = 2^{-23}\frac{r^{19}}{(r-1)!\,(20-r)!}.$$

For $r = 16$ the perturbation is maximal and becomes

$$\delta z_{16} \approx 2^{-23}\frac{16^{19}}{15!\,4!} \approx 287.$$

Wilkinson computed the exact zeros using multiple precision and found e.g. the zeros $16.730 \pm 2.812\,i$. We can easily reconfirm this calculation using Maple:

```
restart;
Digits := 50;
p :=1:
for i from 1 by 1 to 20 do  p := p*(x-i) od:
Z := fsolve( p-2^(-23)*x^19, x, complex,maxsols=20 );
plot(map(z -> [Re(z),Im(z)],[Z]), x=0..22, style=point,symbol=circle);
```

We do not actually need to work with 50 decimal digits, even in Matlab with standard IEEE arithmetic we obtain the same results.

However, we can simulate in Maple the experience of Wilkinson by computing with 7 decimal digits. Of course it is important after expanding to represent the coefficients of the polynomial also as 7-digit numbers:

```
restart;
Digits := 7;
p :=1:
for i from 1 by 1 to 20 do  p := p*(x-i) od:
PP := expand(p);
PPP := evalf(PP);
Z := fsolve( PPP, x, complex,maxsols=20 );
plot(map(z -> [Re(z),Im(z)],[Z]), x=0..28, style=point,symbol=circle);
```

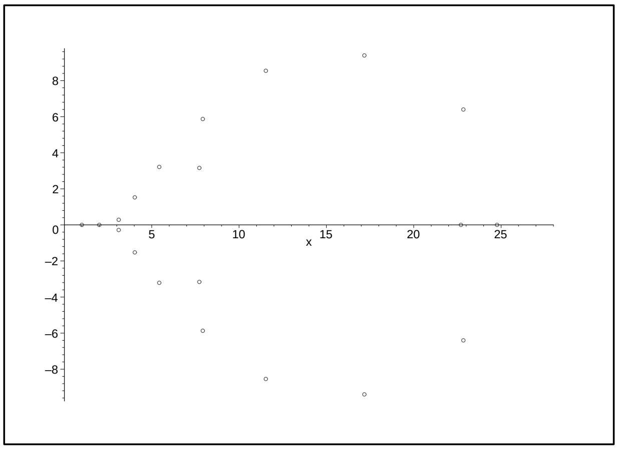

Figure 1: Zeros of Wilkinson's Polynomial computed with 7 digits.

We see from Figure 1 very well how most of the zeros become complex numbers.

In the following MATLAB script we change the second coefficient of Wilkinson's polynomial by subtracting a small perturbation: $p_2 := p_2 - \lambda p_2$ where $\lambda = 0 : 1e{-}11 : 1e{-}8$.

In MATLAB polynomials are represented by

$$P_n(x) = A(1)x^n + A(2)x^{n-1} + \cdots + A(n)x + A(n+1)$$

while in usual mathematical notation one prefers

$$P_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n.$$

In order to switch between both representations we need for the programs the transformation

$$A(i) = a_{n+1-i},\ i = 1, \ldots, n+1 \iff a_j = A(n+1-j),\ j = 0, \ldots, n.$$

Algorithm 1.1: Experiment with Wilkinson's Polynomial

```
axis([-5 25 -10 10])
hold
P = poly(1:20)
for lamb = 0: 1e-11:1e-8
   P(2) = P(2)*(1-lamb);
   Z =  roots(P);
   plot(real(Z), imag(Z),'o')
end
```

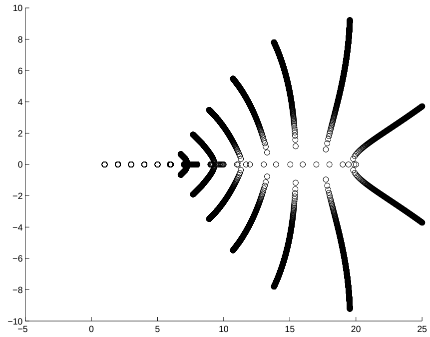

Figure 2: Roots of the perturbed Wilkinson polynomial

By computing the roots and plotting them in the complex plane we can observe in Figure 2 how the larger roots are moving fast away from the real line. We finally remark that multiple roots are always ill-conditioned. To illustrate that, assume that we change the constant coefficient of $P_n(x) = (x-1)^n$ by the machine precision ε. The perturbed polynomial becomes $(x-1)^n - \varepsilon$ and its roots are solutions of

$$(x-1)^n = \varepsilon \quad \Rightarrow \quad x(\varepsilon) = 1 + \sqrt[n]{\varepsilon}.$$

The new roots are all simple and lie on the circle of radius $\sqrt[n]{\varepsilon}$ with center 1. For $\varepsilon = 2.2204e{-16}$ and $n = 10$ we get $\sqrt[10]{\varepsilon} = 0.0272$ which shows that the multiple roots "explode" quite dramatically into n simple ones.

2 The Convergence Rates of Iterative Methods

A fixed point iteration

$$x_{k+1} = F(x_k)$$

converges if the iteration function F is a contraction. That is the case if e.g. $|F'(x)| < 1$ in a interval containing the fixed point s.

For example consider

$$f(x) = xe^x - 1 = 0. \tag{3}$$

A first fixed point iteration is obtained by rearranging and dividing Equation (3) by e^x:

$$x_{k+1} = e^{-x_k}. \tag{4}$$

With the initial guess $x_0 = 0.5$ we obtain the iterates shown in Table 1. Indeed x_k

k	x_k	k	x_k	k	x_k
0	0.5000000000	10	0.5669072129	20	0.5671424776
1	0.6065306597	11	0.5672771960	21	0.5671437514
2	0.5452392119	12	0.5670673519	22	0.5671430290
3	0.5797030949	13	0.5671863601	23	0.5671434387
4	0.5600646279	14	0.5671188643	24	0.5671432063
5	0.5711721490	15	0.5671571437	25	0.5671433381
6	0.5648629470	16	0.5671354337	26	0.5671432634
7	0.5684380476	17	0.5671477463	27	0.5671433058
8	0.5664094527	18	0.5671407633	28	0.5671432817
9	0.5675596343	19	0.5671447237	29	0.5671432953

Table 1: Iteration $x_{k+1} = \exp(-x_k)$

seems to converge to $s = 0.5671432...$

A second fixed point form is obtained from $xe^x = 1$ by adding x on both sides to get $x + xe^x = 1 + x$, factoring the left-hand side to get $x(1 + e^x) = 1 + x$, and dividing by $1 + e^x$:

$$x = F(x) = \frac{1+x}{1+e^x}. \tag{5}$$

This time the convergence is much faster — we need only three iterations to obtain a 10-digit approximation of s:

$$\begin{aligned} x_0 &= 0.5000000000 \\ x_1 &= 0.5663110032 \\ x_2 &= 0.5671431650 \\ x_3 &= 0.5671432904, \end{aligned}$$

In this section we would like to analyze the speed of convergence. Geometrically we observe that convergence is faster if $|F'(s)|$ is smaller.

Definition 2.1 *The error in iteration step k is $e_k = x_k - s$.*

Subtracting the equation $s = F(s)$ from $x_{k+1} = F(x_k)$ and expanding in a Taylor series we get

$$x_{k+1} - s = F(x_k) - F(s) = F'(s)(x_k - s) + \frac{F''(s)}{2!}(x_k - s)^2 + \cdots$$

or expressed in terms of the error:

$$e_{k+1} = F'(s)e_k + \frac{F''(s)}{2!}e_k^2 + \frac{F'''(s)}{3!}e_k^3 + \cdots \tag{6}$$

If $F'(s) \neq 0$ then we conclude from Equation (6) that

$$\lim_{k\to\infty} \frac{e_{k+1}}{e_k} = F'(s).$$

This means that asymptotically $e_{k+1} \sim F'(s)e_k$. *Thus for large k the error is reduced in each iteration step by the factor $|F'(s)|$.* This is called *linear convergence* because the new error is a linear function of the previous one. Similarly we speak of *quadratic convergence* if $F'(s) = 0$ but $F''(s) \neq 0$, because $e_{k+1} \sim (F''(s)/2)\, e_k^2$. More generally we define:

Definition 2.2 *The rate of convergence of $x_{k+1} = F(x_k)$ is*

linear: *if $F'(s) \neq 0$ and $|F'(s)| < 1$,*

quadratic: *if $F'(s) = 0$ and $F''(s) \neq 0$,*

cubic: *if $F'(s) = 0$ and $F''(s) = 0$, but $F'''(s) \neq 0$,*

of order m: *if $F'(s) = F''(s) = \cdots = F^{(m-1)}(s) = 0$, but $F^{(m)}(s) \neq 0$.*

Example 2.1 1. *Consider Iteration (4): $x_{k+1} = F(x_k) = e^{-x_k}$. The fixed point is $s = 0.5671432904$. $F'(s) = -F(s) = -s = -0.567143290$. Because $0 < |F'(s)| < 1$ we have linear convergence. With linear convergence the number of correct digits grows linearly. For $|F'(s)| = 0.5671432904$ the error is roughly halved in each step. So in order to obtain a new decimal digit one has to perform p iteration steps where $0.5671432904^p = 0.1$. This gives us $p = 4.01$. Thus after about 4 iterations we obtain another decimal digit as can also be verified by looking at Table 1.*

2. *If we want to solve Kepler's equation: $E - e\sin E = \frac{2\pi}{T}t$ for E an obvious iteration function is*

$$E = F(E) = \frac{2\pi}{T}t + e\sin E.$$

Because $|F'(E)| = e|\cos E| < 1$ the fixed point iteration generates always a linearly convergent sequence.

3. *Iteration (5) $x_{k+1} = F(x_k)$ with $F(x) = \frac{1+x}{1+e^x}$ converges quadratically because*

$$F'(x) = \frac{1 - xe^x}{(1+e^x)^2} = -\frac{f(x)}{(1+e^x)^2}$$

and thus $F'(s) = 0$. Furthermore one can check that $F''(s) \neq 0$. With quadratic convergence the number of correct digits doubles at each step asymptotically. If $e_k = 10^{-3}$ then $e_{k+1} \approx e_k^2 = 10^{-6}$, and thus we have 6 correct digits in the next step.

The doubling of digits can be seen well when computing with Maple *with extended precision (we separated the correct digits with a *):*

```
Digits := 59;
x := 0.5;
for i from 1 by 1 to 5 do x:= (1+x)/(1+exp(x)); od;

x:= .5
x:= .56*63110031972181530416491513817372818700809520366347554108
x:= .567143*16503486221278651209665969636651343135081870855674 77
x:= .56714329040978*10286995766494153472061705578660439731056279
x:= .567143290409783872999968662 2*0889167130372661165136497337 66
x:= .56714329040978387299996866221035554975381578718651250813513*
```

3 Construction of One Point Iteration Methods

In this section we will show how to transform the equation $f(x) = 0$ *systematically* to a fixed point form $x = F(x)$ with a high convergence rate. We can construct these methods algebraically and geometrically.

3.1 Geometric Construction

The basic idea here is to approximate the function f in the neighborhood of a zero s by a simpler function h. The equation $f(x) = 0$ is then replaced by $h(x) = 0$ which should be easy to solve and one hopes that the zero of h is also an approximation of the zero of f. In general h will be so simple that we can solve $h(x) = 0$ analytically. If one has also to solve $h(x) = 0$ iteratively then we obtain a *method with inner and outer iterations.*

Let x_k be an approximation of a zero s of f. We choose for h a linear function that has for $x = x_k$ the same function value and derivative as f (i.e. the Taylor polynomial of degree one):

$$h(x) = f(x_k) + f'(x_k)(x - x_k) \tag{7}$$

The equation $h(x) = 0$ can be solved analytically:

$$h(x) = 0 \iff x = x_k - \frac{f(x_k)}{f'(x_k)}.$$

Now the zero of h may indeed be better approximation of s, as we can see from Figure 3. We have obtained the iteration

$$x_{k+1} = F(x_k) = x_k - \frac{f(x_k)}{f'(x_k)}, \tag{8}$$

which is called *Newton's Iteration* or often also *Newton-Raphson Iteration.*

Example 3.1 *We return to Kepler's Equation:*

$$f(E) = E - 0.8 \sin E - \frac{2\pi}{10} = 0.$$

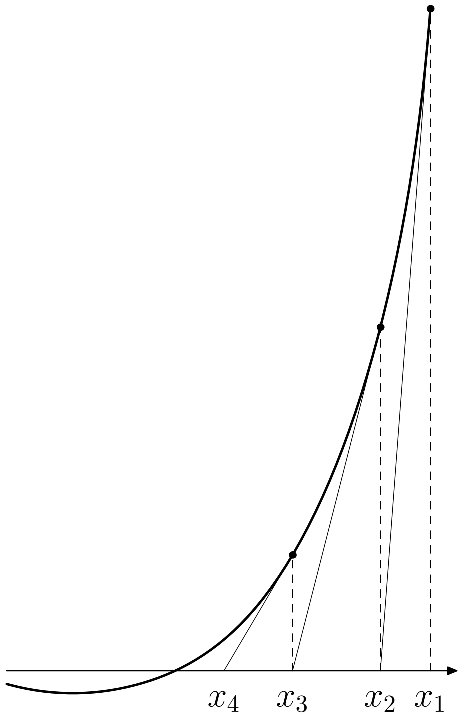

Figure 3: Geometric Derivation of the Newton-Raphson Method

Applying Newton's method, we obtain the iteration

$$E_{k+1} = E_k - \frac{f(E_k)}{f'(E_k)} = E_k - \frac{E_k - 0.8 \sin E_k - \frac{2\pi}{10}}{1 - 0.8 \cos E_K}.$$

Starting this iteration with $E = 1$ we obtain the values

```
1.53102771971995
1.42429107823439
1.41914768835385
1.41913578389432
1.41913578383058
1.41913578383058
```

which clearly indicate the quadratic convergence.

Instead of the Taylor polynomial of degree one we now propose to consider the function

$$h(x) = \frac{a}{x+b} + c. \tag{9}$$

We would like to determine the parameters a, b and c such that again h has the same function value and the same first and second derivatives as f at x_k:

$$\begin{aligned} f(x_k) &= h(x_k) &= \frac{a}{x_k + b} + c \\ f'(x_k) &= h'(x_k) &= -\frac{a}{(x_k+b)^2} \\ f''(x_k) &= h''(x_k) &= \frac{2a}{(x_k+b)^3} \end{aligned} \tag{10}$$

We use MAPLE to define and solve the nonlinear system (10):

```
> h:= x -> a/(x+b) + c;
> solve({h(xk)=f, D(h)(xk)=fs, D(D(h))(xk)=fss, h(xn)=0}, {a,b,c,xn});
```

$$xn = \frac{fss\, xk\, f - 2\, fs^2 xk + 2\, f fs}{fss\, f - 2\, fs^2}$$

Rearranging, we obtain *Halley's Iteration*:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \frac{1}{1 - \frac{1}{2}\frac{f(x_k)f''(x_k)}{f'(x_k)^2}}. \tag{11}$$

3.2 Algebraic Construction

If $f(x) = 0$ then adding x on both sides gives $x = F(x) = x + f(x)$. However, we cannot expect for arbitrary f that $|F'| = |1 + f'|$ is smaller than 1. But we can generalize this idea by premultiplying $f(x) = 0$ first with some function $h(x)$ which still has to be determined. Thus we consider the fixed point form

$$x = F(x) = x + h(x)f(x). \tag{12}$$

Let us try to chose $h(x)$ so to make $|F'(s)| < 1$ or even better $F'(s) = 0$. We have

$$F'(x) = 1 + h'(x)f(x) + h(x)f'(x).$$

For quadratic convergence we need

$$F'(s) = 1 + h(s)f'(s) = 0. \tag{13}$$

To solve Equation (13) for h the condition

$$f'(s) \neq 0 \tag{14}$$

must hold and then

$$h(s) = -\frac{1}{f'(s)}. \tag{15}$$

Condition (14) signifies that s must be a simple zero of f. Furthermore the only condition to chose h is that it must have the function value $-1/f'(s)$ for $x = s$. Since s is unknown the easiest choice for h is

$$h(x) = -\frac{1}{f'(x)}.$$

This choice leads to the iteration

$$x = F(x) = x - \frac{f(x)}{f'(x)}, \tag{16}$$

which is again *Newton's Iteration.* By the algebraic derivation we have proved that *for simple zeros Newton's Iteration generates a quadratically convergent sequence.*

Every fixed point iteration $x = F(x)$ can be regarded as a Newton iteration for some function g. In order to determine g we need to solve the differential equation:

$$x - \frac{g(x)}{g'(x)} = F(x) \tag{17}$$

We conclude from Equation (17)

$$\frac{g'(x)}{g(x)} = \frac{1}{x - F(x)}$$

and by integration

$$\ln|g(x)| = \int \frac{dx}{x - F(x)}$$

therefore

$$|g(x)| = \exp\left(\int \frac{dx}{x - F(x)}\right). \tag{18}$$

Example 3.2 *We consider the second fixed point form (5) of Equation (3). Here we have*

$$F(x) = \frac{1 + x}{1 + e^x}$$

and

$$\int \frac{dx}{x - F(x)} = \int \frac{1 + e^x}{xe^x - 1} dx.$$

Simplifying by canceling with e^x gives

$$= \int \frac{e^{-x} + 1}{x - e^{-x}} dx = \ln|x - e^{-x}|.$$

Thus the fixed point fixed point form

$$x = \frac{1 + x}{1 + e^x}$$

is Newton's iteration for $f(x) = x - e^{-x} = 0$.

We can interpret Halley's Iteration as an "improved" Newton's iteration, since we can write

$$x = F(x) = x - \frac{f(x)}{f'(x)} G(t(x)), \tag{19}$$

with

$$G(t) = \frac{1}{1 - \frac{1}{2}t}, \tag{20}$$

and

$$t(x) = \frac{f(x) f''(x)}{f'(x)^2}. \tag{21}$$

For $G(t) \approx 1 \Leftarrow t(x) \approx 0 \Leftarrow f''(x) \approx 0$, i.e., if f has small curvature, then both methods are similar.

Which function g of Newton's iteration corresponds to Halley's iteration? Surprisingly the differential equation

$$x - \frac{g(x)}{g'(x)} = F(x) = x + \frac{2f'(x)f(x)}{f(x)f''(x) - 2f'(x)^2}$$

has a simple solution. We have

$$\frac{g'(x)}{g(x)} = -\frac{1}{2}\frac{f''(x)}{f'(x)} + \frac{f'(x)}{f(x)},$$

and integration yields

$$\ln|g(x)| = -\frac{1}{2}\ln|f'(x)| + \ln|f(x)| = \ln\left|\frac{f(x)}{\sqrt{f'(x)}}\right|.$$

We have proved the following theorem.

Theorem 3.1 *Halley's iteration for $f(x) = 0$ is Newton's iteration for*

$$\frac{f(x)}{\sqrt{f'(x)}} = 0.$$

Let us now analyze Halley-like iteration forms

$$x = F(x) = x - \frac{f(x)}{f'(x)}H(x) \tag{22}$$

and find conditions for the function $H(x)$ so that the iteration yields sequences that converge quadratically or even cubically to a simple zero of f. With the abbreviation

$$u(x) := \frac{f(x)}{f'(x)},$$

we obtain for the derivatives

$$\begin{aligned} F &= x - uH \\ F' &= 1 - u'H - uH' \\ F'' &= -u''H - 2u'H' - uH'' \end{aligned}$$

and

$$\begin{aligned} u' &= 1 - \frac{ff''}{f'^2} \\ u'' &= -\frac{f''}{f'} + 2\frac{ff''^2}{f'^3} - \frac{ff'''}{f'^2}. \end{aligned}$$

Because $f(s) = 0$ we have

$$u(s) = 0, \quad u'(s) = 1, \quad u''(s) = -\frac{f''(s)}{f'(s)}. \tag{23}$$

It follows for the derivatives of F:

$$F'(s) = 1 - H(s) \tag{24}$$

$$F''(s) = \frac{f''(s)}{f'(s)} H(s) - 2H'(s). \tag{25}$$

We conclude from the Equations (24) and (25) that for *quadratic convergence*

$$H(s) = 1$$

must hold and for cubic convergence we need in addition

$$H'(s) = \frac{1}{2} \frac{f''(s)}{f'(s)}.$$

However, s is unknown and therefore we need to chose H as a function of f and its derivatives:

$$H(x) = G(f(x), f'(x), \ldots).$$

For example, if we choose $H(x) = 1 + f(x)$ then because of $H(s) = 1$ we obtain an iteration with quadratic convergence:

$$x = x - \frac{f(x)}{f'(x)}(1 + f(x)).$$

If we choose

$$H(x) = G(t(x)), \tag{26}$$

with

$$t(x) = \frac{f(x) f''(x)}{f'(x)^2}, \tag{27}$$

then because of

$$t(x) = 1 - u'(x)$$

we get

$$H'(x) = G'(t(x)) t'(x) = -G'(t(x)) u''(x).$$

There follows

$$\begin{aligned} H(s) &= G(0) \\ H'(s) &= -G'(0) u''(s) = G'(0) \frac{f''(s)}{f'(s)}. \end{aligned} \tag{28}$$

We thus have proved a theorem (see [1]) :

Theorem 3.2 *Let s be a simple zero of f and G any function with $G(0) = 1$, $G'(0) = \frac{1}{2}$ and $|G''(0)| < \infty$. Then the sequence generated by the fixed point form*

$$x = F(x) = x - \frac{f(x)}{f'(x)}\, G\left(\frac{f(x) f''(x)}{f'(x)^2}\right)$$

converges cubically to the the simple zero s of f.

Example 3.3 *Many well known iteration methods are special cases of Theorem 3.2. As we can see from the Taylor expansions of $G(t)$ they converge cubically:*

1. *Halley's method*

$$G(t) = \frac{1}{1-\frac{1}{2}t} = 1 + \frac{1}{2}t + \frac{1}{4}t^2 + \frac{1}{8}t^3 + \cdots$$

2. *Euler's Iteration*

$$G(t) = \frac{2}{1+\sqrt{1-2t}} = 1 + \frac{1}{2}t + \frac{1}{2}t^2 + \frac{5}{8}t^3 + \cdots$$

3. *Quadratic inverse interpolation*

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

4 Multiple Zeros

The quadratic convergence of Newton's method was based on the assumption that s is a simple zero and that therefore $f'(s) \neq 0$. We will now investigate the convergence for multiple zeros.

```
 F := x -> x - f(x)/D(f)(x);
dF := D(F)(x);
```

$$F := x \to x - \frac{\mathrm{f}(x)}{\mathrm{D}(f)(x)} \tag{29}$$

$$dF := \frac{\mathrm{f}(x)\, D^{(2)}(f)(x)}{\mathrm{D}(f)(x)^2} \tag{30}$$

Let us now assume that $f(x)$ has a zero of multiplicity n at $x = s$. We therefore define $f(x)$ to be

```
> f := x -> (x-s)^n * g(x);
```

$$f := x \to (x-s)^n\, g(x)$$

where $g(s) \neq 0$. We inspect the first derivative of $F(x)$. If $F'(s) \neq 0$ then the iteration converges only linearly.

```
> dF;
```

$$(x-s)^n\, g(x) \Bigg(\frac{(x-s)^n\, n^2\, g(x)}{(x-s)^2} - \frac{(x-s)^n\, n\, g(x)}{(x-s)^2} + 2\, \frac{(x-s)^n\, n\, \mathrm{D}(g)(x)}{x-s} + (x-s)^n\, (\mathrm{D}^{(2)})(g)(x) \Bigg) \Bigg/ \Bigg(\frac{(x-s)^n\, n\, g(x)}{x-s} + (x-s)^n\, \mathrm{D}(g)(x) \Bigg)^2$$

Taking the limit of the above expression for $x \to s$ we obtain:

```
> limit(%, x=s);
```

$$\frac{n-1}{n}$$

We have proved that *Newton's iteration converges only linearly with factor* $(n-1)/n$ *if* $f(x)$ *has a zero of multiplicity* n. Thus, e.g., for a double root convergence is linear with factor 1/2.

A possible remedy therefore for reestablishing quadratic convergence is to take "double steps" for a double root

$$x_{k+1} = x_k - 2\frac{f(x_k)}{f'(x_k)}$$

or for a root of multiplicity n

$$x_{k+1} = x_k - n\frac{f(x_k)}{f'(x_k)}. \tag{31}$$

However one seldom knows the multiplicity in advance. So we should try to estimate n. An old proposal by Schröder of 1870 is to use

$$n = \frac{f'^2}{f'^2 - ff''}. \tag{32}$$

The resulting iteration looks almost like Halley's iteration (the factor 1/2 is missing):

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \frac{1}{1 - \frac{f(x_k)f''(x_k)}{f'(x_k)^2}}. \tag{33}$$

Interpreting again Iteration (33) as Newton iteration by solving the differential equation

$$\frac{g(x)}{g'(x)} = \frac{f(x)}{f'(x)} \frac{1}{1 - \frac{f(x)f''(x)}{f'(x)^2}}$$

we obtain

$$\frac{g'(x)}{g(x)} = \frac{f' - \frac{ff''}{f'}}{f} = \frac{f'}{f} - \frac{f''}{f'}$$

and

$$\ln|g| = \ln|f| - \ln|f'| \quad \Longrightarrow \quad g(x) = \frac{f(x)}{f'(x)}.$$

Thus Schröder's method (32) *is equivalent to applying Newton's method on* $f(x)/f'(x) = 0$*, thus canceling multiple roots.*

References

[1] Walter Gander, *On Halley's Iteration Method*, The American Mathematical Monthly, Vol. 92, No. 2, February 1985.