A Lecture on

Real Root-Finding

Prof. W. Kahan
Univ. of Calif. @ Berkeley

6 April 1998

Extracted from "Lecture Notes on Real Root-Finding"

http://www.cs.berkeley.edu/~wkahan/Math128/RealRoots.ps and … .pdf
http://www.cs.berkeley.edu/~wkahan/Math128/SolveKey.pdf

Given a real equation $f(z) = 0$, solve it for its real *root* z, a *zero* of function f,

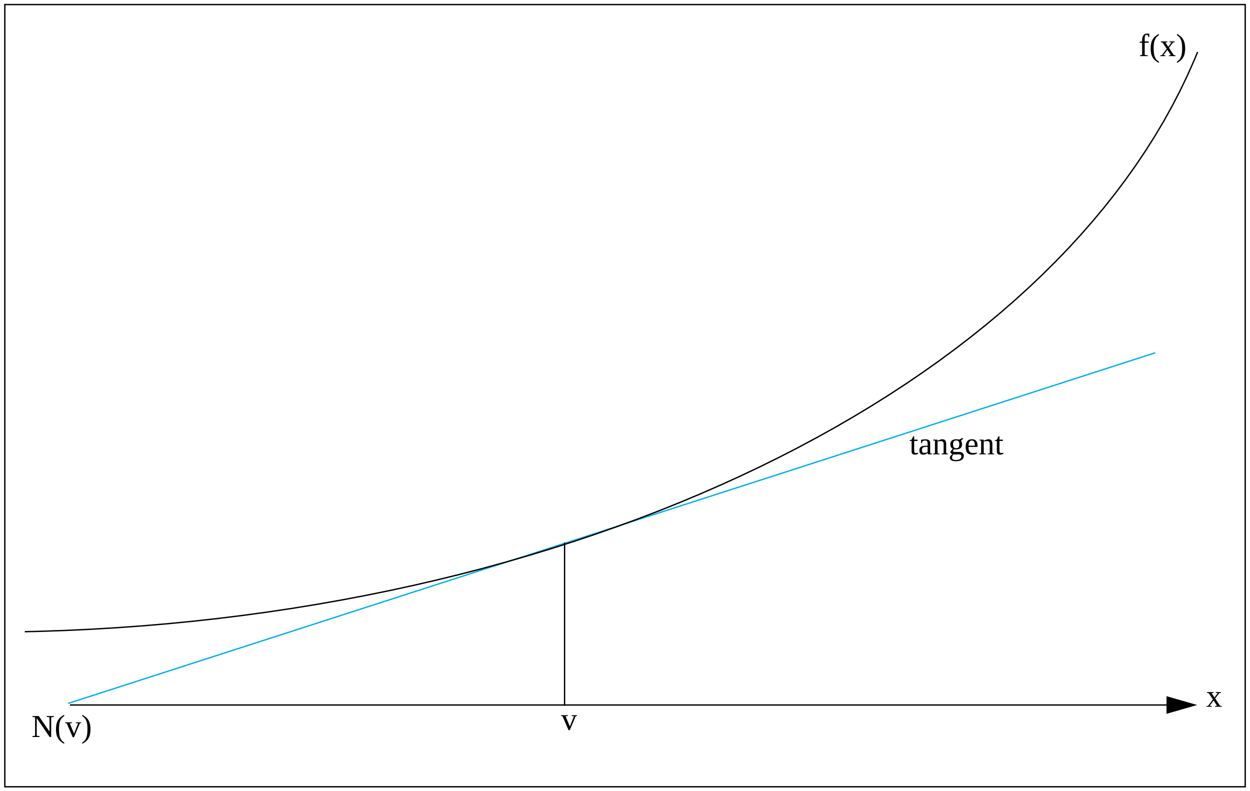
typically by *iteration* : $x_{n+1} := U(x_n) \longrightarrow z = U(z)$, a *fixed point*, as $n \longrightarrow \infty$.

How should the Iterating Function U be constructed?

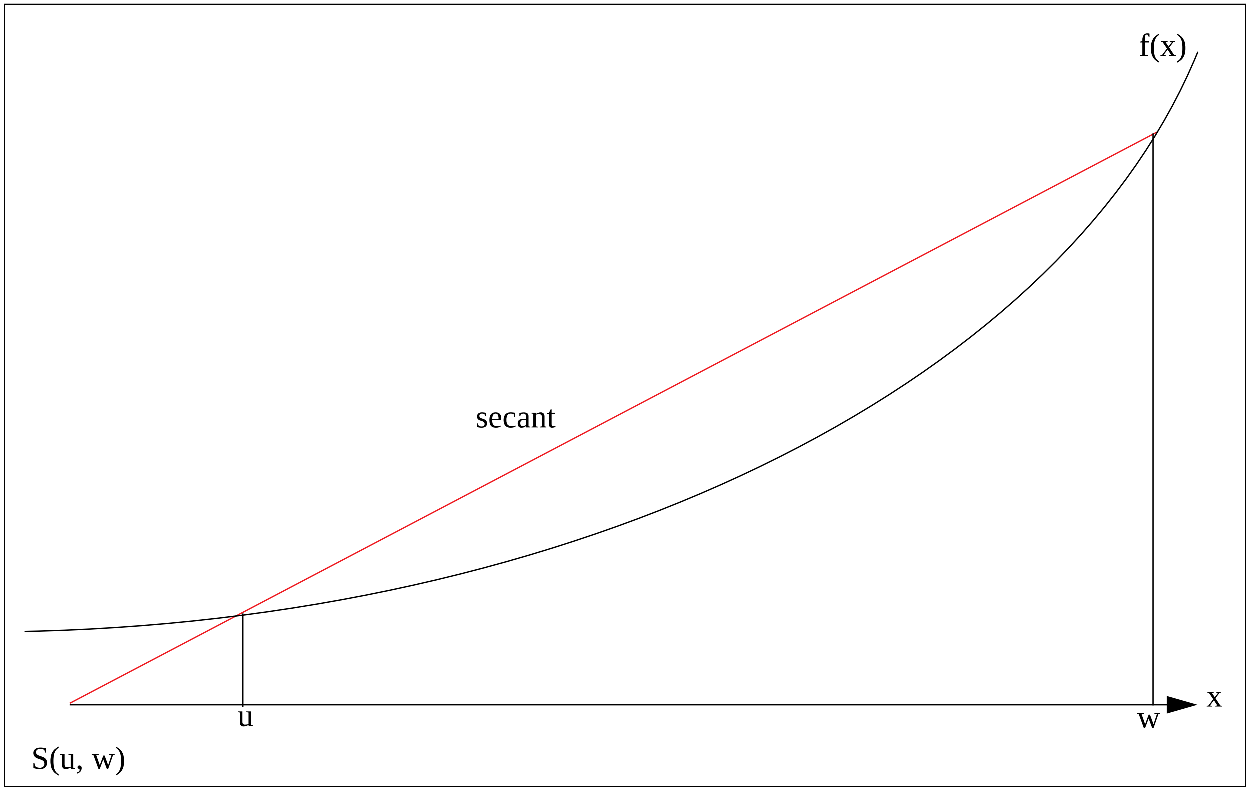
Common iterations:

Newton's: $x_{n+1} := N(x_n)$ where $N(x) = x - f(x)/f'(x)$.

Secant: $x_{n+1} := S(x_n, x_{n-1})$ where $S(x, y) := x - f(x)(x-y)/(f(x) - f(y))$.



Newton's: $x_{n+1} := N(x_n)$ where $N(v) = v - f(v)/f'(v)$.



Secant: $x_{n+1} := S(x_n, x_{n-1})$ where $S(u, w) := u - f(u)(u-w)/(f(u) - f(w))$.

Before a real root z of an equation $f(z) = 0$ can be found, six questions demand attention:

«1» Which equation?
Infinitely many equations, some far easier to solve than others, have the same root z .

«2» What method?
Usually an iterative method must be chosen; there are infinitely many of them too.

«3» Where should the search for a root begin?
A global theory of the iteration's convergence helps compensate for a vague guess at z .

«4» How fast can the iteration be expected to converge?
A local theory helps here. Convergence much slower than expected is ominous.

«5» When should iteration be stopped?
Error-analysis helps here. And the possibility that no z exists may have to be faced.

«6» How will the root's accuracy be assessed?
Error-analysis is indispensable here, and it can be done in more than one way.

These questions lead to others:

Questions:

Is some simple Combinatorial (Homeomorphically invariant) condition both Necessary and Sufficient for convergence of $x \longrightarrow U(x)$? (Yes; §5)

Is that condition relevant to the design of root-finding software? (Yes; §6)

Do other iterations $x \longrightarrow U(x)$ besides Newton's exist? (Not really; §3)

Must there be a neighborhood of z within which Newton's iteration converges if $f'(x)$ and $x - f(x)/f'(x)$ are both continuous? (Maybe Not; §7)

Do useful conditions less restrictive than Convexity suffice Globally for the convergence of Newton's and Secant iteration? (Yes; §8)

Why are these less restrictive conditions not Projective Invariants, as are Convexity and the convergence of Newton's and Secant iterations? (I don’t know; §A3)

Is slow convergence to a multiple root worth accelerating? (Probably not; §7)

Can slow convergence from afar be accelerated with no risk of overshooting and thus losing the desired root? (In certain common cases, Yes; §10)

When should iteration be stopped? (*Not* for the reasons usually cited; §6)

Which of Newton's and Secant iterations converges faster? (Depends; §7)

Which of Newton's and Secant iterations converges from a wider range of initial guesses at z ? (Secant, unless z has even multiplicity; §9)

Why Use Tangents When Secants Will Do? ([1979'], [SOLVE] key on H-P Calculators.)

Do other iterations $x \longrightarrow U(x)$ besides Newton's exist? (Not really; §3)

Thesis 3.1: Newton's Iteration is Ubiquitous

Suppose that U is differentiable throughout some neighborhood Ω of a root z of the given equation $f(z) = 0$. If the iteration $x_{n+1} := U(x_n)$ converges in Ω to z from every starting point x_0 in Ω , then this iteration is Newton's iteration applied to some equation $g(z) = 0$ equivalent on Ω to the given equation; in other words, $U(x) = x - g(x)/g'(x)$, and $g(x) \longrightarrow 0$ in Ω only as $x \longrightarrow z$.

Defense: $g(x) = \pm\exp(\int dx/(x - U(x)))$ with a "constant" of integration that may jump when x passes from one side of z to the other, reflecting the fact that U is unchanged when $g(x)$ is replaced by, say, $-3g(x)$ for all x on one side of z .

Is some simple Combinatorial (Homeomorphically invariant) condition both Necessary and Sufficient for convergence of $x \longrightarrow U(x)$? (Yes; §5)

Theorem 5.1: Sharkovsky's No-Swap Theorem [1964, 5]

Suppose U maps a closed interval Ω continuously into itself; then the iteration $x_{n+1} := U(x_n)$ converges to some fixed–point $z = U(z)$ from every x_0 in Ω if and only if these four conditions, each of which implies all the others, hold throughout Ω :

No-Swap Condition: U exchanges no two distinct points of Ω ; in other words, if $U(U(x)) = x$ in Ω then $U(x) = x$ too.

No Separation Condition: No x in Ω can lie strictly between $U(x)$ and $U(U(x))$; in other words, if $(x - U(x))(x - U(U(x))) \le 0$ then $U(x) = x$.

No Crossover Condition: If $U(x) \le y \le x \le U(y)$ in Ω then $U(x) = y = x = U(y)$.

One-Sided Condition: If $x_1 := U(x_0) \ne x_0$ in Ω then all subsequent iterates $x_{n+1} := U(x_n)$ also differ from x_0 and lie on the same side of it as does x_1 .

This last condition's violation can be detected by a program and thus serves to initiate activity designed to prevent a raw iteration's non-convergence.

Example 5.4: Suppose f is a rational function with simple real interlacing zeros and poles, one of them a pole at ∞. An instance is $f(x) := p(x)/p'(x)$ where $p(x)$ is a polynomial all of whose zeros are real. Another instance is $f(x) := \det(xI - A)/\det(xI - \hat{A}) = \prod_i (x - z_i)/\prod_j (x - \hat{o}_j)$ in which A is an hermitian matrix, $\hat{A}$ is obtained from it by striking off its last row and column, and the I 's are identity matrices; the zeros z_i lie among the eigenvalues of A, and the poles $\hat{o}_j$ are the distinct eigenvalues of $\hat{A}$ that are not also eigenvalues of A. That they interlace, i.e.,

$$z_0 < \hat{o}_1 < z_1 < \hat{o}_2 < z_2 < \ldots < \hat{o}_K < z_K \ ,$$

is a well–known theorem attributed to Cauchy. We do not know the zeros z_i but, like Y. Saad [1974], propose to compute them by running Newton's iteration $x_{n+1} := x_n - f(x_n)/f'(x_n)$. Does it converge? If so, to what? These are thorny questions, considering how spiky is the graph of f, and yet Newton's iteration can be proved to converge to some zero z_i from every real starting value except a countable nowhere–dense set of starting values from which the iteration must converge accidentally (after finitely many steps) to a pole $\hat{o}_j$.

Do useful conditions less restrictive than Convexity suffice Globally for the convergence of Newton's and Secant iteration? (Yes; §8)

Corollary 8.3: A Weak Convexity Condition

Suppose $f = g-h$ is a differentiable difference between two convex functions, one non–decreasing and the other non-increasing, throughout a closed interval Ω . Then Newton's iteration $x_{n+1} := x_n - f(x_n)/f'(x_n)$, started from any x_0 in Ω , either converges in Ω to the zero z of f or leaves Ω ; the iteration cannot meander in Ω endlessly.

Application: Calculation of Internal Rate of Return in Financial Calculators.

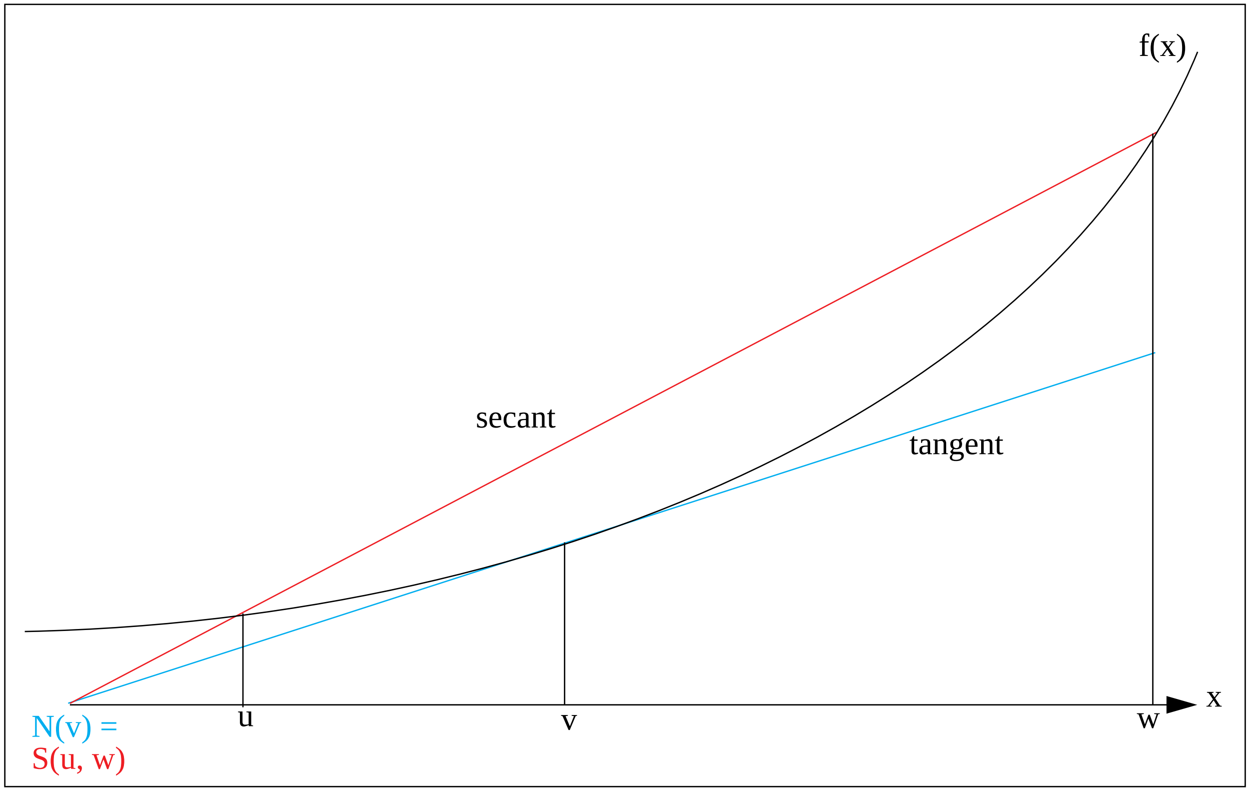
Which of Newton's and Secant iterations converges from a wider range of initial guesses at z ?
(Secant, unless z has even multiplicity; §9)

Theorem 9.2: Suppose f' and N are continuous throughout a closed finite interval Ω strictly inside which f does not vanish without reversing sign there too. If Newton's iteration converges in Ω from every initial x_0 in Ω , then it converges to the sole zero z of f in Ω , and Secant iteration also converges in Ω to z from every two starting points x_0 and x_1 in Ω .

Alas, the proof is very long.

Lemma 9.1: An Intermediate Value

If $S(u, w) := u - f(u)(u - w)/(f(u) - f(w))$ does not lie between u and w, *i.e.* if $f(u)f(w) > 0$, and if $f'(x)$ is finite throughout $u \le x \le w$, then at some v strictly between u and w either $N(v) := v - f(v)/f'(v) = S(u, w)$ or $f(v) = f'(v) = 0$.



Proof: Let $Ø(x) := f(x)/(s-x)$ where $s = S(u,v)$. Since s does not lie between u and $w > u$, $Ø(x)$ and $Ø'(x)$ are finite throughout $u \le x \le w$. And $Ø(u) = Ø(w)$ because of how s was defined, so Rolle's theorem implies $Ø'(v) = 0$ at some v strictly between u and w. Then $Ø'(v) = f'(v)/(s-v) + f(v)/(s-v)^2 = 0$ implies that this v is where either $N(v) = s$ or $f(v) = f'(v) = 0$.

Is slow convergence to a multiple root worth accelerating? (Probably not; §7)

Theorem 7.5: Suppose $|f'(x)|$ increases as x moves away from z through some neighborhood Ω on one side of a zero z of f. Then $0 < (N(x) - z)/(x-z) < 1$ and so Newton's iteration converges monotonically to z from every initial x_0 in Ω. Similarly $0 < (S(x,y) - z)/(x-z) < 1$ for all x and y in Ω and so Secant iteration converges monotonically to z from every initial x_0 and x_1 in Ω.

Theorem 7.6: Under the convexity hypothesis of Theorem 7.5, the iterates x_n may converge to z arbitrarily slowly, though monotonically; but $f(x_n)$ tends monotonically to 0 at least so fast that $\sum_n (2^n f(x_n))^2 \le f(x_0)^2 (x_0 - z)/(x_0 - x_1)$.

There is a special but common case that can be accelerated modestly without overshoot. Define

$N(x) := x - f(x)/f'(x)$ (Newton's iteration function) and

$W(x) := x - 2 f(x)/f'(x)$ (Doubled-Newton's iteration function) .

This $W(x)$ can be iterated with no harm from overshoot in the following circumstances:

Theorem 10.1: Suppose that $f'(y) = 0 \ge f(y)$ at the left-hand end of a finite interval $y \le x \le x_0$ throughout which f'' is a positive nondecreasing function; also assume $f(x_0) > 0$. Then, in that interval, ...

1) The equation $f(z) = 0$ has just one root $z \ge y$, and $W(x) < N(x)$ when $x > z$..
2) $N(x) \ge z$ for all $x > y$, and then $W(x) > y$ unless $W(x) = y = z$.
3) If $x > z$ then $N(W(x)) \le N(x)$, with equality only when $f''' \equiv 0$.