1.1 Arith Using Computer

- **IEEE double** 64-bit floating # $\underline{sc_{10}\cdots c_0 f_{-1}\cdots f_{-52}} = (-1)^s 2^{c-1023}(1+f)$, $c = (c_{10}\cdots c_0)_2$ and $f = (0.f_{-1}\cdots f_{-52})_2$.
- **Underflow** $< 2^{-1022}(1+0)$,
- **Overflow** $> 2^{1023}(2-2^{-52})$,
- **Actual error** $p_{true} - p_{app}$,
- **Absolute error** $|p_{true} - p_{app}|$,
- **Relative error** $|p_{true} - p_{app}|/|p_{true}|$,
- p^* approx. p to ***t* significant digits** if
$$\frac{|p-p^*|}{|p|} \le 5\times 10^{-t}$$
- k-digit chop: rel err $\le 10^{-k+1}$,
- k-digit round: rel err $\le 10^{-k+1}$
- **Finite digit arith** $x \circledast y = fl(fl(x) \star fl(y))$
- **Err growth** $\epsilon_n = O(n)\epsilon_0$ linear; $\epsilon_n = O(C^n)\epsilon_0$ $(C>1)$ exponential
- $|\alpha_n - \alpha| = O(\beta_n) \implies \alpha_n \to \alpha$ with **rate of conv** $O(\beta_n)$

2.1 Err Analysis

- Order of convergence: $p_n \to p$ **of order α with asympt err const λ** if $(\alpha, \lambda > 0)$
$$\frac{|p_{n+1}-p|}{|p_n - p|^\alpha} \to \lambda$$
- $\alpha = 1$, $\lambda = 1$: sublinearly conv,
$\alpha = 1$, $\lambda \in (0,1)$: linearly conv,
$\alpha = 1$, $\lambda = 0$: superlinearly conv,
$\alpha = 2$: quad'ly conv
- $g \in C[a,b]$, $g \in C^1(a,b)$, $|g'(x)| \le {}^\exists k < 1$ on (a,b), $g(p) = p$, $g'(p) \ne 0$, $p_0 \ne p \in [a,b]$ then $p_n = g(p_{n-1}) \to p$ linearly.
- $g \in C^2$, $g(p) = p$, $g'(p) = 0$, $|g''(x)| < M$, then $\exists\delta > 0$ so that $p_0 \in [p-\delta, p+\delta]$ implies $p_n = g(p_{n-1}) \to p$ at least quad'ly with
$$|p_{n+1} - p| < \frac{M}{2}|p_n - p|^2.$$

2.2 Bisection Method

- Stopping crit: $|p_n - p_{n-1}|/|p_n| < \epsilon$, $p_n \ne 0$
- Abs err: $|p_n - p| \le (b-a)\,2^{-n}$.
- $a_n + (b_n - a_n)/2$ is computationally better than $(a_n + b_n)/2$ when $b_n - a_n$ is near the max precision
- $\operatorname{sgn} f(a) \cdot \operatorname{sgn} f(b) < 0$ is better than $f(a)f(b) < 0$ due to over/underflow

2.3 Fixed Point Iteration

- $g \in C[a,b], [a,b] \to [a,b] \implies \exists$ a fixed pt
- $g'(x) < {}^\exists k < 1$ on $(a,b) \implies \exists!$ f.p.,
$|p_n| \le k^n \max\{p_0 - a, b - p_0\}$,
$|p_n| \le k^n |p_1 - p_0|/(1-k)$.

2.4 Newton's Method

- $p_n = p_{n-1} - f(p_{n-1})/f'(p_{n-1})$.
If $f \in C^2[a,b]$, $p \in (a,b)$, $f(p) = 0$, $f'(0) \ne 0$, then $\exists\delta > 0$ s.t. $|p_0 - p| \le \delta \implies p_n \to p$ quadratically.
- **Secant method**
$$p_n = p_{n-1} - f(p_{n-1}) \Big/ \frac{f(p_{n-1}) - f(p_{n-2})}{p_{n-1} - p_{n-2}}$$
- **Method of false position** p_0, p_1 w/ $f(p_0)f(p_1) < 0$,
$$p_2 = p_1 - \frac{f(p_1)(p_1 - p_0)}{f(p_1) - f(p_0)},$$
If $f(p_{n-1}) \cdot f(p_{n-2}) \ge 0$, redefine $p_{n-2} \leftarrow p_{n-3}$. Calculate:
$$p_n = p_{n-1} - \frac{f(p_{n-1})(p_{n-1} - p_{n-2})}{f(p_{n-1}) - f(p_{n-2})}.$$
- **Modified Newton** When p is of multiplicity m, replace f by $\mu(x) = f(x)/f'(x)$ and do the same proc
$$p_n = p_{n-1} - \mu(p_{n-1})/\mu'(p_{n-1})$$

2.5 Aitken's Δ^2 Method

- If $p_n \to p$ linearly, then
$$\hat{p}_n = p_n - \frac{(p_{n+1} - p_n)^2}{p_{n+2} - 2p_{n+1} + p_n} \to p$$
sublinearly.
- **Steffensen's** For a problem finding $g(p) = p$, define $\hat{p}_{-1} = p_0$, $\hat{p}_n = \{\Delta^2\}(\hat{p}_{n-1})$ where
$$\{\Delta^2\}(x) = x - \frac{(g(x) - x)^2}{g(g(x)) - 2g(x) + x}.$$
It is the same with finding a root of $f(x) := g(x) - x$ with the following iterator:
$$s(x) = x - \frac{f(x)}{(f(x+h) - f(x))/h}, \quad h = f(x)$$

2.6 Horner's Method

- $P(x) = a_n x^n + \cdots + a_0$, $b_n = a_n$ and $b_k = a_k + b_{k+1}x_0$ for $k = n-1, \ldots, 0$. Then $b_0 = P(x_0)$ and
$$Q(x) = b_n x^{n-1} + \cdots + b_2 x + b_1$$
satisfies $P(x) = (x - x_0)Q(x) + b_0$ and $P'(x_0) = Q(x_0)$.
- One can repeat this to get an approximate factorization of P.

2.7 Müller's Method

For $(p_i, f(p_i))$ $(i = 0,1,2)$, determine p_3 by a root (closer to p_2) of the quadratic polynomial P agreeing at the given point.
$$f(p_i) = a(p_i - p_2)^2 + b(p_i - p_2) + c,$$
$$p_3 - p_2 = \frac{-2c}{b \pm \sqrt{b^2 - 4ac}}.$$

3.1 Lagrange Interpolation

- $(n+1)$-point interpolation:
$$P(x) = \sum_{k=0}^n f(x_k) L_{n,k}(x)$$
where
$$L_{n,k}(x) = \prod_{j\ne k} \frac{x - x_j}{x_k - x_j}$$
with
$$f(x) = P(x) + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)\cdots(x - x_n)$$
for some $\xi \in \operatorname{int}\operatorname{ConvHull}\{x_0, \cdots, x_n\}$.
- **Generalized Rolle's theorem** $f \in C[a,b]$, n times diff'ble on (a,b). If $f(x) = 0$ at $a \le x_0 < \cdots < x_n \le b$, then $\exists c \in (a,b)$, $f^{(n)}(c) = 0$.
- Runge phenomenon: a problem of oscillation at the edges of an interval that occurs when using polynomial interpolation with polynomials of high degree over a set of equispaced interpolation points.

3.2 Neville's Method

- Letting $P_S(x)$ be the Lagrange interp'n poly agreeing at x_s $(s \in S)$, then $P_{[k]}(x)$ equals to the following: $([k] = \{0, \ldots, k\})$
$$\frac{(x - x_j)P_{[k]-\{j\}}(x) - (x - x_i)P_{[k]-\{i\}}(x)}{x_i - x_j}$$
- With $Q_{i,j} = P_{i-j, i-j+1, \ldots, i}$ $(i \ge j)$, $f(x_i) = Q_{i,0}$ and
$$Q_{i,j} = \frac{(x - x_{i-j})Q_{i,j-1} - (x - x_i)Q_{i-1,j-1}}{x_i - x_{i-j}}$$

3.3 Divided Differences

$$f[x_{a,\cdots,b}] = \frac{f[x_{a+1,\ldots,b}] - f[x_{a,\ldots,b-1}]}{x_b - x_a}$$

For $x = x_0 + sh = x_n + s'h$ $(h = x_i - x_{i-1})$,
$$\begin{aligned} P_n(x) &= \sum_{k=0}^n f[x_0, \ldots, x_k] \prod_{j=0}^{k-1}(x - x_j) \\ &= \sum_{k=0}^n \binom{s}{k} \Delta^k f(x_0) \\ &= \sum_{k=0}^n f[x_n, \ldots, x_{n-k}] \prod_{j=n}^{n-k+1}(x - x_j) \\ &= \sum_{k=0}^n (-1)^k \binom{-s'}{k} \nabla^k f(x_n) \end{aligned}$$
When $n = 2m+1$ is odd, with pts $x_{-m-1,\ldots,0,\ldots,m+1}$ and $x = x_0 + sh$, (Stirling's formula)
$$\begin{aligned} P_n(x) = f[x_0] &+ \frac{sh}{2}(f[x_{-1}, x_0] + f[x_0, x_1]) \\ &+ s^2h^2 f[x_{-1}, x_0, x_1] \\ &+ \frac{s(s^2 - 1^2)h^3}{2}(f[x_{-2,-1,0,1}] + f[x_{-1,0,1,2}]) \\ &+ \cdots \\ &+ s^2(s^2 - 1^2)\cdots(s^2 - (m-1)^2)h^{2m} \cdot \\ &\qquad f[x_{-m}, \cdots, x_m] \\ &+ \frac{s(s^2 - 1^2)\cdots(s^2 - m^2)h^{2m+1}}{2} \cdot \\ &\qquad (f[x_{-m}, \cdots, x_{m+1}] + f[x_{-m-1}, \cdots, x_m]) \end{aligned}$$
where the last term disappears when $n = 2m$ (with points $x_{-m,\ldots,m}$).

3.4 Hermite Polynomials

$f \in C^1[a,b]$, $x_0, \ldots, x_n \in [a,b]$ distinct, the unique poly of least degree agreeing w/ f and f' at $x_0, \ldots, x_n$ is
$$H_{2n+1} = \sum_{j=0}^n f(x_j) H_{n,j} + \sum_{j=0}^n f'(x_j) \hat{H}_{n,j}$$

where $H_{n,j} = [1-(x-x_j)L'_{n,j}(x_j)]L^2_{n,j}(x)$ and $\hat{H}_{n,j}(x) = (x-x_j)L^2_{n,j}(x)$, with the error

$$f(x) = H_{2n+1}(x) + \frac{(x-x_0)^2\cdots(x-x_n)^2}{(2n+2)!}f^{(2n+2)}(\xi),$$

$\xi \in (a,b)$.

- Divided differences: with points $x_0, x_0, x_1, x_1, \dots, x_k, x_k$, use the div diff formula with $f[x_j] := f(x_j)$, $f[x_j, x_j] := f'(x_j)$.

3.5 Cubic Splines

- $x_0, x_1, \dots, x_n$
- $S(x) = S_j(x)$ on $[x_j, x_{j+1}]$, S_j cubic with
$$S_j(x_j) = f(x_j),\ S_j(x_{j+1}) = f(x_{j+1})$$
for $j = 0, \dots, n-1$,
$$S'_{j+1}(x_{j+1}) = S'_j(x_{j+1}),$$
$$S''_{j+1}(x_{j+1}) = S''_j(x_{j+1})$$
for $j = 0, \dots, n-2$, and with one of the following:
- (natural bd) $S''(x_0) = S''(x_n) = 0$,
- (clamped bd)
$$S'(x_0) = f'(x_0), \quad S'(x_n) = f'(x_n).$$
- Construction (natural cubic spl):
$$S_j(x) = a_j + \cdots + d_j(x-x_j)^3,$$
$$a_j = f(x_j), \quad h_j = x_{j+1} - x_j,$$
$$A\mathbf{x} = \mathbf{b} \quad \text{where}$$
$$A = \begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 \\ h_0 & 2(h_0+h_1) & h_1 & \ddots & & \vdots \\ 0 & h_1 & 2(h_1+h_2) & h_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & h_{n-2} & 2(h_{n-2}+h_{n-1}) & h_{n-1} \\ 0 & \cdots & \cdots & 0 & 0 & 1 \end{pmatrix},$$
$$\mathbf{b} = \begin{pmatrix} 0 \\ \frac{3}{h_1}(a_2-a_1) - \frac{3}{h_0}(a_1-a_0) \\ \vdots \\ \frac{3(a_n-a_{n-1})}{h_{n-1}} - \frac{3(a_{n-1}-a_{n-2})}{h_{n-2}} \\ 0 \end{pmatrix},$$
$$\mathbf{x} = [c_0, \dots, c_n]^T,$$
$$b_j = (a_{j+1}-a_j)/h_j - h_j(c_{j+1}+2c_j)/3,$$
$$d_j = (c_{j+1}-c_j)/(3h_j).$$
- Construction (clamped cubic spl):
$$A = \begin{pmatrix} 2h_0 & h_0 & 0 & \cdots & \cdots & 0 \\ h_0 & 2(h_0+h_1) & h_1 & \ddots & & \vdots \\ 0 & h_1 & 2(h_1+h_2) & h_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & h_{n-2} & 2(h_{n-2}+h_{n-1}) & h_{n-1} \\ 0 & \cdots & \cdots & 0 & h_{n-1} & 2h_{n-1} \end{pmatrix},$$
$$\mathbf{b} = \begin{pmatrix} \frac{3}{h_0}(a_1-a_0) - 3f'(a) \\ \frac{3}{h_1}(a_2-a_1) - \frac{3}{h_0}(a_1-a_0) \\ \vdots \\ \frac{3(a_n-a_{n-1})}{h_{n-1}} - \frac{3(a_{n-1}-a_{n-2})}{h_{n-2}} \\ 3f'(b) - \frac{3}{h_{n-1}}(a_n-a_{n-1}) \end{pmatrix},$$

- Error bound for clamped cub spl: $f \in C^4[a,b]$ with $\max_{[a,b]}|f^{(4)}(x)| = M$. Then for the clampled cub spl interpolant to f w.r.t. $a = x_0 < \cdots < x_n = b$,
$$|f(x) - s(x)| \le \frac{5M}{384}\max_{0\le j\le n-1} h_j^4$$
- Not-a-knot condition: $S'''(x)$ is continuous at x_1 and x_{n-1}.
- **Piecewise Cubic Hermite Polynomial**: Guide points (x_0, y_0), $(x_0+\alpha_0, y_0+\beta_0)$, $\underline{(x_1-\alpha_1, y_1-\beta_1)}$, (x_1, y_1). (Be careful for the signs)
$x(i) = x_i$, $x'(i) = \alpha_i$
$$x(t) = [2(x_0-x_1) + (\alpha_0+\alpha_1)]t^3 + [3(x_1-x_0) - (\alpha_1+2\alpha_0)]t^2 + \alpha_0 t + x_0,$$
similarly for y and β.

4.1 Numerical Differentiation

- $(n+1)$-point formula: $x_0, \dots, x_n \in [a,b]$, $f \in C^{n+1}[a,b]$. Then $\exists \eta_1, \dots, \eta_n, \xi(x) \in (a,b)$ s.t.
$$f'(x) = \sum_{k=0}^{n} f(x_k)L'_{n,k}(x) + \frac{f^{(n+1)}(\xi)}{n!}(x-\eta_1)\cdots(x-\eta_n)$$
(where $\eta_i \in (x_{i-1}, x_i)$ at which $f'(\eta_i) = P'(\eta_i)$ by Rolle)

Proof: using
$$g(t) = f'(t) - P'(t) - (f'(x) - P'(x))\prod\frac{t-\eta_j}{x-\eta_j}$$
- Three-pt midpt formula
$$f'(x_0) = \frac{1}{2h}[f(x_0+h) - f(x_0-h)] - \frac{h^2}{6}f^{(3)}(\xi)$$
- Three-pt endpt formula
$$f'(x_0) = \frac{1}{2h}\Big[-3f(x_0) + 4f(x_0+h) - f(x_0+2h)\Big] + \frac{h^2}{3}f^{(3)}(\xi)$$
- Five-pt midpt formula
$$f'(x_0) = \frac{1}{12h}\Big[f(x_0-2h) - 8f(x_0-h) + 8f(x_0+h) - f(x_0+2h)\Big] + \frac{h^4}{30}f^{(5)}(\xi)$$
- Five-pt endpt formula
$$f'(x_0) = \frac{1}{12h}\Big[-25f(x_0) + 48f(x_0+h) - 36f(x_0+2h) + 16f(x_0+3h) - 3f(x_0+4h)\Big] + \frac{h^4}{5}f^{(5)}(\xi)$$
- Second derivative midpt formula
$$f''(x_0) = \frac{1}{h^2}\Big[f(x_0-h) - 2f(x_0) + f(x_0+h)\Big] - \frac{h^2}{12}f^{(4)}(\xi)$$
- Round-off err instability: with the round-off err bound $|e(x_0 \pm h)| \le \varepsilon$ and $|f^{(3)}| \le M$, then the round-off err bound for the 3-pt midpt formula is ($e = f - \tilde{f}$)
$$\left|f'(x_0) - \frac{\tilde{f}(x_0+h) - \tilde{f}(x_0-h)}{2h}\right| \le \frac{\varepsilon}{h} + \frac{h^2}{6}M$$
minimized at $h = \sqrt[3]{3\varepsilon/M}$.

4.2 Richardson's Extrap

- $O(h^j)$ extrapolation N_j:
$$N_j(h) = N_{j-1}\left(\frac{h}{2}\right) + \frac{N_{j-1}(h/2) - N_{j-1}(h)}{2^{j-1}-1}$$
- For even $O(h^{2j})$ extrapolation N_j, replace $2^{j-1}-1$ by $4^{j-1}-1$.

4.3 Newton–Cotes

- $(n+1)$-pt closed Newton–Cotes: $x_0 = a, x_n = b, h = (b-a)/n$, error term is
$$\frac{h^{n+3}f^{(n+2)}(\xi)}{(n+2)!}\int_0^n t^2(t-1)\cdots(t-n)dt$$
for even n and $f \in C^{n+2}[a,b]$; and
$$\frac{h^{n+2}f^{(n+1)}(\xi)}{(n+1)!}\int_0^n t(t-1)\cdots(t-n)dt$$
for odd n and $f \in C^{n+1}[a,b]$.
- Coeff for $n = 1,2,3,4$: $\frac{1}{2}(1,1)$; $\frac{1}{3}(1,4,1)$; $\frac{3}{8}(1,3,3,1)$; $\frac{2}{45}(7,32,12,32,7)$. (Trapezoid, Simpson, Simpson $\frac{3}{8}$, Boole)
- $(n+1)$-pt open Newton–Cotes: $x_{-1} = a, x_{n+1} = b, h = (b-a)/(n+2)$;
for the error term, replace $\int_0^n$ by $\int_{-1}^{n+1}$ in the above.
- Coeff for $n = 0,1,2,3$: (2); $\frac{3}{2}(1,1)$; $\frac{4}{3}(2,-1,2)$; $\frac{5}{24}(11,1,1,11)$. ($n = 0$: <u>midpt rule</u>)
- Compo Simpson err term: $-\frac{b-a}{180}h^4 f^{(4)}(\mu)$
- Compo Trapez err term: $-\frac{b-a}{12}h^2 f''(\mu)$
- Compo midpt rule: $h = (b-a)/(n+2)$, n even,
$$\int_a^b f = 2h\sum_{\text{even } j} f(x_j) - \frac{b-a}{6}h^2 f''(\mu)$$

4.4 Gaussian Quadrature

x_j: roots of n-th Legendre poly,
$$c_i = \int_{-1}^{1}\prod_{j\ne i}\frac{x-x_j}{x_i-x_j}\,dx,$$
$$\int_{-1}^{1}P(x)dx = \sum c_j P(x_j)$$