Density matrix method.

$$P = \sum_{i=1}^{N} |\psi_i\rangle\langle\psi_i| = \Psi \Psi^*$$

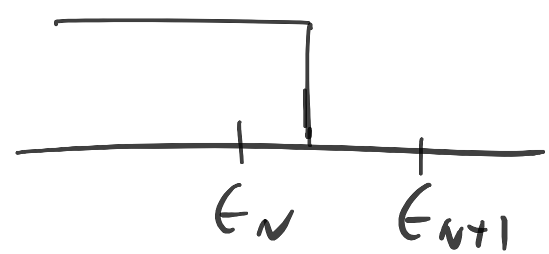
$H \in \mathbb{C}^{d\times d}$ $\quad \Psi = [\psi_1 \cdots \psi_N]$. $\quad \Psi^* \Psi = I$.

$$H \psi_i = \varepsilon_i \psi_i$$

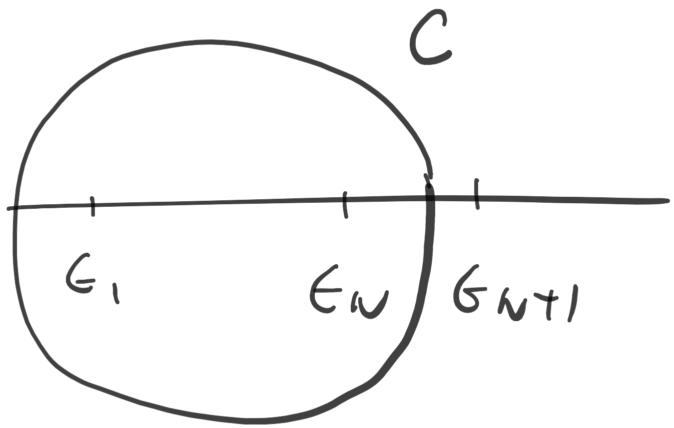
Matrix function

$$f(H) = \sum_{i=1}^{d} f(\epsilon_i)\, \psi_i \psi_i^*$$

$$f(\cdot) = \mathbb{1}_{(-\infty, \mu]}(\cdot)$$



$$P = \mathbb{1}_{(-\infty,\mu]}(H) = \frac{1}{2\pi i}\oint_C \frac{1}{\lambda I - H}\, d\lambda.$$



$$\approx \sum_{\ell} \omega_\ell \underbrace{(z_\ell - H)^{-1}}_{G(z_\ell)}.$$

Optimization perspective. $\varepsilon_N < \varepsilon_{N+1}$

$$E = \sum_{i=1}^{N} \varepsilon_i = \inf_{\substack{P=P^* \\ 0 \le P \le 1 . \\ \operatorname{Tr} P = N}} \operatorname{Tr}[HP]$$

Pf: $H = \sum_{i=1}^{d} \psi_i \varepsilon_i \psi_i^*$. $\tilde{P}_{ij} = \psi_i^* P \psi_j$

$$\operatorname{Tr}(HP) = \sum_{i=1}^{d} \operatorname{Tr}\left[\varepsilon_i \psi_i^* P \psi_i\right] = \sum_{i=1}^{d} \varepsilon_i \tilde{P}_{ii}$$

$$\Rightarrow \inf_{\substack{0 \le \tilde{P}_{ii} \le 1 \\ \sum_i \tilde{P}_{ii} = N}} \sum_{i=1}^{d} \varepsilon_i \tilde{P}_{ii} \Rightarrow \tilde{P}_{ii} = \begin{cases} 1, & i \le N \\ 0, & i > N . \end{cases}$$

$\Rightarrow$ e_i are eigenvectors. $\Rightarrow \tilde{P} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$ $\square$

Courant-Fischer min-max theorem. $A \in \mathbb{C}^{d\times d}$.

$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_N$

$$\lambda_k = \max_{\dim S = k} \min_{\substack{v \in S \\ v^* v = 1}} v^* A v$$

$$= \min_{\dim S = d-k+1} \max_{\substack{v \in S \\ v^* v = 1}} v^* A v .$$

<u>Cor.</u> $$E = \inf_{\substack{P = P^* \\ P^2 = P \\ \operatorname{Tr} P = N}} \operatorname{Tr}[HP] .$$

Finite temperature.

free energy.

$$F = \inf_{P \in \mathcal{D}} \operatorname{Tr}[HP] - \frac{1}{\beta} S[P]$$

$$S[P] = -\operatorname{Tr}\left[P \log P + (1-P)\log(1-P)\right].$$

$$\mathcal{D} = \{P \mid P = P^*, \ \operatorname{Tr} P = N, \ 0 \leq P \leq 1\}.$$

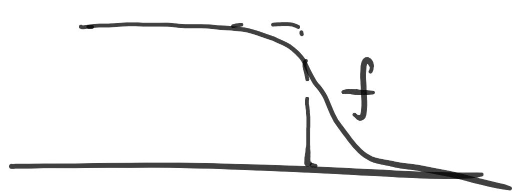
$$\mathcal{L}[P] = \mathrm{Tr}[HP] - \frac{1}{\beta} S[P] - \left(\mathrm{Tr}[P] - N\right)\mu$$

$$\frac{\delta \mathcal{L}}{\delta P} = H + \frac{1}{\beta}\left(\log P + I - \log(1-P) - I\right) - \mu I = 0.$$

$$\Rightarrow \log P(1-P)^{-1} = -\beta(H - \mu I)$$

$$\Rightarrow P = \left(1 + e^{\beta(H-\mu I)}\right)^{-1} = f_\beta(H-\mu)$$

$$f_\beta(x) = \frac{1}{1+e^{\beta x}}, \quad \sum_i f(\epsilon_i - \mu) = N.$$



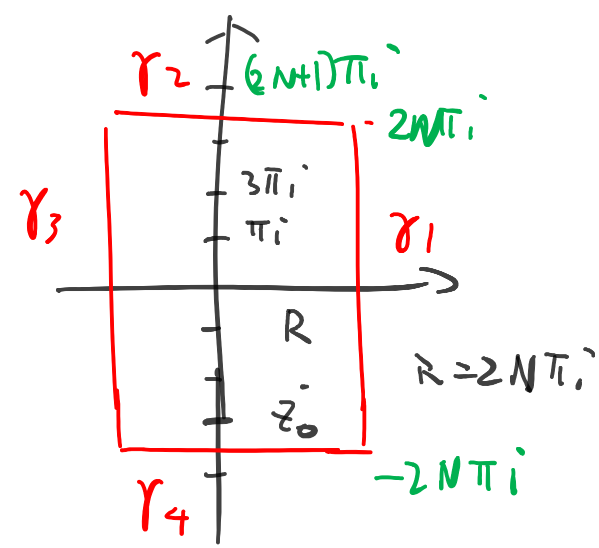
$$|f_\beta(x)| = \infty \Rightarrow \beta x = (2n+1)\pi i \Rightarrow x = \frac{(2n+1)\pi i}{\beta}.$$

Thm. $f_\beta(x) = \frac{1}{2} - \frac{1}{\beta}\sum_{n\in\mathbb{Z}} \frac{1}{x - \frac{(2n+1)\pi i}{\beta}}$ (Matsubara expansion)

Note: conditional conv. see proof below.

pf: partial fraction expansion.

Consider $\beta = 1$, $F(z) = \frac{f(z)}{z - z_0}$



$$\frac{1}{2\pi i}\oint_{C_R} F(z)\,dz.$$

$$= \frac{1}{2\pi i}\oint_{C_R} \frac{f(z)}{z} + \frac{f(z)z_0}{(z-z_0)z}\,dz$$

$C_R = \gamma_1 \cup \gamma_2 \cup \gamma_3 \cup \gamma_4.$

$|f(z)|$ bounded on C.

$$\rightarrow \frac{1}{2\pi i} \oint_{C_R} F(z)\, dz \overset{\text{residue}}{=\!=} f(z_0) + \sum_{n=-N}^{N-1} \frac{1}{(2n+1)\pi i - z_0}$$

$$= f(0) + \sum_{n=-N}^{N-1} \frac{1}{(2n+1)\pi i} + \underbrace{\frac{1}{2\pi i} \oint_{C_R} \frac{f(z) z_0}{z(z-z_0)}\, dz}_{\xrightarrow{R\to\infty} 0}$$

$$\Rightarrow f(z_0) = \frac{1}{2} - \lim_{N\to\infty} \left(\sum_{n=-N}^{N-1} \frac{1}{(2n+1)\pi i - z_0} - \frac{1}{(2n+1)\pi i} \right)$$

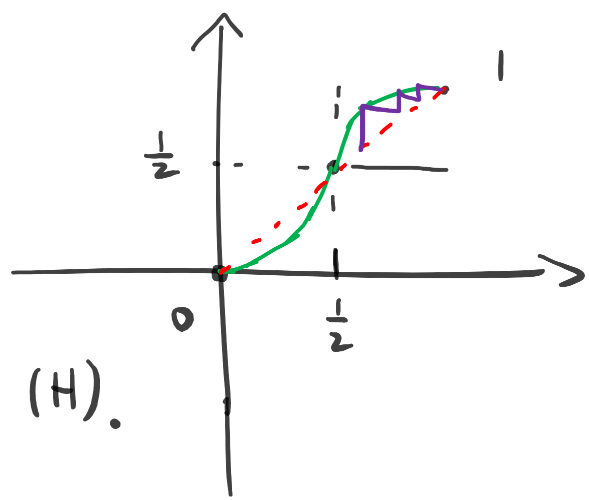
$$= \frac{1}{2} - \sum_{n\in\mathbb{Z}} \frac{1}{(2n+1)\pi i - z_0}$$

Density matrix purification.

$\mathrm{spec}(H) \subseteq [0,1] \setminus \{\tfrac{1}{2}\}$.

McWeeny.

$f(x) = 3x^2 - 2x^3$



$f(x) - x = -x(2x-1)(x-1)$

$$= \begin{cases} >0, & \tfrac{1}{2} < x < 1 \\ <0, & 0 < x < \tfrac{1}{2} \end{cases}$$

$$\lim_{n\to\infty} (\underbrace{f \circ f \cdots \circ f}_{n})(H) = \mathbb{1}_{(\frac{1}{2},1]}(H).$$

general H.

$$P_0 = \frac{1}{2\sigma}(\mu - H) + \frac{1}{2} I, \quad \sigma = \max_i |\varepsilon_i - \mu|$$

Why this particular polynomial?

Newton - Schulz for $\operatorname{sgn}(x) = \begin{cases} 1, & x>0 \\ 0, & x=0 \\ -1, & x<0 \end{cases}$. $\operatorname{spec}(H) \subset (-1,1)$

$\operatorname{sgn}(H)$ from Newton.

$$g(X) = X^2 - I. \qquad X_0 = H.$$

$$X_{k+1} = X_k - (g'(X_k))^{-1} g(X_k) = \frac{1}{2}(X_k + X_k^{-1})$$

Schultz: Newton for

$$h(Y) = Y^{-1} - X. \qquad Y_0 = X.$$

$$Y_{k+1} = Y_k - (h'(Y_k))^{-1} h(Y_k) = 2Y_k - XY_k^2$$

$$\Rightarrow X_{k+1} = \frac{1}{2}(X_k + 2X_k - X_k^3) = \frac{3}{2} X_k - X_k^3. \longrightarrow \text{scaling for McWeeny (exer)}.$$

Evaluation of chemical potential.

① bisection (slow).

② Newton (require derivative. also not so well defined for insulating sys.)

Linear scaling. If matrix-matrix multiplication can be performed at $O(N)$ cost.

near-sightedness. show code.

Recursive expansion

$$\mathbb{1}_{(-\infty,\mu]}(H) = \lim_{n\to\infty} (f_n \circ \cdots \circ f_0)(H).$$

McWeeny: $f_n = \cdots = f_0 = f$

Benefit of choosing multiple different f:

automatically adjust for chemical potential.

second order spectral projection method (SP2)

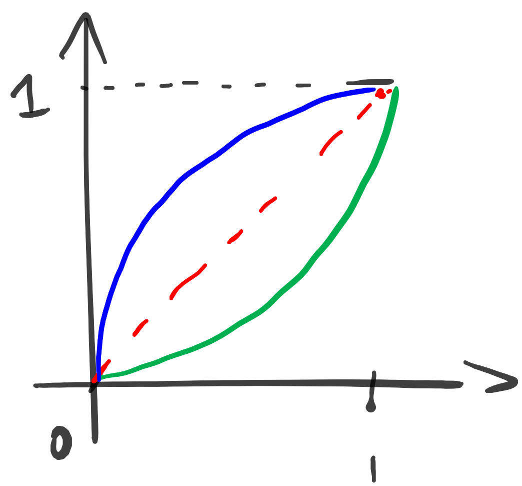
[Niklasson PRB 2002]

$$X_0 = \frac{\epsilon_{max} I - H}{\epsilon_{max} - \epsilon_{min}}$$

no dependence on μ

$$f_n(X_{n-1}) = \begin{cases} X_{n-1}^2 \\ 2X_{n-1} - X_{n-1}^2 \end{cases}$$

i.e. $f_n(x) - x = \begin{cases} x^2 - x \\ x - x^2 \end{cases}$



x^2

$2x - x^2$

trace correction:

$$f_n = \arg\min_{f_n \in \{x^2, 2x - x^2\}} \left| Tr(f_n(X_{n-1})) - N \right|$$

Convergence?

Density matrix minimization.

$$\mathcal{D} = \{ P \mid P^2 = P,\ P = P^*,\ \mathrm{Tr} P = N \}.$$

$$E = \inf_{P \in \mathcal{D}} \mathrm{Tr}[PH].$$

DMM.

$$E = \inf_{P = P^*} \left[(3P^2 - 2P^3)(H - \mu I) \right], \quad P_0 = \frac{1}{2} I. \quad \mathrm{spec}(H) \subseteq [0,1]$$

$$\frac{\delta E}{\delta P} = 6(P - P^2)(H - \mu I)$$ (P commute w. H in each step).

$\Rightarrow$ stationary pt $P = P^2$

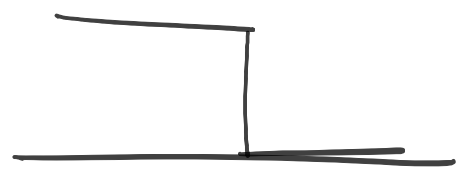
Analyze: at some stage

$$P = \sum_i f_i \psi_i \psi_i^* \quad . \quad 0 \le f_i \le 1.$$

$$\delta E = \sum_i 6 f_i (1 - f_i)(\varepsilon_i - \mu)\, \delta f_i$$

$\varepsilon_i < \mu \Rightarrow \delta f_i > 0.$

$\varepsilon_i > \mu \Rightarrow \delta f_i < 0.$



$\Rightarrow$ gradient descent towards step function.

typically slower than purification.

Fermi operator expansion (FOE).

finite temperature (smoothness)

$$P = f_\beta(H-\mu)$$

$$f_\beta(x) \approx \sum_{n=1}^{m} g_n(x). \Rightarrow f_\beta(H-\mu) \approx \sum_{n=1}^{m} g_n(H-\mu)$$

g_n: polynomial or rational function.

err analysis. $A \in \mathbb{C}^{d\times d}$. spec $A \subseteq [a,b]$

$$\|f(A)-g(A)\|_2 \leq \|f-g\|_{L^\infty([a,b])} := \sup_{x\in[a,b]} |f(x)-g(x)|$$

Pf: $A v_i = \lambda_i v_i$, $i=1,\cdots,d$.

$$\|f(A)-g(A)\|_2 = \sup_{\substack{x\in\mathbb{C}^d \\ \|x\|_2=1}} \|f(A)x - g(A)x\|_2 .$$

$(x = \sum_i c_i v_i)$

$$= \sup_{\substack{x\in\mathbb{C}^d \\ \|x\|_2=1}} \left\| \sum_i \left(f(\lambda_i)-g(\lambda_i)\right) c_i v_i \right\|_2$$

$$\leq \sup_i |f(\lambda_i)-g(\lambda_i)| \cdot \|x\|_2 \leq \|f-g\|_{L^\infty([a,b])} \|x\|_2 .$$

$\square$.

Polynomial expansion.

$$f_\beta(x) \approx \sum_{n=1}^{m} C_n x^{n-1}$$

computing high matrix power is unstable

$\Rightarrow$ Chebyshev polynomial.

$$f_\beta(x) \approx \sum_{n=1}^{m} C_n T_n(x). \qquad x \in (-1,1)$$

$$T_0(x) = 1, \quad T_1(x) = x.$$

$T_{n+1}(x) = 2x T_n(x) - T_{n-1}(x)$. three term recursion.

orthogonality

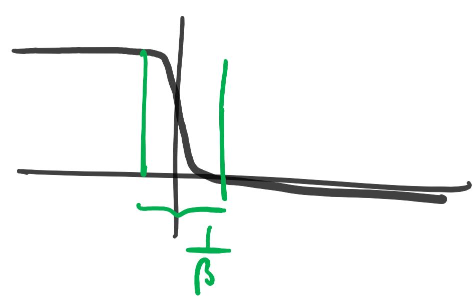
$$\int_{-1}^{1} \frac{1}{\sqrt{1-x^2}}\, T_n(x)\, T_m(x)\, dx = \delta_{nm}.$$

$x = \cos\theta \qquad T_n(\cos\theta) = \cos n\theta.$

$$\int_{0}^{\pi} T_n(\cos\theta)\, T_m(\cos\theta)\, d\theta = \delta_{nm}\gamma_n \qquad \gamma_n = \begin{cases} \pi, & n=0. \\ \frac{\pi}{2}, & n>0 \end{cases}$$

$$c_n = \frac{1}{\gamma_n}\int_{-1}^{1} \frac{1}{\sqrt{1-x^2}}\, f(x)\, T_n(x)\, dx.$$ quadrature.

Q: $f = \frac{1}{1+e^{\beta x}}$. How many terms to achieve error ϵ?



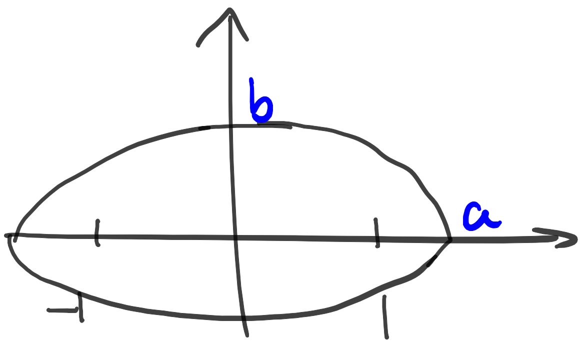
$\Rightarrow O(\beta)$ terms.

Thm. f analytic w. $|f(z)| \le M$ bounded by ellipse w. foci ± 1, ρ is the sum of the sum of the lengths of major and minor semiaxes. then

$$\|f - f_n\|_{L^\infty([-1,1])} \le \frac{2M}{(\rho-1)\rho^n}, \qquad f_n = \sum_n c_n T_n(x).$$

$$c_n = \frac{1}{\gamma_n} \int_{-1}^{1} f(x) T_n(x)\, dx.$$

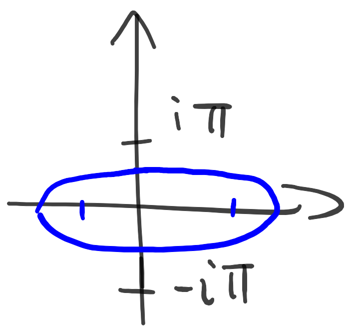
Pf: ex. Trefethen. SIAM Rev. 2008.



$\rho = a+b$, $\rho > 1$.

$a = \sqrt{1+b^2}$, $\rho = b + \sqrt{1+b^2}$

Ex. $f(x) = \dfrac{1}{1+e^x}$,



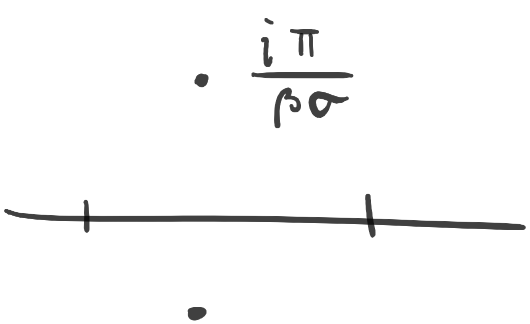
$$\rho = \frac{\pi}{2} + \sqrt{1+\left(\frac{\pi}{2}\right)^2}.$$

exp. converge ρ^{-n}

Ex. $\operatorname{spec}(H-\mu) \subset [-\sigma, \sigma]$

$$f(H-\mu) = \frac{1}{1+e^{\beta(H-\mu)}} = \frac{1}{1+e^{\beta\sigma \tilde{H}}},$$

$\tilde{H} = \frac{1}{\sigma}(H-\mu)$, $\operatorname{spec}(\tilde{H}) \subset [-1,1]$. $\beta\sigma \gg 1$.



$$\rho = \frac{\pi}{2\beta\sigma} + \sqrt{1+\left(\frac{\pi}{2\beta\sigma}\right)^2}$$

$$\approx 1 + \frac{\pi}{2\beta\sigma}$$

$$\rho^{-n} = \left(1+\frac{\pi}{2\beta\sigma}\right)^{\frac{2\beta\sigma}{\pi}\cdot\left(-\frac{n\pi}{2\beta\sigma}\right)} \approx e^{-\frac{n\pi}{2\beta\sigma}} < \epsilon$$

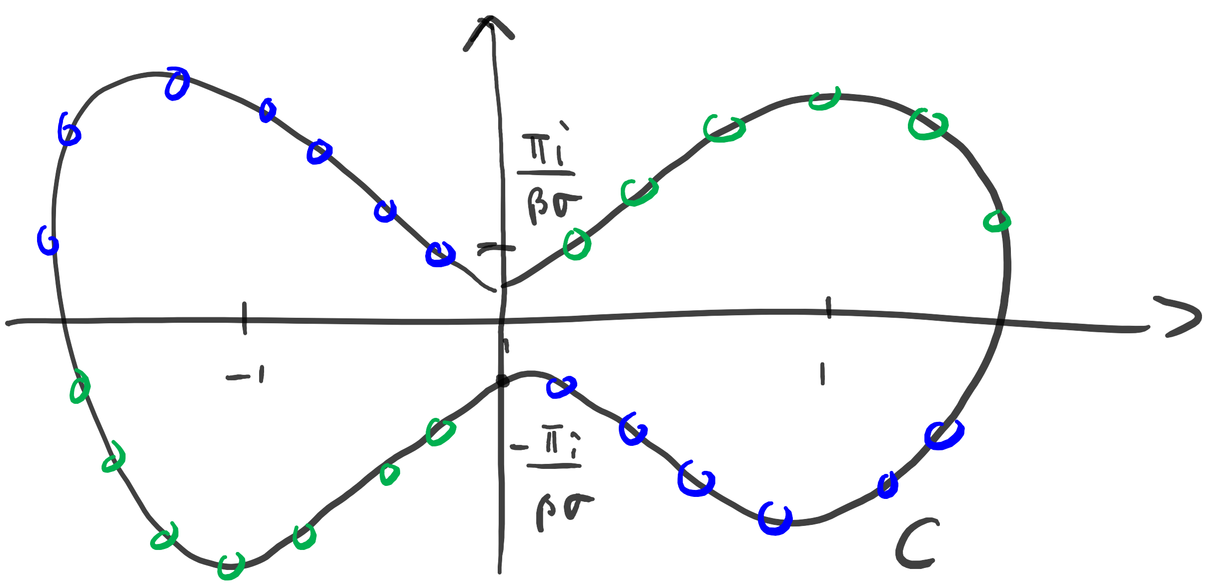
$$\Rightarrow n \sim O(\beta\sigma)$$

Rational expansion.

① truncating Matsubara expansion.

obviously $n \sim O(\beta\sigma)$.

② Discretizing contour integral.



$n \sim \log(\beta\sigma)$

show code.

(perhaps C interface)

$$f_\beta(x) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{z - x}\, dz$$

$$\approx \frac{1}{2\pi i} \sum_l \frac{f(z_l)\Delta z_l}{z_l - x} = \frac{1}{2i} \sum_{l=1}^{2n} \omega_l \, (z_l - x)^{-1}$$

$$\omega_l = \frac{1}{\pi} f(z_l)\Delta z_l .$$

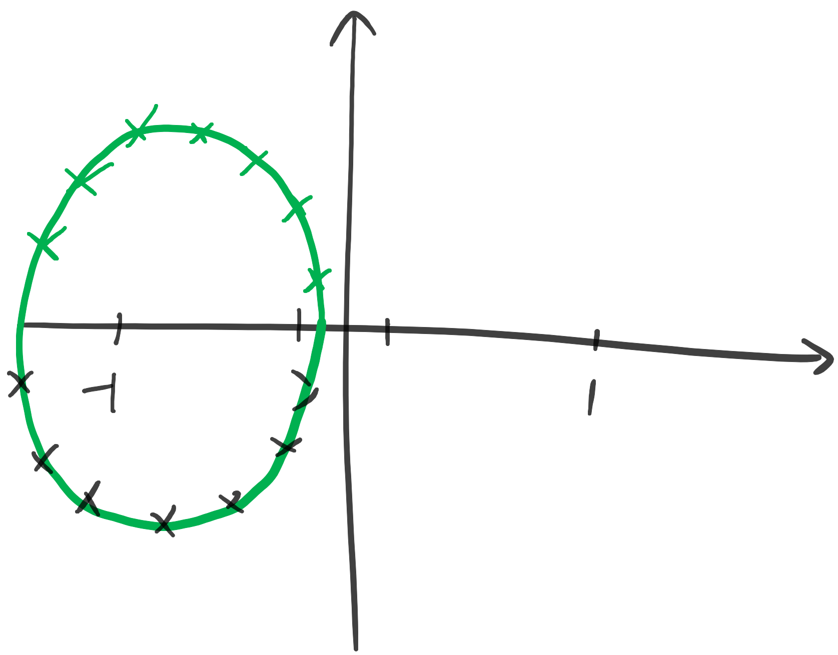
$$z_{l+n} = z_l^*, \qquad \Delta z_{l+n} = -\Delta z_l^* \Rightarrow \omega_{l+n} = -\omega_l^*$$

$$\Rightarrow \frac{1}{2i} \sum_{l=1}^{n} \omega_l (z_l - x)^{-1} - \left[\omega_l (z_l - x)^{-1}\right]^*$$

$$= \mathrm{Im} \sum_{l=1}^{n} \omega_l \, (z_l - x)^{-1}.$$

$$f(H-\mu) \approx \sum_{l=1}^{n} \mathrm{Im}\left[\omega_l \left(z_l - (H-\mu) \right)^{-1} \right].$$

Zero temperature.



Circle & uniform grid along angular direction.

Note: no dependence on E_{max}. Finite temperature can work as well.

Selected inversion.

$A \in \mathbb{R}^{d\times d}$. $A > 0$.

$A = LL^T$ cholesky factorization.

$A \in \mathbb{C}^{d\times d}$. symmetric. (real / complex).

$A = LDL^T$ generalized cholesky factorization.

basis: $U = [\phi_1(r), \cdots, \phi_d(r)]$.

$$\rho(r) = \sum_{ij} \phi_i(r)\phi_j(r) P_{ij}.$$

$\forall r$, only need P_{ij} s.t. $\phi_i(r)\phi_j(r) \neq 0$.

real space discretization : $\phi_i(r_k)\phi_j(r_k) = \delta_{ik}\delta_{jk}\phi_i^2(r_i)$

only need P_{ii}. $\Rightarrow$ diagonal elements of $(z_l - (H-\mu))^{-1}$.

A sparse $\longrightarrow$ compute $\operatorname{diag} A^{-1}$.

Recursive Green's function.

$$A = \begin{pmatrix} a_{11} & a_{12} & & & \\ a_{21} & a_{22} & a_{23} & & \\ & & \ddots & \ddots & \\ & & & & a_{n-1,n} \\ & & & a_{n,n-1} & a_{nn} \end{pmatrix}$$

tridiagonal.

$a_{ij} = a_{ji}$ (can be relaxed).

$$A = \begin{pmatrix} a_{11} & b^T \\ b & \bar{A} \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & \\ l & I \end{pmatrix} \begin{pmatrix} a_{11} & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} 1 & l^T \\ 0 & I \end{pmatrix}.$$

$l = b a_{11}^{-1}$, $S = \bar{A} - b a_{11}^{-1} b^T$.

Schur complement. Assume $a_{11} \neq 0$.

tridiagonal.

$$l = \begin{pmatrix} a_{22} a_{11}^{-1} \\ 0 \end{pmatrix} \qquad b\, a_{11}^{-1} b^T = \left(\begin{array}{c|c} a_{21} a_{11}^{-1} a_{12} & 0 \\ \hline 0 & 0 \end{array}\right)$$

$$S = \begin{pmatrix} a_{22} - a_{21} a_{11}^{-1} a_{12} & a_{23} & & \\ a_{32} & a_{33} & a_{34} & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix}$$

repeat.

$$A = \underbrace{\begin{pmatrix} 1 & \\ l_1 & I \end{pmatrix} \begin{pmatrix} 1 & & \\ & 1 & \\ & l_2 & I \end{pmatrix} \cdots}_{L} \underbrace{\begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}}_{D} \underbrace{\cdots \begin{pmatrix} 1 & & \\ & 1 & l_2^T \\ & & I \end{pmatrix} \begin{pmatrix} 1 & l_1^T \\ & I \end{pmatrix}}_{L^T}.$$

$$L = \begin{pmatrix} 1 & & & & \\ l_{21} & 1 & & & \\ & l_{32} & 1 & & \\ & & \ddots & \ddots & \\ & & & l_{n,n-1} & 1 \end{pmatrix}$$

$A^{-1} = L^{-T} D^{-1} L^{-1}$. dense in general.

$$A^{-1} = \begin{pmatrix} a_{11}^{-1} + l^T S^{-1} l & -l^T S^{-1} \\ -S^{-1} l & S^{-1} \end{pmatrix}.$$

Assume S^{-1} given. only need $(S^{-1})_{11}$. then const cost.

$\xrightarrow{\text{recursive}}$ $(A^{-1})_{nn}$ given. reverse. $O(N)$ cost.

Alg. recursive Green's function. $G = A^{-1}$

$$G_{nn} = d_n^{-1}.$$

for $i = n-1 : -1 : 1$

$$G_{ii} = d_{ii} + l_{i+1,i}\, G_{i+1,i+1}\, l_{i+1,i}$$

end.

Code

selected inversion:

generalize to any sparse A (sym or non-sym).

Cost: 1D : $O(N)$.

2D : $O(N^{1.5})$

3D : $O(N^2)$

asym. same to LDL^T factorization.

PEXSI.