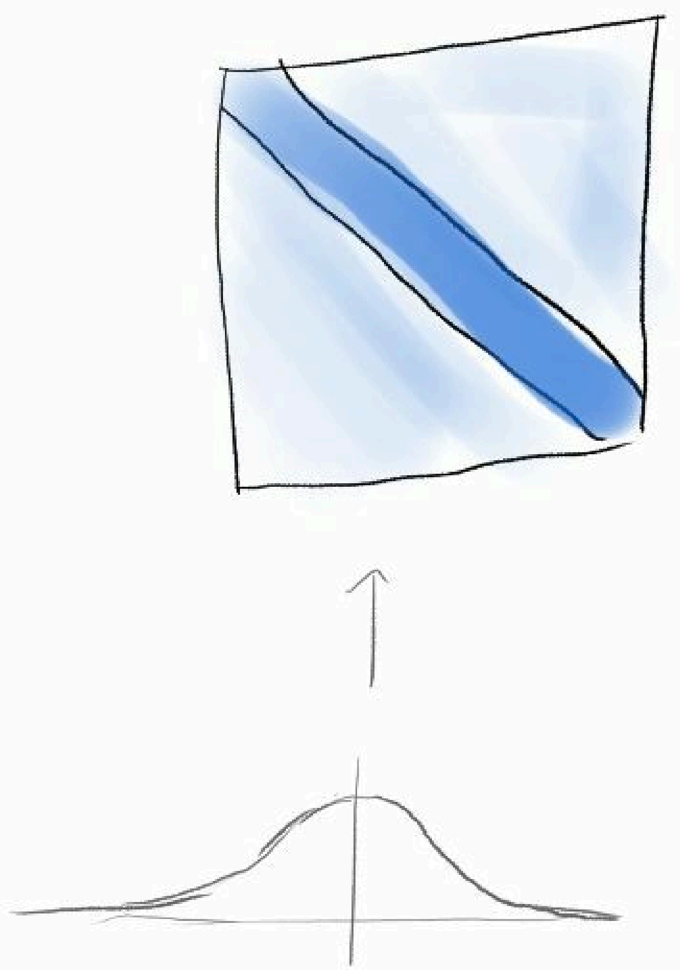
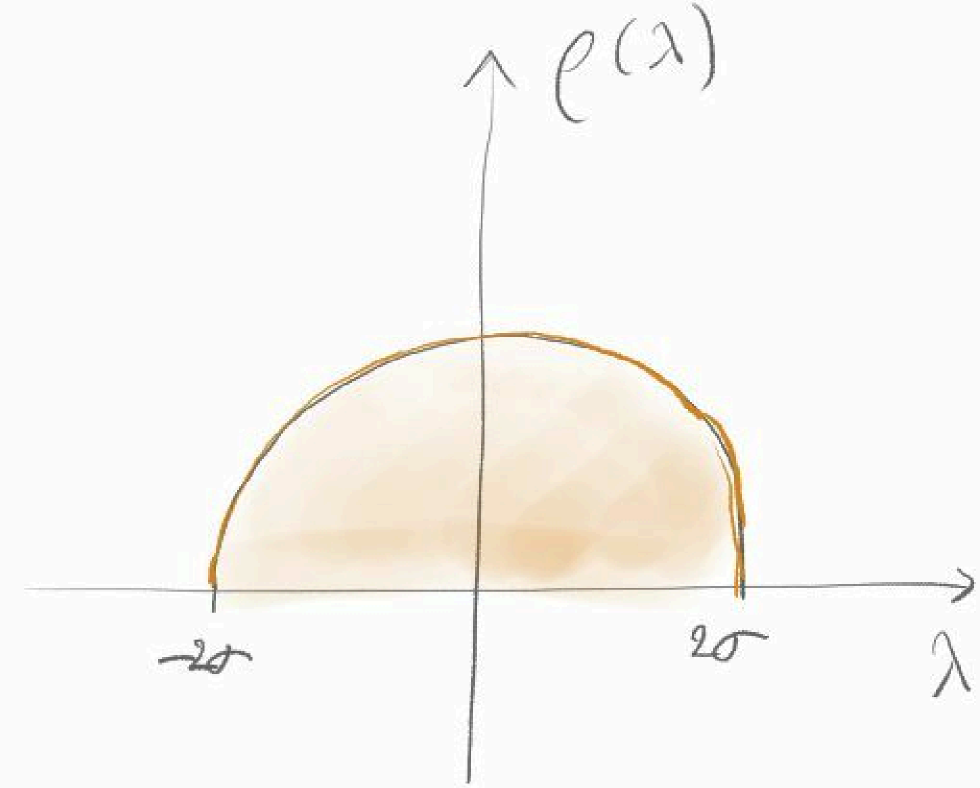
EAI Math Reading Group:

Random Matrix Theory I





Wigner Matrices &

semi-circle law

Basic Problem:

$X \in \mathbb{R}^{N \times N}$ s.t. $X_{ij} \sim Q_{ij}$

probability distribution

Questions:

- What is the distribution of the eigenvalues of X? [as $N \to \infty$]

- What is the distribution of the eigenvectors?

Applications:
- [Bayesian] Statistics (covariance matrices)
- Neural Nets at initialization

Today: The "simplest" matrix distribution

Spectral Theory Refresher

- For a matrix [linear operator] $A \in \mathbb{R}^{N \times N}$, eigenvalues are complex numbers λ s.t. $\lambda I - A$ is not invertible

  ↳ characteristic equation $\underbrace{\det(\lambda I - A) = 0}_{\text{polynomial equation in } \lambda}$

- If there is a vector $v \in \mathbb{R}^N$ s.t

  $$Av = \lambda v \quad \rightarrow \text{eigenvector}$$

- For symmetric matrices $(A = A^T)$, the eigenvalues are all real and there is an orthonormal basis of eigenvectors $\{v_i\}_{i=1}^N$ $\langle v_i, v_j \rangle = \delta_{ij}$

Normalized Trace

$X \sim q \quad \mathbb{E}[X^k]$

Moments of a matrix distribution?

× • $\mathbb{E}[A^k]$ $\rightsquigarrow$ large $(N \times N)$, what happens as $N \to \infty$

✓ • $\tau(A) = \frac{1}{N}\mathbb{E}[\operatorname{tr} A] = \frac{1}{N}\sum_{k=1}^{N}\lambda_k = \langle \lambda \rangle$

For F polynomial/analytic: $\frac{1}{N}\operatorname{tr}(F(A)) = \frac{1}{N}\sum_{k=1}^{N} F(\lambda_k) \equiv \langle F(\lambda) \rangle$

$$\lim_{N\to\infty} \langle F(\lambda) \rangle = \frac{1}{N}\mathbb{E}[\operatorname{tr} F(A)] = \tau(F(A))$$

If the eigenvalues asymptotically have density $\rho(\lambda)$:

$$\tau(F(A)) = \int_{\mathbb{C}} \rho(\lambda) F(\lambda)\, d\lambda$$

Moments of a random matrix

$$m_k = \tau(A^k) = \langle \lambda^k \rangle$$

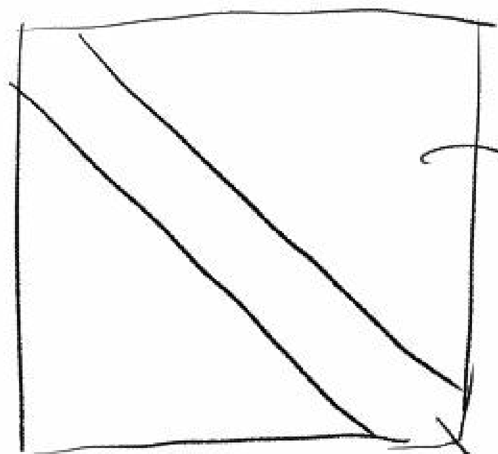
Example

$$m_1 = \frac{1}{N}\mathbb{E}\left[\operatorname{tr}(A)\right] = \frac{1}{N}\sum_{k=1}^{N}\lambda_k \quad \text{(mean eigenvalue)}$$

$$m_2 = \frac{1}{N}\sum_{ij} A_{ij}^2 = \frac{1}{N}\|A\|_F^2 \quad \text{(Frobenius norm)}$$

Wigner Matrices

Symmetric matrix $X = X^T$ with entries $X_{ij} \sim N(0, \sigma_{ij}^2)$



off-diagonal $X_{ij} = X_{ji} \sim N(0, \sigma_{od}^2)$

diagonal elements $X_{ii} \sim N(0, \sigma_d^2)$

Moments:

$$\tau(X) = \frac{1}{N}\mathbb{E}[\operatorname{tr} X] = \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}[X_{ii}] = 0$$

$$\tau(X^2) = \frac{1}{N}\mathbb{E}[\operatorname{tr} XX^T] = \frac{1}{N}\mathbb{E}\Big[\sum_{i,j}^{N} X_{ij}^2\Big] = \frac{1}{N}\Big[N(N-1)\sigma_{od}^2 + N\sigma_d^2\Big]$$

$$= \underbrace{(N-1)\sigma_{od}^2}_{\text{dominates as } N \to \infty} + \sigma_d^2$$

Normalized variances

Pick $\sigma_{od}^2 = \frac{\sigma^2}{N}$ $\sigma_d^2 = \frac{2\sigma^2}{N}$

$\rightsquigarrow$ Gaussian Orthogonal Ensemble (GOE) $\rightarrow \equiv$ distribution

Can be generated as

$$X = H + H^T \quad \text{where } H \in \mathbb{R}^{N \times N} \quad H_{ij} \sim N\left(0, \frac{\sigma^2}{2N}\right)$$

$$\sigma_{od}^2 = Var\left[H_{ij} + H_{ji}\right] = \frac{\sigma^2}{N}$$

$$\sigma_d^2 = Var\left[2H_{ii}\right] = \frac{4\sigma^2}{2N} = \frac{2\sigma^2}{N}$$

- $\tau(X^3) = 0$, $\tau(X^4) = 2\sigma^4 \underset{\neq}{} 3\sigma^4$ $\Rightarrow$ non gaussian eigenvalue distribution

Rotational (Orthogonal) Invariance

A random matrix X is rotationally invariant if

OXO^T has the same law as X

for $O \in \mathbb{R}^{N \times N}$ $OO^T = I$

$$\Leftrightarrow \quad \mathbb{P}[OXO^T \in E] = \mathbb{P}[X \in E]$$

$\rightarrow$ Corollary: the eigenvectors of X are "uniformly distributed" (for symmetric X at least)

More formally: the eigenvectors of X are sampled from the uniform distribution on the Stiefel manifold

GOE matrices are rotationally invariant

Let $X = H + H^T$ with $H_{ij} \sim N\left(0, \frac{\sigma^2}{2N}\right)$

Recall that a multivariate gaussian vector $v \sim N\left(0, \frac{\sigma^2}{N} I\right)$ is rotationally invariant, i.e. $w = Ov \sim N\left(0, \frac{\sigma^2}{N} I\right)$

$\Rightarrow$ The columns and rows of H are rotationally invariant

$$OH \overset{D}{=} H \qquad OHO^T \overset{D}{=} OH$$

$$\Rightarrow OXO^T = O(H + H^T)O^T \overset{D}{=} H + H^T = X$$

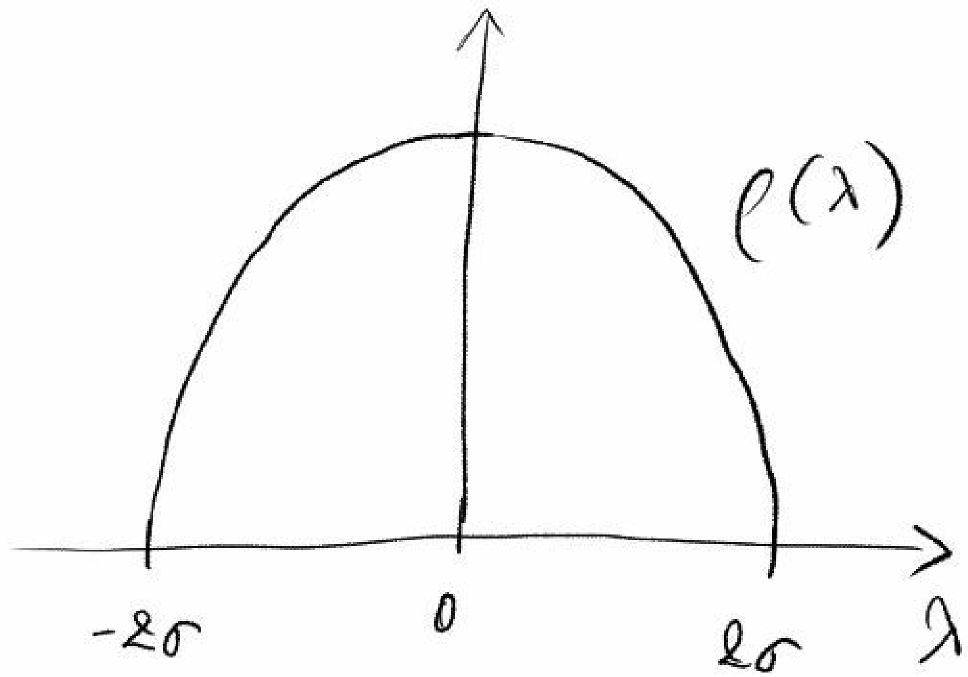
- Alternatively

$$P(\{X_{ij}\}) = \left(\frac{1}{2\pi\sigma_d^2}\right)^{\frac{N}{2}} \left(\frac{1}{2\pi\sigma_{od}^2}\right)^{\frac{N(N-1)}{4}} \exp\left\{-\sum_{i=1}^{N} \frac{X_{ii}^2}{2\sigma_d^2} - \sum_{i<j}^{N} \frac{X_{ij}^2}{2\sigma_{od}^2}\right\}$$

$$\propto \exp\left\{\frac{-N}{4\sigma^2} \underbrace{\operatorname{tr} X^2}_{\text{invariant under } \tilde{X} = OXO^T}\right\} = \exp\left\{-N \operatorname{tr} V(M)\right\}$$

Deriving the eigenvalue density of the GOE

TLDR $\rho(\lambda) = \frac{\sqrt{4\sigma^2 - \lambda^2}}{2\pi\sigma^2}$



Deriving ρ takes a little work,

but is very instructive

Next: "Spelled out" derivation *

*: May contain Complex Analysis and Physicist techniques

Resolvent

For $A \in \mathbb{R}^{N \times N}$ its resolvent is

$$G_A(z) = (zI - A)^{-1} \quad \text{for } z \in \mathbb{C} \setminus \sigma(A)$$

↳ Spectrum of A

Property: For z large enough $(|z| > \|A\|)$

$$G_A(z) = \left[z\left(I - \frac{1}{z}A\right) \right]^{-1} = \frac{1}{z}\left(I - \frac{1}{z}A\right)^{-1} = \frac{1}{z} \sum_{k=0}^{+\infty} \frac{A^k}{z^k}$$

↳ Neumann Lemma

Stieltjes Transform

$$g_N^A(z) = \frac{1}{N} \operatorname{tr}(G_A(z)) = \frac{1}{N} \sum_{k=1}^{N} \frac{1}{z - \lambda_k}$$

The Stieltjes Transform and the eigenvalue density

$$g_N(z) = \int_{-\infty}^{+\infty} \frac{\rho_N(\lambda)}{z-\lambda}\, d\lambda \qquad \rho_N(\lambda) = \frac{1}{N}\sum_{k=1}^{N} \delta(\lambda-\lambda_k)$$

As $z \to \infty$:

$$g_N(z) = \sum_{k=0}^{+\infty} \frac{1}{z^{k+1}} \frac{1}{N} \operatorname{tr}(A^k)$$

Assuming $\frac{1}{N}\operatorname{tr}(A^k) \xrightarrow[N\to\infty]{} \tau(A^k)$ we expect

$$g_N(z) \xrightarrow[N\to\infty]{} g(z) = \sum_{k=0}^{+\infty} \frac{1}{z^{k+1}} \tau(A^k)$$

$\rightsquigarrow$ $g(z)$ contains all information about $\rho(\lambda)$ (via its moments)

Stieltjes Transform of the Wigner Ensemble

Sketch of the approach:

1) Find recurrence relation between g_N and g_{N-1}

↳ Fixed point equation for g as $N \to \infty$

[Physics speech: Cavity method, self consistent equations]

2) Use Stieltjes inversion formula to approximate $\rho(\lambda)$

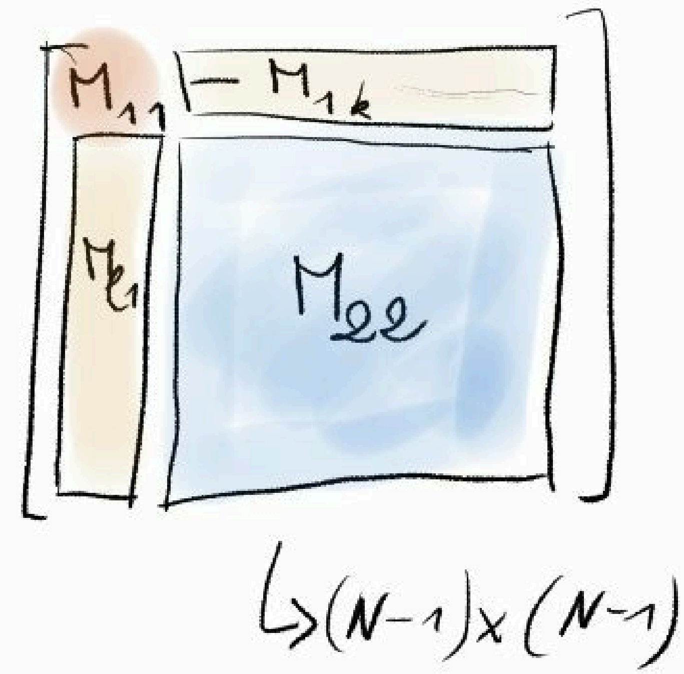
[to be defined later]

$$\begin{tikzcd}
\rho_N(\lambda) \arrow[r, "?"] \arrow[d, "\text{Stieltjes Transform}"'] & \rho(\lambda) \\
g_N(z) \arrow[r, "\text{self-consistent equation}"'] & g(z) \arrow[u, "\text{Stieltjes Inversion Formula}"']
\end{tikzcd}$$

Let $X \in \mathbb{R}^{N \times N}$ be a Wigner matrix, $M = zI - X$

using the Schur complement formula [See Chapter 1]

$$(*) \qquad \frac{1}{(G_X)_{11}} = M_{11} - \sum_{k,l}^{N} M_{1k} (M_{22})^{-1}_{kl} M_{l1}$$



- $\mathbb{E}[M_{11}] = z$
- $\mathbb{E}_{X_{1i}}\left[M_{1i} (M_{22})^{-1}_{ij} M_{1j}\right] = \frac{\sigma^2}{N} (M_{22})^{-1}_{ii} \delta_{ij}$

$$\Rightarrow \mathbb{E}_{X_{1i}}\left[\sum_{k,l}^{N} M_{1k} (M_{22})^{-1}_{kl} M_{l1}\right] = \frac{\sigma^2}{N} \operatorname{tr}\left((M_{22})^{-1}\right)$$

- $\frac{1}{N-1} \operatorname{tr}\left((M_{22})^{-1}\right)$ = Stieltjes transform of a Wigner matrix of size $N-1$ and variance $\frac{\sigma^2 (N-1)}{N}$

$$\Rightarrow \mathbb{E}\left[\frac{1}{N} \operatorname{tr}\left((M_{22})^{-1}\right)\right] \longrightarrow g(z)$$

By the previous points (+ some handwaving)

$\frac{1}{(G_x)_{11}}$ approaches a deterministic limit as $N \to +\infty$

$$\Rightarrow \quad \mathbb{E}\left[\frac{1}{(G_x)_{11}}\right] = \frac{1}{\mathbb{E}\left[(G_x)_{11}\right]}$$

And by rotational invariance [$\Rightarrow$ invariance under permutations]

the diagonal elements of G_x have the same expectation

$$\mathbb{E}\left[(G_x)_{11}\right] = \frac{1}{N}\mathbb{E}\left[\operatorname{tr}(G_x)\right] = \mathbb{E}[g_N] \to g$$

So $(*)$ becomes as $N \to \infty$

$$\frac{1}{g(z)} = z - \sigma^2 g(z) \quad \Rightarrow \quad g(z) = \frac{z \pm \sqrt{z^2 - 4\sigma^2}}{2\sigma^2}$$

$$= \frac{z \pm z\sqrt{1 - \frac{4\sigma^2}{z^2}}}{2\sigma^2}$$

We have two solutions for g, but only

$$g(z) = \frac{z - z\sqrt{1 - 4\sigma^2/z^2}}{2\sigma^2} \underset{z\to\infty}{\sim} \frac{1}{z}$$

is consistent with

$$g(z) = \sum_{k=0}^{+\infty} \frac{1}{z^{k+1}} \tau(A^k)$$

Near the singularities of the function

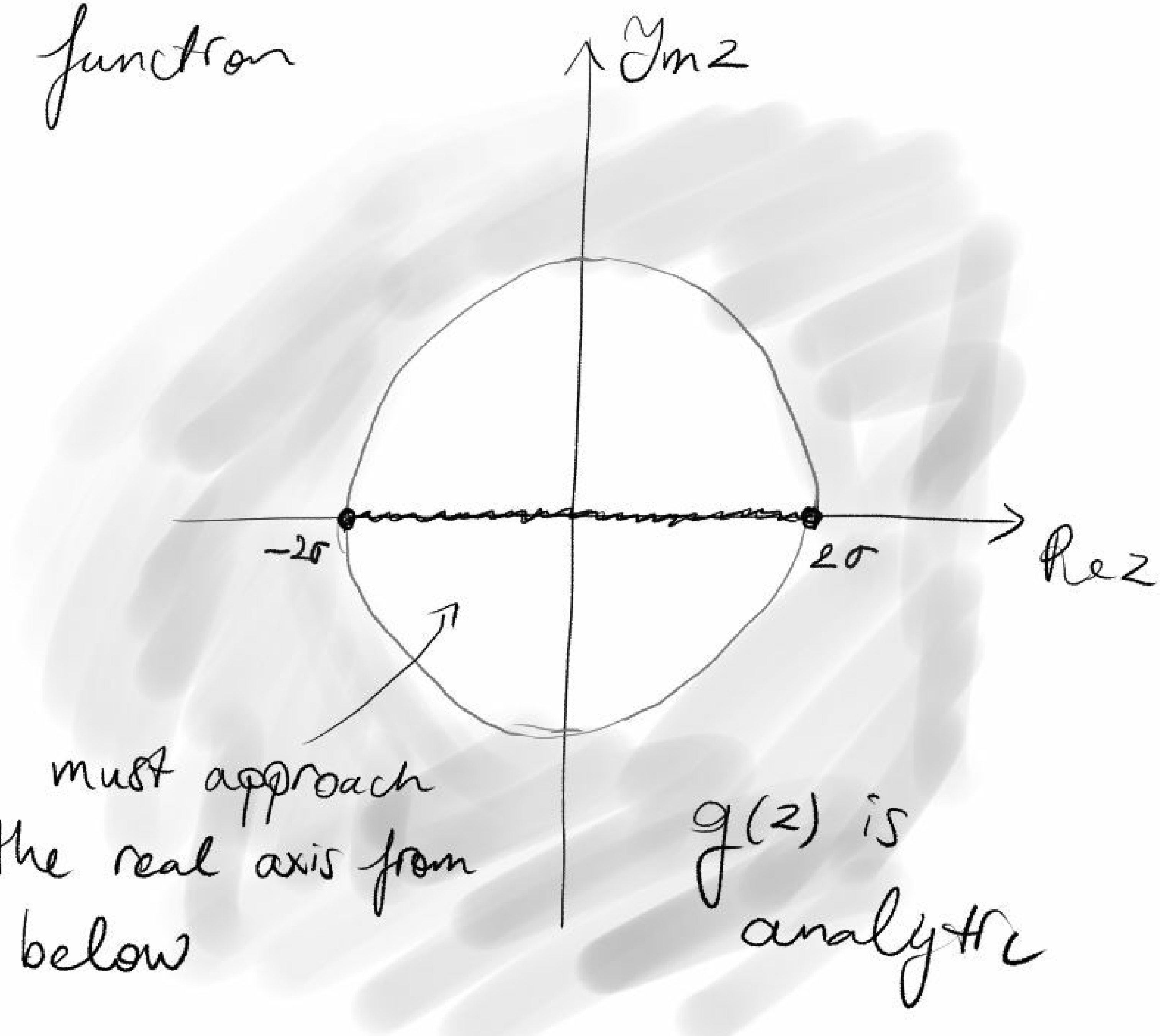
For finite N:

$$g_N(z) = \sum_{k=1}^{N} \frac{1}{z-\lambda_k}$$

poles at λ_k

$$\to \sum_{k=1}^{N} \delta(z-\lambda_k) \sim \rho(z)$$

$$g(z) = \int_{\mathrm{supp}(\rho)} \frac{\rho(x)}{z-x}\, dx$$



Approaching $\rho(x)$ using g

For $z = x - i\eta$ w/ $x \in \operatorname{supp}\{\rho\} = [-2\sigma, 2\sigma]$, $\eta > 0$ small

$$g_N(x - i\eta) = \frac{1}{N}\sum_{k=1}^{N} \frac{1}{x - i\eta - \lambda_k} = \frac{1}{N}\sum_{k=1}^{N} \frac{x - \lambda_k + i\eta}{(x-\lambda_k)^2 + \eta^2}$$

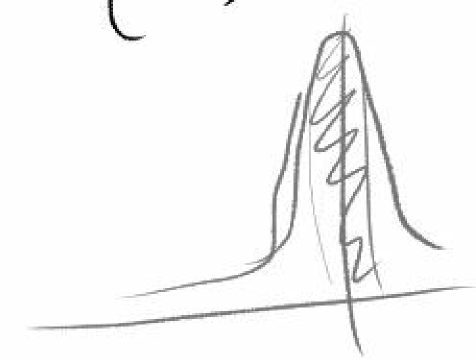
$$\operatorname{Im}\left(g_N(x - i\eta)\right) = \frac{1}{N}\sum_{k=1}^{N} \frac{\eta}{(x-\lambda_k)^2 + \eta^2}$$

$\rightsquigarrow$ Convolution of $\rho_N(x) = \sum_{k=1}^{N} \delta(x - \lambda_k)$ with $\pi K_\eta(x) = \frac{\eta}{x^2 + \eta^2}$

- For η too small ($< \frac{1}{N}$):
  too few eigenvalues contribute to $\operatorname{Im}(g_N(x - i\eta))$ $\to$ too much error

- For $\eta \sim \frac{1}{\sqrt{N}}$ roughly $n \sim N\rho(x)\Delta x$ eigenvalues in interval of size Δx

$$\frac{1}{N}\sum_{k : \lambda_k \in [x-\Delta x, x+\Delta x]} \frac{i\eta}{(x-\lambda_k)^2 + \eta^2} \longrightarrow \int_{x+\Delta x}^{x+\Delta x} \frac{\rho(x)\,\eta\, dy}{(x-y)^2 + \eta^2} \longrightarrow i\pi\rho(x)$$



Stieltjes inversion formula

We can explicitly recover $\rho(x)$ as

$$\lim_{\eta \to 0^+} \operatorname{Im} g(x - i\eta) = \pi \rho(x)$$

For the Wigner distribution, as $\eta \to 0$

$$g(x - i\eta) \to \frac{x - \sqrt{x^2 - 4\sigma^2}}{2\sigma^2}$$

only has an imaginary part if $\sqrt{x^2 - 4\sigma^2}$ is imaginary

$$\Rightarrow \operatorname{Im} g(x - i\eta) \to \frac{\sqrt{4\sigma^2 - x^2}}{2\sigma^2 \pi}$$

$x \in [-2\sigma; 2\sigma]$

Beyond Wigner Matrices:

- Gaussian Unitary Ensemble
  Gaussian Symplectic Ensemble
  } "Symmetric" matrices w/ Complex / quaternion normally distributed entries

  $\rightsquigarrow$ Unitary / Symplectic invariant

generalized eigenvalue density

$$\rho_\beta(\lambda) = \frac{\sqrt{4\sigma^2 - \lambda^2}}{4\pi}$$

$\beta = 1$: GOE
$\beta = 2$: GUE
$\beta = 4$: GSE

- Ginibre ensemble: unsymmetric $X_{ij} \sim N(0, \frac{\sigma^2}{N})$

  $\rightarrow$ Girko circular law $\lambda \sim U[\{z \in \mathbb{C} \mid |z| < \sigma\}]$

- Wishart ensemble $X = HH^T$ $\rightsquigarrow$ singular values $\sim \frac{\sqrt{4\sigma^2 - s^2}}{\pi\sigma^2}$