More generally let $X$ and $Y$ be Banach spaces and let $A: X \to Y$, $B: Y \to X$ be bounded linear operators. Then the following commutation formula is true:

\[(75.1) \quad \frac{1}{\lambda + AB} + \frac{1}{\lambda + BA} = I_Y\]

in the sense that if $0 \neq -\lambda \in \rho(BA) = \text{resolvent set of } BA$, then $-\lambda \in \rho(AB)$ and

\[\frac{1}{2} \left( I_Y - A \frac{1}{\lambda + BA} B \right)\]

is the resolvent of $AB$, and vice versa. In particular we see that (75.1) is true for bounded operators in Banach spaces. (75.1) is also true, suitably interpreted, for certain classes of unbounded operators (see Deift [DjM, 1987]).
We now prove (71.2). Let $\mu$ and $f, g, \in L^2(\mu)$. Then

$$
\int \cdots \int \frac{f(x_1)}{g_1(x_1) \cdots g_N(x_N)} \frac{g(x_1)}{g_1(x_1) \cdots g_N(x_N)} \mu(x_1) \cdots \mu(x_N)
$$

$$
= \sum_{\sigma, (i) \in S_N} \mu_\sigma \mu (f_\sigma \cdots f_N(x_N)) \mu_\sigma (g_1(x_1) \cdots g_N(x_N)) \mu(x_1) \cdots \mu(x_N)
$$

$$
= \sum_{\sigma, (i) \in S_N} \mu_\sigma \mu \left[ \int f_{\sigma_0 \sigma_1}^{-1}(x_1) g_{\sigma_1}^{-1}(x_1) \mu(x_1) \right] \cdots \left[ f_{\sigma_0 \sigma_N}^{-1}(x_N) g_{\sigma_N}^{-1}(x_N) \mu(x_N) \right]
$$

$$
\times \cdots \times \left[ \int f_{\sigma_0 \sigma_N}^{-1}(x_N) g_{\sigma_N}^{-1}(x_N) \mu(x_N) \right]
$$

$$
= \sum_{\sigma} \sum_{\sigma_0 \sigma_1 : \tau \in S_N} \mu_\sigma \mu \left[ \int f_{\sigma_0 \sigma_1}^{-1}(x_1) g_{\sigma_1}^{-1}(x_1) \mu(x_1) \right] \cdots \left[ \int f_{\sigma_0 \sigma_N}^{-1}(x_N) g_{\sigma_N}^{-1}(x_N) \mu(x_N) \right]
$$

$$
\times \cdots \times \left[ \int f_{\sigma_0 \sigma_N}^{-1}(x_N) g_{\sigma_N}^{-1}(x_N) \mu(x_N) \right]
$$

$$
= \sum_{\sigma} \sum_{\sigma_0 \sigma_1 : \tau \in S_N} \mu_\sigma \mu \left[ \int f_{\sigma_0 \sigma_1}^{-1}(x_1) g_{\sigma_1}^{-1}(x_1) \mu(x_1) \right] \cdots \left[ \int f_{\sigma_0 \sigma_N}^{-1}(x_N) g_{\sigma_N}^{-1}(x_N) \mu(x_N) \right]
$$

$$
\times \cdots \times \left[ \int f_{\sigma_0 \sigma_N}^{-1}(x_N) g_{\sigma_N}^{-1}(x_N) \mu(x_N) \right]
$$

$$
= \sum_{\sigma} \sum_{\tau : \tau \in S_N} \mu_\sigma \mu \left[ \int f_{\tau_1}^{-1}(x_1) g_1^{-1}(x_1) \mu(x_1) \right] \cdots \left[ \int f_{\tau_N}^{-1}(x_N) g_N^{-1}(x_N) \mu(x_N) \right]
$$
\[
= \sum_{\delta} \det \left( \int f_i(x) \ g_h(x) \ q(x) \right)_{\delta, h = 1, \ldots, N}
\]

\[
= N! \ \det \left( \int f_i(x) \ g_h(x) \ p(x) \right)_{\delta, h = 1, \ldots, N}
\]

as desired.

The particular case

\[
q(x) = \sum_{k=1}^{N} \delta_{x, x_k}(x)
\]

for a given set of numbers \(x_1, \ldots, x_m\). Then the LHS of (11.2) becomes

\[
\sum_{k_1=1}^{N} \cdots \sum_{k_N=1}^{N} \int \det \left( f_1(x_1) \cdots f_i(x_N) \right) \det \left( g_1(x_1) \cdots g_i(x_N) \right) \delta(x_1 - x_{k_1}) \cdots \delta(x_N - x_{k_N})
\]

\[
= \sum_{k_1=1}^{N} \cdots \sum_{k_N=1}^{N} \det \left( f_1(x_k) \cdots f_i(x_{kN}) \right) \det \left( g_1(x_k) \cdots g_i(x_{kN}) \right)
\]
\[
\sum_{1 \leq k_1 < k_2 < \ldots < k_N \leq M} \det \left( \begin{array}{c}
F_n(\lambda_{k_1}) & \cdots & F_n(\lambda_{k_N}) \\
\vdots & & \vdots \\
F_N(\lambda_{k_1}) & \cdots & F_N(\lambda_{k_N})
\end{array} \right) \\
\times \det \left( \begin{array}{c}
g_1(\lambda_{k_1}) & \cdots & g_1(\lambda_{k_N}) \\
\vdots & & \vdots \\
g_N(\lambda_{k_1}) & \cdots & g_N(\lambda_{k_N})
\end{array} \right)
\]

On the other hand, the roots of (71.2) gives

\[
N! \sum_{j=1}^{N} \det \left( \sum_{i=1}^{N} f_j(\lambda_i) g_k(\lambda_i) \right)
\]

Evaluating these expressions we see that

\[
F = (F_{ij}) = (f_i(\lambda_j)) = \begin{pmatrix} F_1(\lambda_1) & \cdots & F_n(\lambda_M) \\ \vdots & \ddots & \vdots \\ F_N(\lambda_1) & \cdots & F_N(\lambda_M) \end{pmatrix} = N \times M
\]

\[
G = (G_{ij}) = (g_j(\lambda_i)) = \begin{pmatrix} g_1(\lambda_1) & \cdots & g_N(\lambda_M) \\ \vdots & \ddots & \vdots \\ g_N(\lambda_1) & \cdots & g_N(\lambda_M) \end{pmatrix} = M \times N \times M
\]
we obtain the classical Cauchy-Binet formula

\[(79.1) \quad \det FG = \sum_{1 \leq j_1 < \cdots < j_N \leq M} \det \begin{pmatrix} F_{j_1 1} & F_{j_1 2} & \cdots & F_{j_1 N} \\ F_{j_2 1} & F_{j_2 2} & \cdots & F_{j_2 N} \\ \vdots & \vdots & \ddots & \vdots \\ F_{j_N 1} & F_{j_N 2} & \cdots & F_{j_N N} \end{pmatrix} \det \begin{pmatrix} G_{1 j_1} & G_{1 j_2} & \cdots & G_{1 j_N} \\ G_{2 j_1} & G_{2 j_2} & \cdots & G_{2 j_N} \\ \vdots & \vdots & \ddots & \vdots \\ G_{N j_1} & G_{N j_2} & \cdots & G_{N j_N} \end{pmatrix} \]

\[\text{eg: if } F = \begin{pmatrix} a_1 & a_2 & a_3 \\
1 & b_2 & b_3 \end{pmatrix}, \quad G = \begin{pmatrix} c_1 & d_1 \\
1 & c_2 & c_3 \\
1 & c_4 & d_3 \end{pmatrix}\]

\[\det \begin{pmatrix} a_1 & a_2 & a_3 \\
b_1 & b_2 & b_3 \end{pmatrix} \begin{pmatrix} c_1 & d_1 \\
c_2 & c_3 \end{pmatrix} = \det \begin{pmatrix} a_1 & a_3 \\
b_1 & b_3 \end{pmatrix} \det \begin{pmatrix} c_1 & d_1 \\
c_2 & c_3 \end{pmatrix} + \det \begin{pmatrix} a_1 & a_2 \\
b_1 & b_2 \end{pmatrix} \det \begin{pmatrix} c_2 & d_1 \\
c_3 & c_3 \end{pmatrix} + \det \begin{pmatrix} a_1 & a_2 \\
b_1 & b_3 \end{pmatrix} \det \begin{pmatrix} c_1 & d_2 \\
c_4 & c_3 \end{pmatrix} \]

Of course if \(N = 1\), (79.1) is just the familiar fact that

\[\det FG = \det F \det G\]

and for \(N > 1\) both LHS and RHS are 0 [why?]
We will also need the following basic definition.

Let \( m(x) \) be a Borel measure on \( \mathbb{R} \) with all moments finite.

\[
\int |x|^m \, m(x) < \infty, \quad m = 0, 1, 2, \ldots
\]

Suppose that the support of \( m \) is infinite, i.e., \( m \) is not a finite linear combination of delta functions. Then by the Gram-Schmidt procedure

there exist unique monic polynomials

\[
\pi_h(x) = x^h + \cdots, \quad h \geq 0
\]

which are orthogonal with respect to \( m \):

\[
\int_{\mathbb{R}} \pi_h(x) \pi_i(x) \, m(x) = 0 \quad \text{for } i \neq h.
\]

For

\[
\delta_h = \left( \int_{\mathbb{R}} \pi_h(x) \, m(x) \right)^{-\frac{1}{2}} > 0, \quad h = 0, 1, 2, \ldots
\]

set
\( p_k(x) = \delta_k \sigma_k(x), \quad k \geq 0. \)

The polynomials \( \{p_k\} \) are the orthonormal polynomials with respect to \( q(x) \), i.e.,

\[
\int_{\mathbb{R}} p_i(x) p_j(x) q(x) \, dx = \delta_{ij}, \quad i, j = 0, 1, 2, \ldots
\]

The \( p_k \)'s are called the orthonormal polynomials (wrt \( q(x) \)).

and the \( \pi_k \)'s are called the monic orthogonal polynomials (wrt \( q(x) \)).

Exercise: what happens if \( q(x) \) has finite support?

As we will see, the \( p_k \)'s and \( \pi_k \)'s play a central role in RMT.

We now show how to compute a statistic for an invariant unitary ensemble with probability distribution

\[
P_N(m) \, dm = e^{-\frac{1}{4} Q(x_{\infty})} \, dm
\]

where \( Q(x_{\infty}) \to \infty \) sufficiently rapidly as \( x_{\infty} \to \infty \).
In particular, we will compute
\[ \langle f \rangle = \int f(m) P_n(m) \, dm \]
for \( f(m) \) of the form
\[ (82.0) \quad f(m) = \det(I + g(m)) \]
for bounded functions \( g: \mathbb{R} \to \mathbb{R} \).

From the Hermitian analog of (68.1) we have
\[ (82.1) \quad \langle f \rangle = C_N \int \prod_{i=1}^{N} \left( 1 + g(x_i) \right) \prod \left( \lambda_i - x_i \right)^2 \prod_{1 \leq i < j \leq N} W(x_i) \, d^N x \]
\[ = C'_N \int \prod \left( 1 + g(x_i) \right) \prod \frac{1}{\prod W(x_i) \, d^N x} \]
where
\[ (82.2) \quad W(x) = e^{-Q(x)} \]
\[ (82.3) \quad V(\lambda) = \text{vandermonde} = \det \left( \lambda_i^{j-1} \right)_{1 \leq i, j \leq N} \]
and
\[ (82.4) \quad C'_N = \left( \int \prod \frac{1}{\prod W(x_i) \, d^N x} \right)^{-1} \]

As the integrands in (82.1) and (82.4) are
Invariant under all permutations $\sigma$

$$(\lambda_1, \ldots, \lambda_N) \rightarrow (\sigma(\lambda_1), \ldots, \sigma(\lambda_N))$$

We may rewrite (82.1) \((82.4)\) in the following symmetrized form

$$\langle \phi \rangle = C_N \int \prod_{i=1}^{N} (1 + h(\lambda_i)) V(\lambda) \prod_{i=1}^{N} w(\lambda_i) \, d\lambda$$

where

$$C_N = C'_N (N!)^{-1}$$

Now define

$$f_i(x) = g_i(x) = x^{1-1} \quad 1 \leq i \leq N$$

Then

$$V(\lambda)^2 = \det (f_i(\lambda_j)) \quad \det (g_i(\lambda_j))$$

Hence we can apply the generalized Cauchy-Binet formula to (83.1) with measure $du(x) \leq w(x) (1 + g(x))$.

The fact that $g(x)$ is a signed measure, as opposed to a pos. measure, clearly does not affect the proof.
of (83.11). We obtain

\[ \langle \mathbf{r} \rangle = c_n \text{ det } \left( \int f_i(\lambda) g_k(\lambda) q_{n}(\lambda) \right)_{i,k=1}^n \]

where we note from (82.4) that \( c_n \) depends only on \( N \) and \( \omega \), but not on \( q(x) \).

We have

\[ (84.1) \quad \int f_i(\lambda) g_k(\lambda) q_n(\lambda) = \int f_i(\lambda) g_k(\lambda) \left( 1 + g(\lambda) \right) w(\lambda) d\lambda. \]

Let \( p_j(\lambda), j \geq 0, \) be the orthogonal polynomials with respect to the measure \( w(\lambda) d\lambda \) in \( R \).

\[ \int p_i(\lambda) \ p_j(\lambda) \ w(\lambda) d\lambda = \delta_{ij}, \ i, j \geq 0. \]

(Note that \( w(\lambda) d\lambda = e^{-Q(\lambda)} d\lambda \) does not have finite support, also by "sufficient decay" we assume, at least, that \( e^{-Q(\lambda)} d\lambda \) has finite moments. Not that this implies \( \int (V(\lambda))^{-1} w(\lambda) d\lambda < \infty \).)
so that \( P_n(x) dx \) is a normalized prob. measure with finite moments).

Now as we can add rows and columns without changing a determinant, we have (recall \( P_\lambda(x) = \pi_\lambda(x) \))

\[
(85.1) \quad \langle \phi \rangle = C_N \det \left( \int \prod_{j=0}^{N-1} \pi_j(x) \left( 1 + q(x) \right) w(x) dx \right)_{j,k \in \mathbb{N}-1, 0 \leq j \leq k \leq N-1}
\]

Set

\[
(85.2) \quad \phi_j(x) = P_j(x) w(x)^{1/2}, \quad j \geq 0
\]

have \( \int \phi_j(x) \phi_k(x) dx = \delta_{j,k} \) for \( j, k > 0 \).

Then (85.1) takes the form

\[
(85.3) \quad \langle \phi \rangle = C_N' \det \left( \int \prod_{j=0}^{N-1} \phi_j(x) \left( 1 + q(x) \right) dx \right)_{j,k \in \mathbb{N}-1, 0 \leq j \leq k \leq N-1}
\]

\[
= C_N'' \det \left( \delta_{j,k} + \int q(x) \phi_j(x) \phi_k(x) dx \right)_{j,k \in \mathbb{N}-1}
\]

where again \( C_N'' \) does not depend on \( q(x) \). Setting \( q = 0 \),
we have \( f(m) = 1 \) and no LHS of (85.3) = 1.

But \( \text{RHS} = C_N \times \text{det} (\delta_{j,k}) = C_N \), no \( C_N" = 1 \).

Thus

\[
\langle \phi \rangle = \text{det} \left( \delta_{j,k} + \int \phi_j(x) \phi_k(x) g(x) \, dx \right)^{N-1}.
\]

(86.1)

Now we are primarily interested in the situation where the size \( N \) of the matrices become large. We see from (86.1) that \( \langle \phi \rangle \) is expressed in terms of determinants of larger and larger size. Limits of this kind are generally very difficult to control.

Fortunately we can use (71.1) \( \text{det} (I + AB) = \text{det}(I + BA) \) to reduce (86.1) to the (Fredholm) determinant on a fixed space (see below). Such limits are, generally speaking, easier to control.
particular if $k_n$ is a trace class operator on a Hilbert space $H$, and $k_n \to K$ in trace norm, then

$$\det (1 + k_n) \to \det (1 + K)$$

It is just a matter of continuity of the determinant in the trace norm.

Now let $A : L^2(\mathbb{R}) \to \mathbb{C}^N$ denote the bounded operator

$$\tag{87.1} (Ah)_j = \int \phi_j(x) g(x) h(x) \, dx, \quad j = 0, \ldots, N-1, \quad h \in L^2(\mathbb{R})$$

and let $B : \mathbb{C}^N \to L^2(\mathbb{R})$ denote the bounded operator

$$\tag{87.2} (Ba)(x) = \sum_{j=0}^{N-1} \phi_j(x) a_j, \quad a = (a_0, \ldots, a_{N-1})^T \in \mathbb{C}^N$$

Then $AB$ maps $\mathbb{C}^N \to \mathbb{C}^N$ and for $a \in \mathbb{C}^N$

$$\tag{87.3} ((AB)a)_j = \left( A \left( \sum_{k=0}^{N-1} \phi_k(x) a_k \right) \right)_j = \sum_{k=0}^{N-1} a_k (A \phi_k)_j.$$
\[
\sum_{k=0}^{N-1} a_k \left( \int \phi_j(x) s(x) \phi_k(x) \, dx \right)
\]

and we see from (86.1) that

\[(88.1) \quad \langle \Phi \rangle = \det \left( 1 + A \hat{B} \right) \]

On the other hand \(BA : L^2(\mathbb{R}) \to L^2(\mathbb{R})\) and for \(h \in L^2(\mathbb{R})\)

\[(BAh)(x) = (B \left( \sum_{j=0}^{N-1} \int \phi_j(y) g(y) h(y) \, dy \right)) (x) = \sum_{j=0}^{N-1} \int \phi_j(y) g(y) h(y) \, dy = \int k(x,y) g(y) h(y) \, dy \]

where

\[(88.1) \quad k(x,y) = \sum_{j=0}^{N-1} \phi_j(x) \phi_j(y) \]

is the so-called correlation kernel for the ensemble. \(k(x,y)\) plays a central role in RMT.

Thus

\[(88.3) \quad BA = k x_g\]

where \(x_g\) denotes multiplication by \(g(x)\) and \(k\) denotes
The operator with kernel $K$ is $K(x) = \int k(x, y) dy$.

Assembling the above results we obtain the key formula

$$<f> = \text{act} \left( L^2(\mathbb{R}) + K \mathcal{N}_T \right)$$

Note that as $A$ (and also $B$!) is finite rank, $A$ is trace class and the Fredholm determinant in (89.1) is well-defined. There are analogies, but more complicated, formulae for $\beta = 1$ and $\beta = 4$ (see Ref. 31).

If $\mathcal{N}$ is a Borel set in $\mathbb{R}$ and

$$g(x) = -A^{-1}$$

then we see from (89.1) that the gap probability considered before is given by

$$\text{Prob} \left( \text{no eigenvalues in } \mathcal{N} \right) = <\sum_{n} (1 - \chi_n(x)) >$$  

$$= \text{det} \left( L^2(\mathbb{R}) + K \mathcal{N}_T \right)$$

Note that if $\mathcal{N} = \emptyset$, then $\text{RHS} = 1$ i.e. $\mathcal{N}$
\text{Prob} \{ \text{no eigs in } \phi_j \} = 1, \quad \text{as it should be. On the other hand for } \mathbb{R}, \quad \text{Prob} \{ \text{no eigs in } \mathbb{R} \} = 0.

On the other hand for any \( k = 0, 1, \ldots, N-1 \),

\[ k \phi_k(x) = \int_0^1 \sum_{j=0}^{N-1} \phi_j(x) \phi_j(y) \phi_k(y) \, dy \]

\[ = \sum_{j=0}^{N-1} \delta_{j,k} \phi_j(x) \]

\[ = \phi_k(x) \]

Thus \( 1 \) is an eigenvalue of \( k \) (of multiplicity \( N > 0 \)) and so \( \text{det}(1 - k) = 0 \), as it should be.

Now take \( \mathbb{R} = (a, \infty) \) for any \( a \in \mathbb{R} \).

Then clearly

\[ \text{Prob} \{ \text{no eigs in } (a, \infty) \} = \text{Prob} \{ \lambda_{\text{max}} \leq a \} \]

where \( \lambda_{\text{max}} \) is the largest eigenvalue of \( k \).

\begin{equation}
(90.1) \quad \text{Prob} \{ \lambda_{\text{max}} \leq a \} = \text{det} \left( I - \lambda_{\text{max}} \mathbf{1}_{\mathbb{R}} \right)
\end{equation}

This is a key formula in RMT for unitary
ensembles.

**Exercise**

\[ \emptyset = S \cup U \cdots \cup U_{k} \text{ in a union of} \]

disjoint sets in \( \mathbb{R} \), \( S \cap U_{j} = \emptyset \text{ for } j \neq k \),

Then we have computed

\[ \text{Prob} (\text{o ergs in } S, \ldots, \text{o ergs in } S_{k}) = \text{Prob} (\text{no ergs in } S = \cup U_{i}) \]

Derive an explicit formula for

\[ \text{Prob} (\text{n1 ergs in } U_{1}, \ldots, \text{n}_{k} \text{ ergs in } U_{k}) \]

for any non-negative integers \( n_{1}, \ldots, n_{k} \). (see ref. [3], pp 86-88)

We now show how to compute

**correlation functions** for unitary ensembles (recall that 2-point correlation functions arose in the study...
of the non-trivial zeros of the Riemann zeta function in the first lecture.

If \( P_N(x) \, dx \) is the probability distribution for a system of \( N \) identical random particles,

\[
P_N(x_1, \ldots, x_N) \text{ symmetric in the } x_i's
\]

then the \( n \)-point correlation function

\[
R_n = R_n(x_1, \ldots, x_n), \quad 1 \leq n \leq N
\]

for \( P_N(x) \, dx \) is defined by

\[
R_n(x_1, \ldots, x_n) = \frac{N!}{(N-n)!} \int \cdots \int P_N(x_1, \ldots, x_n, x_{n+1}, \ldots, x_N) \, dx_{n+1} \cdots dx_N
\]

Note that \( \int R_n(x_1, \ldots, x_n) \, dx_1 \cdots dx_n \neq 1 \) and hence \( R_n \) is not a probability distribution (more later!).

Correlation functions are useful whenever we want to focus on \( n \leq N \) particles and "average out"