Orthogonal polynomials, random matrices and wireless communications

Brunel University
17-December 2010.

Matthew McKay
Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong

Yang Chen
Imperial College London
United Kingdom
ychen@ic.ac.uk

• “It is not clear what we mean when we say Painlevé equations are integrable”

——Dinner conversation with a colleague.

$y \quad y(/5y \qquad y$

- Information-theoretic quantities in singer and

• The mathematical problem is the computation of the Hankel determinant generated by a deformation of classical weights:

• $w(x) = x^{\alpha} e^{-x} (x + t)^{\lambda}$,

$0 \le x < \infty, \ \alpha > 0, \ t > 0,$

deformed Laguerre, single-user

• $w(x) = x^{\alpha_1}(1 - x)^{\alpha_2} \left(\frac{x+t}{1-x}\right)^{\lambda}$,

$x \in (0, 1), \ \alpha_1 > 0, \ \alpha_2 > 0, \ t > 0,$

deformed Jacobi, multi-users

• Main results. Ladder operators approach to orthogonal polynomials shows that the Hankel determinants in the single user case a particular Painlevé V, and in the multi-users case a particular Painlevé VI.

- Multiple-input multiple-output (MIMO) systems have been at the forefront of wireless communications research and development, e.g., next-generation wireless local area networks (WLAN) and cellular mobile networks. The main reason for this explosion of interest is mainly due to the independent discoveries of Telatar (1990, 1999, 2008, 7157) and Foschini (1990), demonstrated that the fundamental information-theoretic capacity of MIMO systems grows *linearly* with the number of antennas.

- Traditional methods give logarithmic capacity increase (in P). MIMO is a key technology for meeting the ever-increasing demands for higher-rate data-oriented wireless communications applications and services.

• *Ergodic Capacity*, specifies the maximum achievable average mutual information between the transmitter and receiver, assumes that a user's codeword span a large number of "independent channels". Mathematically the expectation value of a certain random variable.

• *Outage Capacity*. Characterizing the communication limits of systems which are not

Methods for studying the outage capacity, which boils down to the Hankel determinant

$$D_n = \det\left(\mu_{i+j}\right)_{i,j=0}^{n-1}$$

generated from the moments of a certain weight function $w(x)$,

$$\mu_k := \int_a^b x^k w(x)dx\ , \quad k$$

• Two di□erent methods from random matrix theory to compute Hankel determinants.

1. Exact expressions employing orthogonal polynomials and ladder operators (Belmehdi,Bonan,Clark,Lubinsky,Magnus...) Logarithmic derivative of Hankel determinants are the $\sigma-$function of certain P_V (Single User) and P_{VI} (Multi User).

2. Large n approximations for these determinants by employing the general linear statistics theorems.
(Essentially Szegö limit theorem.)

•Closed-form expressions for the distribution function. Valid for large dimension; approximations are remarkably accurate for even very small matrix dimensions (2×2).

• Comparison of large n with Painlevé

•Single-user MIMO (deformed Laguerre)
$n_r = n_t$, or $\alpha = 0$. Coulomb fluid gives the distribution of the mutual information, **a Gaussian**, to leading order in n.

• Use P_V to compute the large-n correction terms for the mean, variance, and third cumulant.

• Deviations from Gaussian as P (SNR) increases. Sensitivities of mean, variance, and third moment, with respect to P.

• MIMO communication system has n_t transmit and n_r receive antennas.

• linear model:

Transmitted vector $\mathrm{x}_{n_t \times 1} \quad \mathrm{C}^{n_t}$

Received vector $\mathrm{y}_{n_r \times 1} \quad \mathrm{C}^{n_r}$,

$$\mathrm{y} = \mathrm{H}\mathrm{x} + \mathrm{n} \,.$$

• $\mathrm{n}_{n_r \times 1} \quad \mathrm{C}^{n_r}$ is a complex Gaussian vector

• $E(\mathrm{n}) = 0, \;\; E(\mathrm{n}\mathrm{n}^{\dagger}) = \mathrm{Q}_{\mathrm{n}}$.

• Covariance matrix account for receiver noise and multi-user interference; choice of Q_{n} distinguish between single-user and multi-user MIMO models.

• $\mathrm{H} \quad \mathrm{C}^{n_r \times n_t}$,*channel matrix*, is stochastic, known to the receiver not to the transmitter.

• H complex Gaussian, i.i.d. elements, zero mean and unit variance.(Simplest choice).

• x subject to a power constraint:

$$E(\mathrm{x}^{\dagger}\mathrm{x}) \leq P \,.$$

•Shannon capacity gives highest data rate achievable with negligible errors by any transmission scheme.

Mutual Information between the input and output signals,

$$
\begin{aligned}
I(\mathbf{x};\mathbf{y}|\mathbf{H}) &:= H(\mathbf{y}|\mathbf{H}) - H(\mathbf{y}|\mathbf{x},\mathbf{H}) \\
&= H(\mathbf{y}|\mathbf{H}) - H(\mathbf{n})
\end{aligned}
$$

• $H(\mathbf{y}|\mathbf{H})$ conditional entropy of $\mathbf{y}$, defined by its density $p(\mathbf{y}|\mathbf{H})$:

$$H(\mathbf{y}|\mathbf{H}) = E(-\log p) := -\int_{\mathbb{C}^{n_r}} p(\mathbf{y}|\mathbf{H}) \log p(\mathbf{y}|\mathbf{H}) d\mathbf{y}.$$

• Ergodic capacity (C) relevant for highly dynamic channels; high-mobility wireless applications; $\mathbf{H}$ varies quickly over time, each transmission codeword sees a large number of "independent" channel realizations.

$$C = \max_{p(\mathbf{x})} E_{\mathbf{H}} \left(I(\mathbf{x};\mathbf{y}|\mathbf{H}) \right)$$

where the maximum is taken over all densities $p(\mathbf{x})$ of the input vector $\mathbf{x}$, subject to the power constraint.

- Telatar proved that the optimal input density $p(\mathbf{x})$ is multi-variate complex Gaussian with zero mean.

$$I(\mathbf{x};\mathbf{y}|\mathbf{H}) = \log\det\left(\mathbf{I}_{n_r} + \mathbf{H}\mathbf{Q}\mathbf{H}^{\dagger}\mathbf{Q}_{\mathbf{n}}^{-1}\right)$$

where $\mathbf{Q} = E(\mathbf{x}\mathbf{x}^{\dagger})$ is the input signal covariance.

- **Capacity**

$$C = \max_{\mathbf{Q}\geq 0} E_{\mathbf{H}}\left(I(\mathbf{x};\mathbf{y}|\mathbf{H})\right)$$

subject to $\mathrm{tr}(\mathbf{Q}) \leq P$.

- For this model,

$$\mathbf{Q} = \frac{P}{n_t}\mathbf{I}_{n_t}.$$

In other words, the capacity is achieved by sending independent Gaussian signals from each of the transmit antennas with equal power.

- Outage probability P_{out}

$$
\begin{aligned}
P_{\text{out}}(C_{\text{out}}) &= \Pr\left(I(\mathbf{x};\mathbf{y}) < C_{\text{out}}\right) \\
&= \Pr\left(\log\det\left(\mathbf{I}_{n_r} + \frac{P}{n_t}\mathbf{H}\mathbf{H}^{\dagger}\mathbf{Q}_{\mathbf{n}}^{-1}\right) < C_{\text{out}}\right)
\end{aligned}
$$

with $\mathbf{Q}$ denoting the input covariance which maximizes the mutual information. The outage probability can be calculated via

$$
P_{\text{out}}(C_{\text{out}}) = \frac{1}{2\pi}\int_{-\infty}^{\infty} M(i\omega)\frac{1 - e^{-i\omega C_{\text{out}}}}{i\omega}d\omega ,
$$

where $M(\cdot)$ the moment generating function of the mutual information

$$
\begin{aligned}
M(\lambda) &:= E_{\mathbf{H}}\left(\exp\left(\lambda I(\mathbf{x};\mathbf{y}|\mathbf{H})\right)\right) \\
&= E_{\mathbf{H}}\left(\det\left(\mathbf{I}_{n_r} + \frac{P}{n_t}\mathbf{H}\mathbf{H}^{\dagger}\mathbf{Q}_{\mathbf{n}}^{-1}\right)^{\lambda}\right) ,
\end{aligned}
$$

and $i := \sqrt{-1}$.

- Single-User MIMO and the Deformed Laguerre Weight
- Without loss of generality

$$\mathrm{Q_n} = \mathrm{I}_{n_r}.$$

Due to the normalization of the trace of Q_n, the transmit power P also represents the SNR.

- Moment generating function,

$$M(\lambda) = E_{\mathrm{H}}\left[\det\left(\mathrm{I}_{n_r} + \frac{1}{t}\mathrm{H}\mathrm{H}^{\dagger}\right)^{\lambda}\right], \quad t := \frac{n_t}{P}.$$

Let

$$m := \max\{n_r, n_t\}, \quad n := \min\{n_r, n_t\}, \quad \alpha := m - n$$

and define

$$\mathrm{W} := \begin{cases} \mathrm{H}\mathrm{H}^{\dagger}, & n_r < n_t \\ \mathrm{H}^{\dagger}\mathrm{H}, & n_r \geq n_t \end{cases}.$$

• W is a complex Wishart random matrix with **positive eigenvalues** denoted by $\{x_i\}_{i=1}^{n}$ with j.p.d.f.

$$p(x_1, x_2, \dots, x_n) \quad \prod_{i=1}^{n} w_{\mathrm{Lag}}(x_i) \prod_{1 \leq j < k \leq n} (x_j - x_k)^2,$$

where $w_{\mathrm{Lag}}(x) = x^{\alpha}\mathrm{e}^{-x}$ is the classical Laguerre weight.

• MGF

$$M(\lambda) = E\left[\det\left(\mathrm{I}_n + \frac{1}{t}\mathrm{W}\right)^{\lambda}\right] = E\left[\prod_{k=1}^{n}\left(1 + \frac{x_k}{t}\right)^{\lambda}\right]$$

$$= \frac{\int_{\mathbb{R}_+^n} \prod_{i<j}(x_i - x_j)^2 \prod_{k=1}^{n}\left(1 + \frac{x_k}{t}\right)^{\lambda} w_{\mathrm{Lag}}(x_k)dx_k}{\int_{\mathbb{R}_+^n} \prod_{i<j}(x_i - x_j)^2 \prod_{k=1}^{n} w_{\mathrm{Lag}}(x_k)dx_k}.$$

- Andreief-Heine identity:

$$D_n[w] = \det(\mu_{i+j})_{i,j=0}^{n-1}$$

$$= \quad 1$$

- Facts about OP: Recurrence relations.

$$zP_n(z) = P_{n+1}(z) + \alpha_n P_n(z) + \beta_n P_{n-1}(z).$$

$$P_0(z) = 1, \ \beta_0 P_{-1}(z) = 0.$$

- $p_1(n)$ plays an important role in later developments.

Lemma 1 Suppose $v = -\log w$ has a derivative in some Lipshitz class with positive exponent. The lowering and raising operators satisfy the following:

$$\begin{aligned} P_n'(z) &= -B_n(z)P_n(z) + \beta_n A_n(z)P_{n-1}(z) \\ P_{n-1}'(z) &= [B_n(z) + v'(z)]P_{n-1}(z) - A_{n-1}(z)P_n(z), \end{aligned}$$

where

$$\begin{aligned} A_n(z) &:= \frac{1}{h_n}\int_a^b \frac{v'(z) - v'(y)}{z-y} P_n^2(y)w(y)dy \\ B_n(z) &:= \frac{1}{h_{n-1}}\int_a^b \frac{v'(z) - v'(y)}{z-y} P_n(y)P_{n-1}(y)w(y)dy. \end{aligned}$$

Lemma 2 The functions $A_n(z)$ and $B_n(z)$ satisfy the conditions:

$$B_{n+1}(z) + B_n(z) = (z-\alpha_n)A_n(z) - v'(z) \quad (S_1)$$

$$1 + (z-\alpha_n)[B_{n+1}(z) - B_n(z)] = \beta_{n+1}A_{n+1} - \beta_n A_{n-1}(z) \quad (S_2).$$

Lemma 3 The functions $A_n(z)$, $B_n(z)$, and the sum $\sum_{j=0}^{n-1} A_j(z)$, satisfy the conditions:

$$B_n^2(z) + \mathrm{v}(z) B_n(z) + \sum_{j=0}^{n-1} A_j(z) = \beta_n A_n(z) A_{n-1}(z). \quad (S_2)$$

- If $\mathrm{v}(z)$ rational then

$$\frac{\mathrm{v}(z) - \mathrm{v}(y)}{z - y} = \frac{\alpha}{zy}$$

- Derivation: A Quick Sketch

$$A_n(z) = \frac{1 - R_n(t)}{z} + \frac{R_n(t)}{z + t}$$

$$B_n(z) = -\frac{n + r_n(t)}{z} + \frac{r_n(t)}{z + t}$$

$$R_n(t) = \frac{\lambda}{h_n} \int_0^\infty \frac{[P_n(y)]^2}{y + t} w(y, t) dy$$

$$r_n(t) = \frac{\lambda}{h_{n-1}} \int_0^\infty \frac{P_n(y) P_{n-1}(y)}{y + t} w(y, t) dy.$$

- Sub. into compatibility conditions,

$$\alpha_n = 2n + 1 + \alpha + \lambda - tR_n \qquad (\alpha)$$

$$\beta_n = \frac{1}{1 - R_n} \left[(2n + \alpha + \lambda) r_n + \frac{r_n^2 - \lambda r_n}{R_n} + n(n + \alpha) \right] \qquad (\beta)$$

$$t \sum_{j=0}^{n-1} R_j = n(n + \alpha + \lambda) + \mathrm{p}_1(n, t) \qquad (\text{sum})$$

- Difference equations in n

$$r_{n+1} + r_n = \lambda - R_n(t + 2n + 1 + \alpha + \lambda - tR_n) \qquad (d_1)$$

$$r_n^2 - \lambda r_n = \beta_n R_n R_{n-1} \qquad (d_2)$$

with "initial conditions" $r_0(t) = 0$, $R_0(t) = \text{given}$.

- $t-$ or Toda evolution
- A pair of Riccati Equations

$$
\begin{aligned}
2r_n &= tR_n + \lambda - R_n(t + 2n + \alpha + \lambda - tR_n) \\
tr_n &= \frac{r_n^2 - \lambda r_n}{R_n} \\
&- \frac{R_n}{1 - R_n}\left[(2n + \alpha + \lambda)r_n + \frac{r_n^2 - \lambda r_n}{R_n} + n(n + \alpha)\right].
\end{aligned}
$$

- $R_n = y/(y - 1)$.

$$
\begin{aligned}
y &= \frac{3y - 1}{2y(y - 1)}(y)^2 - \frac{y}{t} + \frac{(y - 1)^2}{t^2}\left[\frac{\alpha^2}{2}y - \frac{\lambda^2}{2y}\right] \\
&+ \frac{(2n + 1 + \alpha + \lambda)y}{t} - \frac{y(y + 1)}{2(y - 1)},
\end{aligned}
$$

- Using the sum,

$$
H_n := t\frac{d}{dt}\log D_n = t\sum_{j=0}^{n-1} R_j = n(n + \alpha + \lambda) + \mathrm{p}_1(n, t).
$$

- Representation of R_n –

- Large n Coulomb Fluid Method

- large n, the ratio approximated by

  $D_n(\lambda)$

- Gaussian distribution with mean and variance given by

$$\mu_{\text{Coul.}} = -S_2(T) - n \log T$$
$$\sigma^2_{\text{Coul.}} = -S_1(T)$$

- Outage probability

$$P_{\text{out}}(C_{\text{out}}) \approx \frac{1}{2}\left(1 + \text{erf}\left(\frac{C_{\text{out}} - \mu}{\sqrt{2\sigma^2}}\right)\right).$$

$$a = 2 + \beta - 2\sqrt{1+\beta}, \quad b = 2 + \beta + 2\sqrt{1+\beta}$$

- In general

$$\log M(\lambda) = \sum_{l=1}^{\infty} \kappa_l \frac{\lambda^l}{l!}.$$

- Beyond the Coulomb Fluid Approximation

H_n

• Summation of the series in P. ($t = n/P =: nT$)

• Look at g_1 in detail:

$$(g_1)^2 - 4n^2 g_1 - 2(t + 2n)g_1 g_1 + (t^2 + 4nt)(g_1)^2 - t^2(g_1)^2 = 0$$

$$(g_1)^2 - 4ng_1 - 2(T + 2)g_1 g_1 + (T^2 + 4T)(g_1)^2 - \frac{T^2}{n^2}(g_1)^2 = 0.$$

$$g_1 = nY_0(T) + \frac{Y_1(T)}{n} + \ldots$$

$$Y_0^2 - 4Y_0 - 2(T + 2)Y_0 Y_0 + (T^2 + 4T)(Y_0)^2 = 0, \quad (\ )$$

$$\begin{aligned} & 2Y_0 Y_1 - 2(2 + T)Y_1 Y_0 - 4Y_1 - 2(2 + T)Y_0 Y_1 \\ + \quad & 2T(4 + T)Y_0 Y_1 - T^2 Y_0 = 0, \quad (\quad) \end{aligned}$$

• Solutions,

$$Y_0(T) = -\frac{4 + T - \sqrt{T(4 + T)}}{4 + T + \sqrt{T(4 + T)}} - 1$$

$$Y_1(T) = -\frac{1}{\sqrt{T}(4 + T)^{5/2}} - \frac{1}{32\sqrt{T}}.$$

• Note from Coulomb Fluid

$$Y_0(T) = -T\frac{d}{dT}(S_2/n - \ln T),$$

satisfies (*) and when sub. into (**) it becomes a *linear* equation in Y_1.

Similarly

$$g_2 = Z_0(T) + \frac{Z_1(T)}{n^2} + \dots.$$

where

$$Z_0(T) = -T \frac{d}{\quad}$$

- Summary
- Cumulants

Note μ_{Coulomb} is LINEAR in n

$$\text{Cumulants} \quad \kappa_l = -l! \int_0^P g_l(n/y)\frac{dy}{y}$$

$$\kappa_1 = \mu_{\text{Coulomb}} + \frac{1}{n}\mu_{\text{Corr.}} + O\left(\frac{1}{n^3}\right)$$

$$\kappa_2 = \sigma^2_{\text{Coulomb}} + \frac{1}{n^2}\sigma^2_{\text{Corr.}} + O\left(\frac{1}{n^4}\right)$$

$$\kappa_3 = \frac{1}{n}\kappa_{3,A} + \frac{1}{n^3}\kappa_{3,B} + O\left(\frac{1}{n^5}\right).$$

- Analysis at large P. Large deviation(?)

$$\mu \qquad a\,n\log P + b\,\frac{\sqrt{P}}{n}, \quad P = O(n^4)$$

$$\sigma^2 \qquad a\,\log P + b\,\frac{P}{n^2}, \quad P = O(n^2)$$

$$\kappa_3 \qquad \frac{a}{n} + b\,\frac{P^2}{n^3}, \quad P = O(n)$$

See slides

End

- A selection of Integrals: Schwinger (1918–1994)

$$\log(a+b) = \log a + \int_0^1 \frac{b\,dx}{a+bx}$$

$$\int_a^b \frac{\log(x+t)}{\sqrt{(b-x)(x-a)}}\,dx = 2\pi \log\left(\frac{\sqrt{t+a}+\sqrt{t+b}}{2}\right)$$

$$\int_a^b \frac{\log(x+t)}{\sqrt{(b-x)(x-a)}\,(x+t)}\,dx$$

$$= -\frac{2\pi}{\sqrt{(t+a)(t+b)}}\log\left(\frac{1}{2\sqrt{t+a}}+\frac{1}{2\sqrt{t+b}}\right)$$

$$\int_a^b \frac{\log(x+t)}{\sqrt{(b-x)(x-a)}\;x}\,dx = \frac{\pi}{\sqrt{ab}}\log\left(\frac{\left(\sqrt{ab}+\sqrt{(t+a)(t+b)}\right)^2-t^2}{\left(\sqrt{a}+\sqrt{b}\right)^2}\right)$$

$$\int_a^b \frac{\log(x+t)}{\sqrt{(b-x)(x-a)}\,(x-1)}\,dx$$