



CENTRO DE INVESTIGACIÓN EN MATEMÁTICAS A.C

FUNCIONALES DEL MOVIMIENTO BROWNIANO Y EL ESPECTRO
DE MATRICES GAUSSIANAS UNITARIAS

T E S I S

QUE PARA OBTENER EL GRADO DE:
MAESTRO EN CIENCIAS CON ESPECIALIDAD EN PROBABILIDAD Y
ESTADÍSTICA

PRESENTA:
MIGUEL ANGEL PLUMA RODRIGUEZ

CO-DIRECTORES DE TESIS:

DR. CHRISTIAN HOUDRÉ
INSTITUTO TECNOLÓGICO DE GEORGIA

DR. VÍCTOR MANUEL PÉREZ ABREU C.
CENTRO DE INVESTIGACIÓN EN MATEMÁTICAS A. C.

2015

Sinodales

Dr. Octavio Arizmendi Echegaray
Institución: CIMAT
Departamento: Probabilidad y Estadística

Dr. Jesús Armando Domínguez Molina
Institución: Universidad Autónoma de Sinaloa
Facultad de Ciencias Físico-Matemáticas

*A mis padres quienes siempre me han apoyado,
especialmente cuando comencé este viaje por las matemáticas.*

Agradecimientos

Agradezco a mi mis padres por el apoyo incondicional que me han brindado a lo largo de mi vida, sin su apoyo nada de esto hubiera sido posible.

También quiero agradecer a mis co-asesores Dr. Christian Houdré y Dr. Víctor Manuel Pérez-Abreu Carrión, por el apoyo que me brindaron en la redacción de la tesis, así como su apoyo para realizar una visita académica al Instituto Tecnológico de Georgia en Estados Unidos. De manera particular quiero mostrar mi agradecimiento al Dr. Víctor Manuel Pérez Abreu Carrión por el tiempo y la confianza que me ha otorgado no sólo en la elaboración de esta tesis, si no en toda la maestría.

Agradezco también a mis sinodales Dr. Jesús Armando Domínguez Molina y Dr. Octavio Arizmeni Echeagaray, por sus valiosos comentarios y el tiempo que le dedicaron a la lectura de esta tesis.

También quiero agradecer al CONACYT por el apoyo económico brindado durante la maestría, al CIMAT por todo lo que me dio tanto económicamente como académicamente y al proyecto del CONACYT con número 222668, por el apoyo brindado.

Índice general

Introducción	6
1. Preliminares	9
1.1. Combinatoria	9
1.1.1. La correspondencia RSK	9
1.1.2. La subsucesión no decreciente más grande	16
1.1.3. LI_n y la correspondencia RSK	17
1.2. Matrices aleatorias y su espectro	21
1.2.1. El ensamble gaussiano unitario (GUE)	22
1.2.2. GUE sin traza y GUE por bloques sin traza	22
1.3. El movimiento browniano y el funcional de Glynn y Whitt	23
1.3.1. El movimiento browniano	23
1.3.2. El funcional de Glynn y Whitt	24
1.4. GUE como límite de diagramas de Young	26
2. Matrices aleatorias y el movimiento browniano	30
2.1. Funcionales del movimiento browniano y el espectro de los menores principales de un GUE	30
2.2. Palabras aleatorias	35
2.2.1. Generalidades	35
2.2.2. El problema de Ulam para palabras aleatorias	38
2.2.3. Forma límite del diagrama de Young de una palabra aleatoria	45
2.3. Comentarios finales	51
Bibliografía	55

Introduccion

En esta tesis se exponen algunos aspectos de las relaciones entre funcionales del movimiento browniano, el espectro de matrices del Ensamble Gaussiano Unitario (GUE) y la distribución límite de la subsucesión no decreciente más grande de palabras aleatorias.

Una primera relación, entre el valor propio más grande D_k de una matriz GUE de dimensión k y el movimiento browniano, establece que

$$D_k \stackrel{d}{=} \rho_k(B) := \sup_{0=t_0 \leq \dots \leq t_k=1} \sum_{i=1}^k (B_i(t_i) - B_i(t_{i-1})), \quad (1)$$

donde $B = (B_1, \dots, B_k)$ es un movimiento browniano estándar de dimensión k , es decir B tiene matriz de covarianza $\Sigma_0 = I_k$, la matriz identidad de dimensión k . Esta relación fue demostrada en 2001 por Baryshnikov [3] y de manera independiente por Gravner, Tracy y Widom [7]. Sin embargo $\rho_k(B)$ fue estudiado por primera vez en 1975 por Pitman [21] en el contexto de procesos de Bessel de dimensión $k = 2$.

Una segunda relación, estudiada recientemente, es la descripción de $\rho_k(B)$ como el límite de la longitud de la subsucesión no decreciente más grande de una palabra aleatoria. En específico, sea $\omega_n = x_1 \dots x_n$ una palabra de longitud n en un alfabeto con k letras, definimos a LI_n la longitud de la subsucesión no decreciente más grande como

$$LI_n = \max\{s : \exists 1 \leq i_1 < i_2 < \dots < i_s \leq n, \text{ con } x_{i_1} \leq x_{i_2} \leq \dots \leq x_{i_s}\}.$$

En 2013 Houdré y Xu [11] demostraron que para el caso de una palabra aleatoria con distribución uniforme en un alfabeto con k letras (caso homogéneo), se cumple que

$$\frac{LI_n - \frac{n}{k}}{\sqrt{\frac{n}{k}}} \Rightarrow \rho_k(B),$$

donde B tiene matriz de covarianza

$$\Sigma_1 = \begin{pmatrix} \frac{k-1}{k} & -\frac{1}{k} & \dots & -\frac{1}{k} \\ -\frac{1}{k} & \frac{k-1}{k} & \dots & -\frac{1}{k} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{k} & -\frac{1}{k} & \dots & \frac{k-1}{k} \end{pmatrix}.$$

Mas aún, Benaych-Georges y Houdré [4] demostraron en 2015 que cuando B tiene matriz de covarianza Σ_1 , el funcional $\rho_k(B)$ es igual en distribución al valor propio

más grande de un GUE sin traza. De hecho, lo anterior es un caso particular de un resultado más general en [4], en donde se toman palabras con letras independientes y tales que no todas tienen la misma probabilidad de aparecer (caso no homogéneo). Específicamente, cuando hay k letras con probabilidad $p \in (0, 1)$ y las otras $n - k$ letras tienen probabilidad de aparecer estrictamente menor que p , se tiene que

$$\frac{LI_n - np}{\sqrt{np}} \Rightarrow \rho_k(B),$$

donde ahora B tiene matriz de covarianza

$$\Sigma_2 = \begin{pmatrix} 1-p & -p & \cdots & -p \\ -p & 1-p & \cdots & -p \\ \vdots & \vdots & \ddots & \vdots \\ -p & -p & \cdots & 1-p \end{pmatrix}.$$

En este caso el funcional $\rho_k(B)$ también es igual en distribución al valor propio más grande de una matriz que se construye a partir de varias matrices GUE sin traza, cuyas dimensiones dependen de las multiplicidades de las probabilidades de las letras del alfabeto.

Una tercera relación, entre la distribución límite de LI_n y su identificación como la distribución del valor propio más grande D_k de matrices aleatorias, fue dada en 1994 por Kerov [18] y redescubierta por 2001 por Tracy y Widom [27], y Johansson [14] para el caso homogéneo, y meses después por Its, Tracy y Widom [12] para el caso no homogéneo. Los métodos utilizados en [12] y [27] incluyen ecuaciones diferenciales parciales (ecuación de Painlevé), determinantes de Fredholm y Toeplitz; en general las matemáticas que están alrededor del problema de Riemann-Hilbert.

Por otra parte, lo novedoso de los trabajos de Houdré y coautores en [4] y [11] es describir la distribución límite de la normalización de LI_n como funcionales del movimiento browniano, además del uso de métodos puramente probabilistas tales como el teorema de Donsker y otros resultados clásicos de convergencia débil de procesos estocásticos.

Una herramienta común en los trabajos [4], [11], [12], [14] y [27], son los diagramas de Young, los cuales tienen asociado un vector cuya primera entrada es la variable LI_n . En general las entradas de este vector son las longitudes de las subsucesiones crecientes ajenas cuya unión es toda la palabra. En este sentido la forma del diagrama de Young de una palabra (o permutación) nos proporciona más información que la variable LI_n .

El objetivo de esta tesis es presentar de manera detallada estas relaciones entre los funcionales del movimiento browniano, el espectro de matrices aleatorias y la distribución límite de LI_n , siguiendo el enfoque probabilístico. Con el fin de hacer autocontenido este trabajo, presentamos el material necesario para estudiar estos resultados, entre los que se encuentran elementos de teoría de representaciones de grupos como los diagramas de Young y la correspondencia RSK.

Los problemas descritos anteriormente sobre la distribución límite de LI_n en el caso de palabras aleatorias son una variante al problema de Ulam, quien propuso en 1961 el estudio de la distribución límite de LI_n en el caso de permutaciones aleatorias, con el fin de ilustrar la utilidad del método Monte Carlo.

En 1999 Baik, Deift y Johansson [2] dieron solución al problema de Ulam [28], dando la relación pionera entre la distribución asintótica de LI_n y la distribución límite del valor propio más grande de una matriz GUE. Esta última distribución límite surgió por primera vez en los trabajos pioneros de Tracy y Widom [25, 26] sobre el límite del valor propio más grande de una matriz aleatoria. Actualmente es una distribución universal en el sentido de que aparece en distintos contextos en estadística, modelos de crecimiento aleatorio, física, comunicación inalámbrica y finanzas [30].

El contenido de la tesis está organizado de la siguiente manera. En el capítulo 1 se presentan diversos preliminares para leer esta tesis. En la sección 1.1 presentamos elementos de combinatoria como permutaciones generalizadas, tabloides y la correspondencia RSK, la cual nos da una relación uno a uno entre permutaciones generalizadas, matrices con entradas enteras y pares de tabloides con la misma forma. También se presenta la forma en que la correspondencia RSK se relaciona con la subsucesión no decreciente más grande y se hace una breve reseña histórica del problema de Ulam. En la sección 1.2 se introducen definiciones y la notación de matrices aleatorias GUE con y sin traza, así como generalizaciones que se usan en esta tesis. En la sección 1.3 se incluye material sobre funcionales del movimiento browniano, incluido el teorema de Donsker, el funcional de Glynn y Whitt y su contexto inicial en teoría de colas. Por último en la sección 1.4 se presenta a detalle un resultado de Johansson [14], conjeturado por Tracy y Widom [27], sobre la convergencia en distribución de las medida inducida por la forma del diagrama de Young asociado a palabras aleatorias homogéneas con distribución uniforme, a la distribución conjunta del espectro ordenado del GUE, cuando la longitud de la palabra tiende a infinito. Este resultado será clave para relacionar al espectro del GUE con el movimiento browniano.

En el capítulo 2 se presentan las relaciones entre: funcionales del movimiento browniano, el espectro de matrices Gaussianas Unitarias y el problema de Ulam para palabras aleatorias. En la sección 2.1 se expone de manera pormenorizada los trabajos de Benaych-Georges y Houdré [4], sobre la relación entre funcionales del movimiento browniano y el espectro de los menores de un GUE (Teorema 2.1.1). En la sección 2.2, se expone la solución al problema de Ulam que fue propuesta por Houdré y Xu [11], en el caso de palabras aleatorias. Ellos, a diferencia de Tracy y Widom [25, 27] utilizan métodos completamente probabilistas para identificar a la forma límite del diagrama de Young asociado a una palabra aleatoria en términos de funcionales del movimiento browniano. Adicionalmente, usando los resultados de la sección 2.1 se muestra cómo se pueden obtener los resultados de Tracy y Widom a partir de los resultados de Houdré y Xu. Como una consecuencia de estos dos enfoques se obtiene una conjetura sobre una igualdad en distribución como la que se muestra en (1), donde ahora D_k es el valor propio más grande de una matriz aleatoria que es una generalización de las matrices GUE y, B es un movimiento browniano con matriz de covarianza distinta a la identidad.

Capítulo 1

Preliminares

1.1. Combinatoria

La herramienta base para relacionar el espectro de una matriz aleatoria con funcionales del movimiento browniano es la correspondencia RSK¹. Este resultado nos da una correspondencia biunívoca entre las permutaciones generalizadas, las matrices con entradas enteras no negativas y el conjunto de pares de tabloides. La forma de construir esta correspondencia es dando explícitamente el algoritmo para pasar de una permutación generalizada a una matriz con entradas enteras no negativas y a un par de tabloides. Por esta razón a veces a la correspondencia RSK se le llama algoritmo RSK.

En esta sección daremos las definiciones necesarias para enunciar la correspondencia RSK, mencionaremos los casos particulares que nos serán de interés y mencionaremos la relación que hay con la subsucesión no decreciente más grande. Para los detalles técnicos referimos al lector a los libros [5] y a [24].

1.1.1. La correspondencia RSK

A continuación definimos las cadenas de caracteres que nos serán de interés: Sean n, p y q enteros positivos, denotaremos $[n] = \{1, 2, \dots, n\}$.

i) Permutaciones

El conjunto de permutaciones de n enteros se define como

$$S_n := \{\sigma : [n] \rightarrow [n] \mid \sigma \text{ es biyectiva}\}.$$

También usaremos la notación

$$S_n \ni \sigma_n = \begin{pmatrix} 1 & \cdots & n \\ j_1 & \cdots & j_n \end{pmatrix}, \quad 1 \leq j_1, \dots, j_n \leq n,$$

con $\sigma_n(k) = j_k$.

¹RSK es el acrónimo de Robinson-Schensted-Knuth, quienes desarrollaron este algoritmo.

ii) **Palabras**

El conjunto S_n^q de palabras de tamaño n en un alfabeto con q letras se define por

$$W_n^q := \{\omega : [n] \rightarrow [q]\} = [q]^n.$$

También usaremos la notación

$$W_n^q \ni \omega_n = \begin{pmatrix} 1 & \cdots & n \\ j_1 & \cdots & j_n \end{pmatrix}, \quad 1 \leq j_1, \dots, j_n \leq q, \quad (1.1)$$

con $\omega_n(k) = j_k$.

iii) **Permutaciones Generalizadas**

El conjunto $GP_n^{p,q}$ de permutaciones generalizadas de p en q elementos y de tamaño n , se define por

$$GP_n^{p,q} := \left\{ \pi : [n] \rightarrow [p] \times [q] \mid \begin{array}{l} \text{Si } \pi_n(k) = (r, s_1) \text{ y } \pi_n(k+1) = (r, s_2) \\ \text{entonces } s_1 \leq s_2. \end{array} \right\} \quad (1.2)$$

También usaremos la notación

$$GP_n^{p,q} \ni \pi_n = \begin{pmatrix} i_1 & \cdots & i_n \\ j_1 & \cdots & j_n \end{pmatrix}.$$

sujeito a $1 \leq i_1 \leq \cdots \leq i_n$ y $1 \leq j_1, \dots, j_n \leq q$, con $i_k = i_{k+1}$ implica $j_k \leq j_{k+1}$.

Observemos que las permutaciones son un caso particular de las palabras, y estas a su vez son un caso particular de las permutaciones generalizadas.

Escribimos

$$M_{p,q}(n) := \{M \in \mathbb{M}_{p \times q}[\mathbb{N}] \mid \sum_{i=1}^p \sum_{j=1}^q M_{ij} = n\} \quad (1.3)$$

donde $\mathbb{M}_{p \times q}[\mathbb{N}]$ es el conjunto de matrices de $p \times q$ con entradas en \mathbb{N} , y

$$M_{p,q} = \bigcup_{n=1}^{\infty} M_{p,q}(n). \quad (1.4)$$

El último ingrediente necesario para enunciar la correspondencia RSK son las tablas y los tabloides, para lo cual necesitamos algunas definiciones preliminares.

Una **partición** λ de n (que denotaremos por $\lambda \vdash n$) es una sucesión $(\lambda_k)_{k \geq 1}$ de enteros no negativos tales que $\lambda_1 \geq \lambda_2 \geq \cdots$. Denotaremos por $\mathcal{P}^{(n)}$ al conjunto de particiones de n , es decir

$$\mathcal{P}^{(n)} = \{\lambda \in \mathbb{N}^{\mathbb{Z}^+} : \lambda_1 \geq \lambda_2 \geq \cdots \text{ y } \sum_{i=1}^{\infty} \lambda_i = n\}, \quad (1.5)$$

y

$$\mathcal{P} = \bigcup_{n=1}^{\infty} \mathcal{P}^{(n)}. \quad (1.6)$$

Cada elemento $\lambda \in \mathcal{P}$ tiene asociadas dos características; su longitud y su peso, las cuales se definen de la siguiente manera.

a) **Longitud**

$$\ell(\lambda) := \text{máx}\{k : \lambda_k > 0\}.$$

b) **Peso**

$$|\lambda| := \sum_{k=1}^{\infty} \lambda_k.$$

Es decir $\lambda \vdash |\lambda|$ y $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_{\ell(\lambda)}, 0, 0, \dots)$, por esta razón usaremos también la notación $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_{\ell(\lambda)})$.

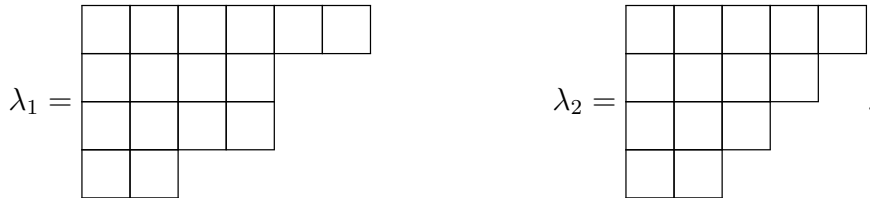
En la sección 1.1.3, veremos que a cada palabra en un alfabeto con m letras le podemos asociar una partición con a lo más m entradas positivas. Por esta razón será conveniente introducir la siguiente notación:

$$\mathcal{P}_m^{(n)} = \{\lambda \in \mathbb{N}^m : \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0 \geq \dots \text{ y } \sum_{i=1}^m \lambda_i = n\}, \quad (1.7)$$

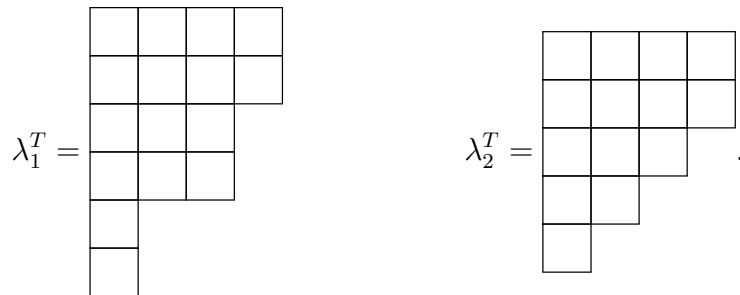
y

$$\mathcal{P}_m = \bigcup_{n=1}^{\infty} \mathcal{P}_m^{(n)}. \quad (1.8)$$

Un diagrama de Young es una representación gráfica de una partición. Esta representación se construye de la siguiente manera: dada $\lambda \in \mathcal{P}$ tomamos $|\lambda|$ bloques- \square y coloquemos λ_1 bloques- \square en un renglón, λ_2 bloques- \square en el segundo renglón justo debajo de los primeros λ_1 bloques- \square , y así sucesivamente hasta colocar $\lambda_{\ell(\lambda)}$ bloques- \square en el último renglón. Por ejemplo los diagramas de Young asociados a $\lambda_1 = (6, 4, 4, 2)$ y $\lambda_2 = (5, 4, 3, 2)$ son



Esta notación nos permite definir la transpuesta λ^T de una partición λ , mediante la transposición del diagrama de Young pensándolo como una matriz. Por ejemplo para los casos anteriores



También usaremos la notación

$$\lambda^T = (\lambda_1^T, \lambda_2^T, \dots).$$

Con esto podemos dar la definición de tablas y tabloides.

Un **tabloide** con forma $\lambda \in \mathcal{P}$ es un arreglo de enteros (a_{ij}) colocados en los bloques- \square del diagrama de Young de λ , y que son no decrecientes de izquierda a derecha en cada renglón y crecientes de arriba a abajo en cada columna.

Una **tabla** de forma $\lambda \in \mathcal{P}$ es un tabloide de la misma forma en el que los enteros sobre el diagrama de Young de λ son estrictamente crecientes de izquierda a derecha en cada renglón.

1	2	2	3	3	5
2	3	5	5		
4	4	6	6		
5	6				

Tabloide

1	3	7	12	13	15
2	5	10	14		
4	8	11	16		
6	9				

Tabla

Para cada $\lambda \in \mathcal{P}$ denotaremos por

$$T_q(\lambda) = \text{“Tabloides con forma } \lambda \text{ y lleno con los números del 1 al } q\text{”}, \quad (1.9)$$

Como los números en los renglones de un tabloide son estrictamente crecientes, se tiene que q tiene que ser mayor o igual al número de entradas positivas en λ . Para cada terna de enteros positivos p, q y n , denotaremos por

$$T_n^{p,q} := \bigcup_{\lambda \in \mathcal{P}(n)} T_p(\lambda) \times T_q(\lambda) = \left\{ \begin{array}{l} \text{Pares de tabloides } (P, Q) \text{ con} \\ \text{forma } \lambda \vdash n, \text{ llenados con los} \\ \text{enteros } \{1, \dots, p\} \text{ y } \{1, \dots, q\} \\ \text{respectivamente} \end{array} \right\} \quad (1.10)$$

Si T es una tabla o tabloide diferente del vacío denotaremos por $sh(T)$ a la forma de la tabla o el tabloide, es decir para toda tabla o tabloide T diferente de vacío se tiene que $sh(T) \in \mathcal{P}$.

El siguiente teorema será fundamental a lo largo de toda la tesis, la demostración completa de este resultado se puede ver en [5]. Aquí se presenta un bosquejo de la misma.

Teorema 1.1.1 (Correspondencia de Robinson-Schensted-Knuth) Sean n, p y q enteros positivos. Existe una relación biunívoca entre los siguientes conjuntos

$$M_{p,q}(n) \Leftrightarrow GP_n^{p,q} \Leftrightarrow T_n^{p,q}.$$

Demostración: (Bosquejo) $GP_n^{p,q} \Rightarrow M_{p,q}(n)$

Sea

$$\pi \in \left(\begin{array}{ccc} i_1 & \cdots & i_n \\ j_1 & \cdots & j_n \end{array} \right) \in GP_n^{p,q}.$$

Construyamos una matriz $M = (m_{ij})$ de la siguiente manera

$$m_{ij} = \text{“Número de veces que aparece } \binom{i}{j} \text{ en } \pi\text{”}.$$

Es claro que $M \in M_{p,q}(n)$.

$\mathbf{M}_{p,q}(\mathbf{n}) \Rightarrow \mathbf{GP}_n^{p,q}$

Dada $M \in M_{p,q}(n)$ construyamos

$$\pi \in \begin{pmatrix} i_1 & \cdots & i_n \\ j_1 & \cdots & j_n \end{pmatrix} \in GP_n^{p,q},$$

revirtiendo el proceso anterior. Recordemos que para π una permutación generalizada, si $i_k = i_{k+1}$ entonces $j_k \leq j_{k+1}$. Esto nos permite asegurar que con este método obtenemos una única permutación generalizada.

$\mathbf{GP}_n^{p,q} \Rightarrow \mathbf{T}_n^{p,q}$

Denotemos por \emptyset al tabloide vacío. Dado un tabloide T construiremos un nuevo tabloide T' mediante la inserción de un tabloide de la forma \boxed{r} o \emptyset . Supongamos que T tiene s -renglones y denotemos a su i -ésimo renglón de la siguiente manera

$$\boxed{R_{i1}} \boxed{R_{i2}} \cdots \boxed{R_{iw_i}}.$$

Recordemos que como T es un tabloide se cumple que

$$R_{i1} \leq R_{i2} \leq \cdots \leq R_{iw_i}, \quad 1 \leq i \leq s,$$

$$w_1 \geq w_2 \geq \cdots \geq w_s.$$

A continuación daremos la descripción del algoritmo inserción por renglón. Con este algoritmo se inserta un bloque- \square a un tabloide con forma de renglón, y después se se generalizará a un tabloide de forma arbitraria. A este algoritmo lo denotaremos de la siguiente forma

$$\boxed{R_{i1}} \boxed{R_{i2}} \cdots \boxed{R_{iw_i}} \leftarrow \boxed{r}$$

El algoritmo genera dos cosas de interés, un nuevo renglón

$$\boxed{R'_{i1}} \boxed{R'_{i2}} \cdots \boxed{R'_{iw_i}}$$

y un tabloide de la forma $\boxed{r'}$ o el tabloide vacío \emptyset , el cual puede ser insertado en otro renglón.

Caso 1: Se inserta el tabloide vacío \emptyset .

$$\boxed{R_{i1}} \boxed{R_{i2}} \cdots \boxed{R_{iw_i}} \leftarrow \emptyset .$$

En es este caso obtenemos el mismo renglón $\boxed{R_{i1}} \boxed{R_{i2}} \cdots \boxed{R_{iw_i}}$ y el tabloide vacío \emptyset .

Caso 2: Si $R_{iw_i} \leq r$ obtenemos el renglón $\boxed{R_{i1}} \boxed{R_{i2}} \cdots \boxed{R_{iw_i}} \boxed{r}$ y el tabloide

vacío \emptyset .

Caso 3: Si existe $1 \leq k \leq w_i$ tal que $r \leq R_{ij}$ para toda $j \geq k$ y $r > R_{ij}$ para toda $j < k$, entonces el renglón se descompone de la siguiente manera

$$\boxed{R_{i1} \quad R_{i2} \quad \cdots \quad R_{iw_i}} = \boxed{S_{i1} \quad \cdots \quad S_{i(k-1)} \quad T_{ik} \quad \cdots \quad T_{iw_i}},$$

Donde r es estrictamente mayor que $S_{i1}, \dots, S_{i(k-1)}$ y menor o igual que T_{ik}, \dots, T_{iw_i} . En este caso el algoritmo genera el siguiente renglón

$$\boxed{S_{i1} \quad \cdots \quad S_{i(k-1)} \quad r \quad \cdots \quad T_{iw_i}},$$

y el tabloide $\boxed{T_{ik}}$.

De modo que para definir $T' = T \leftarrow \boxed{r}$ lo hacemos de manera recursiva. Sean R_1, R_2, \dots, R_s los renglones de T , el primer renglón R'_1 de T' es el que se obtiene en $R_1 \leftarrow \boxed{r}$ y sea $\boxed{r_1}$ el tabloide obtenido. El segundo renglón R'_2 de T' se obtiene en $R_1 \leftarrow \boxed{r_1}$ y sea $\boxed{r_2}$ el tabloide obtenido. Los demás renglones se obtienen de la misma manera.

Si en el último renglón, al hacer $R_s \leftarrow \boxed{r_{s-1}}$ obtenemos el tabloide vacío, terminamos la construcción de T' . Si el tabloide obtenido no es el vacío, agregamos un nuevo renglón a T' con un sólo bloque $\boxed{r_s}$.

Tomemos ahora

$$\pi_n \in \begin{pmatrix} i_1 & \cdots & i_n \\ j_1 & \cdots & j_n \end{pmatrix} \in GP_n^{p,q},$$

usando el algoritmo inserción por renglón vamos a construir, a partir de π_n a un par de tabloides (P, Q) los cuales llenaremos con los enteros $\{1, \dots, p\}$ y $\{1, \dots, q\}$. Construyamos primero a Q insertando bloques- \square con los enteros j_1, \dots, j_n al tabloide vacío. Es decir, en la sucesión de tabloides

$$\begin{aligned} Q_0 &= \emptyset, \\ Q_1 &= Q_0 \leftarrow \boxed{j_1}, \\ Q_2 &= Q_1 \leftarrow \boxed{j_2}, \\ &\vdots \\ Q_n &= Q_{n-1} \leftarrow \boxed{j_n}, \end{aligned}$$

tomaremos a $Q = Q_n$. Observemos que para pasar de Q_{k-1} a Q_k sólo agregamos un tabloide con un sólo bloque, entonces podemos construir otra sucesión de tabloides $P_0 = \emptyset, P_1, \dots, P_n$ de tal forma que los tabloides (P_i, Q_i) tienen la misma forma, pero los bloques de cada P_k están llenos con los números i_1, \dots, i_n y son colocados en ese orden en la casilla que se va agregando. ■

Ejemplos:

a) Sea

$$\pi_6 \in \begin{pmatrix} i_1 & i_2 & i_3 & i_4 & i_5 & i_6 \\ 1 & 2 & 1 & 3 & 2 & 1 \end{pmatrix}$$

Entonces

$$\boxed{1} \rightarrow \boxed{1 \ 2} \rightarrow \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 1 & \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 1 & & \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 1 & 2 & \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 1 & 2 & \\ \hline 1 & & \\ \hline \end{array} = Q.$$

$$\boxed{i_1} \rightarrow \boxed{i_1 \ i_2} \rightarrow \begin{array}{|c|c|} \hline i_1 & i_2 \\ \hline i_3 & \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|} \hline i_1 & i_2 & i_4 \\ \hline i_3 & & \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|} \hline i_1 & i_2 & i_4 \\ \hline i_3 & i_5 & \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|} \hline i_1 & i_2 & i_4 \\ \hline i_3 & i_5 & \\ \hline i_6 & & \\ \hline \end{array} = P.$$

b) Sea

$$\pi_6 \in \begin{pmatrix} 1 & 2 & 2 & 4 & 4 & 5 \\ 3 & 2 & 2 & 1 & 4 & 3 \end{pmatrix}$$

Entonces

$$\boxed{3} \rightarrow \begin{array}{|c|} \hline 2 \\ \hline 3 \\ \hline \end{array} \rightarrow \begin{array}{|c|c|} \hline 2 & 2 \\ \hline 3 & \\ \hline \end{array} \rightarrow \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 2 & \\ \hline 3 & \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|} \hline 1 & 2 & 4 \\ \hline 2 & & \\ \hline 3 & & \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 2 & 4 & \\ \hline 3 & & \\ \hline \end{array} = Q.$$

$$\boxed{1} \rightarrow \begin{array}{|c|} \hline 1 \\ \hline 2 \\ \hline \end{array} \rightarrow \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 2 & \\ \hline \end{array} \rightarrow \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 2 & \\ \hline 4 & \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|} \hline 1 & 2 & 4 \\ \hline 2 & & \\ \hline 4 & & \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|} \hline 1 & 2 & 4 \\ \hline 2 & 5 & \\ \hline 4 & & \\ \hline \end{array} = P.$$

En nuestro bosquejo de la demostración del teorema anterior, nos hemos limitado a dar únicamente los algoritmos que permiten pasar de una permutación generalizada a una matriz, y a un par de tabloides. Para los detalles técnicos relacionados con este teorema remitimos al lector a [5].

En esta sección hemos presentado a la correspondencia RSK en su versión más general. Sin embargo, nosotros utilizaremos un caso particular que es cuando las permutaciones generalizadas son palabras de tamaño n en un alfabeto con k letras. Observemos que cuando una permutación aleatoria es una palabra, su correspondiente matriz es tal que en cada renglón aparece sólo un 1 y sus demás entradas son ceros. Denotemos por $M_q(n)$ al conjunto de todas estas matrices, es decir

$$M_q(n) = \left\{ M \in M_{n,q}(n) : \sum_{i=1}^q M_{ki} = 1, \text{ para toda } 1 \leq k \leq n \right\}.$$

Por otra parte, la correspondencia RSK asigna a cada palabra ω_n de tamaño n un par de tabloides (P, Q) de la misma forma tales que $sh(P) \vdash n$. Pero como P está lleno con los enteros $1, \dots, n$ entonces para el caso de palabras P es una tabla. Con estas observaciones se tiene el caso particular de la correspondencia RSK que nos será de interés.

Teorema 1.1.2 (Correspondencia de Robinson-Schensted-Knuth) Sean n y q enteros positivos. Existe una relación biunívoca entre los siguientes conjuntos

$$M_q(n) \Leftrightarrow W_n^q \Leftrightarrow T_n^{n,q}.$$

1.1.2. La subsucesión no decreciente más grande

La historia del estudio de la subsucesión no decreciente más grande comienza con el matemático Polaco–Estadounidense Stanislaw Ulam (1909-1984). Durante la segunda guerra mundial, John von Neumann sugirió que Ulam se uniera al Proyecto Manhattan en Los Alamos. Durante su estancia, Ulam sugirió el uso de los métodos Monte Carlo para resolver algunos problemas complejos, que surgieron durante la construcción de la bomba atómica. Años después en 1961, Ulam quería mostrar la utilidad de los métodos Monte Carlo [28]. Para ello propuso una variable que es inaccesible con métodos deterministas pero sí con métodos estadísticos. Dicha variable es definida a continuación: sea S_n el conjunto de permutaciones de los enteros $1, 2, \dots, n-1, n$, para $\sigma_n \in S_n$ definamos a la longitud de la subsucesión no decreciente más grande de σ_n por

$$LI(\sigma_n) = \max\{s : \exists 1 \leq i_1 < i_2 < \dots < i_s \leq n, \text{ con } \sigma_{i_1} \leq \sigma_{i_2} \leq \dots \leq \sigma_{i_s}\}.$$

Si queremos calcular el valor promedio de la subsucesión no decreciente más grande tenemos que calcular $LI(\sigma_n)$ para cada elemento de S_n , pero recordemos que la cardinalidad de este conjunto es $n!$, de modo que dicho cálculo sólo sería posible para valores pequeños de n . Por otra parte si calculamos $LI(\sigma_n)$ para solo algunos miles de permutaciones, seleccionadas de manera uniforme, podemos obtener una buena estimación del valor promedio. De esta forma Ulam conjeturó que

$$\mathbb{E}[LI(\sigma_n)] \approx 1.7\sqrt{n} + \text{Gaussiana}, \quad (1.11)$$

conjetura que resultó ser falsa, pero esto no se supo si no hasta 1999 con los trabajos de Baik, Deift y Johansson [2]. De aquí en adelante al problema de encontrar la descripción asintótica de la variable $LI(\sigma_n)$ se le conoce como el problema de Ulam. La solución al problema de Ulam se dió en 1999 y fué dada por Baik, Deift y Johansson en [2]. Ellos demostraron lo siguiente

Teorema 1.1.3 (Daik, Deift y Johansson (1999))

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{LI(\sigma_n) - 2\sqrt{n}}{n^{\frac{1}{6}}} \leq t \right) = F_{TW}(t), \quad \text{para toda } t \in \mathbb{R}.$$

Donde F_{TW} es la distribución de Tracy-Widom.

La distribución de Tracy-Widom surgió por primera vez en los pioneros trabajos de Tracy y Widom [25, 26], como la distribución límite de los valores propios más grandes de matrices GUE^2 . La distribución F_{TW} se define de la siguiente manera

$$F_{TW}(t) = \exp \left\{ \int_t^\infty (x-t)^2 q(x) dx \right\},$$

²Las matrices GUE se definió en la sección 1.2.

donde q es la solución a la ecuación de Painlavé tipo II

$$q''(s) = s q(s) + 2q(s)^3,$$

tal que

$$q(s) \sim Ai(s), \text{ cuando } s \rightarrow \infty,$$

con Ai la función de Airy.

El resultado de Baik, Deift y Johansson nos muestra una relación entre LI_n y los valores propios más grandes de matrices GUE. En el capítulo 2 de esta tesis abordaremos el problema de Ulam, pero en el caso de palabras aleatorias. Intuitivamente, podemos pensar que una palabra de tamaño n en un alfabeto de n letras, es asintóticamente ($n \rightarrow \infty$) una permutación. De modo que una propuesta natural para el límite de LI_n cuando $n \rightarrow \infty$, en el caso de palabras aleatorias en un alfabeto con m letras, es el valor propio más grande de una matriz GUE de dimensión m . Daremos más detalles sobre este resultado en el capítulo 2.

1.1.3. LI_n y la correspondencia RSK

Definamos a la subsucesión no decreciente más grande para permutaciones, palabras y permutaciones generalizadas. Repetiremos la definición para permutaciones con el fin de ilustrar lo análogo de las definiciones.

i) **Permutaciones** Sea $\sigma_n \in S_n$

$$LI(\sigma_n) = \text{máx}\{s : \exists 1 \leq i_1 < i_2 < \dots < i_s \leq n, \text{ con } \sigma_{i_1} < \sigma_{i_2} < \dots < \sigma_{i_s}\}.$$

ii) **Palabras** Sea $\omega_n \in S_n^q$

$$LI(\omega_n) = \text{máx}\{s : \exists 1 \leq i_1 < i_2 < \dots < i_s \leq n, \text{ con } \omega_{i_1} \leq \omega_{i_2} \leq \dots \leq \omega_{i_s}\}.$$

iii) **Permutaciones Generalizadas** Sea $\pi_n \in GP_n^{p,q}$

$$LI(\pi_n) = \text{máx}\{s : \exists 1 \leq i_1 \leq i_2 \leq \dots \leq i_s \leq n, \text{ con } \pi_{i_1} \leq \pi_{i_2} \leq \dots \leq \pi_{i_s}\}.$$

En la sección 1.1.1 vimos que podemos representar a las permutaciones generalizadas como matrices con entradas enteras o como pares de tabloides. Surge entonces la pregunta ¿Cuál es el análogo de la sucesión no decreciente más grande en matrices con entradas enteras y en pares de tablas y tabloides? Quizás la forma mas simple de responder esta pregunta es con el siguiente.

Ejemplo: Para $n = 10$, $p = 4$ y $q = 3$ el algoritmo RSK genera la siguiente

equivalencia

$$GP_n^{p,q} \ni \pi = \left(\begin{array}{cccccccc} 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 4 & 4 \\ 2 & 3 & 3 & \underline{1} & \underline{2} & \underline{2} & 1 & \underline{2} & 1 & \underline{3} \end{array} \right), \quad LI(\pi) = 5,$$

\Updownarrow

$$(P, Q) = \left(\begin{array}{|c|c|c|c|c|} \hline 1 & 1 & 1 & 2 & 3 \\ \hline 2 & 2 & 2 & & \\ \hline 3 & 3 & & & \\ \hline \end{array}, \begin{array}{|c|c|c|c|c|} \hline 1 & 1 & 1 & 3 & 4 \\ \hline 2 & 2 & 2 & & \\ \hline 3 & 4 & & & \\ \hline \end{array} \right), \quad \lambda_1 = 5,$$

\Updownarrow

$$M = \begin{pmatrix} 0 & \rightarrow & 1 & & 2 \\ & & \downarrow & & \\ 1 & & 2 & & 0 \\ & & \downarrow & & \\ 1 & & 1 & \rightarrow & 0 \\ & & & & \downarrow \\ 1 & & 0 & & 1 \end{pmatrix}, \quad \sum_{(i,j) \in \{\text{camino}\}} w_{ij} = 5.$$

El hecho de que las longitudes de la subsucesión no decreciente más grande, la longitud del primer renglón en el diagrama de Young y la suma de las entradas sobre el camino más pesado que va de la entrada (1, 1) a la entrada (4, 3) y que sólo da pasos a la derecha y hacia abajo en nuestro ejemplo sea igual, no es coincidencia. Se puede demostrar que esto siempre ocurre³, de hecho esta relación será muy importante en el siguiente capítulo, pues es lo que permitirá relacionar al movimiento browniano con el espectro de una matriz aleatoria.

Si la longitud del primer renglón en el diagrama de Young es igual a la longitud de la subsucesión no decreciente más grande, ¿se le podrá dar una interpretación similar a las longitudes de los otros renglones? La respuesta a esta pregunta es si, y para hacer esto se generalizará la definición de LI , de la siguiente manera:

Para

$$\pi = \begin{pmatrix} i_1 & \cdots & i_n \\ j_1 & \cdots & j_n \end{pmatrix} \in GP_n^{p,q},$$

denotemos por

$$G_\pi(n, k) = \max \left\{ r : \begin{array}{l} \exists k\text{-subsucesiones no decrecientes} \\ \text{ajenas de } \pi \text{ tales que la suma de sus} \\ \text{longitudes es } r. \end{array} \right\}$$

Observemos que en particular

$$G_\pi(n, 1) = LI(\pi).$$

El siguiente resultado será de gran utilidad en el próximo capítulo, su demostración puede ser consultada en [5].

³Ver [5].

Lema 1.1.1 *Sea $\pi \in GP_n^{p,q}$ y $\lambda = (\lambda_1, \lambda_2, \dots)$ la forma del diagrama de Young que se obtiene al aplicar el algoritmo RSK a π . Entonces para toda k entero positivo se cumple que*

$$G_\pi(n, k) = \lambda_1 + \dots + \lambda_k.$$

También podemos interpretar a $G_\pi(n, k)$ en términos de la matriz $M = (m_{ij})$ que genera el algoritmo RSK al aplicarlo a π . Para ello serán necesarias las siguientes definiciones:

Definición 1.1.1 (Caminos $\downarrow \rightarrow$) *Un camino $\downarrow \rightarrow C$ de longitud l sobre una matriz $M \in M_{n,k}$, es un conjunto $C = \{(i_1, j_1), \dots, (i_l, j_l)\}$ de tal manera que para cualquier $2 \leq r \leq l$ se tiene que $(i_r, j_r) - (i_{r-1}, j_{r-1}) \in \{(1, 0), (0, 1)\}$.*

Observemos que las entradas de M que están en un camino $\downarrow \rightarrow$ inducen, una subsucesión no decreciente. Ilustraremos esto con un ejemplo.

$$M = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ & \downarrow & & & \\ 1 & 0 & 1 & 0 & 1 \\ & \downarrow & & & \\ 0 & 0 & 0 & 1 & 1 \\ & \downarrow & & & \\ 0 & 2 & \rightarrow & 1 & 0 & 0 \end{pmatrix} \quad \pi = \left(\begin{array}{cccccccccccc} 1 & 1 & 2 & 2 & 2 & 3 & 3 & 4 & 4 & 4 \\ \underline{2} & 4 & 1 & 3 & 5 & 4 & 5 & \underline{2} & \underline{2} & \underline{3} \end{array} \right)$$

El camino $\downarrow \rightarrow$ induce la subsucesión subrayada.

Como se observó anteriormente, para encontrar la longitud de la subsucesión no decreciente más grande basta con buscar sobre este tipo de caminos. De modo que $G_\pi(n, k)$ se puede interpretar de la siguiente manera

$$G_\pi(n, k) = \max \left\{ \ell_1 + \dots + \ell_k : \begin{array}{l} \exists k\text{-caminos } \downarrow \rightarrow \text{ ajenos de} \\ \text{longitudes } \ell_1, \dots, \ell_k \end{array} \right\}.$$

Ilustraremos esta interpretación con el siguiente ejemplo.

Ejemplo: Sea $\pi \in GP_{10}^{4,5}$, observemos los siguiente

$$\pi = \left(\begin{array}{ccccccccc} 1 & 1 & 2 & 2 & 2 & 3 & 3 & 4 & 4 & 4 \\ \underline{2} & \underline{4} & \tilde{1} & \hat{3} & \underline{5} & \hat{4} & \underline{5} & \tilde{2} & \tilde{2} & \tilde{3} \end{array} \right), \quad G_\pi(10, 3) = 10,$$

$$\Updownarrow$$

$$(P, Q) = \left(\begin{array}{|c|c|c|c|} \hline 1 & 2 & 5 & 7 \\ \hline 3 & 4 & 6 & 10 \\ \hline 8 & 9 & & \\ \hline \end{array}, \begin{array}{|c|c|c|c|} \hline 1 & 1 & 2 & 3 \\ \hline 2 & 2 & 3 & 4 \\ \hline 4 & 4 & & \\ \hline \end{array} \right), \quad \lambda_1 + \lambda_2 + \lambda_3 = 10,$$

$$\Updownarrow$$

$$M = \begin{pmatrix} 0 & \rightarrow & 1 & \rightarrow & 0 & \rightarrow & 1 & \rightarrow & 0 \\ & & & & & & & & \downarrow \\ 1 & & 0 & & 1 & \rightarrow & 0 & & 1 \\ \downarrow & & & & & & \downarrow & & \downarrow \\ 0 & & 0 & & 0 & & 1 & & 1 \\ \downarrow & & & & & & & & \\ 0 & \rightarrow & 2 & \rightarrow & 1 & & 0 & & 0 \end{pmatrix}, \quad \sum_{(i,j) \in \{\text{caminos}\}} w_{ij} = 10.$$

Es decir, si $\pi \in GP_n^{p,q}$ y $\lambda = (\lambda_1, \dots, \lambda_{\ell(\lambda)})$ es la forma del diagrama de Young que se obtiene al aplicar el algoritmo RSK a π , entonces π se puede descomponer en $\ell(\lambda)$ subsucesiones no decrecientes ajenas de longitudes $\lambda_1, \dots, \lambda_{\ell(\lambda)}$.

Sean C_1, \dots, C_k caminos $\downarrow \rightarrow$ disjuntos en $M \in \mathbb{M}_n^{p,q}$, denotemos por $|C_1|, \dots, |C_k|$ sus pesos, es decir

$$|C_i| = \sum_{(i,j) \in C_i} m_{i,j}. \quad (1.12)$$

Entonces con esta notación vemos que

$$G_\pi(n, k) = \max \left\{ |C_1| + \dots + |C_k| : \begin{array}{l} C_1, \dots, C_k \text{ caminos } \downarrow \rightarrow \text{ disjuntos} \\ \text{en } M. \end{array} \right\}$$

En 2015 Benaych-Georges y Houdré [4] demostraron lo siguiente.

Lema 1.1.2 *Para cualesquiera C_1, \dots, C_k caminos $\downarrow \rightarrow$ existen $\tilde{C}_1, \dots, \tilde{C}_k$ caminos $\downarrow \rightarrow$ que comienzan y terminan en el primer y último renglón de M tales que*

$$|C_1| + \dots + |C_k| = |\tilde{C}_1| + \dots + |\tilde{C}_k|.$$

Para el ejemplo anterior ilustramos los caminos \tilde{C}_1 , \tilde{C}_2 y \tilde{C}_3 .

$$M = \begin{pmatrix} 0 & & 1 & \rightarrow & 0 & & 1 & \rightarrow & 0 \\ \downarrow & & & & \downarrow & & & & \downarrow \\ 1 & & 0 & & 1 & \rightarrow & 0 & & 1 \\ \downarrow & & & & & & \downarrow & & \downarrow \\ 0 & & 0 & & 0 & & 1 & & 1 \\ \downarrow & & & & & & \downarrow & & \downarrow \\ 0 & \rightarrow & 2 & \rightarrow & 1 & & 0 & & 0 \end{pmatrix}$$

Esta clase de caminos juega un papel esencial en la identificación de la distribución asintótica de la subsucesión no decreciente más grande en palabras aleatorias, en términos de un funcional del movimiento browniano. Por este motivo es necesaria la siguiente notación: para enteros positivos n y k denotaremos

$$\widehat{\Pi}_{n,k} = \left\{ \pi : \begin{array}{l} \pi \text{ es un camino } \downarrow \rightarrow \text{ en } \{1, \dots, n\} \times \{1, \dots, k\}, \\ \text{tal que empieza en } \{1\} \times \{1, \dots, k\} \\ \text{y termina en } \{n\} \times \{1, \dots, k\} \end{array} \right\}. \quad (1.13)$$

En el capítulo 2, determinaremos a la distribución asintótica de la forma del diagrama de Young asociado a palabras aleatorias. Para ello, será necesario considerar a varios caminos $\downarrow \rightarrow$ que sean disjuntos. El que los caminos $\downarrow \rightarrow$ sean disjuntos, es una propiedad de los elementos de $\widehat{\Pi}_{n,k}$ que será de importancia fundamental en el capítulo 2, por este motivo presentamos la siguiente definición.

Definición 1.1.2 Sean $\widehat{\pi}_1, \widehat{\pi}_2 \in \widehat{\Pi}_{n,k}$ diremos que $\widehat{\pi}_1 < \widehat{\pi}_2$, si para toda $r = 1, \dots, n$, $(r, s) \in \widehat{\pi}_1$ y $(r, l) \in \widehat{\pi}_2$ se cumple que $s < l$. Es decir, $\widehat{\pi}_1 < \widehat{\pi}_2$ si $\widehat{\pi}_2$ está siempre a la derecha de $\widehat{\pi}_1$.

En términos de $\widehat{\Pi}_{n,k}$ y de la definición 1.1.2 podemos reescribir el lema 1.1.1 de la siguiente manera.

Lema 1.1.3 Denotemos por $\lambda = (\lambda_1, \lambda_2, \dots)$ a la forma del diagrama de Young que se obtiene al aplicar el algoritmo RSK a $\pi \in GP_n^{p,q}$, y por $M \in \mathbb{M}_n^{p,q}$ a su matriz asociada. Entonces para toda k entero positivo se cumple que

$$\lambda_1 + \dots + \lambda_k = \max_{\substack{\widehat{\pi}_1, \dots, \widehat{\pi}_k \in \widehat{\Pi}_{p,q} \\ \widehat{\pi}_1 < \dots < \widehat{\pi}_k}} \sum_{i=1}^k |\widehat{\pi}_i|.$$

La demostración del lema 1.1.3 puede ser consultada en [4].

1.2. Matrices aleatorias y su espectro

Una matriz aleatoria M de dimensión n es un elemento aleatorio que toma valores en un conjunto Q de matrices de dimensiones $n \times n$. En esta tesis sólo consideraremos matrices cuadradas. Recordemos que a Q lo podemos identificar con algún subconjunto de \mathbb{R}^{n^2} (o \mathbb{C}^{n^2}), y también que una función que toma valores en \mathbb{R}^{n^2} (o \mathbb{C}^{n^2}), es medible si y sólo si cada una de sus entradas es medible. Por lo que podemos representar a la matriz aleatoria M como

$$M = \begin{pmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & & \vdots \\ m_{n1} & \cdots & m_{nn} \end{pmatrix},$$

donde las variables $\{m_{ij}\}_{i,j=1}^n$ son variables aleatorias con valores reales o complejas.

1.2.1. El ensamble gaussiano unitario (GUE)

Consideremos dos familias $\{X_{ij}\}_{j \geq i \geq 1}^n$, $\{Y_{ij}\}_{j \geq i \geq 1}^n$ de variables aleatorias independientes e independientes entre si, idénticamente distribuidas como $N\left(0, \frac{1}{2}\right)$.

Definición 1.2.1 *El ensamble gaussiano unitario (GUE) es la sucesión de matrices aleatorias $\{H_n = (h_{ij})_{i,j=1}^n\}_{n \geq 1}$ definidas de la siguiente manera*

$$h_{ij} = \begin{cases} h_{jj} = X_{jj} + Y_{jj} & \text{si } i = j \\ h_{ij} = X_{ij} + iY_{ij} & \text{si } i < j \\ h_{ji} = h_{ij} & \text{si } i > j \end{cases}$$

Usaremos la notación $H_n \sim \text{GUE}$ para hacer referencia a que la matriz H_n es el n -ésimo elemento del GUE.

Otra forma de dar una descripción del GUE es a través de la densidad de cada una de las matrices que lo conforman, que en este caso es

$$f_{\text{GUE}}(H) = \frac{1}{2^{\frac{n}{2}} \pi^{\frac{n^2}{2}}} \exp\left\{-\frac{\text{Tr}(H_n^2)}{2}\right\}.$$

Recordemos que el conjunto \mathbb{H}_n de las matrices hermitianas de dimensión n se define por

$$\mathbb{H}_n = \{H \in M_{n \times n}(\mathbb{C}) : H^* = H\}, \quad (1.14)$$

donde H^* es la transpuesta conjugada de H .

Observemos que si $H_n \sim \text{GUE}$ entonces $H_n \in \mathbb{H}_n$, de modo que H_n es diagonalizable y además sus valores propios son reales. Al conjunto de los valores propios $\lambda_1(H) \geq \dots \geq \lambda_n(H)$ de una matriz M lo llamaremos espectro ordenado de M . Cuando no haya lugar a confusión usaremos la notación $\lambda_1, \dots, \lambda_n$.

También se puede dar una descripción del espectro de una matriz GUE en términos de su densidad conjunta, por ejemplo Anderson, Guionnet y Zeitouni [1], dan la siguiente expresión para el espectro ordenado del GUE:

$$f(\lambda_1, \dots, \lambda_n) = \frac{1}{(2\pi)^{\frac{n}{2}} \prod_{i=1}^n i!} \exp\left\{-\frac{\lambda_1^2 + \dots + \lambda_n^2}{2}\right\} \prod_{j>i} |\lambda_j - \lambda_i|^2,$$

donde c_n es una constante tal que $\int_{\mathbb{R}^n} f_{\text{GUE}}(\lambda_1, \dots, \lambda_n) d\lambda_1 \dots d\lambda_n = 1$.

1.2.2. GUE sin traza y GUE por bloques sin traza

Otro de los conjuntos de matrices aleatorias que han sido de mucha utilidad es el GUE sin traza. Decimos que una matriz \widetilde{H}_n de dimensiones $n \times n$ es un GUE sin traza si

$$\widetilde{H}_n = H_n - \frac{\text{tr}(H_n)I_n}{n}, \quad (1.15)$$

donde $H_n \sim \text{GUE}$. Observemos que $\text{tr}(\widetilde{H}_n) = 0$ esto es lo que le da el nombre GUE sin traza.

1.3. EL MOVIMIENTO BROWNIANO Y EL FUNCIONAL DE GLYNN Y WHITT23

Para los fines de esta tesis será conveniente generalizar este concepto. Diremos que una matriz H es un GUE por bloques sin traza, si existen matrices H_{n_1}, \dots, H_{n_k} GUE sin traza, no necesariamente de la misma dimensión tal que

$$H = \begin{pmatrix} H_{n_1} & & 0 \\ & \ddots & \\ 0 & & H_{n_k} \end{pmatrix}.$$

Otra forma de generalizar al GUE sin traza que además generaliza al GUE por bloques sin traza es la siguiente: consideremos números no negativos p_1, \dots, p_n tales que $p_1 + \dots + p_n = 1$. Denotemos por $p^{(1)}, \dots, p^{(k)}$ a los k valores distintos en p_1, \dots, p_n , supongamos que d_1, \dots, d_k son las diferentes multiplicidades de $p^{(1)}, \dots, p^{(k)}$, respectivamente. Denotemos

$$J = \text{diag}(\sqrt{p_1}, \dots, \sqrt{p_n}),$$

y tomemos

$$H = \begin{pmatrix} H_{d_1} & & 0 \\ & \ddots & \\ 0 & & H_{d_k} \end{pmatrix}.$$

Entonces una matriz \widetilde{H}_n es un GUE por bloques sin traza generalizado (o simplemente block traceless generalizado) si se puede escribir de la siguiente forma

$$\widetilde{H}_n = H_n - \text{tr}(H_n J)J. \quad (1.16)$$

Como caso particular observemos que si $p_1 = \dots = p_n = \frac{1}{n}$ entonces (1.16) y (1.15) coinciden.

1.3. El movimiento browniano y el funcional de Glynn y Whitt

El objetivo de esta sección es introducir al movimiento browniano y a los funcionales que nos serán de interés. Las herramientas combinatorias expuestas en la sección anterior nos permitirán darle una interpretación a nuestros funcionales en términos de teoría de colas.

1.3.1. El movimiento browniano

Recordemos que el movimiento browniano estándar $(B_t)_{t \geq 0}$ es el un proceso con trayectorias continuas que cumple lo siguiente

- a) $B_0 = 0$ con probabilidad 1.
- b) Para toda $t > s$ la variable $B_t - B_s$ es independiente de $(B_u)_{0 \leq u \leq s}$.
- c) Para toda $t > s$ la variable $B_t - B_s$ tiene distribución normal con media cero y varianza $t - s$.

Existe una gran cantidad de bibliografía sobre este proceso estocástico, para el lector interesado en este tema recomendamos [22] o [16]. También diremos que un proceso B con valores en \mathbb{R}^n es un movimiento browniano estándar n dimensional si las coordenadas de B son movimientos brownianos estándar e independientes.

La base de esta tesis es la correspondencia RSK, por lo que necesitaremos una herramienta para pasar de una estructura discreta al movimiento browniano. Esta herramienta nos la da el Teorema de Donsker, el cual nos dice que: bajo ciertas condiciones técnicas las caminatas aleatorias vistas como elementos aleatorios de

$$C[0, \infty) := \{f : [0, \infty) \rightarrow \mathbb{R} \mid f \text{ es continua} \},$$

con respecto a su σ -álgebra de borel $\mathcal{B}(C[0, \infty))$ converge al movimiento browniano.

Teorema 1.3.1 (Donsker) *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, y en él una sucesión $(\xi_n)_{n \geq 1}$ de variables i.i.d. con media μ y varianza $0 < \sigma^2 < \infty$. Definamos*

$$X_t^{(n)} = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^n (\xi_i - \mu) + \frac{(t - \lfloor t \rfloor)\xi_{\lfloor t \rfloor + 1}}{\sigma},$$

y sea P_n la medida que induce $X^{(n)}$ en $(C[0, \infty), \mathcal{B}(C[0, \infty)))$. Entonces la sucesión $(P_n)_{n \geq 1}$ converge débilmente a una medida P_ bajo la cual, el proceso $W_t(\omega) = \omega(t)$ en $(C[0, \infty), \mathcal{B}(C[0, \infty)))$ es un movimiento browniano estándar.*

La demostración del teorema de Donsker puede consultarse en [16].

1.3.2. El funcional de Glynn y Whitt

Consideremos el siguiente esquema. Supongamos que tenemos n servidores en serie y k clientes formados en el primer servidor. Denotemos por ω_{ij} al tiempo de servicio del cliente i en el servidor j . Supondremos que las variables ω_{ij} son i.i.d. con media μ y varianza σ^2 . Los clientes y servidores siguen la siguiente dinámica. El primer cliente entra a servicio y espera un tiempo ω_{11} para salir, cuando sale comienza a ser atendido el segundo cliente en el primer servidor y el primer cliente en el segundo servidor. Si $\omega_{12} > \omega_{21}$ el primer cliente pasa al tercer servidor, si $\omega_{12} < \omega_{21}$ el segundo cliente se forma en el segundo servidor y espera su turno, mientras que el tercer cliente comienza a ser atendido en el primer servidor, etc.

Denotemos por:

$$D(n, k) = \begin{array}{l} \text{Tiempo en que el } k \text{ ésimo cliente termina} \\ \text{de ser atendido en el } n \text{ ésimo servidor} \end{array}$$

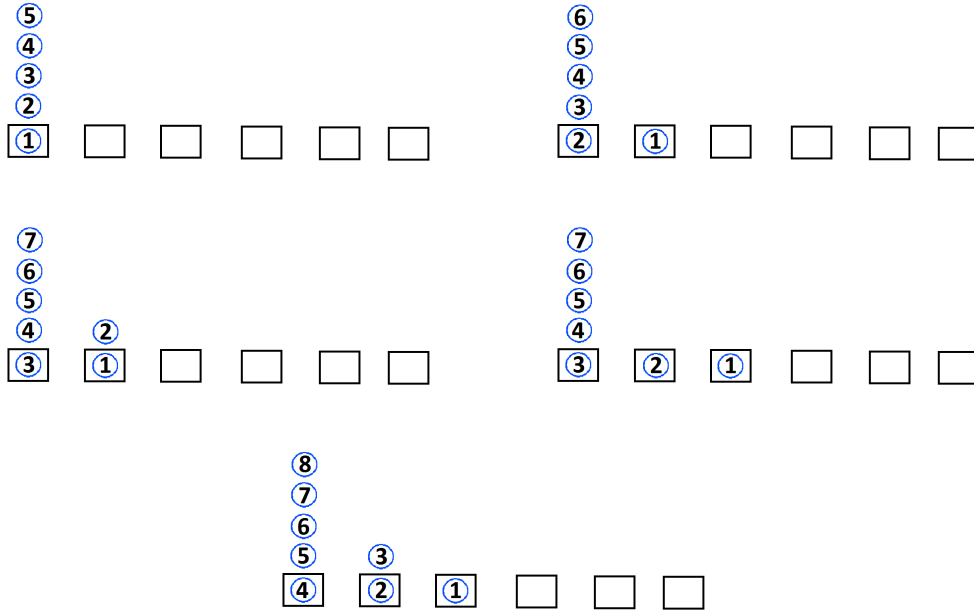
De la dinámica descrita anteriormente se puede deducir que

$$D(n, k) = \max\{D(n-1, k), D(n, k-1)\} + \omega_{nk},$$

tomando $D(n, 0) = D(0, k) = 0$. Usando inducción se puede demostrar⁴ que

$$D(n, k) = \max_{\pi \in \Pi_{n,k}} \left\{ \sum_{(i,j) \in \pi} w_{ij} \right\}, \quad (1.17)$$

⁴Ver [6].



donde $\Pi_{n,k}$ es el conjunto de caminos $\downarrow \rightarrow$ que empiezan en $(1, 1)$ y terminan en (n, k) . Sea

$$D_k^{(n)} = \frac{D(n, k) - \mu n}{\sigma \sqrt{n}}. \quad (1.18)$$

en 1991 Glynn y Whitt [6] probaron el siguiente resultado.

Teorema 1.3.2 *El proceso $D_k^{(n)} = (D_k^{(n)}, k \geq 1)$ definido en (1.18) converge en ley a un proceso $D_k = (D_k, k \geq 1)$, tal que*

$$D_k \stackrel{d}{=} \sup_{0=t_0 \leq \dots \leq t_k=1} \sum_{i=1}^k (B_i(t_i) - B_i(t_{i-1})), \quad (1.19)$$

donde para cada $k \geq 1$ se tiene que (B_1, \dots, B_k) es un movimiento browniano estándar k dimensional.

Al funcional del movimiento browniano que aparece en (1.19) le llamaremos funcional de Glynn y Whitt. En el siguiente capítulo daremos una demostración de este resultado. De hecho, uno de los principales resultados expuestos en esta tesis es una generalización de este teorema.

Diez años después de los trabajos de Glynn y Whitt, Baryshnikov demostró en [3] que el proceso formado por las variables D_k es igual en distribución al proceso que forman los valores propios más grandes de matrices GUE. El resultado preciso es el siguiente.

Teorema 1.3.3 (Baryshnikov) *Sea $(H_n)_{n \geq 1} \sim GUE$ y para $1 \leq k \leq n$ denotemos por μ_1^n al valor propio más grande de H_n . Sea $\{B_n = (B_n(t); t \geq 0)\}_{n \geq 1}$ una sucesión de movimientos brownianos estándar independientes, entonces para toda $n \geq 1$*

$$\mu_1^n \stackrel{d}{=} \sup_{0=t_0 \leq \dots \leq t_k=1} \sum_{i=1}^k (B_i(t_i) - B_i(t_{i-1})).$$

1.4. GUE como límite de diagramas de Young

En esta sección presentamos un resultado de Baryshnikov [3], el cual nos proporciona una forma de aproximar al espectro de los menores principales de una matriz GUE, a partir de la forma del diagrama de Young asociado a una permutación aleatoria. Un primer resultado en esta dirección se debe a Johansson [13], el cual presentamos a continuación.

Proposición 1.4.1 *Consideremos el siguiente diagrama*

$$(\Omega, \mathcal{F}, \mathbb{P}) \xrightarrow{X_W} (M_{n,m}, \Sigma_1, \mathbb{P}_{X_W}) \xrightarrow{sh} (\mathcal{P}_m, \Sigma_2, \mathbb{P}_{sh})$$

Donde $(\Omega, \mathcal{F}, \mathbb{P})$ es un espacio de probabilidad tal que X_W es una matriz aleatoria con valores en $M_{n,m}$, de tal forma que las entradas de X_W sean independientes y con distribución geométrica de parámetro $0 < q < 1$, Σ_1 y Σ_2 son las respectivas σ -álgebras potencia, \mathbb{P}_{X_W} y \mathbb{P}_{sh} son las medidas inducidas por X_W y sh respectivamente. Entonces la variable $sh(X_W)$ se distribuye de la siguiente manera:

$$\mathbb{P}_{sh}[\{\lambda\}] = (1-q)^{nm} \prod_{i=0}^{m-1} \frac{(n-m)!}{j!(n-m+i)!} \prod_{1 \leq i < j \leq m} (\lambda_i - \lambda_j + j - i)^2 \prod_{i=1}^m \binom{\lambda_i + n - i}{\lambda_i + m - i} q^{\lambda_i}, \quad (1.20)$$

donde $\lambda \in \mathcal{P}_m$.

Demostración:

Para la demostración será necesario considerar a la correspondencia RSK, es decir se tiene el siguiente diagrama

$$(\Omega, \mathcal{F}, \mathbb{P}) \xrightarrow{X_W} (M_{n,m}, \Sigma_1, \mathbb{P}_{X_W}) \xrightarrow{sh} (\mathcal{P}_m, \Sigma_2, \mathbb{P}_{sh})$$

$$\downarrow RSK$$

$$(\bigcup_{k=1}^{\infty} T_k^{n,m}, \Sigma_3, \mathbb{P}_{RSK}),$$

donde Σ_3 es la σ -álgebra potencia y \mathbb{P}_{RSK} la medida inducida por la correspondencia RSK.

Observemos que para $A \in M_{n,m}$ se tiene que

$$\mathbb{P}_{X_W}[\{A\}] = \prod_{i=1}^n \prod_{j=1}^m (1-q)q^{a_{ij}} = (1-q)^{nm} q^{|A|},$$

donde

$$|A| = \sum_{i=1}^n \sum_{j=1}^m A_{ij}.$$

De lo cual se sigue que la medida $\mathbb{P}_{X_W}(\cdot | M_{n,m}(k))$ es la medida uniforme en el conjunto $M_{n,m}(k)$. Esta es la razón de pedir que las entradas de X_W sean variables aleatorias geométricas del mismo parámetro. En particular, como la suma de variables aleatorias independientes, con distribución geométrica del mismo parámetro tiene distribución binomial negativa, se tiene que

$$\mathbb{P}_{X_W}[M_{n,m}(k)] = \binom{nm - k - 1}{k} (1-q)^{nm} q^k. \quad (1.21)$$

De la correspondencia RSK se sigue la igualdad de eventos

$$RSK^{-1}(T_k^{n,m}) = M_{n,m}(k). \quad (1.22)$$

Entonces de (1.21) y (1.22) se sigue que

$$\mathbb{P}_{RSK}[T_k^{n,m}] = \mathbb{P}_{X_W}[RSK^{-1}(T_k^{n,m})] = \binom{nm - k - 1}{k} (1 - q)^{nm} q^k. \quad (1.23)$$

Observemos que

$$\#M_{n,m}(k) = \binom{nm + k - 1}{k}. \quad (1.24)$$

Luego por ley de la probabilidad total

$$\begin{aligned} \mathbb{P}_{sh}[\{\lambda\}] &= \mathbb{P}_{X_W}[sh^{-1}(\{\lambda\})] \\ &= \sum_{k=0}^{\infty} \mathbb{P}_{X_W}[sh^{-1}(\{\lambda\}) | M_{n,m}(k)] \mathbb{P}_{X_W}[M_{n,m}(k)] \\ &= \mathbb{P}_{X_W}[sh^{-1}(\{\lambda\}) | M_{n,m}(|\lambda|)] \mathbb{P}_{X_W}[M_{n,m}(|\lambda|)] \\ &= \frac{\#T_n(\lambda) \times T_m(\lambda)}{\binom{nm+|\lambda|-1}{|\lambda|}} \binom{nm + |\lambda| - 1}{|\lambda|} (1 - q)^{nm} q^{|\lambda|} \\ &= \#T_n(\lambda) \times T_m(\lambda) \cdot (1 - q)^{nm} q^{|\lambda|}. \end{aligned} \quad (1.25)$$

Para llegar a la cuarta igualdad, utilizamos (1.21), (1.24) y que la medida $\mathbb{P}_{X_W}[\cdot | M_{n,m}(|\lambda|)]$ es la medida uniforme en $M_{n,m}(|\lambda|)$.

Es posible demostrar que

$$\#T_n(\lambda) = \prod_{1 \leq i < j \leq n} \frac{\lambda_i - \lambda_j + j - i}{j - i}, \quad (1.26)$$

ver [13, 29].

Sin pérdida de generalidad supondremos que $n > m$, luego de la definición de $\mathcal{P}_m^{(n)}$ se sigue que $\lambda_i = 0$ para $m < i \leq n$. De esta observación, haciendo el cambio de variable $h_i = \lambda_i + m - i$ para $i = 1, \dots, m$, y haciendo uso de (1.26), podemos calcular lo siguiente

$$\begin{aligned} \#T_n(\lambda) \times T_m(\lambda) &= \prod_{1 \leq i < j \leq n} \frac{\lambda_i - \lambda_j + j - i}{j - i} \prod_{1 \leq i < j \leq m} \frac{\lambda_i - \lambda_j + j - i}{j - i} \\ &= \prod_{1 \leq i < j \leq m} \left(\frac{\lambda_i - \lambda_j + j - i}{j - i} \right)^2 \prod_{i=1}^m \prod_{j=m+1}^n \frac{\lambda_i + j - i}{j - i} \\ &= \prod_{1 \leq i < j \leq m} \frac{(h_j - h_i)^2}{(j - i)^2} \prod_{i=1}^m \prod_{j=m+1}^n \frac{h_i + j - m}{j - i} \\ &= \frac{1}{\prod_{i=1}^m \prod_{j=m+1}^n (j - i) \prod_{1 \leq i < j \leq m} (j - i)^2} \prod_{1 \leq i < j \leq m} (h_j - h_i)^2 \prod_{i=1}^m \frac{(h_i + n - m)!}{h_i!} \\ &= \prod_{i=0}^{m-1} \frac{1}{j!(n - m + i)!} \prod_{1 \leq i < j \leq m} (h_j - h_i)^2 \prod_{i=1}^m \frac{(h_i + n - m)!}{h_i!}. \end{aligned} \quad (1.27)$$

Para obtener la última igualdad se utilizaron las siguientes dos identidades

$$\prod_{1 \leq i < j \leq m} (j - i) = \prod_{j=0}^{m-1} j!. \quad (1.28)$$

$$\begin{aligned} & \prod_{i=1}^m \prod_{j=m+1}^n (j - i) \prod_{1 \leq i < j \leq m} (j - i) \\ &= \prod_{j=0}^{m-1} j! \prod_{i=1}^m (m + 1 - i)(m + 2 - i)(m + 3 - i) \cdots (n - i) \\ &= m! \frac{(m + 1)!}{1!} \frac{(m + 2)!}{2!} \cdots \frac{(n - 1)!}{(n - m - 1)!} 1! 2! \cdots (m - 1)! \\ &= \prod_{i=0}^{m-1} (n - m + i)!. \end{aligned} \quad (1.29)$$

Sustituyendo (1.27) en (1.25) obtenemos

$$\begin{aligned} \mathbb{P}_{sh}[\{\lambda\}] &= (1 - q)^{nm} q^{|\lambda|} \prod_{i=0}^{m-1} \frac{1}{j!(n - m + i)!} \prod_{1 \leq i < j \leq m} (h_j - h_i)^2 \prod_{i=1}^m \frac{(h_i + n - m)!}{h_i!} \\ &= (1 - q)^{nm} \prod_{i=0}^{m-1} \frac{(n - m)!}{j!(n - m + i)!} \prod_{1 \leq i < j \leq m} (\lambda_j - \lambda_i + j - i)^2 \prod_{i=1}^m \binom{\lambda_i + n - i}{\lambda_i + m - i} q^{\lambda_i}. \end{aligned}$$

■

A la expresión (1.20) se le conoce como el ensamble de Meixner [14]. La importancia de la proposición 1.4.1, es que nos permite demostrar que la forma del diagrama de Young ($sh(X_W)$) asociada a la matriz X_W converge al espectro del GUE.

Corolario 1.4.1 *Bajo las mismas hipótesis de la proposición 1.4.1, se tiene que*

$$\lim_{n \rightarrow \infty} \mathbb{P}_{sh} \left(\frac{\lambda_i - \frac{qn}{1-q}}{\frac{\sqrt{qn}}{1-q}} \leq y_i, \quad i = 1, \dots, m \right) = \frac{\int_A \prod_{1 \leq i < j \leq m} (x_i - x_j)^2 \prod_{i=1}^m e^{-\frac{x_i^2}{2}} dx_i}{\int_{\mathbb{R}^m} \prod_{1 \leq i < j \leq m} (x_i - x_j)^2 \prod_{i=1}^m e^{-\frac{x_i^2}{2}} dx_i},$$

donde $A = (\infty, y_1] \times \cdots \times (-\infty, y_m]$.

En 2001 Baryshnikov [3] publica un resultado que mejora al corolario 1.4.1. Baryshnikov considera a las siguientes variables: para un número entero positivo m , se X_W^k un elemento aleatorio en $M_{n,k}$ $1 \leq k \leq m$, y de tal forma que las entradas de X_W^k sean variables i.i.d. como geométricas de parámetro $0 < q < 1$. Denotemos por $\lambda^k = (\lambda_1^k, \dots, \lambda_k^k)$ a la forma del diagrama de Young que genera el algoritmo RSK cuando lo aplicamos a X_W^k . Definamos

$$\xi_i^k = \frac{\lambda_i^k - \frac{qn}{1-q}}{\frac{\sqrt{qn}}{1-q}},$$

con $1 \leq i \leq k \leq m$. Por otra parte consideremos a una matriz $H_m \sim \text{GUE}$ de dimensión m , denotemos por H_k para $1 \leq k \leq m$ a los menores principales de H_m . Para $1 \leq k \leq m$ denotemos por

$$\mu_1^k \geq \cdots \geq \mu_k^k,$$

al espectro ordenado de H_k .

Proposición 1.4.2 (Baryshnikov) *Con la notación anterior se tiene que*

$$\left(\begin{array}{cccc} & & \xi_1^1 & \\ & \cdots & \vdots & \cdots \\ \xi_1^m & & \cdots & \xi_m^m \end{array} \right) \xrightarrow{n \rightarrow \infty} \left(\begin{array}{cccc} & & \mu_1^1 & \\ & \cdots & \vdots & \cdots \\ \mu_1^m & & \cdots & \mu_m^m \end{array} \right).$$

El resultado de Baryshnikov es el último ingrediente necesario para demostrar el teorema que nos da la relación entre el espectro de H_m y las generalizaciones del funcional de Glynn-Whitt.

Capítulo 2

Matrices aleatorias y el movimiento browniano

En este capítulo se expone de manera pormenorizada los trabajos de Benaych-Georges y Houdré [4], y Houdré y Xu [11], sobre relaciones de funcionales del movimiento browniano y el espectro de los menores de un GUE, y con la distribución límite de la forma del diagrama Young asociado a palabras aleatorias.

Recordemos del capítulo 1, que LI_n coincide con la longitud del primer renglón del diagrama de Young, asociado a la misma palabra que LI_n . En particular, en este capítulo exponemos la forma de identificar a la distribución límite de LI_n como el valor propio más grande de una matriz aleatoria, y como un funcional del movimiento browniano. Al final del capítulo veremos que, la unicidad del límite de LI_n nos da una igualdad en distribución entre un funcional del movimiento browniano y el valor propio más grande de una matriz aleatoria.

2.1. Funcionales del movimiento browniano y el espectro de los menores principales de un GUE

En la sección 1.1.3 usamos a la correspondencia RSK para interpretar a los caminos $\downarrow \rightarrow$ sobre matrices en $M_{n,m}(N)$ como subsucesiones no decrecientes de elementos en $GP_N^{n,m}$. El lema 1.1.2, nos dice que cualquier colección de caminos $\downarrow \rightarrow$ puede ser sustituida por otra colección de caminos que sean elementos del conjunto $\widehat{\Pi}_{n,m}$. En este capítulo nos es de utilidad una versión continua de los caminos en $\widehat{\Pi}_{n,m}$, esta nueva clase de caminos es la siguiente.

$$\Pi_m = \{\pi : [0, 1] \rightarrow \{1, \dots, m\} \mid \pi \text{ es cádlág y no decreciente}\}. \quad (2.1)$$

Los caminos en Π_m pueden ser descritos de la siguiente forma: para cada camino $\pi \in \Pi_m$ existe una partición $0 = t_0 \leq \dots \leq t_m = 1$ del intervalo $[0, 1]$ de tal forma que

$$\pi(\cdot) = \sum_{i=1}^{m-1} i \mathbf{1}_{[t_{i-1}, t_i)}(\cdot) + m \mathbf{1}_{[t_{m-1}, t_m]}(\cdot). \quad (2.2)$$

Usando la representación (2.2), a cada camino $\pi \in \Pi_m$ le asociamos un funcional de la siguiente forma,

$$\Delta_\pi(F) = \sum_{i=1}^m (F_i(t_i) - F_i(t_{i-1})), \tag{2.3}$$

donde $F : [0, 1] \rightarrow \mathbb{R}^m$ con $F = (F_1, \dots, F_m)$.

El siguiente teorema se debe a Benaych-Georges y Houdre [4], y fue publicado en 2015. El teorema 2.1.1 presenta una generalización del teorema de Baryshnikov 1.3.3. La generalización es en el sentido de que nos da una igualdad en ley entre la suma de los valores propios mas grandes de un GUE y un funcional del movimiento browniano.

Teorema 2.1.1

Sea $B(t) = (B_1(t), \dots, B_m(t))_{t \in [0,1]}$ un movimiento browniano estándar m -dimensional y H_m una matriz GUE de dimensión m . Para $k = 1, \dots, m$ denotemos por

$$\mu_1^k \geq \dots \geq \mu_k^k,$$

al espectro del menor principal $k \times k$ de H_m . Entonces se tiene que

$$(\mu_1^k + \dots + \mu_l^k)_{1 \leq l \leq k \leq m} \stackrel{d}{=} \left(\sup_{\substack{\pi_1, \dots, \pi_l \in \Pi_k \\ \pi_1 < \dots < \pi_l \leq k}} \sum_{i=1}^l \Delta_{\pi_i}(B) \right)_{1 \leq l \leq k \leq m}. \tag{2.4}$$

Demostración:

Sea $\omega_{n,m} = (\omega_{ij}; 1 \leq i \leq n, 1 \leq j \leq m)$ un arreglo de variables aleatorias i.i.d. como geométricas de parámetro $0 < q < 1$ con

$$e = \mathbb{E}[w_{11}] \quad \text{y} \quad v = \text{Var}[w_{11}].$$

Para cada $k = 1, \dots, m$ denotemos por $\omega_{n,k} = (\omega_{ij}; 1 \leq i \leq n, 1 \leq j \leq k)$ al subarreglo que se obtiene con las primeras k columnas de $\omega_{n,m}$. Denotemos también por $\lambda^k = (\lambda_1^k, \dots, \lambda_k^k)$ a la forma del diagrama de Young que se obtiene al aplicar el algoritmo RSK al subarreglo $\omega_{n,k}$. Es decir, gráficamente

$$\omega_{n,m} = \begin{pmatrix} \omega_{11} & \dots & \omega_{1k} \\ \vdots & \ddots & \vdots \\ \omega_{n1} & \dots & \omega_{nk} \end{pmatrix} \rightsquigarrow \begin{array}{cccccc} \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square \end{array} = \begin{pmatrix} \lambda_1^k \\ \vdots \\ \lambda_k^k \end{pmatrix}.$$

De la proposición 1.4.2 es claro que

$$(\xi_l^k + \dots + \xi_l^k)_{1 \leq l \leq k \leq m} \xrightarrow{D} (\mu_l^k + \dots + \mu_l^k)_{1 \leq l \leq k \leq m}. \tag{2.5}$$

Recordemos que el lema 1.1.3, nos dice que para $1 \leq l \leq k \leq m$ se tiene que

$$\lambda_1^k + \dots + \lambda_l^k = \max_{\substack{\hat{\pi}_1, \dots, \hat{\pi}_l \in \hat{\Pi}_{n,k} \\ \hat{\pi}_1 < \dots < \hat{\pi}_l}} \sum_{i=1}^l |\hat{\pi}_i|, \tag{2.6}$$

donde $|\cdot|$, $\hat{\Pi}_{n,k}$ y “ $<$ ” se definen en (1.12), (1.13) y en la definición 1.1.2, respectivamente. Es decir, gráficamente

browniano de cada columna del arreglo normalizado, gráficamente

$$\omega_{n,k} = \begin{pmatrix} \frac{\omega_{11}-e}{\sqrt{vn}} & \dots & \frac{\omega_{1k}-e}{\sqrt{vn}} \\ \vdots & \ddots & \vdots \\ \frac{\omega_{n1}-e}{\sqrt{vn}} & \dots & \frac{\omega_{nk}-e}{\sqrt{vn}} \end{pmatrix}$$

Donsker \Downarrow \dots \Downarrow

$$(B_1, \dots, B_k).$$

Luego para obtener el límite del lado derecho en (2.8) usamos el teorema del mapeo continuo con el movimiento browniano obtenido de la normalización del arreglo $\omega_{n,m}$. Ahora para cada $1 \leq r \leq k$ definamos

$$\hat{B}_r^n(t) = \sum_{i=1}^{\lfloor nt \rfloor} \left(\frac{\omega_{ir} - e}{\sqrt{vn}} \right),$$

y denotemos $\hat{B} = (\hat{B}_1, \dots, \hat{B}_k)$.

Observemos que para cada $1 \leq r \leq k$ se tiene que

$$\sum_{(r,s) \in \hat{\pi}_r} \left(\frac{\omega_{r,s} - e}{\sqrt{vn}} \right) = \sum_{i=1}^k (\hat{B}_i^n(t_i) - \hat{B}_i^n(t_{i-1})),$$

para alguna partición $0 = t_0 \leq \dots \leq t_k = 1$ adecuada. Es decir, para cada $1 \leq r \leq k$ tenemos que

$$\sum_{(r,s) \in \hat{\pi}_u} \left(\frac{\omega_{r,s} - e}{\sqrt{vn}} \right) = \Delta_\pi(\hat{B}),$$

para alguna $\pi \in \Pi_k$. Por lo que el lado derecho de (2.8) se puede reescribir de la siguiente manera

$$\max_{\substack{\hat{\pi}_1, \dots, \hat{\pi}_l \in \hat{\Pi}_{n,k} \\ \hat{\pi}_1 < \dots < \hat{\pi}_l}} \sum_{r=1}^l \sum_{(i,j) \in \hat{\pi}_r} \left(\frac{\omega_{i,j} - e}{\sqrt{vn}} \right) - \frac{\mathbf{O}(kl)}{\sqrt{vn}} = \sup_{\substack{\pi_1, \dots, \pi_l \in \Pi_k \\ \pi_1 < \dots < \pi_l}} \sum_{r=1}^l \Delta_\pi(\hat{B}) - \frac{\mathbf{O}(kl)}{\sqrt{vn}}. \quad (2.9)$$

Luego de (2.8) y (2.9) obtenemos que

$$\xi_1^k + \dots + \xi_l^k = \sup_{\substack{\pi_1, \dots, \pi_l \in \Pi_k \\ \pi_1 < \dots < \pi_l}} \sum_{r=1}^l \Delta_\pi(\hat{B}) - \frac{\mathbf{O}(kl)}{\sqrt{vn}}. \quad (2.10)$$

Luego por el teorema de Donsker y el teorema del mapeo continuo se tiene que

$$\sup_{\substack{\pi_1, \dots, \pi_l \in \Pi_k \\ \pi_1 < \dots < \pi_l}} \sum_{r=1}^l \Delta_\pi(\hat{B}) - \frac{\mathbf{O}(kl)}{\sqrt{vn}} \xrightarrow{D} \sup_{\substack{\pi_1, \dots, \pi_l \in \Pi_k \\ \pi_1 < \dots < \pi_l \leq k}} \sum_{i=1}^l \Delta_{\pi_i}(B), \quad (2.11)$$

donde $B = (B_1, \dots, B_k)$ es un movimiento browniano k -dimensional. La demostración del teorema se sigue de (2.5), (2.10) y (2.11). ■

El teorema 2.1.1 nos da la relación entre los funcionales del movimiento browniano y el espectro del GUE. Gráficamente la idea de la demostración del teorema 2.1.1 es la siguiente:

$$\begin{array}{c}
 \omega_{n,k} = \begin{pmatrix} \frac{\omega_{11}-e}{\sqrt{vn}} & \dots & \frac{\omega_{1k}-e}{\sqrt{vn}} \\ \vdots & \ddots & \vdots \\ \frac{\omega_{n1}-e}{\sqrt{vn}} & \dots & \frac{\omega_{nk}-e}{\sqrt{vn}} \end{pmatrix} \xrightarrow[\sim]{\text{RSK}} \begin{array}{|c|c|c|c|c|} \hline & & & & \\ \hline & & & & \\ \hline & & & & \\ \hline & & & & \\ \hline & & & & \\ \hline & & & & \\ \hline \end{array} = \begin{pmatrix} \frac{\lambda_1^k - en}{\sqrt{vn}} \\ \vdots \\ \frac{\lambda_k^k - en}{\sqrt{vn}} \end{pmatrix} \xrightarrow{\text{Johansson}} \begin{pmatrix} \mu_1^k \\ \vdots \\ \mu_k^k \end{pmatrix} \\
 \text{Donsker} \quad \downarrow \quad \dots \quad \downarrow \\
 (B_1, \dots, B_k)
 \end{array}$$

Una consecuencia directa del teorema 2.1.1 es el teorema de Baryshnikov (teorema 1.3.3), el cual enunciamos a continuación

Corolario 2.1.1 *Bajo las mismas hipótesis del teorema 2.1.1 se tiene que*

$$\mu_1^k \stackrel{d}{=} \sup_{0=t_0 \leq \dots \leq t_k=1} \sum_{i=1}^k (B_i(t_i) - B_i(t_{i-1})).$$

Siguiendo con el enfoque probabilista de Benaych-Georges, Houdré y Xu [4, 11], en la siguiente sección identificaremos a la distribución asintótica de la forma del diagrama de Young asociado a una palabra aleatoria, en términos de funcionales del movimiento browniano. La importancia del teorema 2.1.1 es que nos permitirá identificar a la forma límite del diagrama de Young, como el espectro de una matriz aleatoria. Para ello nos será de gran utilidad el siguiente lema.

Lema 2.1.1 *Sea $B(t) = (B_1(t), \dots, B_m(t))_{t \in [0,1]}$ un movimiento browniano estándar m -dimensional y $H_m = (h_{ij})$ una matriz GUE de dimensión m . Para $k = 1, \dots, m$ denotemos por*

$$\mu_1^k \geq \dots \geq \mu_k^k,$$

al espectro del menor principal $k \times k$ de H_m . Se tiene que

$$(\mu_1^j)_{1 \leq j \leq k} \cup (h_{jj})_{1 \leq j \leq k} \stackrel{D}{=} \left(\max_{0=t_0 \leq \dots \leq t_j=1} \sum_{i=1}^j (B_i(t_i) - B_i(t_{i-1})) \right)_{1 \leq j \leq k} \cup (B_j(1))_{1 \leq j \leq k}. \quad (2.12)$$

Demostración:

Observemos que las entradas izquierda y derecha correspondientes a los índices $(l, k) = (1, k)$ en la igualdad (2.4), corresponden a

$$\mu_1^k \quad \text{y} \quad \sup_{\substack{\pi \in \mathcal{P} \\ \pi \leq k}} \Delta_\pi(B) = \sup_{\substack{\pi \in \mathcal{P} \\ \pi \leq k}} \sum_{i=1}^k (B_i(t_i) - B_i(t_{i-1})).$$

Por otro lado, las expresiones correspondientes a la diferencia de entradas $(j, j) - (j - 1, j - 1)$ en (2.4) son, del lado izquierdo

$$h_{jj} = \sum_{i=1}^j \mu_i^j - \sum_{i=1}^{j-1} \mu_i^{j-1}$$

y del lado derecho

$$\begin{aligned} \sup_{\substack{\pi_1, \dots, \pi_j \in \mathcal{P} \\ \pi_1 < \dots < \pi_j \leq j}} \sum_{i=1}^j \Delta_{\pi_i}(B) - \sup_{\substack{\pi_1, \dots, \pi_{j-1} \in \mathcal{P} \\ \pi_1 < \dots < \pi_{j-1} \leq j-1}} \sum_{i=1}^{j-1} \Delta_{\pi_i}(B) &= \sum_{i=1}^j (B_i(1) - B_i(0)) - \sum_{i=1}^{j-1} (B_i(1) - B_i(0)) \\ &= B_j(1). \end{aligned}$$

Luego por el teorema del mapeo continuo se sigue el resultado. ■

2.2. Palabras aleatorias

Con las mismas ideas usadas en la demostración del teorema 2.1.1, podemos abordar el problema de encontrar la distribución asintótica de la longitud de la sucesión no decreciente más grande en palabras aleatorias. En esta sección identificaremos a la distribución asintótica antes mencionada, en términos de funcionales del movimiento browniano y en términos del espectro de una matriz aleatoria, que especificaremos más adelante. La identificación se hará en el caso de palabras aleatorias no homogéneas y como caso particular se obtendrá el límite cuando las letras tienen la misma probabilidad de aparecer.

2.2.1. Generalidades

Sea $(X_n)_{n \geq 1}$ una sucesión de variables aleatorias i.i.d que toman valores en un alfabeto con letras $\{1, \dots, m\}$. Supongamos que las variables aleatorias $(X_n)_{n \geq 1}$ cumplen la siguiente condición: para $j = 1, \dots, m$ se tiene que

$$\mathbb{P}(X_1 = j) = p_j > 0, \tag{2.13}$$

y entonces

$$\sum_{j=1}^m p_j = 1.$$

Sea τ una permutación de $\{1, \dots, m\}$ tal que

$$p_{\tau(1)} \geq \dots \geq p_{\tau(m)}. \tag{2.14}$$

Supongamos que hay K probabilidades distintas $p^{(1)}, \dots, p^{(K)}$ en el conjunto $\{p_1, \dots, p_m\}$, con multiplicidades $d_1 \dots, d_K$, respectivamente. Es decir

$$\sum_{i=1}^K p^{(i)} d_i = 1. \tag{2.15}$$

Sin pérdida de generalidad supondremos que

$$p^{(1)} > \dots > p^{(K)}. \quad (2.16)$$

Sea $m_1 = 0$ y para $k = 2, \dots, K$ denotemos

$$m_k = \sum_{i=1}^{k-1} d_i.$$

Es decir, m_k es el número de probabilidades estrictamente mayores que $p^{(k)}$. En resumen

$$\underbrace{p_{\tau(m_1+1)} = \dots = p_{\tau(m_2)}}_{\substack{d_1 \text{ probabilidades} \\ \text{con valor común } p^{(1)}}} > \underbrace{p_{\tau(m_2+1)} = \dots = p_{\tau(m_3)}}_{\substack{d_2 \text{ probabilidades} \\ \text{con valor común } p^{(2)}}} > \dots > \underbrace{p_{\tau(m_K+1)} = \dots = p_{\tau(m)}}_{\substack{d_K \text{ probabilidades} \\ \text{con valor común } p^{(K)}}} \quad (2.17)$$

A lo largo de este capítulo consideraremos a una palabra aleatoria

$$W_n = X_1 \cdots X_n, \quad (2.18)$$

donde las letras X_1, \dots, X_n de W_n cumplen con las condiciones descritas anteriormente.

La correspondencia RSK genera a partir de W_n a una matriz con entradas enteras no negativas $X_W \in M_{n,m}$, y a un diagrama de Young de forma $\lambda^n = (\lambda_1^n, \dots, \lambda_m^n) \in \mathcal{P}_m$. De modo que una pregunta natural sería ¿qué propiedades probabilistas tienen la matriz X_W y la partición λ^n ? Las propiedades probabilistas de X_W se describen en el siguiente lema.

Lema 2.2.1 *Sea $W_n = X_1 \cdots X_n$ una palabra aleatoria como en (2.18). Denotemos por X_W a la matriz que genera el algoritmo RSK al aplicarlo a W_n . Entonces los renglones de X_W tiene la siguiente matriz de covarianza.*

$$\Sigma_W = \begin{pmatrix} p_1(1-p_1) & -p_1p_2 & \cdots & -p_1p_m \\ -p_2p_1 & p_2(1-p_2) & \cdots & -p_2p_m \\ \vdots & \vdots & \ddots & \vdots \\ -p_m p_1 & -p_m p_2 & \cdots & p_m(1-p_m) \end{pmatrix}. \quad (2.19)$$

Demostración:

Sea $(\mathbf{e}_i)_{i=1, \dots, m}$ la base canónica de \mathbb{R}^m , y sea $\mathbf{V} = (V_1, \dots, V_m)$ un vector aleatorio tal que

$$\mathbb{P}(\mathbf{V} = \mathbf{e}_i) = p_i, \quad j = 1, \dots, m.$$

Observemos que para cada $1 \leq i \leq m$, se tiene que

$$\mathbb{E}(V_i) = p_i, \quad \text{Var}(V_i) = p_i(1-p_i),$$

y que para $i \neq j$, $\text{Cov}(V_i, V_j) = -p_i p_j$. Entonces $\text{Cov}(\mathbf{V}) = \Sigma_W$.

Tomemos $\mathbf{V}_1, \dots, \mathbf{V}_n$ copias independientes de \mathbf{V} , donde para $1 \leq i \leq m$ se

tiene que $\mathbf{V}_i = (V_{i1}, \dots, V_{im})$. Denotemos por \widetilde{X}_W a la matriz de $n \times m$ cuyos renglones son los vectores $\mathbf{V}_1, \dots, \mathbf{V}_n$. Observemos que \widetilde{X}_W es igual en ley a X_W . ■

Una descripción de la distribución de λ^n en términos de polinomios ortogonales discretos se puede ver en [14].

En la sección 2.3 de esta tesis daremos una descripción de la distribución asintótica de λ^n en términos de funcionales del movimiento browniano, y en términos del espectro de una matriz aleatoria. Pero antes, veremos como podemos codificar la estructura probabilista de W_n en términos de una matriz aleatoria.

Sean H_{d_1}, \dots, H_{d_K} matrices aleatorias independientes GUE de dimensiones d_1, \dots, d_K respectivamente. Denotemos por

$$H := \begin{pmatrix} H_{d_1} & & \\ & \ddots & \\ & & H_{d_K} \end{pmatrix} \quad \text{y} \quad J := \begin{pmatrix} \sqrt{p^{(1)}} I_{d_1} & & \\ & \ddots & \\ & & \sqrt{p^{(K)}} I_{d_K} \end{pmatrix}. \quad (2.20)$$

Donde I_d es la matriz identidad de dimensión d . Luego consideremos a la siguiente matriz

$$\widehat{H} := H - \text{tr}(HJ)J = \begin{pmatrix} \widehat{H}_{d_1} & & \\ & \ddots & \\ & & \widehat{H}_{d_K} \end{pmatrix}, \quad (2.21)$$

donde para $i = 1, \dots, K$ se tiene que

$$\widehat{H}_{d_i} = H_{d_i} - \text{tr}(HJ)\sqrt{p^{(i)}} I_{d_i}. \quad (2.22)$$

El siguiente lema nos muestra que las matrices definidas en (2.22), codifican parte de la información probabilista de W_n .

Lema 2.2.2 *Construyamos \widehat{H} como en (2.21) a partir de la colección $H_{d_1} = (h_{ij}^{(1)})_{i,j=1}^{d_1}, \dots, H_{d_K} = (h_{ij}^{(K)})_{i,j=1}^{d_K}$ de matrices GUE y denotemos para $l = 1, \dots, K$ a la matriz $\widehat{H}_{d_l} = (\widehat{h}_{i,j}^{(l)})_{i,j=1}^{d_l}$ como en (2.22). Entonces el vector*

$$(\widehat{h}_{11}, \dots, \widehat{h}_{d_l d_l}),$$

tiene matriz de covarianza

$$\begin{pmatrix} 1 - p^{(l)} & -p^{(l)} & \dots & -p^{(l)} \\ -p^{(l)} & 1 - p^{(l)} & \dots & -p^{(l)} \\ \vdots & \ddots & \dots & \vdots \\ -p^{(l)} & \dots & & 1 - p^{(l)} \end{pmatrix}. \quad (2.23)$$

Demostración:

Observemos que

$$\begin{aligned}\widehat{h}_{ii}^{(l)} &= h_{ii}^{(l)} - \sqrt{p^{(i)}} \operatorname{tr}(HJ) \\ &= h_{ii}^{(l)} - \sqrt{p^{(i)}} \sum_{r=1}^K \sqrt{p^{(r)}} \operatorname{tr}(H_{d_r}).\end{aligned}$$

La matriz (2.23) se obtiene de un cálculo directo. ■

Observemos que si multiplicamos (2.23) por $p^{(l)}$ obtenemos la submatriz de (2.19) correspondiente a las columnas con probabilidad $p^{(l)}$. De hecho la información contenida en las K matrices definidas en (2.23) es la misma que la información contenida en (2.19). En la sección 2.3 veremos que la distribución límite de λ^n corresponde al espectro de la matriz \widehat{H} definida en (2.21).

Siguiendo las ideas de la demostración del teorema 2.1.1, identificaremos a la distribución límite de λ^n como un funcional del movimiento browniano. Para ello daremos a la siguiente definición.

Definición 2.2.1 *Sea $B = (B_1, \dots, B_m)$ un movimiento browniano m -dimensional. Para cada $l = 1, \dots, m$ denotemos por $k = k(l)$ al único entero tal que $1 \leq k \leq K$ y $p_{\tau(l)} = p^{(k)}$. Definamos a los siguientes funcionales de B*

$$L_m^l(B) := \sum_{i=1}^{m_k} B_i(1) + \sup_{\substack{\pi_1, \dots, \pi_{l-m_k} \in \Pi_{d_k} \\ \pi_1 < \dots < \pi_{l-m_k} \leq d_k}} \sum_{i=1}^{l-m_k} \Delta_{\pi_i}(B_{m_k+1}, \dots, B_m), \quad (2.24)$$

donde Π_m y Δ_π se definen como en (1.13) y (2.3).

Observemos que para el caso particular $l = 1$, se tiene que $m_1 = 0$ y entonces L_m^1 se reduce a

$$\begin{aligned}L_m^1(B) &= \sup_{\substack{\pi_1 \in \Pi_{d_1} \\ \pi_1 \leq d_1}} \Delta_\pi(B_1, \dots, B_{d_1}) \\ &= \sup_{0=t_0 \leq \dots \leq t_{d_1}=1} \sum_{i=1}^{d_1} (B_i(t_i) - B_i(t_{i-1})).\end{aligned} \quad (2.25)$$

Observemos que la expresión (2.25) ya ha aparecido en múltiples ocasiones a lo largo de la tesis, por ejemplo si $B = (B_1, \dots, B_m)$ es un movimiento browniano estándar el funcional (2.25) aparece en el teorema 1.3.2 como límite de tiempos de espera en teoría de colas, y en los teoremas 1.3.3 y 2.1.1 aparece una igualdad en distribución entre (2.25) y al valor propio más grande de una matriz GUE. Por esta razón decimos que los funcionales $L_m^l(B)$ son una generalización del funcional de Glynn y Whitt (1.19).

2.2.2. El problema de Ulam para palabras aleatorias

Recordemos que en el problema de Ulam para palabras aleatorias, se trata de encontrar la distribución asintótica de la longitud de la subsucesión no decreciente más grande en una palabra aleatoria.

A continuación veremos que la expresión (2.25) cuando $B = (B_1, \dots, B_m)$ es un movimiento browniano no estándar, se puede ver como la distribución asintótica de la longitud de la subsucesión no decreciente más grande en palabras aleatorias.

Teorema 2.2.1 *Sea $W_n = X_1 \cdots X_n$ una palabra aleatoria como (2.18), y sea*

$$\lambda^n = (\lambda_1^n, \dots, \lambda_m^n),$$

la forma del diagrama de Young que se obtiene al aplicar el algoritmo RSK a W_n . Entonces cuando $n \rightarrow \infty$,

$$\frac{\lambda_1^n - np_{\tau(1)}}{\sqrt{n}} \Rightarrow L_m^1(\tilde{B}) = \sup_{t_0=0 \leq \dots \leq t_{d_1}=1} \sum_{i=1}^{d_1} (\tilde{B}_i(t_i) - \tilde{B}_i(t_{i-1})), \quad (2.26)$$

donde $\tilde{B} = (\tilde{B}_1, \dots, \tilde{B}_{d_1})$ es un movimiento browniano con matriz de covarianza

$$\begin{pmatrix} p_{\tau(1)}(1 - p_{\tau(1)}) & -p_{\tau(1)}^2 & \cdots & -p_{\tau(1)}^2 \\ -p_{\tau(1)}^2 & p_{\tau(1)}(1 - p_{\tau(1)}) & \cdots & -p_{\tau(1)}^2 \\ \vdots & \vdots & \ddots & \vdots \\ -p_{\tau(1)}^2 & -p_{\tau(1)}^2 & \cdots & p_{\tau(1)}(1 - p_{\tau(1)}) \end{pmatrix}, \quad (2.27)$$

donde τ es una permutación como en (2.14).

Demostración:

Sea X_W la matriz que genera el algoritmo RSK cuando lo aplicamos a W_n . Recordemos que los renglones de esta matriz son los vectores de la base canónica en \mathbb{R}^m . Más específicamente, si en la palabra aleatoria $W_n = X_1 \cdots X_n$ se cumple que $X_k = r$, para $1 \leq k \leq n$ y $1 \leq r \leq m$. Entonces el k -ésimo renglón de X_W es el vector $e_r = (0, \dots, 0, 1, 0, \dots, 0)$ de la base canónica en \mathbb{R}^m . Como las letras de W_n son independientes, entonces los renglones de X_W son independientes. Por otra parte, del lema 2.2.1 se tiene que las variables en cada renglón de X_W tienen matriz de covarianza como en (2.19).

Denotemos por $\lambda^n = (\lambda_1^n, \dots, \lambda_m^n)$ a la forma del diagrama de Young que se obtiene al aplicar el algoritmo RSK a X_W . Si en el lema 1.1.3 tomamos $k = 1$ obtenemos que

$$\lambda_1^n = \max_{\hat{\pi} \in \hat{\Pi}_{n,m}} |\hat{\pi}_m|, \quad (2.28)$$

donde $|\cdot|$ y $\hat{\Pi}_m$ se definen como en (1.12) y (1.13), respectivamente. Ahora observemos que

$$\begin{aligned} \lambda_1^n &= \max_{\hat{\pi} \in \hat{\Pi}_{n,m}} \sum_{(i,j) \in \hat{\pi}} (X_W)_{ij} \\ &= \sup_{0=t_0 \leq \dots \leq t_m=1} \sum_{i=1}^m \sum_{j=[t_{i-1}n]}^{\lfloor t_i n \rfloor} (X_W)_{ij}. \end{aligned}$$

Sea

$$J_1 = \{i | p_i = p^{(1)}\},$$

es decir J_1 son los índices de las d_1 -letras con mayor probabilidad. Denotemos por

$$\widehat{\lambda}_1^n = \sup_{\substack{0=t_0 \leq \dots \leq t_m=1 \\ t_{i-1}=t_i, \quad i \notin J_1}} \sum_{i=1}^m \sum_{j=\lfloor t_{i-1}n \rfloor}^{\lfloor t_i n \rfloor} (X_W)_{ij}, \quad (2.29)$$

es decir, $\widehat{\lambda}_1$ corresponde a tomar el camino $\widehat{\pi}$ en X_W con el mayor peso, de tal manera que las entradas en X_W que están sobre $\widehat{\pi}$, son positivas solo en las columnas que corresponden a letras con probabilidad $p^{(1)}$.

Supongamos que

$$J_1 = \{\alpha_1, \dots, \alpha_{d_1}\} \quad \text{con} \quad \alpha_1 < \dots < \alpha_{d_1}.$$

Definamos para $i = 1, \dots, d_1$ y $0 \leq t \leq 1$ a los siguientes procesos

$$\widetilde{B}_i^n(t) = \sum_{j=1}^{\lfloor nt \rfloor} \left(\frac{(X_W)_{j\alpha_i} - p_{\alpha_i}}{\sqrt{n}} \right).$$

Restando $np_{\tau(1)}$ a (2.29) y dividiendo por \sqrt{n} obtenemos

$$\begin{aligned} \frac{\widehat{\lambda}_1^n - np_{\tau(1)}}{\sqrt{n}} &= \sup_{\substack{0=t_0 \leq \dots \leq t_m=1 \\ t_{i-1}=t_i, \quad i \notin J_1}} \sum_{i=1}^m \sum_{j=\lfloor t_{i-1}n \rfloor}^{\lfloor t_i n \rfloor} \frac{(X_W)_{ij}}{\sqrt{n}} - \frac{np_{\tau(1)}}{\sqrt{n}} \\ &= \sup_{\substack{0=t_0 \leq \dots \leq t_m=1 \\ t_{i-1}=t_i, \quad i \notin J_1}} \sum_{i=1}^m \sum_{j=\lfloor t_{i-1}n \rfloor}^{\lfloor t_i n \rfloor} \frac{(X_W)_{ij} - p_{\tau(1)}}{\sqrt{n}} + \frac{\mathbf{O}(m)}{\sqrt{n}} \\ &= \sup_{\substack{0=t_0 \leq \dots \leq t_m=1 \\ t_{i-1}=t_i, \quad i \notin J_1}} \sum_{i=1}^m (\widetilde{B}_i^n(t_i) - \widetilde{B}_i^n(t_{i-1})) + \frac{\mathbf{O}(m)}{\sqrt{n}} \\ &= \sup_{0=t_0 \leq \dots \leq t_m=1} \sum_{i=1}^{d_1} (\widetilde{B}_{\alpha_i}^n(t_i) - \widetilde{B}_{\alpha_i}^n(t_{i-1})) + \frac{\mathbf{O}(m)}{\sqrt{n}} \end{aligned}$$

Luego, por el teorema de Donsker y el teorema del mapeo continuo se tiene que cuando $n \rightarrow \infty$

$$\frac{\widehat{\lambda}_1^n - np_{\tau(1)}}{\sqrt{n}} \Rightarrow \sup_{t_0=0 \leq \dots \leq t_{d_1}=1} \sum_{i=1}^{d_1} (\widetilde{B}_i(t_i) - \widetilde{B}_i(t_{i-1})), \quad (2.30)$$

donde $\widetilde{B} = (\widetilde{B}_1, \dots, \widetilde{B}_m)$ es un movimiento browniano, y por el lema 2.2.1 \widetilde{B} tiene matriz de covarianza (2.34).

Se puede demostrar¹ que cuando $n \rightarrow \infty$

$$\frac{\lambda_1^n}{\sqrt{n}} - \frac{\widehat{\lambda}_1^n}{\sqrt{n}} \xrightarrow{\mathbb{P}} 0. \quad (2.31)$$

¹Ver [9].

En particular

$$\frac{\lambda_1^n - np_{\tau(1)}}{\sqrt{n}} - \frac{\widehat{\lambda}_1^n - np_{\tau(1)}}{\sqrt{n}} \xrightarrow{\mathbb{P}} 0. \quad (2.32)$$

Por lo que sumando (2.30) y (2.32) y usando el teorema de Slutsky se obtiene el resultado. ■

Ejemplo:

La idea principal de la demostración del teorema 2.2.1 es reducir el problema a uno ya conocido. Por ejemplo, si tenemos un alfabeto con nueve letras $1, 2, \dots, 9$, con probabilidades p_1, \dots, p_9 , respectivamente, y las probabilidades cumplen que

$$p_1 = p_3 = p_7 > p_4 = p_8 > p_2 = p_5 = p_9 > p_6,$$

entonces a la palabra

$$W_{14} = 53873132636879,$$

le corresponde la siguiente matriz de enteros

$$\omega_{14,9} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Sea $\lambda^{14} = (\lambda_1^{14}, \dots, \lambda_9^{14})$ la forma del diagrama de Young asociado a W_{14} . Sabemos que λ_1^{14} es la longitud de la subsucesión no decreciente más grande, la cual es igual al máximo de los pesos de los caminos $\downarrow \rightarrow$ sobre $\omega_{14,9}$ que van de la entrada $(1, 1)$ a la entrada $(14, 9)$. Lo que se dice en (2.31) es que asintóticamente, tomar el máximo sobre los caminos $\downarrow \rightarrow$ es equivalente a tomar el máximo sobre un arreglo más pequeño, como el que se muestra a continuación.

$$\begin{array}{cccccccccc}
p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 & p_8 & p_9 & & p_1 & p_3 & p_7 \\
\left(\begin{array}{cccccccccc}
0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\
\downarrow & & & & & & & & & & & & \\
0 & \rightarrow & 0 & \rightarrow & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & & & \\
& & & & \downarrow & & & & & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 \\
& & & & \downarrow & & & & & & & & \\
0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 \\
& & & & \downarrow & & & & & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
& & & & \downarrow & & & & & & & & \\
\mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
& & & & \downarrow & & & & & & & & \\
0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 \\
& & & & \downarrow & & & & & & & & \\
0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
& & & & \downarrow & & & & & & & & \\
0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 \\
& & & & & \downarrow & & & & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & 0 \\
& & & & & & & \downarrow & & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & \rightarrow & \mathbf{1} & 0 & 0 \\
& & & & & & & \downarrow & & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \rightarrow & 0 & \rightarrow & \mathbf{1} \\
& & & & & & & & & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rightarrow & 0 & \rightarrow & \mathbf{1}
\end{array} \right) \rightsquigarrow \left(\begin{array}{ccc}
0 & 0 & 0 \\
\downarrow & & \\
0 & \rightarrow & \mathbf{1} & 0 \\
& & \downarrow & \\
0 & & 0 & 0 \\
& & \downarrow & \\
0 & & \mathbf{1} & 0 \\
& & \downarrow & \\
0 & & 0 & \mathbf{1} \\
& & \downarrow & \\
\mathbf{1} & & 0 & 0 \\
& & \downarrow & \\
0 & & \mathbf{1} & 0 \\
& & \downarrow & \\
0 & & 0 & 0 \\
& & \downarrow & \\
0 & & 0 & 0 \\
& & \downarrow & \\
0 & & \mathbf{1} & 0 \\
& & \downarrow & \\
0 & & 0 & 0 \\
& & \downarrow & \\
0 & & 0 & 0 \\
& & \downarrow & \\
0 & & 0 & \rightarrow & \mathbf{1} \\
& & & \downarrow & \\
& & & & \downarrow & \\
& & & & & \rightarrow & \mathbf{1} \\
& & & & & & \downarrow
\end{array} \right)
\end{array}$$

Es decir, basta tomar el máximo en el arreglo que se forma con las columnas de $\omega_{14,9}$ correspondientes a las probabilidades más grandes p_1 , p_2 y p_3 . De esta forma, podemos usar las mismas ideas que en la demostración del teorema 2.1.1, pues el nuevo arreglo corresponde a una palabra aleatoria cuyas letras tienen distribución uniforme en el alfabeto.

Como un caso particular del teorema 2.2.1 se obtiene la distribución límite de la longitud de la subsucesión no decreciente más grande de una palabra aleatoria, en el caso en que las letras tienen distribución uniforme en el alfabeto.

Corolario 2.2.1 *Sea $W_n = X_1 \cdots X_n$ una palabra aleatoria con letras i.i.d. tales que para toda $r = 1, \dots, m$*

$$\mathbb{P}(X_j = r) = \frac{1}{m}.$$

Supongamos que $\lambda^n = (\lambda_1^n, \dots, \lambda_m^n)$ es la forma del diagrama de Young que se obtiene al aplicar el algoritmo RSK a W_n . Entonces cuando $n \rightarrow \infty$,

$$\frac{\lambda_1^n - \frac{n}{m}}{\sqrt{n}} \Rightarrow \sup_{t_0=0 \leq \dots \leq t_{d_1}=1} \sum_{i=1}^{d_1} (\tilde{B}_i(t_i) - \tilde{B}_i(t_{i-1})), \quad (2.33)$$

donde $\tilde{B} = (\tilde{B}_1, \dots, \tilde{B}_{d_1})$ es un movimiento browniano con matriz de covarianza

$$\begin{pmatrix} \frac{m-1}{m^2} & -\frac{1}{m^2} & \cdots & -\frac{1}{m^2} \\ -\frac{1}{m^2} & \frac{m-1}{m^2} & \cdots & -\frac{1}{m^2} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{m^2} & -\frac{1}{m^2} & \cdots & \frac{m-1}{m^2} \end{pmatrix}. \quad (2.34)$$

Demostración:

Basta con tomar $d_1 = m$ en el teorema 2.2.1. ■

En 2001 Its, Tracy y Widom [12] demostraron que

$$\frac{\lambda_1^n - \frac{n}{m}}{\sqrt{\frac{n}{m}}} \Rightarrow \mu_1, \quad (2.35)$$

donde μ_1 es el valor propio más grande de una matriz GUE sin traza de dimensión m . El cambio de variable $B = \sqrt{m}\tilde{B}$ en el corolario 2.2.1 nos da la normalización usada por Tracy y Widom en (2.35).

Corolario 2.2.2 *Sea $W_n = X_1 \cdots X_n$ una palabra aleatoria con letras i.i.d. tales que para toda $r = 1, \dots, m$*

$$\mathbb{P}(X_j = r) = \frac{1}{m}.$$

Supongamos que $\lambda^n = (\lambda_1^n, \dots, \lambda_m^n)$ es la forma del diagrama de Young que se obtiene al aplicar el algoritmo RSK a W_n . Entonces cuando $n \rightarrow \infty$,

$$\frac{\lambda_1^n - \frac{n}{m}}{\sqrt{\frac{n}{m}}} \Rightarrow \sup_{t_0=0 \leq \dots \leq t_{d_1}=1} \sum_{i=1}^{d_1} (B_i(t_i) - B_i(t_{i-1})), \quad (2.36)$$

donde $\tilde{B} = (\tilde{B}_1, \dots, \tilde{B}_{d_1})$ es un movimiento browniano con matriz de covarianza

$$\begin{pmatrix} \frac{m-1}{m} & -\frac{1}{m} & \dots & -\frac{1}{m} \\ -\frac{1}{m} & \frac{m-1}{m} & \dots & -\frac{1}{m} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{m} & -\frac{1}{m} & \dots & \frac{m-1}{m} \end{pmatrix}. \quad (2.37)$$

Demostración:

Basta con multiplicar por \sqrt{m} en ambos lados de (2.33). ■

Por la unicidad del límite, de (2.35) y (2.36) obtenemos que

$$\mu_1 \stackrel{d}{=} \sup_{t_0=0 \leq \dots \leq t_{d_1}=1} \sum_{i=1}^{d_1} (B_i(t_i) - B_i(t_{i-1})), \quad (2.38)$$

donde (B_1, \dots, B_m) es un movimiento browniano con matriz de covarianza (2.37), y μ_1 es el valor propio más grande de una matriz GUE sin traza de dimensión m . Nosotros utilizaremos el teorema 2.1.1 para obtener (2.38) a partir de (2.26).

Observemos que el funcional $L_m^1(\tilde{B})$ que aparece en (2.26), tiene la misma estructura que el funcional que aparece en el corolario 2.1.1. Sin embargo, el movimiento browniano que aparece en el corolario 2.1.1 es estándar, y el movimiento browniano que aparece en (2.26) tiene una matriz de covarianza distinta a la identidad. Un cambio de variable nos permite expresar al funcional $L_m^1(\tilde{B})$ que aparece en (2.26) en términos de un movimiento browniano estándar.

Lema 2.2.3 Sea $B = (B_1, \dots, B_m)$ un movimiento browniano estándar, y entonces $\text{Cov}(B) = I_m$. Definamos al proceso $\tilde{B} = (\tilde{B}_1, \dots, \tilde{B}_m)$ por

$$\tilde{B}_i = \sqrt{p_{\tau(i)}}B_i - p_{\tau(i)} \sum_{j=1}^m \sqrt{p_{\tau(j)}}B_j. \quad (2.39)$$

Entonces \tilde{B} es un movimiento browniano m dimensional con matriz de covarianza Σ definida por

$$\Sigma = \begin{pmatrix} p_{\tau(1)}(1 - p_{\tau(1)}) & -p_{\tau(1)}p_{\tau(2)} & \cdots & -p_{\tau(1)}p_{\tau(m)} \\ -p_{\tau(2)}p_{\tau(1)} & p_{\tau(2)}(1 - p_{\tau(2)}) & \cdots & -p_{\tau(2)}p_{\tau(m)} \\ \vdots & \vdots & \ddots & \vdots \\ -p_{\tau(m)}p_{\tau(1)} & -p_{\tau(m)}p_{\tau(2)} & \cdots & p_{\tau(m)}(1 - p_{\tau(m)}) \end{pmatrix}. \quad (2.40)$$

Demostración:

Observemos que \tilde{B} es un movimiento browniano, pues se obtiene mediante una transformación lineal de B . La matriz Σ se obtiene mediante un cálculo directo. ■

Si aplicamos el cambio de variable (2.39) en al funcional $L_m^1(\tilde{B})$ que aparece en (2.26) obtenemos que

$$L_m^1(\tilde{B}) = -p_{\tau(1)} \sum_{q=1}^m \sqrt{p_{\tau(q)}}B_q(1) + \sqrt{p_{\tau(1)}} \sup_{0=t_0 \leq \dots \leq t_{d_1}=1} \sum_{i=1}^{d_1} (B_i(t_i) - B_i(t_{i-1})), \quad (2.41)$$

donde $B = (B_1, \dots, B_m)$ es un movimiento browniano estándar.

Corolario 2.2.3 Consideremos a una palabra aleatoria W_n definida como en (2.18), y sea $\lambda^n = (\lambda_1^n, \dots, \lambda_m^n)$ la forma del diagrama de Young que genera el algoritmo RSK. Sea \widehat{H}_{d_1} como en (2.22) y μ_1 su valor propio más grande. Entonces se tiene que

$$\frac{\lambda_1^n - np_{\tau(1)}}{\sqrt{np_{\tau(1)}}} \Rightarrow \mu_1. \quad (2.42)$$

Demostración:

Por el teorema 2.2.1 y por el teorema del mapeo continuo será suficiente demostrar que

$$\frac{L_m^1(\tilde{B})}{\sqrt{p_{\tau(1)}}} \stackrel{d}{=} \mu_1 \quad (2.43)$$

Usando (2.41) y el lema 2.1.1 se sigue que el lado izquierdo de (2.56) se puede escribir como

$$\begin{aligned} \frac{L_m^1(\tilde{B})}{\sqrt{p_{\tau(1)}}} &= -\sqrt{p_{\tau(1)}} \sum_{q=1}^m \sqrt{p_{\tau(q)}}B_q(1) + \sup_{0=t_0 \leq \dots \leq t_{d_1}=1} \sum_{i=1}^{d_1} (B_i(t_i) - B_i(t_{i-1})) \\ &= -\sqrt{p_{\tau(1)}} \sum_{q=1}^m \sqrt{p_{\tau(q)}}h_{qq} + \lambda_1(H_{d_1}), \end{aligned} \quad (2.44)$$

donde $H = (h_{ij})_{i,j=1}^m$ es como en (2.20), y $\lambda_1(H_{d_1})$ es el valor propio más grande de la matriz H_{d_1} definida como en (2.22).

Del lado derecho observemos que si λ es un valor propio de H_{d_1} asociado al vector propio v , entonces

$$\lambda - \text{tr}(HJ)\sqrt{p^{(1)}},$$

es un valor propio de $\widetilde{H}_{d_1} = H_{d_1} - \text{tr}(HJ)\sqrt{p^{(1)}}I_{d_k}$ que tiene a v con su vector propio asociado. En particular se tiene que

$$\begin{aligned} \mu_1 &= -\text{tr}(HJ)\sqrt{p^{(1)}} + \lambda_{\max}(H_{d_1}), \\ &= -\sqrt{p^{(1)}} \sum_{q=1}^m \sqrt{p_{\tau(q)}} h_{qq} + \lambda_{\max}(H_{d_1}). \end{aligned} \quad (2.45)$$

Luego de (2.44) y (2.57) se sigue el corolario. ■

Los corolarios 2.2.5 y 2.2.2 nos dan una demostración de (2.38).

Corolario 2.2.4 *Sea H_m una matriz GUE sin traza de dimensión m , y $B = (B_1, \dots, B_m)$ un movimiento browniano con matriz de covarianza*

$$\begin{pmatrix} \frac{m-1}{m} & -\frac{1}{m} & \dots & -\frac{1}{m} \\ -\frac{1}{m} & \frac{m-1}{m} & \dots & -\frac{1}{m} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{m} & -\frac{1}{m} & \dots & \frac{m-1}{m} \end{pmatrix}. \quad (2.46)$$

Entonces

$$\mu_1 \stackrel{d}{=} \sup_{t_0=0 \leq \dots \leq t_{d_1}=1} \sum_{i=1}^{d_1} (B_i(t_i) - B_i(t_{i-1})), \quad (2.47)$$

donde μ_1 es el valor propio más grande de H_m .

Demostración:

Basta con tomar $d_1 = m$ en el corolario 2.2.5, y el resultado se sigue de la unicidad del límite en distribución en las expresiones (2.55) y (2.36). ■

2.2.3. Forma límite del diagrama de Young de una palabra aleatoria

Es posible extender el teorema 2.2.1 para encontrar la distribución asintótica de toda la forma del diagrama de Young. Esta extensión nos permitirá interpretar a los funcionales L_m^i en términos de caminos sobre X_W .

Teorema 2.2.2 *Sea $W_n = X_1 \dots X_n$ una palabra aleatoria definida como en (2.18), y $\lambda^n = (\lambda_1^n, \dots, \lambda_m^n)$ la forma del diagrama de Young que se obtiene al aplicar el algoritmo RSK a W_n . Entonces cuando $n \rightarrow \infty$,*

$$\left(\frac{\lambda_1^n - np_{\tau(1)}}{\sqrt{n}}, \dots, \frac{\lambda_m^n - np_{\tau(m)}}{\sqrt{n}} \right) \Rightarrow (L_m^1(\widetilde{B}), (L_m^2 - L_m^1)\widetilde{B}, \dots, (L_m^m - L_m^{m-1})\widetilde{B}),$$

donde L_m^1, \dots, L_m^m son los funcionales definidos en (2.24), y \widetilde{B} es un movimiento browniano con matriz de covarianza (2.40).

Demostración:

Denotemos por X_W a la matriz que genera el algoritmo RSK al aplicarlo a W_n . Recordemos que del lema 1.1.3 se tiene que para $1 \leq l \leq m$ tenemos que

$$\lambda_1^n + \cdots + \lambda_l^n = \max_{\substack{\hat{\pi}_1, \dots, \hat{\pi}_l \in \hat{\Pi}_{n,m} \\ \hat{\pi}_1 < \cdots < \hat{\pi}_l}} \sum_{r=1}^l |\hat{\pi}_r|, \quad (2.48)$$

donde $|\cdot|$ y $\hat{\Pi}_m$ se definen como en (1.12) y (1.13), respectivamente. Definamos para $i = 1, \dots, m$ y $0 \leq t \leq 1$ a $\tilde{B} = (\tilde{B}_m, \dots, \tilde{B}_1)$ por

$$\tilde{B}_i^n(t) = \sum_{j=1}^{\lfloor nt \rfloor} \left(\frac{(X_W)_{j, \tau(i)} - p_{\tau(i)}}{\sqrt{n}} \right),$$

por el lema 2.2.1 tenemos que

$$\text{Cov}(\tilde{B}) = \begin{pmatrix} p_{\tau(1)}(1 - p_{\tau(1)}) & -p_{\tau(1)}p_{\tau(2)} & \cdots & -p_{\tau(1)}p_{\tau(m)} \\ -p_{\tau(2)}p_{\tau(1)} & p_{\tau(2)}(1 - p_{\tau(2)}) & \cdots & -p_{\tau(2)}p_{\tau(m)} \\ \vdots & \vdots & \ddots & \vdots \\ -p_{\tau(m)}p_{\tau(1)} & -p_{\tau(m)}p_{\tau(2)} & \cdots & p_{\tau(m)}p_{\tau(m)} \end{pmatrix}. \quad (2.49)$$

Sumando $n(p_{\tau(1)} + \cdots + p_{\tau(l)})$ a (2.48) y dividiendo por \sqrt{n} , obtenemos

$$\begin{aligned} \frac{(\lambda_1^n + \cdots + \lambda_l^n) - n(p_{\tau(1)} + \cdots + p_{\tau(l)})}{\sqrt{n}} &= \max_{\substack{\hat{\pi}_1, \dots, \hat{\pi}_l \in \hat{\Pi}_{n,m} \\ \hat{\pi}_1 < \cdots < \hat{\pi}_l}} \sum_{r=1}^l \sum_{(i,j) \in \hat{\pi}_r} \left(\frac{(X_W)_{ij} - p_{\tau(r)}}{\sqrt{n}} \right) \\ &\quad + \frac{\mathbf{O}(lm)}{\sqrt{n}} \\ &= \sup_{\substack{\pi_1, \dots, \pi_l \in \Pi_m \\ \pi_1 < \cdots < \pi_l}} \sum_{r=1}^l \Delta_{\pi_r}(\tilde{B}) + \frac{\mathbf{O}(lm)}{\sqrt{n}}. \end{aligned} \quad (2.50)$$

Para $l \geq 1$ denotemos

$$J_l = \{i | p_i = p^{(l)}\}, \quad \text{y} \quad \bar{J}_l = J_1 \cup \cdots \cup J_{l-1}$$

recordemos que de la notación c) y d) se tiene que $|J_l| = d_l$ y $|\bar{J}_l| = d_1 + \cdots + d_{l-1} = m_l$. Sea

$$\tilde{\Pi}_l^n = \{\pi | \pi \text{ es un camino } \downarrow \rightarrow \text{ en la lattice } \{1, \dots, n\} \times J_l\}.$$

En la definición de $\tilde{\Pi}_l^n$ hay que tomar en cuenta que importa el orden en la lattice $\{1, \dots, n\} \times J_l$. Nosotros tomaremos el orden usual en los conjuntos.

Para cada $l = 1, \dots, m$ denotemos por $k = k(l)$ al único entero tal que $1 \leq k \leq K$ y $p_{\tau(l)} = p^{(k)}$, y sea

$$G_{l,n} = \sum_{j \in \bar{J}_k} \sum_{i=1}^n (X_W)_{ij} + \max_{\substack{\hat{\pi}_1, \dots, \hat{\pi}_{l-m_k} \in \tilde{\Pi}_k^n \\ \hat{\pi}_1 < \cdots < \hat{\pi}_{l-m_k}}} \sum_{r=1}^{l-m_k} \sum_{(i,j) \in \hat{\pi}_r} (X_W)_{ij}. \quad (2.51)$$

Aplicando en (2.51) la misma normalización que en (2.50) obtenemos

$$\begin{aligned} \frac{G_{l,n} - n(p_{\tau(1)} + \cdots + p_{\tau(l)})}{\sqrt{n}} &= \sum_{j \in \bar{J}_k} \sum_{i=1}^n \left(\frac{(X_W)_{ij} - p_{\tau(j)}}{\sqrt{n}} \right) \\ &\quad + \max_{\substack{\hat{\pi}_1, \dots, \hat{\pi}_{l-m_k} \in \tilde{\Pi}_k \\ \hat{\pi}_1 < \dots < \hat{\pi}_{l-m_k}}} \sum_{r=1}^{l-m_k} \sum_{(i,j) \in \hat{\pi}_r} \left(\frac{(X_W)_{ij} - p_{\tau(j)}}{\sqrt{n}} \right) + \frac{\mathbf{O}(lm)}{\sqrt{n}} \\ &= \sum_{j=1}^{m_k} \tilde{B}_j^n(1) + \sup_{\substack{\pi_1, \dots, \pi_{l-m_k} \in \tilde{\Pi}_k \\ \pi_1 < \dots < \pi_{l-m_k}}} \sum_{r=1}^{l-m_k} \Delta_{\pi_r}(\tilde{B}_{m_k+1}, \dots, \tilde{B}_m) + \frac{\mathbf{O}(lm)}{\sqrt{n}}, \end{aligned}$$

donde

$$\tilde{\Pi}_k = \{\pi: [0, 1] \rightarrow J_k \mid \pi \text{ es no decreciente y cádlág}\}.$$

Luego por el teorema de Donsker y el teorema del mapeo continuo se tiene que

$$\frac{G_{l,n} - n(p_{\tau(1)} + \cdots + p_{\tau(l)})}{\sqrt{n}} \xrightarrow{D} L_m^l(B). \quad (2.52)$$

Por el mismo argumento que en [9] se tiene que

$$\frac{\lambda_1^n + \cdots + \lambda_l^n}{\sqrt{n}} - \frac{G_{l,n}}{\sqrt{n}} \xrightarrow{\mathbb{P}} 0. \quad (2.53)$$

Del teorema de Slutsky y de (2.52) y (2.53) se obtiene el resultado. \blacksquare

Si aplicamos el cambio de variable (2.39) a los funcionales $L_m^1(\tilde{B}), \dots, L_m^m(\tilde{B})$ con \tilde{B} un movimiento browniano con matriz de covarianza (2.40) obtenemos las siguientes expresiones.

$$\begin{aligned} L_m^l &= \sum_{j=1}^{m_k} \sqrt{p_{\tau(j)}} B_j(1) - \sum_{j=1}^l p_{\tau(j)} \sum_{q=1}^m \sqrt{p_{\tau(q)}} B_q(1) \\ &\quad + \sqrt{p_{\tau(m_k+1)}} \sup_{\substack{\pi_1, \dots, \pi_{l-m_k} \in \tilde{\Pi}_{d_k}}} \sum_{j=1}^{l-m_k} \Delta_{\pi_j}(B_{m_k+1}, \dots, B_{m_k+d_k}), \end{aligned} \quad (2.54)$$

donde (B_1, \dots, B_m) es un movimiento browniano estándar m dimensional, y para cada $l = 1, \dots, m$ denotemos por $k = k(l)$ al único entero tal que $1 \leq k \leq K$ y $p_{\tau(l)} = p^{(k)}$.

Las expresiones (2.54) y el teorema 2.1.1 nos permiten identificar a

$$\left(L_m^1(\tilde{B}), (L_m^2 - L_m^1)\tilde{B}, \dots, (L_m^m - L_m^{m-1})\tilde{B} \right),$$

como el espectro de una matriz GUE sin traza generalizado.

Corolario 2.2.5 Consideremos a matrices \widehat{H} y $\widehat{H}_{d_1}, \dots, \widehat{H}_{d_K}$ como en (2.21) y (2.22). Denotaremos por

$$(\mu_1^{d_i}, \dots, \mu_{d_i}^{d_i}) = \text{espectro}(\widehat{H}_{d_i}),$$

donde el orden se está tomando decreciente, es decir $\mu_1^{d_i} \geq \dots \geq \mu_{d_i}^{d_i}$. Sea

$$(\mu_1, \dots, \mu_m) = (\text{espectro}(\widehat{H}_{d_1}), \dots, \text{espectro}(\widehat{H}_{d_K})).$$

Consideremos también a un movimiento browniano $\tilde{B} = (\tilde{B}_1, \dots, \tilde{B}_m)$ con matriz de covarianza (2.40). Entonces cuando $n \rightarrow \infty$ se tiene que

$$\left(\frac{\lambda_1^n - np_{\tau(1)}}{\sqrt{np_{\tau(1)}}}, \dots, \frac{\lambda_m^n - np_{\tau(m)}}{\sqrt{np_{\tau(m)}}} \right) \Rightarrow (\mu_1, \dots, \mu_m). \quad (2.55)$$

Demostración:

Por el teorema 2.2.2 y por el teorema del mapeo continuo será suficiente demostrar que

$$\left(\frac{L_m^1}{\sqrt{p_{\tau(1)}}}, \frac{L_m^2 - L_m^1}{\sqrt{p_{\tau(2)}}}, \dots, \frac{L_m^m - L_m^{m-1}}{\sqrt{p_{\tau(m)}}} \right) \stackrel{d}{=} (\mu_1, \dots, \mu_m). \quad (2.56)$$

Observemos que para $i = 1, \dots, K$, si λ es un valor propio de H_{d_i} asociado al vector propio v , entonces

$$\lambda - \text{tr}(HJ)\sqrt{p^{(i)}},$$

es un valor propio de \widetilde{H}_{d_i} que tiene a v con su vector propio asociado. En particular si denotamos por $\lambda_{\max}(U)$ al valor propio más grande de una matriz U se tiene que

$$\begin{aligned} \mu_1 &= \lambda_{\max}(\widetilde{H}_{d_1}) = -\text{tr}(HJ)\sqrt{p^{(1)}} + \lambda_{\max}(H_{d_1}), \\ &= -\sqrt{p^{(1)}} \sum_{q=1}^m \sqrt{p_{\tau(q)}} h_{qq} + \lambda_{\max}(H_{d_1}), \\ &= \frac{-p^{(1)} \sum_{q=1}^m \sqrt{p_{\tau(q)}} h_{qq} + \sqrt{p^{(1)}} \lambda_{\max}(H_{d_1})}{\sqrt{p^{(1)}}}, \end{aligned} \quad (2.57)$$

donde $H = (h_{ij})_{1 \leq i, j \leq m}$. Observemos que el numerador en (2.57) es muy similar a (2.41).

Ahora para $l = 2, \dots, m$ tenemos dos casos.

Caso 1: $p_{\tau(l)} = p_{\tau(l-1)}$.

Recordemos que existe una única $k \in \{1, \dots, K\}$ tal que $p_{\tau(l)} = p^{(k)}$, y además observemos que $p_{\tau(m_k+1)} = p^{(k)}$. En este caso de (2.54) se sigue que

$$\begin{aligned} L_m^l - L_m^{l-1} &= -p_{\tau(l)} \sum_{q=1}^m \sqrt{p_{\tau(q)}} B_q(1) \\ &\quad + \sqrt{p_{\tau(l)}} \sup_{\pi_1, \dots, \pi_{l-m_k} \in \widetilde{\Pi}_{d_k}} \sum_{j=1}^{l-m_k} \Delta_{\pi_j}(B_{m_k+1}, \dots, B_{m_k+d_k}) \\ &\quad - \sqrt{p_{\tau(l)}} \sup_{\pi_1, \dots, \pi_{l-1-m_k} \in \widetilde{\Pi}_{d_k}} \sum_{j=1}^{l-1-m_k} \Delta_{\pi_j}(B_{m_k+1}, \dots, B_{m_k+d_k}). \end{aligned}$$

Por lo que al dividir por $\sqrt{p_{\tau(l)}}$ obtenemos que

$$\begin{aligned} \frac{L_m^l - L_m^{l-1}}{\sqrt{p_{\tau(l)}}} &= -\sqrt{p_{\tau(l)}} \sum_{q=1}^m \sqrt{p_{\tau(q)}} B_q(1) \\ &+ \sup_{\pi_1, \dots, \pi_{l-m_k} \in \tilde{\Pi}_{d_k}} \sum_{j=1}^{l-m_k} \Delta_{\pi_j}(B_{m_k+1}, \dots, B_{m_k+d_k}) \\ &- \sup_{\pi_1, \dots, \pi_{l-1-m_k} \in \tilde{\Pi}_{d_k}} \sum_{j=1}^{l-1-m_k} \Delta_{\pi_j}(B_{m_k+1}, \dots, B_{m_k+d_k}). \end{aligned} \quad (2.58)$$

Caso 2: $p_{\tau(l-1)} > p_{\tau(l)}$.

En este caso existe una $s = 2, \dots, K$ tal que $l = m_s + 1$, y entonces

$$p^{(s-1)} = p_{\tau(l-1)} > p_{\tau(l)} = p^{(s)}.$$

De (2.54) se sigue que

$$\begin{aligned} L_m^l - L_m^{l-1} &= \left[\sum_{j=1}^{m_s} \sqrt{p_{\tau(j)}} B_j(1) - \sum_{j=1}^{m_{s-1}} \sqrt{p_{\tau(j)}} B_j(1) \right] \\ &- \left[\sum_{i=1}^l p_{\tau(i)} \sum_{q=1}^m \sqrt{p_{\tau(q)}} B_q(1) - \sum_{i=1}^{l-1} p_{\tau(i)} \sum_{q=1}^m \sqrt{p_{\tau(q)}} B_q(1) \right] \\ &+ \left[\sqrt{p_{\tau(m_s+1)}} \sup_{\pi_1, \dots, \pi_{l-m_s} \in \tilde{\Pi}_{d_s}} \sum_{j=1}^{l-m_s} \Delta_{\pi_j}(B_{m_s+1}, \dots, B_{m_s+d_s}) \right. \\ &\left. - \sqrt{p_{\tau(m_{s-1}+1)}} \sup_{\pi_1, \dots, \pi_{l-1-m_{s-1}} \in \tilde{\Pi}_{d_{s-1}}} \sum_{j=1}^{l-1-m_{s-1}} \Delta_{\pi_j}(B_{m_{s-1}+1}, \dots, B_{m_{s-1}+d_{s-1}}) \right], \\ &= \sum_{j=m_{s-1}+1}^{m_s} \sqrt{p_{\tau(j)}} B_j(1) - p_{\tau(l)} \sum_{q=1}^m \sqrt{p_{\tau(q)}} B_q(1) \\ &+ \sqrt{p_{\tau(l)}} \sup_{\pi \in \tilde{\Pi}_{d_s}} \Delta_{\pi}(B_{m_s+1}, \dots, B_{m_s+d_s}) \\ &- \sqrt{p_{\tau(m_{s-1}+1)}} \sup_{\pi_1, \dots, \pi_{l-1-m_{s-1}} \in \tilde{\Pi}_{d_{s-1}}} \sum_{j=1}^{d_{s-1}} \Delta_{\pi_j}(B_{m_{s-1}+1}, \dots, B_{m_{s-1}+d_{s-1}}), \\ &= -p_{\tau(l)} \sum_{q=1}^m \sqrt{p_{\tau(q)}} B_q(1) + \sqrt{p_{\tau(l)}} \sup_{\pi \in \tilde{\Pi}_{d_s}} \Delta_{\pi}(B_{m_s+1}, \dots, B_{m_s+d_s}). \end{aligned}$$

Por lo que al dividir por $\sqrt{p_{\tau(l)}}$ se tiene que

$$\frac{L_m^l - L_m^{l-1}}{\sqrt{p_{\tau(l)}}} = \frac{-\sqrt{p_{\tau(l)}} \sum_{q=1}^m \sqrt{p_{\tau(q)}} B_q(1) + \sup_{\pi \in \tilde{\Pi}_{d_s}} \Delta_{\pi}(B_{m_s+1}, \dots, B_{m_s+d_s})}{\sqrt{p_{\tau(l)}}}. \quad (2.59)$$

Por último de (2.57), (2.58), (2.59) y haciendo uso del teorema 2.1.1 y del lema 2.1.1 se sigue el resultado. ■

Los corolarios 2.2.5 y 2.2.2 nos dan una demostración de (2.38).

Corolario 2.2.6 Sea H_m una matriz GUE sin traza de dimensión m , y $B = (B_1, \dots, B_m)$ un movimiento browniano con matriz de covarianza

$$\begin{pmatrix} \frac{m-1}{m} & -\frac{1}{m} & \dots & -\frac{1}{m} \\ -\frac{1}{m} & \frac{m-1}{m} & \dots & -\frac{1}{m} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{m} & -\frac{1}{m} & \dots & \frac{m-1}{m} \end{pmatrix}. \quad (2.60)$$

Entonces

$$\mu_1 \stackrel{d}{=} \sup_{t_0=0 \leq \dots \leq t_{d_1}=1} \sum_{i=1}^{d_1} (B_i(t_i) - B_i(t_{i-1})), \quad (2.61)$$

donde μ_1 es el valor propio más grande de H_m .

Demostración:

Basta con tomar $d_1 = m$ en el corolario 2.2.5, y el resultado se sigue de la unicidad del límite en distribución en las expresiones (2.55) y (2.36). ■

2.3. Comentarios finales

En esta tesis se ha visto que si $B = (B_1, \dots, B_m)$ es un movimiento browniano tal que $\text{Cov}(B) = I_m$ entonces el funcional

$$L_m^1(B) := \sup_{0=t_0 \leq \dots \leq t_{d_1}=1} \sum_{i=1}^{d_1} (B_i(t_i) - B_i(t_{i-1})), \quad (2.62)$$

aparece en el teorema 1.3.2 como límite de tiempos de espera en teoría de colas, y en los teoremas 1.3.3 y 2.1.1 aparece una igualdad en distribución entre (2.62) y el valor propio más grande de una matriz GUE. Más específicamente, para una matriz hermitiana A denotemos por $\lambda_1(A)$ al valor propio más grande de A . Entonces si H_m es una matriz GUE de dimensión m , el corolario 2.1.1 nos dice que

$$L_m^1(B) \stackrel{d}{=} \lambda_1(H_m). \quad (2.63)$$

Notemos que $H_m = (h_{ij})_{i,j=1}^m$, entonces de la definición 1.2.1 se sigue que las variables en la diagonal de H_m son independientes. Es decir, si

$$D(H_m) := (h_{11}, \dots, h_{mm}) \quad (2.64)$$

entonces $\text{Cov}(D(H_m)) = I_m$. Por lo que

$$\text{Cov}(D(H_m)) = \text{Cov}(B). \quad (2.65)$$

Por otra parte, del corolario 2.2.6 se sigue que

$$L_m^1(B') \stackrel{d}{=} \lambda_1(H'_m), \quad (2.66)$$

donde

$$\text{Cov}(B') = \begin{pmatrix} \frac{m-1}{m} & -\frac{1}{m} & \dots & -\frac{1}{m} \\ -\frac{1}{m} & \frac{m-1}{m} & \dots & -\frac{1}{m} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{m} & -\frac{1}{m} & \dots & \frac{m-1}{m} \end{pmatrix}, \quad (2.67)$$

y H'_m es una matriz GUE sin traza. Siguiendo con la notación (2.64), el lema 2.2.2 nos dice que

$$\text{Cov}(D(H'_m)) = \text{Cov}(B').$$

Un poco más general, si consideramos a una matriz \widehat{H}_{d_1} como la que se definió en (2.22), el lema 2.2.2 nos dice que

$$\text{Cov}(D(\widehat{H}_{d_1})) = \begin{pmatrix} 1 - p^{(1)} & -p^{(1)} & \dots & -p^{(1)} \\ -p^{(1)} & 1 - p^{(1)} & \dots & -p^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ -p^{(1)} & \dots & \dots & 1 - p^{(1)} \end{pmatrix}. \quad (2.68)$$

Por otra parte, el de manera completamente análoga al corolario 2.2.2, y por la unicidad del límite en distribución en el corolario 2.2.5 y el teorema 2.2.1 se puede demostrar que

$$L_m^1(B'') \stackrel{d}{=} \lambda_1(\widehat{H}_{d_1}), \quad (2.69)$$

donde por el lema 2.2.2 se tiene que

$$\text{Cov}(D(\widehat{H}_{d_1})) = \text{Cov}(B'').$$

De estos tres ejemplos se sigue la siguiente

Conjetura:

Sea $H = (h_{ij})_{i,j=1}^m$ una matriz aleatoria tal que las entradas fuera de la diagonal son como las de un GUE, y las entradas en la diagonal tienen matriz de covarianza Σ . Sea $B = (B_1, \dots, B_m)$ un movimiento browniano con matriz de covarianza Σ , entonces

$$\lambda_1(H) \stackrel{d}{=} L_m^1(B),$$

donde $\lambda_1(H)$ es el valor propio más grande de H .

Bibliografía

- [1] Greg W Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*. Number 118. Cambridge University Press, 2010.
- [2] Jinho Baik, Percy Deift, and Kurt Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *Journal of the American Mathematical Society*, 12(4):1119–1178, 1999.
- [3] Yu Baryshnikov. Gues and queues. *Probability Theory and Related Fields*, 119(2):256–274, 2001.
- [4] Florent Benaych-Georges and Christian Houdré. Gue minors, maximal brownian functionals and longest increasing subsequences. *Markov Processes and Related Fields*, 21:109–126, 2015.
- [5] William Fulton. *Young tableaux: with applications to representation theory and geometry*, volume 35. Cambridge University Press, 1997.
- [6] Peter W Glynn and Ward Whitt. Departures from many queues in series. *The Annals of Applied Probability*, pages 546–572, 1991.
- [7] Janko Gravner, Craig A Tracy, and Harold Widom. Limit theorems for height fluctuations in a class of discrete space and time growth models. *Journal of Statistical Physics*, 102(5-6):1085–1132, 2001.
- [8] JM Hammersley. Postulates for subadditive processes. *The Annals of Probability*, pages 652–680, 1974.
- [9] Christian Houdré and Trevis J Litherland. On the limiting shape of young diagrams associated with markov random words. *arXiv preprint arXiv:1110.4570*, 2011.
- [10] Christian Houdré, Trevis J Litherland, et al. On the longest increasing subsequence for finite and countable alphabets. In *High Dimensional Probability V: The Luminy Volume*, pages 185–212. Institute of Mathematical Statistics, 2009.
- [11] Christian Houdré and Hua Xu. On the limiting shape of young diagrams associated with inhomogeneous random words. In *High Dimensional Probability VI*, pages 277–302. Springer, 2013.

- [12] Alexander R Its, Craig A Tracy, and Harold Widom. *Random words, Toeplitz determinants and integrable systems I*, volume 40. Mathematical Sciences Research Institute Publications, 2001.
- [13] Kurt Johansson. Shape fluctuations and random matrices. *Communications in mathematical physics*, 209(2):437–476, 2000.
- [14] Kurt Johansson. Discrete orthogonal polynomial ensembles and the plancherel measure. *Annals of Mathematics*, 153(1):259–296, 2001.
- [15] Kurt Johansson and Eric Nordenstam. Eigenvalues of gue minors. *Electronic Journal of Probability*, 11(50):1342–1371, 2006.
- [16] Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 2012.
- [17] Sergei Kerov. *Asymptotic representation theory of the symmetric group and its applications in analysis*, volume 219. AMS, Translations of Mathematical Monographs, 2003. (Russian edition: D. Sci thesis, 1994), 2003.
- [18] SV Kerov and AM Vershik. Asymptotics of the plancherel measure of the symmetric group and the limiting form of young tableaux. In *Soviet mathematics - Doklady*, volume 18, pages 527–531, 1977.
- [19] Benjamin F Logan and Larry A Shepp. A variational problem for random young tableaux. *Advances in mathematics*, 26(2):206–222, 1977.
- [20] AM Odlyzko and EM Rains. On longest increasing subsequences in random permutations. *Contemporary Mathematics*, 251:439–452, 2000.
- [21] James W Pitman. One-dimensional brownian motion and the three-dimensional bessel process. *Advances in Applied Probability*, pages 511–526, 1975.
- [22] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293. Springer Science & Business Media, 1999.
- [23] Dan Romik. *The surprising mathematics of longest increasing subsequences*, volume 4. Cambridge University Press, 2015.
- [24] Richard P Stanley. Enumerative combinatorics. vol. 2, volume 62 of cambridge studies in advanced mathematics, 1999.
- [25] Craig A Tracy and Harold Widom. Level-spacing distributions and the airy kernel. *Communications in Mathematical Physics*, 159(1):151–174, 1994.
- [26] Craig A Tracy and Harold Widom. On orthogonal and symplectic matrix ensembles. *Communications in Mathematical Physics*, 177(3):727–754, 1996.
- [27] Craig A Tracy and Harold Widom. On the distributions of the lengths of the longest monotone subsequences in random words. *Probability theory and related fields*, 119(3):350–380, 2001.

- [28] Stanislaw M Ulam. Monte carlo calculations in problems of mathematical physics. *Modern Mathematics for the Engineers*, pages 261–281, 1961.
- [29] Pierre van Moerbeke. Random and integrable models in mathematics and physics. In *Random Matrices, Random Processes and Integrable Systems*, pages 3–130. Springer, 2011.
- [30] Natalie Wolchover. At the far ends of a new universal law. <https://www.quantamagazine.org/20141015-at-the-far-ends-of-a-new-universal-law/>, 2014.