Universidad Autónoma de San Luis Potosí.

Facultad de Ciencias

# Matchings Bicromáticos en Conjuntos en Posición Convexa

TESIS

Que para obtener el Grado de

## Maestro en Ciencias Aplicadas

P R E S E N T A :

Ing. Mario Augusto Cetina Guerra

A S E S O R :

Dr. Gelasio Salazar Anaya

San Luis Potosí, S.L.P. Diciembre de 2005.

# Índice

# Capítulo 1

# Introducción

El tema de estudio del presente trabajo es uno de los objetos más clásicos dentro de la geometría combinatoria: las estructuras geométricas inducidas por conjuntos de puntos coloreados en el plano.

En este contexto, el objeto geométrico básico es un conjunto finito $S$ de $n$ puntos en el plano. El conjunto (o *configuración*) $S$ nos induce in *dibujo geométrico* de la gráfica completa $K_n$: la arista incidente con un par de puntos es simplemente el segmento de recta que los une.

## 1.1 Antecedentes: la Conjetura de Erdös para trayectorias alternantes

La clase de problemas que nos concierne ha sido recientemente popularizada en un artículo de revisión de Kaneko y Kano (ver [4]). La idea general consiste simplemente en asignar un color (tradicionalmente rojo o azul) a cada punto de $S$, de esta manera obteniendo un conjunto $R$ de puntos rojos y un conjunto $B$ de puntos azules, de tal modo que $S = B \cup R$. Los problemas tradicionalmente estudiados en este contexto se ocupan de las propiedades de subestructuras *escasas* (es decir, en las cuales el número de aristas es lineal en el número de vértices), tales como trayectorias o árboles (ver [1, 2, 6]).

En este trabajo presentaremos nuestro progreso reciente en un problema cuya historia se remonta a una pregunta de Erdös (ver [5]): si $|R| = |B|$, ¿cuál es el tamaño máximo de una trayectoria *alternante* (ésto es, sus vértices son alternadamente rojos y azules) sin cruces sobre $S$? (recordamos al lector que una trayectoria es una sucesión de vértices y aristas en la cual no se repite ningún vértice, y el tamaño de la trayectoria es su número de vértices).

Erdös propuso el estudio de este problema para lo que podría considerarse el caso más sencillo: los puntos en $S$ se encuentran en *posición convexa*, esto es, el cierre convexo de $S$ es igual a $S$. Erdös conjeturó que siempre existía una trayectoria de tamaño $(3/4)n$.

En lo sucesivo, si $S$ es un conjunto de puntos bicoloreado (donde el número de puntos rojos no es necesariamente igual al número de puntos azules), denotaremos por $\lambda(S)$ al tamaño de una trayectoria alternante máxima en $S$.

Kynčl, Pach, y Tóth [5], e independientemente Abellanas, García, Hurtado y Tejel [3], recientemente hallaron ejemplos de configuraciones en las cuales la trayectoria alternante máxima tiene longitud $(2/3)n + O(\sqrt{n})$. También observaron que es (bastante) sencillo probar que cualquier configuración tiene una trayectoria alternante de tamaño $n/2$, y Kynčl et al. fueron un poco más lejos y probaron la existencia de una trayectoria alternante de tamaño $n/2 + c\sqrt{n/\log n}$, donde $c$ es una constante. Resumimos estos resultados en el siguiente enunciado.

**Teorema 1 ([3, 5])**

$$n/2 + c\sqrt{n/\log n} \leq \min_{S=R\cup B,\,|R|=|B|} \lambda(S) \leq (2/3)n + O(\sqrt{n}),$$

*donde el mínimo se toma sobre todos los conjuntos $S$ bicoloreados en posición convexa con igual número de puntos rojos y azules.*

Aunque estos trabajos demuestran que la conjetura de Erdös es falsa, es claro que de ninguna manera resuelven el problema por completo. La determinación de la constante correcta de proporcionalidad es el centro de una vigorosa actividad de investigación en estos momentos.

## 1.2 Matchings lineales y trayectorias alternantes

Una de las herramientas principales en [3] y [5] es la cercana conexión que existe entre trayectorias alternantes y *matchings lineales bicromáticos*. Recordamos que un *matching geométrico*, o simplemente *matching*, es un conjunto de aristas tales que ningún par de aristas se cruza, y además ningún vértice es incidente con más de una arista del matching. Si el conjunto $S$ subyacente de puntos está en posición convexa, podemos entonces definir un matching *lineal* (también llamado *separable*): este es un matching para el cual existe una línea recta que corta a todas las aristas del matching. Finalmente, un matching lineal es *bicromático* si cada arista es incidente con un punto rojo y con un punto azul.

La pregunta fundamental es fácil de adivinar: dado un conjunto $S$ bicoloreado, ¿Cuál es el tamaño (número de vértices) $\nu(S)$ del matching lineal bicromático máximo sobre $S$?

Aunque como objeto geométrico–combinatorio básico, los matchings lineales bicromáticos tienen interés en sí mismos, [7] la importancia de estas estructuras en el estudio de trayectorias alternantes resulta evidente a partir del siguiente resultado.

**Teorema 2 ([3, 5])**

$$\lim_{n\to\infty} \min_{S=R\cup B, |R|=|B|} \frac{\lambda(S)}{n} = \lim_{n\to\infty} \min_{S=R\cup B, |R|=|B|} \frac{\nu(S)}{n}.$$

Este resultado refrenda el interés en el estudio del tamaño del matching lineal bicromático máximo sobre un conjunto de puntos bicoloreado. El problema central de estudio de nuestro trabajo es la estimación de $\nu(S)$, para el caso en que $S$ es bicoloreado aleatoriamente.

## 1.3 Nuestro problema: matchings lineales en conjuntos de puntos coloreados aleatoriamente

En el proyecto que culminó en este trabajo de tesis, nos dimos a la tarea de estudiar el problema de la estimación de $\nu(S)$ para el caso en el que la coloración sobre $S$ (en posición convexa) es realizada de manera aleatoria. Esto es, el color de cada punto de $S$ es asignado de manera aleatoria.

En concordancia con el problema original de Erdös y los problemas estudiados por Kynčl et al. y por Abellanas et al. , es natural restringir nuestra atención a conjuntos *equicoloreados*, esto es, aquellos en los cuales $|R| = |B|$. Así, nuestro problema fundamental es el siguiente.

**Pregunta 1** *¿Cuál es el valor esperado del tamaño máximo de un matching lineal bicoloreado en una colección aleatoriamente equicoloreada de n puntos en posición convexa?*

Antes de abundar un poco sobre las dificultades de abordar esta pregunta directamente, proponemos la siguiente pregunta relacionada.

**Pregunta 2** *¿Cuál es el valor esperado del tamaño máximo de un matching lineal bicoloreado en una colección aleatoriamente coloreada de n puntos en posición convexa?*

Aunque es relativamente sencillo proponer algoritmos para encontrar matchings lineales bicoloreados, con valores esperados no triviales, una breve reflexión revela que el análisis de un algoritmo sobre colecciones equicoloreadas presenta problemas de fondo: el análisis de tal algoritmo (al menos de los algoritmos más naturales) debe basarse en la perspectiva de que se exploran puntos (cuyo color es aleatorio), y de acuerdo a cierto criterio, se añaden (o no) aristas bicoloreadas al matching lineal. Sin embargo, la hipótesis de aleatoreidad no puede asumirse en cada punto explorado, debido a la restricción de que el punto a ser explorado no tiene necesariamente probabilidad igual de ser rojo que de ser azul: en efecto, si al momento de explorar el vértice hemos revelado un número de vértices rojos mayor al número de vértices azules, entonces en el conjunto de vértices por explorar existen evidentemente más vértices azules que rojos, de modo que un vértice elegido al azar tiene mayor probabilidad de ser azul. Este obstáculo de ninguna manera es trivial, y deberá ser tomado en cuenta en cualquier análisis formal.

Por otra parte, notamos que las dificultades mencionadas desaparecen por completo cuando consideramos el modelo planteado en la Pregunta 2 (quitando la restricción de equicoloración).

Afortunadamente, en un trabajo realizado paralelamente a nuestro proyecto, Moreno [8] recientemente demostró que, en el término relevante, ambas preguntas tienen la misma respuesta.

**Teorema 3** *Los valores esperados respectivos a las Pregutnas 1 y 2 son lineales en $n$ (más términos de menor orden), con el mismo coeficiente de proporcionalidad.*

Este resultado es de enorme importancia para nuestro trabajo, ya que nos permitirá proponer algoritmos para encontrar matchings lineales bicromáticos grandes, y demostrar formalmente cotas sobre el tamaño de los matchings lineales encontrados por estos algoritmos.

## 1.4 Estructura de esta tesis

El resto de este trabajo consiste de cuatro capítulos. El teorema principal (Teorema 4) se enuncia y demuestra en el Capítulo 2. La prueba del Teorema 4 es una simple consecuencia de los Teoremas 6 y 8, demostrados en los Capítulos 3 y 4, respectivamente. En el último capítulo presentaremos las conclusiones y algunos problemas abiertos relacionados con este trabajo.

# Capítulo 2

# El teorema principal: estimando el mayor matching lineal en una equicoloración aleatoria

Los resultados principales de este trabajo se resumen en el siguiente enunciado.

**Teorema 4 (Teorema Principal)** *El valor esperado $\Lambda_{eq}(n)$ del tamaño del matching lineal bicromático máximo en una equicoloración aleatoria en un conjunto de $n$ puntos en posición convexa satisface las siguientes desigualdades:*

$$0.7539n + O(\sqrt{n}) \leq \Lambda_{eq}(n) \leq 0.858n + O(\sqrt{n}).$$

Para demostrar la primera desigualdad, creamos un algoritmo que encuentra un matching lineal bicromático, y probamos que el tamaño esperado de este matching es precisamente $0.7539n + O(\sqrt{n})$. Para la cota superior, presentamos una combinación de argumentos combinatorios ideados *ad hoc* para este problema.

## 2.1 Demostración del teorema principal

*Prueba del Teorema 4 (provista solamente por claridad y completez).* Simplemente señalamos que el Teorema 4 se sigue de los Teoremas 6 y 8, a ser demostrados en los Capítulos 3 y 4, respectivamente. ∎

# Capítulo 3

# La cota inferior

Como se ha mencionado en la Introducción, nuestra principal contribución (Teorema 4) consistió en la demostración de cotas (inferior y superior) para el tamaño del matching lineal máximo en una equicoloración aleatoria. Nuestro objetivo en el presente capítulo es demostrar estas desigualdades.

En vista del Teorema 3, podemos enfocar nuestros esfuerzos a coloraciones (no necesariamente *equi*coloraciones) aleatorias. Esto es, para demostrar el Teorema 4 es suficiente establecer el siguiente enunciado (notamos que el término "ar" en $\Lambda_{ar}(n)$ se utiliza para enfatizar que contemplamos coloraciones *ar*bitrarias, y no necesariamente equicoloraciones, como en el caso de $\Lambda_{eq}(n)$).

**Teorema 5** *El valor esperado $\Lambda_{ar}(n)$ del tamaño del matching lineal bicromático máximo en una coloración aleatoria en un conjunto de $n$ puntos en posición convexa satisface las siguientes desigualdades:*

$$0.7539n + O(\sqrt{n}) \leq \Lambda_{ar}(n) \leq 0.858n + O(\sqrt{n}).$$

## 3.1 La estrategia

Nuestra estrategia para demostrar la cota inferior sigue los siguientes pasos:

**Paso 1** Dar un algoritmo (que llamaremos Algoritmo B) cuya entrada es una coloración de un conjunto $S = \{p_0, p_1, p_2, \ldots, p_{n-1}\}$ de $n$ puntos en posición convexa (etiquetados de acuerdo al orden cíclico contrarreloj en que aparecen en el cierre convexo de $S$), y cuya salida es un matching lineal bicromático.

**Paso 2** Demostrar que el tamaño esperado (asumiendo como entrada una coloración aleatoria) del matching lineal bicromático obtenido por Algoritmo B es al menos $0.7539n + O(\sqrt{n})$.

### 3.1.1 Calentando motores: el Algoritmo A

Antes de describir el algoritmo cuya validez y análisis implican la cota inferior del Teorema 5, consideramos pertinente presentar un algoritmo más débil, pero cuya descripción y análisis son bastante más accesibles. La simplicidad de las ideas detrás de este algoritmo preliminar (que llamaremos Algoritmo A) servirán como preparación para la presentación y análisis del Algoritmo B.

ALGORITMO A

*Entrada :* Un conjunto $S = \{p_0, p_1, p_2, \ldots, p_{n-1}\}$ de $n$ puntos en posición convexa, y una coloración $\chi : S \to \{0, 1\}$ (donde 0 denota el color rojo y 1 el color azul).

*Salida :* Un matching lineal $M$ bicromático sobre $S$.

*Procedimiento :*

1. Fijamos los valores iniciales $p_k := p_0$, $p_\ell = p_{n-1}$, $M = \emptyset$.
2. Preguntamos si $p_k, p_\ell$ tienen diferentes colores (ésto es, si $\chi(p_k) \neq \chi(p_\ell)$). En caso afirmativo:

   (i) Hacemos $M = M \cup \{(p_k, p_\ell)\}$ (ésto es, añadimos la arista $(q, r)$ a $M$).

(ii) Hacemos $p_k \longleftarrow p_{k+1}$ y $p_\ell \longleftarrow p_{\ell-1}$.

En caso negativo, hacemos $p_\ell \longleftarrow p_{\ell-1}$.

3. Si $\ell - k \geq 1$, volvemos al Paso 2. Si no, terminamos el proceso.

En términos coloquiales, el Algoritmo A trabaja de la siguiente manera. Comenzamos tomando el punto $p_k := p_0$, y pregunta los colores de los puntos (en este orden) $p_{n-1}, p_{n-2}, \ldots$, hasta encontrar un punto con diferente color al color de $p_0$. Al encontrar este punto (digamos $p_r$), añadimos la arista $(p_0, p_r)$ al matching $M$ (que en este punto tiene únicamente esta arista), y hacemos $p_k := p_1$ y $p_\ell := p_{r-1}$. Esto es, los puntos $p_0, p_{n-1}, \ldots, p_r$ quedan completamente desechados, y nunca más serán explorados en el algoritmo. Es el turno de encontrar "pareja" para $p_1$: buscamos en orden $p_\ell, p_{\ell-1}, \ldots$, en este orden, hasta encontrar un punto que tiene diferente color a $p_1$. Una vez encontrado tal punto, incluimos en $M$ la arista que une este punto con $p_1$, y procedemos a buscar pareja para $p_2$, etc. Naturalmente, cuando llegamos a un punto para el cual no encontramos pareja (situación que sucede si y solo si todos los puntos no explorados tienen el mismo color), el proceso se detiene.

¿Qué tan eficiente es el Algoritmo A? Esto es, si la entrada es una coloración aleatoria, ¿cuál es el tamaño esperado del matching lineal que arroja el algoritmo?

Para responder a esta pregunta, notamos que es posible visualizar la acción del algoritmo como un proceso de Markov. En efecto, fijemos como conjunto de estados a la colección de las parejas $(t, m)$, donde $t \geq 1$ y $m \geq 0$. En un paso del proceso, el parámetro $t$ indicará el número de vértices "descubiertos" hasta el momento, y $m$ denotará el número de aristas incluidas en nuestro matching lineal.

Para definir las probabilidades de transición, notamos que en el tiempo inicial, hemos descubierto únicamente 1 vértice, $p_0$. De este modo, el estado inicial es $(1, 0)$. De acuerdo al Algoritmo A, el paso siguiente es buscar (de entre $p_{n-1}, p_{n-2}, \ldots$, en este orden) un punto para emparejar a $p_0$. El primer esfuerzo consiste en ver si $p_{n-1}$ y $p_0$ tienen diferente color (lo que sucede con probabilidad $1/2$), en cuyo caso los emparejamos, y el estado se convierte en $(2, 1)$ (dos vértices descubiertos, 1 arista en el matching $M$). Si tienen el mismo color (lo cual sucede con probabilidad $1/2$), $p_{n-1}$ "muere", y el nuevo estado es $(2, 0)$. Ahora, si el estado actual es $(2,1)$, entonces, dado que el paso siguiente será revelar (el color de) $p_1$, podríamos para el mismo efecto decir que $(2,1)$ se transforma siempre en $(3,1)$ (equivalentemente, $(1,0)$ se transforma en $(3,1)$ con probabilidad $1/2$). Si el estado actual es $(2,0)$, seguimos buscando pareja para el punto $p_0$: es fácil ver que $(2,0)$ se convierte en $(3,1)$ con probabilidad $1/2$, y en $(3,0)$ también con probabilidad $1/2$.

Procediendo con el mismo razonamiento, vemos que la acción del algoritmo es modelada perfectamente por la Cadena de Markov con estados $(t,m)$, $t \geq 1, m \geq 0$, estado inicial $(1,0)$, y las siguientes condiciones sobre las transiciones:

$$\begin{aligned} P\Big((t,m) \mapsto (t+2,m+1)\Big) &= \frac{1}{2} \\ P\Big((t,m) \mapsto (t+1,m)\Big) &= \frac{1}{2} \end{aligned} \tag{3.1}$$

El análisis del Algoritmo A se transforma entonces en el análisis de esta Cadena de Markov. La información esencial, en el contexto de la Cadena de Markov, es el valor esperado de $m$ cuando $t$ llega al valor $n$ (recordamos que $n$ es el número de puntos en el conjunto $S$).

Para valores grandes de $n$, es fácil ver que esta información puede obtenerse equivalentemente con el siguiente enfoque: en lugar de interesarnos únicamente en la situación que ocurre cuando $t$ alcanza $n$, ponemos atención a lo que sucede *en promedio* en cada transición. Es evidente que, en promedio, en cada transición (en cada paso del algoritmo) se añaden (descubren) $2 \cdot (1/2) + 1 \cdot (1/2) = (3/2)$ vértices, y se añaden $1 \cdot (1/2) + 0 \cdot (1/2) = (1/2)$ aristas. Esto es, el "costo" de añadir 1 arista al matching es, en promedio de 3 vértices. Este simple razonamiento demuestra el siguiente enunciado.

**Proposición 1** *El valor esperado de vértices en el matching lineal bicromático que se obtiene al aplicar el Algoritmo A a un conjunto de n puntos en posición convexa coloreado aleatoriamente es de* $\big(2/3 + o(1)\big)n$.

Con este sencillo ejemplo como antecedente, pasamos ahora a presentar y analizar el Algoritmo B.

### 3.1.2 Paso 1: el Algoritmo B

Como en el Algoritmo A, en el Algoritmo B en cada paso también añadimos 0 o 1 aristas al matching $M$. Por otra parte, a diferencia del Algoritmo A, en el Algoritmo B no insistimos en tomar un vértice y forzosamente buscar un vértice de color diferente para unirlos con una arista a ser agregada al matching. En lugar de ésto, nos abrimos a la idea de descubrir vértices a ambos lados del casco convexo (y no únicamente a uno, como en el Algoritmo A, donde descubrimos sucesivamente los vértices $p_i, p_{i-1}, \ldots$ hasta encontrar un vértice útil).

Debido a que la aleatoreidad (en la coloración) es el ingrediente fundamental para analizar el Algoritmo B, debemos mantener siempre en mente el siguiente hecho evidente: *una vez que hemos descubierto el color de un vértice, es incorrecto asumir en un paso posterior que su color es rojo con probabilidad* $(1/2)$ *y azul con probabilidad* $(1/2)$. Tomando en cuenta ésto, un elemento esencial de nuestro algoritmo es, en cada paso, tener bien claro cuántos vértices *cuyo color ya ha sido descubierto* aún pueden formar parte del matching.

Otro elemento esencial que distingue al Algoritmo B del Algoritmo A, es que en cada paso no necesariamente descubrimos únicamente 1 vértice: bajo ciertas circunstancias, es posible descubrir 2 vértices en un solo paso.

Procedemos ahora a la descripción formal del Algoritmo B.

ALGORITMO B

En el Algoritmo B, en cada paso tenemos (i) un conjunto $D$ de vértices que ya no se considerarán en futuros pasos para ser incidentes con una arista que se incluirán en $M$ en futuros pasos; aclaramos que es posible que vértices de $D$ *ya* sean incidentes con aristas en $M$); (ii) un conjunto $L$ de vértices, cada uno de los cuales puede en principio ser incidente con una arista que se incluirá en $M$ en un futuro paso; y (iii) un conjunto $R$ de vértices, cada uno de los cuales puede en principio ser incidente con una arista que se incluirá en un futuro paso. Así, los vértices de $L$ y los vértices de $R$ tienen propiedades similares, pero se encuentran en regiones "opuestas" del casco convexo de $S$ (en los lados izquierdo y derecho, respectivamente, en un dibujo típico de $S$; de ahí la notación: $L$ por "left" y $R$ por "right").

En cada paso, se descubren nuevos vértices, ésto es, visitamos un vértice que no hemos considerado anteriormente, y preguntamos su color. Cuántos y cuáles vértices son descubiertos en cada paso depende de los tamaños relativos de $L$ y $R$, como especificamos abajo.

*Entrada :* Un conjunto $S = \{p_0, p_1, p_2, \ldots, p_{n-1}\}$ de $n$ puntos en posición convexa, y una coloración $\chi : S \longrightarrow \{0, 1\}$ (donde 0 denota el color rojo y 1 el color azul).

*Salida :* Un matching lineal $M$ bicromático sobre $S$.

*Procedimiento :*

NOTA 1: En esta descripción, $D, L$, y $R$ son conjuntos de vértices con las siguientes propiedades: $D$ es de la forma $\{p_\ell, p_{\ell-1}, \ldots, p_1, p_0, p_{n-1}, p_{n-2}, \ldots, p_{r-1}, p_r\}$; $L$ es (posiblemente vacío) de la forma $\{p_{\ell+1}, p_{\ell+2}, \ldots, p_{\ell'}\}$; $R$ es (posiblemente vacío) de la forma $\{p_{r-1}, p_{r-2}, \ldots, p_{r'}\}$.

NOTA 2: Como veremos a continuación, si al terminar un paso tenemos que ambos $L$ y $R$ son no vacíos, entonces $p_{\ell+1}$ y $p_{r-1}$ son del mismo color. Aún más: $|L| = |R|$. Reiteramos que esta observación se sigue trivialmente a partir del algoritmo.

1. Fijamos los valores iniciales de acuerdo a las siguientes reglas.

   (i) *Si $p_0$ y $p_{n-1}$ tienen diferente color :* Definimos $M := \{(p_0, p_{n-1})\}$; $D := \{p_0, p_{n-1}\}$; $L := \emptyset$; $R := \emptyset$.

   (ii) *Si $p_0$ y $p_{n-1}$ tienen el mismo color :* Definimos $M := \emptyset$; $L := \{p_0\}$; $R := \{p_{n-1}\}$.

2. Procedemos de acuerdo al número de elementos de $L$ y de $R$:

   *Aclaración :* Realizamos los procedimientos descritos mientras haya puntos sin descubrir. Si en algún punto nos quedamos sin puntos por descubrir, entonces el algoritmo se detiene y se devuelve el matching $M$.

(i) Si $L = R = \emptyset$, *descubrimos* (vemos por primera vez los colores de) $p_{\ell+1}$ y $p_{r+1}$. Si estos puntos tienen distintos colores, entonces añadimos a $M$ la arista $(p_{\ell+1}, p_{r-1})$, añadimos ambos vértices a $D$, y dejamos a $L$ y a $R$ sin cambio. Si tienen el mismo color, entonces hacemos $L = \{p_{\ell+1}\}$ y $R = \{p_{r-1}\}$, y dejamos a $M$ y a $D$ sin cambio.

(ii) Si $R = \emptyset$ y $L$ es no vacío, ésto es, $L = \{p_{\ell+1}, p_{\ell+2}, \ldots, p_{\ell'}\}$, entonces descubrimos los vértices $p_{r-1}, p_{r-2}, \ldots$ hasta encontrar un vértice con diferente color a $p_{\ell+1}$, **sin** buscar más de $|L|$ vértices.

Si encontramos tal vértice $p_i$, entonces añadimos la arista $(p_{\ell+1}, p_i)$ a $M$, añadimos $p_{\ell+1}$ y los vértices recién descubiertos a $D$ (ésto es, hacemos $D = \{p_{\ell+1}, p_\ell, \ldots, p_1, p_0, p_{n-1}, \ldots, p_i\}$), dejamos a $R$ vacío y quitamos $p_{\ell+1}$ a $L$ (ésto es, hacemos $L = \{p_{\ell+2}, p_{\ell+3}, \ldots, p_{\ell'}\}$; notamos que posiblemente $L = \emptyset$).

Si no encontramos tal vértice, entonces dejamos a $L$, a $D$, y a $M$ sin cambio, y conformamos a $R$ con los vértices recién descubiertos, ésto es, hacemos $R = \{p_{r-1}, p_{r-2}, \ldots, p_{r-(\ell'-\ell)}\}$ (notamos que en este caso tendremos $|L| = |R|$).

(iii) Si $L = \emptyset$ y $R$ es no vacío, ésto es, $R = \{p_{r-1}, p_{r-2}, \ldots, p_{r'}\}$, entonces descubrimos los vértices $p_{\ell+1}, p_{\ell+2}, \ldots$ hasta encontrar un vértice con diferente color a $p_{r-1}$, **sin** buscar más de $|R|$ vértices.

Si encontramos tal vértice $p_i$, entonces añadimos la arista $(p_{r-1}, p_i)$ a $M$, añadimos $p_{r-1}$ y los vértices recién descubiertos a $D$ (ésto es, hacemos $D = \{p_i, p_{i-1}, \ldots, p_1, p_0, p_{n-1}, \ldots, p_{r-1}\}$), dejamos a $L$ vacío y quitamos $p_{r-1}$ a $R$ (ésto es, hacemos $R = \{p_{r-2}, p_{r-3}, \ldots, p_{r'}\}$, notamos que posiblemente $R = \emptyset$).

Si no encontramos tal vértice, entonces dejamos a $R$, a $D$, y a $M$ sin cambio, y conformamos a $L$ con los vértices recién descubiertos, ésto es, hacemos $L = \{p_{\ell+1}, p_{\ell+2}, \ldots, p_{\ell+(r'-r)}\}$ (notamos que en este caso tendremos $|L| = |R|$).

(iv) Finalmente, si tanto $L$ como $R$ son no vacíos, esto es, $R = \{p_{r-1}, p_{r-2}, \ldots, p_{r'}\}$ y $L = \{p_{\ell+1}, p_{\ell+2}, \ldots, p_{\ell'}\}$, entonces descubriremos en un solo paso *dos* puntos: $p_{\ell'+1}$ y $p_{r'-1}$. Aún más: descubriremos sucesivamente conjuntos tales de dos puntos, hasta llegar a una condición que en breve especificaremos.

Recordemos el comentario de la Nota 2 (que se verifica fácilmente), $p_{r-1}$ y $p_{\ell+1}$ tienen el mismo color, y $|L| = |R|$ (ésto es, $r - r' = \ell' - \ell$). Procedemos de la siguiente forma: descubrimos las parejas $\{p_{\ell'+1}, p_{r'-1}\}$, $\{p_{\ell'+2}, p_{r'-2}\}$, y así sucesivamente, creciendo efectivamente los conjuntos $L$ y $R$ en cada descubrimiento de pareja (vamos añadiendo a $L$ los puntos $p_{\ell'+1}, p_{\ell'+2}$, etc., y vamos añadiendo a $R$ los puntos $p_{r'-1}, p_{r'-2}$, etc.), hasta que lo siguiente sea cierto para alguna $t > 0$: (A) es posible unir los puntos $p_{\ell+1}, p_{\ell+2}, \ldots, p_{\ell+t}$ con $t$ puntos de $R$ para formar $t$ aristas bicromáticas, pero *no* es posible unir los puntos $p_{r-1}, p_{r-2}, \ldots, p_{r-t}$ con $t$ puntos de $L$ para formar $t$ aristas bicromáticas; o (B) es posible unir los puntos $p_{r-1}, p_{r-2}, \ldots, p_{r-t}$ con $t$ puntos de $L$ para formar $t$ aristas bicromáticas, pero *no* es posible unir los puntos $p_{\ell+1}, p_{\ell+2}, \ldots, p_{\ell+t}$ con $t$ puntos de $R$ para formar $t$ aristas bicromáticas. Cuando llegamos a esta situación, abandonamos el proceso de descubrir parejas, y añadimos las $t$ aristas bicromáticas a $M$. Si el Caso (A) se aplica, removemos $p_{\ell+1}, p_{\ell+2}, \ldots, p_{\ell+t}$ de $L$ (añadiendo estos puntos a $D$), y pasamos todos los puntos de $R$ a $D$, haciendo $R = \emptyset$. Si el Caso (B) se aplica, removemos los puntos $p_{r-1}, p_{r-2}, \ldots, p_{r-t}$ de $R$ (añadiendo estos puntos a $D$), y pasamos todos los puntos de $L$ a $D$, haciendo $L = \emptyset$.

En cualquier caso, al terminar el proceso tendremos que $L = \emptyset$ o $R = \emptyset$.

3. Si aún quedan vértices por descubrir, volvemos al Paso 2. De otra manera, terminamos el algoritmo, y devolvemos el matching resultante $M$.

### 3.1.3 Paso 2: análisis del Algoritmo B

Dada su complejidad, en el sentido del número de posibilidades por analizar, no es sorprendente que un análisis exhaustivo del Algoritmo B es notablemente difícil de realizar. Es más, renunciamos rápidamente a la idea de realizar un análisis exacto, y trabajamos en variantes simplificadas de procesos de Markov que emulan el funcionamiento de este algoritmo. El análisis que presentamos a continuación representa un buen balance entre el resultado (tamaño del matching $M$) que garantiza y la dificultad inherente del modelo.

Como en el Algoritmo 1, nuestra idea es describir la situación en cada paso del Algoritmo 2 como un estado, y encontrar probabilidades de transición entre estados, con la esperanza de obtener un proceso de Markov analizable.

Una descripción completa se obtiene fijando los estados como vectores con toda la información al comenzar cada paso del algoritmo: el tamaño del matching $M$ actual, el tamaño de $L$, el tamaño de $R$ (inclusive los colores ya revelados de los vértices de $L$ y $R$), etc. Sin embargo, es fácil ver que pretender una descripción completa es demasiado ambiciosa, en el sentido de que las probabilidades de transición se vuelven difíciles (virtualmente imposibles) de calcular con exactitud.

Con estos obstáculos en mente, nos dimos a la tarea de buscar una descripción que reflejara suficientemente la complejidad del algoritmo, pero susceptible de ser analizada formalmente.

A continuación describiremos la cadena de Markov que diseñamos para simular el algoritmo, y posteriormente discutiremos cuál es la relevancia de esta cadena para el análisis del Algoritmo B.

### La cadena de Markov

Definimos la cadena de Markov con dos tipos de estados: $(k,t,m)$ y $(k,t,m)^*$, donde $k,t,m \geq 0$.

El estado inicial es $(0,0,0)$.

Las reglas y probabilidades de transición están dadas de la siguiente manera:

$$\begin{aligned} P\Big((0,t,m) \mapsto (0,t+2,m+1)\Big) &= \frac{1}{2} \\ P\Big((0,t,m) \mapsto (1,t+1,m)^*\Big) &= \frac{1}{2} \end{aligned} \tag{3.2}$$

Para $k \geq 1$,

$$
\begin{aligned}
P\Big((k,t,m) \mapsto (k-1,t+r+1,m+1)\Big) &= \frac{1}{2^r}, \ 1 \le r \le k \\
P\Big((k,t,m) \mapsto (k,t+k,m)^*\Big) &= \frac{1}{2^k}
\end{aligned}
\tag{3.3}
$$

También para $k \geq 1$,

$$
\begin{aligned}
P\Big((k,t,m)^* \mapsto (k-1,t+2,m+1)\Big) &= \frac{1}{2} \\
P\Big((k,t,m)^* \mapsto (k,t+2,m+1)^*\Big) &= \frac{1}{4} \\
P\Big((k,t,m)^* \mapsto (k+1,t+1,m)^*\Big) &= \frac{1}{4}
\end{aligned}
\tag{3.4}
$$

**La relación entre la cadena y el Algoritmo B**

Las transiciones entre estados de la cadena de Markov pueden ser interpretados como pasos del Algoritmo B, de la siguiente manera.

Interpretamos a $k$ como $\max\{|L|,|R|\}$, a $t$ como el número de vértices en $D$, y a $m$ como el número de aristas en $M$.

La interpretación de un estado con exponente $*$ es la siguiente: encontrarse en un estado tal corresponde a estar en el Caso (iv) del Paso 2 del Algoritmo B: $|L| = |R|$, y no hemos decidido cuál(es) arista(s) añadir a $M$ (posiblemente ninguna a esta altura).

Nuestro objetivo es demostrar que, siguiendo la evolución de la cadena, el cociente entre $m$ y $2t$ (que corresponde a la "efectividad" de los vértices: cuántos vértices se emplean en alcanzar $m$ aristas) es dominado (es menor o igual que) por $|M|/2|D|$.

Consideremos los estados descritos en (3.2), que corresponden al caso en que $|L| = |R| = 0$. En este caso, siguiendo la descripción del algoritmo encontramos que con probabilidad $1/2$ añadiremos una arista al descubrir la siguiente pareja de puntos, y con probabilidad $1/2$ encontraremos que la siguiente pareja de puntos tienen el mismo color. La primer posibilidad corresponde claramente a una transición al estado $(0, t+2, m+1)$. Por otra parte, la segunda posibilidad corresponde claramente a un estado $*$ (no se ha decidido cuál de los vértices explorados será incidente con una arista del matching). Hemos pasado de la situación $|L| = |R| = 0$ a la situación $|L| = |R| = 1$ (de aquí que corresponda moverse a un estado con $k = 1$), y sin importar cuál punto será incidente con una arista del matching, el otro vértice explorado será desperdiciado (de aquí el razonamiento de llevar la variable $t$ a $t+1$).

Consideremos ahora los estados descritos en (3.3), que corresponden al caso en que (de acuerdo a nuestra descripción anterior), o bien $|L| > 0$ y $|R| = 0$, o bien $|R| > 0$ y $|L| = 0$. Sin pérdida de generalidad, asumimos que $|L| > 0$ y $|R| = 0$. En este caso, de acuerdo a (ii) del Paso 2 del Algoritmo B, nuestra primer intento deberá ser hallar, dentro de los $|L|$ vértices siguientes a $p_r$, un vértice con diferente color al color de $p_{t+1}$. La probabilidad de tener éxito en el $r$–ésimo intento es de $1/2^r$, para $1 \leq r \leq k$; esto corresponde claramente a la primera transición de (3.3), como se ha postulado. Finalmente, si no tenemos éxito en la búsqueda, nos quedamos con los $|L|$ puntos recién descubiertos (que conforman ahora el conjunto $R$), y pasamos al caso $|L| = |R|$: esto sucede con probabilidad $1 - \sum_{r=1}^{k} 1/2^r = 1/2^k$, como se describe verazmente en la segunda transición de (3.3).

Finalmente, analizamos el conjunto de transiciones en (3.4). Este caso, que corresponde a (iv) del Paso 2 del Algoritmo B, se analiza similarmente a los anteriores, con la aclaración de que la sucesión de descubrimiento de parejas se hace *paso por paso* en la cadena de Markov: efectivamente, con probabilidad $1/2$ terminaremos (al revelar los colores de la siguiente pareja) con la situación $|L| = |R|$; equivalentemente, pasaremos al estado $(k, t+2, m+1)$. Con probabilidad $1/4$ continuaremos con la situación $|L| = |R|$ añadiendo (a futuro) una arista al matching (ésto corresponde a la transición a $(k, t+2, m+1)^*$). Y, finalmente, con probabilidad $1/4$ continuaremos con la situación $|L| = |R|$ sin añadir (a futuro) una arista al matching; éste último caso corresponde a la transición a $(k+1, t+1, m)^*$.

Esta discusión se resume en el siguiente enunciado.

**Proposición 2** *El límite del cociente entre la tercera y la segunda componentes del vector–estado de la cadena de Markov es menor o igual a la proporción esperada de vértices incidentes con una arista en un matching obtenido por el Algoritmo B.*

### Análisis de la cadena de Markov

Pretender un análisis directo de la cadena de Markov es una tarea formidable. Afortunadamente, un análisis completo no es necesario para nuestros fines: en vista de la Proposición 2, todo nuestro interés se centra en el cociente entre la tercera y la segunda componentes del vector–estado de la cadena.

Para estimar este cociente (asintóticamente), la primera simplificación necesaria es truncar el número de estados. Esta reducción tiene por objeto, naturalmente, acotar el problema a una cadena de Markov finita. Naturalmente, hay una pérdida asociada en este proceso (las simulaciones numéricas indican que es una pérdida despreciable, en todo caso), pero esta pérdida va en la dirección correcta: subestimaremos el cociente, por lo que el resultado obtenido aún será útil en el contexto de la Proposición 2.

La reducción que hemos realizado consiste en añadir una transición adicional: $P((k_0,t,m)^* \mapsto (k_0,t+2,m)^*) = 1/4$. Esta transición refleja la siguiente simplificación: si nos encontramos en (iv) del Caso 2 del Algoritmo B, y revelamos los colores de una pareja de vértices, podemos darnos el lujo de ignorar esa pareja de vértices, si no nos lleva de (el presente) estado $*$ a un estado sin $*$. Reiteramos que esta simplificación corresponde a un algoritmo más débil, pero hace el proceso analizable. Además, también reiteramos, la simplificación es en la dirección correcta: el resultado que obtengamos será una cota inferior para el valor esperado del matching lineal obtenido por el Algoritmo B.

La transición adicional descrita en el párrafo anterior es válida para cualquier $k_0$. En nuestro caso, encontramos que fijar $k_0 = 10$ resultó en un proceso manejable y analizable con facilidad.

Construimos la correspondiente matriz $Q$ de transición, de tamaño $21 \times 21$. Dada la naturaleza de las transiciones, no es sorprendente que $Q$ es ergódica, como es fácilmente demostrable. Existe entonces un único estado estacionario, que es fácil de calcular (es posible calcularlo numéricamente, y, dado el tamaño de la matriz, verificarlo a mano).

Hacemos uso del estado estacionario de la siguiente manera. La componente $j$-ésima del estado estacionario nos da la probabilidad de encontrarnos (asumiendo un gran número de pasos en el proceso, que corresponde a un conjunto $S$ con un gran número de puntos) en el estado $j$. Encontrándonos en el estado $j$, la matriz de transición $Q$ (equivalentemente, el conjunto de transiciones arriba definido) nos da las probabilidades de transición a otros estados. La observación crucial es que ésto conlleva la información de con qué probabilidad añadiremos $x$ aristas consumiendo $y$ vértices en el proceso. Realizando este análisis para cada entrada del estado estacionario, y ponderando por la probabilidad de encontrarnos en cada estado (ésto es, ponderando con cada entrada del estado estacionario) obtenemos el valor esperado para el cociente entre la tercera y las segunda coordenada del vector–estado de la cadena de Markov.

Nuevamente enfatizamos que realizamos esta labor numéricamente, pero el tamaño manejable del número de estados permite verificar los cálculos a mano.

Los resultados que obtuvimos se presentan en el siguiente enunciado.

**Proposición 3** *En la cadena de Markov, el valor esperado del cociente entre dos veces la tercera y la segunda componentes del vector–estado es estrictamente mayor a* 0.7539.

## 3.2 La cota inferior: el Teorema 6

En vista de la Proposición 2, este último resultado nos implica la siguiente cota inferior para el tamaño esperado del matching que arroja el Algoritmo B.

**Teorema 6** *El valor esperado de vértices en el matching lineal bicromático que se obtiene al aplicar el Algoritmo B a un conjunto de $n$ puntos en posición convexa coloreado aleatoriamente es mayor que* $0.7539n$. ∎

# Capítulo 4

# La cota superior

Nuestro objetivo en este capítulo es demostrar la cota superior en el Teorema 4.

Al igual que en la demostracion de el capítulo 3 hacemos énfasis en que gracias al Teorema 3 analizaremos funciones sobre coloraciones arbritarias $\Lambda_{ar}(n)$.

El resultado que demostraremos es aparentemente menos fuerte, pero un argumento de concentración demuestra que es equivalente a la cota superior en el Teorema 4.

## 4.1 $(P,Q)$–matchings

Consideraremos conjuntos de $2n$ puntos en el plano en posición convexa. El conjunto de puntos es particionado en dos conjuntos $P = \{p_1, p_2, \ldots, p_n\}$ y $Q = \{q_1, q_2, \ldots, q_n\}$, etiquetados de modo tal que los puntos $p_1, \ldots, p_n, q_n, \ldots, q_1$ aparecen en el casco convexo de $P \cup Q$ en el orden cíclico dado. Los puntos en $P$ son *$p$–puntos*, y los puntos en $Q$ son *$q$–puntos*.

Estaremos interesados en *$(P,Q)$–matchings*, esto es, matchings en los cuales cada arista es incidente con un $p$–punto y un $q$–punto.

Cuando cada punto en $P \cup Q$ es coloreado azul o rojo, el $2n$–conjunto de puntos resultante se convierte en una *$(2n)$–configuración*. Entonces, hay $2^{2n}$ posible $(2n)$–configuraciones.

Si $X$ es una $(2n)$–configuración, y $r \in P \cup Q$, entonces $\kappa_X(r)$ denotará el color de $r$ in $X$. Si no hay posibilidades de confusión, entonces omitiremos el subíndice $X$ y simplemente escribiremos $\kappa(r)$.

Un *buen* matching es un $(P,Q)$–matching en el cual cada arista es incidente con puntos de diferentes colores. Notamos que este concepto es (bastante) similar al concepto de matchings bicromáticos, pero en este caso además requerimos que cada arista tenga un extremo en $P$ y el otro en $Q$.

Nuestro objetivo es demostrar lo siguiente.

**Objetivo** *Dado un entero $m_0$, encontrar una cota superior para el n ´umero de $(2n)$–confi guraciones que tienen un buen matching de tama˜no $m_0$ o mayor.*

### 4.1.1 Matchings óptimos y parejas válidas $(X,M)$

Todo $(P,Q)$–matching $M$ en una $(2n)$–configuración $X$ tiene un orden natural de sus aristas. En efecto, las aristas de $M$ pueden ser naturalmente etiquetadas $\{e_1^M, e_2^M, \ldots, e_m^M\}$, de acuerdo a la regla de que el subíndice del $p$–extremo de $e_i^M$ es menor que el subíndice del $p$–extremo de $e_j^M$ si y solo si $i < j$ (notamos que podríamos equivalentemente haber tomado subíndices de $q$–puntos como referencia). Denotamos por $k_1^M, k_2^M, \ldots, k_m^M$ los subíndices de los $p$–extremos de las aristas en $M$: $p_{k_i^M}$ es el $p$–extremo de $e_i^M$. Similarmente, denotamos por $\ell_1^M, \ell_2^M, \ldots, \ell_m^M$ los subíndices de los $q$–extremos de las aristas en $M$.

Es evidente que la notación descrita, usando $M$ en todas partes como un superíndice, es engorrosa, así que la usaremos solamente cuando tengamos más de un matching en consideración. De otra forma, no haremos ninguna referencia a $M$, y simplemente escribiremos $e_i, p_{k_i}, q_{\ell_i}$.

Ahora definiremos un orden parcial sobre el conjunto de los matchings. Sea $M = \{e_1, e_2, \ldots, e_m\}$ un $(P,Q)$–matching. Asociamos a $M$ la secuencia finita $\tau(M) = (t_1, t_2, \ldots, t_m)$, donde $t_i := k_i + \ell_i - 2$. Esto es, Si vemos a $e_i$ como separador de $(P \cup Q) \setminus \{p_{k_i}, q_{\ell_i}\}$ en dos conjuntos, entonces $t_i$ es el tamaño de uno de estos conjuntos (es claro de la definicón cuál de ellos). Esto naturalmente define un orden parcial en el conjunto de $(P,Q)$–matchings en una $(2n)$–configuración: $M \preceq M'$ si y solo si $\tau(M)$ es lexicográficamente menor que, o igual a, $\tau(M')$.

Un buen matching $M$ en una $(2n)$–configuración $X$ es *óptima* si no hay un buen matching más grande en $X$, y $M$ es minimal con respecto a $\preceq$. Si $X$ es una $(2n)$–configuración y $M$ es un matching óptimo de $X$, entonces $(X, M)$ es una *pareja válida* de *orden* $2n$.

Sea $m$ un entero fijo. Denotamos por $\Lambda_{m,n}$ el número de parejas válidas $(X, M)$ de tamaño $2n$ tales que $M$ tiene al menos $m$ aristas.

la siguiente observación crucial es trivial, pues toda $(2n)$–configuración tiene un matching óptimo.

**Observación 1** *El número de $(2n)$–configuraciones que tienen un buen matching de tamaño $m_0$ es a lo más $\sum_{m=m_0}^{n} \Lambda_{m,n}$.*

### 4.1.2 Acotando $\Lambda_{m,n}$

Nuestro objetivo en la presente subsección es encontrar una cota superior para $\Lambda_{m,n}$.

Considérense dos aristas consecutivas $e_i, e_{i+1}$ en una pareja válida $(X, M)$. La *p–lata* definida por $e_i$ y $e_{i+1}$ es el (posiblemente vacío) conjunto de los $p$–puntos "entre" $e_i$ y $e_{i+1}$. Ésto es, la $p$–lata de $e_i$ y $e_{i+1}$ es el conjunto vacío si $k_{i+1} = k_i + 1$, y de otra manera es $\{p_{k_i+1}, \ldots, p_{k_{i+1}-1}\}$.

Si un punto $p_j$ está en la $p$–lata de $e_i$ y $e_{i+1}$, entonces $p_j$ está *entre* estas aristas. También decimos que está *arriba de* $e_{i+1}$ o, equivalentemente, *debajo de* $e_i$. Similarmente, debe ser claro lo que queremos decir cuando mencionamos que hay un punto *arriba de* $e_1$ o *debajo de* $e_m$.

Una serie totalmente análoga de definiciones se cumple para $q$–puntos.

**Proposición 4** *Sea $(X, M)$ una pareja válida de orden $2n$. Entonces lo siguiente se cumple:*

*(i) Si hay un p–punto arriba de (respectivamente abajo de) la arista $e_i$, entonces no puede haber ningún q–punto arriba de (respectivamente abajo de) $e_i$. Recíprocamente, si hay un q–punto arriba de (respectivamente abajo de)*

*$e_i$, entonces no puede haber ningún p–punto arriba de (respectivamente abajo de) $e_i$.*

*(ii) Si hay un p–punto (respectivamente q–punto) abajo de $e_i$, y $p_{k_i}$ y $p_{k_{i+1}}$ tienen el mismo color, entonces no puede haber un q–punto (respectivamente p–punto) arriba de $e_i$.*

*(iii) Si hay un p–punto (respectivamente q–punto) arriba de $e_i$, y $p_{k_i}$ y $p_{k_{i+1}}$ tienen el mismo color, entonces no puede haber un q–punto (respectivamente p–punto) abajo de $e_i$.*

*Prueba.* Si hubiera un $p$–punto arriba de (respectivamente abajo de) $e_i$, y además un $q$–punto arriba de (respectivamente abajo de) $e_i$, entonces claramente $M$ no sería óptimo o $M$ no sería minimal con respecto a $\preceq$, contradiciendo que $(X,M)$ es una pareja válida. Esto prueba (i). Los enunciados (ii) y (iii) se prueban análogamente. ∎

Un conjunto de aristas $\{e_i, e_{i+1}, \ldots, e_{j-1}, e_j\}$, con $j$ estrictamente mayor que $i$, es un *intervalo de p–ocupación* si (i) no hay ningún $p$–punto arriba de $e_i$; (ii) hay un $p$–punto abajo de cada uno de $e_i, e_{i+1}, \ldots, e_{j-1}$; y (iii) no hay ningún $p$–punto abajo de $e_j$. Aquí $e_i$ (respectivamente $e_j$) es la *primera* (respectivamente *última*) arista del intervalo. También decimos que $\{e_1\}$ es un intervalo de $p$–ocupación si hay un $p$–punto arriba de él, y ningún $p$–punto abajo de él; similarmente, $\{e_m\}$ es un intervalo de $p$–ocupación si hay un $p$–punto debajo de ella pero ningún $p$–punto arriba. Un intervalo de $p$–ocupación que incluye o bien $e_1$ o $e_m$ es *extremo*.

El siguiente enunciado es una simple observación.

**Proposición 5** *Si hay dos puntos en la misma p–lata, entonces tienen el mismo color.* ∎

Un intervalo no–extremo de $p$–ocupación $\{e_i, e_{i+1}, \ldots, e_j$ es *monocromático por arriba* si $e_i$ y $e_{i+1}$ tienen el mismo color, y es *monocromático por abajo* si $e_j$ y $e_{j+1}$ tienen el mismo color (nótese que $e_{j+1}$ no pertence al intervalo).

**Proposición 6** *Sea $(X,M)$ una pareja válida de orden $2n$. Supóngase que hay $k$ $p$–latas ocupadas en $(X,M)$, que abarcan $t$ intervalos de $p$–ocupación, $h$ de los cuales son monocromáticos por arriba. Entonces hay a lo más $m-(k+h)$ $q$–latas ocupadas.*

*Proof.* Esta es una consecuencia directa de las proposiciones anteriores, utilizando la hipótesis de que $(X,M)$ es una pareja válida. ∎

Nos encontramos finalmente en posición de establecer la cota superior buscada para $\Lambda_{m_0,n}$.

**Proposición 7**

$$\Lambda_{m_0,n} \leq \sum_{m=m_0}^{n} \sum_{k=0}^{n-m} \sum_{t=0}^{k} \sum_{h=0}^{t} \binom{k}{t}\binom{m-k}{t}\binom{n-m}{k}\binom{t}{h}\binom{n-k-h}{n-m} 2^{m-t}$$

*Proof.* La prueba no es más que un conteo cuidadoso con argumentos combinatorios clásicos. Hay $\binom{k}{t}\binom{m-k}{t}$ posibles formas de tener $k$ $p$–latas ocupadas abarcando $t$ intervalos de ocupación. Hay $\binom{n-m}{k}$ maneras de distribuir los $n-m$ puntos que no son incidentes con aristas en el matching en las $k$ latas. Hay $\binom{t}{h}$ formas de elegir cuáles $k$ intervalos son monocromáticos por arriba. Ahora, notamos que hay $\binom{t}{h}2^{m-t}$ formas de elegir los colores de los $p$–puntos incidentes con aristas en el matching (dado que hay $t$ intervalos, exactamente $h$ de los cuales son monocromáticos por arriba). Finalmente, $n-m$ $q$–puntos no incidentes con aristas en el matching pueden (deben) ser distribuidos en $m-(k+h)$ $q$–latas, y ésto puede hacerse de $\binom{n-k-h}{n-m}$ formas. ∎

### 4.1.3 Acotando el valor esperado del $(P,Q)$–matching máximo en una coloración aleatoria

El resultado principal de esta sección es el siguiente.

**Teorema 7** *El valor esperado del $(P,Q)$–matching máximo en una $(2n)$–configuración aleatoriamente coloreada es a lo más $0.8588n$. La varianza decrece* *exponencialmente con $n$. En particular, si realizamos un número polinomial de coloraciones aleatorias, el valor esperado del máximo $(P,Q)$–matching tomado sobre todas las coloraciones es también a lo más $0.8588n$, salvo términos de orden menor a $n$.*

*Prueba.* Llamémosle $M_0$ al valor esperado del $(P,Q)$–matching máximo en una $(2n)$–configuración aleatoriamente coloreada. Nuestro primer objetivo será acotar $M_0$ superiormente.

La observación crucial es la siguiente: la proporción entre los configuraciones que tienen un matching de tamaño $M_0$ o mayor, entre el total de configuraciones, debe estar acotado inferiormente por alguna constante (no importa qué tan pequeña), para toda $n$ suficientemente grande. Usando la Observación 1, esto se traduce en que debe existir una constante $c > 0$ tal que

$$\lim_{n\to\infty} \sum_{m=M_0}^{n} \sum_{k=0}^{n-m} \sum_{t=1}^{k} \sum_{h=0}^{t} \binom{k}{t} \binom{m-k}{t} \binom{n-m}{k} \binom{t}{h} \binom{n-k-h}{n-m} 2^{m-t} > c \cdot 2^{2n} \tag{4.1}$$

Para obtener una estimación sobre $M_0$, utilizamos la Aproximación de Stirling:

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{1}{12n+1}} < n! < \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{1}{12n}}, \tag{4.2}$$

y hacemos los cambios de variable

$$m = xn, \quad k = yn, \quad t = zn, \quad h = wn.$$

Con ésto, la Ec. (4.1) se transforma en una integral cuádruple, en las variables $x, y, z$, y $w$, con el siguiente integrando:

$$\left(\frac{(x-y)^{x-y}(1-y-w)^{1-y-w}\,2^{x-z}}{(1-x-y)^{1-x-y}(y-z)^{y-z}(x-y-z)^{x-y-z}(z-w)^{z-w}} \cdot \right.$$

$$\left. \frac{1}{(x-y-w)^{x-y-w} z^z\, w^w}\right)^n. \tag{4.3}$$

Para satisfacer la desigualdad, esta integral debe estar acotada por arriba, lo cual se cumple si y solo si el integrando es menor que 1 en todo el intervalo de integración:

$$x_o \leq x \leq 1, \quad 0 \leq y \leq 1-x, \quad 0 < z \leq y, \quad 0 \leq w \leq z.$$

Analizamos numéricamente estas condiciones, y hallamos que las condiciones descritas se satisfacen solo si $x_0 < 0.858$. Equivalentemente, si y solo si $M_0 < 0.858n$. Así, el valor esperado es a lo más $0.858n$, como se postuló. El enunciado sobre la concentración (varianza) es una comunicación privada de J. Balogh. ∎

## 4.2 La cota superior: el Teorema 8

Tomemos una coloración aleatoria de $n$ puntos en posición convexa. Existen menos de $n^2$ maneras de particionar los puntos en conjuntos $P$ y $Q$ como en la primera parte de este capítulo. Para cada manera de particionarlo, buscamos el $(P,Q)$–matching más grande. Notemos que $P$ y $Q$ no tienen necesariamente el mismo tamaño, pero eso no afecta la veracidad del Teorema 7: es fácil ver que el mayor valor esperado se obtiene cuando $P$ y $Q$ tienen igual tamaño. Podemos entonces aplicar el Teorema 7 a todas las (a lo más $n^2$) particiones, obteniendo que el valor esperado del máximo $(P,Q)$–matching (y, por lo tanto, del mayor matching lineal bicromático en el conjunto inicial) es a lo más $0.858n$ más términos de orden menor.

En conclusión, hemos probado lo siguiente.

**Teorema 8** *El valor esperado de vértices en el matching lineal bicromático que se obtiene al aplicar el Algoritmo B a un conjunto de n puntos en posición convexa coloreado aleatoriamente es a lo más* $0.858n$. ∎

# Capítulo 5

# Conclusiones

A modo de conclusión mencionaremos un par de direcciones para futuros proyectos relacionados.

En [3], Abellanas et al. presentan un algoritmo, basado en técnicas de programación dinámica, para generar el mayor matching lineal bicromático en un conjunto coloreado de puntos en posición convexa. Implementamos este algoritmo, y encontramos que el valor esperado para el tamaño del matching máximo es aproximadamente $0.81n$ (alrededor del promedio de nuestras cotas inferior y superior). Naturalmente, un posible proyecto consiste en mejorar cualquiera de nuestras cotas, y acercarse al empírico valor esperado de $0.81n$.

Otro problema abierto es encontrar cotas superior e inferior para el problema de trayectorias bicromáticas. En esta dirección, nuestras simulaciones en computadora revelan que el tamaño de la trayectoria máxima es aproximadamente $0.905n$. Sería muy interesante encontrar una cota inferior que fuera mayor a la cota superior que hemos encontrado para matchings en este trabajo: ello demostraría de una buena vez que las constantes de proporcionalidad son diferentes para caminos y para trayectorias, en el caso de coloraciones aleatorias.

# Bibliografía


[1] M. Abellanas, J. García, G. Hernández, M. Noy and P. Ramos Bipartite embeddings of trees in the plane *Discrete Appl. Math.* **93** (1999) 141–148.

[2] A. Kaneko, M. Kano, K. Yoshimoto Alternating Hamilton Cycles With Minimum Number of Crossings in the Plane *International Journal of Computational Geometry & Applications* **10** (2000) 73–78

[3] M. Abellanas, A. García, F. Hurtado y J. Tejel. Caminos Alternantes *Actas de los X Encuentros de Geometría Computacional*, Sevilla (2003) 7–12.

[4] A. Kaneko, M. Kano Discrete geometry on red and blue points in the Plane - A Survey - *Discrete and Computational Geometry*, Springer-Verlag, Berlin (2004) 551–570.

[5] Jan Kynčl, János Pach, Géza Tóth. Long Alternating Paths in Bicolored Point Sets. *Lecture Notes in Computer Science* **3383** (2004) 340–348.

[6] J. Leaños, C. Merino, G. Salazar and J. Urrutia Spanning trees of multicolored point sets with few intersections *Lecture Notes in Computer Science* **3330** (2005) 113– 122

[7] C. Merino, G. Salazar and J. Urrutia On the Intersection Number of Matchings and Minimum Weight Perfect Matchings of Multicolored Point Sets *Graph and Combinatorics* **21(3)** (2005) 333–341

[8] R. Moreno. Equivalencia de medias sin equivalencia de Ensambles *Tesis de Maestría en Ciencias Aplicadas* (2005)