

Estructuras de ARN y disecciones de polígonos: Una invitación a la combinatoria analítica

Ricardo Gómez Aíza

Instituto de Matemáticas de la
Universidad Nacional Autónoma de México,
Circuito Exterior, Ciudad Universitaria C.P. 04510,
México D.F. México.
rgomez@math.unam.mx

1. Motivación: Estructuras de ARN

El ácido ribonucleico (ARN) es una familia de moléculas que desempeñan diversas funciones en organismos celulares. Su compleja estructura se divide en tres partes fundamentales: primaria, secundaria y terciaria. La **estructura primaria** es la que nos es muy familiar y esencialmente es la respuesta que damos cuando alguien nos pregunta ¿sabes qué es el ARN? Responderemos que el ARN es una *cadena* de cuatro distintas bases nitrogenadas: adenina (A), citocina (C), guanina (G) y uracilo (U). Si nos piden un ejemplo, se nos ocurrirá algo como lo siguiente¹:

$$5\text{'-GCGGAUUUAUUCGCACCA-3'}$$
 (1)

La estructura primaria del ARN es entonces la estructura *lineal* (o 1-dimensional) y está completamente determinada por la sucesión de bases, es decir, por la «palabra» formada por las cuatro letras del alfabeto {A,C,G,U}. El modelo más simple de la estructura primaria del ARN es el **esquema de Bernoulli**, como se describe en la figura 1.

Rápidamente comenzamos a darnos cuenta de la complejidad: el ejemplo anterior (1) consta de tan solo 18 bases y escogimos una de $4^{18} = 68,719,476,736$ posibles cadenas! (claramente, hay 4^n estructuras primarias de cadenas de ARN de longitud $n \in \mathbb{N}$).

¹El 5' y el 3' simplemente establecen una *dirección lineal* a la cadena.

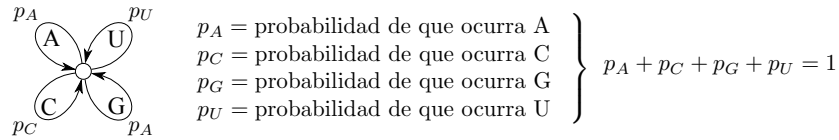


Figura 1. Estructura primaria de ARN modelada con un esquema de Bernoulli: una cadena de longitud n , digamos $x_1, \dots, x_n \in \{A, C, G, U\}^n$, corresponderá a un camino de longitud n en la gráfica dirigida asociada y su probabilidad será $\mathbb{P}(x_1, \dots, x_n) = p_{x_1} \dots p_{x_n}$.

La **estructura secundaria** del ARN es la que resulta cuando una cadena lineal de ARN se «pliega» bajo las reglas de apareamiento de Watson y Crick (las que aprendimos en la escuela: A-U y G-C) y que dan lugar a ciertos *patrones planos* de importancia biológica. Por ejemplo, dos posibles estructuras secundarias para (1) se describen en la figura 2.

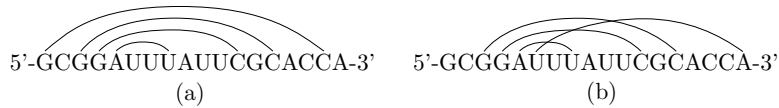


Figura 2. Dos posibles estructuras secundarias de ARN para la estructura primaria descrita en (1). (a) Estructura sin cruces. (b) Estructura con cruces (conocidos como *pseudonudos*, ver [6]).

Los arcos en cada uno de los dos casos de la figura 2 representan apareamientos y se dibujan en la parte superior de la cadena. Estos dos ejemplos se distinguen en un factor importante: en (a) no ocurren cruces mientras que en (b) si hay cruces. Desde luego, una misma cadena de ARN puede dar lugar a muchas estructuras secundarias diferentes y, en contraste con las estructuras primarias, ya no es obvio cuántas posibles estructuras secundarias hay.

Finalmente, la **estructura terciaria** es el resultado de las interacciones *en el espacio* de los átomos que conforman la molécula de ARN (debidas incluso a apareamientos distintos a los propuestos por Watson y Crick).

Ya que la estructura final de la molécula de ARN determina su función, de inmediato surgen preguntas relevantes y difíciles de contestar como:

- ¿Cuántas estructuras primarias, secundarias y terciarias puede tener una molécula de ARN?
- ¿Cuál es la probabilidad de que la estructura secundaria de una molécula de ARN tenga un cierto número de estructuras planas, incluso de un tipo específico?

Para responder satisfactoriamente este tipo de preguntas hay que tomar en cuenta no solo el problema matemático, debemos considerar también factores químicos, biológicos, etc., de allí la enorme dificultad. En el contexto matemático, las técnicas que se han utilizado provienen de la *combinatoria analítica* [5, 6], la cual conjunta tres ramas fundamentales de las matemáticas: análisis combinatorio, análisis complejo y probabilidad.

Nuestro propósito aquí es, tomando como punta de partida las estructuras del ARN —concretamente la primaria y secundaria—, presentar y desarrollar las definiciones y técnicas básicas de combinatoria analítica. Naturalmente consideraremos modelos muy simplificados que en última instancia consistirán en estudiar polígonos convexos y sus disecciones: los vértices corresponderían a las bases y los apareamientos serían las cuerdas de las disecciones (en particular, estaríamos considerando únicamente estructuras secundarias sin cruces, como el caso (a) de la figura 2). El parámetro que estudiaremos (que correspondería a una clase específica de estructuras planas) será simplemente el *número de partes*. Por ejemplo, la disección que correspondería al caso (a) de la figura 2 sería la de un polígono convexo con 18 vértices y tendría 5 partes: un triángulo, un cuadrado, un pentágono y dos heptágonos. Se ilustra en la figura 3 (ponemos las bases como referencia).

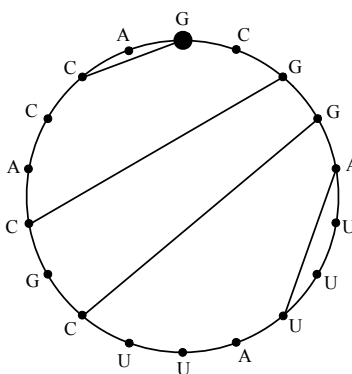


Figura 3. La estructura secundaria (a) de la figura 2 corresponde a una disección de un polígono convexo.

El resto del contenido del artículo es como sigue: En la sección §2 presentamos una breve introducción a la combinatoria analítica. En la sección §3 definiremos a los polígonos y sus disecciones, así como algunas clases auxiliares. En la sección §4 estudiamos el número de partes de disecciones aleatorias de polígonos. Esta sección contiene los resultados centrales del trabajo y la parte con las demostraciones es la más técnica y la que será más demandante al lector. Finalmente

concluimos con la sección §5 al considerar un caso concreto de una molécula de ARN de *Escherichia coli*.

2. Combinatoria analítica

2.1 Clases combinatorias

Definición 2.1. Una **clase combinatoria**, o simplemente una **clase**, consiste de un conjunto \mathcal{A} de **objetos** y de una **función de tamaño** $\tau_{\mathcal{A}}: \mathcal{A} \rightarrow \mathbb{N}$ que tiene la propiedad de que para cada $n \in \mathbb{N}$, existe una cantidad finita de objetos de tamaño n , es decir, si definimos

$$\mathcal{A}_n := \{a \in \mathcal{A} \mid \tau_{\mathcal{A}}(a) = n\} = \tau_{\mathcal{A}}^{-1}(n), \quad (2)$$

entonces $A_n := |\mathcal{A}_n| < \infty$. Diremos que la sucesión $\{A_n\}_{n=0}^{\infty}$ **enumera** a \mathcal{A} .

Entonces una clase es un par $(\mathcal{A}, \tau_{\mathcal{A}})$, sin embargo, para simplificar la notación, nos referiremos únicamente al conjunto de objetos \mathcal{A} como la clase siempre y cuando no haya ninguna ambigüedad en el contexto.

Naturalmente se sigue de la definición que el conjunto de objetos de una clase es un conjunto contable, es decir, es finito o numerable.

Ejemplo 2.2. Si \mathcal{R} es el conjunto de todas las cadenas de ARN (estructuras primarias) y $\tau_{\mathcal{R}}: \mathcal{R} \rightarrow \mathbb{N}$ es la función que regresa la longitud de una cadena, entonces $(\mathcal{R}, \tau_{\mathcal{R}})$ es una clase combinatoria. En este caso, como ya observamos, $R_n = |\mathcal{R}_n| = 4^n$.

Hay que distinguir dos clases fundamentales:

Definición 2.3. Una **clase neutra** consiste de un único objeto ϵ de tamaño cero. Una **clase atómica** consiste de un único objeto \bullet de tamaño uno.

2.2 Funciones generadoras

Asociada a una clase combinatoria \mathcal{A} está su **función generadora**:

$$A(z) = \sum_{n=0}^{\infty} A_n z^n.$$

En otras palabras, la función generadora $A(z)$ de una clase es la serie de Taylor alrededor del cero cuyo n -ésimo coeficiente $[z^n]A(z) := A_n$ es precisamente el número de objetos de tamaño n (por ésto se requiere que haya un número finito de ellos para cada $n \in \mathbb{N}$: en otro caso $f(z)$ no estaría definida como función para ninguna $z \in \mathbb{C}$, excepto quizás para $z = 0$), en particular es una serie de potencias con coeficientes *no-negativos*.

Ejemplo 2.4. La función generadora de la clase del ejemplo 2.2 es

$$R(z) = \sum_{n=0}^{\infty} 4^n z^n = \frac{1}{1-4z}.$$

Ejemplo 2.5. La función generadora de una clase neutra \mathcal{E} es

$$E(z) = 1$$

y la función generadora de una clase atómica \mathcal{Z} es

$$Z(z) = z.$$

Definición 2.6. Dos clases son **isomorfas** si sus funciones generadoras coinciden².

2.3 Suma, producto y sucesiones

Existen diversas operaciones entre clases que admiten «traducciones» a funciones generadoras. Las dos operaciones fundamentales son la *unión disjunta* y el *producto cartesiano*, las cuales se traducen a funciones generadoras como la suma y el producto respectivamente:

Definición 2.7. Sean \mathcal{B} y \mathcal{C} dos clases combinatorias.

- La **suma combinatoria** de \mathcal{B} y \mathcal{C} es la clase cuyo conjunto de objetos es $\mathcal{A} = \mathcal{B} \sqcup \mathcal{C}$ (es decir, la unión ajena de \mathcal{B} y \mathcal{C}) y cuya función de tamaño $\tau_{\mathcal{A}}: \mathcal{A} \rightarrow \mathbb{N}$ está definida para cada $x \in \mathcal{A}$ por la regla

$$\tau_{\mathcal{A}}(x) = \begin{cases} \tau_{\mathcal{B}}(x) & \text{si } x \in \mathcal{B} \\ \tau_{\mathcal{C}}(x) & \text{si } x \in \mathcal{C}. \end{cases}$$

- El **producto combinatorio** de \mathcal{B} y \mathcal{C} es la clase cuyo conjunto de objetos es $\mathcal{A} = \mathcal{B} \times \mathcal{C}$ (es decir, el producto cartesiano de \mathcal{B} y \mathcal{C}) y cuya función de tamaño $\tau_{\mathcal{A}}: \mathcal{A} \rightarrow \mathbb{N}$ está definida para cada $\mathbf{x} = (b, c) \in \mathcal{A} = \mathcal{B} \times \mathcal{C}$ por la regla

$$\tau_{\mathcal{A}}(\mathbf{x}) = \tau_{\mathcal{A}}(b, c) = \tau_{\mathcal{B}}(b) + \tau_{\mathcal{C}}(c).$$

Teorema 2.8 (Teorema I.1 en [5], pp. 27). Sean \mathcal{A} , \mathcal{B} y \mathcal{C} tres clases y $A(z)$, $B(z)$ y $C(z)$ sus funciones generadoras respectivamente.

- Si $\mathcal{A} = \mathcal{B} + \mathcal{C}$, entonces $A(z) = B(z) + C(z)$.
- Si $\mathcal{A} = \mathcal{B} \times \mathcal{C}$, entonces $A(z) = B(z)C(z)$.

Definición 2.9. Sea \mathcal{A} una clase sin objetos de tamaño cero, es decir, $\mathcal{A}_0 = \emptyset$. Definimos la **sucesión** de \mathcal{A} como la clase

$$\text{SUC}(\mathcal{A}) = \mathcal{A}^0 + \mathcal{A}^1 + \mathcal{A}^2 + \dots,$$

²Ver ejemplo 2.12 y también la sección §3.3.

donde \mathcal{A}^0 se define como una clase neutra y

$$\mathcal{A}^n = \underbrace{\mathcal{A} \times \cdots \times \mathcal{A}}_{n\text{-veces}} \quad \forall n \geq 1.$$

Observación 2.10. La definición anterior de sucesión involucra una suma combinatoria «infinita», la cual tiene sentido ya que, en virtud de la definición 2.7, la suma combinatoria de dos clases es una unión ajena, de forma que una suma combinatoria infinita la interpretamos simplemente como una unión ajena infinita; entonces la razón por la cual se requiere que no haya objetos de tamaño cero es que $\text{SUC}(\mathcal{A})$ no es una clase combinatoria si $\mathcal{A}_0 \neq \emptyset$ (¿por qué?).

Teorema 2.11 (Teorema I.1 en [5], pp. 27). *Sea \mathcal{B} una clase sin objetos de tamaño cero y sea $\mathcal{A} = \text{SUC}(\mathcal{B})$. Entonces*

$$A(z) = \frac{1}{1 - B(z)}.$$

Ejemplo 2.12. Sea $\mathcal{B} = \{a, b, c, d\}$ un alfabeto de cuatro símbolos, es decir una clase que consta de cuatro clases atómicas. Entonces $\text{SUC}(\mathcal{B})$ es el lenguaje sobre el alfabeto \mathcal{B} . Ya que $B(z) = 4z$, $\text{SUC}(\mathcal{B})$ es isomorfa a la clase \mathcal{R} del ejemplo 2.2.

Finalmente, podemos definir sucesiones restringidas: si $P \subset \mathbb{N}$, entonces

$$\text{SUC}_P(\mathcal{A}) := \sum_{n \in P} \mathcal{A}^n \quad (3)$$

(la sumatoria en (3) representa suma combinatoria, y si P es finito, entonces no es necesario pedir que $\mathcal{A}_0 \neq \emptyset$), y la función generadora correspondiente es

$$\sum_{n \in P} A(z)^n \quad (4)$$

(la sumatoria en (4) ahora sí es de funciones).

2.4 Parámetros y funciones generadoras bivariadas

Definición 2.13. Sea $(\mathcal{A}, \tau_{\mathcal{A}})$ una clase combinatoria. Un **parámetro** es una función $\rho: \mathcal{A} \rightarrow \mathbb{N}$. De forma similar a (2), para cada $n, k \in \mathbb{N}$, definimos

$$\mathcal{A}_{n,k} := \{a \in \mathcal{A}_n \mid \rho(a) = k\} = \rho^{-1}(k) \cap \mathcal{A}_n$$

y también $A_{n,k} := |\mathcal{A}_{n,k}|$. La **función generadora bivariada de \mathcal{A} con respecto a ρ** es

$$A(z, u) = \sum_{n,k=0}^{\infty} A_{n,k} z^n u^k.$$

Ejemplo 2.14. Continuando con la clase \mathcal{R} del ejemplo 2.2, la función $\rho: \mathcal{R} \rightarrow \mathbb{N}$ que regresa el número de ocurrencias de una base específica, digamos guanina, es un parámetro. En este caso es claro que

$$R_{n,k} = \binom{n}{k} 3^{n-k}$$

y por lo tanto

$$R(z, u) = \sum_{n=0}^{\infty} \sum_{k=0}^n \binom{n}{k} 3^{n-k} z^n u^k = \sum_{n=0}^{\infty} z^n (3+u)^n = \frac{1}{1-z(3+u)}.$$

Observemos que $R(z, 1) = R(z)$ (ver ejemplo 2.4).

2.5 Espacios de probabilidad y variables aleatorias

Sea \mathcal{A} una clase. Para cada $n \in \mathbb{N}$, podemos considerar al conjunto \mathcal{A}_n de objetos de tamaño n como un espacio de probabilidad (discreto) al declarar que todos sus objetos son equiprobables, es decir, definimos en \mathcal{A}_n la medida de probabilidad *uniforme* $\mathbb{P}_n: \mathcal{A}_n \rightarrow [0, 1]$, de forma que para cada $a \in \mathcal{A}_n$,

$$\mathbb{P}_n(a) = \frac{1}{A_n}.$$

De esta manera, un parámetro $\rho: \mathcal{A} \rightarrow \mathbb{N}$ da lugar a una variable aleatoria

$$\rho = \rho|_{\mathcal{A}_n}: \mathcal{A}_n \rightarrow \mathbb{N}.$$

Ejemplo 2.15. Continuando con el ejemplo 2.14, observamos que, para cada $n \in \mathbb{N}$, la función de probabilidad de ρ es

$$\mathbb{P}_n(\rho = k) = \frac{\binom{n}{k} 3^{n-k}}{4^n} = \binom{n}{k} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{n-k},$$

es decir, ρ tiene distribución Binomial($n, 1/4$).

Teorema 2.16 (Proposición III.2 en [5], pp. 158). *Sea \mathcal{A} una clase combinatoria y $\rho: \mathcal{A} \rightarrow \mathbb{N}$ un parámetro. Sea $A(z, u)$ la función generadora bivariada de \mathcal{A} con respecto a ρ . Para cada $n \in \mathbb{N}$, la esperanza y la varianza de $\rho = \rho|_{\mathcal{A}_n}: \mathcal{A}_n \rightarrow \mathbb{N}$ están dadas por*

$$\mathbb{E}_{\mathcal{A}_n}(\rho) = \frac{[z^n] \partial_u A(z, u)|_{u=1}}{A_n} \quad (5)$$

y

$$\mathbb{V}_{\mathcal{A}_n}(\rho) = \frac{[z^n] \partial_u^2 A(z, u)|_{u=1}}{A_n} + \mathbb{E}_{\mathcal{A}_n}(\rho) - \mathbb{E}_{\mathcal{A}_n}(\rho)^2 \quad (6)$$

respectivamente.

Observación 2.17. A la función

$$\Omega(z) := \partial_u A(z, u)|_{u=1}$$

se le conoce como **función generadora acumulativa** de ρ .

Ejemplo 2.18. Continuando con el ejemplo 2.14, observamos que la esperanza del número de guaninas en una cadena de ARN³ de tamaño n es

$$\mathbb{E}_{\mathcal{R}_n}(\rho) = \frac{[z^n] \partial_u R(z, u)|_{u=1}}{R_n} = \frac{[z^n] \frac{z}{(1-4z)^2}}{4^n} = \frac{4^{n-1} \binom{n}{1}}{4^n} = \frac{n}{4},$$

lo cual es, desde luego, intuitivamente claro pero también consistente con la observación del ejemplo 2.15 acerca de la distribución binomial. El lector podrá entonces verificar, usando (6), que

$$\mathbb{V}_{\mathcal{R}_n}(\rho) = \frac{3}{16}n.$$

2.6 Análisis asintótico

Los coeficientes de todas las funciones generadoras que veremos aquí tendrán una *forma explícita* (i.e. una fórmula —no recursiva—). Sin embargo, este no es el caso general, por el contrario, lo común es que *no* exista una fórmula explícita de los coeficientes de una función generadora. Incluso con una fórmula explícita, evaluarla puede ser una tarea muy costosa computacionalmente. En cualquier caso lo que a uno le interesa es determinar el *comportamiento asintótico* de los coeficientes.

2.6.1. Notación asintótica

Sean $\{a_n\}_{n=0}^\infty$ y $\{b_n\}_{n=0}^\infty$ dos sucesiones de números (en general complejos).

- Diremos que a_n es de *orden a lo más* b_n si existe $C > 0$ tal que $|a_n| \leq C|b_n|$ para toda $n \geq 0$, en cuyo caso escribiremos

$$a_n = O(b_n).$$

- Diremos que a_n es *asintóticamente equivalente* a b_n si $\lim_{n \rightarrow \infty} |a_n|/|b_n| = 1$, en cuyo caso escribiremos

$$a_n \sim b_n.$$

- Diremos que a_n es de *orden menor a* b_n si $\lim_{n \rightarrow \infty} a_n/b_n = 0$, en cuyo caso escribiremos

$$a_n = o(b_n).$$

³Bajo un esquema de Bernoulli con $p_A = p_C = p_G = p_U = \frac{1}{4}$ (ver figura 1).

Así, una expresión del estilo

$$a_n = b_n + O(c_n)$$

en donde se cumple que $c_n = o(b_n)$ no solo significa que $a_n \sim b_n$ sino también indica que la «desviación» $a_n - b_n$ es de orden a lo más c_n .

2.6.2. Crecimiento exponencial

Una sucesión de números $\{a_n\}_{n=0}^{\infty}$ es *exponencial de orden* K^n si

$$\limsup_{n \rightarrow \infty} |a_n|^{1/n} = K \neq 1,$$

en cuyo caso escribiremos

$$a_n \asymp K^n.$$

Por otro lado, si ocurre que

$$\limsup_{n \rightarrow \infty} |a_n|^{1/n} = 1,$$

entonces diremos que $\{a_n\}_{n=0}^{\infty}$ es *subexponencial*.

2.6.3. Los dos principios del crecimiento asintótico de coeficientes

Sea $f(z) = \sum_{n=0}^{\infty} f_n z^n$ una serie de potencias con coeficientes en \mathbb{R} . El **radio de convergencia** $R \in [0, \infty]$ de f está definido por

$$R^{-1} = \limsup_{n \rightarrow \infty} |f_n|^{1/n}. \quad (7)$$

Si $R > 0$, entonces f define una función analítica $f: D \rightarrow \mathbb{C}$, donde $D = \{z \in \mathbb{C} : |z| < R\}$. Si $R < \infty$, entonces $f(z)$ debe de tener al menos una *singularidad* en la frontera $\partial D = \{z \in \mathbb{C} : |z| = R\}$. Las singularidades en la frontera de D se conocen como **singularidades dominantes**. El siguiente teorema indica que para funciones generadores de clases combinatorias, el radio de convergencia *es* una singularidad dominante.

Teorema 2.19 (Teorema de Pringsheim, Teorema IV.6 en [5], pp. 240). *Si $f(z) = \sum_{n=0}^{\infty} f_n z^n$ es una serie de potencias con coeficientes en \mathbb{R}^+ y $R \in (0, \infty)$ es su radio de convergencia, entonces $z = R$ es una singularidad de $f(z)$.*

Por lo anterior, el crecimiento exponencial de los coeficientes $\{f_n\}_{n=0}^{\infty}$ lo determina el radio de convergencia R , más precisamente,

$$f_n \asymp \left(\frac{1}{R}\right)^n \quad (8)$$

de forma que existe una sucesión subexponencial $\{\theta(n)\}_{n=0}^{\infty}$ tal que

$$f_n = \theta(n)R^{-n}.$$

Para finalmente determinar la equivalencia asintótica de $\{f_n\}_{n=0}^{\infty}$ hay que encontrar la equivalencia asintótica del factor subexponencial $\theta(n)$.

- **Primer principio de crecimiento asintótico de coeficientes.** La *posición* de las singularidades dicta el crecimiento exponencial de los coeficientes.
- **Segundo principio de crecimiento asintótico de coeficientes.** La *naturaleza* de las singularidades determina el crecimiento subexponencial de los coeficientes.

El primer principio ya lo justificamos con (8). El segundo principio lo conforman un conjunto de resultados de carácter analítico que corresponden a los distintos tipos de singularidades: *polos* (i.e. funciones meromorfas) y singularidades esenciales, concretamente *puntos ramificantes*. La única instancia del segundo principio que usaremos aquí es el Lema 4.6, que corresponde a un tipo específico de singularidades de puntos ramificantes.

3. Polígonos, disecciones, árboles y retículas

3.1 Polígonos

Todos estamos familiarizados con los **polígonos convexos**, los cuales podemos definir como la frontera del casco convexo de un conjunto finito de dos o más puntos (los **vértices**) colocados en posición convexa en el plano Euclidiano. Para ser considerados una clase \mathcal{P} , vamos a definir el **tamaño** de un polígono como el número de vértices menos dos. Como gráficas, un polígono convexo es una arista si su tamaño es cero (el **polígono vacío** \emptyset) o bien un ciclo. Podemos dibujar a los polígonos en el plano de forma **regular**, es decir, colocar los vértices en forma equidistante a lo largo de un círculo, como se ilustra en la figura 4.

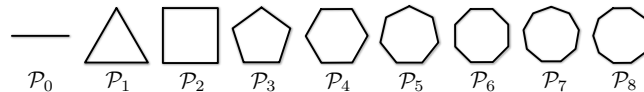


Figura 4. Los primeros polígonos regulares.

Un polígono es **enraizado** si uno de sus vértices (la **raíz**) es escogido junto con una orientación fija (por ejemplo, en la dirección de las manecillas del reloj, de forma que podemos hablar de la **primer arista del polígono**). Dos polígonos convexos enraizados serán considerados

«iguales» si sus tamaños coinciden, ya que en este caso, vistos como gráficas, son isomorfos por medio de un único isomorfismo que identifica a la raíz y preserva la orientación. En otras palabras, para cada $n \geq 0$, existe un único polígono⁴ \mathcal{P}_n de tamaño n , el cual lo podemos identificar con la clase \mathcal{Z}^n , donde \mathcal{Z} es una clase atómica (por definición, $\mathcal{Z}^0 = \mathcal{E}$ es la clase neutra). Así, podemos *especificar* la clase combinatoria de los polígonos convexos enraizados como

$$\mathcal{P} = \mathcal{E} + \mathcal{Z} + \mathcal{Z} \times \mathcal{Z} + \mathcal{Z} \times \mathcal{Z} \times \mathcal{Z} + \dots = \text{SUC}(\mathcal{Z}).$$

Entonces la función generadora de los polígonos es la serie geométrica

$$P(z) = 1 + z + z^2 + z^3 + z^4 + \dots = \frac{1}{1 - z}.$$

3.2 Disecciones

Una **cuerda** de un polígono es una arista entre cualesquiera dos de sus vértices no adyacentes. Una **disección** de un polígono es un conjunto (posiblemente vacío) de cuerdas que no se intersectan entre sí (pueden tener vértices en común). Sea \mathcal{S} la clase de todas las disecciones de los polígonos, donde el tamaño de una disección es el tamaño del polígono disectado, y sea \mathcal{S}_n las disecciones de tamaño $n \geq 0$. Las disecciones de tamaño ≤ 4 se ilustran en la figura 5.

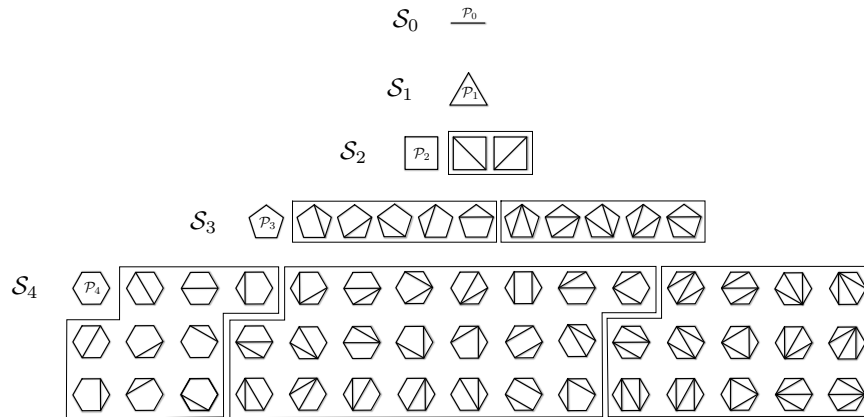


Figura 5. Disecciones de tamaño ≤ 4 . Para cada $n \geq 0$ existe una partición del conjunto de disecciones de tamaño n de acuerdo al número de partes; para cada $n \geq 1$ la partición que se muestra al final corresponde a las triangulaciones, enumeradas por los números de Catalan [2]. Los números de Schröder-Hipparchus enumeran a las disecciones (e.g. [7, 3, 1]).

Para deducir la función generadora $S(z)$ de las disecciones, primero observamos que una disección resulta de un polígono al cual se le «pegan» disecciones a lo largo de las aristas del polígono (excepto a la

⁴De ahora en adelante todos los polígonos son convexos y enraizados.

primera arista adyacente a la raíz). Esta construcción se ilustra en la figura 6.

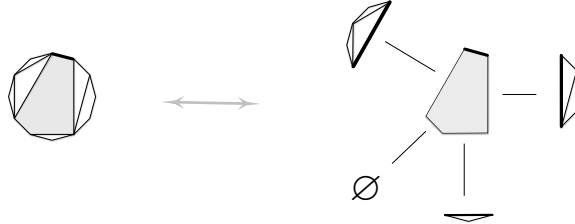


Figura 6. Una disección se define (recursivamente) como un polígono «base» de tamaño $n \geq 1$ junto con una sucesión de longitud $n + 1$ de disecciones, cuyas primeras aristas se identifican con las aristas (distintas a la primera) del polígono base.

Podemos entonces especificar recursivamente⁵ a las disecciones como

$$\mathcal{S} = \mathcal{E} + \mathcal{P}_1 \times \mathcal{S} \times \mathcal{S} + \mathcal{P}_2 \times \mathcal{S} \times \mathcal{S} \times \mathcal{S} + \mathcal{P}_3 \times \mathcal{S} \times \mathcal{S} \times \mathcal{S} \times \mathcal{S} + \dots \quad (9)$$

Entonces $S(z)$ satisface la ecuación funcional

$$S(z) = 1 + \sum_{n=1}^{\infty} z^n S(z)^{n+1} = 1 + \frac{zS(z)^2}{1 - zS(z)}$$

que equivale a

$$2zS(z)^2 - (z + 1)S(z) + 1 = 0 \quad (10)$$

y por lo tanto

$$S(z) = \frac{z + 1 - \sqrt{z^2 - 6z + 1}}{4z} \quad (11)$$

(hemos escogido el signo $-$ al resolver la ecuación de segundo grado (10) porque si escogemos el signo $+$, al menos un coeficiente de la serie de Taylor de la función que resulta es negativo, lo cual no puede ser por definición). La función $S(z)$ es la función generadora de los números de Schröder-Hipparchus, que corresponden a *EIS A001003* en [8]. Entonces

$$S_n = \frac{1}{n + 1} \sum_{k=0}^{n-1} \binom{2n - k}{n} \binom{n - 1}{k}.$$

Los primeros números de Schröder-Hipparchus son como siguen:

n	0	1	2	3	4	5	6	7	8	9	10	...
S_n	1	1	3	11	45	197	903	4279	20793	103049	518859	...

⁵Una especificación de una clase es *recursiva* si involucra a la clase misma en su expresión. En general, una especificación recursiva da lugar a una ecuación funcional de la función generadora de la clase, la cual, en ocasiones, se puede resolver.

3.3 Árboles planos enraizados sin ramas

Sea \mathcal{B} la clase combinatoria de los **árboles planos enraizados sin ramas**, es decir, el conjunto de árboles planos enraizados con todos los nodos internos con al menos dos descendientes (siempre consideraremos a la raíz como un nodo interno), siendo el tamaño de un elemento de \mathcal{B} el número de nodos terminales. Sea $\mathcal{B}_n \subset \mathcal{B}$ el conjunto de todos los árboles planos enraizados sin ramas de tamaño $n \geq 0$. Consideremos la especificación recursiva⁶

$$\mathcal{B} = \text{SUC}_{\geq 2}(\mathcal{Z} + \mathcal{B}). \quad (12)$$

Por (12), deducimos que la función generadora $B(z)$ de \mathcal{B} satisface la ecuación funcional

$$B(z) = \frac{(z + B(z))^2}{1 - z - B(z)},$$

por lo tanto

$$B(z) = \frac{1 - 3z - \sqrt{z^2 - 6z + 1}}{4}.$$

Observemos que $B_0 = B_1 = 0$ y $B_n = S_{n-1}$ para toda $n \geq 2$ ya que $B(z) = z(S(z) - 1)$. Entonces las clases \mathcal{B} y \mathcal{S} son esencialmente las mismas ya que para toda $n \geq 2$, una disección en \mathcal{S}_n corresponde en forma biyectiva a un árbol en \mathcal{B}_{n+1} , o más precisamente, $\mathcal{Z} \times \mathcal{S}_{\geq 2}$ y \mathcal{B} son combinatoriamente isomorfas. Podemos describir un *isomorfismo* entre estas dos clases identificando los nodos internos de un árbol con los polígonos que conforman la disección correspondiente. El tamaño del polígono en cada nodo es el número de descendientes más uno; las aristas entre un nodo y su ancestro corresponden a las adyacencias de las caras de los polígonos en la disección. Este isomorfismo se ilustra en la figura 7.

3.4 Retículas y números de Delannoy centrales

Los números de Delannoy surgen al considerar el total de formas posibles para cruzar una retícula de dimensiones $n \times m$, desde la esquina inferior izquierda hasta la esquina superior derecha, con únicamente tres movimientos: hacia la derecha, en diagonal (de izquierda a derecha y de abajo hacia arriba) y hacia arriba. Los números de Delannoy centrales corresponden al caso cuando $m = n$ (ver figura 8). La función (ver e.g. [9])

$$D(z) = \frac{1}{\sqrt{z^2 - 6z + 1}}$$

⁶Por ≥ 2 nos referimos al conjunto $P = \{n \in \mathbb{N} \mid n \geq 2\}$.

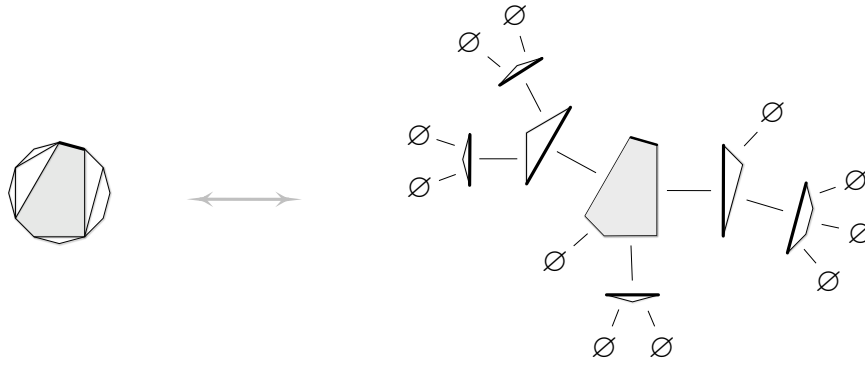


Figura 7. Una disección también es un árbol plano enraizado sin ramas, donde los nodos internos del árbol se identifican con los polígonos que conforman la disección. La raíz del árbol es el único polígono de la disección que contiene a la primera arista del polígono disectado, y las adyacencias de las caras de los polígonos en la disección determinan a los descendientes de cada nodo del árbol. Los nodos internos del árbol tienen al menos dos descendientes (por eso el árbol no tiene «ramas») y los nodos terminales corresponden a las aristas del polígono disectado.

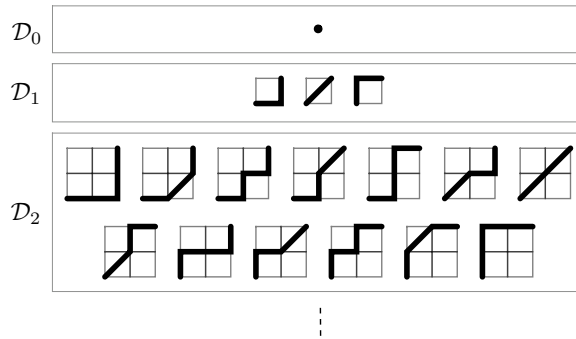


Figura 8. Los números de Delannoy centrales corresponden al número total de formas en las que se puede cruzar una retícula cuadrada desde la esquina inferior izquierda hasta la esquina superior derecha, únicamente con tres movimientos distintos: a la derecha, en diagonal (de izquierda a derecha y de abajo hacia arriba), y hacia arriba.

es la función generadora de los números de Delannoy centrales (que corresponden a *EIS A001850* en [8]). El n -ésimo número de Delannoy central es

$$D_n = \sum_{k=0}^n \binom{n}{k} \binom{n+k}{k}.$$

Los primeros números de Delannoy son como sigue:

n	0	1	2	3	4	5	6	7	8	9	...
D_n	1	3	13	63	321	1683	8989	48639	265729	1462563	...

El n -ésimo polinomio de Legendre es

$$L_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} ((x^2 - 1)^n)$$

y resuelven la ecuación diferencial de Legendre dada por

$$\frac{d}{dx} \left((1 - x^2) \frac{d}{dx} L_n(x) \right) + n(n + 1)L_n(x) = 0.$$

Los polinomios de Legendre satisfacen

$$L(x, z) = \sum_{n=0}^{\infty} L_n(x) z^n = \frac{1}{\sqrt{z^2 - 2xz + 1}}.$$

En particular $D(z) = L(3, z)$ y entonces

$$D_n = L_n(3).$$

Sea $R(z) = \sum_{n=0}^{\infty} R_n z^n$ la función generadora definida por la ecuación

$R(z) = zD(z)^3$, de forma que $R_0 = 0$ y $R_n = L'_n(3)$ para toda $n \geq 0$ (aquí $L'_n(x)$ es el coeficiente de la serie de Taylor $\frac{dL}{dx}(x, z) = \sum_{n=0}^{\infty} L'_n(x) z^n$). La función $D(z)^3$ es la función generadora de la sucesión de números que corresponde a *EIS A002695* en [8]. Los primeros términos de esta sucesión de números son como sigue:

n	1	2	3	4	5	6	7	8	9	10	...
R_n	1	9	66	450	2955	18963	119812	748548	4637205	28537245	...

4. Número de partes en disecciones aleatorias

Sea $\chi: \mathcal{S} \rightarrow \mathbb{N}$ el parámetro en \mathcal{S} que regresa el número de polígonos (o **partes**) en una disección. Las funciones de probabilidad de χ para $n = 5, 10, 15, 20$ y 25 se ilustran en la figura 9.

Teorema 4.1. *La esperanza del número de partes en una disección aleatoria de tamaño n es*

$$\mathbb{E}_{\mathcal{S}_n}(\chi) = \frac{D_n - S_n}{2S_n}$$

donde D_n es el n -ésimo número de Delannoy central y S_n es el n -ésimo número de Schröder-Hipparchus. Más precisamente,

$$\mathbb{E}_{\mathcal{S}_n}(\chi) = \frac{\sum_{k=0}^{n-1} (k+1) \sum_{i=0}^{n-1} 2^i \binom{n-1}{i} \binom{n-k-1}{i}}{\frac{1}{n+1} \sum_{k=0}^{n-1} \binom{2n-k}{n} \binom{n-1}{k}}.$$

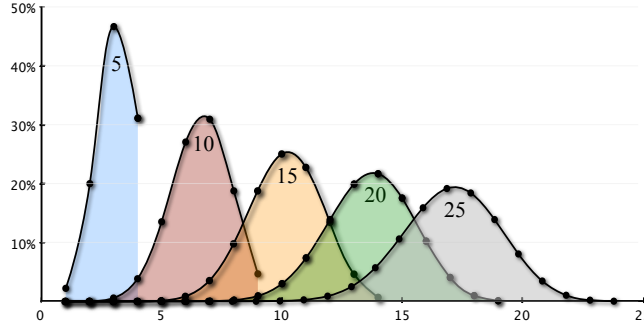


Figura 9. Funciones de probabilidad del número de partes en disecciones aleatorias de tamaño $n = 5, 10, 15, 20, 25$. Las gráficas indican que el número de partes obedece una ley de distribución límite de tipo gaussiano.

Corolario 4.2 ([4]). *La esperanza del número de partes en una disección aleatoria de tamaño n es*

$$\mathbb{E}_{S_n}(\chi) = \frac{n}{\sqrt{2}} \left(1 + \frac{2\sqrt{2}-1}{4\sqrt{2}n} + O\left(\frac{1}{n^2}\right) \right).$$

Teorema 4.3. *La varianza del número de partes en una disección aleatoria de tamaño n es*

$$\mathbb{V}_{S_n}(\chi) = \frac{R_n}{S_n} - \left(\frac{D_n - S_n}{2S_n} \right)^2$$

donde R_n es el n -ésimo coeficiente de $R(z) = zD(z)^3$ y $D(z)$ es la función generadora de los números de Delannoy centrales. Más explícitamente,

$$\mathbb{V}_{S_n}(\chi) = \frac{(n+1) \frac{d^n}{dx^n} ((x^2-1)^{n-1})|_{x=3}}{(n-1)! 2^n \sum_{k=0}^{n-1} \binom{2n-k}{n} \binom{n-1}{k}} - \mathbb{E}_{S_n}(\chi)^2.$$

Corolario 4.4 ([4]). *La varianza del número de partes en una disección aleatoria de tamaño n es*

$$\mathbb{V}_{S_n}(\chi) = \frac{\sqrt{2}n}{8} \left(1 + O\left(\frac{1}{n}\right) \right).$$

Observación 4.5. De estos corolarios podemos observar que la distribución del número de partes en una disección aleatoria de tamaño n es *concentrada* en el sentido de que para toda $\epsilon > 0$ ocurre que

$$\lim_{n \rightarrow \infty} \mathbb{P}_n \left(1 - \epsilon \leq \frac{\sqrt{2}}{n} \chi \leq 1 + \epsilon \right) = 1$$

ya que claramente

$$\lim_{n \rightarrow \infty} \frac{\sigma_{\mathcal{T}_n}(\chi)}{\mathbb{E}_{\mathcal{T}_n}(\chi)} = 0,$$

donde $\sigma_{\mathcal{T}_n}(\chi) = \sqrt{\mathbb{V}_{\mathcal{T}_n}(\chi)}$ es la desviación estándar.

4.1 Demostraciones

Demostración del Teorema 4.1. Para deducir la función generadora bivariada, podemos «marcar» las partes en alguna de las especificaciones combinatorias (9) o (12), digamos la primera. Sea μ una clase neutra y sea

$$\mathcal{S} = \mathcal{E} + \sum_{n=1}^{\infty} (\mu \times P_n) \times \text{SUC}_{n+1}(\mathcal{S}). \quad (13)$$

La función generadora bivariada

$$S(z, u) = \sum_{n, k \geq 1} S_{n, k} z^n u^k$$

tiene a u marcando el número de partes y de (13) deducimos que satisface la ecuación funcional

$$S(z, u) = 1 + \sum_{n=1}^{\infty} uz^n S(z, u)^{n+1} = 1 + \frac{uzS(z, u)^2}{1 - zS(z, u)}.$$

De aquí se sigue que

$$(u + 1)zS(z, u)^2 - (z + 1)S(z, u) + 1 = 0$$

y entonces

$$S(z, u) = \frac{z + 1 - \sqrt{z^2 - (4u + 2)z + 1}}{2(u + 1)z}.$$

Ya que

$$\partial_u S(z, u) = \frac{1}{(u + 1)} \left(\frac{1}{\sqrt{z^2 - (4u + 2)z + 1}} - \frac{z + 1 - \sqrt{z^2 - (4u + 2)z + 1}}{2(u + 1)z} \right),$$

la función generadora cumulativa es

$$\Omega(z) = \partial_u S(z, u)|_{u=1} = \frac{1}{2} \left(\frac{1}{\sqrt{z^2 - 6z + 1}} - \frac{z + 1 - \sqrt{z^2 - 6z + 1}}{4z} \right).$$

En la parte derecha de esta ecuación identificamos a las funciones generadoras de los números de Delannoy centrales $D(z)$ y de los números de Schröder-Hipparchus $S(z)$. Entonces tenemos que

$$\begin{aligned}\Omega_n &= [z^n]\Omega(z) \\ &= \frac{1}{2}(D_n - S_n) \\ &= \frac{1}{2}\left(\sum_{k=0}^n \binom{n}{k} \binom{n+k}{k} - \frac{1}{n+1} \sum_{k=0}^{n-1} \binom{2n-k}{n} \binom{n-1}{k}\right) \\ &= \sum_{k=0}^{n-1} (k+1) \sum_{i=0}^{n-1} 2^i \binom{n-1}{i} \binom{n-k-1}{i}.\end{aligned}$$

Esta sucesión de números es *EIS A035029* en [8], donde ya se describe como el número de regiones en las disecciones de polígonos. Los primeros términos de esta sucesión de números son como sigue:

n	0	1	2	3	4	5	6	7	8	9	10	...
Ω_n	0	1	5	26	138	743	4043	22180	122468	679757	3789297	...

Tenemos entonces que

$$\mathbb{E}_{S_n}(\chi) = \frac{[z^n]\Omega(z)}{[z^n]S(z)} = \frac{D_n - S_n}{2S_n}. \quad (14)$$

Para las demostraciones de los corolarios 4.2 y 4.4 usaremos el siguiente lema, el cual es una versión modificada del Teorema VI.1 en [5], pp. 381.

Lema 4.6. *Sea $\gamma > 1$ un número real y sea $\rho \in \mathbb{C} \setminus \mathbb{Z}_{\leq 0}$. Sea*

$$F(z) = ((1-z)(\gamma-z))^{-\rho}.$$

Entonces

$$\begin{aligned}F_n &:= [z^n]F(z) \\ &= \frac{n^{\rho-1}}{(\gamma-1)^\rho \Gamma(\rho)} \left(1 + \frac{(\rho-1)(\rho-2) - 2(\rho(\gamma-1)^{-1} - 1)(\rho-1)}{2n}\right) \\ &\quad + O\left(\frac{1}{n^2}\right).\end{aligned}$$

Demostración. Usaremos la fórmula integral de Cauchy para derivadas y contornos de Hankel. Tenemos que

$$[z^n]F(z) = \frac{1}{2\pi} \int_{\mathcal{H}(n)} ((1-z)(\gamma-z))^{-\rho} \frac{dz}{z^{n+1}}$$

con $\mathcal{H}(n) = \mathcal{H}^-(n) \cup \mathcal{H}^+(n) \cup \mathcal{H}^\circ(n)$, donde

$$\begin{cases} \mathcal{H}^-(n) &= \{z = w - \frac{i}{n}, w \geq 1\} \\ \mathcal{H}^+(n) &= \{z = w + \frac{i}{n}, w \geq 1\} \\ \mathcal{H}^\circ(n) &= \{z = 1 - \frac{e^{i\phi}}{n}, \phi \in [-\frac{\pi}{2}, \frac{\pi}{2}] \geq 1\}. \end{cases}$$

El cambio de variable

$$z = 1 + \frac{t}{n}$$

implica

$$F_n = \frac{n^{\rho-1}}{2\pi i} \int_{\mathcal{H}} (-t)^{-\rho} \left(\gamma - 1 - \frac{t}{n}\right)^{-\rho} \left(1 + \frac{t}{n}\right)^{-n-1} dt$$

donde $\mathcal{H} = \mathcal{H}(1) - 1$. Recordemos la expansión asintótica

$$\begin{aligned} \left(1 + \frac{t}{n}\right)^{-n-1} &= e^{-(n+1)\log(1+t/n)} \\ &= e^{-t} \left(1 + \frac{t^2 - 2t}{2n} + \frac{3t^4 - 20t^3 + 24t^2}{24n^2} + \dots\right) \end{aligned}$$

y también, por el teorema binomial generalizado de Newton, sabemos que

$$\begin{aligned} \left(\gamma - 1 - \frac{t}{n}\right)^{-\rho} &= (\gamma - 1)^{-\rho} + \rho(\gamma - 1)^{-\rho-1} \left(\frac{t}{n}\right) + \dots \\ &= (\gamma - 1)^{-\rho} \left(1 + \rho(\gamma - 1)^{-1} \left(\frac{t}{n}\right) + \dots\right). \end{aligned}$$

Entonces

$$\begin{aligned} \left(\gamma - 1 - \frac{t}{n}\right)^{-\rho} \left(1 + \frac{t}{n}\right)^{-n-1} &= \frac{e^{-t}}{(\gamma - 1)^\rho} \left(1 + \frac{t^2 + 2(\rho(\gamma - 1)^{-1} - 1)t}{2n} + \dots\right) \end{aligned}$$

y así

$$\begin{aligned}
F_n &= \frac{n^{\rho-1}}{(\gamma-1)^\rho} \frac{1}{2\pi i} \int_{\mathcal{H}} (-t)^{-\rho} e^{-t} \\
&\quad \left(1 + \frac{t^2 + 2(\rho(\gamma-1)^{-1} - 1)t}{2n} + \dots\right) dt \\
&= \frac{n^{\rho-1}}{(\gamma-1)^\rho \Gamma(\rho)} \left(1 + \frac{(\rho-1)(\rho-2) - 2(\rho(\gamma-1)^{-1} - 1)(\rho-1)}{2n}\right. \\
&\quad \left.+ O\left(\frac{1}{n^2}\right)\right). \quad \square
\end{aligned}$$

Corolario 4.7. *Sea $\gamma > 1$. Entonces*

1. $[z^n] \sqrt{(1-z)(\gamma-z)} = -\frac{\sqrt{\gamma-1}}{2\sqrt{\pi n^3}} \left(1 + \frac{3-6(\gamma-1)^{-1}}{8n}\right. \\ \left.+ O\left(\frac{1}{n^2}\right)\right).$
2. $[z^n] \frac{1}{\sqrt{(1-z)(\gamma-z)}} = \frac{1}{\sqrt{(\gamma-1)\pi n}} \left(1 + \frac{2(\gamma-1)^{-1} - 1}{8n}\right. \\ \left.+ O\left(\frac{1}{n^2}\right)\right).$
3. $[z^n] \frac{1}{\sqrt{(1-z)^3(\gamma-z)^3}} = \frac{2\sqrt{n}}{\sqrt{(\gamma-1)^3\pi}} \left(1 + \frac{3-6(\gamma-1)^{-1}}{8n}\right. \\ \left.+ O\left(\frac{1}{n^2}\right)\right).$

Demostación del Corolario 4.2. Comencemos observando que las singularidades de $S(x)$ son $\alpha = 3 - 2\sqrt{2}$, la cual es dominante, y $\beta = 3 + 2\sqrt{2}$. Recordemos el rescalamiento

$$[z^n]f(z) = \zeta^{-n}[z^n]f(z\zeta) = \zeta^{-n}[z^n]g(z)$$

(aquí $g(z) = f(z\zeta)$). Sea $\gamma := \beta/\alpha > 1$. Entonces el n -ésimo coeficiente de $S(z)$ es

$$\begin{aligned}
[z^n]S(z) &= [z^n] \left(\frac{z + 1 - \sqrt{(\alpha - z)(\beta - z)}}{4z} \right) \\
&= \frac{1}{4}[z^{n+1}] \left(z + 1 - \sqrt{(\alpha - z)(\beta - z)} \right) \\
&= -\frac{1}{4}[z^{n+1}] \left(\sqrt{(\alpha - z)(\beta - z)} \right) \quad \forall n \geq 2 \\
&= -\frac{\alpha^{-(n+1)}}{4}[z^{n+1}] \left(\sqrt{(\alpha - \alpha z)(\beta - \alpha z)} \right) \\
&= -\frac{\alpha^{-n}}{4}[z^{n+1}] \left(\sqrt{(1 - z)(\gamma - z)} \right) \\
&= \frac{\sqrt{\gamma - 1}}{8\sqrt{\pi n^3}} \alpha^{-n} \left(1 + \frac{3 - 6(\gamma - 1)^{-1} - 12}{8n} + O\left(\frac{1}{n^2}\right) \right) \\
&= \frac{\sqrt{\sqrt{2}}}{4\sqrt{\alpha\pi n^3}} \alpha^{-n} \left(1 - \frac{12\sqrt{2} + 9}{16\sqrt{2}n} + O\left(\frac{1}{n^2}\right) \right). \tag{15}
\end{aligned}$$

Similarmente,

$$\begin{aligned}
[z^n]D(z) &= [z^n] \frac{1}{\sqrt{1 - 6z + z^2}} \\
&= [z^n] \frac{1}{\sqrt{(\alpha - z)(\beta - z)}} \\
&= \alpha^{-n}[z^n] \frac{1}{\sqrt{(\alpha - \alpha z)(\beta - \alpha z)}} \\
&= \alpha^{-n-1}[z^n] \frac{1}{\sqrt{(1 - z)(\gamma - z)}} \\
&= \frac{1}{\alpha\sqrt{(\gamma - 1)\pi n}} \alpha^{-n} \left(1 + \frac{2(\gamma - 1)^{-1} - 1}{8n} + O\left(\frac{1}{n^2}\right) \right) \\
&= \frac{1}{2\sqrt{\sqrt{2}\alpha\pi n}} \alpha^{-n} \left(1 + \frac{3 - 4\sqrt{2}}{16\sqrt{2}n} + O\left(\frac{1}{n^2}\right) \right).
\end{aligned}$$

Por lo tanto el n -ésimo coeficiente de $\Omega(z)$ es

$$\begin{aligned}
[z^n]\Omega(z) &= \frac{1}{2} ([z^n]D(z) - [z^n]S(z)) \\
&= \frac{1}{2} \left[\frac{1}{2\sqrt{\sqrt{2}\alpha\pi n}} \alpha^{-n} \left(1 + \frac{3-4\sqrt{2}}{16\sqrt{2}n} + O\left(\frac{1}{n^2}\right) \right) \right. \\
&\quad \left. - \frac{\sqrt{\sqrt{2}}}{4\sqrt{\alpha\pi n^3}} \alpha^{-n} \left(1 - \frac{12\sqrt{2}+9}{16\sqrt{2}n} + O\left(\frac{1}{n^2}\right) \right) \right] \\
&= \frac{1}{4\sqrt{\sqrt{2}\alpha\pi n}} \alpha^{-n} \left(1 - \frac{13+4\sqrt{2}}{16\sqrt{2}n} + O\left(\frac{1}{n^2}\right) \right).
\end{aligned}$$

Sustituimos en la ecuación (14) para obtener

$$\begin{aligned}
\mathbb{E}_{\mathcal{S}_n}(\chi) &= \frac{\frac{1}{4\sqrt{\sqrt{2}\alpha\pi n}} \alpha^{-n} \left(1 - \frac{13+4\sqrt{2}}{16\sqrt{2}n} + O\left(\frac{1}{n^2}\right) \right)}{\frac{\sqrt{\sqrt{2}}}{4\sqrt{\alpha\pi n^3}} \alpha^{-n} \left(1 - \frac{12\sqrt{2}+9}{16\sqrt{2}n} + O\left(\frac{1}{n^2}\right) \right)} \\
&= \frac{n}{\sqrt{2}} \left(1 + \frac{2\sqrt{2}-1}{4\sqrt{2}n} + O\left(\frac{1}{n^2}\right) \right). \quad \square
\end{aligned}$$

Demostración del Teorema 4.3. Para el segundo momento primero obtenemos

$$\begin{aligned}
\partial_u^2 S(z, u) &= \frac{1}{(u+1)} \left(\frac{2z}{\sqrt{(z^2 - (4u+2)z + 1)^3}} \right. \\
&\quad \left. - \frac{2}{(u+1)\sqrt{z^2 - (4u+2)z + 1}} \right. \\
&\quad \left. + \frac{z+1 - \sqrt{z^2 - (4u+2)z + 1}}{(u+1)^2 z} \right)
\end{aligned}$$

de forma que

$$\partial_u^2 S(z, u)|_{u=1} = zD(z)^3 - \Omega(z) = R(z) - \Omega(z).$$

La sucesión de número asociada a esta función generadora no tiene registro en [8] y sus primeros términos son como sigue:

n	1	2	3	4	5	6	7	8	9	10	...
$R_n - \Omega_n$	0	4	40	312	2212	14920	97632	626080	3957448	24747948	...

(16)

Tenemos que

$$[z^n]\partial_u^2 S(z, u)|_{u=1} = [z^n]R(z) - \frac{D_n - S_n}{2}$$

y entonces el segundo momento está dado por

$$\mathbb{E}_{\mathcal{S}_n}(\chi^2) = \frac{[z^n]\partial_u^2 S(z, u)|_{u=1}}{[z^n]S(z)} + \mathbb{E}_{\mathcal{S}_n}(\chi) = \frac{R_n}{S_n}$$

y por lo tanto la varianza es

$$\mathbb{V}_{\mathcal{S}_n}(\chi) = \sigma_{\mathcal{S}_n}(\chi)^2 = \frac{R_n}{S_n} - \left(\frac{D_n - S_n}{2S_n}\right)^2. \quad \square$$

Demostración del Corolario 4.4. El n -ésimo coeficiente de $R(z)$ es

$$\begin{aligned} [z^n]R(z) &= [z^n]\frac{z}{\sqrt{(1-6z+z^2)^3}} \\ &= [z^{n-1}]\frac{1}{\sqrt{(\alpha-z)^3(\beta-z)^3}} \\ &= \alpha^{-n+1}[z^{n-1}]\frac{1}{\sqrt{(\alpha-\alpha z)^3(\beta-\alpha z)^3}} \\ &= \alpha^{-n-2}[z^{n-1}]\frac{1}{\sqrt{(1-z)^3(\gamma-z)^3}} \\ &= \frac{2\sqrt{n}}{\alpha^2\sqrt{(\gamma-1)^3}\pi}\alpha^{-n}\left(1 + \frac{3-6(\gamma-1)^{-1}-4}{8n} + O\left(\frac{1}{n^2}\right)\right) \\ &= \frac{\sqrt{n}}{4\sqrt{2}\sqrt{2}\alpha\pi}\alpha^{-n}\left(1 + \frac{4\sqrt{2}-9}{16\sqrt{2}n} + O\left(\frac{1}{n^2}\right)\right). \end{aligned}$$

Entonces

$$\begin{aligned} \frac{R_n}{S_n} &= \frac{\frac{\sqrt{n}}{4\sqrt{2}\sqrt{2}\alpha\pi}\alpha^{-n}\left(1 + \frac{4\sqrt{2}-9}{16\sqrt{2}n} + O\left(\frac{1}{n^2}\right)\right)}{\frac{\sqrt{\sqrt{2}}}{4\sqrt{\alpha\pi n^3}}\alpha^{-n}\left(1 - \frac{12\sqrt{2}+9}{16\sqrt{2}n} + O\left(\frac{1}{n^2}\right)\right)} \\ &= \frac{n^2}{2}\left(1 + \frac{1}{n} + O\left(\frac{1}{n^2}\right)\right) \end{aligned}$$

Tenemos que

$$\mathbb{E}_{\mathcal{S}_n}(\chi)^2 = \left(\frac{D_n - S_n}{2S_n} \right)^2 = \frac{n^2}{2} \left(1 + \frac{2\sqrt{2}-1}{2\sqrt{2}n} + O\left(\frac{1}{n^2}\right) \right).$$

Por lo tanto

$$\begin{aligned} \mathbb{V}_{\mathcal{S}_n}(\chi) &= \frac{R_n}{S_n} - \left(\frac{D_n - S_n}{2S_n} \right)^2 \\ &= \frac{n^2}{2} \left(1 + \frac{1}{n} + O\left(\frac{1}{n^2}\right) \right) - \frac{n^2}{2} \left(1 + \frac{2\sqrt{2}-1}{2\sqrt{2}n} + O\left(\frac{1}{n^2}\right) \right) \\ &= \frac{1}{4\sqrt{2}}n + O(1). \quad \square \end{aligned}$$

5. Conclusión

Motivados por la estructura de las moléculas de ARN, hemos presentado una breve introducción a la combinatoria analítica, estudiando particularmente las disecciones aleatorias de polígonos convexos. Para realmente conectar la motivación con los resultados descritos en la sección §4 sería recomendable ver un ejemplo concreto. Consideremos pues el caso de la estructura secundaria del 16S rARN de *Escherichia coli*, el cual se ilustra en la figura 10. Esta cadena de ARN ribosomal tiene una longitud de $n = 1542$ bases. Los corolarios 4.2 y 4.4 indican que

$$\mathbb{E}_{\mathcal{S}_{1542}}(\chi) \approx \frac{1542}{\sqrt{2}} \approx 1090 \quad \text{y} \quad \sigma_{\mathcal{S}_{1542}}(\chi) \approx \sqrt{\frac{1542\sqrt{2}}{8}} \approx 16.$$

Suponiendo que la estructura primaria es Bernoulli $\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$ y debido a los apareamientos de Watson y Crick (A-U y G-C), es natural considerar el parámetro $\frac{1}{2}\chi$, el cual valdría $\approx 545 \pm 4$. Una inspección exhaustiva (y tediosa) de la ilustración muestra que esta estructura secundaria de ARN ribosomal posee (aproximadamente) 550 partes⁷.

Agradecimientos. Agradecemos profundamente a los árbitros por sus valiosas observaciones y recomendaciones, las cuales mejoraron mucho el contenido, forma y presentación de este trabajo. Agradecemos en particular todo el apoyo y atención de la Coordinadora Editorial de Miscelanea Matemática, la Dra. Ana Meda Guardiola.

⁷De hecho, si observamos cuidadosamente la ilustración, podemos ver que los apareamientos de las bases 9-13 con 21-25 forman un «pseudonudo», es decir, generan cruces, como en el caso (b) de la figura 2. Este pseudonudo produce 5 partes que podemos descartar de nuestra aproximación.

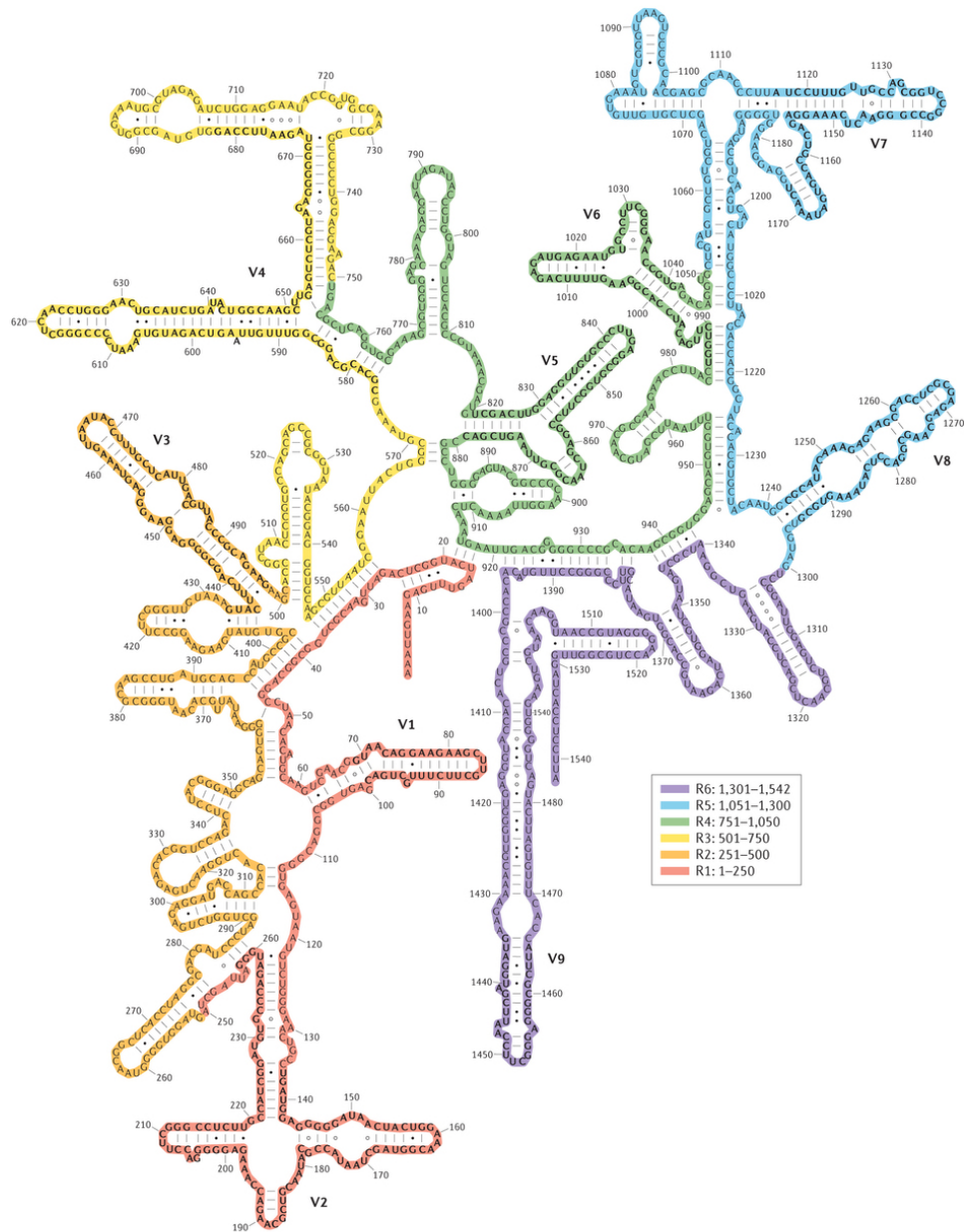


Figura 10. Estructura secundaria del 16S rARN de *Escherichia coli* generada por el programa xrna. Reimpreso bajo permiso de Macmillan Publishers Ltd, Nature Reviews Microbiology, Pablo Yarza et. al. «Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences». **12**, 635-645 (2014).

Bibliografía

- [1] A. Asinowski y T. Mansour, «Dick paths and coloured ascents», *European J. Combin.*, vol. 29, núm. 5, 2008, 1262–1279.
- [2] E. Catalan, «Addition à la note sur une équation aux différences finies, insérée dans le volume précédent, page 508», *Journal de Mathématique Pures et Appliquées*, vol. 4, 1939, 95–99.
- [3] A. Ehrenfeucht, T. Harju, P. ten Pas y G. Rozenberg, «Permutations, parenthesis words, and Schröder numbers», *Discrete Mathematics*, vol. 190, núm. 1-3, 1998, 259–264.
- [4] P. Flajolet y M. Noy, «Analytic combinatorics of non-crossing configurations», *Discrete Mathematics*, vol. 204, 1999, 203–229.
- [5] P. Flajolet y R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, 2009.
- [6] C. Raidys, *Combinatorial computational biology of RNA*, Springer Verlag, 2011.
- [7] L. Shapiro y R. Sulanke, «Bijections for the Schröder numbers», *Math. Mag.*, vol. 73, núm. 5, 2000, 369–376.
- [8] N. Sloane, «The On-Line Encyclopedia of Integer Sequences», 2006, Published electronically at <http://www.research.att.com/~njas/sequences/>.
- [9] R. Sulanke, «Objects counted by the central Delannoy numbers», *J. Integer Seq.*, vol. 6, núm. 1, 2003, 1–19, Article 03.1.5.