Física Teórica 3 — 1er. cuatrimestre de 2019

Extra Guía 2: modelo de Ehrenfest

■ El problema original de la Guía 2 dice lo siguiente (adaptado a la notación que usamos en clase):

> 14. **El modelo de Ehrenfest**. Para ilustrar la tendencia al equilibrio, Tatiana y Paul Ehrenfest propusieron el siguiente modelo: N fichas, numeradas de 1 a N, se distribuyen en dos urnas. En cada paso una ficha se elige al azar y se la cambia de urna. La variable aleatoria que consideraremos es el número s de fichas en la primera urna, y llamaremos $p_i(s)$ a la probabilidad de que luego de i pasos esta urna contenga s fichas.
>
> a) Muestre que el proceso así definido es un proceso de Markov. Para eso escriba la ecuación de evolución para la probabilidad, es decir, una ecuación que dé $p_{i+1}(s)$ en términos de las probabilidades en los pasos anteriores.
>
> b) Mostrar que la solución estacionaria, $p(s)$, es una binomial. *Sugerencia:* la condición de distribución estacionaria quedará en la forma de una ecuación de recurrencia para $p(s)$. Esto puede resolverse con el método de la **función generatriz**, definiendo la función auxiliar $F(x) = \sum x^s p(s)$ y transformando la ecuación de recurrencia para p en una ecuación diferencial para F, sujeta a la condición $F(1) = 1$. Los coeficientes del desarrollo de $F(x)$ en potencias de x son las probabilidades $p(s)$.

En la clase de práctica quisimos mostrar otras propiedades del modelo, pero nos faltó tiempo. Continuamos aquí con una extensión comentada del problema. Para simplificar los cálculos vamos a introducir los siguientes cambios:

- Hay un número par de fichas igual a $2N$. De esta manera la distribución de equilibrio es

$$p(s) = \frac{1}{2^{2N}}\binom{2N}{s}. \tag{1}$$

Notar que esto puede escribirse como un proceso de Bernoulli con probabilidad de éxito $q = 1/2$.

$$p(s) = \binom{2N}{s}\left(\frac{1}{2}\right)^s\left(1-\frac{1}{2}\right)^{2N-s}, \tag{2}$$

lo que muestra que el valor medio y la dispersión de s son, respectivamente,

$$\begin{aligned} \langle s\rangle &= 2Nq = N, \\ \sigma^2 &= \langle s^2\rangle - \langle s\rangle^2 = 2Nq(1-q) = \frac{N}{2}. \end{aligned} \tag{3}$$

Tenemos así una distribución binomial centrada en N y con una desviación estándar $\sigma = \sqrt{N/2}$. El ancho de la campana sobre la que se ubican los valores de $p(s)$ es de orden $\sqrt{N}$. Cuanto más grande sea N, más pequeño será el ancho de la campana en relación al valor medio, es decir, la distribución será más aguda.

- Siempre en referencia a la primera urna: en lugar de contar su número s de fichas, contaremos el número de fichas en exceso que tiene respecto del valor más probable. Es decir, introduciremos la variable aleatoria $m = s - N$. Notar que m puede tomar valores enteros entre $-N$ y N, y que su valor más probable en el equilibrio es $m = 0$. Notar, además, que con esta definición la primera urna tiene $N + m$ fichas y la segunda, $N - m$. El reparto de fichas entre las dos urnas es parejo cuando $m = 0$. En el equilibrio, la probabilidad en términos de m adquiere una forma simétrica respecto de $m = 0$:

$$p(m) = \frac{1}{2^{2N}} \binom{2N}{N+m} = \frac{1}{2^{2N}} \frac{(2N)!}{(N+m)!\,(N-m)!}. \tag{4}$$

La interpretación de la distribución de equilibrio

Una pregunta que surgió en clase es acerca de cómo interpretar la distribución de equilibrio. Teniendo en cuenta que el sistema siempre está cambiando su composición, ¿qué significa que su distribución de probabilidad es la del equilibrio? El significado de que $p(m)$ sea la distribución de equilibrio es que si saben que uno de estos sistemas fue preparado inicialmente asignándole a la primera urna un número de fichas $N + m'$ (desconocido para ustedes) con una probabilidad $p(m')$, y observan al sistema un tiempo después, la probabilidad de que encuentren un número $N + m$ de fichas en la primera urna seguirá siendo $p(m)$.

Una manera alternativa de pensarlo es a través de un ensamble de sistemas. Tienen inicialmente un número $\nu \gg 1$ de sistemas idénticos. En cada uno de estos sistemas colocan cierto número de fichas en la primera urna, con la condición de que la fracción de sistemas en donde colocan $N + m$ fichas sea igual a $p(m)$. Un tiempo después se fijan cuántas fichas hay en la primera urna de cada sistema. Que la distribución $p(m)$ sea la distribución de equilibrio significa que en el límite de $\nu \to \infty$ la fracción de sistemas que tienen $N + m$ fichas en la primera urna seguirá siendo igual a $p(m)$, y esto siempre continuará siendo así. Cada sistema es dinámico. En cada sistema pasan fichas de una urna a la otra, pero el ensamble de sistemas siempre estará compuesto por una población estable de sistemas con un dado número de fichas en la primera urna. El ensamble no cambia cuando su distribución inicial está dada por $p(m)$. Puede ayudar pensar esto en términos de una población que ha alcanzado el equilibrio en la distribución de las edades de sus habitantes. La proporción de habitantes con determinada edad es siempre la misma, pero los habitantes que forman cada grupo etario cambian con el tiempo.

Aproximación por una gaussiana

■ El primer punto que les sugerimos demostrar ahora es que para $N \gg 1$ y $m \gg 1$ la distribución de probabilidad puede aproximarse por la siguiente gaussiana:

$$p(m) \approx \frac{1}{\sqrt{\pi N}}\, e^{-m^2/N}. \tag{5}$$

Deberán emplear la aproximación de Stirling para los factoriales, $\log z! \approx z\log z - z$, y normalizar la distribución obtenida para que sume 1 (como en el problema 10, pero más fácil). La distribución (5) tiene media cero y dispersión $\sigma^2 = N/2$.

Cuando $N \gg 1$, visualmente resulta más informativo trabajar con la fracción $x = m/N$, que toma valores

$$-1,\ -1+\frac{1}{N},\ \ldots,\ \ldots,\ -\frac{2}{N},\ -\frac{1}{N},\ 0,\ \frac{1}{N},\ \frac{2}{N},\ \ldots,\ \ldots,\ 1-\frac{1}{N},\ 1,$$

es decir, los múltiplos enteros de $1/N$ en el intervalo $[-1, 1]$. Si consideramos a x como una variable discreta, su distribución de probabilidad se obtiene a partir de la de m escribiendo $m = Nx$, tanto en la expresión exacta como en la aproximada:

$$p(x) = \frac{1}{2^{2N}}\binom{2N}{(1+x)N} \approx \frac{1}{\sqrt{\pi N}}\, e^{-x^2/N^{-1}}. \tag{6}$$

La última expresión es un tanto incómoda. Tiene la apariencia de una gaussiana pero no es una gaussiana. En el exponente aparece N^{-1}, pero dentro de la raíz aparece N. Antes podíamos escribir

$$\sum_{m=-N}^{m=N} p(m) \approx \frac{1}{\sqrt{\pi N}}\int_{-\infty}^{\infty} dm\, e^{-m^2/N} = 1. \tag{7}$$

En cambio, una fórmula análoga para x no es tan inmediata, porque el intervalo entre valores sucesivos de x no es 1, sino $\Delta x = 1/N$. Detonando... no; denotando al conjunto de valores que puede tomar x por $\{x_i\}$, ahora la condición de normalización se lee como

$$\begin{aligned}\sum_{x_i} p(x_i) &= \frac{1}{\Delta x}\sum_{x_i} p(x_i)\Delta x \\ &\approx \frac{1}{\Delta x}\int_{-\infty}^{\infty} dx\, \frac{1}{\sqrt{\pi N}}\, e^{-x^2/N^{-1}} = \frac{1}{\sqrt{\pi N^{-1}}}\int_{-\infty}^{\infty} dx\ e^{-x^2/N^{-1}} = 1.\end{aligned} \tag{8}$$

Recién en la última expresión recuperamos la función gaussiana correctamente normalizada, con N^{-1} en el exponente y N^{-1} en la raíz. Esto da origen al siguiente comentario.

Con frecuencia en mecánica estadística debemos hacer sumas del tipo $\sum_i f(x_i)$. Muchas veces los puntos x_i están densamente distribuidos en cierto intervalo, o, más precisamente, la separación entre los puntos es mucho menor que la longitud característica de variación de $f(x)$. En el ejemplo que acabamos de ver, teníamos los puntos x_i apretados en el intervalo $[-1, 1]$, a distancia $1/N$ uno del otro. A medida que N aumenta, los puntos x_i cubren el intervalo de manera cada vez más densa. A los efectos prácticos, dependiendo de la precisión con la que uno mida los números de fichas, llega una instancia en la que es como si x pudiera tomar cualquier valor dentro del intervalo. Llegado ese momento, es más cómodo tratar a x como una variable continua en vez de discreta. La Ec. (8) muestra como aproximar sumas por integrales. En el párrafo que sigue veremos cómo transformar la distribución de probabilidad discreta para los x_i, en una densidad de probabilidad para la variable continua x.

Recuerden que para distribuciones continuas, la función de densidad de probabilidad $f(x)$ se define como

$$p(X \leqslant x \leqslant X + dX) \equiv f(X)dX, \tag{9}$$

Nuestro objetivo es tratar de llegar a una expresión similar a la anterior a partir de la distribución discreta.

Teniendo a la vista el primer miembro de la Ec. (9), empezaremos por escribir la probabilidad de que los x_i discretos estén en cierto intervalo. Puesto que $x = m/N$, la probabilidad de que x tome valores en el intervalo entre X y $X + \delta X$ es igual a la probabilidad de que m tome valores entre XN y $XN + N\delta X$,

$$p\left(X \leqslant x \leqslant X + \delta X\right) = \sum_{m=XN}^{XN+N\delta X} p(m). \tag{10}$$

El intervalo δX debe ser lo suficientemente grande como para contener un número de puntos x_i mucho mayor que 1, pero lo suficientemente chico como para que $p(m)$ varíe poco en el intervalo $N\delta X$. La primera condición asegura que al variar X la probabilidad varíe de manera suave, sin saltos súbitos debido a la entrada o salida de puntos al intervalo $[X, X + \delta X]$. La segunda condición hace que todos los términos en la suma (10) tengan esencialmente el mismo valor, $p(XN)$. Las dos condiciones pedidas se satisfacen si tomamos

$$\frac{1}{N} \ll \delta X \ll \frac{1}{\sqrt{N}}, \tag{11}$$

ya que $1/N$ es el espacio entre los puntos x_i y $\sqrt{N}$ es la longitud típica de variación de la función $p(m)$. Puesto que hay $N\delta X$ términos en la suma (10) y todos toman aproximadamente el valor $p(XN)$, obtenemos

$$p\left(X \leqslant x \leqslant X + \delta X\right) \approx p(XN)\, N\delta X. \tag{12}$$

Si hacemos la identificación $f(X) = p(XN)N$, recuperamos la forma de la Ec. (9). De este modo, la variable continua que aproxima la distribución de los valores discretos x_i tiene una densidad de probabilidad

$$f(x) = Np(Nx) = \frac{1}{\sqrt{\pi N^{-1}}}\, e^{-x^2/N^{-1}}. \tag{13}$$

Es una gaussiana centrada en cero y con una dispersión $\sigma^2 = 1/2N$. La ventaja de esta representación, respecto de la Ec. (5), es que ahora es evidente que la distribución se hace más aguda conforme N aumenta. Al emplear la forma dada por la Ec. (5), lo que disminuye con N no es el ancho de la campana sino el ancho relativo respecto del valor máximo que puede tomar m, que es N.

Reversibilidad

■ El segundo punto que les sugerimos que analicen es más interesante, porque está relacionado con la reversibilidad del proceso aleatorio. Muestra cómo un proceso que, en sus pasos elementales, no diferencia entre una y otra dirección del tiempo, puede sin embargo exhibir una tendencia decidida hacia la media, lo que parecería señalar, después de todo, una dirección de evolución privilegiada. Esto que acabamos de escribir, recuérdenlo cuando en la teórica analicen las paradojas asociadas al teorema H de Boltzmann.

En definitiva, las cosas se disponen del siguiente modo: a tiempo $i = 0$ se prepara uno de estos sistemas según la distribución de equilibrio $p(m)$. Esto quiere decir que a la primera urna se le asigna inicialmente un número de fichas $N + m$ con una probabilidad $p(m)$. Para $N \gg 1$, cuando observen el sistema en el futuro, lo más probable es que lo encuentren con un número parejo de fichas en cada urna, $x \approx 0 \pm 1/\sqrt{N}$, donde $x = m/N$.

Supongan que un tiempo después de preparado el sistema observan tres pasos sucesivos de su evolución, digamos, los pasos $i-1$, i y $1+i$. Consideremos la probabilidad de que en esos tres pasos sucesivos vean al sistema alejarse desde los estados más probables hacia la zona de los menos probables. Suponiendo que $m > 0$, la probabilidad de que vean al sistema alejándose de la configuración más probable está dada por la probabilidad conjunta

$$p\big(m-1, i-1\ ;\ m, i\ ;\ m+1, i+1\big) \equiv p_{m-1\to m\to m+1}, \tag{14}$$

que se lee así: la probabilidad de que en el paso $i-1$ haya $N+m-1$ fichas en la primera urna, de que en el paso i haya $N+m$ y de que en el paso $i+1$ haya $N+m+1$. En cada paso, una ficha más, y cada vez más lejos del estado con $m = 0$. Asimismo, la probabilidad de que vean al sistema acercándose a la configuración más probable desde una configuración menos probable, asumiendo $m > 0$, está dada por

$$p\big(m+1, i-1\ ;\ m, i\ ;\ m-1, i+1\big) \equiv p_{m+1\to m\to m-1}. \tag{15}$$

Ahora el número de fichas de la primera urna pasa de $N+m+1$ a $N+m$ y por último a $N+m-1$. En cada paso, una ficha menos, y cada vez más cerca del estado con $m = 0$.

En clase alcanzamos a calcular la probabilidad de la primera serie de eventos. Ahora repetimos ese cálculo. Escribiendo la probabilidad conjunta como un encadenamiento de probabilidades condicionales, resulta

$$\begin{aligned} p_{m-1\to m\to m+1} &= p\big(m+1, i+1 \mid m, i\ ;\ m-1, i-1\big)\, p\big(m, i \mid m-1, i-1\big)\, p(m-1, i-1) \\ &= p\big(m+1, i+1 \mid m, i\big)\, p\big(m, i \mid m-1, i-1\big)\, p(m-1). \end{aligned} \tag{16}$$

Para entender su significado tienen que leer estas ecuaciones en voz alta, si no, no vale de nada. La última igualdad tiene en cuenta dos cosas. Primero, tiene en cuenta que el proceso es de tipo markoviano:

$$p\big(m+1, i+1 \mid m, i\ ;\ m-1, i-1\big) = p\big(m+1, i+1 \mid m, i\big). \tag{17}$$

La probabilidad de un evento a tiempo $i+1$, sujeto a la condición de que haya ocurrido una serie de eventos a tiempos i, $i-1$, $i-2$, etc., sólo depende del evento inmediatamente anterior. Es decir, la probabilidad de que la primera urna reciba o pierda una ficha en el paso $i+1$ sólo depende de cuántas fichas tiene la urna en el paso i, y no de cuántas tuvo en tiempos anteriores a i. La segunda cosa tenida en cuenta al escribir la Ec. (16) es que $p(m-1, i-1) = p(m-1)$, ya que, por hipótesis, el sistema fue preparado según la distribución de equilibrio, y esa distribución no depende del tiempo.

Finalmente, a partir de la Ec. (16), la probabilidad de ver que el sistema se aleja hacia estados menos probables es

$$\begin{aligned} p_{m-1\to m\to m+1} &= \left(\frac{N-m}{N}\right)\left(\frac{N-m+1}{N}\right)\frac{(2N)!}{2^{2N}(N+m-1)!\,(N-m+1)!} \\ &= \frac{(2N)!}{2^{2N}N^2}\,\frac{1}{(N+m-1)!\,(N-m-1)!}. \end{aligned} \tag{18}$$

Queda para ustedes demostrar que la probabilidad de ver que el sistema se acerca a la región de estados más probables, empezando desde $m+1$, pasando por m y llegando a $m-1$, da exactamente lo mismo. Es decir, comprueben que

$$p_{m+1\to m\to m-1} = p_{m-1\to m\to m+1}. \tag{19}$$

Los resultados anteriores muestran la **reversibilidad** del proceso en el equilibrio. Es tan probable verlo ir en una dirección como verlo ir en la otra. Ahora bien, si es tan probable verlo acercarse como verlo alejarse respecto de la configuración más probable, podrían preguntarse cómo es posible que el sistema esté casi todo el tiempo en un entorno de la configuración más probable.

Uno puede demostrar resultados mucho más alarmantes. Por ejemplo, las siguientes probabilidades

$$P(m=0,0\,;\ m=N,t) \quad y \quad P(m=N,0\,;\ m=0,t)$$

son iguales en el equilibrio: la probabilidad de ver que la segunda urna se vacía es igual a la probabilidad de verla llenarse desde cero. Recuerden que $m=0$ significa un reparto parejo entre las dos urnas, y $m=N$ significa que todas las fichas están en la primera urna.

Si piensan en las urnas como si fueran la división de una habitación en dos volúmenes iguales, y a las fichas como si fueran las moléculas de aire, estamos diciendo que la probabilidad de ver que todas las moléculas pasan a amontonase en la mitad izquierda de la habitación es igual a la probabilidad verlas pasar de estar todas amontonadas en la mitad izquierda a ocupar uniformemente la habitación. Debido a que, intuitivamente, la segunda probabilidad parecería ser muy cercana a 1, estamos frente a la contradicción de que nunca vemos ocurrir el proceso por el cual todas las moléculas se amontonan en una mitad, que debería tener también una probabilidad cercana a 1.

Antes que nada queremos llevar tranquilidad a la población. El error en el argumento expuesto más arriba está en suponer que la probabilidad de ver llenarse media habitación desde cero es cercana a 1. En verdad las dos probabilidades consideradas en el párrafo anterior son igualmente insignificantes si $N \gg 1$. Cada una de las probabilidades consideradas demanda que, o bien inicialmente, o bien en la situación final, una de las urnas esté vacía. La probabilidad de que esto ocurra es tan pequeña que casi no importa con qué otro suceso la asociemos: cualquier probabilidad conjunta que incluya entre sus términos que una de las urnas esté vacía es *a priori* despreciable.

Obtuvimos una respuesta que nos pareció contradictoria porque tal vez no estábamos contestando la pregunta que queríamos responder. Hay dos clases de preguntas y cada una tiene su propia respuesta. El error está en tomar una pregunta por otra o una respuesta por otra. Lo que nos preguntamos, quizás, no era lo que teníamos en mente, por eso la respuesta nos desconcertó. De todas maneras llegamos a demostrar una propiedad fundamental del modelo en el equilibrio: la simetría respecto a la dirección temporal.

Las otras preguntas que pudimos habernos hecho son: ¿cuál es la probabilidad de que todas las fichas se amontonen en una sola urna *dado* que inicialmente estaban distribuidas uniformemente?, y ¿cuál es la probabilidad de que las fichas se distribuyan uniformemente *dado* que inicialmente una de las urnas estaba vacía? Dicho de otra manera, cómo se comparan las dos probabilidades condicionales

$$P\big(m = N, j \mid m = 0, 0\big) \quad y \quad P\big(m = 0, j \mid m = N, 0\big).$$

Calculen ustedes la relación entre estas dos probabilidades y verifiquen que el primer evento tiene una probabilidad despreciable frente al segundo, cuando $p(m)$ es la distribución de equilibrio y $N \gg 1$. Este es el resultado intuitivo. Es mucho más probable que el estado en el que una urna está vacía dé lugar al estado en el que las dos urnas están parejas, a que este último estado dé lugar al primero.

Grandes desviaciones

■ Para ser más gráficos, supongamos que el sistema de las dos urnas se reemplaza por dos perros que comparten entre sí 2N pulgas. Cada segundo una pulga salta de un perro al otro. Pongamos por caso que hay 100 pulgas. En determinado momento uno de los perros tiene 90 pulgas, lo que, de acuerdo a la distribución de equilibrio, es un suceso muy improbable, con

$$p(m = 40) \sim 10^{-17}. \tag{20}$$

Nos ahora preguntamos cuál de estas situaciones es más probable:

- Que el proceso esté avanzando hacia un estado más improbable aún.
- Que el proceso esté regresando hacia un estado más probable.
- Que el proceso haya venido desde y esté regresando hacia un estado más improbable.
- Que el proceso haya venido desde y esté regresando hacia un estado más probable.

Dicho en fórmulas, las cuatro probabilidades que les pedimos calcular son las siguientes:

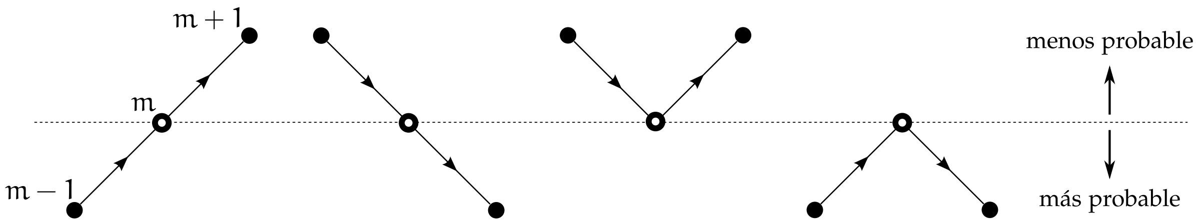
$$p\big(m+1, i+1\ ;\ m-1, i-1\ \big|\ m, i\big), \tag{21}$$

$$p\big(m-1, i+1\ ;\ m+1, i-1\ \big|\ m, i\big), \tag{22}$$

$$p\big(m+1, i+1\ ;\ m+1, i-1\ \big|\ m, i\big), \tag{23}$$

$$p\big(m-1, i+1\ ;\ m-1, i-1\ \big|\ m, i\big), \tag{24}$$

suponiendo $m > 0$. La figura muestra de manera gráfica las cuatro cadenas de eventos:



El cálculo no es difícil. Luego vean qué números da en el caso particular de los dos perros y sus 100 pulgas, cuando uno de los perros acapara 90 pulgas.

A propósito de este resultado general del modelo de Ehrenfest, Kac anota lo siguiente:

> This little analysis provides an interpretation of the seemingly meaningless and paradoxical statement of Boltzmann that every point of the H-curve is a maximum.

Esto no les dirá mucho, pero no dejará de resultarles intrigante. Investiguen más ahora o recuérdenlo para cuando vean el teorema H de Boltzmann.

Bibliografía recomendada

- Kac, M. *Probability And Related Topics In Physical Sciences*, Interscience Publishers (1959). Sección III-7. Amigo de Feynman. Muy entretenido. Se pueden leer las secciones sueltas sin perder la ilación. Tiene muchos resultados y técnicas interesantes. No se enreden con la matemática. Busquen los párrafos más conversados.

- Balian, R. *From Microphysics to Macrophysics vol. 1*, Springer (2007). Sección "The Ehrenfests' Urn Model".

- Kelly, F. P. *Reversibility and Stochastic Networks*, Wiley (1979). Sección 1.4.

- Grinstead, C. y Snell, L. *Introduction to Probability*, American Mathematical Society (2012). Ejemplo 11.28.

- Del artículo de Paul y Tatjana Ehrenfest sólo pude encontrar su versión original en alemán. Si alguien encuentra una versión traducida, o si quiere traducirlo, somos todo oídos. https://www.lorentz.leidenuniv.nl/IL-publications/sources/Ehrenfest_07b.pdf