

Simulaciones de Montecarlo en Mecánica Estadística

Introducción

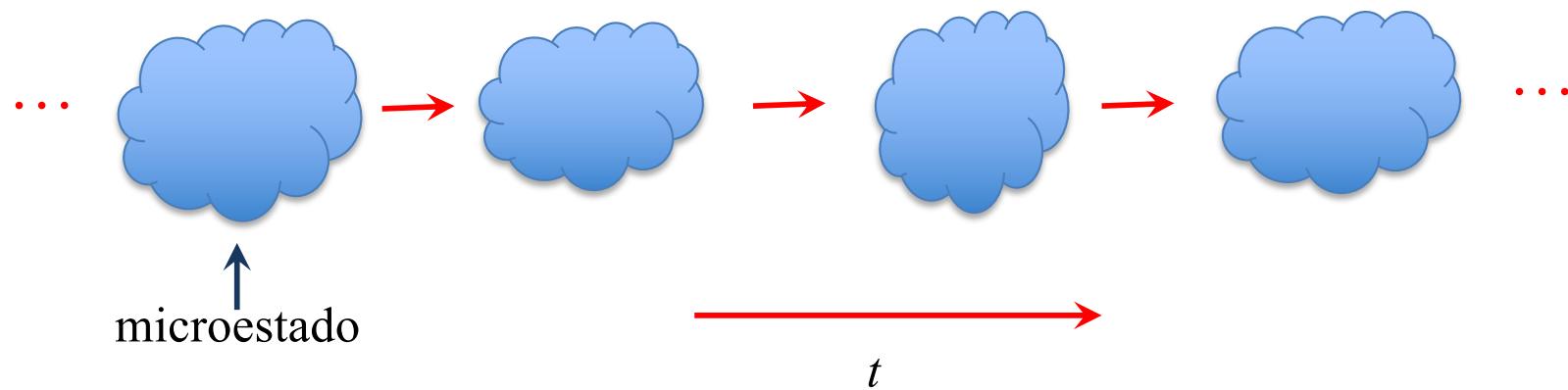
A cualquier simulación por computadora que usa números aleatorios para simular un proceso físico “aleatorio” o no y de esta manera estimar cantidades relacionadas con la salida de este proceso se le llama simulación de Montecarlo

Mayor desarrollo en Mecánica Estadística

Muestreo de importancia en mecánica estadística

Sistema físico en equilibrio:

Ergodicidad: todo estado del sistema es alcanzable desde otro por una secuencia apropiada de movimientos



Ejemplo:

NVT fijas
Ensemble
Canónico

$$P(E_i) = \frac{e^{-\beta E_i}}{Z} \quad \text{Probabilidad de que el sistema se encuentre en el microestado } i$$

$$Z = \sum_i e^{-\beta E_i} \quad \text{Función de Partición}$$

$$\beta = \frac{1}{k_B T}, \quad k_B = 1.38 \times 10^{-23} \text{ J/K} \quad \text{Constante de Boltzmann}$$

Muestreo de importancia en mecánica estadística

¿Qué nos interesa?

$$\langle X \rangle = \sum_i X_i P(E_i)$$

Muchísimos microestados posibles: para un sistema de 10^{23} spines fijos tenemos 10^{70} microestados.

Se puede hacer de manera analítica para muy pocos casos.
Hay que hacerlo de manera numérica. Pensemos en las integrales de Montecarlo.

$$\langle X \rangle \simeq \frac{\sum_{i=1}^N X_i P(E_i)}{\sum_{i=1}^N P(E_i)}$$

¿Qué problemas trae este procedimiento?

Tomamos N
microestados al azar

Muestreo de importancia en mecánica estadística

Solución: Muestreo de importancia

Sea el siguiente promedio ponderado:

$$\langle g \rangle_w = \frac{\sum_i w_i g_i}{\sum_i w_i}$$

$$\text{Sea } g_i = \frac{X_i P(E_i)}{w_i}$$

$$\text{Por lo tanto: } \left\langle \frac{XP(E)}{w} \right\rangle_w = \frac{\sum_i X_i P(E_i)}{\sum_i w_i} = \frac{\langle X \rangle}{\sum_i w_i}$$

$$\text{Despejando } \langle X \rangle = \left\langle \frac{XP(E)}{w} \right\rangle_w \sum_i w_i.$$

$$\text{Por otro lado: } \langle g \rangle_w \simeq \frac{1}{N} \sum_{k=1}^N g_k \quad \text{a partir de: } p_i = \frac{w_i}{\sum_j w_j}$$

$$\text{Por lo tanto: } \langle X \rangle \simeq \frac{1}{N} \sum_{k=1}^N \frac{X_k P(E_k)}{w_k} \sum_j w_j$$

$$\text{Escogiendo: } w_i = P(E_i), \quad \sum_j P(E_j) = 1.$$

$$\text{Tenemos finalmente: } \langle X \rangle \simeq \frac{1}{N} \sum_{k=1}^N X_k \quad \xleftarrow{\text{Generados a partir de:}} \quad p_k = P(E_k)$$

Suma sobre todas
las
configuraciones

Muestreo de importancia en mecánica estadística

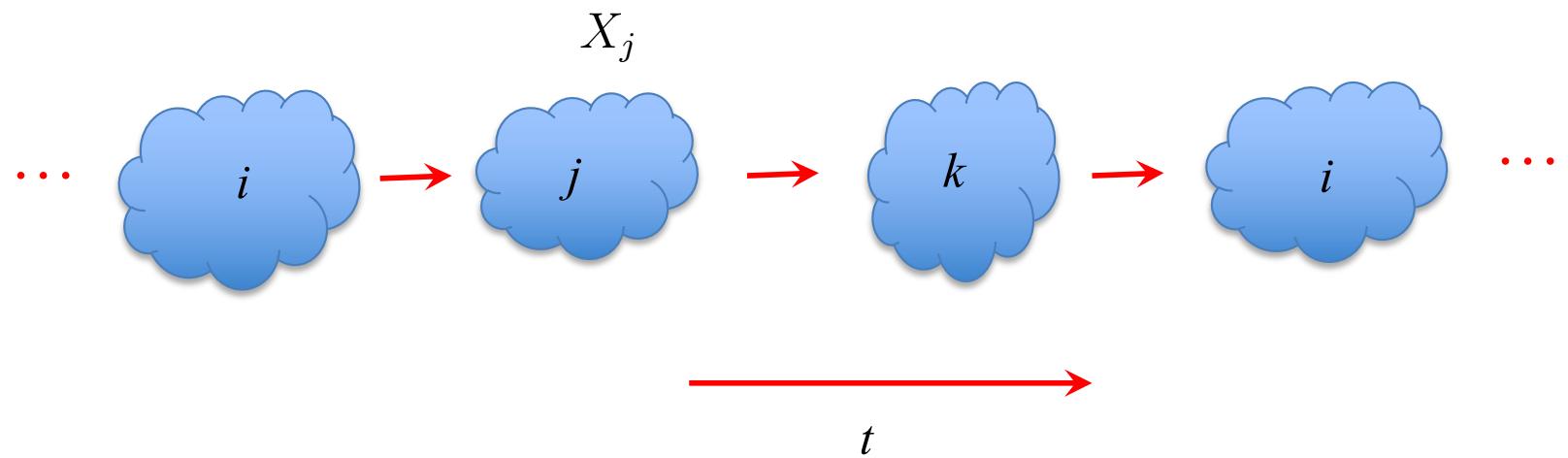
- ¿Cuál es el principal problema de este método?
 - ✓ No se conoce la función de partición.
- La alternativa, mimetizar lo que ocurre en los sistemas reales

Simulaciones de Montecarlo

Se emulará lo que el sistema hace en realidad, es decir, pasando por microestados relativamente cercanos unos de otros.

Método de la cadena de Markov

El objetivo de la cadena de Markov es mimetizar lo que ocurre en la realidad



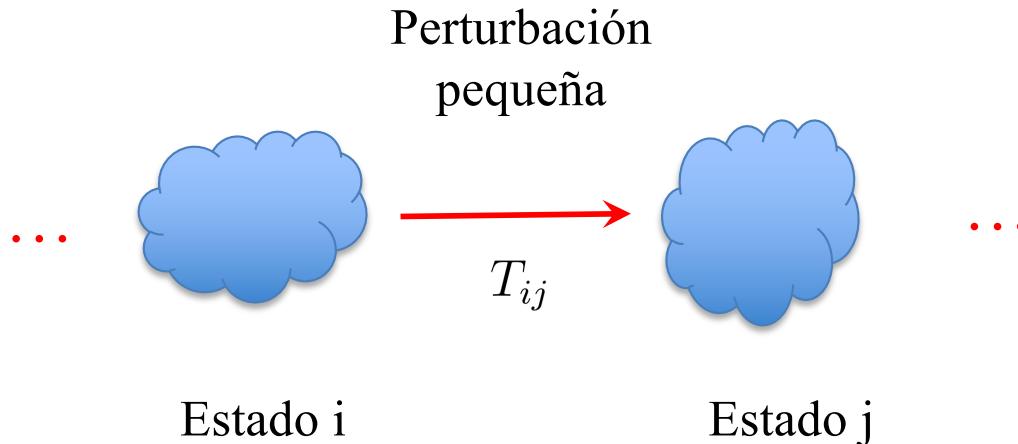
La probabilidad de que el sistema de encuentre en el microestado i es $P(E_j)$

Estimamos el valor de X_j en cada paso de tiempo y calculamos

$$\langle X \rangle \simeq \frac{1}{N} \sum_{t=t_0}^{t_f} X_t$$

Muestreo de importancia

Método de la cadena de Markov



T_{ij} : probabilidad de transición del estado i al j $\sum_j T_{ij} = 1$

Queremos que en equilibrio la probabilidad de visitar un estado particular i sea $P(E_i)$.

$$\frac{T_{ij}}{T_{ji}} = \frac{P(E_j)}{P(E_i)} = \frac{e^{-\beta E_j}/Z}{e^{-\beta E_i}/Z} = e^{-\beta(E_j - E_i)}$$

Balance detallado

$$p_i T_{ij} = p_j T_{ji}$$

Consecuencia:

$$\sum_i T_{ij} P(E_i) = \sum_i T_{ji} P(E_j) = P(E_j) \sum_i T_{ji} = P(E_j) \quad \rightarrow \quad \mathbf{p}_B \hat{\mathbf{T}} = \mathbf{p}_B$$

- ✓ La distribución de Boltzmann es un punto fijo de la cadena de Markov

Método de la cadena de Markov

¿Qué pasa si partimos de otra distribución?

$p_i(t)$: probabilidad de que la cadena de Markov visite el estado i en el tiempo t

$$p_j(t+1) = \sum_i T_{ij} p_i(t)$$

$$\mathbf{p}(t+1) = \hat{\mathbf{T}} \mathbf{p}(t) \quad (\text{el elemento } ij \text{ de } \hat{\mathbf{T}} \text{ es } T_{ji})$$

$$\mathbf{p}(t) = \hat{\mathbf{T}}^t \mathbf{p}(0)$$

$$\mathbf{p}(0) = \sum_k c_k \mathbf{v}_k \quad (\mathbf{v}_k: \text{ autovectores derechos de } \hat{\mathbf{T}})$$

$$\mathbf{p}(t) = \hat{\mathbf{T}}^t \sum_k c_k \mathbf{v}_k = \sum_k c_k \lambda_k^t \mathbf{v}_k = \lambda_1^t \sum_k c_k \left(\frac{\lambda_k}{\lambda_1} \right)^t \mathbf{v}_k$$

λ_k : autovalores λ_1 : el de mayor magnitud

$$\lim_{t \rightarrow \infty} \frac{\mathbf{p}(t)}{\lambda_1^t} = c_1 \mathbf{v}_1$$

$$\sum_i T_{ij} P(E_i) = P(E_j) \Rightarrow \hat{\mathbf{T}} \mathbf{p}_B = \mathbf{p}_B \quad (\mathbf{p}_B: \text{ autovector con las probabilidades de Boltzmann})$$

Solamente falta probar que \mathbf{p}_B es \mathbf{v}_1

$$\sum_j T_{ij} = 1 \Rightarrow \mathbf{1}^T \hat{\mathbf{T}} = \mathbf{1}^T \quad (\mathbf{1}^T: \text{ autovector izquierdo de } \hat{\mathbf{T}} \text{ con autovalor 1})$$

El resto de los autovectores derechos de $\hat{\mathbf{T}}$ deben tener al menos una componente positiva

Supongamos que partimos en $t = 0$ de un $\mathbf{p}(0)$ con todas las componentes positivas. Debido a que $\hat{\mathbf{T}}$ es definida positiva, al aplicarlo t veces sobre $\mathbf{p}(0)$ se obtendrá un vector con todas las componentes positivas. Por lo tanto \mathbf{v}_1 debe ser \mathbf{p}_B

¿Qué pasa si λ_1 es degenerado?
Ergodicidad

Método de la cadena de Markov

El algoritmo de Metropolis

1. Escoger un estado inicial aleatorio.
2. Escoger un movimiento al azar de un cierto conjunto de movimientos permitidos, tal como cambiar una molécula de estado
3. Calcular el valor de la probabilidad de aceptación P_a .

$$P_a = \begin{cases} 1 & \text{si } E_j \leq E_i \\ e^{-\beta(E_j - E_i)} & \text{si } E_j > E_i \end{cases}$$

4. Aceptar el movimiento con probabilidad P_a .
5. Estimar la cantidad de interés X y almacenarla.
6. Repetir desde el punto 2.

Método de la cadena de Markov

El algoritmo de Metropolis

¿Por qué sabemos que este algoritmo convergerá la distribución de equilibrio (Boltzmann)?

La probabilidad T_{ij} de moverse de un estado i a uno j es la probabilidad de escogencia entre todas las posibilidades, $1/M$, multiplicada por la probabilidad de que aceptemos el movimiento P_a .

Caso: $E_j > E_i$

$$T_{ij} = \frac{1}{M} \times e^{-\beta(E_j - E_i)} \quad T_{ji} = \frac{1}{M} \times 1$$

$$\frac{T_{ij}}{T_{ji}} = \frac{e^{-\beta(E_j - E_i)}/M}{1/M} = e^{-\beta(E_j - E_i)}$$

Caso: $E_i \geq E_j$

$$T_{ij} = \frac{1}{M} \times 1 \quad T_{ji} = \frac{1}{M} \times e^{-\beta(E_i - E_j)}$$

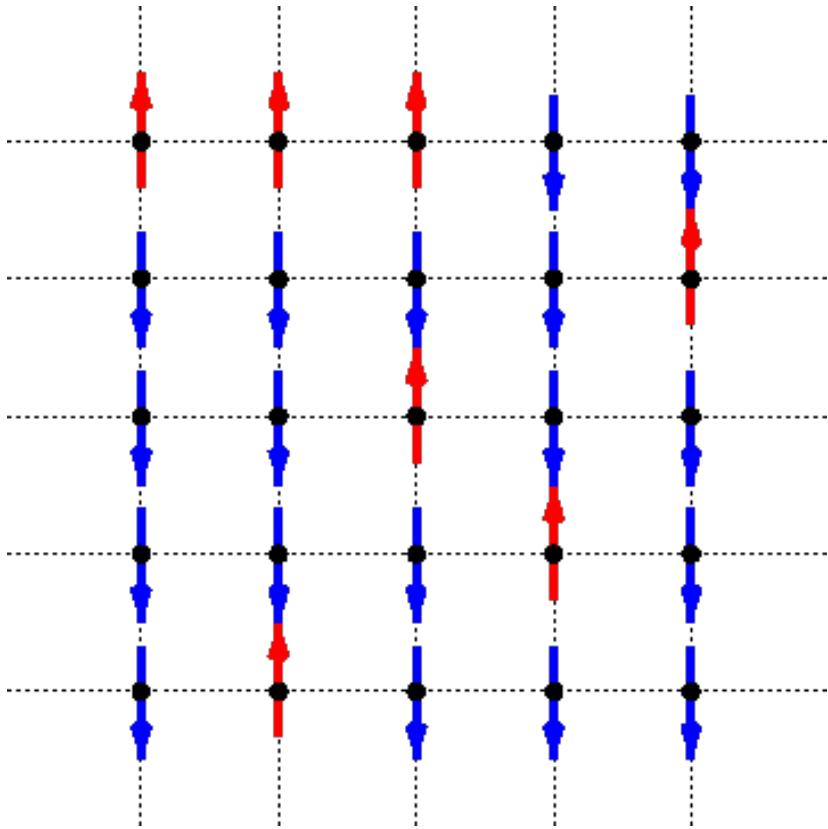
$$\frac{T_{ij}}{T_{ji}} = \frac{1/M}{e^{-\beta(E_i - E_j)}/M} = e^{-\beta(E_j - E_i)}$$

Se cumple
Balance detallado

El modelo de Ising

Ferromagnetismo

Es el modelo más simple para estudiar el ferromagnetismo y el antiferromagnetismo.



$$E = -J \sum_{\langle ij \rangle} s_i s_j$$

- J : Magnitud de la interacción.
Interacción de intercambio.
- s_i : Spin de la prtícula i . $s_i \in \{-1, 1\}$.
- $\sum_{\langle ij \rangle}$: Suma sobre primeros vecinos.

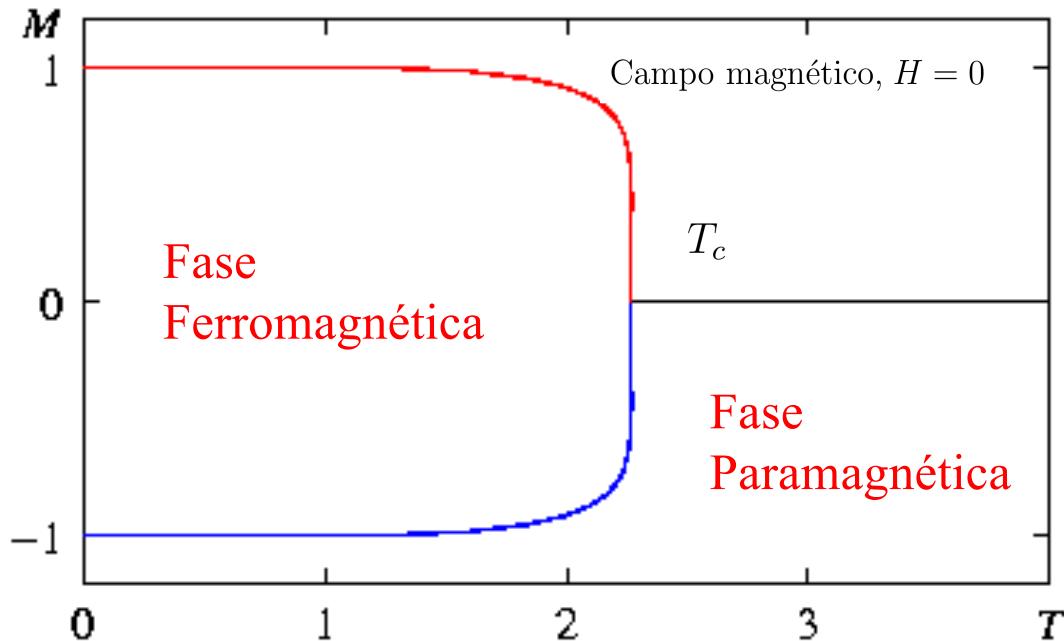
A bajas temperaturas, debajo de la temperatura de Curie
 $J > 0$: ferromagnetismo. $J < 0$: antiferromagnetismo

El modelo de Ising

$d=2$

Ferromagnetismo

$$\frac{k_B T_c}{J} = \frac{2}{\ln(1 + \sqrt{2})} \approx 2.26918531421\cdots$$



Transición de fase de segundo orden

$$M = \frac{1}{N} \sum_{i=1}^N s_i \quad \text{Magnetización por partícula}$$

El modelo de Ising

Detalles de la simulación, $d=2$

- Escogemos un spin al azar, cambiamos su valor de s_1 a s_2 .
- Observamos el cambio en la magnetización y en la energía

Cambio en la magnetización:

$$\Delta M = s_2 - s_1 = \pm 2$$

Cambio en la energía:

$$\Delta E = -J(s_2 - s_1) \sum_{j=1}^4 s_j = -J\Delta M \sum_{j=1}^4 s_j$$

$$\Delta E = J \times \{\pm 8, \pm 4, 0\}$$

El modelo de Ising

Fluctuación disipación.

Capacidad calorífica y susceptibilidad magnética

$$C = \frac{\partial \langle E \rangle}{\partial T} = \frac{1}{k_B T^2} (\langle E^2 \rangle - \langle E \rangle^2)$$

$$\chi = \left. \frac{\partial \langle M \rangle}{\partial H} \right|_{H=0} = \frac{1}{k_B T} (\langle M^2 \rangle - \langle M \rangle^2)$$

Estimando las fluctuaciones de la energía y de la magnetización obtenemos la capacidad calorífica y la susceptibilidad magnética

El modelo de Ising

Fluctuación disipación.
Susceptibilidad magnética y capacidad calorífica

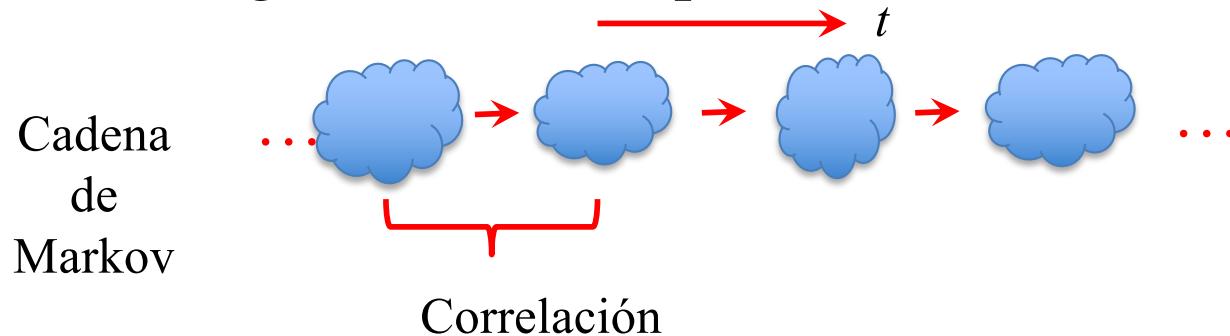
$$\begin{aligned}
 C &= \frac{\partial \langle E \rangle}{\partial T} = -\frac{1}{k_B T^2} \frac{\partial \langle E \rangle}{\partial \beta} \\
 &= -\frac{1}{k_B T^2} \frac{\partial}{\partial \beta} \frac{\sum_i E_i e^{-\beta E_i}}{Z}, \quad Z = \sum_i e^{-\beta E_i} \\
 &= -\frac{1}{k_B T^2} \left(-\frac{1}{Z^2} \frac{\partial Z}{\partial \beta} \sum_i E_i e^{-\beta E_i} - \frac{1}{Z} \sum_i E_i^2 e^{-\beta E_i} \right) \\
 &= -\frac{1}{k_B T^2} \left[\frac{1}{Z^2} \left(\sum_i E_i e^{-\beta E_i} \right)^2 - \frac{1}{Z} \sum_i E_i^2 e^{-\beta E_i} \right] \\
 &= \frac{1}{k_B T^2} (\langle E^2 \rangle - \langle E \rangle^2)
 \end{aligned}$$

Energía del microestado i en presencia de un campo magnético H

$$\begin{aligned}
 E_i &= -J \sum_{\langle lk \rangle} s_i^l s_i^k - H \sum_k s_i^k = -J \sum_{\langle lk \rangle} s_i^l s_i^k - H M_i \\
 \chi &= \frac{\partial \langle M \rangle}{\partial H} = \frac{\partial}{\partial H} \frac{\sum_i M_i e^{-\beta E_i}}{Z}, \quad Z = \sum_i e^{-\beta E_i} \\
 &= -\frac{1}{Z^2} \frac{\partial Z}{\partial H} \sum_i M_i e^{-\beta E_i} - \frac{1}{Z} \sum_i \beta M_i^2 e^{-\beta E_i} \\
 &= \beta \left[\frac{1}{Z^2} \left(\sum_i M_i e^{-\beta E_i} \right)^2 - \frac{1}{Z} \sum_i M_i^2 e^{-\beta E_i} \right] \\
 &= \frac{1}{k_B T} (\langle M^2 \rangle - \langle M \rangle^2)
 \end{aligned}$$

El modelo de Ising

Configuraciones independientes



- En el paso i estimo una cantidad A_i . ¿Cuánto tiempo pasa para que una nueva medida de A sea independiente de la anterior?
- Sistema en equilibrio

Función de autocorrelación temporal:

$$C_A(t) = \frac{\langle A(t + t_0)A(t_0) \rangle - \langle A \rangle^2}{\langle A^2 \rangle - \langle A \rangle^2}$$

Para tiempos suficientemente grandes:

$$\langle A(t + t_0)A(t_0) \rangle \rightarrow \langle A(t + t_0) \rangle \langle A(t_0) \rangle = \langle A \rangle^2$$

Es decir: $C_A(t) \rightarrow 0$ cuando $t \rightarrow \infty$

Para tiempos cortos:

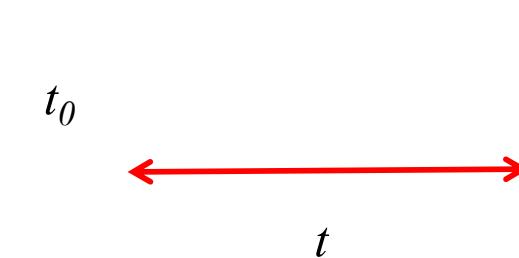
$$C_A(t) \rightarrow 1 \text{ cuando } t \rightarrow 0$$

El modelo de Ising

Configuraciones independientes

- ¿Cómo se calcula?
- De nuestro programa calculamos una cierta cantidad física como función del tiempo (mcs). Ya en equilibrio.

M_0	M_1	M_2	\cdots	M_{N-3}	M_{N-2}	M_{N-1}
-------	-------	-------	----------	-----------	-----------	-----------



- Variamos t_0 (i) entre 0 y $N-1$. Calculamos C_A para t entre 0 y $N-i$.
- Promediamos para cada t para todos los t_0 .

El modelo de Ising

Configuraciones independientes

Cuidado con las normalización:

Para el procedimiento explicado en clases la función de autocorrelación temporal viene dada por:

$$C_a(t) = \frac{\sum_{k=0}^{N'-1} (A_k - \langle A \rangle)((A_{k+t} - \langle A \rangle))}{\sum_{k=0}^{N-1} (A_k - \langle A \rangle)^2}$$

donde $N' = N - t$.

Desarrollando esta expresión tenemos:

$$C_a(t) = \frac{\sum_{k=0}^{N'-1} A_k A_{k+t} - \langle A \rangle \sum_{k=0}^{N'-1} A_k - \langle A \rangle \sum_{k=0}^{N'-1} A_{k+t} + N' \langle A \rangle^2}{\sum_{k=0}^{N-1} (A_k - \langle A \rangle)^2}$$

$$C_a(t) = \frac{N'}{N} \frac{\frac{1}{N'} \sum_{k=0}^{N'-1} A_k A_{k+t} - \frac{\langle A \rangle}{N'} \sum_{k=0}^{N'-1} (A_k + A_{k+t}) + \langle A \rangle^2}{\langle A^2 \rangle - \langle A \rangle^2}$$

Recordemos que:

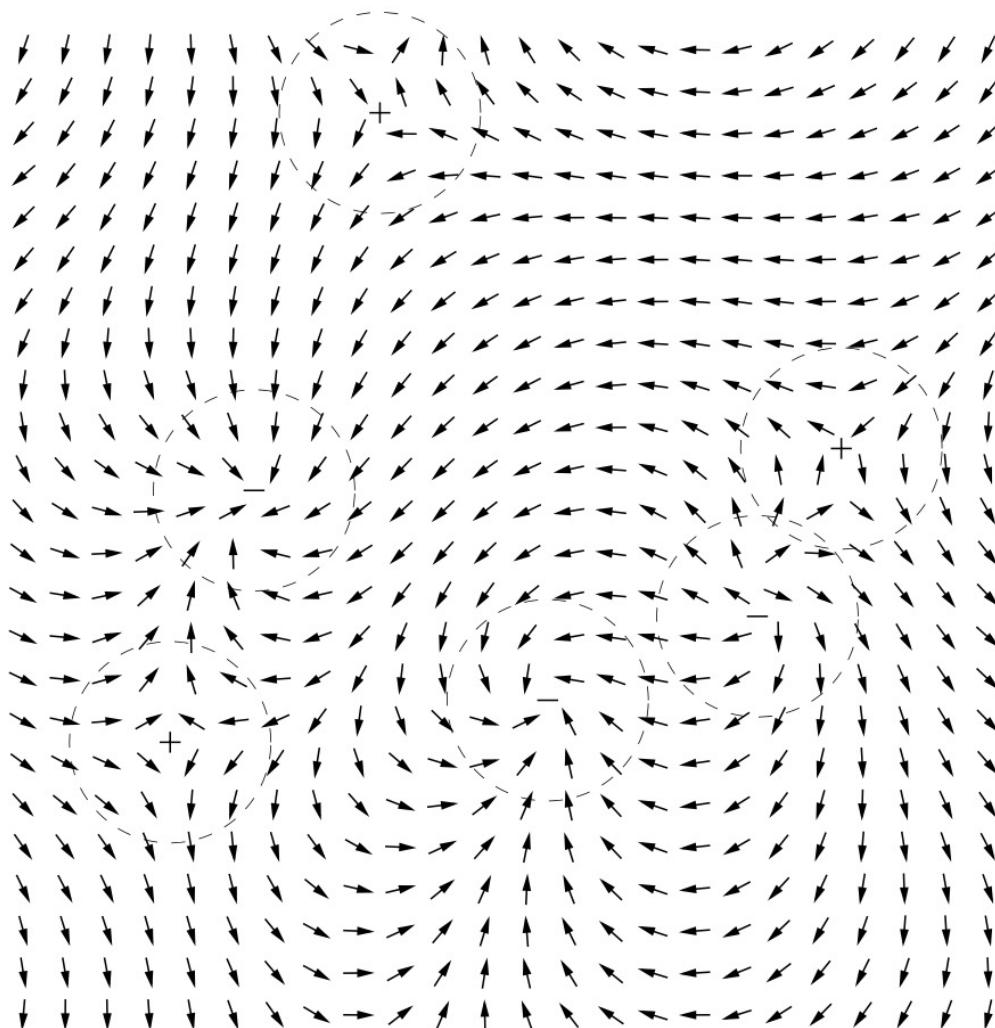
$$\langle A \rangle = \frac{1}{N} \sum_{k=0}^{N-1} A_k, \quad \langle A^2 \rangle = \frac{1}{N} \sum_{k=0}^{N-1} A_k^2$$

Las fórmulas son idénticas en el límite $N \gg 1$ y $t \ll N$.

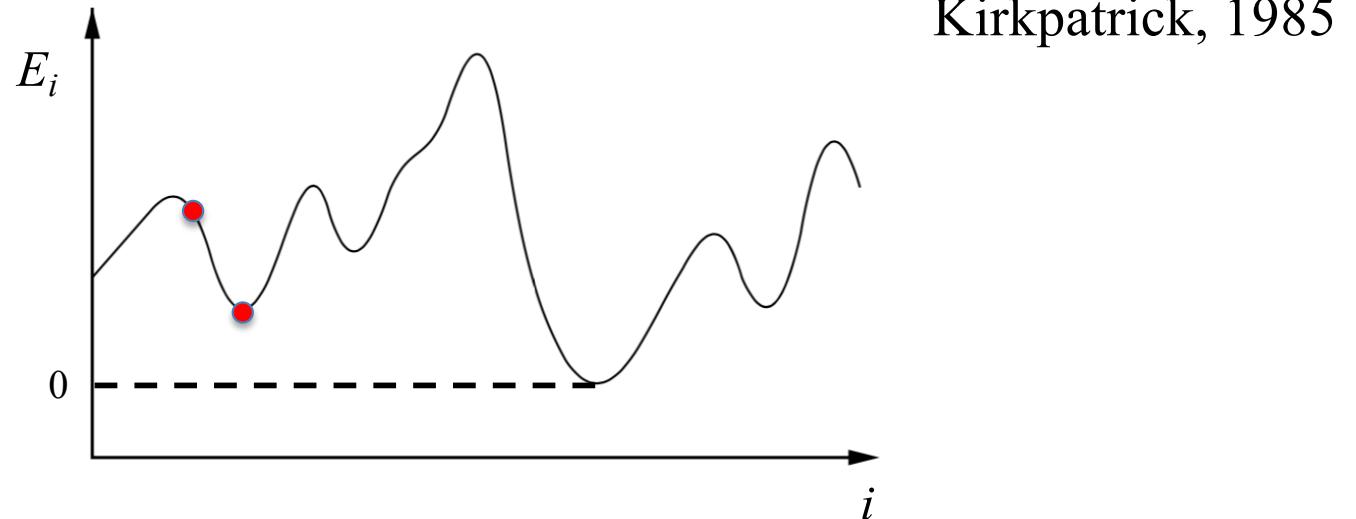
Otros modelos magnéticos

El modelo XY

$$\begin{aligned} E &= -J \sum_{i,j=\text{nn}(i)} \vec{S}_i \cdot \vec{S}_j \\ &= -J \sum_{i,j=\text{nn}(i)} (S_{i,x} S_{j,x} + S_{i,y} S_{j,y}) \end{aligned}$$



Recocido simulado



¿Cómo conseguir el mínimo global de una función?

De la mecánica estadística: $P(E_i) = \frac{e^{-\beta E_i}}{Z}$

Para $T = 0$:

$$P(E_i) = 0, \quad \text{para } E_i \neq E_{min}$$

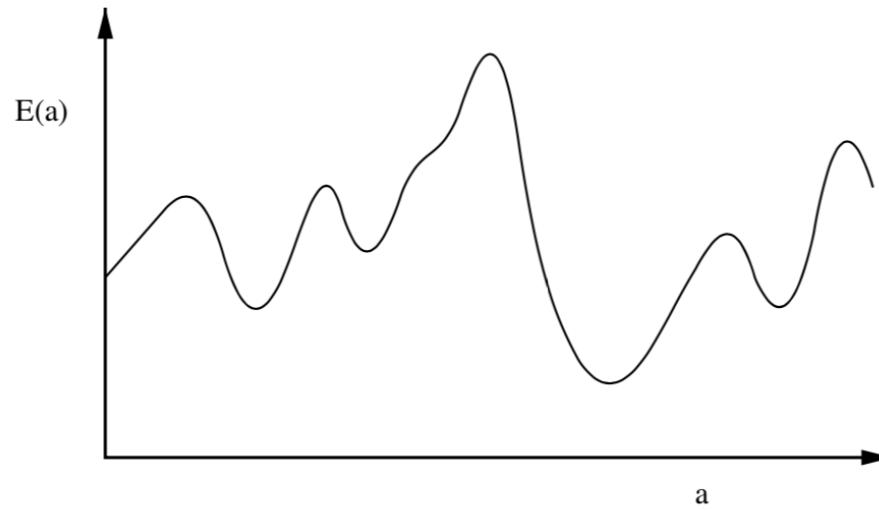
$$P(E_i) = 1, \quad \text{para } E_i = E_{min} = 0$$

¿Qué pasa si aplicamos Metropolis en $T=0$? Mínimo local

¿Cómo resolvemos el problema?

Lo que se aplica en la industria, recocido simulado

Recocido simulado



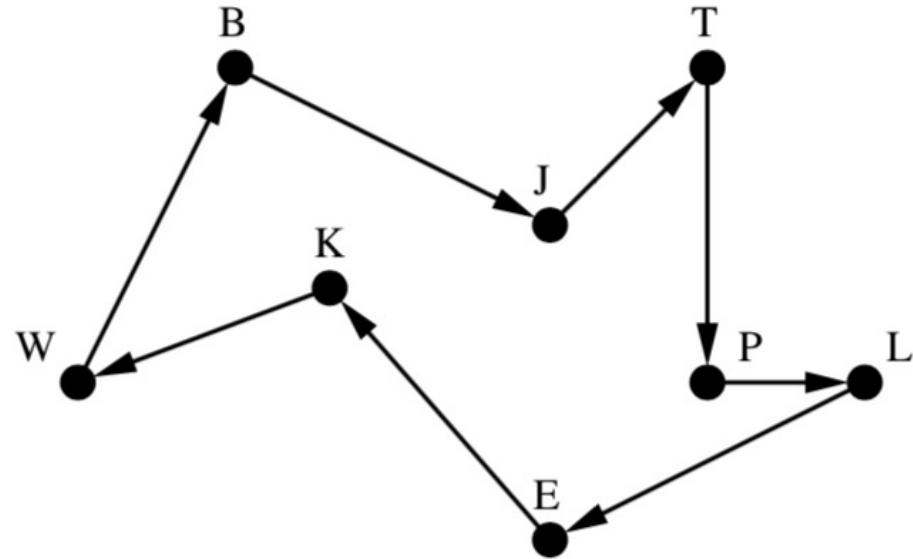
Kirkpatrick, 1985

¿Cómo conseguir el mínimo global de una función?

- Escogemos un valor de a
- Generamos un cambio pequeño de a , da
- Si $\Delta E = E(a+da) \leq E(a)$ aceptamos el cambio, si $\Delta E > 0$ lo aceptamos con probabilidad $p = e^{-\Delta E/T}$, T es la temperatura efectiva.
- Al principio se escoge T para que los movimientos sean aceptados. Luego se va disminuyendo hasta un valor muy pequeño. Mientras más lento se baja la temperatura mayor es la posibilidad de llegar a un mínimo global.

Recocido simulado

El problema del agente viajero

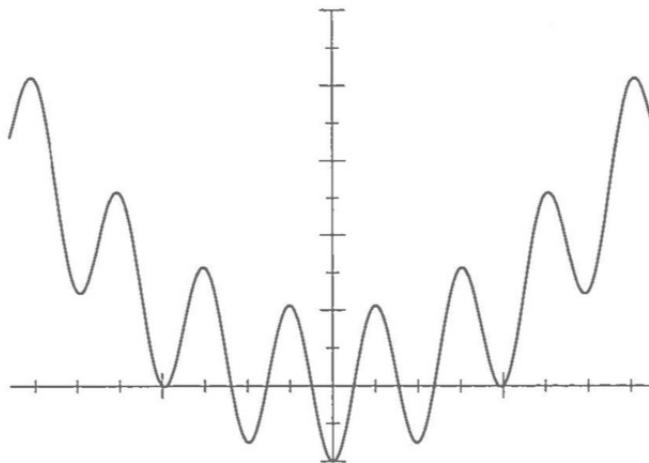


$$D = \sum_{i=0}^{N-1} |\vec{r}_{i+1} - \vec{r}_i|$$

Recocido simulado

Exercise 10.10: Global minimum of a function

Consider the function $f(x) = x^2 - \cos 4\pi x$, which looks like this:



Clearly the global minimum of this function is at $x = 0$.

- Write a program to confirm this fact using simulated annealing starting at, say, $x = 2$, with Monte Carlo moves of the form $x \rightarrow x + \delta$ where δ is a random number drawn from a Gaussian distribution with mean zero and standard deviation one. (See Section 10.1.6 for a reminder of how to generate Gaussian random numbers.) Use an exponential cooling schedule and adjust the start and end temperatures, as well as the exponential constant, until you find values that give good answers in reasonable time. Have your program make a plot of the values of x as a function of time during the run and have it print out the final value of x at the end. You will find the plot easier to interpret if you make it using dots rather than lines, with a statement of the form `plot(x, ".")` or similar.
- Now adapt your program to find the minimum of the more complicated function $f(x) = \cos x + \cos \sqrt{2}x + \cos \sqrt{3}x$ in the range $0 < x < 50$.

Hint: The correct answer for part (b) is around $x = 16$, but there are also competing minima around $x = 2$ and $x = 42$ that your program might find. In real-world situations, it is often good enough to find any reasonable solution to a problem, not necessarily the absolute best, so the fact that the program sometimes settles on these other solutions is not necessarily a bad thing.