



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

**FACULTAD DE CIENCIAS**

**ANÁLISIS DE EARLY WARNINGS EN EL MODELO  
DE ISING EN 2D**

**T E S I S**

**QUE PARA OBTENER EL TÍTULO DE:**

**FÍSICO**

**P R E S E N T A:**

**CARLOS CALDERÓN ÁNGELES**



**DIRECTOR DE TESIS:  
DR. IRVING OMAR MORALES AGISS  
2016**

Ciudad Universitaria, D. F.



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

**1. Datos del Alumno**

Calderón  
Ángeles  
Carlos  
Universidad Nacional Autónoma de México  
Facultad de Ciencias  
Física  
307005153

**2. Datos del tutor**

Dr.  
Irving Omar  
Morales  
Agiss

**3. Datos del Sinodal 1**

Dr.  
Alejandro  
Frank  
Hoeflich

**4. Datos del Sinodal 2**

Dr.  
Del Castillo  
Mussot  
Marcelo

**5. Datos del Sinodal 3**

Dra.  
Lucía  
Medina  
Gómez

**6. Datos del Sinodal 4**

Dr.  
Escobar  
Sotomayor  
Juan Valentín

**7. Datos del trabajo escrito**

Análisis de Alertas Tempranas en el Modelo de Ising en 2d  
115p  
2016

# Resumen

En este trabajo se realiza por primera vez en la literatura un análisis del modelo de Ising a partir del estudio de alertas tempranas encontrando con esto una gran cantidad de propiedades que permiten ver al modelo de Ising desde un punto de vista del análisis de los sistemas dinámicos. Esta forma de ver al modelo como un sistema dinámico que presenta alertas tempranas permite ampliar el panorama no sólo del modelo mismo sino de las transiciones de fase continuas de los sistemas físicos. Esta forma de analizar a los sistemas enriquece su entendimiento y sus posteriores aplicaciones por lo que un análisis desde esta perspectiva es de suma importancia. Persiguiendo este fin, el trabajo muestra como el modelo de Ising puede ser clasificado dinámicamente dentro del universo de las bifurcaciones supercríticas de Pitchfork. Además, por medio de las alertas tempranas es posible determinar la evolución temporal del sistema y en base a eso poder hacer conjeturas sólidas del porqué el sistema presenta o no cambios en esos instantes del tiempo. El tema de alertas tempranas es relativamente nuevo y está cobrando gran auge debido a su posibilidad de aplicación en diversas áreas de la ciencia. . De esta manera este trabajo recopila información al respecto del análisis de alertas tempranas así como de las transiciones de fase para en base a esto encontrar aplicación en un modelo relativamente sencillo y conocido.

Los datos fueron obtenidos de manera computacional utilizando el método de MonteCarlo con el algoritmo de Metrópolis programados en NetLogo 5.1. Para su análisis se utilizaron programas hechos con Mathematica 10. Como fruto de este trabajo se publicó el artículo *“Behavior of Early Warnings near the Critical Temperature in the Two-Dimensional Ising Model”* en la revista Plos One el cuál se anexa al final de esta tesis,

# Índice

• Introducción .....	4
1. Transiciones de Fase y Modelo de Ising .....	8
1.1 Transiciones de Fase .....	8
1.2 Modelo de Ising .....	19
1.3 Simulaciones Monte Carlo .....	34
2. Análisis de Series de tiempo y Alertas Tempranas...	40
2.1 Series de Tiempo .....	40
2.2 El Espectro de Potencias .....	41
2.3 Conceptos Estadísticos Básicos .....	51
2.4 Alertas Tempranas .....	54
2.5 Consideraciones Prácticas .....	62
3. Procedimiento, Resultados y Análisis .....	65
3.1 Procedimiento .....	65
3.2 Resultados y Análisis .....	68
4. Discusión y Conclusiones .....	82
• Referencias .....	96

- Apéndice (artículo "Behavior of Early Warnings near the Critical Temperature in the Two-Dimensional Ising Model").....99

# Introducción

Las transiciones de fase pueden definirse como cambios abruptos en alguna (o algunas) de las propiedades del sistema. Visto de esta manera tan general, el estudio de transiciones de fase podría servir como preámbulo para poder explicar cambios abruptos que ocurren en los sistemas complejos, es decir, sistemas con muchos agentes donde la interacción entre ellos es no lineal; en todo caso, el estudio en el espacio de los parámetros del sistema de las zonas cercanas a los cambios abruptos ha incrementado mucho en los últimos años [1, 29]. En particular es de interés la aparición de ciertos cambios en algunas de las propiedades asociadas al sistema dinámico siempre que ocurre algún cambio abrupto. Algunos de estos cambios en las propiedades reciben el nombre de alertas tempranas [1, 29, 30] que, como su nombre lo indica, muestran cambios repentinos en el comportamiento del sistema en puntos cercanos de alguna transición dinámica, a tal grado que para ciertos sistemas, el saber que está cambiando alguna de sus propiedades es un preámbulo para indicar que el sistema mismo se dirige hacia un cambio abrupto. Muchas de estas alertas tempranas son independientes del modelo que se elija y prácticamente cualquier sistema dinámico con alguna bifurcación las presenta. Es importante recalcar que recientemente se ha encontrado [1] que gran parte de las alertas tempranas pueden detectar incluso, desde puntos alejados de alguna bifurcación dinámica, cambios continuos y muy suaves que sufrirá el sistema. Entender y estudiar las alertas tempranas se vuelve entonces una herramienta muy útil para el estudio de los sistemas dinámicos, ya que ayudarían a determinar comportamientos futuros del sistema sin necesidad de conocer a fondo el comportamiento particular del mismo, sino solamente algunas de sus propiedades. Esto no sólo podría ayudar a entender el sistema mismo, sino también a regularlo, y con esto tener cierto control sobre ellos tanto para que se mantengan en una zona estable como para precipitar su

cambio.

En el presente trabajo se pretende estudiar algunas alertas tempranas en el modelo de Ising en dos dimensiones (2d). El modelo de Ising es tal vez el modelo de transición de fase más conocido y estudiado en toda la literatura científica y para el cual se conoce la solución exacta (si el campo externo aplicado  $H=0$ ) para una, dos e infinitas dimensiones. El hecho de que sea un modelo a partir del cual se basen también muchos modelos físicos (los modelos de lattice gas y de spin glass, por ejemplo), biológicos [2-6] y sociales [7, 8], le da al modelo una importancia aún mayor dentro de la ciencia; sin embargo, carece de un análisis profundo basándose solamente en sus alertas tempranas. El realizar un análisis de alertas tempranas del modelo de Ising es útil en dos sentidos: primero, ayudaría a comprender mejor el comportamiento dinámico del modelo de Ising mismo y con esto entender al modelo de Ising desde una perspectiva un tanto distinta de como se aprende comúnmente entre los físicos. En efecto, el modelo de Ising entre los físicos es visto generalmente como un modelo de transición de fase que presenta una dinámica inducida por algoritmos computacionales y rara vez se le ve como un sistema dinámico propiamente dicho. Un análisis desde los cambios que presenta a través del tiempo permite ver al modelo de Ising desde otra perspectiva y que se complementa a la perfección con la forma puramente física de entender el problema. En segundo lugar, dado que el modelo de Ising es bien conocido, es posible hacer un análisis físico del por qué aparecen las alertas tempranas y con esto comprender muchos de los comportamientos de dichas alertas en otros sistemas dinámicos cerca del punto crítico. El propósito último de este trabajo es mostrar la utilidad de las alertas tempranas para la investigación tanto de transiciones de fase como de sistemas complejos y recalcar que, a pesar de que muchos sistemas puedan provenir de disciplinas a primera vista muy distantes como la sociología, la biología o la economía, la



universalidad de las alertas tempranas permite salirse de las particularidades de cada modelo y analizar solamente cómo cuando el sistema como un todo cambia, cambian también muchas de las propiedades mensurables del mismo. Enmarcarlo en un modelo tan conocido como el de Ising permite además analizar ciertos comportamientos particulares de las alertas tempranas que no resultan tan evidentes al estar tratando con sistemas con una dinámica más compleja.

La investigación se divide en cuatro capítulos:

1. Análisis del modelo de Ising. Este capítulo comienza con un pequeño resumen de las transiciones de fase para posteriormente explicar el modelo de Ising. Se analizan las propiedades principales del modelo de Ising así como los resultados más importantes para el modelo en 2d. Posteriormente se aborda la forma en que el modelo puede ser simulado computacionalmente.
2. Series de tiempo y alertas tempranas. En este capítulo se abordan algunas de las particularidades de las series de tiempo así como el procedimiento para obtener el espectro de potencias de una serie. Se presenta el teorema de Wiener-Kinchin, que permite relacionar las correlaciones del sistema con el espectro de potencias. Se muestran entonces algunos resultados importantes del espectro de potencias del modelo de Ising. Más adelante se habla de algunas definiciones estadísticas importantes para después poder hacer la presentación formal de las alertas tempranas que se usarán en el análisis del sistema, El capítulo concluye con algunas consideraciones prácticas a tomar en cuenta siempre que se estudien las alertas tempranas.
3. Procedimiento, resultados y análisis. En este capítulo se muestra el procedimiento seguido en la investigación así como los resultados de la misma. En el análisis se ahonda en ciertas características particulares de cada resultado obtenido

4. Conclusiones y discusión. Este capítulo resume toda la investigación realizada además de que trata de responder a la pregunta ¿qué tan buenos son las alertas tempranas propuestos y qué explicación física tiene el que aparezcan? El capítulo concluye con mostrar posibles investigaciones a futuro sobre el tema.

# Capítulo 1

## 1 Transiciones de Fase y Modelo de Ising

### 1.1 Transiciones de fase

Una transición de fase es un cambio abrupto en las propiedades de un sistema. Ocurre cuando existe una singularidad en la energía libre o en sus derivadas. Si hay una singularidad en la energía libre que genere una discontinuidad en la primera derivada, se dice que la transición es de primer orden. Si la primera derivada es continua pero la segunda derivada es discontinua o infinita en algún punto se habla de que es una transición de orden superior, una transición continua o una transición crítica. El término “transición de segundo orden” está en desuso aunque proviene de la clasificación original propuesta por Ehrenfest en 1933 en la cual el orden de la transición era el orden de la derivada que mostraba una singularidad [16]. A las transiciones de orden superior se les relaciona con una divergencia en la susceptibilidad, una divergencia en la longitud de correlación y una ley de potencias asociada a esta transición. Posteriormente se ahondará más en este punto.

Para analizar algunas propiedades básicas es necesario recordar algunas definiciones de la mecánica estadística. Recuérdese primero la función de partición:

$$Z = \sum e^{-\beta E_r} \quad (1)$$

Donde la suma va sobre todos los estados energéticos con energía  $E_r$  y  $\beta = 1/kT$ , con  $k$  la constante de Boltzmann y  $T$  la temperatura. Dado que el objetivo de esta breve introducción a las transiciones de fase es estudiar el

modelo de Ising, se tomará como simplificación el hecho de que la función de partición dependa solamente del campo magnético externo aplicado  $H$  y de la temperatura  $T$ , es decir  $Z \rightarrow Z(H, T)$ . Esto no quita la generalidad respecto al análisis de las transiciones de fase pues el desarrollo es idéntico en caso de la función de partición  $Z$  dependa de más variables. A partir de esto, se define la energía libre como:

$$F = -kT \ln Z(H, T) \quad (2)$$

Es decir  $F \rightarrow F(H, T)$  La energía libre tiene la cualidad de que todas las propiedades termodinámicas se desprenden de derivar ésta con respecto a las variables del sistema. Recordando el hecho de que una función  $f(x)$  es cóncava en caso de que  $(df)^2 < 0$ , se tiene

$$\left( \frac{\partial^2 F}{\partial T^2} \right)_H = -\frac{C_H}{T} \quad (3)$$

$$\left( \frac{\partial^2 F}{\partial H^2} \right)_T = -\chi_T \quad (4)$$

De la tercera ley de la termodinámica se desprende que el calor específico debe ser positivo ( $C_H > 0$ ) y por otro lado, en muchos casos, la susceptibilidad puede tomarse como positiva ( $\chi_T > 0$ ). Aunque en general esta condición sobre la susceptibilidad se cumple para muchos sistemas, puede demostrarse ([9]) que si el hamiltoniano del sistema tiene la forma

$$\mathcal{H} = \mathcal{H}_0 - HM \quad (5)$$

se cumple que la susceptibilidad es positiva y por ende la energía libre  $F$  es una función cóncava respecto a las variables  $H$  y  $T$ . El hamiltoniano presentado

en la ecuación (5) es de suma importancia para el presente trabajo ya que es la forma general del modelo de Ising. En la sección 1.2 se ahondará más al respecto.

Para terminar el recordatorio de mecánica estadística se mencionará aquí la función de correlación, la cual ayuda a comprender las relaciones microscópicas que surgen en los sistemas. La función de correlación es dependiente de la longitud de correlación  $\xi$ . La longitud de correlación es una magnitud que dice que tanto afectan al sistema las fluctuaciones de las partes por las que está compuesto y aunque suene lógico tratar de definirla como una función del vector de posición  $\vec{r}$ , puede considerarse como una constante para valores de  $r$  lo suficientemente grandes cuando el sistema se aproxima a la región crítica [9]. Exactamente en el punto crítico pequeños cambios en alguna región del sistema generan cambios en todo el sistema, provocando que la longitud de correlación sea infinita. Cuando el valor de la longitud de correlación es finito, una fluctuación dentro de una región de largo  $\xi$  no tiene efecto afuera de ésta, por lo que para regiones de análisis muy grandes ( $l \gg \xi$ ), los detalles microscópicos se muestran sólo con los promedios de algunas propiedades termodinámicas. De igual manera el sistema es invariante ante traslaciones si se toman distancias  $l \gg \xi$ . En general, puede demostrarse que la medida de cualquier observable en distintas “cajas” de tamaño lineal  $l$  está distribuida de manera gaussiana y la varianza disminuye como  $l/\xi$  (debido al teorema del límite central) [15].

Si el largo de correlación es infinito, no sólo se tiene una invariancia ante traslación sino también una invariancia ante reescalamientos, es decir, el sistema es invariante de escala. A su vez, una fluctuación en una región de tamaño  $l$ , tendrá efectos sobre todo el sistema cualquiera que sea el valor de  $l$ .

En las transiciones de fase de primer orden, la longitud de correlación siempre tiene un valor finito y muy pequeño. Por otro lado, en las transiciones de fase de segundo orden, el valor de  $\xi$  es pequeño para valores de  $T < T_c$ , se vuelve

infinito cuando  $T = T_c$  y vuelve a tomar un valor finito cuando  $T > T_c$  [9].

Se analizará entonces la forma de la función de correlación para un modelo de espines en una dimensión cualquiera. La función de correlación en el caso de espines en las posiciones  $i$  y  $j$  es:

$$\Gamma(\vec{r}_i, \vec{r}_j) = \langle (s_i - \langle s_i \rangle) \langle (s_j - \langle s_j \rangle) \rangle \quad (6)$$

Donde  $\vec{r}_i$  denota la posición del vector en el punto  $i$  y  $\langle \rangle$  es el promedio térmico. Esta función establece cuán relacionados entre sí se encuentran dos espines cualesquiera del sistema. Si el sistema es invariante ante traslaciones no importa cual posición tomemos sobre éste, por lo que  $\langle s_i \rangle = \langle s_j \rangle$  para toda  $i$  y toda  $j$ . Con esto,  $\Gamma$  depende sólo de la distancia ( $\vec{r}_i - \vec{r}_j$ ). Simplificando la expresión de la función de correlación, se llega a que

$$\Gamma(\vec{r}_i, \vec{r}_j) = \Gamma(\vec{r}_i - \vec{r}_j) = \Gamma_{ij} = \langle s_i s_j \rangle - \langle s \rangle^2 \quad (7)$$

Por ende se esperaría que las correlaciones decaigan exponencialmente a cero lejos del punto crítico (tanto por arriba como por abajo del mismo). De igual manera las correlaciones deberían caer a cero mientras mayor sea la distancia entre los espines. Por lo mismo se podría aproximar la función de correlación de una manera empírica de la siguiente forma:

$$\Gamma(\vec{r}) \sim r^{-\tau} \exp^{-r/\xi} \quad (8)$$

Donde  $\tau$  y  $\xi$  son parámetros ajustables dependiendo el modelo que se emplee [13]. Si se aproximan al punto crítico, empieza a incrementarse la correlación hasta llegar a infinito y la ecuación (8) diverge. Tanto la evidencia experimental como algunos modelos que dan la solución exacta del sistema demuestran que la función de correlación para modelos de espines decae siguiendo una ley de

potencias de la forma

$$\Gamma(\vec{r}) \sim \frac{1}{r^{d-2+\eta}} \quad (9)$$

Con  $d$  la dimensión del sistema y  $\eta$  una constante dependiente del sistema [9]. Esto es general: todo sistema que presenta una transición de fase continua sigue una ley de potencias tal cual como lo establece la ecuación (9) ([9], [23], [25]). Dentro de las muchas conclusiones a las que puede llevar este resultado, una de las más importantes es que dentro de una transición de fase continua existe invariancia de escala, lo que en términos espaciales indica que no importa que tan lejos o qué tan cerca se vea al sistema que se analiza (suponiendo que éste sistema es infinito), el sistema presenta fractalidad espacial, o lo que es lo mismo, el sistema presenta las mismas características estadísticamente sin importar qué tan lejos o qué tan cerca se vea al sistema.

### 1.1.1 Transiciones para sistemas magnéticos

En general, la teoría de transiciones de fase es muy similar independientemente de si el sistema es o no magnético; sin embargo, para muchos casos es más sencillo trabajar con sistemas magnéticos pues su interpretación es más sencilla. Para empezar, se define la magnetización como la densidad de momentos dipolares magnéticos que presenta un sistema, esto es:

$$\bar{M} = \frac{d\bar{m}}{dV} \quad (10)$$

Los momentos dipolares magnéticos pueden aparecer de manera intrínseca por las partículas que componen el sistema como es el caso de los espines o por el movimiento mismo de partículas cargadas, como es el caso del momento angular magnético ([10, 11]). Este movimiento de partículas cargadas, a su vez puede ser ocasionado de manera externa o puede existir intrínsecamente en el material como es el caso de los materiales ferromagnéticos. En general, cuando se habla

de transiciones de fase magnéticas se refiere a la transición del sistema mismo desde un punto donde no presenta magnetización hasta un punto en donde la presenta [8], es decir desde un punto donde  $M = 0$  a un punto donde  $M \neq 0$ . En este sentido, dependiendo de la propiedad intensiva que se varíe del sistema, habrá un punto crítico a partir del cual esto se cumpla. Si la variable que cambia es la temperatura, existe un valor crítico en la temperatura o una temperatura crítica  $T_c$  a partir de la cual se lleva a cabo la transición de fase.

El siguiente desarrollo mostrará la relación entre la función de correlación, la magnetización y la susceptibilidad magnética. Las fluctuaciones en la magnetización se pueden expresar como

$$\langle (M - \langle M \rangle)^2 \rangle = \langle M^2 \rangle - \langle M \rangle^2 = (kT)^2 \frac{\partial^2}{\partial H^2} \ln Z = kT \chi_T \quad (11)$$

En un sistema en donde la magnetización se debe solamente a los espines, se tiene que la magnetización total del sistema será la suma total de espines:

$$\langle (M - \langle M \rangle)^2 \rangle = \sum_i (s_i - \langle s_i \rangle) \sum_j (s_j - \langle s_j \rangle) = \sum_{ij} \Gamma_{ij} \quad (12)$$

Si el sistema es invariante ante traslaciones

$$\sum_{ij} \Gamma_{ij} = N \sum_i \Gamma_{i0} \sim N \int \Gamma(r) r^{d-1} dr \quad (13)$$

La última aproximación sólo es válida muy cerca del punto crítico ya que en ese punto, al ser todas las escalas importantes, la estructura de la red deja de ser relevante. Combinando las ecuaciones (10) y (12):



$$\chi_T \sim N \int \Gamma(r) r^{d-1} dr \quad (14)$$

Como en el punto crítico  $\chi_T$  diverge, la función de correlación debe aumentar tanto que la integral también diverge (lo que también nos dice que las fluctuaciones en la magnetización divergen). Esto a la vez pone un límite superior al valor de  $\eta$  en la ecuación (9). En efecto, ya que el sistema diverge en la integral el valor máximo que puede alcanzar el exponente es de  $\eta = 2$ . Nótese además que este mismo desarrollo es general para cualquier sistema que dependa de una variable (en este caso fue  $H$ ).

### 1.1.2 Exponentes Críticos

Se define, sólo como notación

$$t = (T - T_c)/T_c \quad (15)$$

lo cual es una medida que dice qué tan alejado se encuentra el sistema de la temperatura crítica. En este caso se puede definir un exponente asociado a la función  $F(t)$ :

$$\lambda = \lim_{t \rightarrow 0} \frac{\ln |F(t)|}{\ln |t|} \quad (16)$$

Siempre y cuando ese límite exista. Es decir, se puede aproximar a la función  $F(t)$  con una expresión polinomial con parámetros ajustables dependiendo de la información que se tenga del sistema.

$$F(t) = A|t|^\lambda (1 + bt^{\lambda_1} + \dots) \quad (17)$$

Con  $\lambda_i > 0$ . A los términos  $\lambda_i$  se les llama exponentes críticos. Se ha supuesto claramente que tales valores de los exponentes críticos son constantes sin considerar el que el sistema se encuentre por arriba o por abajo de  $T_c$ . Puede demostrarse que estos exponentes son constantes aunque esto va más allá de los fines de la presente tesis; de cualquier manera, en las referencias [9, 13] puede revisarse todo el desarrollo para llegar a esta conclusión. La importancia de los exponentes críticos radica en que en muchos casos dependen de muy pocos parámetros además de que son independientes de factores como las vibraciones atómicas (como es el caso de la temperatura); con esto, puede decirse que tales factores presentan un carácter universal pues dependen más del modelo que se emplea para describir el fenómeno que ocurre en el sistema que de las propiedades particulares del sistema mismo.

### 1.1.3 Métodos numéricos para la obtención de los exponentes críticos<sup>1</sup>

La única forma para obtener analíticamente el valor de los parámetros de orden es la aplicación del grupo de renormalización sin embargo, en muchas ocasiones resulta muy complicado si no es que imposible encontrar los valores de los parámetros por este método. Afortunadamente, existen muchos otros métodos para aproximar el valor de los parámetros de orden de manera numérica, los cuales se pueden englobar en 5 grandes categorías:

#### 1. Ajuste de los exponentes críticos a resultados experimentales:

los exponentes críticos son universales por lo que tienen el mismo valor para distintos sistemas que puedan ser descritos por el mismo modelo. Para emplear este método basta obtener una cantidad considerable de resultados experimen-

---

<sup>1</sup>Casi la totalidad de lo presentado en este apartado proviene de un resumen de las referencias [9, 10, 13], excepto cuando se indique lo contrario

tales y hacer el experimento con muchos materiales, posteriormente se trabaja en ajustar la ecuación (17) a todos los datos ([9, 13]).

**2. Teorías de campo medio:** En general todas las teorías de campo medio parten de la desigualdad de Bogoliubov:

$$F \leq \Phi = F_0 + \langle H - H_0 \rangle_0 \quad (18)$$

Donde  $F$  es la energía libre del sistema,  $H_0$  un hamiltoniano conocido,  $F_0$  la energía libre asociada a dicho hamiltoniano. La energía de campo medio se define minimizando  $\Phi$  respecto al parámetro variacional  $H_0$ , es decir:

$$F_{cm} = \min_{H_0}(\Phi) \quad (19)$$

Esto nos da la mejor aproximación posible a la verdadera energía libre para  $H_0$  ya que la desigualdad (18) dice que la energía libre  $F_{cm}$  estará siempre por arriba de  $F$ . A este grupo de soluciones numéricas corresponde también la teoría de Landau, el cual propone que la energía libre  $F_{cm}$  siempre es simétrica respecto al parámetro de orden con el que se trabaje. En este caso si es la magnetización, se usará la letra  $M$ , por lo que se puede hacer un desarrollo en series de potencias alrededor del 0 de la forma<sup>2</sup>:

$$F = F_0 + a_2 M^2 + a_4 M^4 \quad (20)$$

Llama la atención que gran cantidad del comportamiento del sistema pueda ser entendido con las teorías de campo medio (en particular con el método de Landau). Esto ocurre porque el valor de la magnetización que minimiza la

---

<sup>2</sup>Aunque de una manera general debería escribirse la ecuación con el resto de las potencias pares de  $m$ , se puede demostrar que valores mayores no alteran en gran medida el comportamiento crítico del sistema.

energía libre es a la vez una función singular de la expansión de coeficientes que depende del campo y de la temperatura (ver ecuación (20)). Dado que las teorías de campo medio ignoran las fluctuaciones, su alcance es muy limitado para sistemas con procesos que presentan gran cantidad de fluctuaciones como el caso del modelo de Ising. Para ver esto, aproximemos la energía libre de fluctuación como un parámetro inversamente proporcional a  $\xi^{-d}$ , con  $d$ , la dimensión en la que se encuentre el sistema que se estudia :

$$F_{fluc} \sim \xi^{-d} \sim |t|^{\nu d} \quad (21)$$

Con  $\nu$  un exponente crítico del sistema. Por otro lado, el calor específico dejando el campo magnético externo constante es de la forma  $C_H \sim |t|^{-\alpha}$ , con  $\alpha$  un exponente crítico; por lo que integrando dos veces puede verse que  $F \sim |t|^{2-\alpha}$ . Para que las fluctuaciones sean despreciables  $F \gg F_{fluc}$ , por lo que  $d\nu > 2 - \alpha$ , reescribiendo:

$$d > \frac{2 - \alpha}{\nu} \quad (22)$$

Para el modelo de Ising, puede demostrarse [8] que  $\alpha = 0$ ;  $\nu = 1/2$ , por lo que el límite dimensional para que los exponentes sean exactos al aplicar los métodos de campo medio sobre Ising es  $d > 4$ . Sin embargo, algunos aspectos del modelo pueden entenderse fácilmente a partir de las teorías de campo medio, cómo se verá en los capítulos 3 y 4.

**3. Ajuste de matrices de transferencia:** La idea general de este método es construir una matriz en la que vengan todos los estados de la función de partición del sistema. Las propiedades termodinámicas del sistema se obtienen entonces de diagonalizar la matriz. La energía libre por espín dependerá solamente del valor propio más grande y la longitud de correlación dependerá

solamente de los valores propios más altos. Para el modelo de Ising en 1d con un nivel finito de vecinos y un número finito de valores posibles de espín éste método da una solución exacta sin embargo, para dimensiones mayores la matriz se vuelve infinita y el método solamente tiene solución numérica aproximada.

**4. Expansión en series:** Ésta, junto con las teorías de campo medio, son las técnicas de solución aproximada más utilizadas dentro de la literatura científica, con la diferencia de que la expansión en series ha dado mejores resultados a través de los años para cualquier sistema termodinámico empleado [9]. La idea principal consiste en encontrar una manera sistemática de calcular las contribuciones a la función de partición y esperar que sucesivas aproximaciones pueden ser extrapoladas para que den información de las propiedades del sistema en el punto crítico. Al hacer las aproximaciones se toman por lo general dos casos:  $T \rightarrow 0$  y  $T \rightarrow \infty$ . Para analizar el punto crítico se espera que la expansión en series se comporte de tal manera que se pueda obtener información de las singularidades tanto en la aproximación cuando  $T \rightarrow 0$  como en la aproximación  $T \rightarrow \infty$ .

**5. Aproximaciones mediante métodos computacionales:** Es tal vez el mejor método actual para sistemas con grandes fluctuaciones (como el modelo de Ising) y se basa principalmente en generar algoritmos computacionales que reproduzcan la dinámica planteada por el hamiltoniano del sistema. Dentro de este tipo de métodos es fundamental el uso de las simulaciones Monte Carlo si se pretende poder simular modelos que presentan una dinámica particular. En la sección (1.3) se ahondará más sobre este método ya que es un tema fundamental para el desarrollo del presente trabajo.

## 1.2 El Modelo de Ising

El modelo de Ising es uno de los modelos más importantes y más estudiados de la física estadística. Fue propuesto por el físico Wilhem Lenz en 1920 el cual lo concibió como un problema para la tesis doctoral de su entonces alumno Ernst Ising [10]. Lenz pidió que Ising mostrara que el problema presentaba una transición de fase en 1d. Después de 4 años en trabajar arduamente en su tesis doctoral, Ising demostró que en 1d el sistema no presenta transición de fase a  $T > 0$  y por mera transitividad llegó a la conclusión de que el sistema no presentaba transición de fase ni en 2 ni en 3 dimensiones. Este resultado desmotivó a Ising, llevándolo a abandonar sus estudios en física estadística. Sin embargo, el modelo siguió siendo arduamente estudiado en los años posteriores. Debido a que Ising fue el primero en estudiarlo a fondo, el modelo resuelto en 1d fue denominado como “modelo de Ising” y por costumbre, años después, los casos de dimensiones mayores fueron llamados de la misma forma. Exactamente veinte años después de la presentación de la tesis doctoral de Ising, en 1944, el futuro premio nobel Lars Onsager resolvió de manera analítica el modelo de Ising en 2d y demostrando además que el sistema efectivamente presenta una transición de fase [14,16].

El modelo de Ising es un modelo que representa una red compuesta de microsistemas los cuales tienen un número finito de valores de magnetización (espín) permitidos y simétricos respecto a cero. La interacción de cada microsistema con sus alrededores es uno a uno y a su vez, ésta se da solamente entre vecinos cercanos de acuerdo a la siguiente regla: *“el microsistema analizado toma el valor de magnetización (espín) que minimice la energía en un sistema que se componga del microestado y sus vecinos cercanos”*.

Hay distintas formas de plantear el sistema de Ising dependiendo tanto de la dimensión en la que se plantea el sistema, de la cantidad de valores de espín

permitidos y otro conjunto de propiedades que se analizarán en el siguiente apartado. Esto hace que la forma del hamiltoniano cambie dependiendo de los supuestos. El modelo más sencillo en 2 dimensiones es aquel donde el espín sólo tiene dos valores  $s = \pm 1/2$ <sup>3</sup>. El hamiltoniano que describe al sistema es :

$$H = -J \sum_{\langle ij \rangle} s_i s_j - H \sum_i s_i \quad (23)$$

donde el primer término es la interacción uno a uno entre los espines y el segundo es el término paramagnético con  $H$  el campo externo aplicado.  $J$  es conocida como la constante de interacción y depende del material. Claramente si  $J > 0$ , y debido a la regla dicha anteriormente, el estado que minimiza la energía es aquel que favorece la uniformidad en el valor de la magnetización de cada microsistema, es decir, el material favorece el ferromagnetismo. De manera contraria, el valor  $J < 0$  favorece valores contrarios de magnetización en cada microsistema o lo que es lo mismo, favorece el antiferromagnetismo. En términos de espines, la mínima energía representa el hecho de que los espines se encuentren todos en la misma dirección si  $J > 0$  y favorece el que los espines estén desalineados entre sí si  $J < 0$ .

El modelo de Ising tiene solución exacta para las dimensiones  $d = 1$ . Para  $d = 2$  y  $d = \infty$  tiene también solución exacta si  $H = 0$  sin embargo, tanto si  $d = 2$  y  $H \neq 0$  como para  $d = 3$  y  $H$  cualquiera, se han desarrollado potentes aproximaciones numéricas que aproximan muy bien a los datos experimentales.

La importancia del modelo de Ising radica principalmente en los siguientes aspectos:

1. Es un modelo relativamente “sencillo” que presenta transiciones de fase

---

<sup>3</sup>En caso de otros valores de espines, el hamiltoniano agregaría más términos de interacción. En la presente tesis se trabajará solamente con el modelo de Ising con espín-1/2 y las simulaciones computacionales estarán hechas para el sistema en 2d.

(en  $d = 1$  se presenta a temperatura cero mientras que en  $d = 2, \infty$ , la presenta a temperaturas finitas). Además, la transición de fase muestra un comportamiento colectivo que se comprende bien y que tiene muy fuertes evidencias experimentales.

2. Es uno de los primeros modelos de transición de fase y además uno de los más conocidos de la física en general; a partir de él se desprenden otros modelos de transición de fase importantes como el XY y el de Heisenberg [9,13, 15]. Además, modelos en otras ramas de la ciencia como la sociología o la economía se han creado basándose en el modelo de Ising con muy buenos resultados [2-7].
3. Es también una forma de aplicar los métodos de los procesos estocásticos a la física estadística y de mostrar con esto una dinámica temporal aleatoria dentro de un sistema físico [15]. También es un modelo que muestra la formación de costras o islas (del inglés “coarsening”) en sistemas y, en el caso de  $d = 2$ , permite el análisis de la fractalidad tanto espacial como temporal en zonas cercanas a la temperatura crítica .

### **1.2.1 Algunos resultados importantes del modelo de Ising en $1d$ y $2d$ .**

Se puede trabajar con muchas variantes del modelo. Un aspecto importante es definir la cantidad de “primeros vecinos” y eso depende mucho de como se defina la red en con la que se pretenda trabajar. Si se considera una red periódica e infinita se puede, además, hablar de que el sistema es invariante traslacional. En caso de que se quieran poner algunas particularidades a la red,

---

<sup>4</sup>La mayoría de los resultados que aquí se presentan fueron obtenidos de las referencias [9, 10], [13], [15-17]



ésta podría perder su invariancia traslacional. La ventaja de que el sistema sea invariante ante traslaciones es que trabajar en una zona que contenga al espín  $i$ , es equivalente a trabajar en una zona que contenga al espín  $j$ , por ejemplo.

Una variante igual de importante es decidir si hay o no un campo externo y en caso de que haya, decidir la forma en la que actúa. Otro aspecto a considerar es el de definir los bordes del sistema: se pueden tomar como periódicos (lo que en un red de  $N$  espines en la dimensión  $d = 1$  quiere decir que el espín  $N+1$  es en realidad el espín 1) o como puntos aislados. Por último se debe decir si el material con el que se trabaja es ferromagnético o antiferromagnético en cuyo caso, se debe decir si el valor de  $J$  es positivo o negativo respectivamente y en cuyo caso decir que tan intenso es. Ya que la física en general es prácticamente la misma, se toman por lo general los valores más sencillos: La mínima cantidad de vecinos (dependiendo de la dimensión, el mínimo número de vecinos es dos veces el valor de la dimensión, con esto, si  $d = 1$ , la cantidad de vecinos cercanos se toma como 2; para  $d = 2$ , la cantidad de vecinos cercanos es 4, etc), red invariante ante traslaciones, campo externo nulo ( $H = 0$ ), bordes con condiciones periódicas y  $J$  con valor positivo. Aunque en realidad la intensidad de  $J$  no afecta mucho en los cálculos, muchos autores (si no es que la mayoría), deciden tomar  $J = 1$ . De igual forma, ya que muchos resultados importantes quedan en términos de  $J/k_B$ , en general el cociente es tomado como 1 además que los valores de espín se toman como  $\pm 1$  en vez de  $\pm 1/2$ . El presente trabajo no es una excepción y en adelante, a menos que se especifique lo contrario, se tomarán dichos supuestos para desarrollar el modelo.

**Modelo de Ising en 1d** Se empezará con el modelo de Ising en 1 dimensión. Para este modelo, todas las ecuaciones pueden obtenerse de manera explícita. La figura 1, muestra el modelo de Ising unidimensional.



**Figura 1.** Modelo de Ising en una dimensión. Las condiciones periódicas establecen que en un modelo de  $N$  espines, el espín  $N+1$  es el espín 1.

Usaremos, por ser el método más cómodo, el método de matrices de transferencia:

Imaginemos una cadena de  $N$  espines enlazados unos con otros. Los vecinos de cada espín son, por ende, el espín que tiene a la derecha y el que tiene a la izquierda. Para cumplir con las condiciones periódicas se necesita que:

$$s_{N+1} = s_1 \quad (24)$$

El hamiltoniano, tomándolo de la ecuación (22) es:

$$\mathcal{H} = -J \sum_{i=1}^N s_i s_{i+1} - \frac{H}{2} \sum_{i=1}^N (s_i + s_{i+1}) \quad (25)$$

Nótese que el último término se dividió entre dos para evitar la doble suma. En tal caso, la función de partición vendría dada por:

$$Z_N = \sum_{\{s_i\}} \exp \left[ k \sum s_i s_{i+1} + \frac{h}{2} \sum (s_i + s_{i+1}) \right] \quad (26)$$

Donde  $k = -J\beta$  y  $h = H\beta$ . Se define entonces:

$$\Lambda(s_i, s_{i+1}) = \exp \left[ k s_i s_{i+1} + \frac{h}{2} (s_i + s_{i+1}) \right] \quad (27)$$

Reescribiendo la función de partición como:

$$Z_N = \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \cdots \sum_{s_N=\pm 1} \prod_{i=1}^N \Lambda(s_i, s_{i+1}) \quad (28)$$

Por las propiedades de la función exponencial, se puede reducir la ecuación anterior a

$$Z_N = \sum_{s_1 = \pm 1} \Lambda^N(s_1, s_1) \quad (29)$$

Se propone la matriz de transferencia

$$\Lambda = \begin{pmatrix} \Lambda(+1, +1) & \Lambda(+1, -1) \\ \Lambda(-1, +1) & \Lambda(-1, -1) \end{pmatrix} = \begin{pmatrix} e^{k+h} & e^{-k} \\ e^{-k} & e^{k-h} \end{pmatrix} \quad (30)$$

Uniendo esto último con la ecuación (28), la función de partición no es más que la traza de la matriz de transferencia elevada a la N-ésima potencia:

$$Z_N = \text{Tr}(\Lambda^N) \quad (31)$$

El problema se vuelve entonces un problema de valores propios. No es difícil ver que:

$$\lambda_{\pm} = e^k \cosh(h) \pm \sqrt{e^{2k} \sinh^2(h) + e^{-2k}}$$

Reescribiendo

$$Z_N = \lambda_+^N \left[ 1 + \left( \frac{\lambda_-}{\lambda_+} \right)^N \right] \quad (32)$$

El segundo término dentro del paréntesis es  $< 1$ . Al aplicar entonces el límite termodinámico a la energía libre por partícula se tiene:

$$f(T, H) = \lim_{N \rightarrow \infty} \left[ -\frac{k_B T}{N} \ln Z_N \right] = -k_B T \ln \lambda_+ \quad (33)$$

Derivando una vez respecto a  $H$ , obtenemos la magnetización:

$$m(T, B) = \langle s \rangle = -\left(\frac{\partial f}{\partial H}\right)_T = \frac{e^k \sinh(h)}{\sqrt{e^{2k} \sinh^2(h) - e^{-2k}}} \quad (34)$$

La cual claramente es una función continua para toda  $T \neq 0$ . Esto quiere decir que el sistema no presenta transición de fase discontinua. Por otro lado, la segunda derivada de  $f$  respecto a la temperatura  $T$  nos da el calor específico a campo constante:

$$C_H = T\left(\frac{\partial s}{\partial T}\right) = -T\left(\frac{\partial^2 f}{\partial T^2}\right)_{H=0} = \frac{J^2}{k_B T^2} \left[ \frac{1}{\cosh^2\left(\frac{J}{k_B T}\right)} \right] \quad (35)$$

De nuevo, la función es continua en todo punto de  $T \neq 0$ , por lo que tampoco presenta una transición de fase continua.

Utilizando las ecuaciones de la sección (1.1), se puede obtener la longitud de correlación:

$$\xi^{-1} = -\ln \left\{ \frac{e^k \cosh(h) - \sqrt{e^{2k} \sinh^2(h) + e^{-2k}}}{e^k \cosh(h) + \sqrt{e^{2k} \sinh^2(h) + e^{-2k}}} \right\} \quad (36)$$

Véanse los aspectos importantes relacionados con el cambio de temperatura. Si  $T \rightarrow 0$ ,  $\beta \rightarrow \infty$  y por ende  $k$  y  $h$  también lo hacen. Dado que  $e^k \cosh(h) > \sqrt{e^{2k} \sinh^2(h) + e^{-2k}}$ , la expresión dentro de los corchetes de la ecuación (35) tiende a 1 y entonces  $\xi^{-1} \rightarrow 0$ , ocasionando un largo de correlación infinito a temperatura  $T = 0$ . Por ende el sistema presenta transición a  $T = 0$ . Si ahora  $T \rightarrow \infty$ , ocurre todo lo contrario y la expresión entre paréntesis tiende a infinito y  $\xi^{-1} \rightarrow 0$ .

**Modelo de Ising en 2d** Aquí se toman 4 vecinos cercanos (una malla cuadrada) tal cual lo indica la imagen (2) con  $H = 0$ . Dada la complejidad de los cálculos, sólo se muestran los resultados finales. Aplicando el límite termodinámico, la energía libre por espín es:

$$f(T) = \frac{1}{\beta} \left[ -\ln(2 \cosh 2k) - \frac{1}{2\pi} \int_0^\pi d\phi \ln \left[ \frac{1}{2} \left( 1 + \sqrt{1 - \kappa \sin^2 \phi} \right) \right] \right] \quad (37)$$

En esta última ecuación se usó de nuevo  $k = \beta J$  y además se define:

$$\kappa = \frac{2 \sinh(2k)}{\cosh^2(2k)} \quad (38)$$

La energía interna que se obtiene tiene la forma:

$$u(T) = -J \coth(2k) \left[ 1 + \frac{2}{\pi} \kappa' K_1(\kappa) \right] \quad (39)$$

Con  $\kappa'^2 = 1 - \kappa$  y

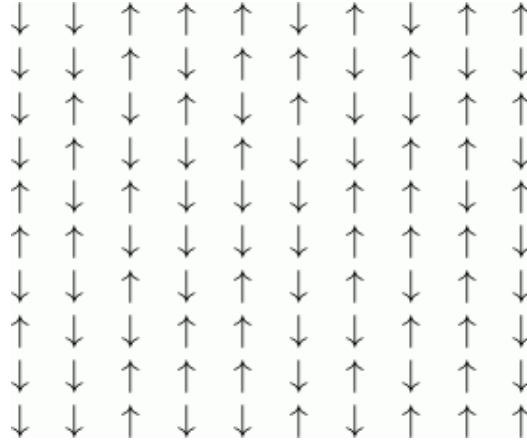
$$K_1(\kappa) \equiv \int_0^{\pi/2} \frac{d\phi}{\sqrt{1 - \kappa^2 \sin^2(\phi)}} \quad (40)$$

Que es una integral elíptica completa de primera especie. Derivando respecto a  $T$ , se obtiene el calor específico:

$$c(T) = \frac{2k_B}{\pi} (k \coth(2k))^2 \left\{ 2K_1(\kappa) - 2E_1(\kappa) - (1 - \kappa') \left[ \frac{\pi}{2} \kappa' K_1(\kappa) \right] \right\} \quad (41)$$

y se definió:

$$E_1(\kappa) \equiv \int_0^{\frac{\pi}{2}} d\phi \sqrt{1 - \kappa^2 \sin^2 \phi} \quad (42)$$



**Figura 2.** Modelo de Ising en 2 dimensiones con una malla cuadrada. Los vecinos cercanos son cuatro: derecha, izquierda, arriba y abajo. De nuevo, las condiciones periódicas establecen que en una malla cuadrada de  $N \times N$  espines, el espín en la posición  $(i, N+1)$  es el espín  $(i, 1)$  mientras que el espín  $(i, N+1)$  está representado por el  $(i, 1)$ .

Que es una integral elíptica de segunda especie. La integral  $K_1(\kappa)$  tiene una singularidad en  $\kappa = 1$  (o lo que es lo mismo,  $\kappa' = 0$ ), y la temperatura crítica sale de la condición:

$$\kappa = \frac{2 \sinh(2k_c)}{\cosh^2(2k_c)} = \frac{2 \sinh(2\beta_c J)}{\cosh^2(2\beta_c J)} = 1 \quad (43)$$

De donde resulta:

$$2 \tanh(2k_c) = 1 \quad (44)$$

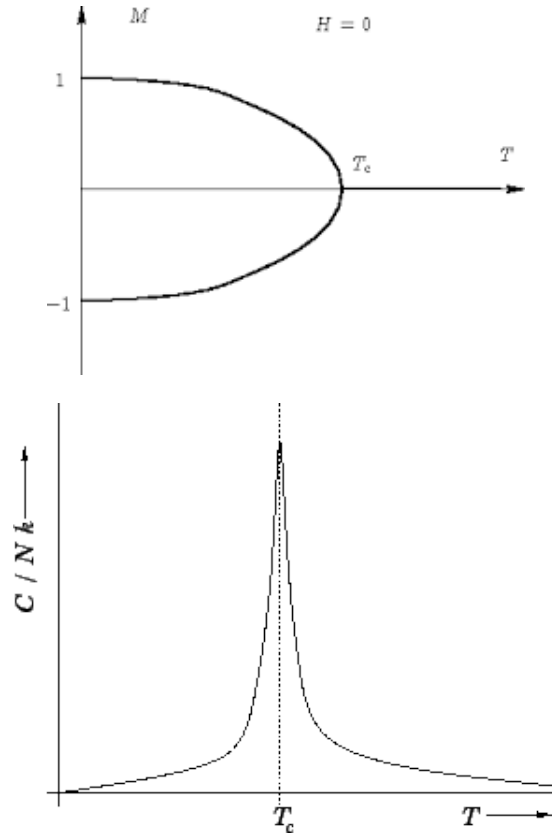
Aproximando se tiene:

$$T_c \approx \frac{J}{k_B} \left[ \frac{2}{\ln(1 + \sqrt{2})} \right] \approx 2.269185 \left( \frac{J}{k_B} \right) \quad (45)$$

Calculando el calor específico, y usando la ecuación anterior, se puede definir al mismo en términos de la temperatura crítica:

$$c(T) = \frac{2k_B}{\pi} \left( \frac{2J}{k_B T_c} \right)^2 \left[ -\ln\left(1 - \frac{T}{T_c}\right) + \ln\left(\frac{k_B T_c}{2J}\right) - \left(1 + \frac{\pi}{4}\right) \right] \quad (46)$$

En esta última expresión, puede verse que si  $T \rightarrow T_c$ , el calor específico se indefiniría, teniendo así una transición de fase continua. Veamos ahora el comportamiento de la magnetización:



**Figura 3.** Gráficas de a) la magnetización y b) del calor específico para el modelo de Ising en 2d (Imágenes tomadas de las referencias [42] y [43] respectivamente).

$$m(T) = \begin{cases} \{1 - [\sinh(2\beta J)]^{-4}\}^{1/8} & \text{si } T \leq T_c \\ 0 & \text{si } T > T_c \end{cases} \quad (47)$$

Lo cual indica la transición del orden al desorden si el campo magnético externo  $H = 0$ . Nótese que la función tiene una discontinuidad en la magnetización en el punto  $T = T_c$ , lo que también nos habla de una transición de fase. Las figuras 3a y 3b muestran las gráficas de la magnetización y del calor específico respectivamente.

### 1.2.2 Entre el orden y el desorden.

Los resultados de la sección anterior arrojan una conclusión por demás interesante: mientras que el modelo de Ising en una dimensión no presenta una transición de fase para temperaturas distintas a cero, el mismo modelo en dos dimensiones presenta una clara transición de fase a una temperatura distinta de cero. ¿Por qué sucede esto? Sabemos de antemano que un sistema en equilibrio debe minimizar su energía libre  $F$ . A temperatura cero, la energía libre de cualquier sistema es claramente igual a la energía total del sistema  $U$ . En un sistema de interacción magnética como lo es el modelo de Ising, la energía total se minimiza cuando todos los momentos magnéticos son paralelos. Por ende la configuración preferida del sistema es el orden.

Supongamos ahora lo contrario, un sistema con temperatura infinita. La entropía tiene que alcanzar un máximo, es decir, todos los momentos magnéticos están colocados al azar. ¿Esto nos dice que existe un punto medio en donde el orden y el desorden se estabilizan? No necesariamente ya que para que exista tal punto medio, resulta necesario que la longitud de correlación se pueda mantener estable respecto a las fluctuaciones (o cambios de momentos magnéticos)



del sistema. En efecto, supongamos que tenemos  $N$  casillas y la longitud de correlación es igual o menor al tamaño de una casilla. En tal caso, si vamos de un sistema absolutamente ordenado e incrementamos la temperatura, por más que cambiemos el momento magnético de algunas casillas del sistema, esta fluctuación no podrá transmitirse y el sistema quedará prácticamente igual. Si por el contrario vamos del desorden que proporciona la temperatura infinita hasta una temperatura cero, el sistema tampoco tendrá la capacidad de transmitir el hecho de que es preferible estar completamente alineados. Es así que el modelo de Ising es sorprendente pues puede mostrar, si las condiciones así lo permiten, un punto intermedio en donde el orden y el desorden encuentran un equilibrio.

Nótese que el argumento aquí expuesto es completamente independiente de la dimensión y que sólo se habla de un sistema magnético más no específicamente del modelo de Ising. Hay un bello y sencillo argumento con el que se puede mostrar se que el modelo de Ising en 1d no tiene transición de fase a partir de pensar solamente en la energía libre del sistema y con esto evitar el pesado camino que se recorrió en la sección anterior<sup>5</sup>:

Tomemos un modelo de Ising con  $N$  espines acomodados de manera lineal apuntando todos hacia la misma dirección y sin que se aplique un campo magnético externo con condiciones periódicas. Cuando el sistema está completamente ordenado, la energía es:

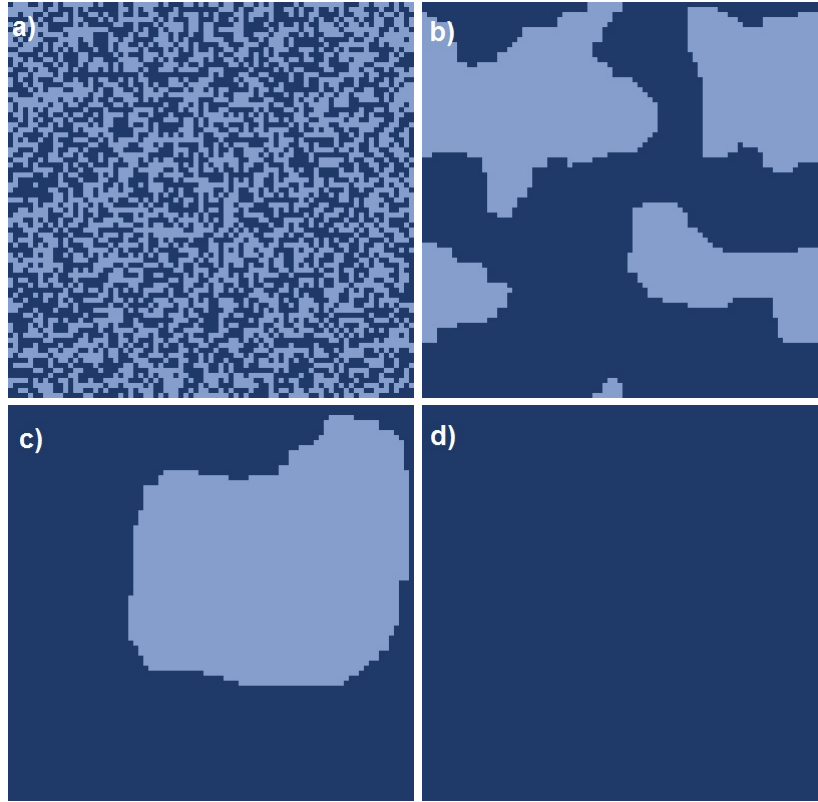
$$U_0 = -JN \tag{48}$$

Como hay dos posibles configuraciones en este estado (todos los espines hacia arriba o todos hacia abajo), la entropía total tiene valor:

$$S_0 = \ln 2 \tag{49}$$

---

<sup>5</sup>Aunque este argumento se toma de [11], se le atribuye originalmente a Landau



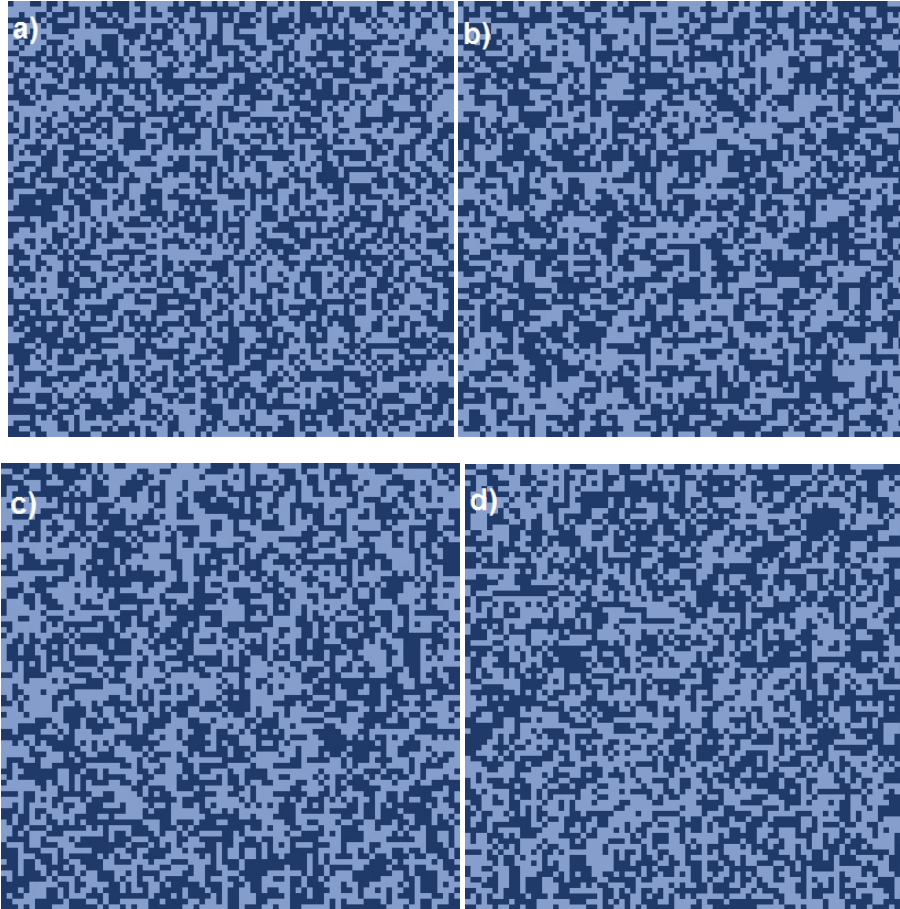
**Figura 4.** Evolución temporal del modelo se Ising por computadora usando el método Monte Carlo. La temperatura en este caso es 0. La imagen a) muestra la configuración inicial del sistema (al azar), la imagen b) muestra al mismo sistema pero 2 millones de unidades de tiempo (pasos) después, la c) muestra 4 millones de pasos después y por último la d) muestra al sistema a 6 millones de pasos. Nótese como todos los espines tienden a alinearse hacia una misma dirección.

Por ende, la energía libre del sistema es:

$$F_0 = -JN - T \ln 2 \quad (50)$$

Tomamos al azar un bloque de dos espines y lo volteamos, de tal manera que la energía total del sistema vale ahora:

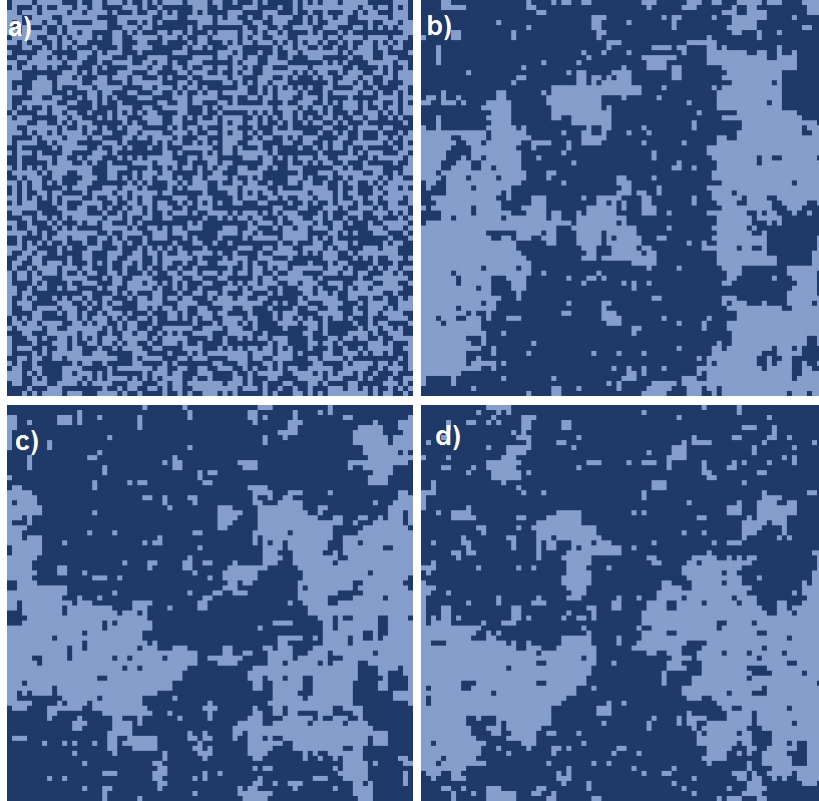
$$U_2 = -J(N - 2) + 2J = -J(N - 4) \quad (51)$$



**Figura 5.** El sistema ahora se muestra a  $T=10$ . En este caso las imágenes b) (2 millones de pasos), c) (4 millones) y d) (6 millones) parecen tan al azar como la configuración inicial a) .

Y dado que que podemos colocar esta alteración donde sea y son dos estados degenerados, la entropía del sistema es:

$$S_2 = \ln(2N) \quad (52)$$



**Figura 6.** Sistema a  $T = T_C$ . Al igual que en los otros casos, la imagen a) es la configuración inicial, la b) es la configuración 2 millones de pasos después, la c) es 4 millones y la d) 6 millones. Puede verse en las imágenes b), c) y d) la formación de islas de una configuración de espines y dentro de éstas se encuentran islas más pequeñas de la configuración opuesta.

Finalmente la energía libre es entonces:

$$F_2 = -J(N - 4) - T \ln(2N) \quad (53)$$

La variación de la energía libre tiene la forma:

$$\Delta F = F_2 - F_0 = 4J - T \ln(N) \quad (54)$$

Si tomamos el límite termodinámico haciendo tender  $N \rightarrow \infty$ , se tiene que  $\Delta F < 0$ . Así, aunque hay un incremento en la energía interna del sistema, el incremento en la entropía es tal que la energía libre es más favorable a cualquier temperatura finita  $T > 0$ . Por ende, voltear espines es favorable y el orden es destruido a cualquier temperatura finita, es decir, no hay una temperatura crítica  $T > 0$  donde ocurra una transición de fase. En cambio, este comportamiento no se muestra en el modelo de Ising en dos dimensiones por lo que Ising en 2d si tiene transición de fase a una temperatura finita distinta de cero.

Las figuras 4, 5 y 6 representan la evolución temporal de la configuración de espines del modelo de Ising a temperatura  $T = 0$ ,  $T = T_c$  y  $T = 10$  a distintos tiempos cada una.

### 1.3 Simulaciones Monte Carlo

En física estadística, se define el observable de una cantidad como el valor promedio de dicha cantidad respecto a todos los valores que pueda tener ésta:

$$\langle A \rangle = \frac{\sum_{\{s\}} A e^{-\beta H}}{\sum_{\{s\}} e^{-\beta H}} \quad (55)$$

Claramente, dado que cada espín en el modelo de Ising de espín 1/2 puede tener dos valores (arriba-abajo), la cantidad de estados que puede tener el

sistema es de  $2^N$ . Al depender tan fuertemente del valor  $N$ , es difícil poder tener una medida de los observables únicamente resolviendo la ecuación (55). Para calcular los observables del sistema de una manera sencilla y sin tener que evaluar toda la ecuación (55), se recurre al método de simulación Monte Carlo.

El método de simulaciones Monte Carlo<sup>6</sup> (MCS por sus siglas en inglés) fue uno de los primeros modelos de simulación propuestos y dada su sencillez, es uno de los más empleados en la actualidad. La idea general del método consiste en crear un algoritmo que simule la dinámica del sistema usando cadenas de Markov donde cada probabilidad de transición obedezca tanto la forma del Hamiltoniano como la ley de minimización de la energía (o maximización de la entropía) [9,15]. Además, dicha cadena de Markov debe cumplir con las siguientes propiedades:

1. Satisfacer la ecuación de balance detallado:

$$p^{eq}(a)p_{ab} = p^{eq}(b)p_{ba} \quad (56)$$

Donde  $p_{ab}$  denota la probabilidad de ir del estado  $a$  al estado  $b$  y  $p^{eq}(x)$  es la probabilidad de que, estando en el equilibrio, el sistema se encuentre en el estado  $x$ .

2. Obedecer el teorema ergódico, es decir, que el sistema sea capaz de alcanzar todos los estados posibles y que estos estados no sean recursivos.

En la referencia [15] puede encontrarse un resultado importante para el método Monte Carlo: cualquier sistema dinámico discreto con un número finito de estados descrito por una cadena de Markov que cumple tanto con el teorema ergódico como con la ecuación de balance detallado (ecuación (56)) tiende a una distribución de equilibrio. Esto asegura que para tiempos muy grandes:

$$\bar{A} \approx \langle A \rangle \quad (57)$$

---

<sup>6</sup>El método Monte Carlo que se va a describir es el método discreto aunque puede ser generalizado a una forma continua.

Lo que quiere decir que después de un tiempo cada ligero cambio en el sistema describa correctamente al sistema en equilibrio. Para que esto suceda, se tiene que esperar el tiempo suficiente para que los estados del sistema dejen de estar lo suficientemente correlacionados y por ende sean independientes unos de otros.

### 1.3.1 Algoritmo de metrópolis

El hecho de que se encuentre una cadena de Markov que describa al sistema y que cumpla con las 2 propiedades del apartado anterior (ser ergódica y satisfacer las ecuaciones de balance detallado) no quiere decir que ésta sea la única cadena de Markov que pueda emplearse. Dependiendo del sistema hay una gran cantidad de cadenas que pueden describirlo y que cumplen con las propiedades anteriormente detalladas. Dado que las probabilidades de transición definen a la cadena de Markov y éstas a su vez definirán el algoritmo computacional que simulará al sistema, cada tipo de cadena de Markov que se emplee es un algoritmo computacional distinto. Por supuesto, hay algoritmos que se han utilizado con éxito durante muchos años y dentro de ellos el más famoso es el algoritmo de Metrópolis<sup>7</sup>.

Dicho algoritmo consiste en definir las probabilidades de transición de un estado inicial  $\{s\}_i$  a un estado final  $\{s\}_f$  como:

$$P(\{s\}_i \rightarrow \{s\}_f) = \begin{cases} e^{-\beta(E_f - E_i)} & \text{si } E_f > E_i \\ 1 & \text{si } E_f \leq E_i \end{cases} \quad (58)$$

Con  $E_i$  la energía del estado inicial y  $E_f$  la del estado final.

Se explicará brevemente como es que trabajará la cadena de Markov propuesta por el algoritmo de metrópolis usando un sistema de espines:

---

<sup>7</sup>Una explicación más detallada de esto puede encontrarse en [8, 9], aquí sólo se muestran las características más importantes

1. Se toma un espín de la red (puede ser al azar o siguiendo un patrón particular) y cambia su estado actual. Posteriormente se calcula  $r = e^{-\Delta E/kT}$ , donde la energía final se elige al azar entre las posibilidades de cambio del espín (en un sistema de espín 1/2, la única posibilidad de cambio es voltear el espín, por lo que se calcula el cambio en la energía como el cambio energético por voltear el espín, sin embargo en otros modelos puede haber más opciones).
2. Si el cambio en energía es negativo, el espín cambia su estado. Si el cambio resulta positivo, se genera un número al azar  $z$  entre 0 y 1.
3. El espín cambia su estado si  $z < r$ , en otro se queda en la posición original.
4. Se repiten los pasos 1-3 con el nuevo sistema de configuración de espines, guardando en cada paso el valor promedio de la cantidad la magnetización.

### 1.3.2 Tiempo de relajación de la simulación

¿Cuántos pasos son necesarios para el análisis del modelo? Si queremos medir las propiedades termodinámicas del sistema, es necesario que el sistema llegue a un equilibrio, de lo contrario no dará de forma correcta los resultados esperados. El tiempo  $\tau$  que tarda el sistema en partir de una configuración aleatoria a una de equilibrio se le conoce como “tiempo de relajación” (del inglés relaxation time) y depende fuertemente del número de espines del sistema. Cabe aclarar también que si el sistema es tratado cerca del punto crítico al desequilibrio inicial se le tiene que sumar el desequilibrio causado por el aumento en la longitud de correlación. En efecto, ya que la longitud de correlación aumenta considerablemente cerca de un punto crítico, una alteración del sistema empieza a afectar cada vez a más espines a su alrededor, lo que aleja cada vez más al sistema con respecto al equilibrio. Esto ocasiona que en el punto crítico, el tiempo de relajación sea infinito y por ende en la simulación computacional la dispersión en los datos sea infinita, es decir, no se pueden tener valores concretos de las



propiedades termodinámicas del sistema.

En general, el tiempo de relajación  $\tau$  guarda la siguiente relación con la longitud de correlación:

$$\tau \sim \xi^z \tag{59}$$

Con  $z \sim 2$  para el modelo de Ising. Diversos métodos alternos son propuestos para poder reducir el tiempo de relajación del sistema, sin embargo, tales métodos resultan más tardados (tanto en programar como en computar) y complejos [19].

Por supuesto que en el sistema se pueden tomar en cuenta otros errores importantes si se quieren obtener valores muy precisos, tales como el hecho de trabajar con una configuración finita y en tal caso se tiene que agregar un término de error asociado a no considerar los efectos de tomar los bordes periódicos o tomarlos en cuenta desde un principio. En todo caso, basta con que éstos errores sean menores a los errores estadísticos para que no sean tomados en cuenta.

Se puede demostrar también que mientras más próximo se esté al punto crítico mayor es el tiempo de relajación  $\tau$  e incluso, el método resulta inservible exactamente en el punto crítico, por lo que es inservible para medir las propiedades termodinámicas del sistema. Esto se debe principalmente a que la correlación es tan alta que no se pueden obtener dos estados independientes sin importar la escala temporal que se utilice. De cualquier manera, es importante recalcar el punto de que esto es relevante sólo si se quieren medir las propiedades termodinámicas del sistema más no si se va a trabajar con la evolución temporal del mismo. En la presente tesis se pretende trabajar con las series de tiempo emanadas de la evolución temporal del sistema. Para esto, cabe decir que el sistema de Ising no contiene en sí mismo un cambio temporal, es decir las

ecuaciones del sistema que se presentan en las secciones anteriores no presentan en sí mismas variaciones temporales. En este sentido el método Monte Carlo no describe la dinámica del modelo de Ising, sin embargo, con el propósito de hacer valer la primera ley de la termodinámica, se puede inferir que el modelo de Ising presentará variaciones que minimicen la energía. El método Monte Carlo toma esto en cuenta para poder inferir una dinámica al modelo y convertir a Ising en un modelo dinámico con variaciones estocásticas en el momento de elegir si sí o no un espín se voltea y que se vuelven cada vez más importantes mientras la temperatura se aleja del cero.

Como el propósito último de esta tesis no es medir las propiedades termodinámicas del sistema, no será necesario esperar a que el sistema alcance el tiempo de relajación para que se pueda trabajar con él. De hecho, la importancia de este trabajo radica precisamente en esto: observar la evolución temporal del sistema haciendo un análisis entre configuraciones cercanas en el tiempo ya que es esto lo que produce las alertas tempranas. Por lo mismo se priorizará el análisis de la reconfiguración temporal del sistema a partir de alguna configuración inicial aleatoria dada una temperatura fija. En este sentido el modelo de Ising cumple una particularidad muy importante: la configuración espacial se refleja al mismo tiempo en el valor temporal de la magnetización, es decir, lo que sucede “visualmente” con el sistema sucede con el valor temporal de la magnetización. Esto es claro, si viendo el sistema podemos ver que muchos espines están “volteándose hacia abajo (toman un valor negativo)” claramente la serie de tiempo de la magnetización deberá indicar que el sistema en promedio es más negativo que positivo con respecto a un tiempo anterior. Este hecho, por más que obvio, es la base fundamental del trabajo y ayudará mucho a comprender el por qué de muchos cambios temporales en el sistema. Tanto en los resultados como en la discusión y las conclusiones se hará especial énfasis en este punto.

# Capítulo 2

## 2 Análisis de Series de Tiempo y Alertas Tempranas.

### 2.1 Series de tiempo

Una serie de tiempo es una colección de datos observados de manera secuencial a través del tiempo. Ejemplos de series de tiempo se encuentran en prácticamente cualquier rama de la ciencia y el análisis de las mismas comprenden muchas ramas del análisis funcional y de la estadística siendo el análisis de series de tiempo (AST ó más comúnmente en la literatura TSA, por sus siglas en inglés) la rama de la ciencia que se encarga de su estudio. Los objetivos del AST son primordialmente los siguientes:

1. Descripción de fenómenos dinámicos que presente el sistema
2. Explicación del porqué de sus variaciones temporales
3. Predicción de su futuro comportamiento
4. Control absoluto de la dinámica del sistema

Las técnicas empleadas en su análisis son múltiples y dependen mucho de la cantidad de datos disponibles, la estacionariedad de los mismos, su distribución espectral y la correlación entre ellos, por lo que dada una serie de tiempo no hay una técnica única con la que puede analizarse. En las referencias [20-22] pueden encontrarse una gran cantidad de análisis distintos que pueden ser aplicados a una serie de tiempo; en cualquier caso, una gran cantidad de métodos se basa en

analizar como se distribuyen las frecuencias de las series de tiempo independientemente de lo que se quiera obtener de éstas. Por ende muchos de los análisis se hacen dentro del espacio de las frecuencias en vez de el espacio temporal ya que dentro del primero es posible determinar si hay algún patrón dentro de la serie de tiempo. Más que eso, el teorema de Wiener-Kinchin demuestra que el espacio de análisis de frecuencias (en específico, el espectro de potencias) y la autocorrelación guardan cierta relación entre sí por lo que para muchos sistemas el análisis de series de tiempo implica, en el fondo, un análisis mismo de las correlaciones dentro del sistema. En el siguiente apartado se analizará qué es el espectro de potencias, algunas particularidades teóricas y prácticas para la obtención de este espectro y finalmente se ahondará en cómo el espectro de potencias se relaciona con la función de correlación. Posteriormente, dentro de ese mismo apartado se hará un pequeño análisis del espectro de potencias para el modelo de Ising en 2d.

## 2.2 El espectro de potencias

El espectro de potencias es una transformación de una serie de tiempo en términos de frecuencias de senos y cosenos que tiene como fin el poder analizar la relevancia de ciertas frecuencias en la misma serie. En este sentido, el espectro de potencias trata de responder a la pregunta ¿qué tanto aporta una frecuencia  $\omega$  a la periodicidad de la serie? Con esto en mente se pensarán dos ejemplos sencillos. Supóngase que se tiene una nota musical pura (sin armónicos) por lo que su ecuación quedará descrita de la forma  $\text{sen}(kx + \omega_0 t)$ . En tal caso, la única frecuencia de la nota es  $\omega_0$ , por lo que el espectro de potencias será una delta de Dirac de la forma  $\delta(\omega - \omega_0)$  sobre el espacio de las frecuencias. Por otro lado supóngase ruido blanco. Aquí todas las frecuencias son importantes por lo

que ninguna resalta encima de otra, dando como resultado que el espectro será una gran “mancha” en donde todas las frecuencias son representadas de igual manera en términos estadísticos.

Suponiendo que se tiene un conjunto de datos  $y(t)$  infinitos y continuos, el espectro de potencias será proporcional al cuadrado del valor absoluto de la transformada de Fourier del conjunto de datos:

$$P(\omega) \propto |\tilde{y}(\omega)|^2 = \left| \int_{-\infty}^{\infty} y(t)e^{-i\omega t} dt \right|^2 \quad (60)$$

sin embargo para cuestiones prácticas, el tener un conjunto de datos “infinitos y continuos” es imposible. Una condición más aceptable sería el tener un total de  $N$  datos sobre un intervalo de tiempo de longitud  $T$ . Uno estaría tentado a definir una transformada de Fourier “discreta”, cambiando la integral por suma y completando los “espacios” en donde no hay datos continuos con repeticiones “continuas” del mismo dato durante intervalos de tiempo de tamaño  $\Delta$ . En tal caso, la transformada de Fourier “discreta” tomaría la forma:

$$\tilde{y}_k = \sum_{j=0}^{N-1} y_j \exp(i\omega_k t_j) \quad (61)$$

donde  $\omega_k = 2\pi k/T$  y  $t_j = j\Delta$ . Para extender el resultado a valores negativos, como lo exige el espectro de potencias, se podría hacer, sin pérdida de generalidad  $\tilde{y}_k = \tilde{y}_{-k}$  y con fines de mantener la periodicidad y tener un número infinito de datos,  $\tilde{y}_k = \tilde{y}_{N+k}$ .

Con esto, el espectro de potencias para datos discretos sería:

$$P(\omega) \propto \left| \sum_{j=0}^{N-1} y_j \exp(i\omega_k t_j) \right|^2 \quad (62)$$

Nótese que la igualdad se da con una normalización apropiada al conjunto de datos que se tienen. En el presente caso, la elección de la normalización se toma

de tal manera que  $P(\omega)$  sea el cuadrado de la amplitud promedio de la señal. En general, la elección de cómo normalizar y cómo discretizar la transformada de Fourier define los distintos métodos para calcular el espectro de potencias. El que aquí se presenta, aunque no es el único, es el más usado en la literatura. Otros métodos pueden consultarse en las referencias [20-22].

¿Por qué puede hacerse esta aproximación que a primera vista parece muy intuitiva? Para analizar esto supóngase que se tiene una función  $y(t)$  con una transformada de Fourier asociada  $\tilde{y}(\omega)$ . Tal función se modificará para crear una nueva función  $\bar{y}(t)$  la cual tendrá las siguientes características: 1) sólo reflejar un intervalo de longitud  $T$  de la función total; 2) estar definida de  $+\infty$  a  $-\infty$ , 2) continua por tramos de longitud  $\Delta \ll T$  y con la condición de que  $\bar{y}(n\Delta) = y(n\Delta)$ . Esta transformación es lo que en el fondo se propone cuando se aplica la transformada de Fourier a un conjunto discreto de datos. La función que se construye es la siguiente<sup>8</sup>:

$$\bar{y}(t) = [y(t)H(t, T) \otimes S(t, T)] S(t, \Delta) \quad (63)$$

Aquí,  $H(t, T)$  es el filtro *top hat* (la función que vale 1 en la región  $0 < t < T$ , y 0 en cualquier otro punto),  $\otimes$  es el operador de convolución y  $S(t, T)$  es la función *spike*, que vale cero en todo punto que no sea múltiplo entero de  $T$ .

Analícese la ecuación (62). La convolución de  $y(t)H(t, T)$  con  $S(t, T)$  repite periódicamente el primer término (con un periodo  $T$ ), es decir, da una cantidad infinita de datos sólo reflejando un intervalo de tamaño  $T$  respecto a la función original. Por otro lado, la última multiplicación por  $S(t, \Delta)$  alarga por un intervalo  $\Delta$  el valor de los datos, que era la condición que faltaba cumplir. En tal caso, el espectro de potencias de los nuevos datos “modificados” vendría dado por la transformada de Fourier de la ecuación (4), esto es:

---

<sup>8</sup>Este análisis fue tomado de la referencia [22]

$$\bar{y}(\omega) \sim \left[ (\tilde{y}(\omega) \otimes \tilde{H}(\omega, \frac{2\pi}{T})) S(\omega, \frac{2\pi}{T}) \right] \otimes S(\omega, \frac{2\pi}{\Delta}) \quad (64)$$

Con  $\tilde{H}(\omega, \frac{2\pi}{T}) = \text{sen}(\frac{1}{2}\omega T)/(\frac{\omega}{2})$ , que es un seno periódico con “colas” que descienden como  $1/\omega$  y que alcanza un valor máximo en  $\omega = 0$ . Véase entonces que la transformada de Fourier de la ecuación (62) es similar a convolucionar la transformada de la función  $y(t)$  y insertar el resultado a cada intervalo de tamaño  $\frac{2\pi}{T}$  para finalmente repetirlo cada intervalo de longitud  $\frac{2\pi}{\Delta}$ . Sin pérdida de generalidad, esto quiere decir que se puede reducir el intervalo a analizar a cualquiera de tamaño  $\frac{2\pi}{\Delta}$ , ya que el resto serán repeticiones de lo mismo. Si el término  $1/\omega$  desciende lo bastante rápido de tal manera que sea despreciable cuando termina el intervalo de longitud  $\frac{2\pi}{T}$ , entonces las ecuaciones (61) y (64) serán prácticamente las mismas, siempre y cuando se tome en cuenta que: 1) el tiempo entre mediciones debe de ser demasiado pequeño para que el término  $\pi/\Delta$  (encontrado muchas veces en la literatura como “frecuencia de Nyquist”), sea tan pequeño que puedan ser minimizados los efectos de borde al hacer infinitos los datos, 2) que efectivamente el término  $1/\omega$  sea despreciable cuando termine el intervalo de tamaño  $\frac{2\pi}{T}$ <sup>9</sup>, lo cual puede hacerse tomando datos durante un intervalo de tiempo  $T$  bastante largo.

Otro punto a tomarse en cuenta sobre todo con datos que distan muchos de ser periódicos es que la aproximación del espectro dado por la ecuación (61) pueden agregar mucho “ruido” (frecuencias poco importantes en los datos reales que obtienen, dada las limitaciones de la aproximación, altos valores de amplitud) a la medición. Para eliminar el ruido se recomienda realizar varias veces la toma de datos y promediar el espectro obtenido. Claramente esto puede

---

<sup>9</sup>Sólo como comentario: para muchos casos, el término  $1/\omega$  no llega a converger al final de intervalo por lo que se proponen otras funciones que lleguen a converger mucho más rápido y con esto poder seguir usando la misma idea de transformada de Fourier discreta. A este tipo de procedimientos se le denomina transformadas de ventaneo de Fourier (“*Fourier windowed transforms*”).

cumplirse para eventos repetibles en cantidades enormes como es el caso de las simulaciones en computadora.

### 2.2.1 Teorema de Wiener-Khinchin

El teorema de Wiener-Kinchin muestra de qué manera se liga la correlación con el espectro de potencias. De antemano, era de esperarse que existiera alguna relación entre estas dos variables ya que ambas provienen de una medición temporal del sistema. En el caso de la correlación, se trata de cuánto ha cambiado el sistema con respecto a sus estados anteriores, mientras que una serie de tiempo mide simplemente las propiedades del sistema durante un periodo determinado de tiempo. El enunciado del teorema es el siguiente:

“Considérese un proceso aleatorio en el tiempo con función de correlación definida como:

$$C(\tau) = \langle x(t) x(t + \tau) \rangle \quad (65)$$

Suponiendo que el proceso estacionario débil (i.e. los valores de promedio pueden considerarse independientes del tiempo) y que la función  $C(\tau)$  tenga transformada de Fourier entonces, el espectro de potencias del sistema será:

$$P(\omega) = \int_{-\infty}^{\infty} C(\tau) e^{i\omega\tau} d\tau = \tilde{C}(\omega) \quad (66)$$

Cabe aclarar que este enunciado es para sistemas continuos. Para sistemas discretos (como es el caso más común en la práctica, donde se tiene un conjunto discreto de datos) la demostración es un poco más complicada matemáticamente pero el trasfondo físico es el mismo. La demostración para sistemas discretos puede ser consultada en la referencia [22] aunque la siguiente ecuación muestra la igualdad para un conjunto de datos discretos:



$$P(\omega) = \sum_{n=-\infty}^{\infty} C(\tau) e^{i\omega n} \quad (67)$$

En términos prácticos, el valor de  $\tau$  es conocido como el retraso o “lag” con respecto al que se compara el sistema. Por ejemplo, tomando mediciones temporales de manera discreta, se pueden comparar los datos actuales con el conjunto de datos inmediatamente anteriores, a lo que se dice que se calcula la correlación a  $\tau = 1$ . Por otro lado, el comparar los datos con respecto a dos unidades de tiempo anteriores se le llama correlación a  $\tau = 2$ , etc. Aunque es posible comparar respecto a cualquier valor de  $\tau$  (siempre y cuando se tengan datos durante un tiempo  $t \gg \tau$ ), para el análisis de alertas tempranas en el presente trabajo se usará el valor de  $\tau = 1$ .

Para la demostración del teorema piense en lo siguiente: al ser datos independientes, obtener el promedio es equivalente a obtener la suma de la multiplicación de todos los datos de la serie de tiempo que se encuentren separados por una distancia  $\tau$ , es decir, no hay términos cruzados. Por otro lado, dado que cada dato proviene de la transformada inversa de Fourier sobre el espacio de las frecuencias entonces cada dato representaría por sí mismo una suma sobre el espacio de las frecuencias. Se tiene entonces la suma sobre todos los datos de la multiplicación de la transformada inversa de Fourier por otra transformada inversa de Fourier separada por una distancia temporal  $\tau$ . Como es sobre todos los datos, se podrían acomodar los datos de tal suerte que se pueda tener que la correlación se equivalente a la transformada inversa de Fourier al cuadrado. Aplicando la transformada de Fourier, se tendría entonces que el espectro de potencias es igual a la transformada de Fourier de la correlación. Todo esto en términos matemáticos sería lo siguiente:

$$C(\tau) = \langle x(t) x(t + \tau) \rangle = \int_{-\infty}^{\infty} x(t) x(t + \tau) dt$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} \tilde{x}(\omega) e^{-\pi i \omega \tau} d\omega \right] \left[ \int_{-\infty}^{\infty} \tilde{x}(\omega) e^{-\pi i \omega' (t+\tau)} d\omega' \right] dt \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\tilde{x}(\omega)|^2 \delta(\omega' - \omega) e^{-2\pi i \omega' t} d\omega' d\omega \\
&C(\tau) = \int_{-\infty}^{\infty} |\tilde{x}(\omega)|^2 e^{-2\pi i \omega t} d\omega \tag{68}
\end{aligned}$$

Aplicando la transformada inversa a esta última ecuación, se llega a la ecuación (66).

Nótese que la hipótesis de independencia solamente fue usada al inicio y aún suponiendo que los datos no fueran independientes, el espectro de potencias seguiría siendo proporcional a la integral de la correlación.

### 2.2.2 Análisis espectral para el modelo de Ising

Se examinará cómo es el espectro de potencias para el modelo de Ising en 2d de manera teórica y se hará una corroboración práctica de lo mismo. Lo que se presenta aquí es un resumen sencillo pero completo del análisis que se hace en las referencias [23-25]. Se empezará empleando la aproximación a cuarto orden de la teoría de campo medio vista en el capítulo anterior. La ecuación resultante es:

$$\frac{\partial \phi}{\partial t} = \Gamma \nabla^2 \phi - A \phi - g \phi^3 + \varepsilon \tag{69}$$

Donde  $A = T - T_c$ ,  $\Gamma$  y  $g$  son constantes del sistema,  $\phi$  es la magnetización local del sistema y  $\varepsilon$  se asocia con el ruido blanco de manera que su promedio sea directamente proporcional a la temperatura, es decir:

$$\langle \varepsilon(x, t)\varepsilon(x', t') \rangle = \Gamma k_B T \delta^d(x - x') \delta(t - t') \quad (70)$$

Ahora, si la magnetización total de un sistema la definimos de la siguiente manera:

$$m(t) = \frac{1}{V} \int_V d^d x \phi(x, t) \quad (71)$$

Entonces el espectro de potencias será:

$$P(\omega) = \int dt e^{i\omega t} \langle m(t)m(0) \rangle \quad (72)$$

Aquí se emplea la definición de la función de auto-correlación dada en el capítulo 1 y en la sección anterior. Ésta en el punto crítico tiene la forma general:

$$C(x, t) = x^{-(d-2+\eta)} \bar{g}(t/x^z) \quad (73)$$

Donde  $z$  y  $\eta$  son exponentes críticos del sistema vistos en el capítulo anterior. De aquí una larga serie de pasos y simplificaciones<sup>10</sup> da como resultado la forma general del espectro de potencias en el punto crítico:

$$P(\omega) = \frac{1}{V} \omega^{-\mu} \bar{h}(\omega L^z) \quad (74)$$

Con  $\mu = 1 + \frac{2-\eta}{z}$ ,  $L$  la escala del sistema y con la función  $\bar{h}$  comportándose en general de la forma:

$$\bar{h}(x) = \begin{cases} cte & x \gg 1 \\ x^\mu & x \ll 1 \end{cases} \quad (75)$$

En la referencia [18], se muestra que la función de auto-correlación para el

<sup>10</sup>Los cuales pueden verse en la referencia [23]

modelo de Ising tiene la forma:

$$C(k, \omega, T) = k^{-(2-\mu+z)} f(\xi k, \xi^z \omega) \quad (76)$$

Con  $\xi(T) = |T - T_c|^{-\nu}$  la longitud de correlación vista en el capítulo anterior. De aquí se sigue que en general la forma del espectro de potencias:

$$P(\omega) = \lim_{k \rightarrow 0} C(k, \omega, T) \sim \omega^{-\mu} g(\omega/\xi^z) \quad (77)$$

donde la función  $g(x)$  está relacionada con la función  $f(x, y)$ . Nótese que si  $T = T_c$ , la ecuación se reduce a la ecuación (74). A partir de esto, se encontrará el comportamiento para  $T \gg T_c$ . De la ecuación (69), se puede eliminar el término a tercer orden  $g\phi^3$ . Como consecuencia, aplicando Fourier, la ecuación de correlación queda como:

$$C(k, \omega, T) = \langle \phi(k, \omega) \phi(-k, -\omega) \rangle = \frac{2\Gamma k_B T}{(\Gamma k^2 + A)^2 + \omega^2} \quad (78)$$

Lo que lleva a que el espectro a temperaturas altas sea:

$$P_T(\omega) \sim \frac{T}{(T - T_c)^2 + \omega^2} \quad (79)$$

Cabe destacar que si la frecuencia es muy pequeña  $P_T(\omega)$  será prácticamente constante. Por otro lado, si la frecuencia aumenta predominará el comportamiento de la forma  $1/\omega^2$ . Esto indica que básicamente existen dos regiones en el logaritmo del espectro de potencias: una en donde el espectro se comporta de manera constante y otra en donde se comporta como una recta con pendiente negativa. Nótese que mientras  $T$  se aproxime más a  $T_c$ , la región en donde el logaritmo del espectro se comporta como una constante será cada vez menor y por ende, la región en donde se comporta como una recta será mayor. Para

analizar por qué sucede esto, se regresará al teorema de Wiener-Kinchin. En él dice que el espectro de potencias es la transformada de Fourier de la correlación. Por ende, si el sistema tiene una zona con un espectro de potencias constante entonces la función de correlación se comporta como ruido blanco, es decir, no hay correlación. El resultado que se obtuvo anteriormente indica que el comportamiento del espectro como una constante se da para frecuencias bajas, o lo que es lo mismo, a largo plazo el sistema no se parece a sí mismo. En cambio, el hecho de que para frecuencias altas, el sistema se comporte como  $1/\omega^2$  indica que a corto plazo el sistema se parece a sí mismo. Uniendo esto, puede decirse que si el sistema se ve por periodos muy cortos de tiempo, no habrá cambios importantes en su configuración espacial mientras que, si se deja de ver por periodos largos, el sistema ya no se parecerá en nada al visto anteriormente. Para el punto crítico, la ecuación (74) indica que el logaritmo del espectro de potencias se comporta como una recta, lo que dice que el sistema se parece a sí mismo en todas las escalas. A esto se le conoce como que el sistema es autosimilar en el tiempo, o lo que es lo mismo, el sistema tiene una especie de “fractalidad temporal”, en donde el sistema tiene las mismas propiedades estadísticas, sin importar la escala temporal en la que se analice.

Después de una revisión de la literatura, no se encontró un desarrollo matemático similar para el caso de temperaturas bajas; sin embargo, después de hacer algunas visualizaciones del espectro de potencias para este rango de temperaturas se encontró una gran similitud entre éstos y los espectros de potencias a temperaturas altas, por lo que puede suponerse para temperaturas bajas se pueden encontrar los dos comportamientos del espectro de potencias: constante para frecuencias bajas y una recta con pendiente negativa para frecuencias altas.

## 2.3 Conceptos estadísticos básicos

El presente apartado pretende mostrar algunas definiciones estadísticas que resultarán muy útiles en lo posterior. En general, esto es un resumen de lo que se presenta en las referencias [26-28].

### Valor esperado y promedio

El valor esperado es, como su nombre lo indica, el valor que en promedio toma el sistema después de repetir por periodos muy largos de tiempo el mismo experimento. Se define como la suma de cada valor posible que pueda obtener el sistema multiplicado por la probabilidad de que el sistema tenga ese mismo valor<sup>11</sup>:

$$E[X] = \sum_{i=1}^S x_i p(x_i) \quad (80)$$

Donde  $S$  es el número de estados del sistema.

El promedio (o media aritmética), en cambio, es el valor característico de una propiedad que posee un conjunto de datos. Para calcularlo de manera discreta se toma la suma de todos los valores de la propiedad que se quiera medir del sistema (en el modelo de Ising, lo que se pretende medir es la magnetización) y se divide entre el número total de datos:

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i \quad (81)$$

Nótese que mientras la definición de valor esperado es meramente probabilística, la definición de promedio es un concepto estadístico. El teorema ergódico permite hacer una equivalencia entre ambas:

---

<sup>11</sup>Dado que el modelo de Ising es un modelo discreto, todas las definiciones de este apartado se darán en términos discretos. Para el caso de modelos continuos pueden consultarse las fórmulas en las referencias [26-28].

$$E[X] \approx \bar{X} \quad (82)$$

En general resulta muy interesante analizar cuánto se alejan los datos respecto al valor promedio. Para esto se calcula el valor esperado de cualquier potencia del término  $(X - \bar{X})$ , es decir se calcula  $E[(X - \bar{X})^n]$ . Al conjunto de todos éstos valores esperados se le conoce como los momentos de la distribución alrededor de la media. El valor de  $n$  indica el número del momento de distribución. Así por ejemplo el segundo momento es el valor esperado de  $(X - \bar{X})^2$ , el tercero es el valor esperado de  $(X - \bar{X})^3$ , etc.

### **Varianza**

La varianza es el segundo momento de la distribución. Esta da una aproximación respecto a cuánto varían los datos respecto al promedio. Teniendo  $N$  datos, puede calcularse empleando la siguiente fórmula:

$$Var[X] = E[(X - \bar{X})^2] = E[X^2] - 2\bar{X}E[X] + (\bar{X})^2 = E[X^2] - (\bar{X})^2$$

$$Var[X] = \left( \frac{1}{N} \sum_{i=1}^N x_i^2 \right) - \bar{X}^2 \quad (83)$$

A partir de la varianza, se define la desviación estándar:

$$\sigma = \sqrt{Var[X]} \quad (84)$$

### **Asimetría estadística**

La asimetría estadística (o skewness), se define como el tercer momento de

una distribución entre la desviación estándar al cubo, esto es:

$$Skew(X) = \frac{E[(X - \bar{X})^3]}{\sigma^3} = \frac{E[X^3] - 3\bar{X}\sigma^2 - \bar{X}^3}{\sigma^3} \quad (85)$$

donde  $\bar{X}$  es el promedio,  $N$  el número de datos y  $\sigma$  la desviación estándar. Lo importante de la asimetría estadística es que mientras más simétrica sea la distribución de datos analizados, la asimetría tiende a cero. Por ejemplo, para la distribución normal,  $Skew = 0$ . Valores negativos de la asimetría indican que hay más datos a la izquierda del promedio, mientras que  $Skew > 0$  indica que los datos están a la derecha. Si la distribución es bimodal, el signo de la asimetría dice cual es el modo más importante de la distribución. Por lo general (y dado que en general sólo interesa saber si el valor es positivo o negativo), se pueden utilizar fórmulas similares que reducen el tiempo del cálculo, conocidos popularmente como estimadores. Uno de los estimadores más utilizados es el estimador  $b$ :

$$b[X] = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{X})^3}{\left[ \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2 \right]^{3/2}} \quad (86)$$

En la literatura, existen otros estimadores de la asimetría estadística, pero para propósitos del presente trabajo, basta con conocer este.

### **Curtosis**

La curtosis es el cuarto momento de la distribución respecto a la media entre la desviación estándar a la 4 potencia. Ésta da información respecto a que tan cercanos o alejados están los datos entre sí, o lo que es lo mismo, dice cuan “puntiaguda” es una distribución. Matemáticamente se define como:

$$Kurt = \frac{E[(X - \bar{X})^4]}{\sigma^4} = \frac{\sum_{i=1}^N (X_i - \bar{X})^4 / N}{\sigma^4} \quad (87)$$



Es común que siempre se tome a la distribución normal como la distribución con respecto a la cual se hacen todas las comparaciones, por lo que se postula que ésta distribución tiene todos sus momentos igualados a 0. En este caso, dada la definición anterior de curtosis (ecuación [86]), la distribución normal tiene  $Kurt(normal) = 3$ . Esto origina que en casi toda la literatura se emplee también la siguiente definición de curtosis:

$$Kurt = \frac{E[(X - \bar{X})^4]}{\sigma^4} - 3 \quad (88)$$

Tomando ésta última definición pueden verse 3 comportamientos:

- Si  $Kurt > 0$ , entonces el conjunto de datos tiene una punta más afilada y con colas más anchas respecto a la normal. A esto se le conoce como curva leptocúrtica.
- Si  $Kurt < 0$ , el conjunto de datos tiene una punta más ancha y colas más delgadas respecto a la normal, denominándolo curva platicúrtica.
- Si  $Kurt = 0$ , tiene una forma muy parecida a la normal, es decir, es una curva mesocúrtica.

## 2.4 Alertas Tempranas

Muchos de los sistemas dinámicos tienen la cualidad de que la variación temporal que presentan conlleva a puntos en dónde aparecen cambios abruptos. Ejemplos de este tipo de sistemas son la generación y desaparición de especies y los cambios en los ecosistemas [31-34], el clima [35-37], los mercados bursátiles [38, 39] y por supuesto las transiciones de fase en sistemas físicos [9, 29, 30]. Aunque predecir en qué momento exacto del tiempo se da un punto de cambio abrupto del sistema es aún complicado, en los últimos años se han observado

patrones del comportamiento del sistema antes de que el cambio ocurra. Tales patrones son independientes del sistema dinámico que se estudie, lo que permite hablar de que existe un conjunto de “avisos anticipados del cambio” asociados a cualquier cambio abrupto del sistema. A tales “avisos anticipados de cambio” se les conoce por el nombre de alertas tempranas<sup>12</sup>. Cabe destacar que en general, muchos de los sistemas estudiados en la naturaleza muestran, en muchos casos, variaciones temporales mínimas que hacen imposible predecir un cambio abrupto; sin embargo, se han detectado variaciones importantes en algunos de los parámetros del sistema. Sobre tales parámetros es donde se busca encontrar las alertas tempranas.

En general las alertas tempranas pueden dividirse en dos tipos: las alertas tempranas basadas en mediciones y las alertas tempranas basadas en modelos. Las primeras trabajan con los datos del sistema y buscan cambios abruptos sobre la serie de tiempo que arroja el sistema al medir de éste alguna característica en particular. Por otro lado, las segundas muestran los cambios en el sistema cuando éste es ajustado a un modelo razonable [29].

De igual manera, dentro de las alertas tempranas basadas en las mediciones hay dos tipos, las basadas en las propiedades estadísticas del sistema y las basadas en la memoria del sistema mismo. Aquellas que se basan en propiedades estadísticas están relacionadas con la variación de propiedades estadísticas como la varianza, el promedio, la curtosis y la asimetría. Mientras que las que se basan en la memoria del sistema son aquellas que se muestran al comparar el estado actual del sistema con estados previos. En este rubro la correlación del sistema

---

<sup>12</sup> el término aún no ha sido traducido al español o si existe no sé encontró traducción alguna en la literatura utilizada en el presente trabajo. Para evitar mayores confusiones y para estar en correspondencia con la escasa literatura del tema se optó usar la traducción literal del término: alertas tempranas.

a distintos tiempos es lo más usado [29].

La mayoría de las variaciones que se buscan dentro del sistema están relacionadas profundamente con lo que se conoce en los sistemas dinámicos como ralentización del sistema (del inglés *critical slowing down*), es decir un sistema con fluctuaciones temporales cada vez más “lentas” comparadas con el comportamiento “común” del sistema. Aunque la teoría es muy similar para gran parte de las bifurcaciones, por simplicidad, se analizarán las bifurcaciones de plegamiento caóticas<sup>13</sup>. En una bifurcación de plegamiento (figura 7), la distancia entre los valores de equilibrio tiende a cero en los puntos de bifurcación. Esto significa que mientras que el sistema se aproxime más a tales puntos le cuesta cada vez más trabajo recuperarse de las perturbaciones pequeñas. Un análisis de varios modelos de este tipo de bifurcación muestra que esta ralentización del sistema en muchas ocasiones comienza a presentarse en regiones considerablemente lejanas al punto de bifurcación [1], por lo que, como una primera buena alerta temprana se propondrá la observación de una ralentización en el sistema. Para ejemplificar este ralentizamiento véase siguiente<sup>14</sup>:

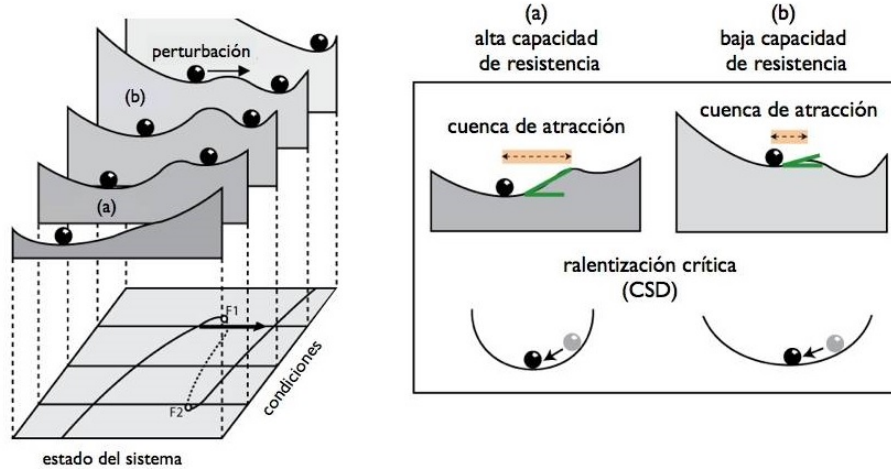
Supóngase que se tiene un sistema dinámico cuya dinámica está gobernada por la siguiente ecuación:

$$\frac{dx}{dt} = \gamma(x - a)(x - b) = f(x) \quad (89)$$

---

<sup>13</sup> Aquí, no está de más mencionar que no sólo el caos asociado a una transición da como resultado un ralentizamiento del sistema, sino que cualquier cambio en el mismo (sea caótico o no) generará una ralentización. Esto se ha mostrado válido tanto para bifurcaciones que llevan al caos como las de desdoblamiento (del inglés “fold bifurcation”), las transcíticas (del inglés “transcritical bifurcations”), las bifurcaciones Hopf y las bifurcaciones de Pitchfork (de las cuales el modelo de Ising es parte) como para modelos continuos con cambios muy suaves en sus propiedades (ver referencia [1]).

<sup>14</sup> Este desarrollo fue tomado de [29]



**Figura 7.** Ejemplo de bifurcación de plegamiento. Se muestra como un sistema con dos puntos atractores puede ir modificándose hasta llegar a un sistema donde sólo hay un punto atractor, dando pie al fenómeno de ralentización crítica (*critical slowing down*) (imagen tomada de la referencia [40])

Este modelo tiene la sencillez de tener dos puntos de equilibrio:  $x = a$  y  $x = b$ , uno de ellos estable y otro inestable. Si se cumple que  $a = b$ , el tipo de equilibrio colisiona y se tiene sólo un punto estable. Supongamos que  $a$  es un punto de equilibrio, en tal caso podemos ver como se comporta el sistema cerca del equilibrio haciendo:

$$\frac{d(a + \epsilon)}{dt} = f(a + \epsilon) \quad (90)$$

Haciendo una expansión de Talador a primer orden se tiene:

$$\frac{d(a + \epsilon)}{dt} \approx f(a) + \left. \frac{\partial f}{\partial x} \right|_{x=a} \epsilon \quad (91)$$

Que puede simplificarse aproximando el lado izquierdo a  $f(a) + \frac{d\epsilon}{dt}$  :

$$\frac{d\epsilon}{dt} = \left. \frac{\partial f}{\partial x} \right|_{x=a} \epsilon = -\gamma(b-a)\epsilon = \lambda_1 \epsilon \quad (92)$$

Por lo tanto, tenemos un valor propio para el equilibrio. Tomando de igual forma el punto  $b$ , llegamos al otro valor propio del sistema:

$$\lambda_2 = \gamma(b-a) = -\lambda_1 \quad (93)$$

Nótese que estos valores propios dan una aproximación de las tasas de recuperación del sistema. Si  $b > a$ ,  $\lambda_1 < 0$  y por ende es estable. Lo importante aquí es ver la dependencia del cambio respecto a la lejanía del punto crítico. Si se toma un valor grande de  $\epsilon$  (es decir, si se toma un valor muy lejano al punto crítico), el sistema se recuperará más rápido que si se toma un valor pequeño de  $\epsilon$ . Véase que si  $a = b$ , las tasas de recuperación son cero y el sistema jamás se recuperará de una perturbación. Muy lejos del punto, las tasas de recuperación dependen linealmente del parámetro  $(b-a)$ . Aunque esta es una simplificación muy burda para demostrar el ralentizamiento del sistema en puntos cercanos al crítico, muchos de los sistemas que se han estudiado en la literatura responden también de una manera lineal y monotónica [29].

Puede demostrarse que cerca de un punto de bifurcación, algunas propiedades que caracterizan al sistema cambian debido al ralentizamiento del mismo. En particular existe un aumento de la autocorrelación y de la varianza, por lo que estos dos parámetros pueden usarse como alertas tempranas del sistema. Se analizará primero por qué es que el ralentizamiento da origen a una alerta temprana en la autocorrelación. Piense en esto: si el sistema se recupera más lentamente cerca del punto crítico, esto quiere decir que el sistema cambia cada vez menos en el tiempo o, lo que es lo mismo, el sistema aumenta su autocorrelación de corto alcance significativamente cerca de un punto crítico. Por lo mismo,

un análisis del espectro de potencias estaría fuertemente ligado a un análisis de la correlación ya que en el fondo, ambos están asociados con la semejanza del sistema en el tiempo, por lo que se pueden inferir indirectamente las propiedades del tiempo de relajación del mismo. El aumento en el valor de la correlación es un comportamiento que se ha visto tanto en los modelos más simples como en los más complejos [29].

Íntimamente ligado al aumento de la correlación del sistema se encuentra el aumento de la varianza. Para esto, véase que mientras más cerca se encuentre el sistema de algún punto crítico, el sistema se aleja cada vez más frecuentemente de su estado “promedio”, es decir hay un aumento de la varianza. Una explicación matemática para que esto ocurra puede verse así:

Supóngase que el sistema es perturbado después de un cierto periodo de tiempo  $\Delta t$  y que el sistema tiene una rapidez de recuperación  $\lambda$ . El sistema de autoregresión quedaría:

$$x_{n+1} - x = e^{\lambda\Delta t}[x_n - x] + \sigma\varepsilon_n$$

$$y_{n+1} = e^{\lambda\Delta t}y_n + \sigma\varepsilon_n \quad (94)$$

Donde  $y_n$  es la desviación de la variable de estado  $x$ ,  $\varepsilon_n$  es un número aleatorio de una distribución normal estándar y  $\sigma$  es la distribución estándar. Suponiendo que  $\lambda$  y  $\Delta t$  son independientes de  $y_n$ , el modelo podría ser reescrito como:

$$y_{n+1} = \alpha y_n + \sigma\varepsilon_n \quad (95)$$

Donde  $\alpha$  está relacionado con la correlación del sistema, es decir, es cero para ruido blanco y cerca de 1 para una autocorrelación grande. Calculando la

esperanza de ambos lados:

$$E(y_{n+1}) = \alpha E(y_n) + \sigma E(\varepsilon_n) \implies \mu = 0$$

Dado que  $\mu = 0$  es la única solución a la ecuación. De esta manera, la varianza es:

$$Var(y_{n+1}) = E(y^2) - 0 = \frac{\sigma^2}{1 - \alpha^2} \quad (96)$$

Puede verse entonces que para la bifurcación dinámica del sistema,  $\lambda$  se aproxima a cero, la correlación por ende aumenta y por el término de la varianza incrementa significativamente. Otra forma de explicarlo intuitivamente es que si  $\lambda$  (la rapidez de recuperación del sistema) en la bifurcación dinámica se aproxima a cero, el sistema no se recupera rápidamente de las perturbaciones haciendo que fluctúe entre varios valores por muchos periodos de tiempo. De esta forma el hecho de que exista una ralentización del sistema implica la existencia de un aumento en la varianza, por lo que esta última puede usarse también como alerta temprana.

Hay también otras propiedades que cambian abruptamente cuando el sistema cambia pero que, a diferencia de las anteriores, no provienen directamente de la ralentización del sistema. Tal es el caso de la asimetría estadística por lo que ésta puede ser también propuesta como una buena alerta temprana del sistema. Una explicación de por qué cambia abruptamente proviene del hecho de que en una bifurcación catastrófica, el sistema empieza a acercarse cada vez más a un nuevo atractor en vez de estar fluctuando entre los distintos atractores, lo que provoca una desviación grande de la media y por ende la asimetría de la distribución estadística se vea afectada.

El valor esperado de algún estado accesible del sistema es una propiedad

que también cambia cuando se acerca a un punto crítico. La explicación del por qué de este cambio es simplemente que el sistema pasa de un estado a otro, por lo que la probabilidad de que el sistema elija un estado antes poco probable aumenta en perjuicio de la probabilidad de que se encuentre en cualquier otro estado. A esta propiedad se le denomina *flickering* en la literatura [29].

La curtosis también puede establecerse como otra alerta temprana. Esto puede explicarse por la relación entre la varianza y la curtosis aunque también del hecho de que la curtosis mide la cercanía total de los datos a la media.

El total de alertas tempranas usadas en la literatura es amplísimo, aunque las hasta aquí mencionados son las más comunes y son los que se usarán en el presente trabajo.

#### **2.4.1 Propuesta de alerta temprana para el modelo de Ising**

El presente trabajo analizará todos las alertas tempranas expuestas en el apartado anterior aplicados al modelo de Ising así como también propondrá una nueva. Es bien conocido que por lo general, las transiciones de fase se caracterizan por ser libres de escala y su espectro de potencias da una línea recta. Sin embargo, esta línea recta en coordenadas logarítmicas dentro del espectro de potencias no aparece de forma instantánea sino de una manera gradual, es decir, la forma general del logaritmo del espectro de potencias va tomando poco a poco la forma de una línea recta mientras más cerca se esté del punto crítico. En general y como se mencionó en la sección 1.2.2, el logaritmo del espectro de potencias del modelo de Ising se comporta como una constante a frecuencias bajas y como una recta a frecuencias altas. Mientras se pueda aproximar a la temperatura crítica, el espectro de potencias va tomando la forma de una recta con pendiente negativa asociada a la invariancia de escala. Esto es debido a que



las correlaciones temporales van haciéndose cada vez más notorias a frecuencias bajas mientras más nos acerquemos al punto crítico. Es de esperarse entonces que haya un punto medio (punto de corte o crossover,  $C_f$ ) a partir del cual el espectro deje de comportarse como una constante para comportarse como una recta. Se espera entonces que, analizando el punto de corte cerca del punto crítico, se pueda obtener una nueva alerta temprana. Es importante resaltar que esta alerta temprana está ampliamente relacionada con las correlaciones del sistema a varias escalas temporales. En efecto, dado que el ajuste del espectro de potencias a una constante y a una recta con pendiente negativa para las frecuencias bajas y altas, respectivamente, indica que en el sistema prevalecen las correlaciones para frecuencias altas; esta alerta temprana ayudará a descubrir cambios en las correlaciones temporales del sistema, por lo que, mientras sea menor el punto de corte, más frecuencias serán ajustadas por una recta con pendiente negativa y por ende mayor será la correlación del sistema. En resumen, esta alerta temprana nos dice qué tan “invariante de escala” se está volviendo el sistema como un todo. Esta alerta temprana no se ha encontrado antes en la literatura sobre Ising y, aunque las referencias [24-25] ya habían comentado respecto del cambio del punto de corte en puntos cercanos al crítico, no ahondaron más en su posibilidad de establecerlo como una alerta temprana y no mostraron un análisis al respecto. En el siguiente capítulo se analizará esta propuesta con mayor detalle.

## 2.5 Consideraciones prácticas.

Muchos análisis de alertas tempranas se basan en analizar las tasas de recuperación del sistema y con eso saber si se encuentran o no cerca de un punto crítico. Esto tiene implicaciones prácticas que podrían ser devastadoras si se dan como

válidos ciertos resultados falsos o incompletos asociados al hecho de una mala elección de la escala temporal. ¿Cuál es una buena escala temporal para poder tomar como válidos los avisos de las alertas tempranas? Hasta el momento es imposible saberlo y depende ampliamente del sistema que se esté analizando.

Otra consideración para tomarse en cuenta en el análisis de alertas tempranas es la aparición de falsos positivos y falsos negativos. Los primeros ocurren cuando hay un cambio importante en la señal que anuncia la alerta temprana pero ésta no conlleva a un cambio abrupto en el sistema. Por lo general esto se debe a un análisis incompleto del sistema (asociado a su vez a una mala elección de la escala temporal), una tendencia confusa dentro del sistema o en sus alrededores o por mera coincidencia. Por otro lado, el falso negativo aparece cuando, sin que haya un cambio en las alertas tempranas, el sistema padece un cambio abrupto. Esto puede estar asociado a muchos factores; por ejemplo, un cambio tan súbito que no lo puedan detectar las alertas tempranas o un evento extremadamente raro que lo haya llevado a tal cambio de una manera no gradual. En cualquier caso también puede afectar de gran manera el que no se tomen a una frecuencia necesaria las alertas tempranas y esto a su vez tiene que ver con la escala temporal elegida para medir el sistema. Otro falso negativo también puede aparecer si el régimen de perturbaciones externas cambia con el tiempo; en cualquier caso, la elección de una escala temporal y un buen análisis de series de tiempo siempre son fundamentales. Muchos proponen nuevas técnicas para filtrar las series de tiempo y hacerlas más sensibles a las perturbaciones externas [32, 33]. La investigación actual en alertas tempranas y series de tiempo va en camino a establecer cuál es una buena escala temporal y cuándo un cambio en parámetros es significativo. Algunos resultados, sin embargo, son claros: se tiene que, al menos, eliminar el ruido estocástico para poder detectar efectivamente la autocorrelación. En casos como el modelo de Ising computacional es sencillo

eliminarlo, basta con correr el mismo programa varias veces bajo las mismas condiciones iniciales y hacer promedios de los resultados del parámetro que se mide. Nótese que eliminar el ruido ayuda también a tener un buen espectro de potencias al calcular la transformada de Fourier, por lo que calcular el espectro de potencias promedio a partir del espectro de potencias de cada serie es un paso clave para obtener resultados más precisos. Para modelos más complicados, como el análisis de ecosistemas, el problema de eliminar el ruido estocástico se vuelve un poco más complicado. Las referencias [35, 36] muestran cómo es posible tratar este tipo de sistemas.

Un aspecto último por resaltar respecto a las alertas tempranas es que, aunque efectivamente se haga una buena elección de la escala temporal y se conozca prácticamente a la perfección el sistema y todas sus alertas tempranas asociadas, es imposible determinar con las herramientas actuales el momento exacto de la transición. En efecto, se puede, a lo más, saber que el sistema va en camino hacia un cambio debido a una modificación en alguna de las propiedades del sistema. Con esto puede saberse que en el sistema está pasando “algo”, mas no puede saberse el punto exacto en donde este cambio dará lugar a un cambio abrupto en todo el sistema, cuánto tiempo durará ese periodo de cambio hasta llegar a una relativa “nueva” estabilidad y, en algunos casos, tampoco las consecuencias totales de dicho cambio. Aún más: no se sabe incluso si las consecuencias de dicho cambio serán o no caóticas ya que las alertas tempranas aquí propuestas se han demostrado útiles para medir cualquier cambio del sistema sin importar si este cambio lleve o no al caos.

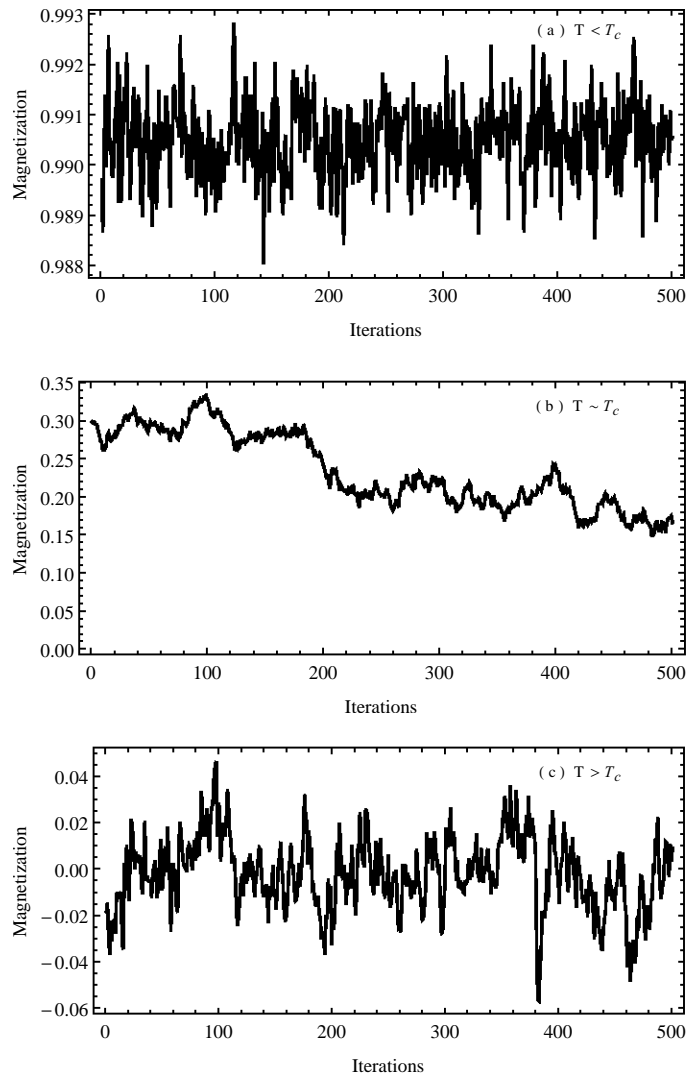
# Capítulo 3

## 3 Procedimiento, resultados y análisis

### 3.1 Procedimiento

Se propone entonces crear simulaciones computacionales del modelo de Ising en 2D para un conjunto basto de temperaturas. Con esto, dada una temperatura fija se obtendrá la serie de tiempo asociada a la magnetización del sistema y se dejará correr por un periodo de tiempo largo (5 mil repeticiones del método Monte Carlo por simulación) obteniendo entonces una serie de 5 mil datos de magnetización por cada simulación. En el presente trabajo se define una unidad de tiempo para la simulación como el número de pasos en el que en promedio, cada elemento de la malla del modelo de Ising haya sido afectada por el método Monte Carlo. El tamaño de la malla es de  $100 \times 100$ . Como se mencionó en el capítulo 1, no es necesario esperar a que se cumpla el tiempo de relajación ya que no se medirán las propiedades termodinámicas del sistema sino la evolución dinámica del mismo, sin embargo se tomará un tiempo  $t = 1000$  para que el sistema elimine las primeras fluctuaciones. El tiempo  $t = 1000$  es el tiempo que en promedio es necesario para que la temperatura más baja que se analiza en este trabajo ( $T = 1.42$ ) alcance una magnetización promedio  $\bar{M} = \pm 1$ , por lo que el total de datos a analizar por serie de tiempo será solamente de 4000. La serie de tiempo que se obtenga será analizada con los métodos presentados en el capítulo 2 y con esto se obtendrán las alertas tempranas del sistema, a saber: el promedio, la varianza, la asimetría estadística (*skewness*), la curtosis, la auto-correlación a  $\tau = 1$  y el únto de corte del ajuste de dos rectas al logaritmo del espectro de potencias. En orden de eliminar ruido no deseado, para cada temperatura se hicieron mil series de tiempo y posteriormente se trabaja

con cada una de ellas (es decir, hace estadística con alrededor de 4 millones de datos por temperatura). los resultados que se obtienen son el promedio de cada uno de los valores computados para cada serie. El rango de temperaturas estudiado es de  $T = 1.42$  hasta  $T = 3.12$ . La temperatura va aumentando 0.05 en cada medición. El software usado para la simulación del modelo de Ising es la simulación del mismo nombre que viene incluida en la librería de NetLogo 5.0.5. Este emplea el método Monte Carlo con el algoritmo de Metrópolis por lo que no se vio la necesidad de crear un nuevo programa aunque sí hacerle algunas modificaciones para los propósitos del trabajo. Las series de tiempo de la magnetización son extraídas y posteriormente analizadas en Mathematica 10. En el siguiente apartado se mencionarán brevemente las generalidades de cada programa. Por lo general en cada resultado se remarcarán tres zonas fácilmente distinguibles, una zona de temperatura baja comparada con la crítica ( $T < T_c$ ) que se toma con temperaturas entre  $T = 1.42$  y  $T = 2.22$ , una donde la temperatura es muy cercana a la crítica ( $T \sim T_c$ ), con rangos de temperatura entre  $T = 2.22$  y  $T = 2.32$  y por último, una zona con temperaturas mayores a la crítica ( $T > T_c$ ), que comprende un rango de temperaturas entre  $T = 2.32$  y  $T = 3.12$ . Aquí cabe decir que aunque teóricamente la temperatura crítica es  $T_c = 2.27$ , en la práctica los efectos de finitud de la malla pueden recorrer la temperatura crítica a otras zonas cercanas. Para asegurar que el resultado es robusto, se probó con distintos tamaños de malla, pudiendo percibir algunos cambios en el valor de la temperatura crítica pero con resultados equivalentes para cada tamaño, es por esto que en adelante no se hablará de una temperatura crítica sino en general de una “zona crítica” en la cual se producen los cambios en las propiedades del sistema. Un ejemplo de las series de tiempo obtenidas con el modelo de NetLogo y después graficadas en Mathematica son las mostradas en la figura 8. Como puede verse cada una fluctúa entre distintos valores. A



**Figura 8.** Magnetización total del modelo de Ising en  $2d$  como función del tiempo para 3 distintas temperaturas. Se toma solamente un periodo de tiempo de 0 a 500. La imagen a) muestra la serie para la región  $T < T_c$ , la b) para la región  $T \sim T_c$  y por último la imagen c) para la región  $T > T_c$ . Nótese el cambio de las escalas en cada gráfica.

temperaturas muy bajas la gráfica fluctúa muy cerca del valor de magneti-

zación <sup>15</sup>, en la zona crítica las fluctuaciones son mayores y recorren un amplio rango. Por último, a temperaturas altas, la serie de tiempo fluctúa alrededor del cero.

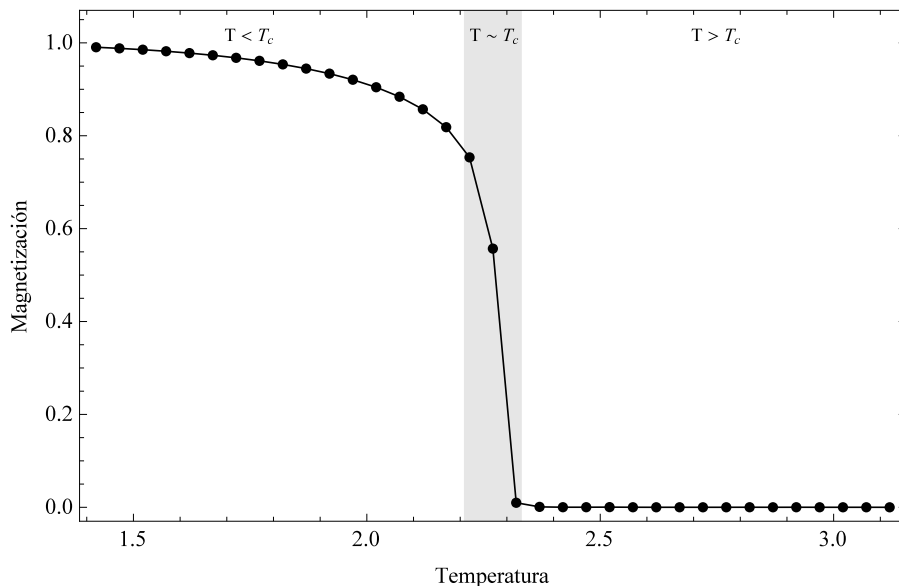
## 3.2 Resultados y análisis

### 3.2.1 Promedio

El promedio es el primer momento de la distribución. El programa empleado para analizar las series de tiempo solamente usa la ecuación [80] del capítulo anterior desde un tiempo  $t = 1000$  (que es cuando el sistema a temperaturas baja llega al promedio  $\bar{M} = 1$ ) hasta  $t = 5000$ , tomando el promedio de todos los tiempos, se tiene la magnetización asociada a la temperatura. La figura 9 muestra el promedio de la magnetización como función de la temperatura. Era de esperarse un comportamiento del estilo ya que, como se vio en el capítulo 1, pasando la temperatura crítica la magnetización total es prácticamente cero, mientras que en valores lejanos a la temperatura crítica, la magnetización sólo puede tener valores de  $M = \pm 1$ . Nótese que en este caso, la gráfica muestra que la magnetización va desde un valor de +1 hasta llegar al cero, es decir, se tomó el valor absoluto de la magnetización. Nótese además cómo los tres regímenes pueden encontrarse fácilmente. Cuando  $T < T_c$ , recordamos del capítulo 1, se forman grandes costras que poco a poco van llenando todo el espacio. Hay posibilidades de que existan estados metaestables donde grandes costras lleguen a un equilibrio mutuo pero esos estados son muy difíciles de encontrar; lo más común es que una gran costra de espines hacia arriba o hacia abajo llene todo el espacio hasta el punto de que después de cierto tiempo indeterminado, se acaba

---

<sup>15</sup>El sistema podría fluctuar también alrededor de  $-1$ , sin embargo en el presente trabajo y con el fin de homogeneizar resultados sólo se toma el valor absoluto de la magnetización.



**Figura 9.** Magnetización promedio del sistema como función de la temperatura. Se muestran los tres regímenes de temperatura. Es destacable la caída abrupta del valor de la magnetización en la zona crítica (marcada con gris).

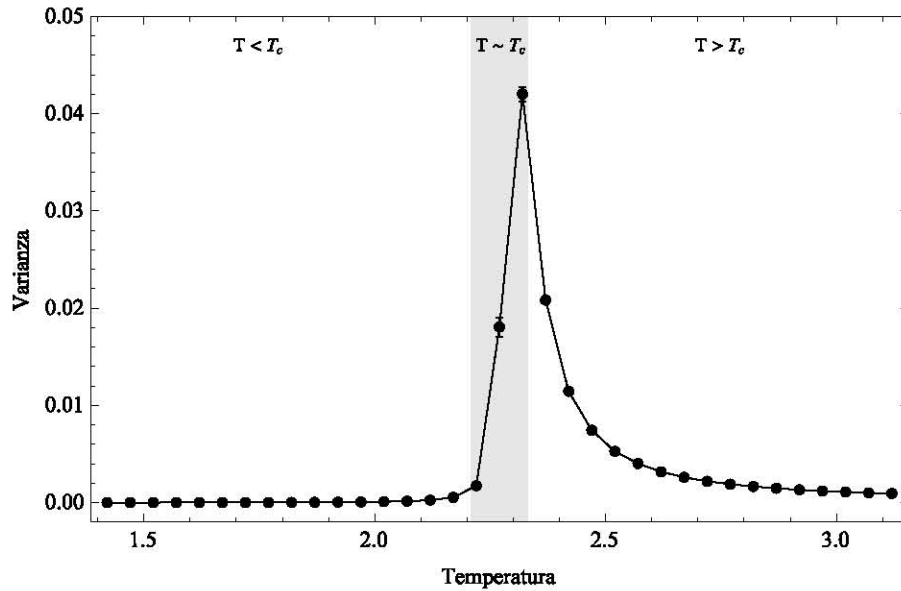
toda la variabilidad dinámica del sistema. Por otro lado, el efecto de  $T > T_c$  es totalmente el opuesto. El ruido estocástico no permite la generación de costras y prácticamente cada espín se mueve independiente de sus vecinos. Esto genera que en promedio la magnetización sea  $M = 0$ . La zona donde  $T \sim T_c$ , muestra una combinación de los dos efectos. Ya desde antes de la zona  $T \sim T_c$ , se puede ver una pequeña disminución del valor  $M = 1$ , de igual manera, poco después de que termina la misma zona, la magnetización se encuentra ligeramente arriba del valor  $M = 0$ . Lo sorprendente es que en el rango de temperaturas señalado en la figura 9, la magnetización cae abruptamente. Como se mencionó también en el capítulo 1, se pueden tener costras de todos los tamaños (hay invariancia de escala espacial) a tal grado de que una ligera variación del sistema afecta gravemente a todo el sistema, es decir el sistema es crítico en el sentido de que



todos los miembros del sistema se influyen entre sí.

### 3.2.2 Varianza

La varianza es el segundo momento de la distribución de probabilidad. Como tal, el programa sólo utiliza la ecuación [82] vista en el capítulo 2 evaluada desde un tiempo  $t = 1000$  hasta  $t = 5000$ . El resultado del cambio de la varianza respecto a la temperatura se muestra en la figura 10. Claramente se pueden distinguir tres zonas:  $T < T_c$ ,  $T \sim T_c$  y  $T > T_c$ . Cuando  $T < T_c$ , la gráfica se muestra prácticamente constante alrededor del cero, mientras que la zona  $T > T_c$  muestra una caída desde un punto máximo hasta cero, por lo que se tienen dos comportamientos distintos. Si se trabaja en un modelo de Ising en donde la temperatura va ascendiendo desde un punto inferior a  $T_c$  y del sistema sólo se conoce la varianza, no existirían indicios de cuando se está próximo a la temperatura crítica. Por otro lado, si la temperatura va descendiendo desde un valor mayor a la temperatura crítica y el único dato que se conoce del sistema es la varianza, se puede saber que se está cerca del punto crítico por un aumento en el valor de la varianza. Una explicación física del porque esto ocurre se dará en la parte de la discusión pero por el momento puede entenderse que a temperaturas muy bajas el sistema está alineado casi siempre hacia la misma dirección, por lo que la varianza será muy baja y en caso de ser estrictamente positiva estará relacionada con las fluctuaciones estocásticas del sistema. En cambio, a temperaturas muy altas las fluctuaciones estocásticas predominan por lo que el sistema no varía mucho en los valores de magnetización. En el punto crítico, las fluctuaciones empiezan a tomar gran importancia por lo que la varianza diverge pues hay más estados a los que es posible que acceda el sistema.

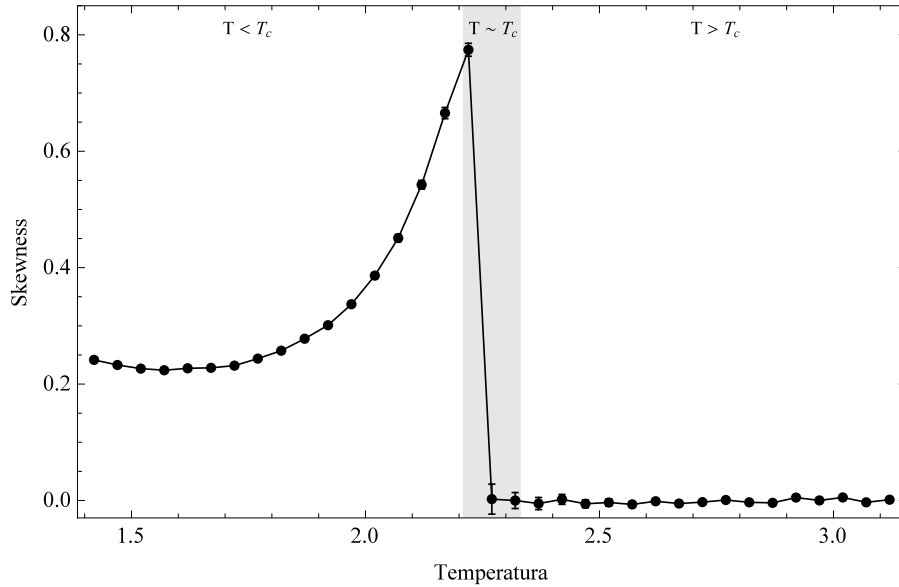


**Figura 10.** *Varianza como función de la temperatura. Comportamiento de la varianza como función de la temperatura. Se muestran 3 zonas  $T < T_c$ ,  $T \sim T_c$  y por último  $T > T_c$ .*

### 3.2.3 Asimetría estadística

Para calcular el término de skewness se toman una temperatura fija  $T$  de la cual se elijen las mediciones de la magnetización desde un tiempo  $t = 1000$  hasta un tiempo  $t = 5000$  calculando la asimetría estadística a cada una, usando para esto la ecuación 84. Posteriormente se hace un promedio entre todos los datos y ese valor resultante es el que se presenta como el skewness asociado a la temperatura  $T$ . Hay que recalcar que aunque esta ecuación da la asimetría de manera exacta, existen varias formas de tener un valor aproximado de asimetría estadística y por lo general son sencillas y fáciles de emplear cuando se tienen

gran cantidad de datos. Una de estas ecuaciones es la aproximación  $b$  definida en la ecuación (86) del capítulo anterior. Los resultados de calcular la asimetría se muestran en la figura 11.



**Figura 11.** Asimetría estadística (*skewness*) como función de la temperatura. En la imagen a) se muestra el resultado calculado con la definición formal de asimetría estadística, mientras que en la imagen b) se muestra el resultado con la aproximación  $b$ . En ambas imágenes se muestran 3 zonas  $T < T_c$ ,  $T \sim T_c$  y  $T > T_c$ .

De nuevo, pueden distinguirse básicamente 3 zonas. Una de crecimiento gradual ( $T < T_c$ ), una de caída abrupta ( $T = T_c$ ) y una de valor constante alrededor del cero ( $T > T_c$ ). Todos estos resultados eran los esperados puesto que en la zona  $T < T_c$  el sistema pasa de tener grandes costras “peleando” entre sí por el espacio a tener una sola gran costra, originando así una asimetría en la distribución de magnetización que se va haciendo cada vez más evidente conforme aumenta  $T$ . Este aumento en la asimetría de la distribución responde a que la

aparición de espines al azar hacen que descienda el promedio, que es un término negativo en la aproximación  $b$ . De la misma manera en  $T > T_c$ , el sistema está completamente al azar, por lo que la distribución de magnetización es parecida a ruido blanco, dando como resultado que sea prácticamente simétrica e igual para toda  $t$ , por lo que no se presentan asimetrías ni cambio alguno en la medición, es decir, el skewness es cero. En la zona crítica, la caída abrupta en la asimetría obedece a un aumento en la cantidad de regiones al azar, las cuales hacen disminuir el promedio. En la discusión se ampliará el análisis de estos datos usando la aproximación de la asimetría dada en la ecuación (86).

### 3.2.4 Curtosis

El programa, al igual que los anteriores, toma las mediciones de  $t = 1000$  hasta  $t = 5000$  y se hace un promedio entre las curtosis obtenidas a mil series de tiempo. Para calcular la curtosis de cada serie se usa la ecuación (87) del capítulo 2. Los resultados se muestran en la figura 12. Claramente, pueden verse de nuevo las 3 regiones de temperatura. La curtosis se comporta de una forma simétrica siendo la región  $T \sim T_c$  el punto de simetría. Nótese que tanto para temperaturas cercanas a cero o a infinito, la curtosis tiende al valor 3, es decir al valor que tiene la normal. Aproximándose a la temperatura crítica desde temperaturas bajas, la curva tiene una forma leptocúrtica, mientras que en la región con temperaturas altas, se obtiene una distribución platicúrtica.

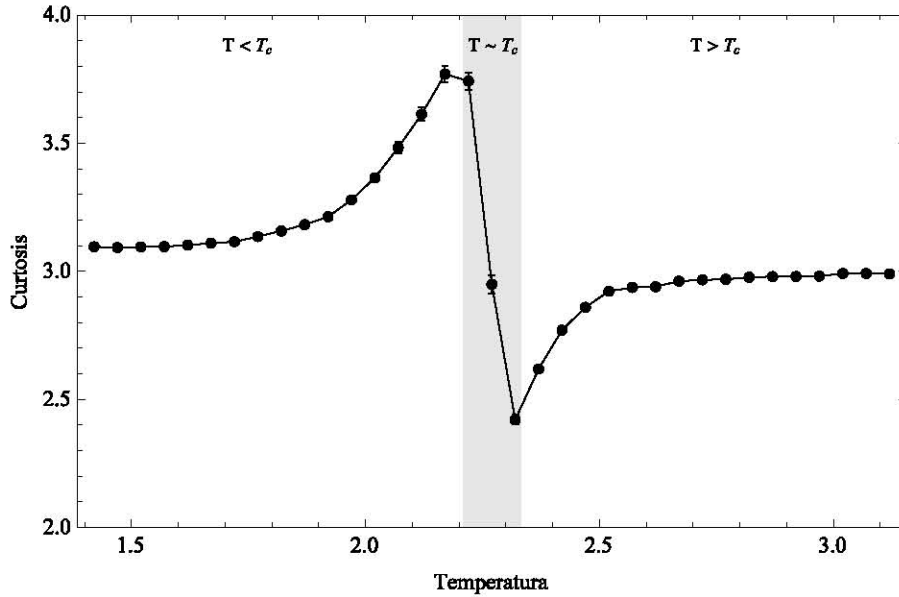


Figura 12. Curtosis como función de la temperatura. Comportamiento de la curtosis como función de la temperatura. 3 zonas recalçadas:  $T < T_c$ ,  $T \sim T_c$  y  $T > T_c$ .

### 3.2.5 Autocorrelación a $\tau = 1$

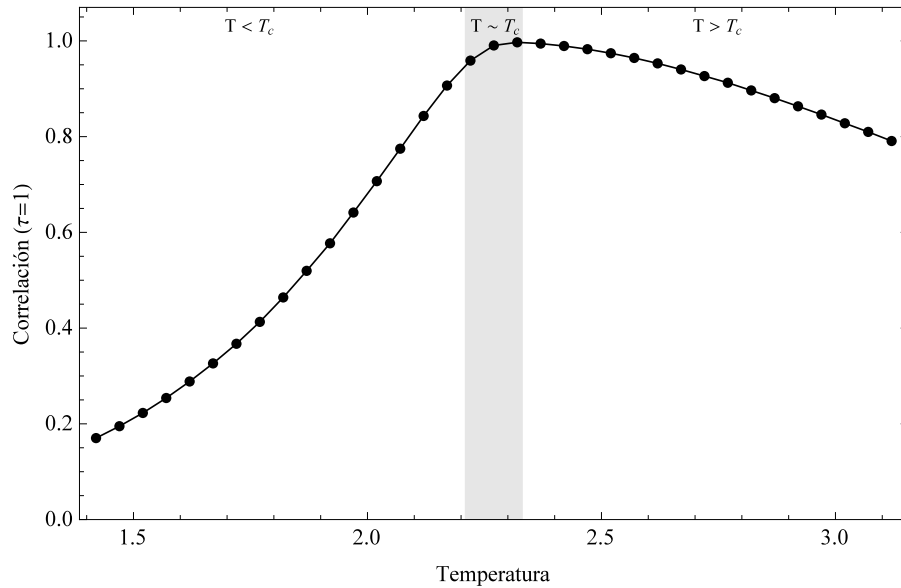
En esta parte, se analizó la correlación a  $\tau = 1$  del sistema. Para esto se usó la siguiente fórmula de correlación:

$$C_{\tau=1}(x) = \frac{1}{N} \sum_{n=0}^{N-1} x(n)x(n-1) \quad (97)$$

aplicándola a cada una de las 1000 series de tiempo, para posteriormente mostrar sólo el valor promedio de la correlación para cada temperatura. Como se mencionó en el capítulo 2, es de esperarse que la correlación sea muy alta para la temperatura crítica y que descienda conforme se aleje de ella. La figura 13 muestra la gráfica de los resultados. Puede verse que para temperaturas muy

bajas la correlación a  $\tau = 1$  es extremadamente baja comparada con la correlación a temperaturas muy altas. Además, puede apreciarse que la correlación es prácticamente 1 en la región crítica, esto se debe a la autosimilaridad del sistema en esta región. El que a temperaturas muy bajas el sistema no se parezca a sí mismo sorprende a primera vista por el hecho de que la varianza del sistema es prácticamente cero así como que el promedio se mantiene en un valor constante. Esto se debe primordialmente a que el sistema es en general estable aunque tenga asociadas siempre fluctuaciones estocásticas debido a las reglas de transición definidas por el método Monte Carlo, lo que ocasiona que cada uno de los estados del sistema pueda considerarse independiente del anterior, por lo que no hay correlación. Recuérdese en este punto el Teorema de Wiener-Kinchin, el cual establece que los cambios en la correlación se relacionan directamente con cambios en espectro de potencias. Se puede inferir entonces que el aumento de correlación de corto alcance está ampliamente relacionado con la aproximación a la ley de potencias que puede verse en el espectro de potencias de la figura 14. Las leyes de potencias se asocian generalmente a la criticalidad, específicamente a la invariancia de escala temporal. Como se ha mencionado el sistema muestra fractalidad espacial en el punto crítico, esto es, hay una invariancia de escala espacial. Es entonces de resaltar que al mismo tiempo, el sistema alcanza una invariancia de escala temporal, esto es que, estadísticamente, las series de tiempo son prácticamente las mismas independientemente de la escala temporal que se utilice. Tan pronto como el sistema se aleja del punto crítico tanto hacia temperaturas más bajas como a temperaturas más altas, el sistema deja de presentar las propiedades de invariancia espacial y temporal ocasionando que deje de ser dependiente de configuraciones anteriores y por ende, la correlación a  $\tau = 1$  disminuye. Esto está directamente relacionado con que el espectro de potencias se aplana en la zona de frecuencias bajas, pero eso se analizará a fondo

en el próximo apartado.



**Figura 13.** Autocorrelación a  $\tau = 1$  como función de la temperatura.

*Comportamiento de la correlación como función de la temperatura. Se recalcan 3 zonas:  $T < T_c$ ,  $T \sim T_c$  y  $T > T_c$ .*

### 3.2.6 Punto de corte del ajuste de dos rectas al espectro de potencias.

Recuérdese la discusión teórica del capítulo 2 respecto al espectro de potencias del modelo de Ising. En esta se menciona que para temperaturas por encima de la crítica pero cercanas a la misma, el logaritmo del espectro de potencias asociado a la serie de tiempo de la magnetización en el modelo de Ising puede dividirse en dos regiones:

- Una región en donde el espectro permanece prácticamente constante

- Una región en que puede ser ajustada una recta con pendiente negativa.

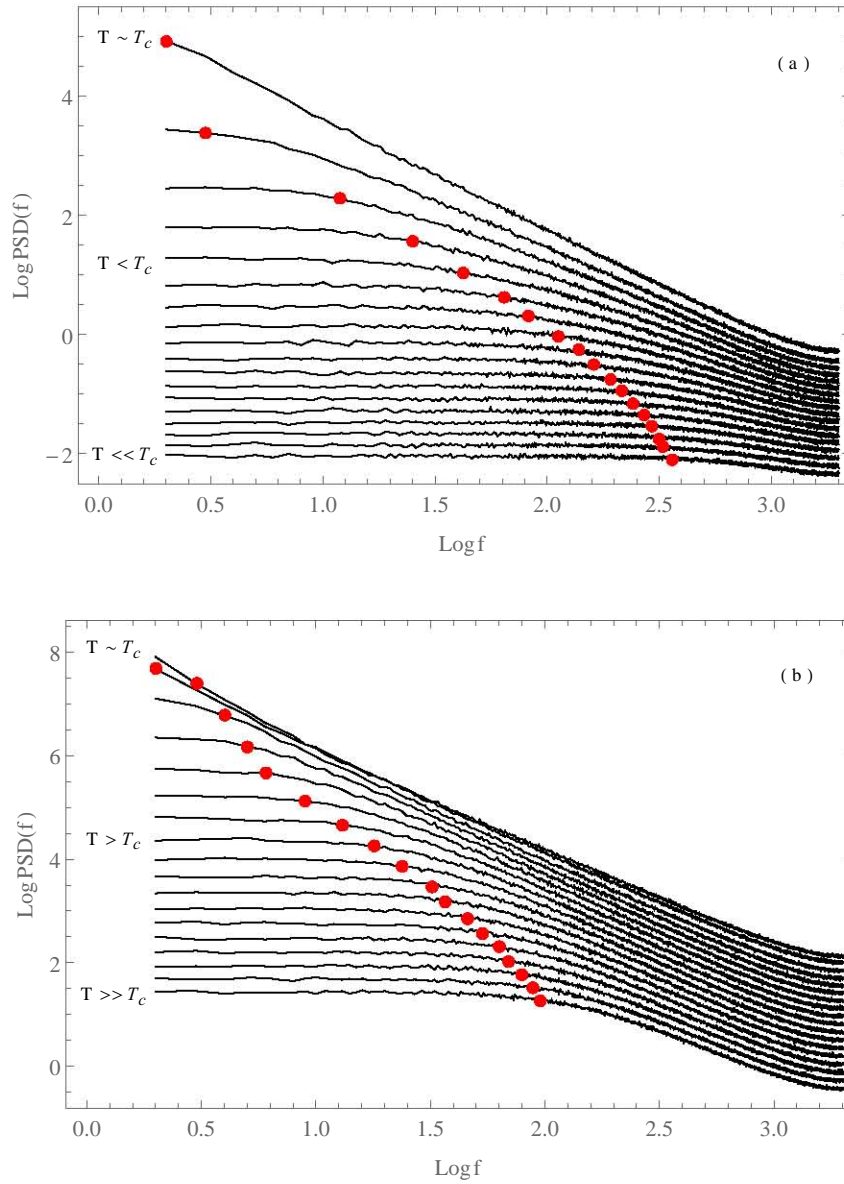
En ese momento se mencionó que aunque no se encontró en la literatura una justificación para que esto se siguiera cumpliendo para temperaturas por debajo de la crítica; sin embargo, viendo varios espectros de potencias del modelo de Ising se encontraron grandes parecidos, por lo que se supondría que las mismas dos regiones se podrían encontrar también para temperaturas menores pero cercanas a la crítica. La figura 14 muestra el espectro de potencias para distintas temperaturas. La figura fue acomodada para que puedan apreciarse todos los espectros en la misma figura.

Con esto en mente, el programa consiste en ajustar estas dos regiones al espectro promedio de 1000 corridas llevadas desde un tiempo  $t = 1000$  hasta un tiempo  $t = 5000$ . Aquí se decidió dividir el programa en dos subprogramas. El primero se enfoca en ajustar una constante  $k$  y una recta de la forma  $ax + b$  a todos los datos del espectro de potencias. El programa empieza partiendo los datos del espectro de potencias en 2 subgrupos de datos: uno con los primeros 4 datos del espectro y el otro con el resto ( $4000 - 4 = 3996$  datos)<sup>16</sup>. A este primer subgrupo de 4 datos, el programa le ajusta una recta constante y al resto una recta. Posteriormente se calcula la incertidumbre asociada al ajuste de las dos rectas y el resultado se guarda en una lista. Se procede a hacer el mismo ajuste de una constante pero ahora con 5 datos iniciales y al resto de datos se les rea-

---

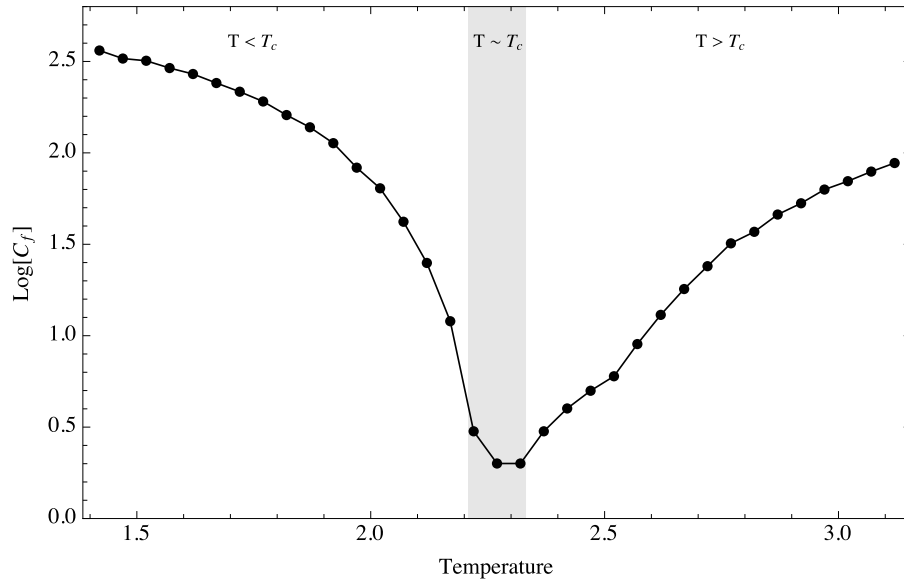
<sup>16</sup> El elegir que se empiece ajustando una constante a 4 datos fue a base de prueba y error: ajustando 2 datos, el ajuste era trivial y la incertidumbre asociada era claramente igual a cero. Ajustando 3 datos iniciales, la incertidumbre asociada a esto era muy próxima a cero y gran parte de los puntos de corte se quedaban en 3 como la mejor aproximación. A partir de tomar 4 como el punto de corte inicial, los datos comienzan a dar un resultado mucho más razonable por lo que se decidió empezar el ajuste de la constante primeramente con 4 puntos





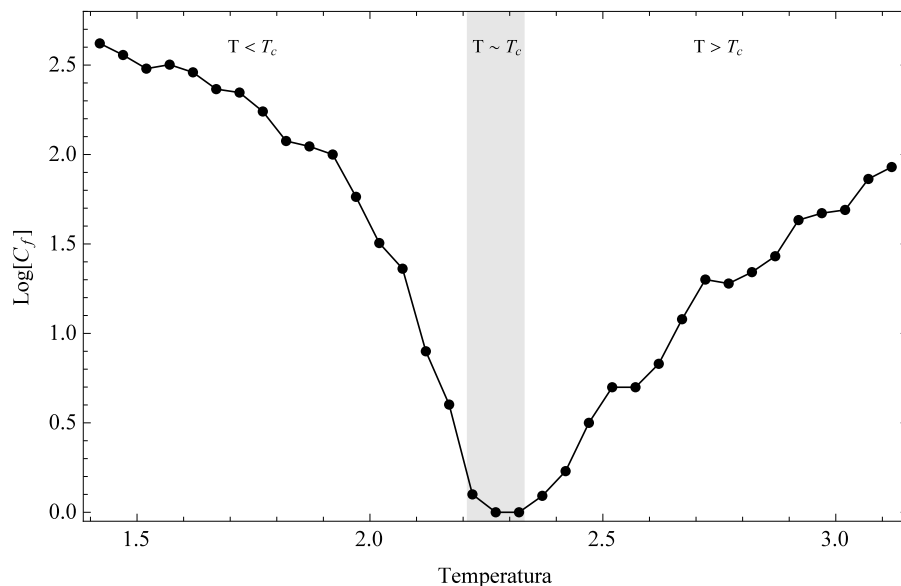
**Figura 14.** Espectro de potencias del modelo de Ising. Comportamiento del espectro de potencias del modelo de Ising como función de la temperatura. La imagen a) muestra el comportamiento para  $T \leq T_c$ ; mientras que la imagen b) lo muestra para  $T \geq T_c$ . Los puntos remarcados son los puntos de corte que se encontraron al ajustar dos rectas al espectro.

liza un ajuste de una recta, se calcula la incertidumbre y se guarda en la misma lista donde se guardó la incertidumbre anterior. Después se hace el ajuste de una constante a 6 datos iniciales y así sucesivamente hasta llegar a ajustar una recta constante a 3996 datos y a los últimos 4 una recta<sup>17</sup>. Tomando la lista donde se guardaron las incertidumbres de todos los ajustes, se elige la incertidumbre mínima y el punto en donde tal ajuste pasa de ser una constante a una recta se le denomina el punto de corte  $C_f$  del espectro de potencias.



**Figura 15.**  $C_f$  como función de la temperatura en el caso del ajuste de una constante y una recta. Se muestran 3 zonas:  $T < T_c$ ,  $T \sim T_c$  y  $T > T_c$ .

<sup>17</sup> La razón de elegir detener el programa en este punto es la misma que en el pie de página anterior



**Figura 16.**  $C_f$  como función de la temperatura en el caso del ajuste de dos rectas.

Se muestran 3 zonas:  $T < T_c$ ,  $T \sim T_c$  y  $T > T_c$ .

El segundo subprograma hace exactamente lo mismo que el anterior pero en vez de ajustar una constante y una recta se ajustan dos rectas. Los resultados de los dos subprogramas se muestran en las imágenes 15 y 16. Para el caso de la figura 16, el ajuste de dos rectas deja de funcionar para algunos valores dentro de la zona crítica, por lo que manualmente se ajustaron tales puntos a cero dentro de esa región. De nuevo, como en los casos anteriores se pueden distinguir marcadamente 3 zonas:  $T < T_c$ ,  $T \sim T_c$  y  $T > T_c$ . El hecho de que haya una clara disminución de los valores de  $C_f$  al aproximarse a la zona crítica tanto por arriba como por abajo está asociada a la invariancia temporal que alcanza el sistema, como se discutió en la sección anterior. Tan pronto como se sale de la región crítica, el sistema deja de estar correlacionado a tiempos largos y el espectro de potencias se aplanan, ocasionando que el punto  $C_f$  vaya aumentando de posición en el espectro de potencias. En principio, tómesese en consideración que este

programa, al tomar en cuenta una gran cantidad de ajustes de rectas para una gran cantidad de datos, es muy susceptible a variaciones, lo que le puede dar a las curvas un aspecto no tan suave como se había mostrado en el resto de las imágenes del presente trabajo, sobre todo en el caso de ajuste de dos rectas, en donde el programa se vuelve aún más susceptible a variaciones. Cabe decir que la teoría menciona de antemano la existencia de las dos regiones para  $T > T_c$  y de una recta negativa cuando  $T \sim T_c$  asociada a la invariancia de escala, por lo que era de esperarse que hubiera un cambio negativo en el valor  $C_f$  en los dos ajustes si se recorría el sistema desde un punto por encima de la temperatura crítica hasta llegar a la zona crítica; esto debido a que mientras la temperatura se aproxime más a la temperatura crítica, la ecuación (79) del capítulo 2 va reduciendo el término  $(T - T_c)^2$  y por ende, se alcanza el comportamiento de recta a frecuencias cada vez menores hasta llegar a un punto (el crítico), en donde el espectro de potencias sea una recta, razón por la cual  $C_f \approx 0$ . Lo que sorprende aquí es que la propuesta hecha en el capítulo 2 respecto a que el espectro de potencias a temperaturas bajas tiene las mismas regiones que el espectro para temperaturas altas; las imágenes 15 y 16 verifican la propuesta puesto que fue posible ajustar dos regiones al espectro y ver un cambio gradual de  $C_f$  conforme se aumenta  $T$  en esta zona. Con esto se demuestra que tanto a temperaturas bajas como a altas, el sistema tiene una cierta similitud en intervalos de tiempo muy cortos. Sin embargo, mientras mayor es el tiempo el sistema pierde toda la similitud consigo mismo, por lo que la configuración del sistema es completamente independiente de la configuración del sistema a un tiempo anterior muy largo. Con esto es posible ver que el sistema no tiene memoria a largo alcance para esta zona de temperaturas. La zona crítica es, entonces, la zona donde el sistema adquiere memoria tanto a largo como a corto plazo, alcanzando con esto la invariancia de escala.

# Capítulo 4

## Discusión y conclusiones.

Como pudo verse en todas las figuras del capítulo anterior, existe una marcada diferencia entre los tres regímenes de temperaturas analizados:  $T < T_c$ ,  $T \sim T_c$  y  $T > T_c$ , lo que permite establecer que en general había tres comportamientos en cada una de las figuras, por lo que dependiendo del cambio de los valores de cada una de las alertas tempranas en cada región se puede establecer su utilidad como alerta temprana. Por ejemplo, si se proviene de una temperatura muy baja y se aproxima a la crítica, es de muy poca utilidad como alerta temprana la varianza, mientras que para la misma región lo podría muy bien ser la curtosis o la auto-correlación. La tabla 1, muestra un resumen de las alertas tempranas propuestas en el presente trabajo indicando qué alertas tempranas son mejores para cuando se aproximan a la temperatura crítica por debajo, por arriba y aquellas alertas tempranas que funcionan bien para ambos casos (son simétricas). Esta división se realiza solamente analizando las gráficas, una división mejor que pueda indicar cuál es la mejor alerta temprana para el modelo de Ising necesitaría otros métodos aunque bien podría obtenerse una buena aproximación observando como va cambiando la pendiente de cada uno de las alertas tempranas.

Con todo esto, puede concluirse que a excepción solamente del promedio (en donde el comportamiento del sistema no muestra cambios tan abruptos antes o después de la zona crítica), todas las alertas tempranas propuestas podrían funcionar bien como indicadores de la proximidad de un cambio abrupto del sistema desde regiones incluso lejanas al cambio mismo. Recuérdese aquí que las alertas tempranas de la correlación y de la varianza reflejan un comportamiento

de ralentización del sistema. Dado que el punto de corte ( $C_f$ ) proviene de un análisis del espectro de potencias y que éste último está asociado a la correlación, puede decirse que esta alerta temprana proviene también de la ralentización del sistema. Es de recalcar que mientras los primeros son distribuciones marcadamente asimétricas, el último es más bien una alerta temprana que puede ser usado tanto si se viene por arriba de la temperatura crítica como si se viene por debajo de la misma. Es interesante darle una interpretación a las figuras 15 y 16 en términos de la transmisión de información. Se había dicho ya que el ajuste de una constante y una recta indica que tan “invariante de escala” resulta el sistema. Para esto, que el valor de  $C_f$  baje indica que el sistema está correlacionándose temporalmente más para periodos de tiempo cada vez mayores. En un sentido estricto, una variación ligera del sistema, empieza a ser relevante del sistema para periodos de tiempo cada vez mayores o lo que es lo mismo, una característica del sistema, por más pequeña que esta sea tiende a afectar a gran escala el sistema por completo. Se puede pensar en esto como que el sistema permite transmitir información y esta información tiende a propagarse cada vez a distancias mayores. Las figuras 15 y 16 indican por ende que para zonas cercanas a la crítica, la información puede ser transmitida a todas las escalas. Este comportamiento de transmisión de información se va perdiendo poco a poco tanto si se baja la temperatura a partir de la crítica como si se aumenta. Aún más, dado que esta alerta temprana no es completamente simétrica, según lo obtenido, la información tiene una probabilidad mayor de ser transmitida si está por encima de la temperatura crítica que si se está por debajo. Esto tiene grandes implicaciones físicas, así como grandes repercusiones en el caso de que el modelo de Ising quiera ser extrapolado a modelos sociales, económicos o biológicos, ya que medir qué tanto se expande la información en un sistema puede indicar si se encuentra por encima o por debajo del punto crítico .

Dada las similitud entre las figuras 15 y 16 se podría pensar que ambas son intercambiables y que tal vez, una aproximación de dos rectas entre ambas imágenes sería algo más realista. En ambos casos se calcularon los errores de los ajustes obteniendo incluso un valor de error ligeramente menor en los ajustes de dos rectas respecto al ajuste de una constante y una recta. Además, la forma de ambas gráficas es prácticamente la misma. Sin embargo mientras que en la figura 16 algunos valores de la zona crítica debieron ser ajustados a cero, los valores del punto de corte en la zona crítica de la figura 15 salieron de forma natural y sin tanto ruido de fondo, es decir, son más suaves.

Early Warning Utilidad	$T_c > 0$	$T_c < 0$	Simétricos
Promedio	X		
Varianza		X	
Curtosis			X
Asimetría estadística, b	X		
Autocorrelación a $\tau = 1$	X		
$C_f$			X

**Tabla 1.** Utilidad de las alertas tempranas dependiendo la forma de acercamiento.

En la primera columna se dice qué alertas tempranas son mejores si la temperatura es menor que la crítica pero se acerca a la crítica (acercamiento a  $T_c$  por la izquierda).

La segunda indica las alertas tempranas que resultan mejores si se realiza una acercamiento desde una temperatura muy alta hasta la crítica (acercamiento a  $T_c$  por la derecha). La última columna indica las alertas tempranas que muestran variación similar tanto por la izquierda como por la derecha.

Cabe aclarar que en este caso se supuso que la pendiente que ajusta la parte de frecuencias bajas para el espectro de potencias podría ser aproximada a una constante, es decir, tenía un pendiente igual a 0, sin embargo esto no podría ser siempre cierto en el caso de que se trabajen con datos experimentales, por lo que el uso de los dos métodos es lo más apropiado para poder analizar el punto de crossover del sistema en caso de que presente invariancia de escala temporal. El determinar que método es mejor dependerá mucho del sistema que se pretenda analizar por lo que está fuera del alcance del presente trabajo.

Hay que tomar en cuenta que muchas de las propiedades del modelo de Ising (incluyendo la temperatura crítica del sistema) pueden cambiar dependiendo de el número de espines que se usen en la simulación. Se eligieron mallas de tamaños 60x60, 80x80, 120x120 y 140x140. Para cada una se repitieron las mismas mediciones que se hicieron para la malla de 100x100. En todos los casos, el patrón que siguieron las alertas tempranas fue prácticamente el mismo salvo una diferencia en la escala. Esto demuestra que el comportamiento de las alertas tempranas es independiente del tamaño del sistema. Para el espectro de potencias, la conclusión es la misma, incluso el punto de corte se encuentra exactamente en el mismo lugar en todas las escalas.

Resulta ilustrativo hacer un análisis de los resultados en términos del potencial de bifurcación. En principio, véase que en la aproximación de campo medio puede llegarse, sin mucha complicación a que el modelo de Ising satisface, la ecuación diferencial:

$$\frac{dM(t)}{dt} = -M(t) + \tanh[\beta M(t)] \quad (98)$$

Haciendo una expansión hasta segundo orden se llega a que:

$$\frac{dM(t)}{dt} \approx -(\beta_c - \beta)M(t) - \frac{1}{3}(\beta M(t))^3 \quad (99)$$



Esta aproximación muestra que el modelo de Ising puede ser aproximado a una bifurcación supercrítica de Pitchfork, muy común en problemas físicos que presentan simetría. A pesar de no ser una bifurcación no catastrófica, se ha demostrado [1] que presenta algunas de las principales alertas tempranas como el promedio, la varianza y la auto-correlación  $\tau = 1$ . La gráfica de la bifurcación que representa la ecuación (98) es muy parecida a la mostrada en el capítulo 1 respecto a la magnetización en el modelo de Ising (figura 3a) así como a la del valor de la magnetización promedio mostrada en el capítulo 3 (figura 9). Por otro lado, recordando también la aproximación de campo medio, en el caso particular del modelo de Ising la energía libre toma la forma particular:

$$F(M) \approx C + \frac{1}{2}(\beta_c - \beta)M^2 + \frac{1}{12}\beta^3 M^4 \quad (100)$$

Donde de igual forma hay dos puntos estables por debajo de la temperatura crítica y un punto estable ( $M=0$ ) por encima de la misma. Por lo mismo, el sistema se vuelve biestable para temperaturas inferiores a la crítica mientras que para temperaturas por encima el sistema se vuelve estable. En el punto crítico se da el paso del sistema biestable al sistema con un único punto estable. La figura 17 muestra el comportamiento del flujo para la ecuación general de Pitchfork supercrítica ( $\frac{dx}{dt} = \mu t - t^3$ ). Los puntos en donde la función corta al eje X, son los puntos estables del sistema. Véase que la ecuación tiene sólo una raíz real cuando  $\mu \leq 0$ , mientras que para  $\mu > 0$ , hay 3 raíces. Extrapolando esto al modelo de Ising, hay que recalcar dos aspectos. En el capítulo 2 se demostró que la teoría de campo medio era inservible para obtener resultados exactos en dimensiones menores a 4, por lo que este análisis es meramente descriptivo y no indica los puntos exactos de equilibrio o desequilibrio del sistema. Por otro lado, la aproximación en el punto crítico se indefine por lo que tampoco es bueno esperar resultados exactos para todas las temperaturas con esta aproximación.

Sin embargo, es una forma esquemática y que describe el porqué hay un cambio abrupto en la dinámica del sistema cuando se pasa por el punto crítico, ya que abajo de la temperatura crítica, el sistema está en una zona biestable y por ende tiende a irse por completo hacia la rama positiva o hacia la negativa. Claramente, por la definición misma del cambio de espín dada en el capítulo 2, mientras más alejado esté el sistema del crítico, más rápido se llegará a un punto fijo del sistema. Por encima del punto crítico el sistema tiene un punto crítico estable en el cero. Por otro lado, por debajo del crítico hay dos puntos fijos estables. El cero también es un punto fijo en el sistema en este caso pero es inestable. El caso que analiza el modelo de Ising por ende se vuelve sumamente interesante: el cero, un punto fijo claramente estable debajo del crítico pasa a ser un punto fijo inestable. Con esto puede entenderse el porqué de la variación en el promedio magnético desde un punto de vista más ligado a la teoría de sistemas dinámicos que desde el punto de vista físico. Además, la asimetría del potencial ayuda también a entender el porqué la varianza y la auto-correlación a  $\tau = 1$  son asimétricas. En efecto, dado que a partir del punto de bifurcación pase de ser biestable a tener sólo punto estable hace que la aproximación desde temperaturas bajas a la temperatura crítica sea completamente diferente que la aproximación de temperaturas altas a la temperatura crítica, creando con esto la asimetría en estas dos alertas tempranas.

Con el análisis del potencial es posible también ver que existe un importante conexión entre el punto de bifurcación y la temperatura crítica. En el caso del modelo de Ising, resulta que ambos puntos son el mismo por lo que en muchos caso para este modelo puede hablarse indiferentemente del punto de bifurcación y el punto crítico aunque es importante recalcar que el punto de bifurcación dinámica del modelo de Ising está referido al punto en donde el sistema pasa de ser biestable a tener sólo un punto de estabilidad, mientras que el punto crítico

está referido como el punto a partir del cual hay una transición de fase en el sistema.

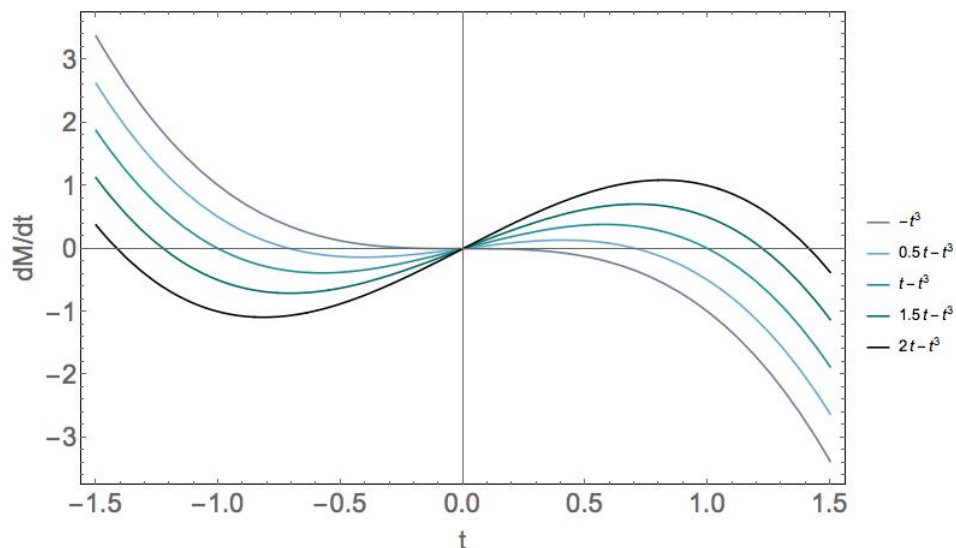
El caso de la varianza en el sistema es también interesante si se analiza desde el punto de vista de las gotas de Peierls [19]. Las gotas de Peierls son “islas” de espines que apuntan hacia la misma dirección. Supongamos que existe una gota tal que en una vecindad grande fuera de ella todos los espines apuntan hacia una dirección opuesta a la dirección que apuntan los de la gota. En teoría, si el perímetro de la gota es de  $L$ , al sistema creará esa gota con un costo de energía  $\Delta E = 2JL$ , dado que sólo le costará modificar a los espines del perímetro. Para encontrar la entropía de generar esa gota, se debería preguntar cuántas gotas de perímetro  $L$  se podrían generar. Es difícil responder dado que las gotas podrían tener formas arbitrarias pero se puede encontrar que el número posible de tales configuraciones está acotado tanto por encima como por debajo. Si se quieren formar gotas en 2 dimensiones con la única posibilidad de ir arriba, abajo, izquierda y derecha (esto es, consideramos que estamos en una malla cuadrada), se podrían originar  $4^L$  caminatas aleatorias, pero ese sería un mal estimador. De hecho, al elegir alguna dirección en realidad sólo nos quedan, a lo más, tres posibles elecciones para que la gota no se “regrese” sobre sí misma mientras que en un caso extremo, sólo nos quedarían 2 direcciones a elegir.

Esto dice que el número de configuraciones posibles para una gota de perímetro  $L$ ,  $n_L$  está acotado por:

$$2^L < n_L < 3^L \quad (101)$$

Sea  $C$  el número de elecciones promedio que podría tener la gota al estar formándose, en tal caso, el cambio en la energía libre sería:

$$\Delta F = \Delta E - T\Delta S = L(2J - k_B T \log C) \quad (102)$$



**Figura 17.** *diagrama de flujo de la forma general de la bifurcación de Pitchfork supercrítica para los valores de  $\mu = 0, 0.5, 1, 1.5$  y  $2$ . Los puntos en donde corta al eje  $X$  son los puntos fijos del sistema. Puede verse que para  $\mu = 0$  el  $0$  es un atractor, mientras que para los demás casos se convierte en repulsor. En los demás casos, el sistema presenta dos puntos fijos más y cada uno es estable. Esto ayuda a comprender el porqué de la bifurcación en la magnetización en el modelo de Ising*

El signo de  $\Delta F$  estará dado por el término  $2J - k_B T \log C$ . Claramente si

$$T > \frac{2J}{k_B \log C} \quad (103)$$

$\Delta F < 0$  por lo que se favorece el crear islas de cualquier tamaño, destruyendo cualquier magnetización espontánea. Eventualmente estas gotas se convierten en islas enormes de espines irán juntándose entre si y el sistema elegirá alguno de los dos puntos fijos mencionados anteriormente.

Por otro lado, si

$$T < \frac{2J}{k_B \log C} \quad (104)$$

entonces  $\Delta F > 0$  y el sistema será adverso a la creación de islas de cualquier tamaño. Nótese que esto se cumple apenas por encima de la temperatura crítica, mientras que en el otro caso, si el sistema está cercano a la temperatura crítica pero por debajo de la misma, las magnetizaciones espontáneas pueden jugar un papel importante. Por último está el caso de:

$$T = \frac{2J}{k_B \log C} \quad (105)$$

Puede demostrarse que la ecuación (105) coincide con la temperatura crítica del sistema por lo que el sistema es libre de formar y destruir gotas de todos los tamaños [41].

Este pequeño análisis en términos de gotas ayuda a entender bastante el comportamiento de la varianza en el sistema. A temperaturas menores que la crítica, el sistema tiende a crear islas grandes. Mientras más se aproxime a la temperatura crítica por debajo las fluctuaciones estocásticas del sistema serán cada vez más importantes, creando así un aumento en la varianza del sistema. Justo en la temperatura crítica, el desastre en el sistema es de tal magnitud que las grandes gotas conviven con espines al azar haciendo que la varianza aumente en gran medida. Después de la temperatura crítica los espines al azar se encuentran dentro de todo el sistema, por lo que la varianza del mismo vuelve a ser cero. Exactamente esto mismo fue lo obtenido en la figura 10.

De igual forma, puede hacerse un análisis más a fondo de la asimetría estudiando la aproximación b dada por la ecuación (86). Para la región  $T < T_c$ , se espera que a tiempo  $t = 1000$  existan regiones de gran tamaño de espines totalmente alineados con la preponderancia de regiones con espín +1, debido a

lo mencionado anteriormente. Aún con esto, para este tiempo es de esperarse que continúen existiendo regiones importantes de espines  $-1$ . Para tiempos mayores, las regiones de espín  $-1$  desaparecen para que toda la malla se encuentre prácticamente alineada con espín  $+1$ . Supóngase que a un tiempo  $t$ , se tienen  $q_t$  espines negativos. Como  $T$  es baja respecto a la crítica,  $\bar{M} \approx 1$ , y la ecuación de estimación de la asimetría  $b$  puede aproximarse como:

$$b_t [M] = \frac{\frac{1}{n} \sum_{i=1}^n (m_i - 1)^3}{\left[ \frac{1}{n-1} \sum_{i=1}^n (m_i - 1)^2 \right]^{3/2}} \approx \frac{\frac{1}{n} \sum_{i=1}^{q_t} (-2)^3}{\left[ \frac{1}{n-1} \sum_{i=1}^{q_t} (-2)^2 \right]^{3/2}}$$

$$b_t [M] \approx -\frac{(n-1)^{3/2}}{n} \sqrt{\frac{1}{q_t}} \quad (106)$$

Si  $n$  es muy grande, puede aproximarse a

$$b_t [M] \approx -\sqrt{\frac{n}{q_t}} < 0 \quad (107)$$

De esta forma puede verse que el parámetro  $b$  resulta ser negativo para valores inferiores a la temperatura crítica. Véase que mientras más vaya aumentando la temperatura se espera que el valor promedio de espines negativos vaya aumentando y por ende el valor de  $b$  se aproxime cada vez más a cero. Posteriormente, a temperaturas cada vez mayores, el promedio se aleja de 1 y los espines positivos empiezan a tomarse en cuenta causando el cambio abrupto para  $T \sim T_c$ . Por otro lado, para la región  $T > T_c$ ,  $\bar{M} \approx 0$ , por lo que la asimetría a tiempo  $t$  toma la forma:

$$b_t [M] \approx \frac{\sqrt{n} \left[ \left( \sum_{i=1}^{p_t} (1)^3 \right) + \left( \sum_{i=1}^{q_t} (-1)^3 \right) \right]}{\left[ \sum_{i=1}^n (\pm 1)^2 \right]^{3/2}} = \frac{p_t - q_t}{n} \quad (108)$$

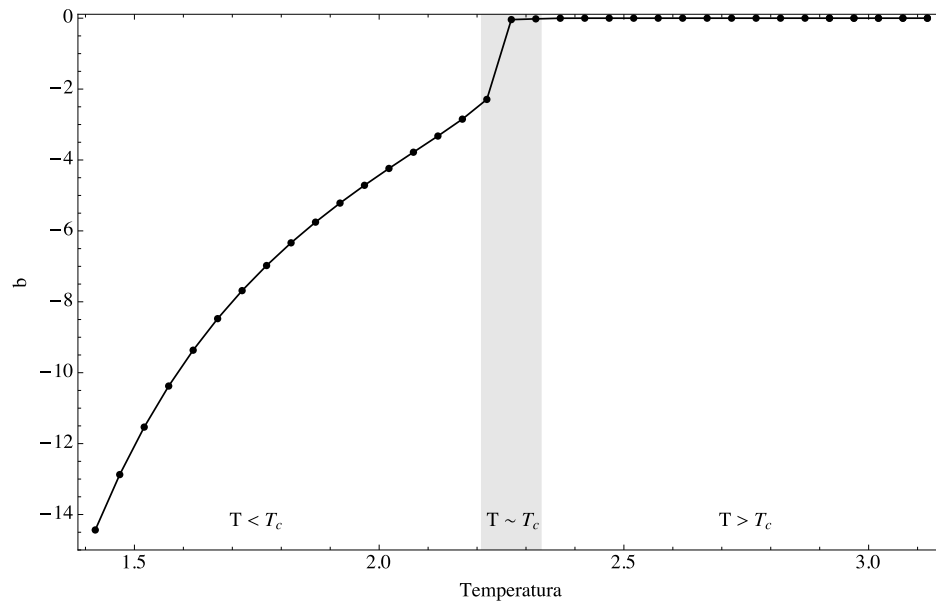
Con  $p_t$  el número total de espines hacia arriba y  $q_t$  el de espines hacia abajo. Como en esta zona hay prácticamente el mismo número de espines hacia arriba

que hacia abajo,  $p_t - q_t = 0$  y por ende el skewness total se vuelve cero. En la zona crítica se da un descenso abrupto de la magnetización lo que provoca un incremento hasta llegar al punto en donde  $b = 0$ . Con los datos obtenidos, puede también hacerse una serie de tiempo de  $b$  y demostrarse que también puede funcionar como una alerta temprana. La figura 18 muestra la gráfica de  $b$ . Nótese que al igual que en la gráfica del skewness, hay un cambio drástico de los valores en la región donde  $T < T_c$  mientras que hay un cambio nulo para la región donde  $T > T_c$ . Nótese aquí algo interesante. En ambos casos los valores van cambiando mientras la temperatura se aproxima a la crítica por debajo, sin embargo, cerca del punto crítico, la pendiente del skewness aumenta considerablemente mientras que para  $b$  la pendiente disminuye. Así que, aunque las dos gráficas representen una manera de medir la asimetría estadística, si el parámetro utilizado es el skewness entonces al llegar cerca del crítico por debajo habrá un incremento considerable de la pendiente, mientras que para  $b$  la gráfica se irá haciendo cada vez más plana. De cualquier manera ambas son buenas alertas tempranas para zonas inferiores a la crítica y el método a emplear dependerá mucho de la escala de resolución con la que se cuente para medir las propiedades del sistema (para  $b$  la escala es mucho mayor que para el skewness) y del tiempo de cómputo que se necesite (el cálculo de  $b$ , resulta en muchos casos mucho más sencillo debido a que no necesita calcular desviaciones estándar sino trabajar con el promedio).

Con lo visto hasta el momento puede llegarse a la siguiente conclusión: el papel que juegan las fluctuaciones estocásticas es primordial. Éstas hacen que la correlación sea baja en zonas alejadas del punto crítico, así como originan la asimetría en la varianza. También son las que, debajo de la temperatura crítica, deciden qué ramal de la bifurcación de Pitchfork elegirá el sistema, por lo que en todas las alertas tempranas basadas en la métrica jugarán un papel

fundamental.

Muchas investigaciones futuras pueden acompañar el presente trabajo. En primer lugar un análisis de las alertas tempranas del modelo de Ising en dos dimensiones con espín = 1 puede ser de gran utilidad, así como un análisis de modelos como el XY o el de Heisenberg. Sería interesante poder comparar los resultados obtenidos con los resultados en esos modelos y verificar si es que existió algún cambio importante en las alertas tempranas. Posteriormente sería interesante un análisis más a fondo tomando en cuenta la distribución espacial de los espines.



**Figura 18.** Gráfica de  $b$  como función de temperatura. Puede verse que para temperaturas inferiores a la crítica y al igual que el skewness, el valor de  $b$  puede servir como una buena alerta temprana del sistema.

Una particularidad del modelo de Ising es que todo indicador espacial puede



ser analizado también de manera temporal. En muchos casos, como se hizo en la magnetización, la varianza, la auto-correlación y la asimetría, se trató en este trabajo de dar una explicación espacial a lo que sucede temporalmente. Esto con el fin de entender por qué físicamente sucede lo que se ve en las series de tiempo pero también para poder dar pie a generar alertas tempranas espaciales. Actualmente existen técnicas de análisis de imágenes en donde se podrían analizar directamente las configuraciones espaciales de los espines y utilizar esto como una alertas temprana. Una propuesta de alerta temprana que estaría directamente asociado a la varianza (y por ende, a la correlación del sistema) sería utilizar las gotas de Piereles como una alerta temprana. En efecto, el objetivo consistiría en elaborar un programa que pueda contar la cantidad de gotas y el tamaño de las mismas que tiene la configuración de Ising a una temperatura dada y utilizar eso como una alertas temprana. Esta alerta podría incluso ser de gran utilidad para a otras ramas de la ciencia ya que muchos de las alertas tempranas en la literatura [4-6, 31-32] son propuestas de cambio en la configuración espacial. Por ejemplo, las referencias [31, 32] trabajan con un análisis espacial de la distribución de parches de vegetación y de cambios dinámicos en ecosistemas. Para esto, se crean autómatas celulares o modelos de sistemas dinámicos que permitan simular en una región el comportamiento espacial del sistema, por lo que una alerta temprana basada en el análisis de imágenes resultaría muy útil para este tipo de casos incluso cuando esta alerta temprana no proviniera propiamente de modelar Ising. Medir, por ejemplo, el tamaño promedio de las manchas de vegetación como un indicador de que tan cercano o alejado se encuentra un sistema de un punto estable o un punto crítico. Regresando al modelo de Ising, sería interesante analizar la forma en que se comporta esta alerta temprana aunque se esperaría, dada la discusión mostrada anteriormente que a temperaturas grandes el tamaño promedio de las gotas tienda a cero, mientras

que a temperaturas bajas tienda al área de la malla.

Otro posible trabajo a futuro es poder corroborar lo aquí presentado en algún modelo real. Como ya se ha mencionado, muchos sistemas en otras ramas de la ciencia pueden ser aproximados usando el modelo de Ising. Comparar cómo suceden los cambios abruptos en algún modelo de Ising construido a partir de datos experimentales (emanados tanto de experimentos a sistemas ferromagnéticos como de datos de desarrollo del cáncer [4], electroencefalogramas [2-3], electrocardiogramas [6] o de dinámica social [7-8]) respecto al modelo de Ising “idealizado” aquí presentado sería un resultado bastante enriquecedor y una prueba contundente de la aplicación de todos los resultados del presente trabajo.

Por último y a manera de conclusión, el análisis de las alertas tempranas en el modelo de Ising ayudó mucho a comprender el comportamiento general de éstas bajo circunstancias muy controladas. El propósito final de esto es que entendidos bien estos comportamientos, puedan ser utilizadas en sistemas que se encuentren lejos de poder ser controlados y en donde no se conozca de antemano la cercanía al punto crítico.

## Referencias

1. Kéfi, Sonia, et. al. Early Warnings signals also precede non-catastrophic transitions.
2. Marinazzo D, et. al. (2014), Information transfer and criticality in the ising model on the human coennectome, PLoS ONE 9(4): e93616.
3. Dast TK, Adeyasinghe PM, Crone JS, et. al. (2014) Highlighting the structure-function relationship of the brain wuth the ising model and graph theory. Biomed Res Int: 237898
4. Torquato S (2011), Toward an ising model of cancer and beyond. Phys Biol 8(1): 015017
5. Liu Y, Dilger JP(1993) Application of the one and two dimensional ising model to studies of cooperativity between ion channels. Biophys J 64(1): 26-35
6. Rice JJ, Stolovitzky G, Tu Y, Tombe PP (2003), Ising Model of cardiac filament activation with neares- neighbor cooperative interactions. Biophys J 84:897-909
7. Jarman, M et. al. (2015), The Critical Few: Anticonformist at the Crossroads of Minority Opinion Survival and Collapse, J. Artificial Sci. and Soc. Sim.18(1): 6
8. Castellano C, Fortunato S, Loreto V (2009), Statistical physics of social dynamics. Rev Mod Phys 81: 591-646
9. Yeomans, J M (1992), Statistical Mechanics of Phase Transitions, Oxford University Press.
10. Callen , H (1985), Thermodynamics and Introduction to Thermostatistics, Second Edition, John Wiley & Sons
11. Saxon, DS (1968), Elementary Quantum Mechanics, Holden Day
12. Landau, LD, Lifshitz, EM (1958), Course of Theoretical Physics, Vol. 5, Statistical Physics, Pergamon Press LTD.
13. Solé, Ricard V (2011), Phase Transitions, Princenton University Press
14. Brush G. Stephen (1967), History of Lenz- Ising Model, Rev Mod Phys, 4: 883-893
15. Cugliandolo, Leticia F. (2009), Satatistical Mechanis and Stochastic Processes, Université Pierre et Marie Curie- Páris VI
16. Gelover-Santiago, Angélica (2005), Simulación del Modelo de Ising con el Método Monte Carlo, 2da ed. Las Prensas de Ciencias

17. Chandler, David (1987), Introduction to Modern Statistical Mechanics, Oxford University Press
18. Sethna, James P. (2006), Entropy, Order Parameters and Complexity, Oxford University Press.
19. Gonsalves, Richard J. (2011), “5.2 Statistical Physics: Ising Model and Classical Gas”, visto en <http://www.physics.buffalo.edu> (consultado el 20 de Octubre del 2015)
20. Brockwell Peter J, David Richard (1987), Time Series: Theory and Methods, Spirnger Series in Statistics, Springer- Verlag
21. Chatfield Chris (1995), The Analysis of Time Series an Introduction, 5ta Edición, Chapman & Hall/CRC
22. Champeney DC (1989), A Handbook of Fourier Theorems, Cambridge University Press
23. Boekgaard K (1994), Scaling and Dynamics of Laplacian Growth, Random Lattices, Ising Models & Interfaces, Aarhus University, Denmark
24. S. K. Ma (1976), Modern Theory of Critical Phenomena, Frontiers in Physics 46
25. Larsen Jan (2009), Correlation Functions and Power Spectra, 8th ed, Technical University of Denmark
26. Knill Oliver (2009), Probability Theory and Stochastic Processes with Applications, Overseas Press
27. Sheldon M. Ross (2010), Introduction to Probability Models, 10th Ed, Academic Press
28. Lemons Don S. (2002), An Introduction to Stochastic Processes in Physics, The John Hopkins University Press
29. Scheffer M, Bascompte J, Brock WA, et. al. (2009) Early warning signals for critical transitions, Nature 461: 53-59
30. Sceffer et. al. (2012), Anticipating Critical Transitions, Science 338: 344-348
31. Donagelo Raul, Fort Hugo, Scheffer M, Dakos Vasili (2010), Early warning for catastrophic shifts in ecosystems: comparison between spatial and temporal indicators, Int J of Bif and Chaos, 20 (2): 315-321
32. Rietker Max, Dekker Stefan C, de Ruiter Peter C, van de Koppel Johan (2004), Self-Organized Patchiness and Catastrophic Shifts in Ecosystems, Science 305: 1926- 1929

33. Kéfi S, Guttal V, Brock WA, et. al. (2014) Early warning signals of ecological transitions: Methods for spatial patterns. *PLoS ONE* 9(3): e92097
34. Chisolm, RA & Filotas (2009), E; Critical slowing down as an indicator of transitions in two-species models, *J Theor. Biol.* 257: 142-149
35. Kleinen Thomas, Held Hermann, Petschel-Held Gerard (2002), The potential role of spectral properties in detecting thresholds in the Earth system: application to the thermaline circulation, *Ocean Dynamics* 53: 53-63
36. Dakos V et. al. (2008) Slowing down as an early warning signal for abrupt climate change, *Proc Natl Acad Sci USA*, 104: 14308-14312
37. Lenton T M et. al. (2009). Using GENIE to study a tipping point: the climate system. *Phil. Trans. R. Soc A* 367: 871-884
38. Arveinud, E (2002), Calm before the storm? Low volatility often precedes market downturn, *Barron's*, (Enero 28, 2002)
39. Lebaron, B. (1992), Some relations between volatility and serial correlations in stock-market returns, *J. Bus.* 65: 199-219
40. Dakos, V, "Why Should we expect Early Warnings before Critical Transition?", Recuperado de: <http://www.early-warnings-signals.org> (Consultado el: 20 de Octubre 2015), Traducido por Joel Mendoza
41. Gonsalves, Richard J(2008), "Statistical Mechanics, Phase Transitions and the Ising Model", Recuperado de <http://www.physics.buffalo.edu> (Consultado el 20 de Octubre del 2015)
42. Veytsman, Boris (1997), "Phase Transitions of Ising Model", Recuperado de <http://www.boris.lk.net> (Consultado el 20 de Octubre del 2015)
43. Fitzpatrick, Richard (2006), "The Ising Model", Recuperado de <http://www.farside.ph.utexas.edu> (Consultado el 20 de Octubre del 2015)

# Apéndice

Gran parte de los resultados de la presente tesis se desprende del artículo “*Behavior of Early Warnings near the Critical Temperature in the Two-Dimensional Ising Model*”, del cual el autor de este trabajo es coautor. El artículo fue entregado a la revista PlosOne en Octubre del 2014 y en Junio del 2015 fue aceptado para su publicación. En las siguientes páginas se reproduce íntegramente.

RESEARCH ARTICLE

# Dynamical Signatures of Collective Quality Grading in a Social Activity: Attendance to Motion Pictures

Juan V. Escobar<sup>1,2\*</sup>, Didier Sornette<sup>3</sup>

**1** Instituto de Física, Universidad Nacional Autónoma de México, P.O. Box 20-364, México City, 04510, México, **2** Physics Department, Universidad Autónoma Metropolitana-Iztapalapa, México City, 09340, México, **3** ETH Zurich, Department of Management, Technology, and Economics, Switzerland and Swiss Finance Institute, Zürich, Switzerland

\* [escobarjuanvalentin@gmail.com](mailto:escobarjuanvalentin@gmail.com)



OPEN ACCESS

**Citation:** Escobar JV, Sornette D (2015) Dynamical Signatures of Collective Quality Grading in a Social Activity: Attendance to Motion Pictures. PLoS ONE 10(1): e0116811. doi:10.1371/journal.pone.0116811

**Academic Editor:** Dante R. Chialvo, National Scientific and Technical Research Council (CONICET), ARGENTINA

**Received:** August 16, 2014

**Accepted:** December 15, 2014

**Published:** January 22, 2015

**Copyright:** © 2015 Escobar, Sornette. This is an open access article distributed under the terms of the [Creative Commons Attribution License](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Data are available from <http://www.imdb.com>, <http://www.rottentomatoes.com/> and <http://www.boxofficemojo.com/>.

**Funding:** JVE acknowledges DGAPA's postdoctoral fellowship (CJIC/CTIC/2135/2009) and CONACYT's postdoctoral fellowship (project #177679, CVU 37905). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

We investigate the laws governing people's decisions and interactions by studying the collective dynamics of a well-documented social activity for which there exist ample records of the perceived quality: the attendance to movie theaters in the US. We picture the flows of attendance as impulses or "shocks" driven by external factors that in turn can create new cascades of attendances through direct recommendations whose effectiveness depends on the perceived quality of the movies. This corresponds to an epidemic branching model comprised of a decaying exponential function determining the time between cause and action, and a cascade of actions triggered by previous ones. We find that the vast majority of the ~3,500 movies studied fit our model remarkably well. From our results, we are able to translate a subjective concept such as movie quality into a probability of the deriving individual activity, and from it we build concrete quantitative predictions. Our analysis opens up the possibility of understanding other collective dynamics for which the perceived quality or appeal of an action is also known.

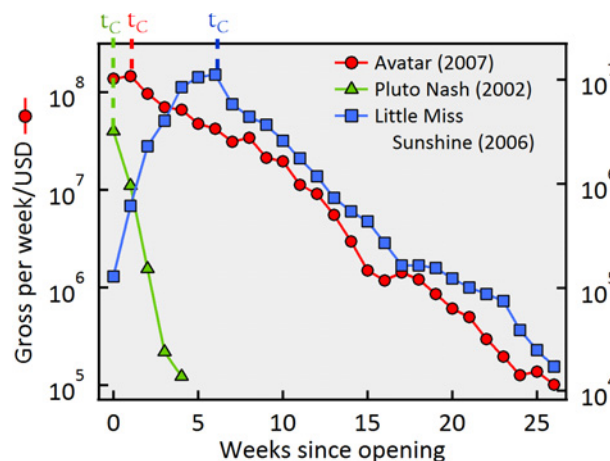
## Introduction

Despite the complexity of human interactions, understanding and ultimately predicting individual and collective human behavior would have deep implications in Economics and Psychology not to mention Finance and Advertising [1–5]. To this end, a fruitful approach is based on a parallel that exists in some cases between social and physical systems: when a portion of a social/natural system originally in some sort of equilibrium is driven out of it by an external perturbation (or "shock"), the laws governing its constituting people/particles determine how the information about the shock is spread within the system [1, 6, 7]. Studying the response function to given shocks, the laws behind people's decisions and interactions can be uncovered. Many aspects of collective social behavior are actually dominated by diffusion-like processes in which the topology of a social network determines how information is transmitted

[8, 9]. Dramatic evidence for this can be found in the online activity dynamics during the recent social movements of the so called Twitter revolution [9–12]. Of particular interest is to find the conditions that can bring the system to an “explosive” or “viral” state where information travels quickly and reaches almost the whole network. While government officials may fear a pandemic breakout or a social movement going viral, discovering exactly how to trigger such a state would constitute the wholly grail of advertisement [13, 14]. In this respect, compared to broadcasting, being exposed to direct opinions or recommendations from acquaintances tends to have a much deeper and longer lasting impact [8]. Indeed, word of mouth has been recognized to play a paramount role in turning social movements effervescent [8–12], on-line videos viral [6], books into best sellers [7], songs into hits [15], and movies into blockbusters [16, 17] or DVD commercial successes [18]. However, experience tells us that not every recommendation carries the same weight: the better we find a product to be or the more strongly we feel about a social cause, the more enthusiastically will we spread the word about it [19]. Thus, perceived quality is bound to play a fundamental role in the dynamics of social systems.

In this work, we attempt to gain insight into the nature of individual and collective social behavior by investigating the influence that the average perceived quality has on the aggregated dynamics of the attendance to ~3,500 movies in the US. Here, the information being propagated is the perceived quality of a movie, and we make the reasonable assumption that its dynamics is reflected directly in the box-office performance. Indeed, it has been shown that reviews have an effect on people’s decisions to watch a movie [20]. A concrete example of the effect of the popularity on the dynamics of this social system is displayed in Fig. 1, in which we have labeled  $t_c$  the number of weeks that have elapsed from the date of the opening to the week of the maximum attendance. In contrast to the worldwide blockbuster *Avatar* (2009), the attendance to a notorious flop such as *The Adventures of Pluto Nash* (2002) decays sharply after the opening week until it is finally taken out the theaters on week 5.

A characteristic normally used to group similar types of aggregated dynamics into classes is whether a burst of activity is driven by a large external (or Exogenous shock) event, or rather it is the cumulative result of smaller positive factors (or Endogenous shock) [21–25]. In Fig. 1, the time series of *Little Miss Sunshine* is an example of an Endogenous shock in this system, which relied on continuous recommendations from former viewers to create a multiplicative



**Figure 1. Attendance dynamics for movies representing different shocks and qualities.** Weekly gross vs. time for *Little Miss Sunshine* (squares, Endogenous), *Avatar* (circles, Exogenous, good movie) and *The adventures of Pluto Nash* (triangles, Exogenous, bad movie).

doi:10.1371/journal.pone.0116811.g001



effect that brought the audience to a maximum level many weeks after the opening day. The other two movies shown belong to the Exogenous shocks class, as the peak of attendance occurs on the first or second week. It is expected that the dynamics arising from exogenous and endogenous excitations will present different signatures from which key information about the social system can be extracted and theoretical models be validated.

We show below that the laws behind the relaxation dynamics exemplified in Fig. 1 can be understood, quantified and therefore predicted in terms of the movie's popularity. In a nutshell, the key concept behind our findings is quite intuitive: the better the movie, the more likely it is that a first-generation viewer will recommend it to her peers, creating an avalanche of attendances that will shape the dynamics of the attendance. The remarkable fact is that this qualitative intuition leads to precise reproducible quantitative predictions. Given that going to the movies is a widespread social activity in the U.S. (in which as many as fifty million people in total can watch a single movie), this activity constitutes an ideal social system to quantify the impact of the perceived quality of a product or the appeal of an action on the dynamics of the spreading of information in human activities.

The reader may question why we need to invoke an agent-to-agent microscopic mechanism describing a kind of epidemic interaction, while perhaps a simple mean-field or representative agent model such as the Bass diffusion model [26] might be sufficient. In this respect, de Vany and Lee [27] have shown the relevance of information cascades in the dynamics of motion picture's revenue, and have proposed an agent-based model to deal with the rich and intermittent nature of the dynamics of movie attendance. Our message is that the class of epidemic branching models that we use here is actually well-known to be of the "mean-field" class, allowing simple and exact analytical treatment. And to describe the dynamics of social influence, we use the simplest hierarchical decomposition provided by a branching framework that allows one to decompose the dynamics to exhibit important relationships between activity before and after peaks. Such predictions are beyond the reach of the Bass diffusion model or similar models, because they do not distinguish between generations of activity that are deeply associated with the quality of the product and the propensity for propagation in the community.

## Methods

### Dataset and Perceived Quality Function

The initial dataset consists of a collection of time series of the weekly revenue of about 10,000 movies that played in the U.S. during the period from 1970 to 2010 obtained from <http://www.boxofficemojo.com/> now owned by the Internet Movie Data Base (IMBD). In order to be able to perform an analysis of statistical significance, we study the post-peak dynamics of the attendance to those 3,469 (~35%) movies that played for at least four consecutive weeks in no less than 50 theaters. This data is complemented with the corresponding audience ratings given by a collection of American letter-system grades (*A*, *B*, *C*, *D*, *F*) provided by subscribers to this website. We assign to each movie a single normalized grade  $G \equiv (0.5A' + (7/20)B' + 0C' - (7/20)D' - 0.5F') + 0.5$ , where the symbols with apostrophes are the normalized number of votes a movie obtained for that particular grading letter. With this definition,  $G$  always lies between 0 and 1. Applying the simple selection criterion described above ensures that 99% of the movies analyzed have at least 60 ratings, a number that validates the statistical significance of the average normalized grade. Details on data filtering as well as the general trends of the data can be found in [S3 Appendix](#) and [S7 Appendix](#) respectively. The linear-logarithmic scales of [Fig. 1](#) suggest that the time series of activity  $\lambda(t)$  of the post-peak weekly gross for each movie can be

fitted to a decaying exponential [28] given by:

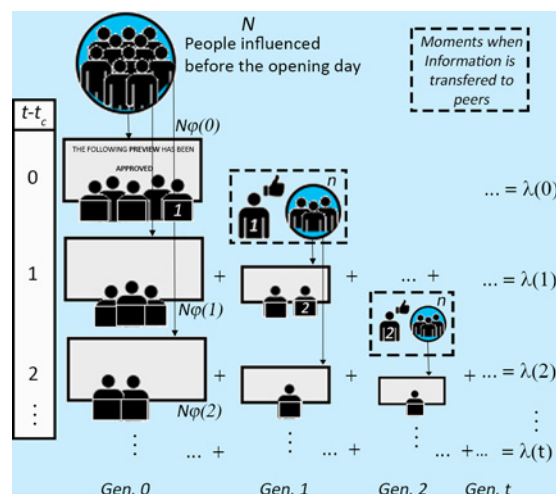
$$\lambda(t) = \lambda(t_c)e^{-(t-t_c)/\tau_0} \tag{1}$$

where  $\lambda(t_c)$  is the maximum activity at the peak occurring at time  $t_c$  and  $(1/\tau_0)$  is the decay rate of total activity for each movie equal to the inverse of their typical lifetime  $\tau_0$  in theatres. An algorithm similar to that used in other studies designed to fit noisy data [6] was employed to fit the time series to equation 1 (see complete methods in S11 Appendix). The data set is in general very well fit to equation 1, with an average  $R^2$  coefficient of 0.98 (see details on data filtering in S3 Appendix).

### Theoretical Model

To describe the dynamics of the attendance, we implement an epidemic branching model with latency effects [6, 7, 19, 29] whereby former viewers influence or “infect” acquaintances to watch the same movie with a probability that depends on its perceived quality. This model is known as the self-excited conditional Poisson process or “Hawkes” process [29], and is composed of two main ingredients. First, it is assumed that former viewers can influence new potential ones within their social network. Once they have already watched it and formed their own opinion, people from this new generation may recommend the movie themselves to their own acquaintances, and so on. This constitutes an epidemic branching process whose end result is the formation of a cascade of attendances characterized by a single parameter  $n$  called the “branching ratio” (see Fig. 2). The branching ratio is defined as the average number of events triggered in direct lineage by a single mother event. For our case of study, we expect the perceived quality of a movie to be correlated with  $n$ , as  $n$  quantifies the propensity for an epidemic of influences to propagate.

The second ingredient of the model incorporates the fact that, once a person has made a decision to perform an action, she will not do it right away, but rather in some future time. This latency effect is characterized by a function that gives the probability that she will perform an



**Figure 2. Schematics of the epidemic branching model, eq. 3.** Arrows point to the progressive attendance due to the latency effects given by  $\varphi(t)$ , while the dashed squares shows the moments when previous viewers influence or “infect” new potential viewers that constitute new generations. The total attendance at time  $t$  (right column) is the sum of the contributions from all previous generations. Only interactions from viewers labeled “1” and “2” are shown here, but all previous viewers influence an average of  $n$  peers. See S1 Appendix for details.

doi:10.1371/journal.pone.0116811.g002

action at a time  $t$ , given that she was influenced or “infected” at a previous time  $t_i$ . This function  $\varphi(t-t_i)$  is called the “bare kernel”, or “bare propagator” and is interpreted as the average time between cause and action of a single individual, which may vary depending on the specific human activity. In [30], this distribution is referred to as the “law of procrastination” resulting from an optimization by human agents to maximize the utility derived from their activities. From general priority queuing theory, far from criticality (where criticality is defined by the equality between the rate of incoming tasks and the rate of solving them), one can expect  $\varphi(t-t_i)$  to be an exponential function. Indeed, while in most social systems studied so far [5–7, 19, 31–34], the bare kernel is given by  $\varphi(t-t_i) \propto (t-t_i)^{-(1+\theta)}$  (where  $0 < \theta < 1$ ), here we find that a decaying exponential form gives an excellent description of the activity:

$$\varphi(t-t_i) = \varepsilon e^{-(t-t_i)/\tau}, \tag{2}$$

where  $0 < 1/\tau$  and  $\varepsilon = e^{1/\tau} (1-e^{-1/\tau})$  ensures the normalization of  $\varphi(t)$  to sum to unity over the discrete time steps. The exponential form of  $\varphi(t)$  implies that our problem lives in the regime where the rate of arrival of new movies is significantly larger than the rate at which most people go to the theatres to watch them, i.e., in general people do not try to maximize the number of movies they watch, and missing a specific one does not have important repercussions. Combining the bare kernel with the branching process, the proposed self-excited model gives the instantaneous rate of attendance at a time  $t$ :

$$\lambda(t) = S(t) + \sum_{i,t_i \leq t} n_i \varphi(t-t_i) \tag{3}$$

where  $n_i$  is the “fertility” of a previous attendant, defined as the total number of viewers of first generation (i.e. directly influenced by the attendant  $i$ ) that will watch that movie at any future time and who were influenced by that single person  $i$  who previously saw the movie at time  $t_i$ . The term  $S(t)$  embodies all the Exogenous shocks that may include the spontaneous attendances triggered by publicity or by new theaters showing the movie. These external sources are not directly related to previous viewers.

For the particular kernel proposed in equation 2, equation 3 can be solved in closed form for discrete time steps (see S1 Appendix for the derivation) to give:

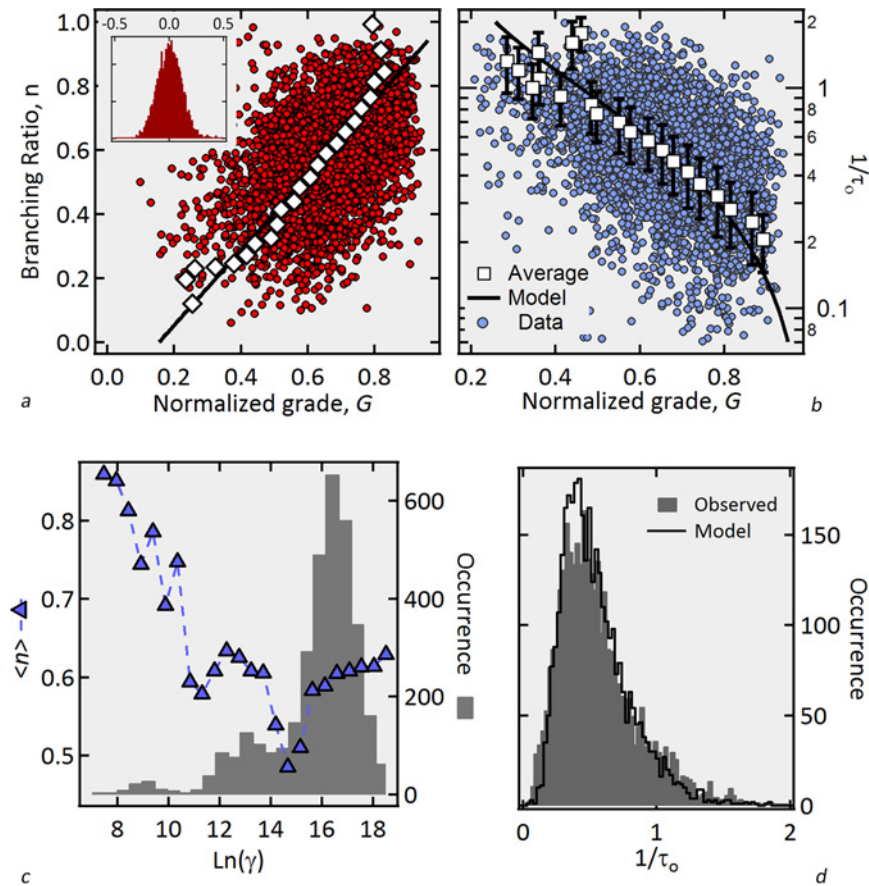
$$\lambda(t) \propto e^{-\frac{t}{\tau}(1-\theta)}, \tag{4}$$

where

$$\theta(n) \equiv \text{Log}(n(e^{1/\tau} - 1) + 1)^\tau. \tag{5}$$

In the linear-log representations used in Fig. 1, such an exponential law (eq. 4) qualifies as a straight line, which is approximately supported by the data illustrated in Fig. 1. Equation 4 implies that larger values of the branching ratio lead to slower relaxation functions that are nonetheless still exponential, i.e., the branching process renormalizes the bare kernel while keeping the exponential shape of the decay of attendance as a function of time elapsed since the peak. The observed decay rates  $1/\tau_0$  (like those presented in Fig. 3) are simply related to the branching ratio and the bare kernel as follows:

$$\frac{1}{\tau_0} = \frac{1}{\tau} (1 - \theta(n)). \tag{6}$$



**Figure 3. Results: Correlation between the branching ratio  $n$  and normalized perceived quality grade.** a) Observed branching ratio (dots, calculated from eq. 7) as a function of the perceived quality  $G$ . Solid line corresponds to the axis of symmetry of the data (eq. 8 with  $\alpha = -49.5^\circ$  and  $\Delta y = 0.118$ ; see S4 Appendix). Squares symbols are the average of  $n$  around this axis (bin size =  $1/35$ ). b) Observed decay rate  $1/\tau_0$  (circles), prediction curve (line) and averages of the data (squares) as a function of  $G$ , where error bars are the standard deviation and the constant  $(1/\tau) = 4$  was used for all movies. c) Average observed  $n$  vs.  $\ln(\gamma)$  (triangles) where  $\gamma$  is the maximum revenue for each movie. The histogram gives the number of movies in each bin of the variable  $\ln(\gamma)$ . d) Distribution of observed (solid bars) and predicted (black line) decay constants obtained from eqs. 5–8. Inset of Fig. 4a is the distribution of deviations around the symmetry axis, with standard deviation equal to 0.11

doi:10.1371/journal.pone.0116811.g003

From equations 5 and 6, the branching ratio can be solved in terms of the observed decay constant  $1/\tau_0$  of each movie and the bare kernel decay constant  $1/\tau$ , yielding

$$n = \frac{e^{\theta/\tau} - 1}{e^{1/\tau} - 1} = \frac{e^{(1-\frac{\tau}{\tau_0})/\tau} - 1}{e^{1/\tau} - 1}. \tag{7}$$

## Results and Discussion

### Branching ratio vs. $G$

Assuming that  $1/\tau$  is the same constant for all movies as in other social dynamic activities [2, 3], Fig. 3a shows the branching ratio  $n$  as a function of the perceived quality grade  $G$  for the whole dataset as obtained with eq. 7, evidencing a significant correlation between these variables. The

solid line in Fig. 3a, defined by two constants,  $\alpha$  and  $\Delta y$  for the whole dataset, is the axis of symmetry of the data around which the mean is zero and the skewness is minimized, i.e., standard deviation is almost constant as a function of  $G$  (See S4 Appendix). Similar results would have been obtained using principal component analysis. The deviations around this axis are very well described by a Normal function (inset Fig. 3a) with standard deviation equal to 0.11. This axis of symmetry is the best prediction for  $n$  as a function of  $G$  and is given by:

$$n_{pred}(G) = G \tan(-\alpha) - \frac{\Delta y}{\cos(-\alpha)}. \tag{8}$$

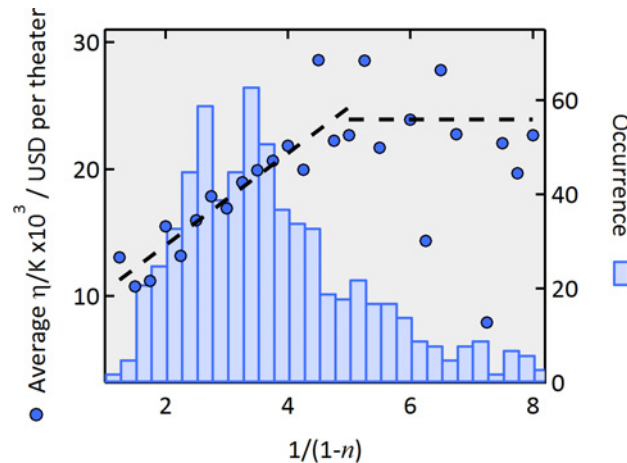
Inserting  $n_{pred}$  from equation 8 back into the expression for  $\theta(n)$  (eq. 5), we obtain a prediction of the observed decay rate for each movie (eq. 6). In Fig. 3b, we plot these predicted values along with the observed values of  $1/\tau_0$  and their corresponding averages as a function of  $G$ . The distribution of the difference between the observed and predicted decay constants are well fit to a Normal distribution with standard deviation of 0.26 (not shown). As a consequence of this good fit, not only the peak (as in refs.[6, 7, 19]) but also the complete distributions of the observed and predicted values of  $1/\tau_0$  are in good agreement with each other as shown in Fig. 3d. Fig. 3c shows the average branching ratio  $\langle n \rangle$  as a function of  $\ln(\gamma)$ , where  $\gamma$  is defined as the maximum revenue for each movie, (i.e.  $\gamma \equiv \lambda(t_c)$ ). We observe that  $n$  is a decreasing function of  $\ln(\gamma)$  for  $\ln(\gamma) \leq 15.5$  after which point it increases again. Since a larger  $n$  implies smaller  $1/\tau_0$ , this figure means that the larger the audience on the opening week, the easier the information will be transmitted about the quality of big production movies ( $\ln(\gamma) > 15.5$ ).

Even though the correlation shown in Fig. 3a is clearly present, correlation does not necessarily imply causation. The question arises whether what we interpret as a rising branching ratio resulting from an avalanche of recommendations and attendances could in fact be due to some other mechanism that is associated with the perceived quality. To further validate our interpretation of these results, we turn our attention to the accumulated weekly revenue up to the week in which the maximum activity is attained for the whole Endogenous class. This accumulated revenue is henceforth denoted by  $\eta$ . While the branching ratio  $n$  is calculated by fitting the data after the time  $t_c$  of the peak of activity, we stress that  $\eta$  is independently calculated using data up to  $t_c$ . As we show below, our cascade model predicts that  $\eta$  should be proportional to  $1/(1-n)$ . The independent estimation of these two variables and their proportionality shown in Fig. 4 thus serves as a remarkable independent evidence of the validity of our model. We now explain the reasoning in details.

Consider a small production movie played in  $\kappa_0$  theaters on the opening week with initial audience  $m_0$  proportional to the initial activity  $S_0 = \lambda(0)$ . By definition, in our avalanche model, every single attendant influences an average of  $n$  new ones to watch that movie. If such movie was allowed to play indefinitely, the total number of viewers influenced by the  $m_0$  initial ones would be proportional to  $S_0(1+n+n^2+n^3+\dots) = S_0/(1-n)$ . On the other hand, a larger number of movie theaters samples geographically a correspondingly larger number of potential movie goers. Actually, for the Endogenous class, the number of theaters grows in time (Fig. 5b) up to a maximum reached at  $t = t_c$ , and each one of these  $\kappa_t$  new theaters may induce a corresponding activity  $S_t$  that does not derive from any personal recommendation. These “external sources” are contemplated by our general equation, eq. 3 and will in turn bring a new cascade of activity each with total revenue  $S_0/(1-n)$ , should the movie play indefinitely. If, as a first approach, we assume  $S_t$  to be proportional to the corresponding number  $\kappa_t$  of new theatres ( $S_t = \alpha \kappa_t$  where  $\alpha$  is a constant), then the total revenue  $\eta$  up to  $t = t_c$  can be approximated as:

$$\eta \equiv \sum_{t=0}^{t_c} \lambda(t) \approx S_0 \sum_{t=0}^{\infty} n^t + S_1 \sum_{t=1}^{\infty} n^{t-1} + \dots + S_{t_c} \sum_{t=t_c}^{\infty} n^{t-t_c} = \frac{(S_0 + S_1 + \dots + S_{t_c})}{(1-n)} = \frac{\alpha K}{(1-n)} \tag{9}$$

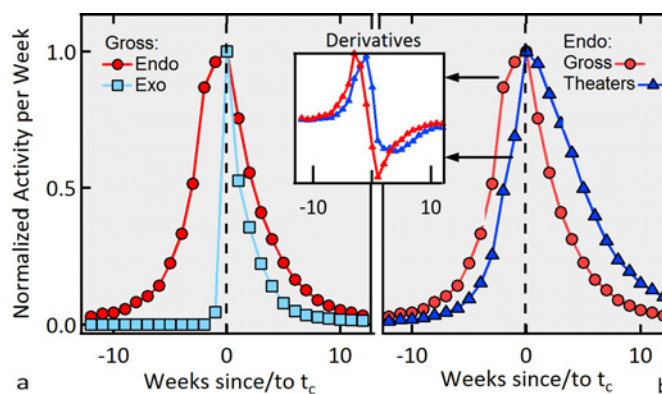




**Figure 4. Independent validation of the epidemic branching model.** Average accumulated weekly revenue  $\eta$  as a function of  $1/(1-n)$ . Dashed diagonal line is the best fit to a linear relation in the range [1, 5], which validates the epidemic branching model. Dashed horizontal line represents the average of  $\eta$  in the range [5, 8]. The blue bars quantify the occurrence number in each bin.

doi:10.1371/journal.pone.0116811.g004

where  $K = \kappa_0 + \kappa_1 + \dots + \kappa_{t_c}$ . Note that the sum of the new theaters is actually equal to the number of theaters playing that movie at  $t = t_c$ . Now, if our epidemic branching model is correct, then  $\eta/K$  should be proportional to  $1/(1-n)$ . This is shown to be the case in Fig. 4 for values of  $1/(1-n)$  in the range [1, 5] (i.e. for  $n$  in the range [0, 0.8]), where the vast majority of the data reside. Furthermore, the infinite limit approximation used in eq. 9 gives a better description of  $\eta$  when  $n$  is closer to zero ( $1/(1-n)$  between 1 and 5), than when  $n$  is closer to 1 (for  $1/(1-n) > 5$ ), because the formula is less sensitive to finite size effects when only a few generations are involved. For large values of  $1/(1-n)$ , a truncation due to several finite size effects is expected, including the finite number of theatres, movie-goers and limited time a movie is presented. This rationalizes the saturation observed in Fig. 4 for  $1/(1-n) > 5$ .



**Figure 5. Non-symmetric precursory growth.** a) Normalized accumulated gross per week as a function of time, centered on the day of the maximum activity for Endogenous (circles) and Exogenous (squares). b) Normalized accumulated gross per week (circles) and theaters (triangles) for the Endogenous class. Inset) Corresponding time derivatives in arbitrary units vs. time. The peak for the activity for the Endogenous class is non-symmetric due to the new external shocks brought by the theaters that open up once a movie proves itself successful. Note that the derivative of the normalized number of theatres playing a given Endogenous movie lags by about 2–4 weeks the derivative of the normalized gross per week.

doi:10.1371/journal.pone.0116811.g005

A null-hypothesis could be stated according to which the decay constant measured after  $t_c$  is not related of the total attendance measured before  $t_c$ , excluding our proposed mechanism in terms of the avalanches mechanism that are claimed to relate the attendance dynamics both before and after  $t_c$ . As we have just shown, [Fig. 4](#) disproves this null-hypothesis. A second null-hypothesis could be that the correlation found between  $n$  and  $G$  of [Fig. 3a](#) is a spurious result arising from pure chance. To disprove this null-hypothesis, we randomly shuffle the grades for all movies and calculate the Pearson's correlation coefficient  $r$  between the sets given by  $n$  and  $G_{random}$ . We repeat this method for 10,000 different sets of  $G_{random}$  and obtain the distribution of  $r$ , which is very well fitted to a Normal distribution with standard deviation  $\sigma = 0.017$  (not shown). In contrast, the actual sets of  $n$  and  $G$  give  $r = 0.4$ . Thus, the probability that we obtained such a correlation between  $n$  and  $G$  is essentially zero, i.e. the “ $p$ -value” for the null hypothesis to be true is essentially zero.

We recognize that the social system under study is extremely complex, in which many factors can in principle determine the success of a movie. Some of these factors may include competition, finiteness of resources for advertising, imitation vs. resistance vs. resilience in the public, or fashion waves. While investigating in detail the effects of each one of these factors would be of great interest, such a task lies outside of the scope of this article. However, as an example, in [S10 Appendix](#) we briefly explore how the dynamics of the attendance is affected by the competition between movies.

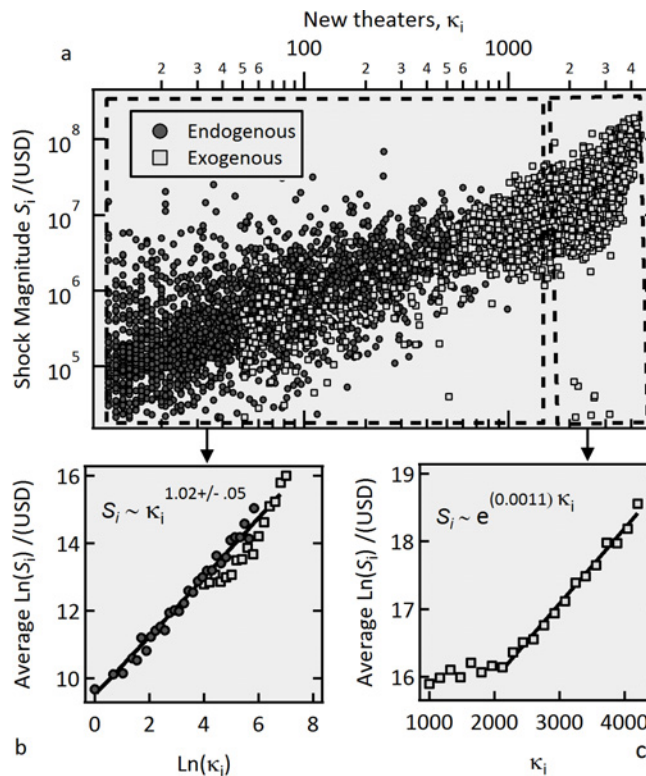
## Endogenous shocks mechanism

Endogenous shocks are usually the cumulative result of small positive perturbations that, in the case of a power law kernel, renormalize the activity yielding a slower decay [[6](#), [7](#), [19](#), [35](#)]. A second important feature arising from the action of these small perturbations is that the activity is symmetric around the peak. Note in [Fig. 1](#), however, that this is not the case for *Little Miss Sunshine*, nor is it in general for the rest of the Endogenous movies as can be appreciated in the accumulated normalized activity displayed in [Fig. 5a](#). These facts suggest that, in this social system, there exists a different mechanism responsible for the pre-peak growth we find, as anticipated by the analysis performed in the previous section. The key to finding this mechanism can be determined from an important difference between the attendance to movies and other social activities such as watching videos or buying books online: the accessibility of the product. While the latter two activities can be performed at all times (assuming the book is not off-print), in the case of movies, the specific film in which we are interested must be playing nearby our usual whereabouts for us to go and watch it. It is then reasonable to assume that the number of theaters playing a movie at different times during the lifetime of a movie will be a factor in determining the dynamics of the attendance. To test this hypothesis, in [Fig. 5b](#), we show both the normalized accumulated gross per week and number of theaters for the whole Endogenous class. One can observe that, once a movie begins to perform well (according to some standard), new theaters will decide to play that movie in an effort to profit from this proven success, i.e., supply adaptation is dynamic [[19](#)]. The inset of [Fig. 5](#) displays the corresponding time derivatives. Note how, even though the response from the supply side (the available theaters) is slow at the beginning, eventually it becomes faster than the growth rate of the revenue itself until a maximum number of theaters is reached. At this point, when no more theaters open up, the activity of the attendance reaches its maximum and the system relaxes thereafter following an exponential decay with constant  $(1/\tau_0)$ . If this picture is correct, then new theaters serve as sources of new external shocks  $S(t)$ , as contemplated in our general [equation 3](#) for the activity. Let the time series of these new theaters be  $\{\kappa_i\}$ , which is also known for each movie. To obtain the series of new shocks  $\{S_i\}$ , given the activity  $\lambda(t)$ , we assume that every new

generation  $g$  of movie goes evolves in time following [eq. 1](#) with the same  $1/\tau_0$  observed after the peak for that Endogenous movie. Then, the external sources can be obtained recursively as:

$$S(t) \equiv S_t(0) = \lambda(t) - \sum_{g=0}^{t-1} S_g(t-g) \tag{10}$$

where  $S_g(t-g) = S_0(0)e^{-(t-g)/\tau_0}$ , and  $S_0(0) \equiv \lambda(0)$ . For a given Endogenous movie, the time series  $S_i$  is obtained using [equation 10](#) assuming the measured  $1/\tau_0$  for that movie ([Fig. 3b](#)) applies to every new generation (see [S2 Appendix](#) for the derivation of [eq. 10](#), and [S8 Appendix](#) for an example of its application). To investigate the response function of the system, we analyze the relation between  $\kappa_i$  and the corresponding  $S_i$  for the whole set of precursors of the Endogenous shocks (3586 such events were extracted) in [Fig. 6](#). The circles in [Fig. 6a](#) and the corresponding averages shown in [Fig. 6b](#) strongly suggest there exists a power law relationship between the number of new theaters and the magnitude of the shocks the produce. We find a power law exponent of  $1.01 \pm 0.05$ , which corresponds to a linear relation, i.e.  $S_i \propto \kappa_i$ , as we had anticipated in the analysis of [Fig. 4](#). We realize that the proposed proportionality between  $\kappa_i$  and  $S_i$  can only hold on average, because it is the eventual weakening of the response (or in mathematical terms, the eventual reduction of the susceptibility) from the public to a given “impulse” what brings the activity to a maximum. In that respect, further analyses of the time



**Figure 6. External sources of new shocks.** a) Magnitude of the induced shocks  $S_i$  vs. number of new theaters  $\kappa_i$ . The red circles depict  $S_i$  as calculated with [equation 9](#) for the precursors’ activity of the whole Endogenous class. In the calculation of  $S_i$ , the activity from previous generations was properly subtracted (see [S2 Appendix](#)). The squares represent the maximum activity  $\gamma_i$  vs. the number of theaters that played the movies on their opening week for the Exogenous class. (b) Average  $\ln(S_i)$  as a function of  $\ln(\kappa_i)$  for the Endogenous precursors, showing a linear dependence. (c) Average  $\ln(S_i)$  vs.  $\kappa_i$  for the Exogenous class. For  $\kappa_i > 2000$  the dependence becomes exponential.

doi:10.1371/journal.pone.0116811.g006



evolution of the average  $S_i/\kappa_i$  would provide insights about the response function of this social system.

We have postulated that the dynamics of the relaxation in this social system is universally given by eqs. 3–5, regardless of the type of shock. It is then interesting to verify if there exists a difference in the excitation process between Endogenous and Exogenous shocks. To this end, we perform the equivalent  $S_i$  vs.  $\kappa_i$  analysis now with the set of Exogenous initial shocks. Noting that, in this case, there are no precursors, we take as single sources of excitation the number of theaters  $\kappa_i$  from the opening week and the corresponding maximum gross per week  $\gamma_i$  as the magnitude of the shock (squares, Fig. 6a).

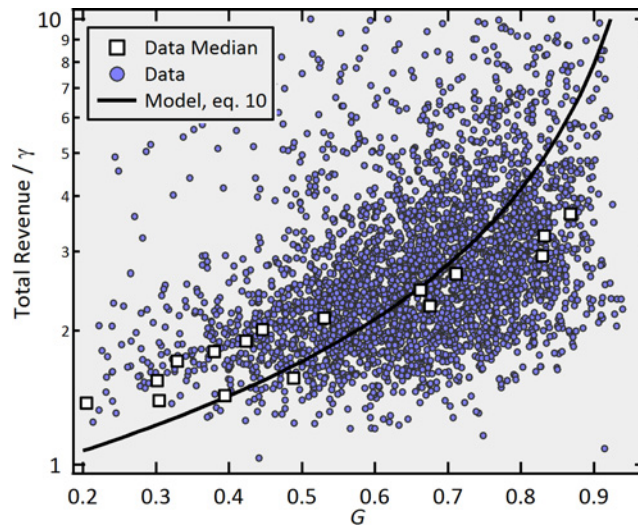
Fig. 6b shows that the events from the Exogenous shocks extend the power law trend of the Endogenous precursors up to a limit  $\kappa_i = 2000$ . In other words, new theaters serve as linear sources of new shocks irrespective of whether a movie belongs to the Endogenous or the Exogenous class. Nevertheless, for larger events than those accessible by Endogenous shocks, ( $\kappa_i > 2000$ ), the log-linear graph on Fig. 6c suggests that the relation between  $\kappa_i$  and  $S_i$  transitions to an exponential function, signaling a change has taken place to an exponential regime. Note in the y-axis of Fig. 6c that this transition happens to movies with  $\ln(\gamma) > 16.2$ . This corresponds roughly to the same threshold we find in Fig. 3c after which the average branching ratio becomes an increasing function of  $\ln(\gamma)$ . These two pieces of evidence suggest that there exists a critical level of connectivity [10, 36] in the social network above which the interaction between people becomes non-linear and as a result, much stronger. Taking the average ticket price to be \$10, this threshold can be estimated to be of the order of  $(e^{16.2}/10) \approx 1$  million people.

### Total revenue vs. G

Even though the strong linear correlation that we find between the perceived quality and the branching ratio is quite revealing, what is utterly relevant for the movie industry is the impact that this relation has on the total revenue. As it turns out, the effect of the perceived quality on the revenue is also non-linear. To illustrate this, take two of the movies analyzed in Fig. 1, *Little Miss Sunshine* and *The adventures of Pluto Nash* that presented roughly the same maximum weekly revenue. While the former ended up making \$60M, the later made only 4.4M at the box office. Indeed, *Little Miss Sunshine* was an Endogenous smash-hit for which every former viewers brought to the theaters an average of  $n(G = 87\%) \approx 0.83$  new ones, whereas those who saw *Pluto Nash* influenced on average only  $n(G = 20\%) \approx 0.05$  new viewers. Note that even though the ratio of their grades is  $(0.87/0.20) \approx 4$ , the ratio of their earnings is almost 14. This non-linear effect can be understood by writing explicitly the approximate total revenue (after the maximum activity was reached) assuming the movies are exhibited indefinitely in theaters:

$$TotalRevenue \approx \gamma \sum_{t=t_c}^{\infty} e^{-(t-t_c)/\tau_0} = \gamma \left( \frac{1}{1 - e^{-1/\tau_0}} \right) = \gamma \left( \frac{1}{1 - e^{-(1-\theta)/\tau}} \right) \approx \gamma \left( \frac{1}{1 - e^{-(1-n)/\tau}} \right), \quad (11)$$

where  $n = n(G)$  is a function of the perceived quality given by equation 8, and the approximation  $\theta \approx n$  was used. Recall that  $\gamma$  is the peak revenue of the corresponding movie. The error obtained by using the infinity limit in equation 11 is actually less than 5% (see S5 Appendix). According to equation 11, as the branching ratio tends to 1, the revenue becomes infinite. However, the relation between  $n$  and  $G$  (eq. 8) sets a limit on the maximum  $n$  attainable. With the parameters used to fit the data in Fig. 3 ( $\alpha = -49.5^\circ$  and  $\Delta y = 0.11$ ), then  $n(G = 100\%) = 0.986$ , which in turn implies that the total revenue is at most equal to  $82\gamma$ , but may be as small as  $1.08\gamma$  for  $n(G = 20\%) = 0.05$ .



**Figure 7. Revenue ratio vs. G.** Total revenue divided by the peak revenue  $\gamma$  for all movies belonging to the Exogenous class (dots) and the prediction of the model (line, eq. 10). The total revenue is the actual profit of these movies as found on the website.

doi:10.1371/journal.pone.0116811.g007

Fig. 7 shows the total revenue earned after  $t_c$  divided by the corresponding peak revenue  $\gamma$  as a function of  $G$  for each movie belonging to the Exogenous class and their median values, along with the prediction of our theoretical model. Through eq. 11 and Fig. 7, our model explains why it is so important for the film industry that audiences like and recommend the movies it produces: since the box office earnings on the first week is of the order of  $e^{16.5} \approx 15$  million dollars (see S7 Appendix), a “must see” movie ( $G = 90\%$ ) can make  $(11.9 \times 15 \text{ million}) = 178$  million dollars while a “don’t watch it” one ( $G = 20\%$ ) will only make  $(1.08 \times 15 \text{ million}) = 16.2$  million, even if both movies were equally successful during the opening week. This analysis becomes even more relevant if we consider that the average movie costs about \$100 million dollars to produce [37]. These results help explain the conclusions of other studies that have found that launching advertising campaigns after the movie has already opened does not have a strong impact on the final revenue [20]. Broadcasting can only go so far.

## Conclusions

We have shown how “perceived movie quality” translates into a probability of watching that movie and the power of direct connections to spread information within social networks. While this relation is linear, its effect on the total number of people involved in this activity is exponential.

Further studies regarding the transition we find in the response of the system to external shocks from linear to exponential (Fig. 6) could give information about the topology of the network. In turn, this could help predict the dynamics of other social processes in which over 1 million people serve as the nucleation threshold for global information propagation. These activities may range from picking candidates in governmental or presidential elections and supporting humanitarian causes, to starting a social movement or buying the newest smart phone.

The results presented in this work open up the possibility of understanding and predicting the dynamics of a wide range of social systems for which the perceived quality is recorded. The advent of the internet has made it as easy as pressing a key to express our assessment of millions of products, activities, social causes, political views, and many others, in a society that is addicted to ranking [38] tagging [39] and betting [40]. Similar methods to the ones we have

applied here could provide concrete quantifiable metrics of peoples opinion's concerning readily available or future social activities or commercial products.

## Supporting Information

**S1 Appendix. Hawkes process with exponential kernel: Response function for Exogenous Shocks.**

(PDF)

**S2 Appendix. New theaters as sources of Exogenous shocks: precursors for Endogenous shocks.**

(PDF)

**S3 Appendix. On data filtering.**

(PDF)

**S4 Appendix. Finding the axis of symmetry.**

(PDF)

**S5 Appendix. Total revenue vs. Grade.**

(PDF)

**S6 Appendix. Correlation between the normalized grade  $G$  used in the present paper, and the "Audience Average Rating" obtained from the website [www.rottentomatoes.com](http://www.rottentomatoes.com).**

(PDF)

**S7 Appendix. General Trends of the data.**

(PDF)

**S8 Appendix. An example of how new theaters serve as a source of external shocks.**

(PDF)

**S9 Appendix. Note on medians.**

(PDF)

**S10 Appendix. Correlation between then number of movies playing, and  $n$  and  $n/G$ .**

(PDF)

**S11 Appendix. Methods.**

(PDF)

## Author Contributions

Conceived and designed the experiments: JVE DS. Performed the experiments: JVE. Analyzed the data: JVE. Wrote the paper: JVE DS. Theoretical models: JVE DS.

## References

1. Roehner BM, Sornette D, Andersen JV (2004) Response functions to critical shocks in social sciences: An empirical and numerical study. *Int. J. Mod. Phys. C* 15(6):809–834. doi: [10.1142/S0129183104006236](https://doi.org/10.1142/S0129183104006236)
2. Sornette D (2002) Predictability of catastrophic events: Material rupture, earthquakes, turbulence, financial crashes, and human birth. *Proc. Natl. Acad. Sci. U. S. A.* 99:2522–2529. doi: [10.1073/pnas.022581999](https://doi.org/10.1073/pnas.022581999) PMID: [11875205](https://pubmed.ncbi.nlm.nih.gov/11875205/)
3. Sornette D, Malevergne Y, Muzy JF (2003) Tail risk: What causes crashes? *Risk* 16 (2), 67–71.
4. Schweitzer F, Fagiolo G, Sornette D, Vega-Redondo F, Vespignani A et al. (2009) Economic Networks: The New Challenges. *Science* 325(5939):422–425. PMID: [19628858](https://pubmed.ncbi.nlm.nih.gov/19628858/)

5. Barabasi AL (2005) The origin of bursts and heavy tails in human dynamics. *Nature* 435(7039):207–211. doi: [10.1038/nature03459](https://doi.org/10.1038/nature03459) PMID: [15889093](https://pubmed.ncbi.nlm.nih.gov/15889093/)
6. Crane R, Sornette D (2008) Robust dynamic classes revealed by measuring the response function of a social system. *Proc. Natl. Acad. Sci. U. S. A.* 105(41):15649–15653. doi: [10.1073/pnas.0803685105](https://doi.org/10.1073/pnas.0803685105) PMID: [18824681](https://pubmed.ncbi.nlm.nih.gov/18824681/)
7. Sornette D, Deschatres F, Gilbert T, Ageon Y (2004) Endogenous versus Exogenous shocks in complex networks: An empirical test using book sale rankings. *Phys. Rev. Lett.* 93 (22) 228701. doi: [10.1103/PhysRevLett.93.228701](https://doi.org/10.1103/PhysRevLett.93.228701) PMID: [15601125](https://pubmed.ncbi.nlm.nih.gov/15601125/)
8. Borge-Holthoefer J, Meloni S, Goncalves B, Moreno Y (2013) Emergence of Influential Spreaders in Modified Rumor Models. *J. Stat. Phys.* 151(1–2):383–393. doi: [10.1007/s10955-012-0595-6](https://doi.org/10.1007/s10955-012-0595-6)
9. Gonzalez-Bailon S, Borge-Holthoefer J, Moreno Y (2013) Broadcasters and Hidden Influentials in Online Protest Diffusion. *American Behavioral Scientist* 57(7):943–965. doi: [10.1177/0002764213479371](https://doi.org/10.1177/0002764213479371)
10. Gonzalez-Bailon S, Borge-Holthoefer J, Rivero A, Moreno Y (2011) The Dynamics of Protest Recruitment through an Online Network. *Sci. Rep.* 1:197. doi: [10.1038/srep00197](https://doi.org/10.1038/srep00197) PMID: [22355712](https://pubmed.ncbi.nlm.nih.gov/22355712/)
11. Borge-Holthoefer J, Rivero A, Moreno Y (2012) Locating privileged spreaders on an online social network. *Phys. Rev. E* 85(6):6. doi: [10.1103/PhysRevE.85.066123](https://doi.org/10.1103/PhysRevE.85.066123)
12. Borge-Holthoefer J, Rivero A, García I, Cauhé E, Ferrer A, et al. (2011) Structural and Dynamical Patterns on Online Social Networks: The Spanish May 15th Movement as a Case Study. *PLoS ONE* 6 (8):8. doi: [10.1371/journal.pone.0023883](https://doi.org/10.1371/journal.pone.0023883)
13. How companies are marketing online: A McKinsey Global Survey, July 2007, McKinsey Quarterly survey of business executives.
14. Social Marketing: How Companies Are Generating Value from Customer Input. Available: <http://knowledge.wharton.upenn.edu/article.cfm?articleid=1864>.
15. Borghesi C, Bouchaud JP (2007) Of songs and men: a model for multiple choice with herding. *Quality & Quantity* 41(4):557–568. doi: [10.1007/s11135-007-9074-6](https://doi.org/10.1007/s11135-007-9074-6)
16. Moretti E (2011) Social Learning and Peer Effects in Consumption: Evidence from Movie Sales. *Review of Economic Studies* 78(1):356–393. doi: [10.1093/restud/rdq014](https://doi.org/10.1093/restud/rdq014)
17. Moul CC (2007) Measuring word of mouth's impact on theatrical movie admissions. *Journal of Economics & Management Strategy* 16(4):859–892. doi: [10.1111/j.1530-9134.2007.00160.x](https://doi.org/10.1111/j.1530-9134.2007.00160.x)
18. McKenzie J (2010) How do theatrical box office revenues affect DVD retail sales? Australian empirical evidence. *J. Cult. Econ.* 34:159–179.
19. Deschatres F, Sornette D (2005) The Dynamics of Book Sales: Endogenous versus Exogenous Shocks in Complex Networks, *Phys. Rev. E* 72, 016112. doi: [10.1103/PhysRevE.72.016112](https://doi.org/10.1103/PhysRevE.72.016112)
20. De Vany A (2003) "Hollywood Economics: How Extreme Uncertainty Shapes the Film Industry", (Routledge Press, UK).
21. Sornette D, Helmstetter A (2003) Endogenous versus Exogenous shocks in systems with memory. *Physica A-Statistical Mechanics and Its Applications* 318(3–4):577–591. doi: [10.1016/S0378-4371\(02\)01371-7](https://doi.org/10.1016/S0378-4371(02)01371-7)
22. de Menezes MA, Barabasi AL (2004) Separating internal and external dynamics of complex systems. *Phys. Rev. Lett.* 93(6):4.
23. Lambiotte R, Ausloos M (2006) Endo- vs. Exogenous shocks and relaxation rates in book and music "sales". *Physica A-Statistical Mechanics and Its Applications* 362(2):485–494. doi: [10.1016/j.physa.2005.08.047](https://doi.org/10.1016/j.physa.2005.08.047)
24. Valverde S (2007) Crossover from Endogenous to Exogenous activity in open-source software development. *Europhys. Lett.* 77(2):6. doi: [10.1209/0295-5075/77/20002](https://doi.org/10.1209/0295-5075/77/20002)
25. Johansen A, Sornette D (2000) Download relaxation dynamics on the WWW following newspaper publication of URL. *Physica A-Statistical Mechanics and Its Applications* 276(1–2):338–345.
26. Bass F (1969) A new product growth model for consumer durables. *Management Science* 15 (5), 215–227. doi: [10.1287/mnsc.15.5.215](https://doi.org/10.1287/mnsc.15.5.215)
27. De Vany A, Lee C (2001) Quality signals in information cascades and the dynamics of the distribution of motion picture box office revenues, *Journal of Economic Dynamics & Control*, 25 593–614. doi: [10.1016/S0165-1889\(00\)00037-3](https://doi.org/10.1016/S0165-1889(00)00037-3)
28. Sitabhra S, Pan RK (2005) Blockbusters, Bombs and Sleepers: The income distribution of movies. In *Econophysics of Wealth Distributions, New Economic Windows*, pp 43–47 (Springer, New York, USA).
29. Hawkes AG, Oakes D (1974) Cluster process representation of a self-exciting process. *Journal of Applied Probability* 11(3):493–503.

30. Saichev A, Sornette D (2009) Effects of Diversity and Procrastination in Priority Queuing Theory: the Different Power Law Regimes, *Phys. Rev. E* 81, 016108.
31. Oliveira JG, Barabasi AL (2005) Human dynamics: Darwin and Einstein correspondence patterns. *Nature* 437(7063):1251–1251. doi: [10.1038/4371251a](https://doi.org/10.1038/4371251a) PMID: [16251946](https://pubmed.ncbi.nlm.nih.gov/16251946/)
32. Filimonov V, Sornette D (2011) Self-excited multifractal dynamics. *Europhys. Lett.* 94, 46003. doi: [10.1209/0295-5075/94/46003](https://doi.org/10.1209/0295-5075/94/46003)
33. Saichev A, Sornette D (2009) Effects of Diversity and Procrastination in Priority Queuing Theory: the Different Power Law Regimes, *Phys. Rev. E* 81, 016108. doi: [10.1103/PhysRevE.81.016108](https://doi.org/10.1103/PhysRevE.81.016108)
34. Maillart T, Sornette D, Frei S, Duebendorfer T, Saichev A (2011) Quantification of deviations from rationality from heavy-tails in human dynamics, *Phys. Rev. E* 83, 056101. doi: [10.1103/PhysRevE.83.056101](https://doi.org/10.1103/PhysRevE.83.056101)
35. Saichev A, Sornette D (2010) Generation-by-Generation Dissection of the Response Function in Long Memory Epidemic Processes. *Eur. Phys. J. B* 75, 343–355. doi: [10.1140/epjb/e2010-00121-7](https://doi.org/10.1140/epjb/e2010-00121-7)
36. Sornette D (2006), *Critical Phenomena in Natural Sciences: Chaos, Fractals, Self-organization and Disorder: Concepts and Tools*, PP 179, (Springer, New York, USA).
37. Eliashberg J, Hui SK, Zhang ZJ (2007) From Story Line to Box Office: A New Approach for Green-Lighting Movie Scripts, *Management Science*, 53, 6 (issn 0025–1909). doi: [10.1287/mnsc.1060.0668](https://doi.org/10.1287/mnsc.1060.0668)
38. Blumm N, Ghoshal G, Forró Z, Schich M, Bianconi G et al. (2012) Dynamics of Ranking Processes in Complex Systems. *Phys. Rev. Lett.* 109(12). doi: [10.1103/PhysRevLett.109.128701](https://doi.org/10.1103/PhysRevLett.109.128701) PMID: [23005999](https://pubmed.ncbi.nlm.nih.gov/23005999/)
39. Cattuto C, Loreto V, Pietronero L (2007) Semiotic dynamics and collaborative tagging, *Proc. Natl. Acad. Sci. U. S. A* 104(5):1461–1464. doi: [10.1073/pnas.0610487104](https://doi.org/10.1073/pnas.0610487104) PMID: [17244704](https://pubmed.ncbi.nlm.nih.gov/17244704/)
40. Saiz A, Simonsohn U (2007) Predictions and perceptions: downloading wisdom from online crowds Available: <http://ssrn.com/abstract=990021>.